Modeling Bitcoin Prices and Returns using ARIMA Model

David Aidoo[1] and Kofi Agyarko Ababio[1]

[1]*Department of Statistical Sciences, Kumasi Technical University, Kumasi - Ghana*

Corresponding author's email: aidoodavid86@gmail.com

Abstract

This paper aims to model and forecast daily prices and returns of Bitcoin using the Box-Jenkins technique. The study used daily Bitcoin prices and returns from 19th June 2020 to 22nd January 2022 comprising 572 observations. The results indicated that the best models for predicting the daily prices and returns of Bitcoin, respectively, were ARIMA (0, 1, 0) and ARIMA (0, 0, 0), according to Akaike's Information Criterion (AIC). Based on the study's findings, investors are advised to avoid the temptation of over reliance on asset prices and returns forecasts in financial markets, especially the Bitcoin market.

Keywords: Cryptocurrency, Bitcoin, Akaike's Information Criterion, Box-Jenkins technique, Asset prices

1.0 INTRODUCTION

Cryptocurrencies are digital money that is processed anonymously over a decentralised network, Tarasova et al., (2020). The cryptocurrency was first introduced to the world's financial markets about a decade ago. They act outside centralised financial institutions by finding additional money and investment options. Over 2000 cryptocurrencies are in use, with most of them being traded anonymously via blockchain technology. The most widely used cryptocurrencies are Bitcoin, USD coin, Dogecoin, Chainlink, XRP, Tether, Cardano, Polkadot, and Stellar. However, the first cryptocurrency, Bitcoin, which accounts for most market capitalisation was the subject of this research.

Forecasting plays a vital significant role in finance, education, health, agriculture and marketing. Several of these forecasts employ the statistical model known as the Autoregressive Integrated Moving Average (ARIMA), which is widely used to forecast stationary datasets. For example, the significant volatility in pricing is the key challenge researchers face in predicting the price of cryptocurrencies. Using the ARIMA model, the researchers addressed this issue by rendering the datasets stationary over many time series dimensions.

The price of Bitcoin (BTC) has been discussed by many academics despite its rapid fluctuations. According to Adcock and Gradojevic (2019) and (Chen, 2020), technical indications can be used to anticipate BTC prices. Other recent studies have employed machine learning-related techniques for daily price forecasting and price increase/decrease forecasting (Mallqui, 2019). According to Adcock and Gradojevic (2019), accuracy can reach 63 percent. Mallqui (2019) discovered that daily price forecasts had a 98 percent success rate.

To develop financial plans, finance specialists utilise projections. They also employ them to communicate their revenue projections. Investors put their money into stocks/securities to make a profit. As a result, financial professionals and investors benefit from having a strong understanding of future share price movement. They can increase their confidence by knowing what the future holds by consulting and investing. Financial professionals and investors will undoubtedly be interested in forecasting methodologies that will accurately predict future share price movements with the smallest possible error margin. The Box-Jenkins method, the Black-Scholes model, and the binomial model are just a few forecasting strategies that can be used to forecast changes in stock values. In this study, however, bitcoin prices and returns predictions were made using the Box-Jenkins approach.

This research employed the ARIMA model to predict the price and returns of bitcoin. Accordingly, the study seeks to address three specific objectives: to determine which moving average (q) and autoregressive (p) parameters best match the Bitcoin price and return series, to identify the best differencing order (d) that makes the time series stationary and to use the ARIMA model to predict Bitcoin prices and returns for a period of 365 days. Again, this study's findings will assist financial portfolio managers, investors, and other shareholders in making better decisions about cryptocurrency investments. This will also instill greater trust in financial sector investors, allowing them to embark on more risky transactions. Furthermore, the study will be beneficial to investors, shareholders, directors, regulators, other investment businesses, and university academics.

According to Nakamoto (2008), many academic papers have been published on Bitcoin since its inception in 2008. Traditional econometric approaches were used by some authors to examine the future prospect of bitcoin and its influence on additional economic factors using conventional economic techniques. In terms of hedging capabilities and exchange benefits, the GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model revealed significant parallels between gold and the US dollar.

Dyhrberg (2015) revealed that bitcoin has apparent advantages for risk-averse investors because it sits in the middle between gold and US currency in terms of financial markets and project management. A cointegration analysis was employed by various authors to look into the impact of specific variables on bitcoin's price. In this regard, Ciaian et al., (2015) used Barro's gold standard model to discover that the Dow Jones index, currency rate, and oil price only significantly influenced global macro-financial growth in the short run. Over time, these factors rarely have an impact on bitcoin's price.

Zhu et al., (2017) discovered that using the VEC (Vector Error Correction) model, it was found that the Consumer Price Index, the Dow Jones Industrial Average, and the Federal Funds Rate had a long-term negative impact on the price of bitcoin.

Namin et al., (2018) revealed that the results for rolling ARIMA and LSTM from the financial time series were 511.481 and 64.213, respectively. The results were computed using the average Rooted Mean Squared Error (RMSE). The error rate was reduced by 87.445 percent. Both models had RMSE values of 5.999 and 0.936, respectively. With the RMSE data, it is clear that LSTM outperforms ARIMA by a large margin.

This is strong evidence that the LSTM is considerably superior at forecasting time series, particularly Bitcoin price projections, which are more difficult to anticipate due to high volatility (NAMIN, 2018). According to Mallqui (2019), the behaviour of Bitcoin is constantly unpredictable, and price prediction has a vital role because money may quickly lose value depending on this price forecast.

Mallqui (2019) employs the GARCH model to estimate prices for the Ethereum, Bitcoin, and Dash markets because the price of Bitcoin is highly correlated with the number of transactions and the average value of each transaction.

The error reduction rates of ARIMA and LSTM were recommended by Ensemble A findings. The error reduction rate was estimated to be between 50.05 and 52.78 percent, LSTM won by a sliver of a point, demonstrating its supremacy in ARIMA.

Some writers have compared machine-learning approaches to classical models to estimate bitcoin values. In terms of accuracy and RMSE, McNally (2018) found that neural non-linear network models (RNN and LSTM) outperformed conventional ARIMA model forecasting. Additionally, the LSTM slightly outperformed the RNN, but the difference was not significant because the LSTM had greater accuracy (53%) and a slightly larger RMSE (7%) than the RNN. However, due to the various activation functions and equations that must be resolved, the LSTM requires 3.1 times more to train with the same network settings. Nevertheless, a number of academics used machine-learning optimization to examine the impact of a variety of new features on bitcoin pricing.

Huisu et al., (2018) developed a rolling window LSTM model to predict the price of bitcoin using macroeconomic, global currency ratio, and blockchain data as inputs. According to Yamak et al., (2019), the LSTM model with rolling window settings outperforms the traditional LSTM and neural network models using MAPE and RMSE. Other authors have quantified the impact of different indicators on bitcoin pricing, such as tweets about bitcoin or web search media results, using machine learning optimization.

In order to predict if the price of bitcoin would change in the near future, Stenqvist and Lonno (2017) looked through 2.27 million tweets about cryptocurrency. Based on the intensity of sentiment swings from one period to the next, the estimated model states that compiling tweet sentiments over a 30-minute period with four shifts forward and a sentiment-limited fluctuation of 2.2 percent would result in an accuracy estimate of 83 percent. The author suggests using machine learning to look at the correlation between Bitcoin and Twitter data further. Matta et al., (2015) indicated that the quantity of tweets and/or search engine results was compared to the price of bitcoin.

There have been various studies on bitcoin price predictions that make use of machine learning (LSTM), GARCH, and other sophisticated statistical methods. To forecast bitcoin prices and returns, the researcher is investigating the feasibility of using the traditional time series forecasting method, ARIMA. It should be emphasised that earlier studies failed to consider the significance of forecasting bitcoin prices and returns using the same model and then assessing the model that was employed to forecast the same cryptocurrency. This study took into account Bitcoin prices and returns in order to anticipate potential Bitcoin prices and returns. To each, an ARIMA model was applied, and the dataset was trained using daily time series.

2.0 Materials and Methods

2.1 Source of Data

The research employed daily price and returns of Bitcoin from 19th June, 2020 to 22nd January, 2022 for a period of 2 years. Information on the daily prices and returns of bitcoin for the underlying period of the study was obtained from Bloomberg and was used for this research. This dataset consists of 572 observations for the study.

2.2 Time Series Models

In order to anticipate how a factor would affect the predicted variable, a time series model (ARIMA) uses the predicted previous behavior of the variables to generate predictions about its future behavior. According to Howrey (1980), time series models are based on an unstable nonparametric formulation. He concentrated more on data analysis to simplify the model. In time series data, numerous patterns or trends can be noticed.

However, due to the volatility of the cryptocurrency and the numerous extrinsic factors that affect its price, it can be extremely difficult to model Bitcoin prices and returns. In this situation, a number of factors may influence the application of ARIMA against alternative cutting-edge techniques: the GARCH model primarily handles the clustering of volatility in financial returns. In order to capture historical patterns in bitcoin return patterns, the researcher used the ARIMA model. Again, the ARIMA was reasonably utilized because of the time frame for the data, even though sophisticated techniques like LSTM require a huge amount of data for training to prevent overfitting. Additionally, ARIMA was employed because, unlike more complex methods like machine learning, GARCH, it is superior at stabilizing non-stationary data and allowing for straightforward result interpretation.

2.2.1 The Random Walk Model

The general equation for random walk model is given as follows:

$Y_t = Y_{t-1} + e_t \qquad t \in Z,$

where "e_t" is also the white noise or error term and is normally distributed. Y_t is the random walk and Y_{t-1} is the immediate past value of a time series determines its future value in the simplest random walk process.

2.2.2 Autoregressive Model, AR(p)

An autoregressive model combines one or more previous values to describe the current value of a time series. It shows how one value is related to its immediate preceding values. The order term, p, in an autoregressive model indicates how many previous values should be used in the difference equation to obtain the present value.

The general equation of autoregression (p) is written as follows:

$$Z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots\ldots\ldots\ldots + \phi_p z_{t-p} + \varepsilon_t \quad t > P \qquad (1)$$

Where

Z_t: response variable at time t

$z_{t-1}, z_{t-2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots z_{t-p}$: response variable at time t lags

ϕ_i: coefficient to be estimated as; i = 1,2, 3…... p $-1 < \phi p < 1$

ε_i: random error with mean zero and constant variance i.e. $\varepsilon_i \approx N(0, \sigma_i^2)$

2.2.3 Moving Averages MA (q)

Random shocks in a noisy environment affect a time series. As a result, the random shocks that existed in past values have an impact on the present value of the series. The effect of earlier random shocks on future values is captured using moving average terms.

The MA(q) process is given by;

$$Z_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} \ldots\ldots\ldots\ldots\ldots + \theta_q\varepsilon_{t-q} \qquad (2)$$

Z_t: response variable at time t

ε_i : random error with mean zero and constant variance i.e. $\varepsilon_i \approx N(0, \sigma_i^2)$

$\varepsilon_{t-1}, \varepsilon_{t-2} \ldots\ldots\ldots\ldots \varepsilon_{t-q}$: errors in previous time periods that are incorporated in the response Z_t

2.2.4 Mixed Models/ARMA (p q)

The AR/MA model is used to generate the Autoregressive Moving Average model. Any stationary process may be approximated with any degree of precision using an autoregressive model or a moving average.

$$Z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots\ldots\ldots + \phi_p z_{t-p} + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \ldots\ldots\ldots\ldots\ldots + \theta_q\varepsilon_{t-q} \qquad (3)$$

2. 2. 5 ARIMA (pdq)

The mixed models (ARMA) model presupposes the stationarity of time series data (i.e., their statistical qualities do not vary over time). However, actual data is not always stationary. The differencing method makes time series data stationary.

The equation for the first order difference of Z_t is given as $Z_t = Z_t - Z_{t-1.}$ (4)

The Integrated Autoregressive Moving Average (ARIMA) model uses an ARMA time series that has been rendered stationary by a differencing technique.

The ARIMA model breaks down historical data into three processes: an autoregressive (AR) process that records past events, an integrated (I) process that makes data stationary for forecasting, and a moving average (MA) process that calculates prediction error.

ARIMA (p, d, q)'s general equation is given by;

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots\ldots + \phi_p z_{t-p} + \ldots + dZ_{t-p-d} + \ldots + \theta_q\varepsilon_{t-q} + \varepsilon_t \qquad (5)$$

Furthermore, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) methods was used to determine if an ARIMA (p, d, q) model was a good statistical fit for data.

2.3 Functions of Autocorrelation and Partial Autocorrelation

The ACF is taken into consideration when the linear dependence between y_t and its previous values, y_{t-1}, is of importance. The coefficient of sample autocorrelation between y_t and y_{t-1} is denoted by r_k, which is a function of k only under the weak condition of stationarity.

$$r_k = \frac{\sum_{t=k+1}^{n}(y_t - \bar{y})(y_{t-k} - \bar{y})\sum_{t=k+1}^{n}(y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{n}(y_t - \bar{y})^2 \qquad \sum_{t=1}^{n}(y_t - \bar{y})^2} \tag{6}$$

Again, the PACF, which is a function of the ACF, assesses the degree of correlation between a variable and a lag of itself that is not accounted for by correlations at any lower-order lags, taking the AR model into account.

2.4 Box-Jenkins Methodology

A series of steps known as the Box-Jenkins technique can be used with time series data to choose, fit, and validate ARIMA models.

Model Identification

Estimation and testing

Diagnostic check

Forecasting

2.4.1 Model Identification

The process of model identification entails evaluating the statistical fit of a given model with certain p, d, and q parameters. Unit root tests such as the ADF and KPSS tests, sample partial autocorrelation function (PACF) and sample autocorrelation function (ACF) tests were used in this research to determine the d parameter. The next step is to discover a suitable ARMA form to represent the stationary series once the proper order of differencing necessary to make the bitcoin price and return series stationary has been established. The p autoregressive and q lagged error parameters with the best fit to the bitcoin price and returns data were determined using the ACF, PACF, and the consequent correlograms as well as objective penalty function statistics. Alternative objective methods for identifying ARMA models were essential because of the very subjective nature of the methodologies of ACF and PACF. Akaike Information Criterion (AIC), and Bayesian Information Criterion are the penalty function statistics used in this research.

2.4.2 Estimation and testing

Model estimation means finding the most accurate estimations for the specified model's parameters. The forecast model for this study is estimated as follows:

$$\Delta Z_t = c + \phi_1 \Delta z_{t-1} + \ldots\ldots\ldots + \phi_p \Delta z_{t-p} + \ldots\ldots\ldots + \theta_1 \varepsilon_{t-1} + \ldots\ldots\ldots\ldots\ldots + \theta_q \varepsilon_{t-q} \tag{7}$$

2.4.3 Diagnostic check

Diagnostic checking involves validating the predicted model's reasonable fit to the data. The residuals produced from the estimated model were checked for autocorrelation using the ACF, PACF, and Ljung-Box Q statistics in this research to make sure the model is a reasonable fit to the bitcoin data. If the residuals show any evidence of autocorrelation, a different model should be used. Equally, if the residuals are white noise, it indicates a good fit. Another crucial check will be to estimate the chosen model over a range of time periods and see how robust it is. If the parameter estimates remain constant throughout time, the fit is satisfactory. Otherwise, extra thought will be needed. Finally, the model created was used to predict the price of bitcoin for 365 days in advance.

2.5 Validation Matrices

Validation measures evaluate the effectiveness of time series models in predicting and identifying reliable models. When utilizing ARIMA or other forecasting techniques, choosing the most pertinent statistic for particular applications or business goals is essential. With regard to this, the researcher considered the following error matrices.

RMSE $=\sqrt{\frac{1}{n}}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2\sqrt{\frac{1}{n}}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$

MAPE $=\frac{100}{n}\frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_1-\hat{y_1}}{y_i}\right|\sum_{i=1}^{n}\left|\frac{y_1-\hat{y_1}}{y_i}\right|$

MSE $=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$

y_i *is the actual observation of time t*

$\hat{y}_i$ *is the predicted observation at the using the model*

n is the number of the observations

2.6 Assumptions and limitations for using the ARIMA Model

For the ARIMA model to accurately predict bitcoin prices and returns, the researcher must acknowledge some assumptions and built-in limitations, as follows: the model assumed that the statistical characteristics would not change over time and that there should be a linear relationship between the variables. Notwithstanding this, bitcoin price dynamics, being influenced by a multitude of factors, might not always exhibit linearity.

One key limitation was that the model disregarded external factors. Because ARIMA models are univariate, they can only forecast future values of the series itself based on past values. Yet a number of external factors, like the state of the global economy, adjustments to the law, or technological advancements, can have an impact on the price and returns of Bitcoin. These factors were not captured in the model.

3.0 RESULTS

3.1 Test for Stationarity

To fit the model of the data, the stationarity of the variable was first tested. The stationarity was checked by plotting the data, after which the ADF and KPSS tests were further used to test whether the data is stationary.

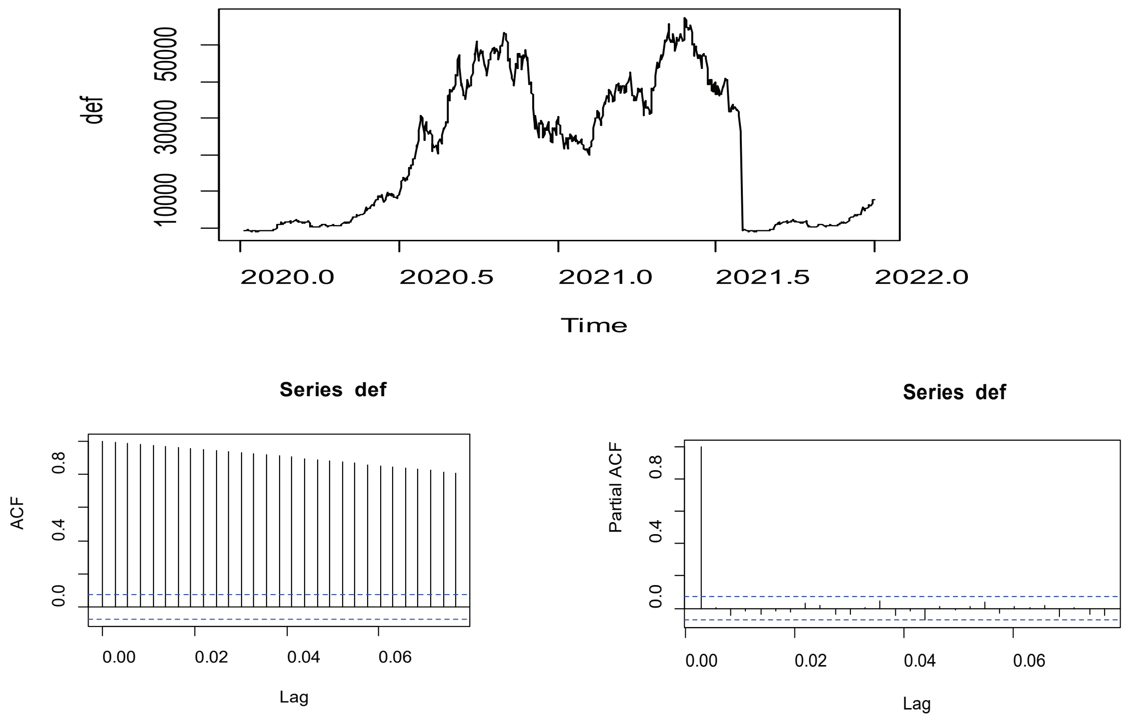

Figure 1: ***Plot of Bitcoin Prices showing Stationarity***

The figures above show that there are irregular (residuals) variations in the Bitcoin prices. Again, it was observed from the ACF and PACF plot that there is high autocorrelation in the Bitcoin prices. This is an indication that the data for Bitcoin prices is non-stationary. The high autocorrelation of the bitcoin price is indicating a strong link between recent and historical values. This knowledge is essential for traders, analysts, investors and econometricians to choose the best models to predict future prices in the volatile cryptocurrency market.

Table 1: ADF test for Bitcoin Prices

TEST	TEST STATSITC	P-VALUE
ADF	-1.3555	0.8511

From Table 1 it can be deduced that the p-value (0.8511) exceeds the conventional significance level of 0.05 indicating that the null hypothesis cannot be rejected. It is therefore concluded that there is no stationarity in the data. The non-stationarity shows the tendency of Bitcoin prices to shift over time in terms of its statistical features like mean and variance. Due to the possibility that historical trends no longer hold true, forecasting future prices becomes more difficult, and this may be a reflection of external variables like changes in legislation, technological improvements, political affairs, or alterations in how people view cryptocurrencies. It could be beneficial to investors if they monitor the movement before they make any investment decision.

Table 2: KPSS stationarity test results for Bitcoin Prices

TEST	TEST STATISTIC	P-VALUE
KPSS	2.0381	0.01

The results of the KPSS test have a p-value of 0.01 and are less than the traditional significant level of 0.05, rejecting the null hypothesis in support of the alternative hypothesis and indicating that there is no trend stationarity in the price of bitcoin. This finding confirms that of ADF test which implies that the actual data for Bitcoin price is not stationary and it can be very challenging to forecast the future price.

Table 3: Model Selection Criteria for Bitcoin Prices

	AIC	AICc	BIC
ARIMA (2,1,2)	12852.18	12852.3	12879.7
ARIMA (0,1,0)	12846.28	12846.3	12855.45
ARIMA (1,1,0)	12849.26	12849.29	12863.01
ARIMA (0,1,1)	12848.26	12848.29	12862.01
ARIMA (0,1,0)	12844.31	12844.32	12848.9
ARIMA (0,1,0)	12859.21	12859.22	12863.8

The estimated results from Table 3 showed that ARIMA (0,1,0) model with zero mean is selected as the best model because it has lowest AIC and BIC values. The findings indicate that the ARIMA (0, 1, 0) model with a zero mean best captures the characteristics of the cryptocurrency market over the period under study. It says the market does not rely heavily on historical data and that the best forecast for price movement tomorrow is typically zero. However, the dynamic nature of cryptocurrency markets necessitates routine re-evaluation of the fit of models and frameworks.

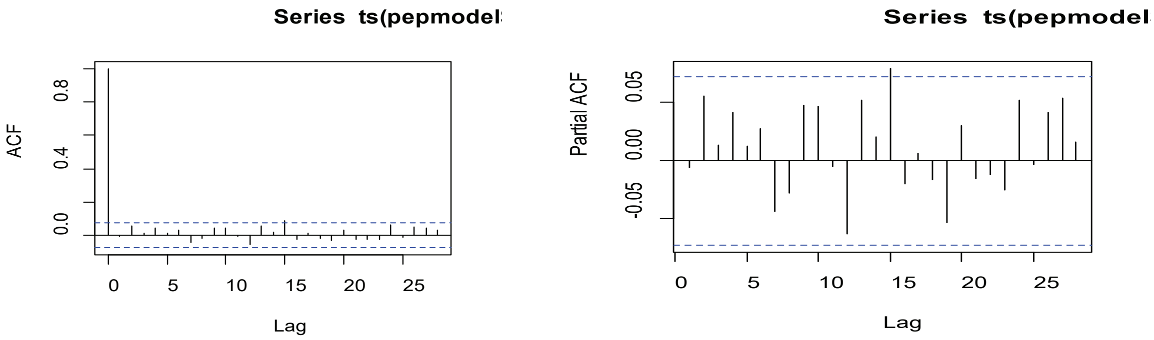

Figure 2: *ACF and PACF Residual for Bitcoin Prices*

Table 4: Box-Ljung test for Residual Correlation for Bitcoin Prices

Test	Chi-Square test	Df	P-value
Box-Ljung test	9.2308	10	0.5104

It is observed from Table 4 that the Chi-Square value of 9.2308 and the p-value of 0.5104 are both greater than standard significant level of 0.05. In line with this, we do not have enough evidence to reject the null hypothesis indicating that the residuals for the observed values are independently distributed.

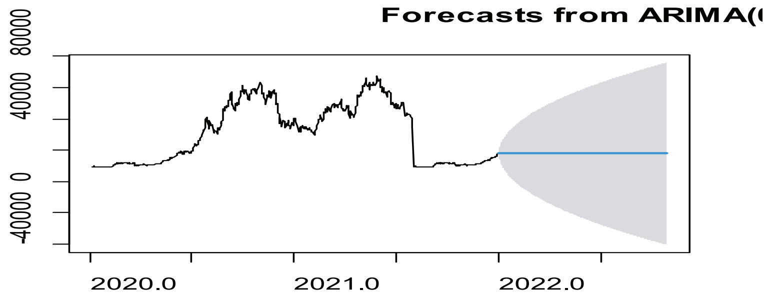

Figure 3: ***The Daily Bitcoin Price forecast***

From the figure above, we can observe that the forecast of the original Bitcoin price shows a constant fluctuation in the upcoming days.

Table 5: Accuracy of forecasting the daily Bitcoin Price

Model	ARIMA (0,1,0)
ME	-11.76083
RMSE	1715.42
MAE	892.0642
MPE	-0.17383
MAPE	2.86705
MASE	0.02592
ACF1	-0.00570

Table 5, shows the forecast accuracy for the original daily price of Bitcoin time series using ARIMA (0, 1, 0). From the results, the different error measurement shows that the forecast accuracy of our model is accepted. This is due to the fact that the MASE value is less than 1, and it implies that ARIMA (0,1,0) performs better than a naïve model in this case study. Again, with an MAPE of 2.86705%, the predictions of the model are generally accurate. Particularly for volatile cryptocurrencies, a MAPE < 10% is regarded favourably. Also, the projections of the model for Bitcoin's price have an average error of 3.43%, or a difference of $1715.42. This estimate was made since the price range at the time the study was conducted was around $50,000.00.

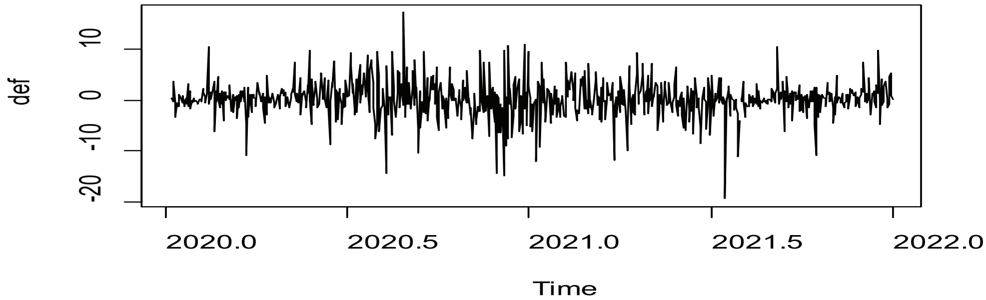

Figure 4: ***Plot of Bitcoin Returns showing stationarity***

The figure above shows that there is a constant mean in the returns. This signifies that the return for Bitcoin is stationary over the time.

Table 6: KPSS test for Stationarity for Bitcoin Returns

TEST	TEST STATISTIC	P-VALUE
KPSS	0.2453	0.1

The results of the KPSS test have a p-value of 0.1, which is higher than the traditional significant level of 0.05, which means the null hypothesis cannot be rejected, indicating that there is trend stationarity in the returns of bitcoin.

Table 7: Box-Ljung test for Residual Correlation for Bitcoin Returns

Test	Chi-Square test	Df	P-value
Box-Ljung test	11.5	10	0.3199

The findings from Table 7 indicated that the Chi-Square value of 11.5 and the p-value of 0.3199 are both greater than significant level of 0.05. This implies that there is no enough basis to reject the null hypothesis indicating that the residuals for the observed values are independently distributed.

Table 8: Model Selection Criteria for Bitcoin Returns

	AIC	AICc	BIC
ARIMA (2,0,2)	4012.925	4013.402	4040.434
ARIMA (0,0,0)	3909.601	3909.618	3918.771
ARIMA (1,0,0)	3912.751	3912.784	3926.505
ARIMA (0,0,1)	4008.149	4008.182	4021.903
ARIMA (0,0,0)	3911.785	3911.79	3916.37

The output from Table 8 indicates that ARIMA (0,0,0) with zero mean is selected as a best model because it has least AIC and BIC values. The ARIMA(0,0,0) model suggests that Bitcoin's returns during the study period exhibited random behaviour, making it challenging to predict future returns using only historical data as the market behaves like white noise.

Table 9: Accuracy for forecasting daily Bitcoin Returns

Error Matrices	Bitcoin Returns
ME	0.27333
RMSE	3.60091
MAE	2.49341
MPE	-14.3370
MAPE	462.8137
MASE	0.65259
ACF1	-0.03420

Table 9, shows the forecast accuracy for the original daily returns of Bitcoin time series using ARIMA (0,0,0). From the results, the different error measurement shows that the forecast accuracy of our model is accepted. This is due to the MASE value (0.6526) which is less than 1, and it implies that ARIMA (0, 0,0) performs better than a naïve model in this case study.

4.0 DISCUSSION

The first approach to Box-Jerkins technique is to determine if the order of differencing is important to make the data stationary for forecasting. It was observed from Figure 1 that ACF are all statistically significant since all the spikes are outside the 95% confidence bounds. The non-stationarity of the data showed that the prices of Bitcoin are unstable. This finding supports that of Mallqui (2019), who attested to the fact that the behaviour of Bitcoin is constantly unpredictable, and price prediction has a vital role because money may quickly lose value depending on this price forecast.

Again, after the first differencing, both ACF and PACF were found to be statistically significant which means that first difference is enough to render the dataset stationary. In line with this ARIMA (0,1,0) was recommended as the best model to forecast the future prices of bitcoin. This finding is in agreement with Stenqvist and Lonno (2017), who searched through 2.27 million tweets on bitcoin to see if they could predict a change in the cryptocurrency price in the coming days. Based on the intensity of sentiment swings from one period to the next, the anticipated model indicates that aggregating tweet sentiments over a 30-minute period with four shifts forward and a sentiment-limited fluctuation of 2.2 percent will result in an 83 percent accuracy estimate. NAMIN (2018) indicated that the LSTM is considerably superior at forecasting time series, particularly Bitcoin price projections than ARIMA due to high volatility. Finally, this study also predicted the daily returns of Bitcoin and realized that ARIMA (0,0,0) was the best model to forecast since the actual data was stationary. This implies that the data for returns of Bitcoin under the period of study is white noise (random shock). Again, the forecasts of the price of bitcoin would provide essential information to traders, investors, decision-makers, and the cryptocurrency industry. They support risk management, return maximization, and well-informed investment choices. These estimates would also guide investors on how to select viable businesses in the cryptocurrency market.

However, the linear forecasting technique ARIMA will not adequately capture the nonlinear patterns in Bitcoin price changes or the presence of any seasonality. In order to address these challenges, the researchers recommended more advanced techniques such as GARCH models, neural networks, Prophet, ARIMA with Exogenous Variables (ARIMAX), etc. for future research. These models take into consideration seasonal impacts, nonlinear patterns, and fluctuating variances.

REFERENCES

Adcock, R., & Gradojevic, N. (2019). Non-fundamental, non-parametric Bitcoin forecasting. *Physica A: Statistical Mechanics and its Applications*, *531*, 121727.

Chen, Z., Li, C., & Sun, W. (2020). Bitcoin price prediction using machine learning: An approach to sample dimension engineering. *Journal of Computational and Applied Mathematics*, *365*, 112395.

Ciaian, P., Rajcaniova, M., & Kancs, D. A. (2016). The economics of BitCoin price formation. *Applied economics*, *48*(19), 1799-1815.

Dyhrberg, A. H. (2016). Bitcoin, gold and the dollar–A GARCH volatility analysis. *Finance Research Letters*, *16*, 85-92.

Howrey, E. P. (1980). The role of time series analysis in econometric model evaluation. In *Evaluation of econometric models* (pp. 275-307). Academic Press.

Huisu, J., Lee, J., Ko, H., & Lee, W. (2018, August). Predicting bitcoin prices by using rolling window LSTM model. In *Proceedings of the KDD data science in Fintech Workshop, London, UK* (pp. 19-23).

Mallqui, D. C., & Fernandes, R. A. (2019). Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques. *Applied Soft Computing*, *75*, 596-606.

Matta, M., Lunesu, I., & Marchesi, M. (2015, June). Bitcoin Spread Prediction Using Social and Web Search Media. In *UMAP workshops* (pp. 1-10).

McNally, S., Roche, J., & Caton, S. (2018, March). Predicting the price of bitcoin using machine learning. In *2018 26th euromicro international conference on parallel, distributed and network-based processing (PDP)* (pp. 339-343). IEEE.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Decentralized business review*, 21260.

Siami-Namini, S., & Namin, A. S. (2018). Forecasting economics and financial time series: ARIMA vs. LSTM. *arXiv preprint arXiv:1803.06386*.

Stenqvist, E., & Lönnö, J. (2017). Predicting Bitcoin price fluctuation with Twitter sentiment analysis.

Tarasova, T., Usatenko, O., Makurin, A., Ivanenko, V., & Cherchata, A. (2020). Accounting and features of mathematical modeling of the system to forecast cryptocurrency exchange rate. *Accounting*, *6*(3), 357-364.

Yamak, P. T., Yujian, L., & Gadosey, P. K. (2019, December). A comparison between arima, lstm, and gru for time series forecasting. In *Proceedings of the 2019 2nd*

International Conference on Algorithms, Computing and Artificial Intelligence (pp. 49-55).

Zhu, Y., Dickinson, D., & Li, J. (2017). Analysis on the influence factors of Bitcoin's price based on VEC model. *Financial Innovation*, *3*(1), 1-13.