

A Correcting Note on Forecasting Conditional Variance Using ARIMA vs. GARCH Model

Mohammad Naim Azimi¹ & Seyed Farhad Shahidzada¹

¹ Faculty of Economics, Kabul University, Kabul, Afghanistan

Correspondence: Mohammad Naim Azimi, Faculty of Economics, Kabul University, Kabul, 25000, Afghanistan.
Tel: 93-790-060-073. E-mail: naeem.azimi@ku.edu.af

Received: April 6, 2019

Accepted: April 28, 2019

Online Published: April 30, 2019

doi:10.5539/ijef.v11n5p145

URL: <https://doi.org/10.5539/ijef.v11n5p145>

Abstract

In this study, we demonstrate that a common approach in using the Autoregressive Integrated Moving Average model is not efficient to forecast all types of time series data and most specially, the out-of-sample forecasting of the time series that exhibits clustering volatility. This gap leads to introduce a competing model to catch up with the clustering volatility and conditional variance for which, we empirically document the efficient and lower error use of the Generalized Autoregressive Conditional Heteroscedasticity model instead.

Keywords: ARIMA, GARCH, root mean squared error, conditional variance, forecasting

1. Introduction

In forecasting time series variables, it is imperative to analyze the pattern of series as this is a matter that remains at forefront of all the academic research in financial econometrics by finding the least square based -forecast of the time series to be biased and inefficient. The dynamic behavior of the time series variables exhibiting clustering volatility, issue greater challenges in accurately estimating and forecasting the time series. It is rare to find studies concentrating on comparative analysis of competing models in financial econometrics that cover both theoretical and empirical findings. Although a vast academic literature shows that forecasting of time series data has commonly focused on linear and stationarity assumptions (Tealab, Hefny, & Badr, 2017) in using the common models like Autoregressive Integrated Moving Average (ARIMA), Autoregressive Conditional Heteroscedasticity (ARCH), Generalized Autoregressive Conditional Heteroscedasticity (GARCH) and much commonly the Vector Autoregressive (VAR) model (Zhang, 2003) to forecast the time series variables, a comparative study on the empirical result of ARIMA and GARCH models in terms of variance forecast and optimal estimation of time series variables that exhibit volatility is still left untouched. Gunter and Önder (2015) compare the Bayesian and Classical ARMA and ADLM in forecasting the international city tourism demand on a set of panel data while Khandelwal, Adhikari, and Verma (2015) use ARIMA and Artificial Neural Network modeling to separately recognize and predict the reconstructed detailed components. Many other researchers attempt in determination of models in forecasting both the linear and nonlinear time series variables using ARIMA, VAR or alternatively the hybrid methodology but ignore the least conditional variance forecast optimization and accuracy for time series data that present clustering volatility throughout the period of concern (see for instance, Aladag, Egrioglu, & Kadilar, 2009; Contreras, Espínola, Nogales, & Conejo, 2003; Khashei & Bijari, 2012; Chen & Wang, 2007). Interestingly, few other authors present comparative studies on ARMA and ARIMA for forecasting but overlooking the pattern of time series movement across the time in accounting for volatility before attempting the accuracy of its out-of-sample forecasting (see for instance, Valipour, Banihabib, & Behbahani, 2013; Nochai, 2006). Among a sheer number of studies, Liu and Shi (2013) suggest the ARMA-GARCH model for accurate prediction of mean and volatility of electricity pricing in managing the relevant risk associated with bid or hedge forecast while, Vilasuso (2002) argues on the accuracy of GARCH and suggests to apply Integrated Generalized Autoregressive Conditional Heteroscedasticity (IGARCH) for accurate out-of-sample forecasting of a times series variable instead. Laterally, the IGARCH model is a non-stationary GARCH model and almost a similar one to ARIMA in which, the persistence impact of the previous shocks $\eta_{t-i} = \sigma_{t-i}^2$ for $i > 0$ on σ_t^2 is clearly evidenced (Caporale, Pittis, & Spagnolo, 2003). In passing, one should note that not all types of time series data fit the least square and conditional heteroscedasticity on which, day-ahead

forecast can be based and thus, it deserves a greater care and due conservatism. In light of the empirical gaps stated above, the overarching objective of this paper is to answer the following main question:

Does ARIMA (p, d, q) used as a common approach for forecasting the time series account for volatility and conditional variance while attempting out-of-sample forecasting or GARCH (1,1) can be applied as an accurate model instead? This question will be answered in light of the approach to catch up with volatility of time series variables and experimental data simulation. The originality of this paper is due to its comparative analysis among almost all the literature to date in financial econometric modeling. The rest of this paper is organized as: section 2 offers the theoretical framework of the comparative model for forecasting; section 3 presents the simulative data on which the empirical analysis is based; section 4 presents the paper conclusion that is followed by acknowledgement and the list of references used in this study.

2. Comparative Modeling

2.1 ARIMA Approach

Literally, the proportion of AR and MA model is to be integrated with differencing the non-stationary time series for which, the ARIMA(p, d, q) equation for forecasting can be expressed as follow:

$$\hat{Y}_t = \mu + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (1)$$

where ϕ s are the parameters of the AR process and θ s are those of the moving average process (see for instance, (Wooldridge, 2015)). In finding the appropriate ARIMA model for non-stationary time series forecasting, we should keep differencing (d) in stationarizing the series for removing the gross feature of seasonality in the series as if $d = 0$: $y_t = Y_t$, $d = 1$: $y_t = Y_t - Y_{t-1}$, $d = 2$: $y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2})$ (Mills & Markellos, 2008) while, ϕ and higher – order of MA coefficients require a higher order of differencing the series though, non-invertibility requires otherwise. Such convention is empirically proved to be worse for post-sample forecasting (Asteriou & Hall, 2011).

2.2 GARCH Approach

In contrast, Engle (1982) introduced a more powerful model that accounts for volatility of the time series and that allows for past conditional variance in the current conditional variance equation and an example relating to uncertainty of the inflation rate was presented by (Bollerslev, 1986). In simplest form, to keep the mean equation (2) the same, we can modify the variance equation as:

$$\sigma_t^2 = \lambda_0 + \lambda_1 u_{t-1}^2 + \lambda_2 \sigma_{t-1}^2 \quad (2)$$

As a matter of fact, uncertainty is disliked by almost all investors so the conditional variance and volatility should be forecasted for the period ahead, that is:

$$\sigma_{t+1}^2 = \lambda_0 + \lambda_1 u_t^2 + \lambda_2 \sigma_t^2 \quad (3)$$

Therefore, (3) can easily be generalized to give forecast volatility for j period or period ahead forecast (Gujarati, 2004) as:

$$\sigma_{t+j}^2 = \lambda_0 + (\lambda_1 + \lambda_2) \sigma_{t+j-1}^2 \quad (4)$$

But in long periods, the steady state variance can be obtained by equating all the variance terms as follow:

$$\sigma^2 = \frac{\lambda_0}{(1 - \lambda_1 - \lambda_2)} \quad (5)$$

Since, almost a major set of financial time series data exhibit volatility, in theoretical comparison, the latter is a powerful model in estimating and forecasting the time series data while the former can only be used under rare empirical assumptions (also see, Ardia & Hoogerheide, 2019; Engle & Patton, 2001; Christoffersen & Jacobs, 2002). In addition to what the literal comparison offers the strength of a model to forecast the out-of-sample time series, we use the following two methods of error forecast comparison 1) the Root Mean Squared Error (RMSE); and 2) Mean Absolute Error (MAE) to support our both theoretical and empirical claims:

$$RMSE = \left(m^{-1} \sum_{h=0}^{m-1} \hat{e}_{n+h+1}^2 \right)^{1/2} \quad (6)$$

which is the sample standard deviation of the forecast error without any adjustment for the degree of freedom and the MAE which is expressed as follow:

$$MAE = m^{-1} \sum_{h=0}^{m-1} |\hat{e}_{n+h+1}| \quad (7)$$

that is preferred to obtain a smaller MAE in competing model for forecasting.

3. Estimating Alternative

3.1 Data

For empirical proof of the comparative estimation and forecasting of a time series variable and choosing between the least variance and statistically optimal value, we use a set of time series data relevant to the United State Economy on Unemployment Rate (UR_t) and Percentage Change in Unemployment Rate (ΔUR_t) which is arranged on daily basis from 2016-09-02 to 2019-02-01. The data is retrieved from the U.S. Bureau of Labor Statistics 02.02.2019 and the pattern of the rate both for actual and percentage change is graphically shown in figure 1 and figure 2. The mean value of the unemployment rate throughout the stated period is shown to be 5.8 percent while the percentage change in the unemployment rate of the U.S. is 2.5 percent increase in average. The standard deviation of the unemployment rate and the percentage change in unemployment rate is 1.633 and 24.89 respectively. This paper uses both variables for analysis but the unemployment rate for day-ahead-forecast only.

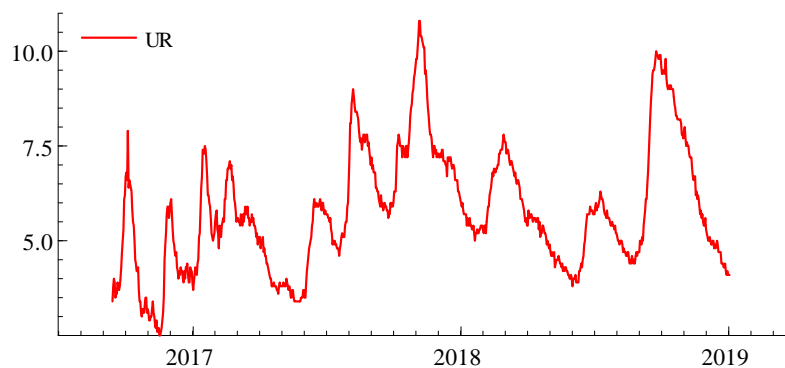


Figure 1. $UR_t, t = 1, 2, \dots, n$

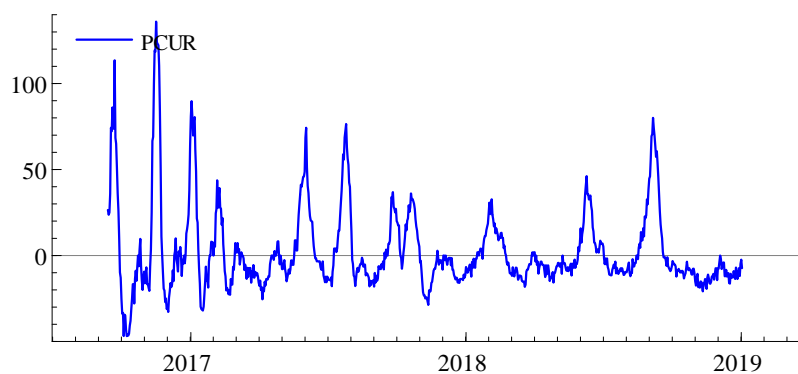


Figure 2. $\Delta UR_{t-1}, 1 = \text{year ago}$

Figure 1 presents the percentage of civilian unemployment rate through the period under study while figure 2 presents the percentage change in the unemployment rate on annual basis with a mean value of 1.78 percent increase throughout the stated period.

3.2 Data Analysis and Findings

This section presents the estimation of simulative data considering both ARIMA and GARCH models.

3.2.1 ARIMA Model

Figure 3 shows the forecast values obtained from [1] for 30 days ahead. The forecast pattern indicates an increase in the Unemployment Rate of the U.S. based on the actual data of the previous days with $n=842$ observation in the sample data.

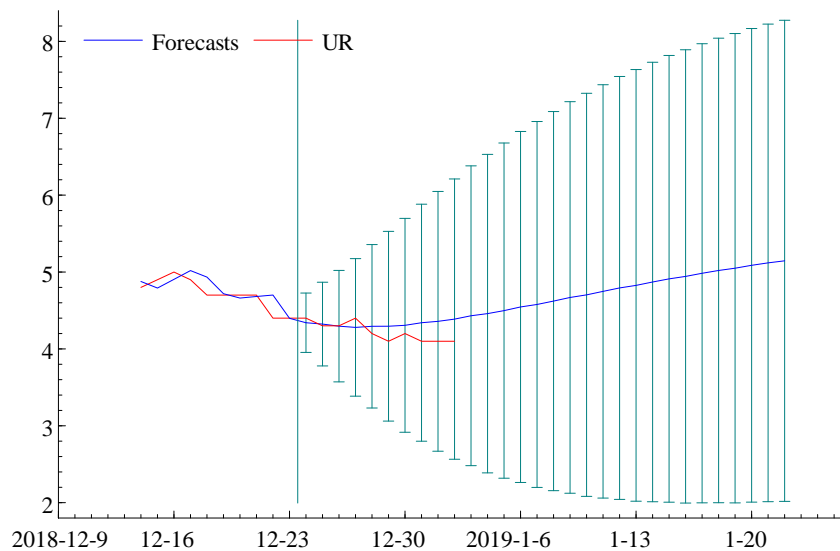


Figure 3. ARIMA day-ahead forecast

The estimated forecast value for 30 days ahead by ARIMA shows that the unemployment rate shall increase by 0.98 percent while the actual unemployment rate for the stated period shows an invert behavior and decrease by 1 percent (see, table 1 for calculation and United States Unemployment Rate).

3.2 GARCH Model

Here, the forecast values obtained from the GARCH model is presented in figure 4 where it consists of three parts for conditional mean forecast [fig. 4-A], conditional variance forecast [fig. 4-B] and VaR Confidence both for 0.05 and 0.95 [fig. 4-C].

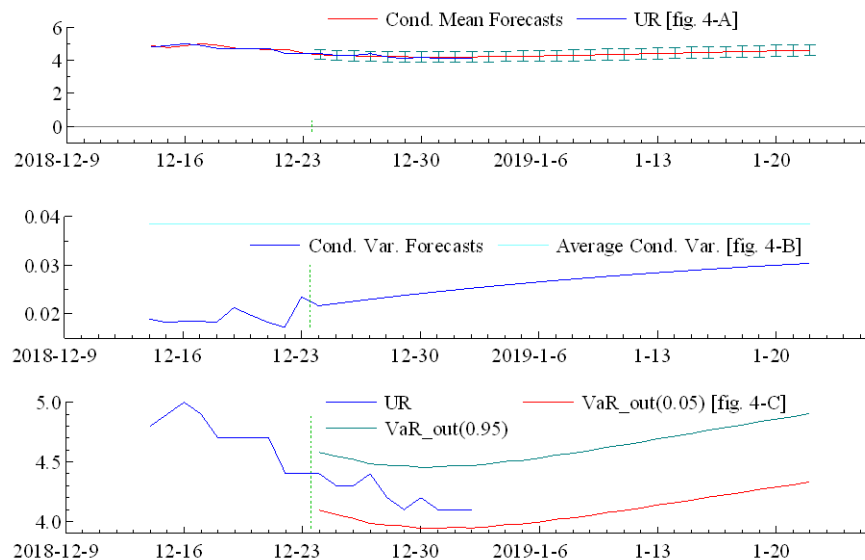


Figure 4. GARCH day-ahead-forecast

This figure (our current interest is fig. 4-B) shows that the forecast values obtained from conditional variance GARCH model using the set of simulative data for the United States unemployment rate is decreasing for the 30 days ahead forecast while its conditional variance increases by 0.03 percent. The decrease in forecast value is by 0.25 percent as shown by actual unemployment rate for the stated days by the United States (see again, table 1 and United States Unemployment Rate).

Table 1. Comparative out-of-sample forecast

| Day-ahead Forecast | GARCH Model | | ARIMA Model | |
|-----------------------|--------------------------|----------------------|--------------------------|----------------------|
| | Forecast value UR_t | Std. Error UR_t | Forecast Value UR_t | Std. Error UR_t |
| 2019-01-03 | 4.337 | 0.1472 | 4.340 | 0.1930 |
| 2019-01-04 | 4.302 | 0.1487 | 4.323 | 0.2721 |
| 2019-01-05 | 4.274 | 0.1501 | 4.296 | 0.3623 |
| 2019-01-06 | 4.233 | 0.1515 | 4.279 | 0.4473 |
| 2019-01-07 | 4.221 | 0.1529 | 4.294 | 0.5312 |
| 2019-01-08 | 4.217 | 0.1542 | 4.295 | 0.6169 |
| 2019-01-09 | 4.198 | 0.1554 | 4.307 | 0.6953 |
| 2019-01-10 | 4.201 | 0.1566 | 4.340 | 0.7707 |
| 2019-01-11 | 4.211 | 0.1577 | 4.358 | 0.8447 |
| 2019-01-12 | 4.207 | 0.1589 | 4.387 | 0.9114 |
| MES | -0.11401 | | -0.10230 | |
| RMSE | 0.12545 | | 0.16868 | |

Values presented in table 1 are the forecast values for the 10 days-ahead from January 03 to January 12, 2019 measured by GARCH and ARIMA Models. The comparative values show that GARCH offers much optimal forecast than the ARIMA one as the actual unemployment rate in January has been 4 percent and kept declining to 3.9 percent throughout the February 2019. The GARCH estimation shows a decrease in the forecast values as presented in figure 5 but the estimates of ARIMA indicates otherwise. The standard error of the ARIMA model forecast shows an incremental pattern while the GARCH conditional $std.Error = \sqrt{var}$ shows almost a constant pattern (see table and 1 and figure 5). We further use [6] and [7] for Root Mean Squared Error and Mean Error Squared respectively in selecting the model that exhibit lower error under which, GARCH shows lower error than ARIMA.

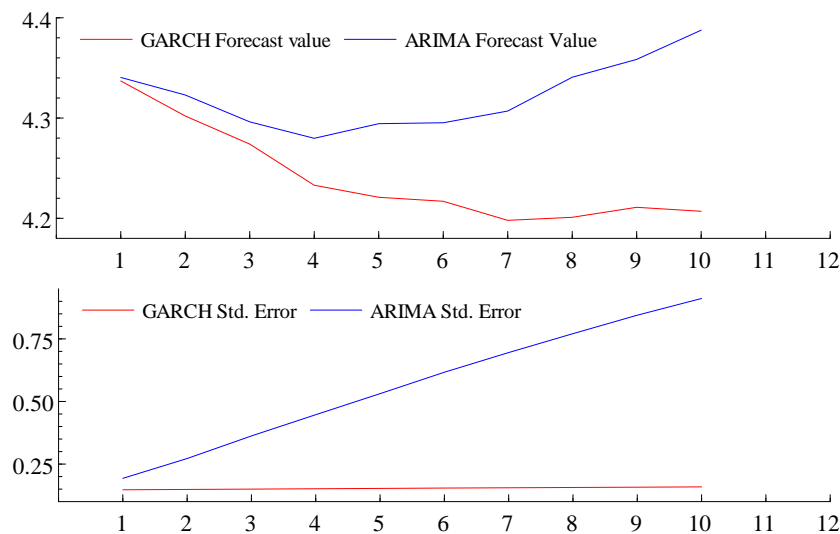


Figure 5. Comparison of forecast values obtained from GARCH and ARIMA models

Finally, we present the clustering volatility of the conditional variance estimated from the GARCH model in figure 6 to insist on the appropriate selection of the model and the accuracy of the forecast estimated by using the competing model of GARCH.

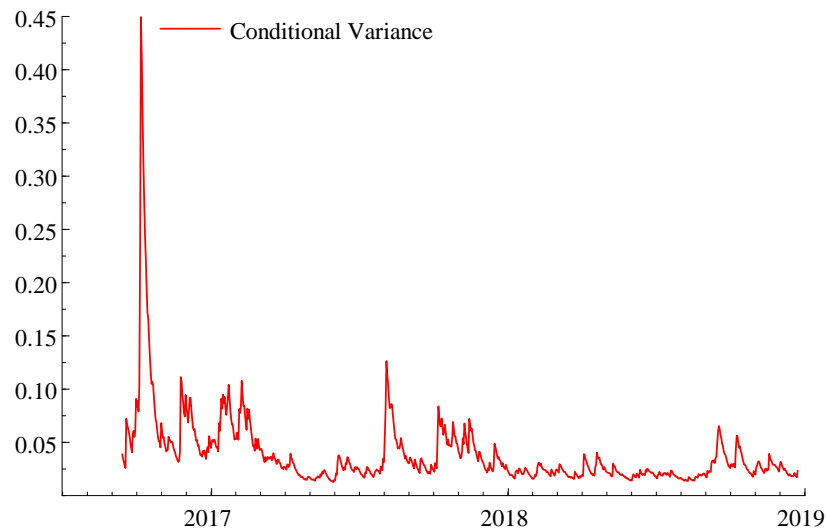


Figure 6. Conditional variance of GARCH

Figure 6 shows that the previous day's volatility influences the Unemployment Rate and fits the model with a nice magnitude indicating a p-value of <0.05 to significantly explain the variable.

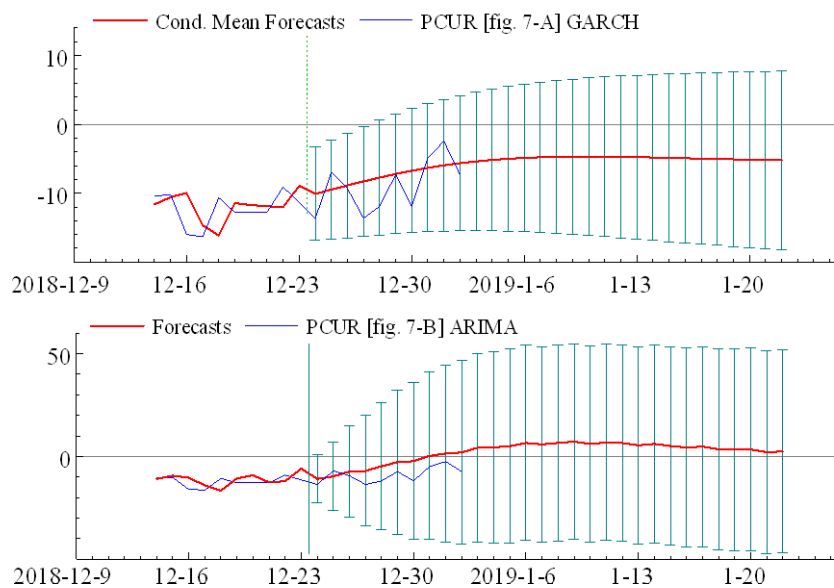


Figure 7. Comparative forecast of GARCH and ARIMA

As an instance, figure 7 shows the comparative analysis of the day-ahead forecast for percentage change in unemployment rate. Part 7-A presents -15.63 percent decline for day 1 forecast while part 7-B shows -10.92 percent decrease in the percentage change in unemployment rate. The results obtained from GARCH forecast is much optimal than the one achieved from ARIMA model.

4. Conclusion

It is empirically shown and documented that not all types of time series data fit with ARIMA model to be used for efficient estimation and accurate forecasting purposes. In this paper, we use the U.S. unemployment rate data arranged on daily basis from 2016-09-13 to 2016-01-02 retrieved from the Bureau of Labor Statistics and we present an empirical analysis of two competing models namely ARIMA and GARCH models for out-of-sample forecasting in which, the ARIMA day-ahead forecast shows a greater standard error and far afield from the actual values while, GARCH model day-ahead forecast presents a comparatively lower standard error and closer values of forecast to the real life values.

Acknowledgements

We acknowledge the assistance of Kabul University Research Center (KRC) in facilitating us in retrieval of the reliable data for this paper. We extend our thanks to the Ministry of Higher Education of Afghanistan for providing the financial assistance for this study.

References

- Aladag, C. H., Egrioglu, E., & Kadilar, C. (2009). Forecasting nonlinear time series with a hybrid methodology. *Applied Mathematics Letters*. <https://doi.org/10.1016/j.aml.2009.02.006>
- Ardia, D., & Hoogerheide, L. (2019). Bayesian Estimation of the GARCH(1,1) Model with Student-t Innovations. *The R Journal*. <https://doi.org/10.32614/rj-2010-014>
- Asteriou, D., & Hall, S. G. (2011). ARIMA Models and the Box–Jenkins Methodology. *Applied Econometrics* (2nd ed.). Palgrave MacMillan.
- Bollerslev, T. (1986). Bollerslev, Generalized Autoregressive Conditional Heteroskedasticity.pdf. *Journal of Econometrics*. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1)
- Caporale, G. M., Pittis, N., & Spagnolo, N. (2003). IGARCH models and structural breaks. *Applied Economics Letters*. <https://doi.org/10.1080/1350485032000138403>
- Chen, K. Y., & Wang, C. H. (2007). A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan. *Expert Systems with Applications*. <https://doi.org/10.1016/j.eswa.2005.11.027>
- Christoffersen, P., & Jacobs, K. (2002). Which GARCH Model for Option Valuation? *SSRN*. <https://doi.org/10.2139/ssrn.306843>
- Contreras, J., Espínola, R., Nogales, F. J., & Conejo, A. J. (2003). ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems*. <https://doi.org/10.1109/TPWRS.2002.804943>
- Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*. <https://doi.org/10.2307/1912773>
- Engle, R. F., & Patton, A. J. (2001). What good is a volatility model? *Quantitative Finance*. <https://doi.org/10.1088/1469-7688/1/2/305>
- Gujarati, D. N. (2004). *Basic Econometric* (4th ed.). New York. <https://doi.org/10.1126/science.1186874>
- Gunter, U., & Önder, I. (2015). Forecasting international city tourism demand for Paris: Accuracy of uni- and multivariate models employing monthly data. *Tourism Management*. <https://doi.org/10.1016/j.tourman.2014.06.017>
- Khandelwal, I., Adhikari, R., & Verma, G. (2015). Time series forecasting using hybrid arima and ann models based on DWT Decomposition. In *Procedia Computer Science*. <https://doi.org/10.1016/j.procs.2015.04.167>
- Khashei, M., & Bijari, M. (2012). A new class of hybrid models for time series forecasting. *Expert Systems with Applications*. <https://doi.org/10.1016/j.eswa.2011.09.157>
- Liu, H., & Shi, J. (2013). Applying ARMA-GARCH approaches to forecasting short-term electricity prices. *Energy Economics*. <https://doi.org/10.1016/j.eneco.2013.02.006>
- Mills, T. C., & Markellos, R. N. (2008). *The econometric modelling of financial time series*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511817380>
- Nochai, R. T. N. (2006). ARIMA MODEL FOR FORECASTING OIL PALM PRICE. In *Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and Applications*, Universiti Sains Malaysia.
- Tealab, A., Hefny, H., & Badr, A. (2017). Forecasting of nonlinear time series using ANN. *Future Computing and Informatics Journal*. <https://doi.org/10.1016/j.fcij.2017.05.001>
- Valipour, M., Banihabib, M. E., & Behbahani, S. M. R. (2013). Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *Journal of Hydrology*. <https://doi.org/10.1016/j.jhydrol.2012.11.017>
- Vilasuso, J. (2002). Forecasting exchange rate volatility. *Economics Letters*. [https://doi.org/10.1016/S0165-1765\(02\)00036-8](https://doi.org/10.1016/S0165-1765(02)00036-8)
- Wooldridge, J. M. (2015). Basic regression analysis with time series data. In *Introductory Econometrics: A modern approach*.

Zhang, P. G. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*.
[https://doi.org/10.1016/S0925-2312\(01\)00702-0](https://doi.org/10.1016/S0925-2312(01)00702-0)

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).