Applying Hybrid time series models for modeling bivariate time series data with different distributions for forecasting unemployment rate in the USA

Firas Ahmed Mohammed\textsuperscript{1} , Moamen Abbas Mousa\textsuperscript{2}

\textsuperscript{1}Ass.Prof. Department of statistics, College of Management and Economics, Baghdad University, Baghdad, Iraq.

\textsuperscript{2}PhD student. Department of Statistics, College of Management and Economics, Baghdad University, Baghdad, Iraq.

E-mail : \textsuperscript{1}drfirasmohana@gmail.com, \textsuperscript{2}saidmoamen@gmail.com

Abstract

Unemployment rate forecasting has become a particularly promising field of research in recent years because it's an important problem in state planning and management. Since the time series data are rarely pure linear or nonlinear obviously, sometimes contain both components jointly. Therefore, this study introduces new hybrid models contain three commonly used, first is the Stochastic Linear Autoregressive Moving Average with eXogenous variable (ARMAX) model for modeled the relationship between the unemployment rate and exchange rate, second and third are a nonlinear Generalized Autoregressive Conditional Heteroscedasticity (GARCH) and GARCH with eXogenous variable (GARCHX) used When the assumption of homoscedasticity error variance is violated for the purpose of capture the volatility in the residuals of ARMAX model and to enhance the Forecasting ability of ARMAX model by combining it with other nonlinear models. In this case, to have a better forecasting efficiency, we introduce a hybrid methodology of amalgamating the forecasts from a linear time series model (ARMAX) and from a nonlinear time series model (GARCH, GARCHX) with three different distributions (Normal Distribution, Student’s t-distribution and General Error Distribution (GED)), the last two distributions for capturing fat-tailed properties in residuals time series. The hybrid approach specifically (ARMAX-GARCH) and (ARMAX-GARCHX) have been used for modeling and forecasting the unemployment rate in the USA. Diverse approaches have been employed in the parameters vector estimation. A comparison evaluation was as well been done based on mean absolute error (MAE), mean absolute percentage error (MAPE), and as well as Root mean square error (RMSE) between the hybrid and their individual rival model in accordance with forecasting performance. From investigational results, it is perceived that the hybrid model (ARMAX-GARCHX) is more effective than other twin hybrid and individual rival models for the data under study. MATLAB, SAS, and EViews software packages have been used for the data analysis.

Keywords : ARMAX, GARCH, GARCHX, Normal distribution, Student-t distribution, General Error distribution (GED), Hybrid model, Unemployment rate, Exchange rate.
I. Introduction

Time series forecasting stands for an imperative statistical analysis procedure as a basis for manual and automatic planning in numerous application fields[XII]. The econometric investigation of economic and business time series stands for a main researchfield and application. The last few decades have perceived agrowing interest in theoretic and experimental expansions in building time series models and in their essential application in forecasting [VIII]. Time series forecasting stands for an imperative field of forecasting in which observing the same variable is a sequence of taken observations successively in time. Numerous sets of data act as time series: a once-a-month sequence, day-to-day sequence, hourly sequence, etc., are analyzed to develop and construct a model depicting the fundamental relationship [X]. Forecasting rules can play an important role in many areas such as business, industry, and intergovernmental organizations. Many times we have a few knowledge about the underlying data generating process. So, in this cases modeling methodology becomes advantageous.

Apart from several factors namely the interest rate, Gross Domestic Product (GDP), and inflation. The exchange rate and unemployment are one of the major factors that keeping importance in economic growth advancement. To modeling these two factors, the time series literature provided various linear time series models. One of the major and commonly used approaches for analysis of bivariate time series data is the Autoregressive Moving Average with exogenous variable (ARMAX) model. The ARMAX model is probably the second most popular linear model after the ARX model. Compared with the ARX, the ARMAX model is a more flexible class because it possesses an extended noise model and due to its statistical properties [XVIII]. ARMAX is a flexibility class of models including mixed pure autoregressive (AR) and moving average (MA) models with additional external input called exogenous variable. But one of the main constraints of ARMAX models is the linearity structure of the models. This assumption of linearity restrictions the application of the ARMAX model to real-time series data. There are many studies that have discussed the application of this model such as (Yiu and Wang, 2007), (Aldemğr and Hapoğlu, 2015) [XXV][I].

Linear models have no likelihood to refer to any instability in the actual conditional variance in the real time series data or in The residuals of ARMAX linear model. To overwhelm this tricky issue, Engle (1982) has proposed Autoregressive Conditional Heteroskedasticity (ARCH) statistical model, for the purposing capturing the volatility in time series data that depicts the variance of the current error term or innovation as a function of the actual sizes of the preceding time periods error terms [V]. The ARCH model can besuitableas the error variance in a time series adheres to an autoregressive (AR) model and has a drawback with a huge required number of parameters in construction of the forecast model. Consequently, Bollerslev (1986) suggests a more economical method is the Generalized ARCH (GARCH) model appropriate as the mixed autoregressive moving average (ARMA) model has presumed for the error variance [III]. There are many empirical studies confirm that the nonlinear models have a good performance for long term forecasting whereas the linear models are appropriate for short term forecasting, as well as the real time-series data often composed of linear and nonlinearities compound [XVI][XVII].
So there is a necessity of hybridization the linear and nonlinear models in one hybrid model in order to obtain more efficient forecast. And for the process of building these model we use the nonlinear GARCH and GARCHX model with three different distributions (Normal Distribution, Student’s t-distribution and General Error Distribution (GED)) in hybridization method with linear ARMAX model. The Student’s t and GED distribution are applied for capturing fat-tailed properties in residuals time series. Numerous hybrid models have verified and the optimal model has been selected based on model selection criteria like Akaike information criteria (AIC), Bayesian information criteria (BIC), Final Prediction Error (FPE), in addition to Hannan and Quinn (HQ) criteria. Nevertheless, there is no guaranteed that the final selected model will give the perfect forecast if there are some effectual factors such as the model type and distributions of random errors used in the hybridization method. To deal with this problem different forecasting models can be mixed together in a single hybrid model to get the optimum composition of the final forecast. Much of the problems in the real world is complicated in nature and any single model is incapable to capture different patterns regularly, that means, no single forecasting model will be the better choice in every case. Consequently, the mixing of different single models is influential and important to increase the chance of capturing different structures and yield more efficient forecasting performance. Many researchers (Hickey et al., 2012) [XII], and (Porshnev et al., 2016) [XX] and others have studied the combined ARMAX model with GARCH and showed the improvement of the hybrid model in forecasting accuracy. In the present study, ARMAX-GARCH and ARMAX-GARCHX models with three different distributions along with the individual ARMAX model have been applied to the real data set. MATLAB, SAS, and EViews software packages have been used for the data analysis.

**Research Significance**

Some methodology adopted for a mathematical model was taken into account for developing the regression analysis to predict 28 days and 56 days compressive and splitting tensile strengths of GPC, which may serve as the useful tools in the civil engineering optimization problems such as optimization of concrete mixtures. In certain applications toughness property is needed. Experimental data were statistically analyzed for developing mathematical models considering more influencing factors which are not considered by the early researchers, and also the models were verified for its performance/suitability [XVI].

**Statistical Analysis for Strength Prediction**

The strengthening of concrete is a complex process involving many external factors. A number of improved prediction techniques have been proposed by including empirical or computational modeling, statistical techniques. Many attempts have been made for modeling this process through the use of computational techniques such as finite element analysis. While, a number of research efforts have concentrated on using multivariable regression models to improve the accuracy of predictions. Statistical models have the attraction that once fitted they can be used to perform predictions much more quickly than other modeling techniques, and are correspondingly simpler to implement in software. Apart of its speed, statistical modeling has the advantage over other techniques that it is mathematically rigorous and can be used to define confidence interval for the predictions. This is especially true when comparing statistical modeling with other models. Statistical analysis can also provide insight into the key factors.
influencing 28 days compressive strength through regression analysis. For these reasons, statistical analysis was chosen to be a technique for strength prediction of this paper [VIII], [XV].

II. Material and Methods

In this section, We'll describe two types of forecasting time series models used in the hybridization method:

II.i.

In modeling the mean equation, one of the most important linear models for modeling bivariate time series is Autoregressive Moving Average with exogenous variable (ARMAX). In the single-input single-output (SISO) ARMAX model structure specification the endogenous variable is modeled as difference equation in Eq. (1 and 2) [XXIII][XXII]:

\[ y_t = \mu + \sum_{i=1}^{n_p} \phi_i y_{t-i} - \sum_{j=1}^{n_q} \theta_j \varepsilon_{t-j} + \sum_{k=1}^{n_b} \Phi_k x_{t-n_k} + \varepsilon_t, \]  

(1)

Or in the compact form:

\[ \Phi_{n_p}(L) y_t = \phi_{n_b}(L) x_{t-n_k} + \Theta_{n_q}(L) \varepsilon_t, \]  

(2)

Where:

- \( y_t \): endogenous variable (Model output at time \( t \)).
- \( x_{t-n_k} \): exogenous variable, Previous and delayed inputs on which the current output depends.

The parameters \( n_p, n_b, \) and \( n_q \) are the orders of the ARMAX model (the order of autoregressive, exogenous variable and moving average respectively) 1 , and \( n_k \) is the delay time2. \( \Phi_{n_p}(L), \Theta_{n_q}(L), \phi_{n_b}(L) \) are the polynomial of lag operator \( L \) of order \( n_p, n_q \) and \( n_b \) respectively with root outside the unit circle such that :

1. \[ \Phi_{n_p}(L) = 1 + \phi_1 L + \phi_2 L^2 + \cdots + \phi_{n_p} L^{n_p} = 1 + \sum_{i=1}^{n_p} \phi_i L^i \]  

2. \[ \Theta_{n_q}(L) = 1 - \theta_1 L^1 - \theta_2 L^2 - \cdots - \theta_{n_q} L^{n_q} = 1 - \sum_{j=1}^{n_q} \theta_j L^j \]  

3. \[ \phi_{n_b}(L) = \phi_1 L^{-1} + \phi_2 L^{-2} + \cdots + \phi_{n_b} L^{-n_b} = \sum_{k=1}^{n_b} \Phi_k L^k \]  

(2.a, 2.b, 2.c)

In equation (1), \( \varepsilon_t \) represents GARCH innovations for modeled the variance equation, and the next section illustrates this term.

\[^1\text{The orders (} n_p, n_b, \text{ and } n_q \text{) of ARMAX model specified using (PACFs, ACFs, EACFs) by MATLAB and SAS.}\]

\[^2\text{To specify the order of delay time (} n_k \text{) we use the cross-correlation function between } y_t \text{ and } x_t.\]
1 The orders \((n_p, n_d, \text{and } n_q)\) of ARMAX model specified using (PACFs, ACFs, EACFs) by MATLAB and SAS.

To specify the order of delay tie \((n_k)\) we use the cross-correlation function between \(y_t\) and \(x_t\).

### II.ii. Conditional Variance Models:

For modeling the residuals of the ARMAX model in the mean equation (1), the non-linear (ARCH/GARCH/GARCHX) model is applied in the variance equation, as in The following sections:

#### II.ii.a The ARCH model:

Linear time series models are unable to interpret volatility (conditional variance) which is introduced in several real time series data or in its residuals, to deal with this problem Engle (1982) has introduced the ARCH models. The ARCH method refers to a process in which volatility has fluctuations in a specific way. Let us an normal ARCH(r) model for a series \(\epsilon_t\):

\[
\epsilon_t = h_t \eta_t, \quad t = 1, 2, \ldots, n, \quad \eta_t \sim \text{i.i.d } N(0, 1)
\]

\[
h_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i \epsilon_{t-i}^2, \quad \alpha_0 > 0, \quad \sum_{i=1}^{r} \alpha_i < 1,
\]

\(h_t\) is called the conditional variance of \(\epsilon_t\).

But the ARCH model has a disadvantage that, when the order of the ith ARCH model is very large, estimation of a large number of parameters is required which is a really complicated process. Also, the ARCH model is not a parsimonious model\[8,16\].

#### II.ii.b. The GARCH model:

To overwhelm these problematical difficulties of the ARCH model. Bollerslev (1986) offers a more parsimonious method termed as Generalized ARCH (GARCH) model that Conditional variance has a linear function of the square of the error term and as well a linear function of its specific lags. For modeling the conditional variance of ARMAX residuals \((\epsilon_t)\) can analyze by using the GARCH \((r,s)\) model represented in the following equations\[VIII\] [III]:

\[
\epsilon_t = h_t \eta_t, \quad \eta_t \sim \text{i.i.d } N(0, 1)
\]

Where:

\(\epsilon_t\) : represents GARCH innovations.

\(\eta_t\) : The random variable \(\eta_t\) is an innovation term which is typically assumed to be independent and identically distributed (i.i.d) with mean zero and unit variance.

\(h_t\) : Conditional variance \(h_t\) is modeling as GARCH\((r,s)\), as in Eq. (4):

\[
h_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{s} \theta_j h_{t-j}
\]

---

Copyright reserved © J. Mech. Cont.& Math. Sci.
Firas Ahmed Mohammed et al

355
where r, s are the orders of GARCH model, and parameters $\alpha_i$ and $\theta_j$ estimate for the ARCH and GARCH effects of ith and jth orders respectively. The parameters $(\alpha_0, \alpha_1, \ldots, \alpha_r, \theta_1, \ldots, \theta_s)$ are restricted such that ($h_t > 0$) for all t, which is ensured when:

$$\alpha_0 > 0, \quad \alpha_i \geq 0 \text{ for } i = 1, 2, 3, \ldots, r, \quad \theta_j \geq 0, \quad \text{for } j = 1, 2, 3, \ldots, s$$

From Eq. (4) the simple GARCH (1,1) model, that is the most prevalent method for modeling volatility. We write this model as:

$$\varepsilon_t = h_t \eta_t, \quad \eta_t \sim \text{i.i.d } N(0, 1)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \theta_1 h_{t-1}$$

(5)

The conditional variance in Eq.(4) is modeled by the past shock $\varepsilon_{t-1}^2$ and its own lagged value $h_{t-1}$. For $\alpha_0 \geq 0, \alpha_1 > 0, \theta_1 > 0$ and $\alpha_1 + \theta_1 < 1$[VI].

II.i.c. The GARCH with Exogenous Variable (GARCHX) model:

The second model used in the hybridization method after the GARCH model is GARCH with Exogenous Variable (GARCHX). Expression (4) of GARCH(r,s) model is typically extended to be more complex and involve an additional explanatory variable known as exogenous variable $x_t$ in volatility equation. The volatility equation (5) can be extended and rewritten for the GARCHX(1,1,1) model and represented as in Eq. (6) [21,12]:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \theta_1 h_{t-1} + \omega_1 x_{t}^2$$

(6)

for exogenous variable $x_t$ which is squared to ensure that ($h_t > 0$). The including of the additional exogenous variable $x_t$ helps to explain the volatilities of exchange rate series and tend to lead to improve in-sample fit and out-of-sample forecasting perform.

II.i.d. Building a GARCH forecasting model for a hybridization method:

ARCH and GARCH models have become important tools in the analysis of time-series data. The building process of GARCH forecasting model usually includes the following steps:

a. The first step of the test in the GARCH effect is to estimate the mean equation (1) to obtain the estimated residual time series and then run them on lagged squared terms is to check for ARCH effect of the squared residual series which is known as the ARCH test [XIV]. It also was used BDS test for (Brock, Dechert, and Scheinkman[IV]) to check the nonlinearity test for extracted residuals of the fitted ARMAX linear model.

b. the second is a GARCH model identification, to identify a proper model, appropriate r and s should be carefully chosen, according to (AIC, BIC, HQ) In our study, the GARCH (1,1) and GARCHX (1,1,1) outperforms the GARCH models with larger r and s.

c. The third step is a parameters estimation of the hybrid model; in this study a two-step procedure can be adapted to estimate the vector of unknown parameters:

1. Step one: Estimate the parameters of the mean equation in Eq. (1). In this situation we used the (Ordinary Least Square Method (OLS),
Recursive Least Square Method with Exponential Forgetting Factor (RLS-EF), RLS with Kalman Filter Factor (RLS-KF) \cite{18,15,22}.

2. Step two: From residuals of Eq. (1), estimate the parameters of the GARCH and GARCHX model using the Maximum Likelihood Method.

3. After validating model efficiency, apply the GARCH and GARCHX model to forecast volatility.

II.i.i.e. GARCH models specified under heavy-tailed distributions:

The other main purpose of this study applied heavy-tailed distributions for (ARMAX-GARCH) and (ARMAX-GARCHX) hybrids models. There is a need to study the effect of distributions the GARCH innovations. To captures volatility on financial series, the studies applied the GARCH model, and The GARCH process exists on the assumption of normal distribution. Normal distribution was found to be not useful in capturing the tail behavior of the series. Therefore, Bollerslev (1987) and Nelson (1991) proposed Student t distribution and GED distribution to capture the long-tail behavior of the process. In practice, error term $\eta_t$ of Eq. (4) is assumed to follow the normal distribution or non-normal distributions. These non-normal distributions have been proved to perform well in modeling the fatter tails (leptokurtic) observed in GARCH residuals in several studies \cite{7,9,24}. Thus, to obtain more forecast efficient of hybrid models, the hybrid ARMAX-GARCH and ARMAX-GARCHX models imply on Student t distribution proposed in Bollerslev (1987) and Generalized Error Distribution (GED) by Nelson (1991)\cite{2}\cite{XIX}\cite{XXIV}:

- The standardized Normal distribution is:
  From Eq. (4) the error term can be rewritten as: $\eta_t = \frac{\epsilon_t}{\sqrt{h_t}} \sim N(0,1)$, When applied standardized Normal distribution to the GARCH model, the corresponding density functions of $\epsilon_t$ are described below:

$$f(\epsilon_t) = \frac{1}{\sqrt{2\pi h_t}} \exp \left( -\frac{1}{2} \frac{\epsilon_t^2}{h_t} \right), \quad \epsilon_t \sim N(0,1),$$

- The standardized Student t distribution proposed in Bollerslev (1987) is given as:

$$f(\epsilon_t) = \frac{\Gamma \left( \frac{\nu+1}{2} \right)}{\Gamma \left( \frac{\nu}{2} \right) \sqrt{\pi (\nu-2)h_t}} \times \left[ 1 + \frac{\epsilon_t^2}{(\nu-2)h_t} \right]^{-(\nu+1)/2}$$

$\nu$: degree of freedom ($2 < \nu < \infty$).

- The standardized GED proposed in Nelson (1991) is given as:

$$f(\epsilon_t) = \frac{\nu \exp \left[ -\frac{1}{2} |z/\lambda|^\nu \right]}{\lambda \left[ 2 \left( 1 + \frac{1}{\nu} \Gamma \left( \frac{1}{\nu} \right) \right) \right]}, \quad (-\infty < z < \infty),$$

Where $\lambda$ define as:

Copyright reserved © J. Mech. Cont. & Math. Sci. 
Firas Ahmed Mohammed et al
$$\lambda = \left[ 2^{(-2/\nu)} \Gamma \left( \frac{1}{\nu} \right) \Gamma \left( \frac{3}{\nu} \right) \right]^{1/2}$$

$\nu$: Tail thickness parameter ($0 < \nu \leq \infty$), $\Gamma$: Gamma function.

And this distributional specification makes hybrid (ARMAX-GARCH) and (ARMAX-GARCHX) models special among other time series models.

II.iii. The Zhang hybrid methodology:

The hybrid method supposes that the time series process decomposes as a mixture of both linear and non-linear components. This follows the Zhang (2001)[XXVI] hybrid approach, consequently, the relationship between linear and nonlinear components can be expressed as follows:

$$y_t = L_t + N_t \quad \rightarrow \quad \text{for ARMAX} - \text{GARCH}$$

(7)

$$y_t = L_t + N_t \quad \rightarrow \quad \text{for ARMAX} - \text{GARCHX} \quad ,$$

(8)

Where $L_t$ and $N_t$ represent the linear and nonlinear components present in the time series data, these two components are to be estimated for ARMAX and (GARCH, GARCHX) models respectively. This hybrid approach of combining forecasting values for our models (ARMAX-GARCH) and (ARMAX-GARCHX) has the following steps[16,26]:

1- First, fitted a linear ARMAX time series model for the data.

2- In the second step, the residuals time series are extracted from the fitted ARMAX linear model. The residuals will contain only the nonlinear components. Let $e_t$ denotes the residual at the time $t$ from the linear model, then:

$$e_t = y_t - \hat{L}_t$$

(9)

Where $\hat{L}_t$: The optimal $(\ell - \text{step})$ ahead forecast of the meanequation (1) for the ARMAX model, which can be represented as in Eq.(10)[XXIII]:

$$\hat{L}_t(\ell) = E(y_{t+\ell}) = \sum_{i=1}^{p} \phi_i y_{t+i} - \sum_{j=1}^{q} \theta_j \hat{y}_{t-j} + \sum_{k=1}^{q} \phi_k x_{t+\ell-k}, (\ell = 1, 2, ... )$$

(10)

Where:

$$\hat{y}_t(\ell - i) = y_{t+\ell-i} \quad \text{if} \quad \ell - i \leq 0 \quad \text{or} \quad i \geq \ell$$

$$\hat{e}_t(\ell - j) = \begin{cases} e_{t+\ell-j} & j \geq \ell \\ 0 & j = 0, 1, ..., \ell - 1 \end{cases}$$

1- Use BDS and ARCH tests for the diagnosis of residuals to check for nonlinearities structures left in the residuals.

2- If residuals confirm the nonlinearity, then the residuals are modeled using a nonlinear GARCH and GARCHX models. Subsequently, obtain the forecast values $\hat{N}_t$ for the residual series use the optimal nonlinear GARCH and GARCHX models. When modeled using a GARCH model, the optimal $(\ell - \text{step})$ ahead forecast of the volatility equation (5) for the GARCH (1,1) model can be represented as in Eq. (11)[VIII]:

Copyright reserved © J. Mech. Cont. & Math. Sci.
Firas Ahmed Mohammed et al
In the final step, the linear $\tilde{L}_t(\ell)$ and nonlinear $\tilde{N}_t(\ell)$ forecasted components are combined to obtain the pooled forecast values for the hybrid model (ARMAX-GARCH) as in Eq.(12):

$$\hat{y}_t(\ell) = \tilde{L}_t(\ell)_{\text{ARMAX}} + \tilde{N}_t(\ell)_{\text{GARCH}}$$

And the pooled forecast values for the hybrid model (ARMAX-GARCHX) as in Eq. (13):

$$\hat{y}_t(\ell) = \tilde{L}_t(\ell)_{\text{ARMAX}} + \tilde{N}_t(\ell)_{\text{GARCHX}}$$

The hybrid approach for (ARMAX-GARCH) model can be graphically represented as in Fig. 1.

And the hybrid approach for (ARMAX-GARCHX) model can be graphically represented as in Fig. 2:

**II.iv. Forecasts Evaluation**

Forecasting is a necessary purpose of this study. Given a sample of actual data, it is usually divided into a training (Calibration) dataset and test (Validation) dataset, the training dataset is used to build the model and estimate the parameters. The model then is applied to the test...
dataset to check its performance. Consequently, the comparison process between the efficient of individual and hybrid models in the predicting of the future unemployment rate values depends on the following forecasts evaluation criteria:

- **Mean Square Error (MSE):**
  \[
  \text{MSE} = \frac{1}{n} \sum_{t=1}^{n} (e_t)^2 = \frac{1}{n} \sum_{t=1}^{n} (y_t - \hat{y}_t)^2
  \] (14)

- **Mean Absolute Error (MAE):**
  \[
  \text{MAE} = \frac{1}{n} \sum_{t=1}^{n} |e_t| = \frac{1}{n} \sum_{t=1}^{n} |y_t - \hat{y}_t|
  \] (15)

- **Mean Absolute Percentage Error (MAPE):**
  \[
  \text{MAPE} = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{e_t}{y_t} \right| \times 100\%
  \] (16)

Where:
- \( e_t \): Forecast error.
- \( y_t \): Denote the actual value of unemployment rate at time \( t \).
- \( \hat{y}_t \): Denote the forecast value of unemployment rate at time \( t \).
- \( n \): is the sample size for the hold out data. In the present investigation \( n \) is (12, months).

### III. Case Study

This research mainly focused on applying hybrid approaches with three different distribution models. Hybrid approaches were tested on the actual and residuals datasets, and then use the individual and hybrid models to a future forecast and comparing them using (MSE, MAE, MAPE) to choose the optimum model.

#### III.i. Data sets in the experiment

The data for this study represent a bivariate time-series data of the unemployment rate and the exchange rate in the USA, as monthly measurements for the period from (January 2000 to December 2017) as a training dataset for parameter estimation and then model building and the last twelve observations of the 2018 year (from January 2018 to December 2018) considered as a testing set is used for obtaining the out-of-sample forecast and also for validation (forecast accuracy evaluation). collected from (Source; fred.stlouisfed.org).

#### III.ii. Empirical results

In this section, we build the linear ARMAX model based on actual data and building the nonlinear GARCH and GARCHX models based on the residuals time series are extracted from the fitted ARMAX linear model. Figure 3a and b show the plot of bivariate time-series of the unemployment rate and the exchange rate:
Fig. 3. The monthly a. Unemployment rate and b. The exchange rate for original series from (January 2000 to December 2017)

III.ii.a Fitting of Linear ARMAX($n_p, n_q, n_r, n_k$) models

The suitable ARMAX model was selected based on minimum AIC, BIC and FPE criteria and observing the significance of autocorrelation (ACF), partial autocorrelation (PACF), extended autocorrelation function (EACF) and cross-correlation (CCF) functions to identify the initial model 4. And then choosing the best ARMAX model based on the over-fitting and under-fitting models diagnosing. Accordingly, ARMAX (2,1,1,0) mode is selected for modeling the dynamic relation between the unemployment rate and the exchange rate. The methods (OLS, RLS-EF, and RLS-KF) of parameter estimates of the fitted ARMAX (2,1,1,0) model along with its parameter estimates and Root mean square error (RMSE) represented as in table. (1):

| Method of Estimation | Estimated values of ARMAX (2, 1, 1, 0) model parameters | Error measure |
|----------------------|---------------------------------------------------------|---------------|
| OLS                  | $\phi_1 = -0.06111$, $\phi_2 = -0.05856$, $\theta_1 = -0.04793$, $\phi_1 = 0.05871$ | MSE = 5.5296 |
| RLS – EF             | $\phi_1 = -0.5391$, $\phi_2 = -0.4329$, $\theta_1 = -0.1341$, $\phi_1 = 0.1182$ | MSE = 0.0282 |
| RLS – KF             | $\phi_1 = -0.3967$, $\phi_2 = -0.4017$, $\theta_1 = 0.2525$, $\phi_1 = 0.6948$ | MSE = 0.3209 |

*Cross-correlation functions (CCF) implementation using the SAS 9.1 statistical package and it is found that delay time equal to zero.*

Copyright reserved © J. Mech. Cont.& Math. Sci.

Firas Ahmmed Mohammed et al
From the table (1), We conclude that the Recursive Least Square Method with Exponential Forgetting Factor (RLS-EF) is the optimum method in estimation compared with OLS and RLS-KF methods according to minimum (MSE=0.0282). The parameter estimates of fitted ARMAX (2,1,1,0) model using RLS-EF method are furnished in table (2) along with their Standard error, t-value, significance level (p-value):

| Mean equation | AR(1) | φ₁ | -0.5391 | 0.00793 | 67.9823 | 0.00001 |
|---------------|-------|----|---------|---------|---------|---------|
| ARMAX (2, 1, 1, 0) | AR(2) | φ₂ | -0.4329 | 0.00791 | 54.7282 | 0.00001 |
| MA(1)        | θ₁   | -0.1341 | 0.01030 | 13.0194 | 0.00001 |
| x(1)         | φ₁   | 0.1182 | 0.00945 | 12.5079 | 0.0001  |

φ₁ denote parameter of exogenous variable. And according to the (p-value < 0.05), all parameters are significant and have an effect in the model. And the polynomial equations as discussed in Section 2.1 of the ARMAX(2,1,1,0) model can be written from above table as:

\[ \Phi (L) = 1 - 0.5391L - 0.4329L^2 \]
\[ \Theta (L) = 1 + 0.1341L \]
\[ \phi (L) = 0.1182 \]

III.iib. Testing for ARCH effects (ARCH-Test)

ARCH-test was performed on the squared residuals of the optimum ARMAX(2,1,1,0) model to test the null hypothesis (Ho: that no ARCH-effects exist). Engle’s ARCH-test performed before the estimation of the GARCH model, to supports that there exist heteroskedasticity and the result of this test expressed in table 3:

| Test | ARCH (Lag) | LM – Statistic | p – value |
|------|------------|----------------|-----------|
| ARCH | (k = 10)   | 174.94         |           |
| ARCH | (k = 15)   | 172.31         |           |
| ARCH | (k = 20)   | 168.19         |           |

ARCH-effects are tested on lags (10-20). With these p-values equal to zero we reject the hypothesis that no ARCH-effects exist, (the ARCH effect is significant up to 20 lags).

III.ii.c. Testing for nonlinearity (BDS-Test)

BDS test has been used to test the presence of any remaining nonlinearities structure in the residuals ARMAX(2,1,1,0) model to test the null hypothesis (Ho: linearity in \( \bar{\varepsilon}_t \) exist). And the result of this test expressed in table 4:
## Table (4). BDS test for nonlinearities of ARMAX (2,1,1,0) residuals

| Test | dimension | BDS – Statistic | p – value |
|------|-----------|----------------|----------|
| BDS  | m = 2     | 0.181994       | 0.0000   |
| BDS  | m = 3     | 0.305099       | 0.0000   |
| BDS  | m = 4     | 0.386709       | 0.0000   |

According to (p-value) the results of the BDS-test indicate to exist a nonlinear pattern in the residuals of ARMAX(2,1,1,0) model for different dimensions.

### 3.2.4. Fitting of ARMAX-GARCH and ARMAX-GARCHX hybrid models

Accordingly, for modeling and capture the heteroscedasticity and nonlinearity in the conditional variance of ARMAX, GARCH and GARCHX models are applied and then employed them to forecast volatility values. The suitable GARCH and GARCHX models were selected based on minimum AIC, BIC and HQ criteria and it is found that the GARCH(1,1) and GARCHX(1,1,1) are the best models and possible to use in hybridization methodology with ARMAX (2,1,1,0) model. The hybrid models with three different distributions (Normal, Student’s t and GED Distributions) are furnished in table (5) along with their AIC, BIC and HQ criteria:

### Table (5). Comparative criteria for selecting the best hybrid model (ARMAX-GARCH and ARMAX-GARCHX) with random distribution

| $\varepsilon_t$ Distribution | Hybrid Models | Criteria |
|------------------------------|---------------|----------|
|                              | ARMAX – GARCH | AIC | BIC | HQ |
| Normal Distribution          | ARMAX (2,1,1,0) – GARCH(1,1) | $-1.16354$ | $-1.116667$ | $-1.144606$ |
| Student Distribution          | ARMAX (2,1,1,0) – GARCH(1,1) | $-1.15427$ | $-1.091767$ | $-1.129020$ |
| Generalized Error Distribution| ARMAX (2,1,1,0) – GARCH(1,1) | $-1.40602$ | $-1.343519$ | $-1.380772$ |
| $\varepsilon_t$ Distribution | ARMAX – GARCHX | AIC | BIC | HQ |
| Normal Distribution          | ARMAX (2,1,1,0) – GARCHX(1,1,1) | $-1.16627$ | $-1.103767$ | $-1.141020$ |
| Student Distribution          | ARMAX (2,1,1,0) – GARCHX(1,1,1) | $-1.15365$ | $-1.075526$ | $-1.122092$ |
| Generalized Error Distribution| ARMAX (2,1,1,0) – GARCHX(1,1,1) | $-1.42846$ | $-1.350330$ | $-1.396899$ |

From the table (5) and according to the minimum values of AIC, BIC and HQ criteria the best forecasting hybrid models namely ARMAX (2,1,1,0) – GARCH(1,1) and ARMAX (2,1,1,0) – GARCHX(1,1,1) models error whose conditional variance follows a GED distribution.

### 3.2.5. Estimation of hybrid models

---

5Implementation using the MATLAB,2018a and EViews 9.

Copyright reserved © J. Mech. Cont. & Math. Sci.
Firas Ahmmed Mohammed et al

363
For building hybrid models, we use the two-step procedure of estimation. The estimation of the parameters of mean equation for ARMAX (2,1,1,0) model is given in section 3.2.1 (table (2)), and the estimation of the parameters of variance equation for the GARCH(1,1) and GARCHX(1,1,1) models as discussed in Section 2.2.2 and 2.2.3, when the whose error follow GED-distribution using maximum likelihood estimation are furnished in the table (6) along with their Standard error, t-value, significance level (p-value) :

\[ y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 \epsilon_{t-1} + \beta_4 x_t + \epsilon_t \]

\[ \epsilon_t = \eta_t \sqrt{h_t} \quad , \quad \eta_t \sim \text{GED}(\tilde{v} = 4.971) \]

The variance equations follow GARCH(1,1) and GARCHX(1,1,1) models from the table (6) as:

\[ h_t = 0.000061 + 0.609532 \epsilon_{t-1}^2 + 0.271026 h_{t-1} \quad , \quad \text{for GARCHX - GED} \]

\[ h_t = 0.002267 + 0.622778 \epsilon_{t-1}^2 + 0.241299 h_{t-1} + 0.002686 x_{t-1}^2 \quad , \quad \text{for GARCHX - GED} \]
The above equations represent the hybrid models and employed in the forecasting of the unemployment rate.

4. Evaluation of forecasting performance

The prediction abilities of the individual ARMAX model and the hybrid models namely ARMAX-GARCH-GED and ARMAX-GARCHX-GED are compared based on MSE, MAE, and MAPE mentioned in section 2.4 for last twelve months of the 2018 year (i.e. for last twelve observations). The Actual and forecast values obtained using individual ARMAX model using equation 10, ARMAX-GARCH-GED and ARMAX-GARCHX-GED hybrid models using equation 12 and 13 respectively as in section (2.3) of Zhang hybrid approach, this values are furnished in the table (7) along with their Forecasts Evaluation Criteria namely MSE, MAE, and MAPE of individual and hybrid models:

| Table (7). Actual and forecast values of individual and hybrid models along with forecasts evaluation criterion |
|---------------------------------------------------------------|
| **Forecast values**                                           |
| $\hat{y}_t(\ell) = \hat{L}_t + \hat{\bar{N}}_t$             |
| **ARMAX**                                                     |
| **Forecast values**                                           |
| $\hat{y}_t(\ell) = \hat{L}_t + \hat{N}_t$                   |
| **ARMAX-GARCH**                                               |
| **Forecast values**                                           |
| $\hat{y}_t(\ell) = \hat{L}_t + \hat{\bar{N}}_t$             |
| **ARMAX-GARCHX**                                              |
| **Jan/2018**                                                 |
| 4.1                                                          |
| 4.1591                                                       |
| 4.2162                                                       |
| 4.2152                                                       |
| **Feb/2018**                                                 |
| 4.1                                                          |
| 4.1416                                                       |
| 4.1977                                                       |
| 4.1967                                                       |
| **Mar/2018**                                                 |
| 4.0                                                          |
| 3.3575                                                       |
| 3.4148                                                       |
| 3.4138                                                       |
| **Apr/2018**                                                 |
| 3.9                                                          |
| 3.7569                                                       |
| 3.8059                                                       |
| 3.8169                                                       |
| **May/2018**                                                 |
| 3.8                                                          |
| 3.6261                                                       |
| 3.6731                                                       |
| 3.6751                                                       |
| **Jun/2018**                                                 |
| 4.0                                                          |
| 3.7342                                                       |
| 3.7716                                                       |
| 3.7729                                                       |
| **Jul/2018**                                                 |
| 3.9                                                          |
| 3.7326                                                       |
| 3.7647                                                       |
| 3.7553                                                       |
| **Aug/2018**                                                 |
| 3.8                                                          |
| 3.7728                                                       |
| 3.8040                                                       |
| 3.7946                                                       |
| **Sep/2018**                                                 |
| 3.7                                                          |
| 3.8026                                                       |
| 3.8310                                                       |
| 3.8214                                                       |
| **Oct/2018**                                                 |
| 3.8                                                          |
| 3.8206                                                       |
| 3.8474                                                       |
| 3.8376                                                       |
| **Nov/2018**                                                 |
| 3.7                                                          |
| 3.8551                                                       |
| 3.8705                                                       |
| 3.8697                                                       |
| **Dec/2018**                                                 |
| 3.7                                                          |
| 3.8497                                                       |
| 3.8649                                                       |
| 3.8644                                                       |
| **Forecasts Evaluation Criteria**                             |
| MSE = 0.0521                                                  |
| MAE = 0.1624                                                 |
| MAPE = 4.16%                                                 |
| MSE = 0.0447                                                  |
| MAE = 0.1585                                                 |
| MAPE = 4.061%                                                 |
| MSE = 0.0444                                                  |
| MAE = 0.1564                                                 |
| MAPE = 4.006%                                                 |

We conclude from the above table. The hybrid models are more efficient than the individual model in forecasting according to the minimum values of MSE, MAE and MAPE criteria. And the following graph shows the actual and forecasts values using the hybrid model.
ARMAX(2,1,1,0)-GARCHX(1,1,1) for twelve future values of the unemployment variable from January to December (2018):

Fig. (4) Actual and forecasts values using ARMAX (2,1,1,0)-GARCHX(1,1,1)-GED for 12 months for the unemployment rate.

V. Conclusions

The main purpose of this study uses a hybrid method with different distributions that decomposes bivariate time series into its linearity and nonlinearity part and then modeling each part individually before they are aggregate for getting final forecast using Zhang hybrid methodology. And Based on the results obtained in this study for analysis bivariate time series using individual and hybrid models one can conclude the following outcomes:

1. The classical linear time series models such as ARMAX are not always sufficient for modeling bivariate time series that consists of linear and nonlinear structures. And when the assumptions of linearity and homoscedastic error variance are violated which are two most critical assumptions of linear time series model (ARMAX model). Therefore, the hybrid approach which combines linear and nonlinear models performs better as compared to classical time series models under Heteroscedasticity problem.

2. The residuals time series are extracted from the fitted ARMAX linear model was tested using ARCH and BDS tests which reveals that nonlinearity pattern exists in the residual time series. And then based on this residual series, the nonlinear GARCH and GARCHX models are building and employed them to forecast the volatility and in a hybridization method with ARMAX model.

3. The individual and hybrid models have been applied in forecasting the unemployment rate in the USA. And the comparison of forecasts performance between the individual ARMAX model, ARMAX-GARCH and ARMAX-GARCHX hybrid models have been accomplishing, it is
obvious that the hybrid models can be an effective way to improve forecasting efficiency and perform better than the individual ARMAX model according to minimum MSE, MAE, and MAPE values as mentioned in table (7).

4. And the process of including the exogenous variable in the volatility equation for GARCH model and build GARCHX model gave more improved performance in forecasting and this evident from the lowest values of the evaluation criterion of ARMAX-GARCHX model when error whose conditional variance follow general error distribution (GED), and yield the convergence between actual and forecast values as mentioned in table (7) and Fig (2).

5. Finally, the residuals from the finally fitted ARMAX-GARCH and ARMAX-GARCHX hybrid models that follow GED distribution are examined and it is found that no ARCH effect exists and distributed as i.i.d, which are the main assumptions of ensuring the sufficiency of the model selected.

References:

I. Aldem migli, A and Hapoğlu, H.(2015)." Comparison of ARMAX Model Identification Results Based on Least Squares Method". IJMTER, Vol. (02), No.(10), PP(27-35).

II. 2Bollerslev, T. (1987). "A conditional heteroscedastic time series model for speculative prices and rates of return". Review of Economics and Statistics, 69, 542-547.

III. 3Bollerslev, Tim,(1986). "Generalized autoregressive conditional heteroscedasticity,". Journal of Econometrics, Vol. (31), No (3), pp (307-327).

IV. 4Brock,.W. A., Dechert,.W., Scheinkman,.J., and LeBaron,.B. (1996)."A test for independence based on the correlation dimension". Economic Reviews, Vol(15),No.(3).pp(197–235).

V. 5Engle, R. (1982)."Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation". Econometrica ,Vol.(50),No(4),pp (987-1007).

VI. 6Engle, R. (2001). "GARCH 101: The use of ARCH/GARCH models in applied econometrics". Journal of Economic Perspective, Vol. (15), No. (4), pp.(157-168).

VII. 7Feng, L., & Shi, Y. (2017). "A simulation study on the distributions of disturbances in the GARCH model". Cogent Economics and Finance, Vol.(5), No.(1), pp (1-19).

VIII. 8Franses, P. H., van Dijk, D. J. C., and Opschoor, A. (2014). "Time Series Models for Business and Economic Forecasting", 2nd Edition. Cambridge University Press.

IX. 9Gao,Y. Zhang,.C and Zhang, L . (2012). "Comparison of GARCH Models based on Different Distributions". Journal of Computers, VOL. (7), NO. (8), pp (1967-1973).
X. 10George E. P. Box; Gwilym M. Jenkins; Gregory C. Reinsel; Greta M. Ljung. (2015). "Time Series Analysis Forecasting and Control", Fifth Edition, John Wiley & Sons Inc. Hoboken, New Jersey

XI. 11Gooijer, J. G. D., & Hyndman, R. J. (2006). "25 years of time-series forecasting. International Journal of Forecasting". Vol.(22),No.(3), pp (443-73).

XII. 12Han, H., and Kristensen, D. (2014). "Asymptotic Theory for the QMLE in GARCH-X Models With Stationary and Nonstationary Covariates". Journal of Business and Economic Statistics , Vol.(32),No.(3), pp(416–429).

XIII. 13Hickey, E., Loomis, D. G., & Mohammadi, H. (2012). "Forecasting hourly electricity prices using ARMAX–GARCH models: An application to MISO hubs". Energy Economics, Vol (34), No(1), pp(307–315).

XIV. 14Lee, J.H.H., (1996), "A Lagrange Multiplier Test for GARCH models". Econometric Letters, Vol. (37), pp (256-271).

XV. 15Ljung, L., (1999)."System Identification Theory for user",2nd ed. Prentice Hall Upper Saddle River N.J. London UK.

XVI. 16Mitra, D., & Paul, R. K. (2017). Hybrid time-series models for forecasting agricultural commodity prices. Model Assisted Statistics and Applications, Vol.(12),No(3), pp(255–264).

XVII. 17Moeeni, H., & Bonakdari, H. (2017). Impact of Normalization and Input on ARMAX-ANN Model Performance in Suspended Sediment Load Prediction. Water Resources Management, Vol(32), No(3), pp(845–863).

XVIII. 18 , O. (2001). "Nonlinear System Identification from Classical Approaches to Neural Networks and Fuzzy Models". Springer, New York.

XIX. 19Nelson, D. B. (1991). Conditional heteroscedasticity in asset returns: A new approach. Econometrica, Vol(59), pp(347-370).

XX. 20Porshnev, A., Valeria ,L. and Ilya,R,(2016). "Could Emotional Markers in Twitter Posts Add Information to the Stock Market ARMAX-GARCH Model". Higher School of Economics Research Paper No. WP BRP 54/FE.

XXI. 21Rachev, S., Mittnik, S., Fabozzi, F., Focardi, S., and Jasic, T. (2007). "Financial econometrics: From basics to advanced modeling techniques". John Wiley & Sons, Inc. New York.

XXII. 22Soderstrom, T. and Stoica, P. (2001) ."System Identification" Prentice-Hall International, Hemel Hempstead, U.K.

XXIII. 23Tsay, Ruey S.(2013). "An introduction to analysis of financial data with R". John Wiley & Sons .Hoboken.

XXIV. 24Yaya, O.S., Olubusoye, O.E. and Ojo, O.O. (2014). "Estimates and Forecasts of GARCH Model under Misspecified Probability Distributions: A Monte Carlo Simulation Approach". Journal of Modern Applied Statistical Methods. Vol (13), No (2). pp (479-492).

XXV. 25Yiu,J and Wang,S.(2007)."Multiple ARMAX modeling scheme for forecasting air conditioning system performance". Energy Conversion and Management, Vol (48),No.(8),pp(2276–2285).

XXVI. 26Zhang, G.P. (2003). "Time series forecasting using a hybrid ARIMA and neural network model". Neurocomputing. Comput., Vol.(50), pp(159-175) .