IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 22, Issue 5, Ser. I (Sep. – Oct. 2020), PP 43-48
www.iosrjournals.org

SWGARCH Model for Bitcoin Forecasting

Mohammed Zaki Shbier[1], Zakaria Khaleel Alkayyali[2], Naji Shukri Alzaza[3]
[1]*University of Palestine Gaza, Palestine*
[2]*Al Quds Open University Gaza, Palestine*
[3]*University of Palestine Gaza, Palestine*

Abstract
Nowadays, predicting the price change of bitcoin(BTC) has a main role in the worldwide economy. Selecting the proper model is very measurable to maximize the performance measurements used for comparison. However, the Generalized Autoregresive Conditional Heteroskedasticity (GARCH) is one of the most popular time series models that can be used for time series forecasting. The computation of the long run variance in the GARCH model is based on the historical data that does not reflect the influence of the recent variance. This study aims to utilize the slide window technique with GARCH model to overcome the limitation in the variance of GARCH model. The sliding window technique is solely to estimate the variance in the SWGARCH model.A performance evaluation of SWGARCH was performed on Bitcoins index dataset and compared with GARCH models in terms of mean absolute percentage error. The experimental results showed that the performance of SWGARCH is better than GARCH, which confirmed that SWGARCH can be used for BTCtime series forecasting.
Keywords: Enhanced GARCH, Time Series Forecasting, Bitcoin, Sliding Window

---

Date of Submission: 12-09-2020 Date of Acceptance: 29-09-2020

I. Introduction

Time series analysis utilization has become one of the most interestsfor making statistical inferences and forecasts of future values [1][2][3]. Time series forecasting can be found in several fields that including economics, astronomy, physics, agriculture, disaster, medicine, genetic engineering, and commerce [4][5][6].

Analysing and forecasting time series data and forecasting future values of a time series are among the most important issues that analysts face in many fields,including industry andbusiness, environmental sciences, government, economics, social science, medicine, politics, and finance. Forecasting problems are often classified as short-term, medium-term, and long-term. Short-term forecasting problems involve predicting events only a few time periods (days, weeks, and months) into the future[1]. Short-term is employed in this study, since it is common for BTC.

BTCas a digital currency is quickly turning into a fundamental piece of the worldwide economy [1][3]. It is a cryptographic protocol and distributed network that enables users tostore, and transfer digital currency[2]. However, the measure of research about cryptographic money is still rare and expanding[4].

Furthermore, Severalstudies have used GARCH model for time series forecasting. For instancethe ARIMA-GARCH model utilized for predicting stock [7], three different variants were applied of the ARIMA model in predicting temperature [8], the ARIMA-GARCH model utilized in predicting traffic flow [9]. Moreover, ARIMA-ANN model utilized for predicting short-term electricity prices [10].

The ARIMA-GARCH model utilized for the predictions of the Indian stock [7]. In their model, the moving average filter is used to decompose the given data into two components: low volatile components for the ARIMA model, and highly volatile components for the GARCH model. The ARIMA-GARCH model is compared against ARIMA, trend-ARIMA, wavelet-ARIMA, and GARCH models. The performance measurements used for comparison are mean absolute percentage error (MAPE), maximum absolute percentage error, and root mean square error. The results obtained confirmed that the prediction accuracy of the ARIMA-GARCH model is better as compared to the other models.

A study conducted by [8], ARIMA, trend-based ARIMA, and wavelet-based ARIMA models have been used to predict the average global temperature. The data consisted of global temperatures from the period of 1880 to 2010. The ARIMA model used in this study consisted of three steps. The first step is to make the data stationary by performing the differencing operation. The second step is to identify the suitable values for model order by the autocorrelation function and the partial autocorrelation function. The third step is to predict future values. The prediction model comprised ARIMA, trend-based ARIMA, and wavelet-based ARIMA models. Out of all the three models, it has been observed that the trend-based ARIMA method has better performance than ARIMA and wavelet-based ARIMA in terms of MAPE and mean absolute error.

The hybrid ARIMA-GARCH model for traffic flow prediction [9]. The ARIMA-GARCH model utilized to capture both the conditional mean and conditional heteroscedasticity of traffic flow data. The performance of the hybrid model is compared with the ARIMA model in terms of mean absolute error, mean square error, and mean relative error. The results show that ARIMA-GARCH performed slightly better ARIMA.

In the study conducted by [10], a combination of ARIMA and ANN models was used for predicting short-term electricity prices. The data of the Australian national electricity market in the New South Wales regional in 2006 has been used as a case study. A comparison of the forecasting performance with ARIMA and ARIMA-ANN models in terms of MAPE, MAE, and RMSE showed that the ARIMA-ANN model performed better than ARIMA and ANN models.

The GARCH model uses a long run variance as one of the weights in its calculation of the variance. The long run variance is calculated using the whole series. Nevertheless, using all the series does not reflect the influence of daily variance. In other words, the variance of one month is similar to the variance of the previous one daywhich is not convenience with BTCvolatility.This study aims toutilize the sliding window technique with GARCH model to overcome the limitation in the variance of the original GARCH model

II. The Swgarch Model

The weakness of the GARCH model is fixed by changing a component of the long run average variance with a component called window variance (V_W). The window sliding technique is used to produce the window variance. The enhanced GRACH model, known as SWGARCH, calculates the variance rate (σ_n^2) using window variance as well as from the recent return (u_{n-1}^2) and recent variance rate (σ_{n-1}^2) as shown in (1).

$$\sigma_n^2 = \gamma V_w + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2 \quad (1)$$

whereγ, α, and β are the parameters, and n is the day. Window variance can be calculated as shown in (2).

$$V_w = W_1 \cdot u_{n-1}^2 + W_2 \cdot u_{n-2}^2 + \cdots + W_z \cdot u_{n-z}^2 (2)$$

where u_{n-i} is the return, and W_z (z = 1, 2, 3, …,) is the window weight subject to

$$W_z + W_{z-1} + W_{z-2} + \cdots + W_1 = 1$$

and z is the window size. The return (u_i)is defined as the continuously compounded return during day i (between the end of day i-1 and end of day i). In other words, return is the gain or loss in a particular period. The general process in calculating the variance in the SWGARCH model involves four steps. The first step is to estimate the parameters, and the second step is to compute the return. The sliding window variance is calculated in the third step, and the last step is to calculate the recent variance.

A. Estimating SWGARCH Parameters

The parameters γ, α, and β are estimated from the sample of historical data using the maximum likelihood method (MLM). It involves choosing values for the parameters that have maximum chances or likelihoods of the data [3]. MLM is an iterative searching tool to find the parameters in the model that maximize the expression as shown in (3).

$$\sum_{i=1}^{m} \left[-ln(v_i) - \frac{u_i^2}{v_i} \right] \quad (3)$$

where u_i is the return, and vi is the variance at observation i and is calculated as

$$v_i = (S_i - S_{i-1}) / S_{i-1}$$

where S_i is the actual value at day i. The likelihood measure (L) is given by (4) as follows:

$$L = -ln(v_i) - v_i^2 / v_i (4)$$

where vi is variance of day i. The sum of γ , α, and β is equal to 1.

B. The Return Computation

The return value (u_i^2) at day i for each observation is computed by squaring the period return (μ_i), which is calculated as shown in (5).

$$\mu_i = ln\left(\frac{s_i}{s_{i-1}}\right) \tag{5}$$

where s_i is the actual value at day i. Data other than the ones that have been used in estimating the parameters will be used in the return calculation to avoid bias.

C. Computation of Sliding Window Variance

The SWGARCH model introduces a new method by using sliding window variance in obtaining the weights for forecasting the future value. This method is composed of three steps:

a) Estimate the window size.
b) Calculate the weight of each observation within the window.
c) Multiply each weight by the return, and then compute the sum of the multiplied values.

A principal component analysis has been used to estimate the window size. The window size is identified where there is a big drop in the scree plot of the variance. The weight of each observation within the window is composed of two steps. The first step is to calculate the total of window size weights (TL) using (6).

$$TL = w_z + w_{z-1} + \ldots + w_1 = 1 + 2 + 3 + \cdots + z(6)$$

where w_1 is the first weight, and z is the window size. The second step is to normalize the weight as shown in (7).

$$W_i = \frac{w_i}{TL}(7)$$

whereW_i is the normalized weight, whereas w_i is the weight for each observation for i = 1, 2, …, z. Multiplying each weight by the return, and computing the sum of the multiplied values as shown in (2) are performed to obtain the result for the window variance.

D. Recent Variance

The recent variance is used as the third component of the SWGARCH model. The SWGARCH variance calculation is a recursive procedure approach.

The forecasted variance ($E[\sigma_{n+t}^2]$) for period n + t can be calculated as shown in (8).

$$E[\sigma_{n+t}^2] = V_W + (\alpha + \beta)^t + (\sigma_n^2 - V_W) \tag{8}$$

where t is the time slice and $\alpha + \beta < 1$ is for stable forecasting. The information available at period n is used for this purpose. Finally, the forecasted value can be calculated as follows:

$$ForecastS_{n+t} = S_n + (S_n * E[\sigma_{n+t}^2])$$

III. Evaluation Of The Swgarch Model

The performance of the SWGARCH model is compared with GARCH. These are common forecasting models. MAPE are used as performance metrics computed as follows:

$$MAPE = \frac{1}{n}\sum_{t-1}^{n}\frac{|e_t|}{V_t} \cdot 100.$$

$\hat{Y}$ is a vector of n[th] predictions and Y is the vector of observed values corresponding to the inputs to the function that generates the predictions.

The BTC index dataset, is a cryptocurrency and worldwide payment system. It is the first decentralized digital currency, as the system works without a central bank or single administrator. The dataset contains 365 daily stock indices between the period of 1st January 2017 and 31th December 2017, and is available athttps://www.coindesk.com/price/.

A. Estimating SWGARCH Parameters

The dataset used in this paper consist of 365 days and available on https://finance.yahoo.com/quote/BTC-USD from 1/1/2019 to 31/12/2019. The first 100 days used for parameter estimation (training dataset) while the rest of the dataset used for forecasting.Table 1 shows a sample of ten days' data and how the calculation could be organized in estimating the parameters. The first and second columns show the day and the index(S_i) for the day respectively. The third column records the change in rate (v_i) at the end of day_i where$v_i = (S_i - S_{i-1})/ S_{i-1}$. The fourth column records the estimation of variance, ($v_i = \sigma_i^2$), for day_i based on the change rate. The fifth column tabulates the likelihood measure (L) and this can be obtained using (4). This paper is interested in choosing $\gamma, \alpha,$ and β to maximize the sum of the numbers in the fifth column. This involves an iterative search procedure. In this dataset, the optimal values of the parameters are $\gamma = 0.2415, \alpha = 0.5482,$ and $\beta = 0.2103$. The maximum value of the function in (3) is 713.86.

Table 1: Estimation of parameters in the SWGARCH model

Date	Day_i	S_i	u_i	$v_i = \sigma_i^2$	L
Jan 01 2017	1	997.69			
Jan 02 2017	2	1018.05	0.02041	0.00042	0.00042
Jan 03 2017	3	1030.82	0.01254	0.00016	0.00016
Jan 04 2017	4	1129.87	0.09609	0.00923	0.05684
Jan 05 2017	5	1005.81	-0.10980	0.01206	0.05685
Jan 06 2017	6	895.67	-0.10950	0.01199	0.05685
Jan 07 2017	7	905.17	0.01061	0.00011	0.05685
Jan 08 2017	8	913.52	0.00922	0.00009	0.05685
Jan 09 2017	9	899.35	-0.01551	0.00024	0.05685
Jan 10 2017	10	904.79	0.00605	0.00004	0.05685

B. The Return Calculation

Table 2 displays another five (5) days of BTC index, where the fourth column displays the period return (μ_i). The return (u_i^2) of each day of the index is computed by squaring the period return.

Table 2: Computation of Return

Date	Day_i	S_i	μ_i	u_i^2
Sep 13 2017	256	3874.26		
Sep 14 2017	257	3226.41	-0.18298	0.03348
Sep 15 2017	258	3686.90	0.13342	0.01780
Sep 16 2017	259	3678.74	-0.00222	0.00000
Sep 17 2017	260	3672.57	-0.00168	0.00000

C. Computation of Sliding Window Variance

The component numbers and variance explained for the BTC Index are displayed in Table 3 as well as in Figure 1. There is a big drop in variance explained between components 1 and 2 as shown by the steep slope. Thus, the window size is identified as 2.

Table 3: BTC Index Variance

Component Number	Variance Explained by PCA
1	31.9946
2	20.678
3	18.2899
4	15.1412
5	13.8713
6	0.0245

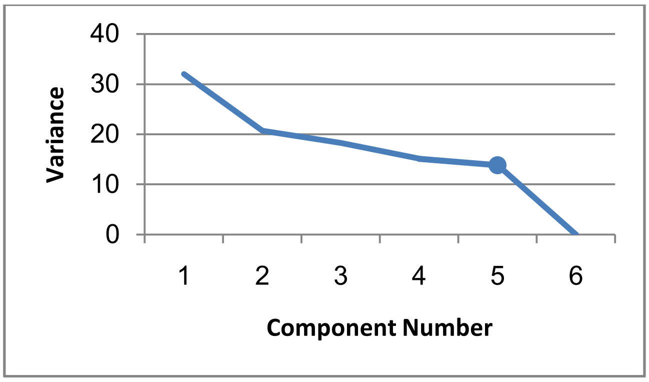

Figure 1: Scree Plot for BTC Index dataset

The total of window size weight (TL) is calculated as $\mathrm{TL} = \mathrm{w}_5 + \mathrm{w}_4 + \mathrm{w}_3 + \mathrm{w}_2 + \mathrm{w}_1 = 1 + 2 + 3 + 4 + 5 = 15$. Each weight is normalized as shown below.

$$W_1 = \frac{w_1}{TL} = \frac{1}{15} = 0.07$$
$$W_2 = \frac{w_2}{TL} = \frac{2}{15} = 0.13$$
$$W_3 = \frac{w_3}{TL} = \frac{3}{15} = 0.2$$
$$W_4 = \frac{w_4}{TL} = \frac{4}{15} = 0.27$$
$$W_5 = \frac{w_5}{TL} = \frac{5}{15} = 0.33$$

Further, the sliding window variance for day 258 is calculated as follows:

$$V_W = (0.07 \times 0.033) + (0.13 \times 0.017) + (0.2 \times 0.0) + (0.27 \times 0.0) + (0.33 \times 0.011) = 0.0223$$

D. Recent Variance

The recent variance for day 258 (σ^2_{258}) is set to equal to the return of day 258 since there is no recent variance. Thus, from Table 2, $\sigma^2_{259} = 0.00023$. The SWGARCH variance for day 259 is calculated using (1) as follows:

$$\sigma^2_{259} = (0.2415 \times 0.0223) + (0.5482 \times 0.0178) + (0.2103 \times 0.0178) = 0.01492$$

E. The Forecasting

The estimated variance of day 258 will have to be calculated in forecasting the index for the same day. This can be done by using (8), where n is equal to 258 and t is equal to 1. Thus, the expected variance for day 259 is given as below:

$$E[\sigma^2_{259}\} = 0.0067 + (0.25473 + 0.20472)^1 \text{ x } (0.0024 - 0.0067) = 0.00346$$

The forecasted index of day 259 is computed as follows:

$4067.08 + (4067.08 \times 0.00346) = 4067.774$

Table 5 depicts the results of forecasting using the SWGARCH model for ten days' data of the BTC dataset. The error is computed as |Actual value – forecasted value|.

Table 5: Sample Model Performance for BTC Dataset

Date	Day	Index	Forecast Index	| Error |
Sep 17 2017	260	3672.57	3719.35	
Sep 18 2017	261	4067.08	4076.77	46.78
Sep 19 2017	262	3897.00	3910.51	9.69
Sep 20 2017	263	3858.09	3883.87	13.51
Sep 21 2017	264	3612.68	3620.67	25.78
Sep 22 2017	265	3603.31	3609.49	7.99

Sep 23 2017	266	3777.29	3787.17	6.18
Sep 24 2017	267	3662.12	3666.20	9.88
Sep 25 2017	268	3927.50	3934.05	4.08
Sep 26 2017	269	3895.51	3904.11	6.55
			Average Error	**14.50**

F. Benchmark SWGARCH model comparison with GARCH

In terms of evaluation the SWGARCH with its master (i.e. GARCH), the performance criteria MAPE was utilized. Table 6 shows that SWGARCH's performance is better than GARCH for the BTC index with MAPE difference=0.30791. The model's propagated forecasted values are almost as accurate as the actual dataset. The results show that the SWGARCH model is able to forecast with betteraccuracy in terms of MAPE.

Table 6: Experimental Results of SWGARCH and GARCH

	SWGARCH	GARCH
MAPE	0.30791	8.91685

IV. Conclusion

BTC critical variance along time series need a properate forecasting model. MAPE is the performance criteria that used to compare between forecasting models.The proposed SWGARCH model has overcome the weakness of the original GARCH model by replacing the component of long run variance with sliding window variance. SWGARCH is a hybrid model where hybridization is done between the GARCH model and the sliding window technique. In the SWGARCH model, variance is based on the most recent observation of return. This would incorporate more recent returns, which will provide greater weight. Indeed, SWGARCH can be utilized for forecasting the BTC price change with more accurate results.

References

[1]. B. Narendra and R. Eswara. 2014. Selected Indian stock predictions using a hybrid ARIMA-GARCH model. In *2014 International Conference on Advances in Electronics Computers and Communications*, 1–6.

[2]. F. Tobar, M. Orchard, D. Mandic, and A. Constantinides. 2014. Estimation of financial indices volatility using a model with time-varying parameters. In *2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics*, 318–324.

[3]. Z. Jingwen, C. Li, and P. Qin. 2015. Modeling non-stationary stochastic systems with generalized time series models. In *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 1061–1067.

[4]. S. Banik and K. Khan. 2015. Forecasting US NASDAQ stock index values using hybrid forecasting systems. In *The 18th International Conference on Computer and Information Technology*, 282–288.

[5]. J. Hernandez, M. Ovando, F. Acosta, and P. Pancardo. 2016. Water level meter for alerting population about floods. In *2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA)*, 879–884.

[6]. Y. Pan, M. Zhang, Z. Chen, M. Zhou, and Z. Zhang. 2016. An ARIMA based model for forecasting the patient number of epidemic disease. In *2016 13th International Conference on Service Systems and Service Management (ICSSSM)*, 1–4.

[7]. C. N. Babu and B. E. Reddy. 2014. Selected Indian stock predictions using a hybrid ARIMA-GARCH model. In *The International Conference on Advances in Electronics Computers and Communications*, 1-6.

[8]. C. N. Babu and B. E. Reddy. 2012. Predictive data mining on average global temperature using variants of ARIMA models. In *The International Conference on Advances in Engineering, Science and Management*, 256–260.

[9]. H. Chen, F. Li, Q. Wan, and Y. Wang. 2011. Short term load forecasting using regime-switching GARCH models. In *2011 IEEE Power and Energy Society General Meeting*, 1–6.

[10]. P. Areekul, T. Senjyu, H. Toyama, and A. Yona. 2009. Combination of artificial neural network and ARIMA time series models for short term price forecasting in deregulated market. In *2009 Transmission & Distribution Conference & Exposition: Asia and Pacific*, 1–4.

Shbier, Mohammed Zaki, Zakaria Khaleel Alkayyali, and Naji Shukri Alzaza. "SWGARCH Model for Bitcoin Forecasting." *IOSR Journal of Computer Engineering (IOSR-JCE)*, 22(5), 2020, pp. 43-48.