EXPLORING BITCOIN PRICE PREDICTION MODELS

Mojmír Sabolovič, Emina Hukic

Abstract: *Bitcoin is a digital currency based on Blockchain technology and not controlled by any government. In this paper, we undertake economic and econometric modeling of Bitcoin prices. The main assumption of the paper is that Bayesian regression should forecast future values of Bitcoin with greater accuracy than ARIMA (1, 2, 2). The predictions yielded the return of 89%. By turning away from conventional use towards different approach such a Bayesian approach is could result in models with greater predictive accuracy that would be significant to the financial world. Obtained forecasts are compared using Mean Squared Error of Prediction and Mean Absolute Percentage Error. In both categories, Bayesian linear regression provides better results, i.e. smaller deviations from actual values for a given period. However, the Bayesian model which used only time series of Bitcoin Close price yielded the worst results among the three models. ARIMA (1, 2, 2) ranked second but with errors 7 times higher than of Bayesian regression. Models could be further improved by incorporating external variable into the modeling historical Bitcoin price data.*

Keywords: *Bitcoin, Price Prediction, ARIMA, Bayesian linear regression*

JEL classifications: *G1, C22, C5*

1 INTRODUCTION

Bitcoin is very complex topic, covering cryptography, economics and software engineering. Bitcoin is a digital, decentralized, partially anonymous currency, not backed by any government or other legal entity, and not redeemable for gold or another commodity (Grinberg, 2011). It was created in 2007 by man under pseudonym Satoshi Nakamoto. Two years later, on January 12, 2009, the first Bitcoin transaction took place. The exchange rate for Bitcoin was established in October 2009 where at that time one U.S. dollar equaled 1,309.3 Bitcoin. Bitcoin operates through a peer-to-peer network system without any control imposed by person nor institution, such as the central bank or government. The code is open source, meaning that it belongs to the public domain, therefore it can be controlled by "anyone". Kristoufek (2013) inspects relationship between Bitcoin price and interest in currency measured by online searches. Relationship between the price of Bitcoin and financial indicators such as Dow Jones Index and oil price, and supply and demand forces of this cryptocurrencies is studied by Ciaian, Rajcaniova and Kancs (2014). Volatility of Bitcoin has mainly been investigated using different Generalized Autoregressive Conditional Heteroskedasticity models by Katsiampa (2017), and Cermak and Chen et al. (2016). Bouoiyour and Selmi (2014) analyzed daily Bitcoin prices using GARCH-optimal model and concluded that volatility has decreased in 2015 compared with earlier years of Bitcoin. The heart of the Bitcoin network is Blockchain technology, which is an open public ledger with the purpose of recording transactions. All financial information and transactions occurring are publicly available, except the identities of parties involved in the transaction. With blockchain technology, contracts are transparent and protected from tampering, revision, and deletion. Distributed ledger system keeps all of the data synchronized between millions of systems, meaning that database in one central location that can be hacked is the matter of the past. In the past several years Bitcoin has caught the attention of the general public, governments, and investors due to its efficiency, low transaction costs, and Blockchain technology. The popularization of Bitcoin has brought important questions and polemics on issues of this cryptocurrency. Investors and governments are interested in defining the Bitcoin and discovering it's proper use in financial markets and portfolios. Therefore, understanding of Bitcoin price movement and volatility is the crucial aspect of creating regulations for its formal use in economies worldwide. With the discovery of the driving forces of Bitcoin, the risk of using and investing with this cryptocurrency can be reduced. Furthermore, by virtue of the public ledger system, that Bitcoin uses, greater transparency in financial transactions can be achieved. Therefore it is highly significant to comprehend Bitcoin and to truly understand the way it works in order to be formally useThe rising interest to research on properties and price formation of Bitcoin, as well as various statistical analyses of Bitcoin, is identified since Bitcoin was created the decade back. Kristoufek (2013) inspects the relationship between Bitcoin price and interest in currency measured by online searches. The relationship between the price of Bitcoin and financial indicators such as the Dow Jones Index and oil price, and supply and demand forces of this cryptocurrencies is studied by Ciaian, Rajcaniova, and Kancs (2014). The volatility of Bitcoin has mainly been investigated using different Generalized Autoregressive Conditional Heteroskedasticity models by Katsiampa (2017), and Cermak and Chen et al. (2016). Bouoiyour and Selmi (2014) analyzed daily Bitcoin prices using

GARCH-optimal model and concluded that volatility has decreased in 2015 compared with earlier years of Bitcoin. Research is conducted to predict the price movement of Bitcoin using Autoregressive Integrated Moving Average Model and Bayesian linear regression model. The motivation for the research is derived from the survey on using Bayesian regression for predicting the price of and Bitcoin by Shah and Zgang (2015). By utilizing Bayesian inference for “latent source model” was developed the trading algorithm which identifies patterns and trades accordingly. Research experiment yielded a successful trading strategy where in 50 days the return was around 89% with a Sharpe ratio of 4. Research is conducted to predict the price movement of Bitcoin using Autoregressive Integrated Moving Average Model and Bayesian linear regression model. Motivation for the research is derived by survey on using Bayesian regression for predicting the price of and Bitcoin by Shah and Zgang (2015). By utilizing Bayesian inference for “latent source model” was developed trading algorithm which identifies patterns and trades accordingly. Research experiment yielded successful trading strategy where in 50 days the return was around 89% with Sharpe ratio of 4.10. Autoregressive Integrated Moving Average model (ARIMA) is the one of the most popular and frequently used for prediction of stochastic time series (Asteriou & Hall, 2007). The idea behind this paper is to compare between ARIMA and Bayesian linear regression models when it comes to Bitcoin price prediction. Particularly, the interest of the research is to explore which model has better predictive power and better overall fit to time series data on Bitcoin price. The main assumption of the paper is that Bayesian linear regression has better predictive accuracy than ARIMA models.

2 MATERIAL AND METHODS

Ciaian, Rajcaniova, & Kancs (2015) identified characteristics of Bitcoin as currency with low transaction costs, learning spillover effects, high anonymity, privacy and no inflationary pressures. Limitations of Bitcoin arising from the nature of Bitcoin were identified as the absence of an institution enforcing dispute resolution, the absence of Bitcoin-denominated credits, deflationary pressure, extremely high price volatility, and issues with cybersecurity. Buchholz et al. (2012) argue that Bitcoin price is determined as the outcome of the interaction between supply and demand. Kristoufek (2013) in his study states that the Bitcoin price formation cannot be described by standard economic theories because demand for Bitcoin is driven by investors’ speculative behavior and because Bitcoin is not issued by a government or central bank, therefore, detaching it from the real economy. Wjik (2013) analyzes the role of global financial development on Bitcoin price formation. Ciaian, Rajcaniova, & Kancs (2015) concluded that market forces (supply and demand) are key drivers of Bitcoin price formation, in particular, demand-side drivers, such as the size of Bitcoin economy and velocity of its circulation have the greatest impact on Bitcoin price. The hypothesis that speculation and attractiveness of Bitcoins to investors affects its price was not rejected. Speculative trading is beneficial activity in terms of absorbing excess risk and providing liquidity to the market. The crucial finding is that macro-financial indicators are not supported as Bitcoin price drivers. Kristoufek (2013) concluded that standard fundamental factors (usage in trade, money supply, and price level) have the significant role in Bitcoin price over the long term. The interest in cryptocurrency by investors is one of the main drivers of its price movement, having an asymmetric effect during the bubble formation and bursting. During the bubble formation, interest boosts the prices further, and during the bursting, it pushes them lower. Balcilar, Bouri, Gupta, & Roubaud (2017) employed non-parametric causality-in-quantiles test in order to analyze the causal relation between trading volume and Bitcoin returns and volatility, over the whole of their respective conditional distributions. When the market is operating around normal mode, a volume can predict returns and provide the investors with valuable predictive information.

When the market is in bull or bear phase, information about volume does not offer a relevant prediction. Garcia, Tessone, & Perony (2014) used autoregression techniques and identified two positive loops that led to price bubbles. One feedback loop was driven by word of mouth and second by new Bitcoin adopters. Spikes in information search, associated with external events, precede drastic price declines. Amjad & Shah (2016) used historical time series for trading strategy and price prediction. Authors developed of the theoretical framework for time series analysis based on generic properties of a time series (stationarity and mixing), and design of the real-time algorithm for prediction and training that yielded high prediction accuracy and highly profitable returns on investment. Then Later, through the paper, the comparison between ARIMA and Bayesian linear regression is done in order to investigate which model has a greater predictive power of Bitcoin prices.

2.1 Autoregressive Integrated Moving Average Model

An Autoregressive Integrated Moving Average process (ARIMA) is the mathematical model used for forecasting time series (Box & Jenkins, 1976). ARIMA is derived from Mixed Autoregressive Moving Average; it is a combination of Autoregressive process AR(p) and Moving Average process MA(q). In AR(p) component, the future value of a variable is assumed to be the linear combination of p past observations and random error term together with the constant term. The number p in parenthesis denotes the order of the autoregressive process and therefore the number of lagged dependent variables that the model will have (Asteriou & Hall, 2007).

AR(p) process can mathematically be estimated by equation (1) (Montgomery, 2008)

$$y_{t=\delta+}\phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_{t=} \delta + \sum_{t=1}^{p} \phi_i y_{t-i} + \varepsilon_t \ , \tag{1}$$

where y_t represents actual value and ε_t relates to random error or random shock at time period t. Model parameters are represented by φ_i *(1,2,3...p)* and δ is a constant term (Montgomery, 2008). The implication of *AR (p)* model is that behavior of y_t is determined to large extent by its own value in preceding period *t-1*. Moving average *(q)* process uses past error terms as explanatory variables.

MA (q) model represents linear regression of the current observation of time series against the random shocks of one or more prior observations. *MA (q)* model implies that future value of time series *(y_t)* is largely determined by random process. Random shocks are assumed to be a sequence of independent and identically distributed random variables with zero mean and constant variance; random shocks are assumed to be white noise process following the normal distribution. The model is articulated by equation (2)

$$y_t = \mu + \theta_1 \epsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} = \sum_{t=1}^{q} \theta_i \varepsilon_{t-1}, \tag{2}$$

where *Bitcoin* represents actual value, θ_i *(1,2,3...p)* is the model parameter and ε_{t-i} represents a random error, shocks. Equation (2) implies that value of time series depends on random shock of past observations. Because any *MA(q)* process is, by definition, an average of q stationary white-noise processes, it follows that every moving average model is stationary, as long as q is finite (Asteriou & Hall, 2007). Stationary time series with complex autocorrelation behavior are more adequately modeled by ARMA processes than by either pure

AR(p) or *MA(q)* process (Ruppert & Matterson, 2015). ARMA (p,q) model is expressed by equation (3)

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} = \delta + \sum_{t=1}^{p} \phi_i y_{t-i} + \varepsilon_t - \sum_{t=1}^{q} \theta_i \varepsilon_{t-1} \quad (3)$$

In reality, much economic time series behave as though they had no fixed mean (e.g. stock prices). These types of industrial and economic time series are demonstrating specific kind of homogenous nonstationary behavior. This kind of time series can be represented by the stochastic model modified form of the autoregressive moving average process. The first difference of time series ($w_t = y_t - y_{t-1} = (1 - L)y_t$) or higher order differences ($w_t = (1 - L)^d y$) produce stationary time series (Montgomery, 2008). The mathematical formulation of ARIMA(p,d,q) model using lag polynomials is represented by following equations (4)

$$\left(1 - \sum_{i=1}^{p} \phi_i L^i\right)(1 - L)^d y_t = \left(1 + \sum_{j=1}^{q} \theta_j L^j\right)\varepsilon_t, \quad (4)$$

where *p,d* and *q* are integers greater than or equal to zero and refer to the order of autoregressive, integrated and moving average parts respectively. Integer *d* refers to differencing of time series and controls level of differencing.

2.2 Bayesian statistical approach

Bayesian approach requires sampling model and prior distribution on all unknowns in the model including missing data and parameters. Prior distribution and likehood are then used to compute posterior distribution, i.e. conditional distribution of the unknowns given the observed data (Carlin & Louis, 2009). Following mathematical expression (5) represents Bayes' rule

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{P(A|B_1)P(B_1) + \cdots + P(A|B_K)P(B_K)}, \quad (5)$$

where *A* and *B* are events and probability $P(B_j) \neq 0$, $P(B_j|A)$ is conditional probability where the likelihood of event *A* occurring given that *B* is true.

Widely used in practical application and well known non-informative prior is Jeffreys prior. This prior is invariant under reparametrization of θ and is defined as proportional to the square root of the determinant of the Fisher information matrix (6) (Koduvely, 2015)

$$P(\theta) \propto \sqrt{\det I\ (\theta)}. \tag{6}$$

After determining a prior distribution of data, the key step in the Bayesian analysis is the use of Bayes's theorem to combine the prior knowledge about $\boldsymbol{\theta}$ with the information in the data. The likelihood is defined in the same way in a non-Bayesian analysis, but in Bayesian statistics, the likelihood has a different interpretation—the likelihood is the conditional distribution of the data $\boldsymbol{\theta}$**.** (Ruppert & Matterson, 2015**).** The likelihood function is written as f(*Bitcoin*| θ). The joint density of θ and Bitcoin is the product of prior and the likelihood (7) (Ruppert & Matterson, 2015)

$$f(y,\theta) = \pi(\theta)f(y|\theta). \tag{7}$$

The marginal density of BITCOIN is found by integrating θ out of joint density (8)

$$f(y) = \int \pi(\theta)f(y|\theta)d\theta. \tag{8}$$

The conditional density of θ given *BITCOIN* in following equation represents form of Bayes's theorem where density on left side represents posterior density. That posterior distribution gives the probability distribution of θ after observing the *data (BITCOIN)*, see equation (9)

$$\pi(\theta|Y) = \frac{\pi(\theta)f(Y|\theta)}{f(Y)} = \frac{\pi(\theta)f(Y|\theta)}{\int \pi(\theta)f(y|\theta)d\theta} \tag{9}$$

Recent developments in computing methods and statistical software, especially advancements in Monte Carlo computing methods allow accurate computations of complex integrals, thus permitting advanced Bayesian analysis to be done. Estimation and uncertainty analysis in Bayesian approach is based upon the posterior distribution. Commonly used summaries of location are the mean, median, and mode(s) of the distribution; variation is commonly summarized by the standard deviation, the interquartile range, and other quantiles. (Charlin &

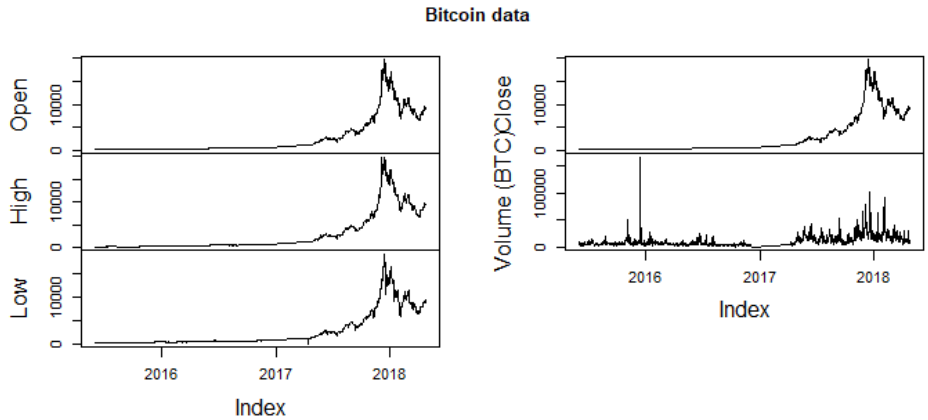
Louis, 2009). Mean is the posterior expectation of the parameter, while mode can be understood as "most likely" value given the data and the model. The mode is called maximum posterior estimator (MAP).

When it comes to hypothesis testing, the Bayesian approach is much simpler and more sensible in principle than traditional hypothesis testing. In the Bayesian hypothesis testing, there can be more than two hypotheses taken into consideration, and they do not necessarily stand in an asymmetric relationship. (Levy, 2007) Bayesian analysis generates probability values that are used to study relative support for one hypothesis over another. Briefly, Bayesians seek probability support for hypothesis while frequentist is searching for significance. A version of a t-test, a probability of H0 and alternative hypothesis, in Bayesian approach is statistics called Bayes factor (BF) which represents a ratio that compares the likelihood of one model over another.

3 DATA

Time series data on the price of Bitcoin were obtained from Coinbase Exchange. The exchange is available in 33 countries and as of 2017, it was the World Largest Bitcoin broker. In order to apply methods and techniques for forecasting time series, raw data is divided into two parts. Training set were observations of Bitcoin Close price with daily frequency from May 30, 2015, to May 30, 2018. The test set comprises from daily observations of Bitcoin Close price from May 31, 2018, to April 30, 2018.

Figure 1: Historical Bitcoin daily data



Source: author's elaboration

As can be noted from Figure 1. Bitcoin had the substantial increase in its price from the mid- 2017, with record high at the end of 2017 of $ 19,650. This digital currency has begun a year with the price under $ 1,000, experiencing growth in value by more than 1300%. Increased interest in Bitcoin started in May 2017, period known as the summer of bulls. A decision by the government of Japan on April 1, that year to declare Bitcoin as legal currency unquestionably assisted growth in price. Rising Significance of BITCOIN was recognized by Commodities Futures Trading Commission on December 1, 2017. Namely, CFTC approved Bitcoin future which allows investors to speculate about future value without "touching" the coin. According to Bloomberg, the fact that there will be finite supply added to increased investments in Bitcoin, in order not to miss the opportunity. However, after reaching high and period of substantial growth, at the beginning of the 2018 price of Bitcoin started to decline. For example, on April 5 dropped to $ 6,600, and in the period from February price of Bitcoin struggles around 7,000 USD. Increased regulation on Bitcoin that multiple countries have pursued and bankruptcy of Mt. Gox exchange contributed to decrease in the price of the coin. Furthermore, the rumour that Finance exchange has been hacked additionally shook the stability of the Bitcoin price.

Table 1: Summary of Bitcoin training data

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Close Price	211.2	386.1	597	1938.6	2347.5	19650
Volume	683.8	5053.9	7063.5	10527.1	12264.7	165542.8

Source: author's elaboration

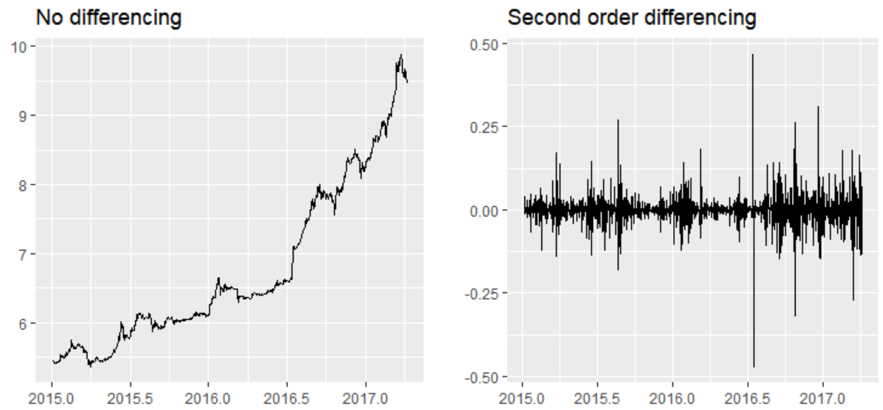
Furthermore, summary statistics in Table 1. displays broad range of Bitcoin Close price data for inspected period, implying great interest for Bitcoin that evolved in short period. Values of standard deviation and variance, 3224.687 and 10398607 respectively are showing substantial dispersion of USD values of Bitcoin prices. These behavior that data exhibits support the need for logarithmic transformation of series. In order to get more homogenous variance across sample, logarithmic transformation of the data is performed to stabilize variance hence getting more adequate model for forecasting.

Results

3.1 ARIMA (1,2,2) Model

ARIMA was created by following Box-Jenkins Methodology. The stated methodology does not assume any specific pattern in historical time series observation but uses an iterative approach which comprises model identification, parameter estimation, and diagnostic checking (Box & Jenkins, 1976). The modelling is done using software RStudio (Box & Jenkins, 1976), specifical packages such as series, TTR, forecast. Quandl, dev tools, ggplot2, etc. Model identification refers to the inspection of the time series to determine an order of p, d, and q components in the ARIMA model. By performing Augmented Dickey-Fuller (ADF) test and Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) test log Bitcoin price series was tested for stationarity and used to determine the level of difference, i.e. d component of ARIMA (p, d, q) model. ADF tests for non-stationarity of time series by following procedure based on the presence of unit root. Results of test on log Bitcoin price, both suggest that the observed time series is not-stationary. After first order differencing of time series data, results of tests are offering mixed results. ADF rejects a null hypothesis of the presence of unit root with the p-value lower than 0.01, while KPSS rejects the hypothesis of stationarity of time series. Since results of ADF and KPSS test were inconclusive, second order differencing was performed, see Figure 2. Both KPSS and ADF tests yielded results that Bitcoin price historical data appears to be stationary. Therefore, the d component of the ARIMA model equals 2. Box Jenkins methodology provides a way to identify the ARIMA model according to autocorrelation and partial autocorrelation graph of the series, making AFC and PACF the core of ARIMA modelling.

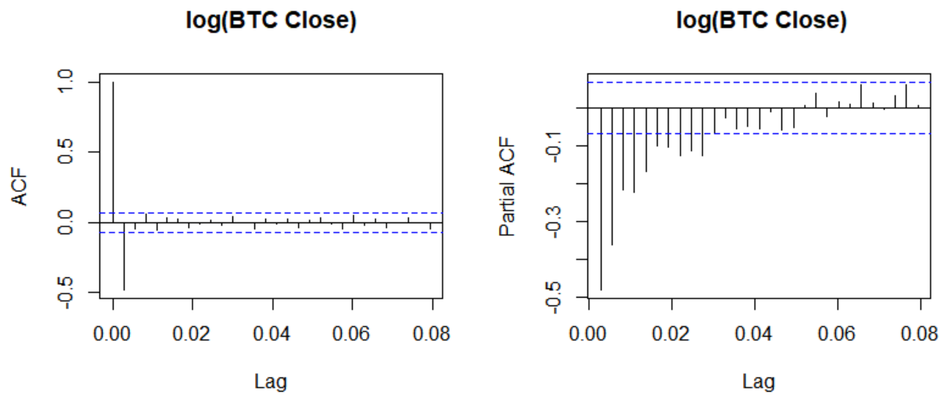
Figure 2: Representation of differentiation of training data



Source: author's elaboration

By examining AFC a PACF plot (Figure 4) on log Bitcoin Close price to identify ARIMA (p, d, q) model, several possibilities of p and q values are emerging. However, ACF and PACF plots are narrowing the choice of models on ARIMA (1,2,2) model or ARIMA (2,2,2) model. ACF plot cuts off after lag *2* and PACF plot is exponentially decaying implying that order of moving average operator is *2*, i.e. *q=2*. However, regarding the autoregressive operator, it is unclear whether p should be 1 or 2, because of the large spikes on first and second lag. Therefore, to examine which model is the better fit for the data Akaike Information Criterion (AIC) is examined further.

Figure 4: Autocorrelation and partial autocorrelation plots of log (BITCOIN Close) price



Source: author's elaboration

Model selection step refers to the choice of statistical model that best describes data among several competing models (Sinharay, 2010). Comparing AIC values from different models, the one with lowest value of Akaike Information Criterion is considered to be “best fit”. Yang (2005) suggests that AIC is asymptotically optimal in selection of the model, under the assumption that true model is not in the candidate set, as virtually it is always the case in the practice (Snipes, 2016). It is important to note that AIC score is ordinal and means nothing on its own. AIC score is calculated as follows (Box & Jenkins, 1976):

$$AIC_{p,q} = \frac{-2\ln(maximized\ likelihood)+2r}{n} \approx \ln\left(\widehat{\sigma^2}\right) + r\frac{2}{n} + constant, \tag{10}$$

where $\widehat{\sigma^2}$ is the maximum likelihood estimate of variance and r $(r = p+q+1)$ is the number of estimated parameters, including the constant term. Model selection step refers to the choice of statistical model that best describes Using command arima, under package stats, several ARIMA models have been explored in order to find the model that provides the best fit to the historical observations of Bitcoin prices. Results of the ARIMA modelling, with AIC, are given in Table 2. Results are showing that ARIMA (1,2,2) model is the better fit to the historical data of Bitcoin close price since it has the lowest value of AIC. However the values of AIC criterion for ARIMA (1,2,2) and ARIMA (2, 2, 2) are not differing that much, -3235.47 and -3134.78 respectively, which can be explained that ARIMA (2,2,2) can be used as well for explaining the behaviour of the Bitcoin price movement from May 30, 2015, to April 30, 2018. Other models from Table… have comparable AIC values as well and could be taken into consideration for modelling Bitcoin price. But, using the criteria stated above that model with the lowest AIC value should be taken as “best” model, which in this case is ARIMA (1, 2, 2).

Table 2: Summary statistics of ARIMA (1, 2, 2)

```
Coefficients:

          ar1      ma1      ma2
      -0.3747  -0.6264  -0.3736
s.e.      NaN      NaN      NaN

sigma^2 estimated as 0.001627:  log likelihood = 1681.85,   aic = -3355.7

Training set error measures:
                       ME        RMSE         MAE         MPE       MAPE       MASE
ACF1
Training set 0.0008944974 0.04029053 0.02614424 0.01183182 0.3536577 0.9943218 -
0.009569967
```

Source: author's elaboration

In order to validate the model further, residuals of ARIMA (1, 2, 2) are examined. To assume that model is the true process generating the data, then the observed residuals should be realized values of white noise sequence.[60] Representation of residuals in Figure 4. shows that residuals from ARIMA (1, 2, 2) are exhibiting random behavior similar to white noise. Next, the autocorrelation function plot and partial autocorrelation plots of observed residuals should lie within the $\pm 1.96/\sqrt{n}$ roughly 95% of the time. If the correlations are substantially more than 5% outside of the range, then the better-fitting model should be introduced.[61]The interval is marked as dashed blue line. ACF and PACF of residuals, Figure 5.; are showing that for observed residuals that resulted from ARIMA (1,2,2) model correlation spikes are arranged within the desired range. Furthermore, compatibility of the distribution of the residuals with normal distribution or t-distribution is checked by examining corresponding density plot. In addition, the Box-Ljung test provides the different approach to double check the model. Box-Ljung test is meant to test the autocorrelation in which it should be verified whether the autocorrelations of a time series are different from 0. The test is applied to the residuals of fitted time series by ARIMA (p, d, q) model. Since Box-Ljung test examines autocorrelation o the residuals, is said that model does not exhibit lack of fit to the data if values of autocorrelations are very small.[62] This means that there still remains serial correlation in the series and that modification of the model is necessary. The null hypothesis is that the model does not exhibit a lack of fit.

Table 3: Box-Ljung test of goodness of fit

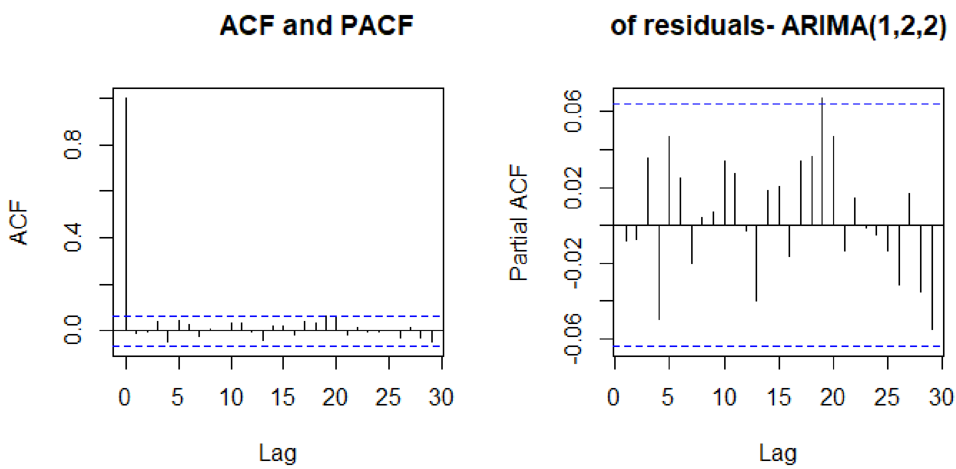
```
        Box-Ljung test

data:  ARRRR$residuals
X-squared = 5.8067e-05, df = 1, p-value = 0.9939
```

Source: author's elaboration

Table 3. represents test statistics of residuals from ARIMA (2, 2, 2) model. Large p-value, p=0.9939, indicates that the null hypothesis is no rejected, meaning that this model does not require further modification and that there is a fit between Bitcoin price data and tested model.

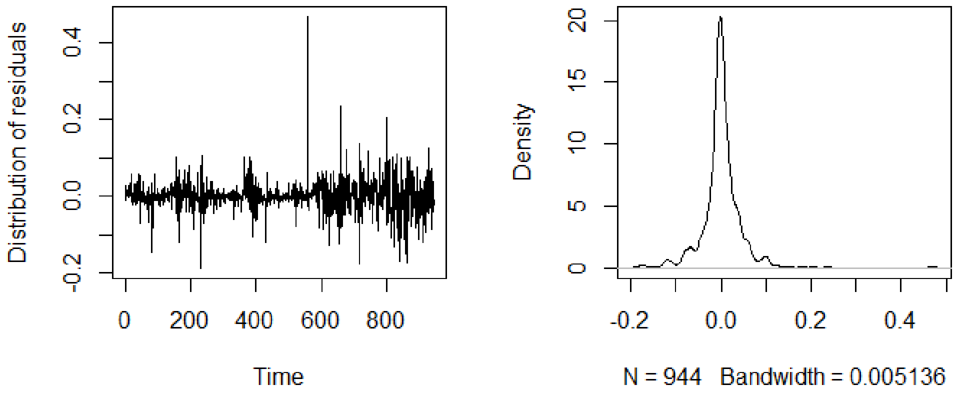
Figure 5: Autocorrelation and partial autocorrelation plots of residuals



Source: author's elaboration

After the identification phase, parameters of the model are estimated. Prediction of the future value of Bitcoin close price was done with RStudio, particularly with command predict. The forecast was done to estimate 30 future values of Bitcoin close price. However, since the original time series was transformed by natural logarithm, inverse logarithm was applied to predicted values. The reason is that it can be easily compared with real values. Also, at this point of modeling of ARIMA (1,2,2) to Bitcoin close price, test dataset will be used, comprising of actual values.

Figure 6: Distribution and density plot of residuals

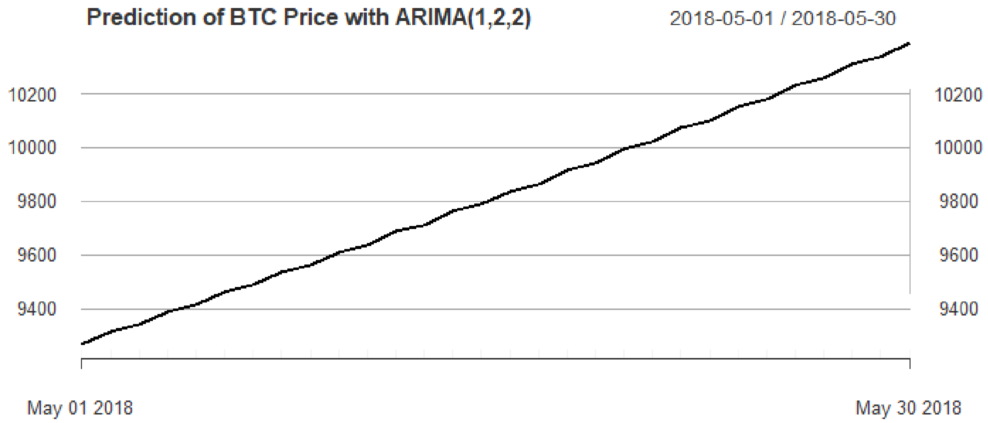


Source: author's elaboration

Figure 6 represents the density plot of observed residuals which shows that mean is close to zero and the shape of distribution suggests that the assumption of Gaussian white noise process is not so unreasonable.

Figure 7. graphically represents forecasted values for a given period. As can be noted from the figure, the prediction of ARIMA (1,2,2) on Bitcoin close price, yields results that are not expected. Namely, the forecast of the Bitcoin price for 30 days is exhibiting an upward linear trend. Obtained forecast seem bit unrealistic, especially when considering the movement of historical Bitcoin price which does not reflect the behavior of estimated values. These predicted values are much of surprise, firstly because according to the test criteria performed on the ARIMA (1, 2, 2,), the model satisfies the criteria and results from tests stated above are implying that this particular model provides the good fit to the data. Despite results from test statistics, it seems that the model needs some kind of transformation and modification to be done. However, it is unclear which type of change model requires in order to obtain valid and more accurate predictions. One thought is that the time interval used in training data, for modeling the ARIMA, maybe somehow long for prediction of Bitcoin. That is, time horizon used can be the reason why prediction failed, because it incorporates a period of "price stagnation" together with a period of the rapid increase in the price of Bitcoin from May 2017. Another possibility is that there is another type of model, other than ARIMA, that can represent and predict the better behavior of Bitcoin price.

Figure 7: Forecast of BITCOIN price



Source: author's elaboration

Bakar & Rosbi (2017) shown that classical time series regression algorithms such as ARIMA could be used to forecast price changes, yet they have poor prediction performance of Bitcoin time series. Instead, the article proposes two different approaches for forecasting Bitcoin price, classification algorithms and directly learning empirical conditional distribution (EC). Both of the proposed models outperformed the ARIMA model. In model used for forecasting method produces reliable forecasting model.

3.2 Analysis and prediction of Bitcoin price using Bayesian models

Bayesian approach has unique characteristic over standard, frequentist approach. This approach has yielded very impressive results, where predictions based on Bayesian approach succeeded to nearly double the investment in less than 60-day period. The most widely used technique for Bayesian analysis simulates a Markov chain whose stationary distribution is posterior, then the sample from this chain is used for Bayesian inference. This method is known as Markov Chain Monte Carlo or MCMC. Majority of Bayesian MCMC computing is done by using one of the two basic algorithms, the Metropolis-Hastings, and the Gibbs sampler. MCMC methods are designed to successfully simulate values of X vector based on a strategy designed to eventually draw these values from the target, posterior distribution. (West & Harrison, 1997) A sequence of simulated values X_1, X_2,…, is generated by firstly specifying starting value, then sample successive values from specified transition distribution with density $f(X_i|X_i\text{-}1)$, for $i = 2, 3, . . .$; X_i is generated conditionally independently of X_{i-2}, X_{i-3},...

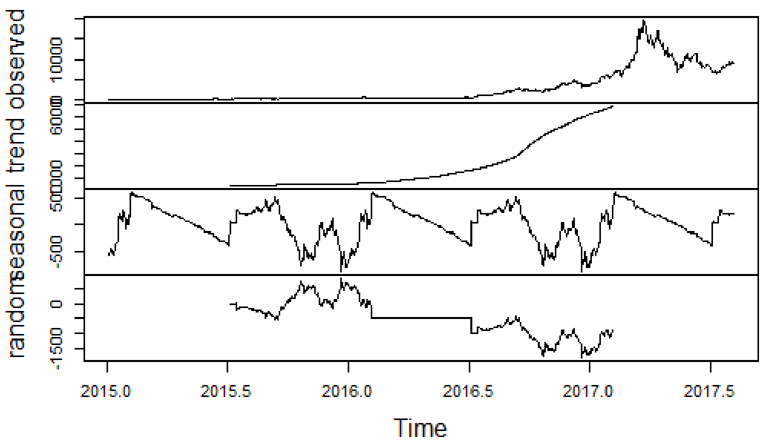
Modeling is done with help of RStudio using package but which performs time series regression using dynamic linear models fit using Markov Chain Monte Carlo. Firstly, the Bayesian model which only uses the time series of the Close price of Bitcoin. The second model is simple Bayesian regression using only a single model. Following the regression equation (11) was used

$$y_t = \mu_t + \tau_t + \beta^T X_t + \varepsilon_t, \quad (11)$$

where μ_t represents trend term, τ_t is seasonal component and $\beta^T X_t$ refers to regression component. Residuals are designated as ε_t, and they are assumed to follow Gaussian distribution.

When first model was created, seasonal trend is added, particularly 3 seasonal periods were added. From Figure 8 it can be noted that Bitcoin price exhibits seasonal behavior, one per year. Also, linear trend was added to the model because decomposition graph shows that Bitcoin Close Price experiences overall growth trend.

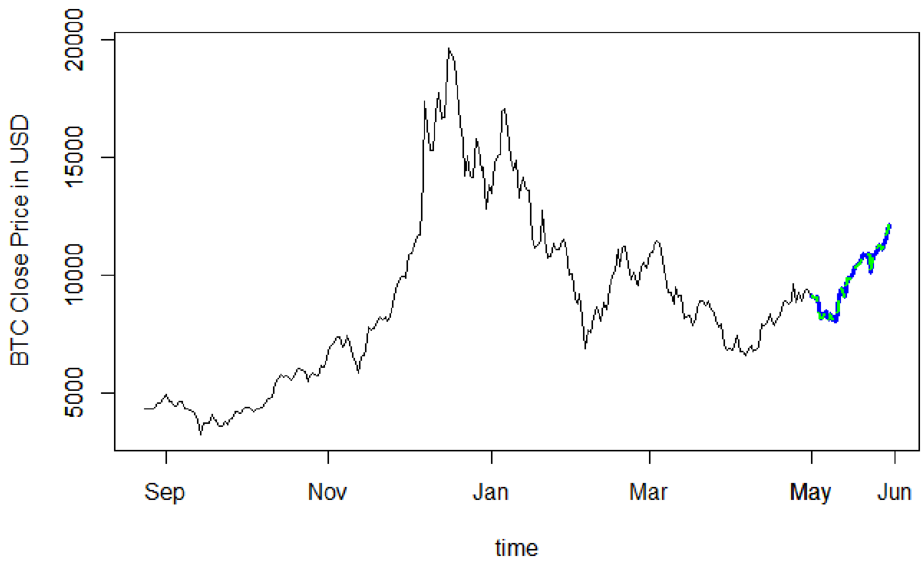
Figure 8: Decomposition of BITCOIN time series



Source: author's elaboration

After 50 MCMC iterations are performed to form a Bayesian model which uses just Close price observations, the model predicted for 30 periods, in order to reflect the length of the test data. Summary statistics of the model states that R squared value is 0.9994511, meaning that this model explains 99% of data behavior, which implies that the model itself fits Bitcoin price data. The standard deviation of residuals is 96.1782, and standard deviation of prediction is 330.7511. Figure 9 represents the plot of Bitcoin price data together with predictions for the following 30 days. Plotted predictions are implying that the price of Bitcoin should exhibit continuation of the downtrend at the beginning of May. However, predicted values are showing that during the May, as values are approaching June, Bitcoin should experience growth. To be exact predicted value for the end of May is $ 12,161.542.

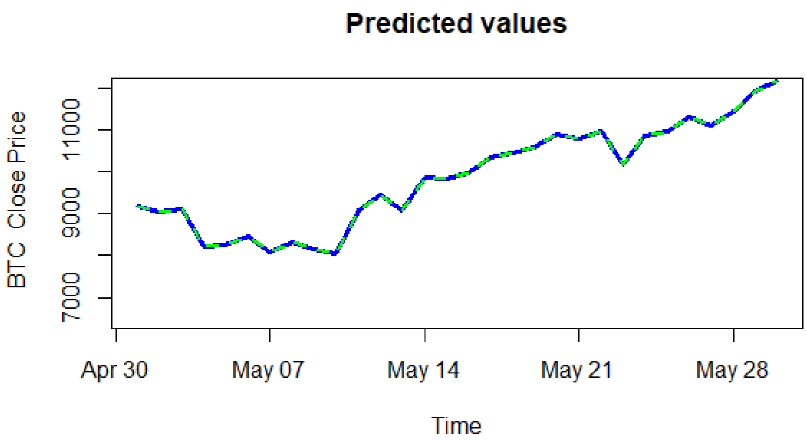
Figure 9: Prediction od Bitcoin price with Bayesian model



Source: author's elaboration

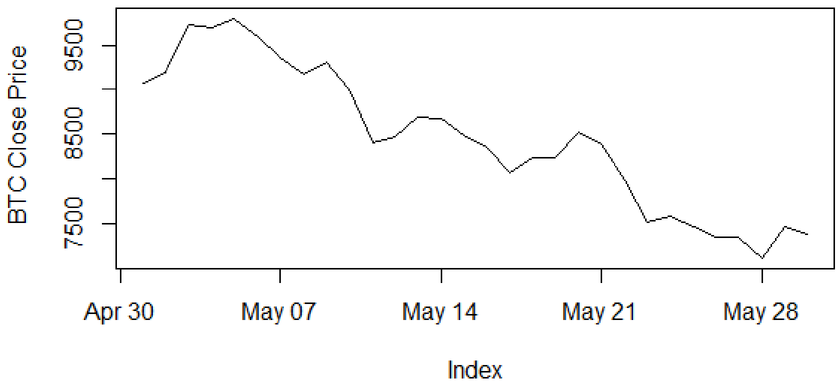
Figure 10 and 11 represents plots of predicted values of Bitcoin close price (upper figure), while lower graph represents actual values of Bitcoin close price for the period May 1 – May 30, 2018, which is test data.

Figure 10: Predicted Values



Source: author's elaboration

Figure 11: Test (actual) data

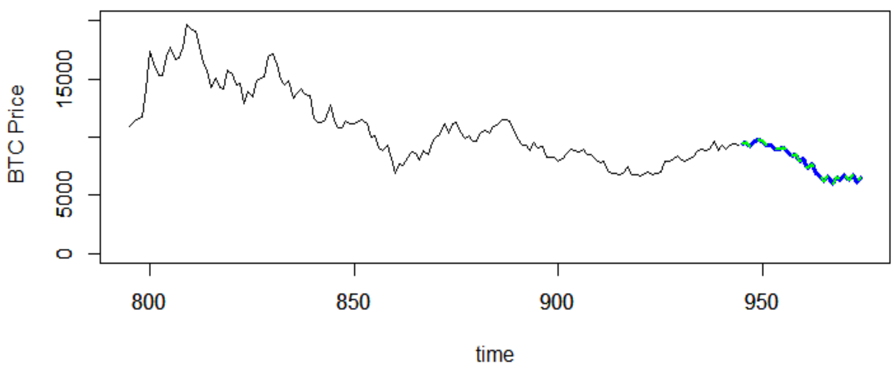


Source: author's elaboration

As Figure 11shows, for the first 7 days of May prediction values indicated that the value of Bitcoin is decreasing, the price dropped around $ 8,000. But, when looking at actual data, the Bitcoin price in that period actually went up, reaching almost $10,000. When plots are inspected further, it is clear that predicted values are not reflecting the behavior of test data. In fact, the model predicts an increase in the value of Bitcoin, while actual observed Bitcoin prices are exhibiting downtrend. The inaccuracy of predictions of the previous model may be caused by inappropriate time horizon of observed data or inadequate frequency of Bitcoin historical price data. In chosen time period for training data, May 30, 2015- April 30, 2018, Bitcoin price exhibits sudden growth in last quarter of 2017, after which in first quarter of 2018 loses almost double of its value. These sharp movements of price in long-term period probably affected the model resulting in the prediction that went in the opposite direction from test values. Furthermore, modeling was done with daily Bitcoin close price data, but the better solution may be to use more frequent observations, such as hourly or 15-min intervals. In the paper, Bayesian Regression and Bitcoin, authors have used frequency of 1-min intervals, which has yielded very accurate predictions resulting in “profitable“ trading strategy. However, this type of data is usually unavailable for modeling.

The second model, Bayesian linear regression, with a single model yielded somewhat better results than the previous model. For this model seasonal component and trend, the component was added as well. When comes to determination of the prior, modeling was done using non-informative prior. In other words, it is assumed that there is not enough information to determine prior distribution. Again 50 MCMC iterations were done prior to building the stated model. R- square statistics for the model is very high, 0.9998768, implying that model is the good fit to the data and explains the behavior of past observations. A standard deviation of residuals is 45.56947, which is the lower value than of the first Bayesian model. However, a standard deviation of prediction is slightly higher than for the first model and it is 334.6618. Following the plot, Figure 12 represents the prediction of the simplest Bayesian linear regression for 30 periods.

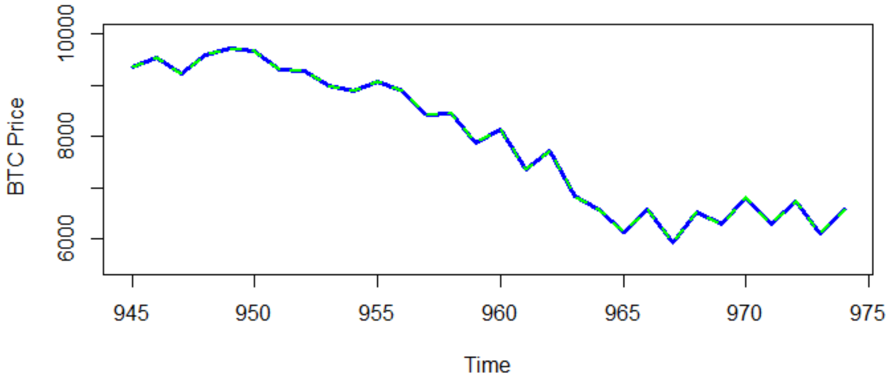
Figure 12: Bitcoin price prediction with Bayesian linear regression



Source: author’s elaboration

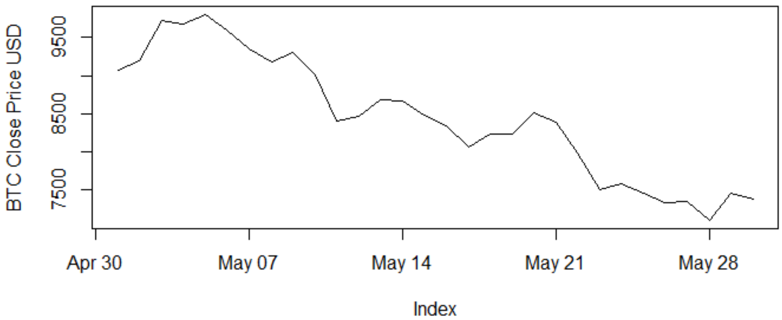
Figure 12 displays that predicted Bitcoin Close price values are exhibiting downtrend, forecasting that at the end of May price of Bitcoin should decline to $ 6,577. 026. When predictions are plotted against test data for the stated period, it is observed that forecast using Bayesian linear regression yielded results that are “reflecting” true values of Bitcoin price. Figure 13 represents Bitcoin price prediction (upper plot) while the second plot represents the actual values of the price of cryptocurrency.

Figure 13: Predicted Values



Source: author’s elaboration

Figure 14: Test (actual) Values



Source: author’s elaboration

By inspecting Figures 13 and 14 it is clear that so far Bayesian linear regression offered superior results, which are reflecting behavior of actual data for period May 1-May 31, 2018. Model predicted decline of Bitcoin value, from $ 9,359.663 to $ $6,577.026 at the end of the period. The value of Bitcoin using actual data shows that at the end of stated period 35 was $ 7,380.01. However, even though predictions are not precise, model predicted down trend for price of Bitcoin, which has happened with actual values. The Bayesian linear regression turns out to model Bitcoin price data accurately by predicting price movements that reflected actual observations.

In addition, model could be improved further by using more frequent data observations for Bitcoin close price and by incorporating some other regressor that influences behavior of Bitcoin. Because of the increased interest for Bitcoin and its increased use in business operations it is possible that some external factor, such as macro financial indicator or trending of Bitcoin among investors, should be added as explanatory variable in the model.

4 COMPARISON OF MODELS AND DISCUSSION

This section is dedicated to comparison of the ARIMA (1,2,2) model, the Bayesian model which uses just Bitcoin close price time series for prediction, and simple Bayesian linear regression. When the main purpose of the model is prediction then it is reasonable to select Mean square error (MSEP) of prediction and Mean Absolute Percentage Error (MAPE) as criteria for determining model quality (Wallach & Goffinet, 1989).

MSEP represents the average squared difference between the quantity of interest and the model prediction of that particular quantity. In other words, it is a measure of the predictive accuracy of the model. Mean square error of prediction is calculated using formula (12)

$$MSEP = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2, \tag{12}$$

where $(Y_i - \widehat{Y}_i)^2$ is squared difference between ith actual value of Y and the corresponding model forecast and n represents number of observations. The Table 7. represents Mean square error of prediction values for models presented in the paper, while Table 6. represents squared errors of prediction and absolute percentage error. As can be seen from the table the highest value of MSEP or lowest predictive accuracy belongs to the Bayesian model which only uses BTC time series for prediction. ARIMA (1, 2, 2) seems to have somewhat smaller MSEP but compared to the Bayesian linear regression model, its value is still relatively high meaning that predictions of the model were not so accurate as well. It is clear that simple Bayesian regression has the greatest accuracy on the prediction of Bitcoin price. MSEP of the model is more than 7 times smaller than of ARIMA (1, 2, 2). Therefore, under the MSEP criteria, it can be concluded that Bayesian regression fitted the Bitcoin price data the best, yielding results with highest predictive accuracy. Also, Bayesian regression model succeeded to predict the overall price movement of Bitcoin, while the other two models predicted movements in opposite direction from actual ones.

The Table 4 represents Mean square error of prediction values for models presented in the paper, while Table 6. represents squared errors of prediction and absolute percentage error. As can be seen from the table the highest value of MSEP or lowest predictive accuracy belongs to the Bayesian model which only uses BTC time series for prediction. ARIMA (1, 2, 2) seems to have somewhat smaller MSEP but compared to the Bayesian linear regression model, its value is still relatively high meaning that predictions of the model were not so accurate as well.

Table 4: Box-Ljung test of goodness of fit

Mean Square Error of prediction (MSEP)			Mean Absolute Percentage Error		
ARIMA (1,2,2)	Bayesian 1	Bayesian regression	ARIMA (1,2,2)	Bayesian 1	Bayesian regression
3117683.922	5834738.392	804994.9931	18.40436%	25.44960%	8.46927%

Source: author's elaboration

It is clear that simple Bayesian regression has the greatest accuracy on a prediction of Bitcoin price. MSEP of the model is more than 7 times smaller than of ARIMA (1, 2, 2). Therefore, under the MSEP criteria, it can be concluded that Bayesian regression fitted the Bitcoin price data the best, yielding results with highest predictive accuracy. Also, Bayesian regression model succeeded to predict the overall price movement of Bitcoin, while the other two models predicted movements in opposite direction from actual ones.

MAPE is widely used in practice because of its intuitive interpretation in terms of relative error. It measures the size of the absolute error in percentage terms. MAPE has an advantage of being scale independent allowing comparison of forecast performance of different data set. It is calculated using the following formula (13)

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \frac{|Y_i - \widehat{Y_i}|}{Y_i}, \tag{13}$$

where $|Y_i - \widehat{Y_i}|$ is absolute error between actual observation and forecasted value, i.e. Y_i is ith observation of real observed value and n is the number of observations.

Table 4. also summarizes MAPE for the three models discussed in the paper. MAPE values are confirming the evaluations of the models with MSPE. Again, a Bayesian model which uses only historical observations for prediction proved

to have lowest predictive power, with MAPE of 25.44960%. ARIMA (1,2,2) shows that its forecast is deviating on average 18.40436% from test data. Bayesian linear regression has the lowest MAPE value, confirming its position of the model with best predictive accuracy. When looking at absolute errors of predictions, Bayesian regression has very low percentage error, especially in the first 10 forecasted periods where absolute errors do not exceed 8%. price.

Considering MSEP and MAPE as criteria for determining the best model, it can be concluded that Bayesian linear regression has the greatest predictive power and accuracy, over the other two models discussed in the paper. Also, following logical reasoning, model forecasts mirror price movements of Bitcoin. Since the beginning of 2018, Bitcoin has lost almost 50 % of its value. Many debates are at the place about the cause of this drastic fall in Bitcoin price. The central argument in the financial world is that this sudden decline in the price of most influential cryptocurrency is tied to launch of Bitcoin futures contract on Chicago Mercantile Exchange. According to FED's researches, the launch of these types of futures allowed pessimists to enter the market, which contributes to a reversal of Bitcoin price dynamics. When taking into the consideration recent happenings and their effect on Bitcoin price movements, modeling of Bitcoin price further in research should include external factors. Research and studies about financial indicators as price drivers of Bitcoin are proposing mixed results. For example van Wijk in his study " What can be expected from Bitcoin?" concludes that several financial indicators such as Dow Jones Index, the euro-dollar exchange rate, and WTI oil price have the significant effect on the value of Bitcoin price in the long run. However, P. Cianian et al. (2015) found that global macro-financial developments do not significantly affect the price of Bitcoin. Furthermore, Yermack (2014) argues that Bitcoin's price is not responsive to macroeconomic variables and therefore is not effective as a tool for risk management. This means that Bitcoin cannot be hedged against other assets that are driven by macroeconomic developments. In recent period Bitcoin drove the attention of the public and investors. More and more people are buying Bitcoins and more and more firms are accepting Bitcoin as a medium of payment. Because of the recent popularity of Bitcoin among investors, it is important to understand what drives Bitcoin and which factors are influencing its price movements. Because of previously stated reasons, further research on Bitcoin should focus on determining external factors, especially on macroeconomic indicators, that are important for modeling and prediction of Bitcoin price. Even though there were no any significant results about a connection of financial indicators and Bitcoin in previous years, recent interest and developments showed that Bitcoin is integrating itself in a global financial system and pretends to become one of the global players in financial markets. Due to the integration of Bitcoin into the world markets, external factors are certainly starting to influence Bitcoin and its price movements.

5 CONCLUSIONS

Bitcoin price dynamic has been the live issue since cryptocurrencies caught attention and increase the interest of a wide audience. It is the most successful virtual currency in terms of its impressive growth in price as well as the number of currency users. As a result of recent developments in Bitcoin exchanges, there is an increasing need for understanding the behavior of Bitcoin and underlying characteristics. Identification of these factors would contribute to the efficient use of Bitcoin in financial markets. The paper approaches the modeling and prediction of Bitcoin Close price from two perspectives, by comparing the predictive accuracy of ARIMA (1, 2, 2) and Bayesian methods. The main assumption of research was that Bayesian linear regression better fit the Bitcoin price data and therefore it provides more accurate estimates of short-term future values. Some interesting findings and points have emerged. The first model discussed, ARIMA (1, 2, 2), produced unexpected results in the prediction on Bitcoin price. Nevertheless, according to test statistics that have been used for model validation, such as ACF, PACF, and ADF, the model showed to provide a good fit to logarithmically transformed BITCOIN Close price data. Further, ARIMA (1, 2, 2) has the lowest value of AIC which further implied that the model should not be modified and adequate for modeling the data. However, predictions that were obtained were far from true values. Moreover, the model predicted that the price of Bitcoin should experience growth in the predicted period exactly opposite from actual price movements for the stated period. Its Mean squared error of predictions is seven times MSE of Bayesian linear regression and with Mean Absolute Percentage Error of 18.40436%. A second model, the Bayesian method that Close 39 Bitcoin price time series for prediction and by far produced the most inaccurate results. Value of R square statistics was unreasonably high implying the strong explanatory power of the model. On contrary, the MSE of the model was largest and its value of MAPE is 25.44960%, which were the poorest results in both criteria. Forecasts implied linear growing trend for Bitcoin Close price. This model proved to be most deficient compared to the other two models discussed. On the other side, the Bayesian linear regression produced results that closely reflected the behavior of actual Bitcoin price movements. This simple regression has the value of R square close to one as well, but its standard deviation was much lower than of the model previously mentioned. The model proved to have the best predictive accuracy and explanation of data on both criteria, MSE and MAPE. It has the very low value of MAPE, 8.46927%, with absolute percentage errors that were around 1% in the first few periods.

After completion of the models, the assumption that Bayesian linear regression should have better predictive accuracy when Bitcoin price data is modeled is confirmed for the period, May 30, 2015- April 30, 2018. With doing better on both criteria than ARIMA (1, 2, 2) further implies that the focus of the research of Bitcoin should be on applying Bayesian statistics to model its historical data. However, the model could be improved and developed additionally by adding external macro-financial factors when modeling the Bitcoin price. This could improve model and predictions because of the rapid development and integration of Bitcoin into the global financial system. Another point on future research is that increasing frequency of the data, using hourly or lower intervals, can yield better forecasts and explanation of price movements of Bitcoin. The significance of knowledge about Bitcoin's characteristics, drivers and behavior are recognized in Academia as well as in the financial world. The better understanding would bring the efficient use of Bitcoin together with its application in business risk management sphere.

AFFILIATION

The research contained in the article was supported by International University of Sarajevo, Bosnia and Herzegovina.

REFERENCES

[1] Carlin, B. P., & Louis, T. A. (2009). *Bayesian Method for Data Analysis* (3rd. vyd.). 6000 Broken Sound Parkway NW, Suite 300: CHAPMAN & HALL/CRC.

[2] Charlin, B. P., & Louis, T. A. (2009). *Bayesian Methods for Data Analysis* . 6000 Broken Sound Parkway NW, Suite 300: CHAPMAN & HALL/CRC.

[3] Ciaian, P., Rajcaniova, M., & Kancs, d. (2015). The Economics of BitCoin Price Formation. *Applied Economics*, *48*.

[4] Coinbase. (05. June 2018). Coinbase Developers. (Coinbase, Editor) Coinbase.

[5] Amjad, M. J., & Shah, D. (2016). Trading Bitcoin and Online Time Series Prediction. *NIPS 2016 Time Series Workshop*, (stránky 1-15).

[6] Asteriou, D., & Hall, S. G. (2007). *Applied Econometrics: A modern approach.* New York: PALGRAVE MACMILLAN.

[7] Bakar, N. &. (Nov 2017). Autoregressive Integrated Moving Average (ARIMA) Model for Forecasting Cryptocurrency Exchange Rate in High Volatility Environment: A New Insight of Bitcoin Transaction. *International Journal of Advanced Engineering and Science, 4*(11), 130-137.

[8] Bouoiyour, J. S. (2014). *Is BitCoin business income or speculative bubble? Unconditional vs. conditional frequency domain analysis.* Získáno 03. Jan 2018, z Munich Personal RePEc Archive: https://mpra.ub.uni-muenchen.de/59595/1/MPRA_paper_59595.pdf

[9] Box, G. E., & Jenkins, G. M. (1976). *Time series anlysis : forecasting and control.* San Francisco: Holden- Day.

[10] Garcia, D., Tessone, C. J., & Perony, P. M. (2014). The digital traces of bubbles: Feedback cycles between socio-economic signals in the Bitcoin economy. *Journal of the Royal Society, Interface / the Royal Society, 11*.

[11] Grinberg, R. (2011). Bitcoin: An Innovative Alternative Digital Currency. *Hastings Science & Technology Law Journal, 4*, 160-208.

[12] Koduvely, D. H. (2015). *Learning Bayesian Models with R.* Birmingham B3 2PB, UK: Packt Oublishung Ltd.

[13] Levy, R. (5. November 2007). Lecture 9: Bayesian hypothesis testing. Získáno 15. May 2018, z http://idiom.ucsd.edu/~rlevy/lign251/fall2007/lecture_9.pdf

[14] Montgomery, D. J. (2008). *Introduction to Time Series Analysis and Forecasting.* New Jersey: Wiley Blackwell.

[15] Ruppert, D., & Matterson, D. S. (2015). *Statistics and Data Analysis for Financial Engineering* (2nd. vyd.). New York: Springer Text in Statistics.

[16] Shah, D., Zgang, K. (2015). *Bayesian regression and Bitcoin.* Získáno 03. Feb 2018, z IEEE Xplore: 10.1109/ALLERTON.2014.7028484

[17] Sinharay, S. (2010). An overview of Statistics in Education. *International Encyclopedia of Education*.

[18] Snipes, M. &. (2016). Model selection and Akaike Information Criteria: An example from wine ratings and prices. *Wine Economics and Policy, 3*(1), 3-9.

[19] Wallach, D., & Goffinet, B. (January 1989). Mean squared error of prediction as a criterion for evaluating and comparing system models. *Ecological Modeling*, *44*(3-4), 299-306.

AUTHORS

Assoc. Prof. Mojmír Sabolovič, Ph.D., Department of International Business and Finance, International University of Sarajevo, Hrasnicka cesta 15, 71210 Sarajevo, BiH, e-mail: msabolovic@gmail.com.

Emina Hukic, BA., Faculty of Business and Administration, International University of Sarajevo, Hrasnicka cesta 15, 71210 Sarajevo, BiH, e-mail: emina.hukic@hotmail.com.