HAL open science

Non-fundamental, non-parametric Bitcoin forecasting

Robert Adcock, Nikola Gradojevic

▶ To cite this version:

Robert Adcock, Nikola Gradojevic. Non-fundamental, non-parametric Bitcoin forecasting. Physica A: Statistical Mechanics and its Applications, 2019, 531, pp.121727 -. 10.1016/j.physa.2019.121727 . hal-03484525

HAL Id: hal-03484525

https://hal.science/hal-03484525v1

Submitted on 20 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Version of Record: https://www.sciencedirect.com/science/article/pii/S0378437119309859
Manuscript_109aeb7086b218bfbc3e715ce9bf8a67

Non-fundamental, Non-parametric Bitcoin Forecasting

Robert Adcock[a], Nikola Gradojevic[*,a,b,c]

[a] *University of Guelph, Canada*

[b] *Lille Catholic University, IESEG School of Management (LEM-CNRS 9221), France*

[c] *University of Novi Sad, Faculty of Technical Sciences, Serbia*

Abstract

Bitcoin is the largest cryptocurrency in the world, but its lack of quantitative qualities makes fundamental analysis of its intrinsic value difficult. As an alternative valuation and forecasting method we propose a non-parametric model based on technical analysis. Using simple technical indicators, we produce point and density forecasts of Bitcoin returns with a feedforward neural network. We run several models over the full period of April 2011 – March 2018, and four subsamples, and we find that backpropagation neural networks dominate various competing models in terms of their forecast accuracy. We conclude that the dynamics of Bitcoin returns is characterized by predictive local non-linear trends that reflect the speculative nature of cryptocurrency trading.

Keywords: Bitcoin; Neural Networks; Technical Analysis; Density Forecasting.

JEL classification: C45; F31; G1.

[*] Corresponding author. Tel. +1 5198244120 (ext. 52357); E-mail address: ngradoje@uoguelph.ca; Department of Economics and Finance, Lang School of Business and Economics, University of Guelph, 50 Stone Road East, Guelph, ON N1G 2W1, Canada. The authors are grateful to the Editor and the two anonymous Referees for their constructive and insightful comments.

© 2019 published by Elsevier. This manuscript is made available under the CC BY NC user license
https://creativecommons.org/licenses/by-nc/4.0/

1. Introduction

Within the past ten years, the cryptocurrency market has emerged and its leading currency, Bitcoin, has captured global attention. There have been attempts to forecast Bitcoin prices using artificial neural networks (ANNs), but few papers have combined it with technical analysis, and none have pursued the combination presented in this paper. There is no physical good that gives Bitcoin value, so it is impossible to perform fundamental analysis on this currency. Instead, technical analysis could serve as a potential alternative method.

The contribution of this paper is threefold: 1) we build a non-linear ANN forecasting model based on technical analysis in the spirit of Gençay (1998) and compare it to an array of parametric and non-parametric models; 2) we demonstrate its predictive accuracy for the BTC/USD exchange rate in terms of both point and density forecasting, and; 3) we assess the impact of market volatility on predicting Bitcoin returns. Since all the predictors being applied in this paper are public information, in essence, we test whether Bitcoin prices are efficient in the weak and semi-strong forms (Fama, 1970).

Our findings reveal that the ANN model with technical indicators as its inputs, in general, dominates the random walk model regarding the out-of-sample mean-squared prediction error (MSPE) and directional forecasts. The overall accuracy of the directional test was as high as 65%, while the MSPE improvements relative to the random walk model were 5-6%. It is important to note that the two most recent subsamples (October 1, 2014 – March 2, 2018), marked by a strong upward trend, produced substantial improvements in the MSPE statistic that ranged from 11% to 44%. The addition of the volatility index to the set of ANN inputs did not increase the out-of-sample predictive accuracy of the model. In all, the results indicate that Bitcoin prices followed local non-linear speculative trends that were most pronounced during the past 3-4 years.

Our comparative forecasting analysis shows that the proposed comprehensive ANN model dominates the competing parametric and non-parametric approaches (ARIMAX, GARCH-M, linear

regression, quantile regression ANN, and kernel regression). The only comparable predictive accuracy is achieved by the recurrent ANN model. This suggests the existence of certain temporal dynamics and short-term memory characteristics in Bitcoin time series, which is consistent with Lahmiri and Bekiros (2019). With regards to density forecasting, the comprehensive ANN model appears to be well-calibrated at the daily horizon and it estimates uncertainty around point forecasts appropriately. The competing models such as quantile regression ANN and ARIMAX exhibit a degree of bias in their predictions.

In all, this paper provides valuable insights into the dynamics of the BTC/USD exchange rate. First, it appears that the observed non-linearities in the BTC/USD return series require attention when modelling the conditional mean of the BTC/USD returns. Next, the BTC/USD returns exhibit sporadic periods of local trends (or bubbles) that might be predictable by using technical analysis. Finally, the accuracy of the proposed ANN model with technical indicators is sensitive to the period of consideration, with the superior predictive ability achieved during recent years (2014 – 2018) that were characterized by lower volatility. A possible explanation for this phenomenon is that, in a higher volatility market regime, moving average indicators ([1,50] and [1,200]) are more prone to mistiming shorter trend turning points[1].

The structure of the paper is as follows. In Section 2, we review the relevant literature. Next, in Section 3, we present the data, followed by the description of methodology in Section 4. The empirical results are reported and discussed in Section 5. Section 6 concludes and offers some directions for future research.

2. Literature Review

Balcilar et al. (2017) find that Bitcoin's volume is a good predictor of returns only when normality of returns is observed. Adding social signals such as Twitter activity, polarization measures, and search engine volume increases predictive power (Garcia and Schweitzer, 2015). Similar literature suggests that

[1] Gradojevic and Gençay (2013) show that moving average indicators are imperfect filters with a nonzero *phase shift*. Essentially, they may omit the temporal memory necessary to identify local trends and turning points.

market sentiment measures offer profitable trading strategies and that by using expert media it is possible to predict short-term Bitcoin price movements (Karalevicius et al., 2018). Using the above inputs and adding block size, exchange rates, mining rate and revenue, and several market indexes, Jang and Lee (2018) applied a Bayesian ANN and created a model capable of explaining and forecasting Bitcoin fluctuations. Furthermore, Bariviera et al. (2017) revealed time-varying long memory properties in Bitcoin returns. Finally, an ANN-GARCH model with several technical indexes and past Bitcoin returns as inputs dominated the benchmark model at forecasting volatility (Kristjanpoller and Minutolo, 2018).

If the random walk model produces the best prediction of future spot prices then a market is weak-form efficient (Fama, 1970). However, moving averages and trading-range breaks based on historical data were found to increase predictive power over the benchmark models of random walk as well as AR(1), EGARCH, and GARCH-M, thus rebutting the efficient market hypothesis (Brock et al., 1992). Lo and MacKinlay (1988) also strongly reject the random walk model in their research covering the 1962-1985 period using a simple specification test.

Kuan and White (1994) establish the framework for neural network applications within financial economics, laying down the basics that Gençay (1998) followed, and thus we followed. Further research concerning the optimal number of hidden layers finds that a one or two-layer network is sufficient; however, as the number of inputs increases, the size of hidden layers may need to increase (Kuan and Liu, 1995; Hornik et al., 1989; Gallant and White, 1988).

The combination of ANNs and technical analysis has been rigorously studied. It has been shown that ANNs have lower out-of-sample MSPE's compared to the random walk when forecasting stock prices and exchange rate series (Gençay and Stengos, 1998; Gençay, 1998; Gençay, 1999; Gradojevic and Yang, 2006; Gradojevic, 2007). Using past buy and sell signals from the Dow Jones Industrial Average Index (DJIA), Gençay and Stengos (1998) find that the feedforward ANN model improves on the benchmark over a 25 year period. Gençay and Stengos' (1995) results also suggest that performance generally increases with more lags, and that it varies between subsamples. Additional study of the DJIA

over a longer period of time suggests that ANN provides at least a 10% forecast improvement upon the random walk during volatile years of the Great Depression, with less pronounced results during calmer years when there is no strong positive or negative direction (Gençay, 1998). A similar study performed on foreign exchange rates supports these results (Gençay, 1999). The findings of Gençay (1999) are supported by Gradojevic and Yang (2006) who find that ANN improves out-of-sample MSPE and sign prediction for the Canadian-US Dollar (CAD/USD) exchange rate, and that ANN is more accurate for short term forecasts over the random walk. Gradojevic (2007) also introduces a fuzzy logic controller into the ANN model and finds, even with transaction costs accounted for, that the model is more profitable than a simple buy and hold strategy when trading CAD/USD.

Recently, Nakano et al. (2018) assessed the economic value of forecasting high-frequency (15-minute) Bitcoin returns with ANNs and technical indicators over the 2016-2018 period. We complement this work by focusing on the statistical value and the directional value of ANN forecasts that are based on technical indicators. Also, we utilize a longer data set of daily BTC/USD observations (2010-2018) and analyze the impact of market volatility (i.e., market sentiment) on the model's forecast performance. Most importantly, given that we concentrate exclusively on trend indicators, our approach uncovers certain stylized facts about Bitcoin price fluctuations that might be of value to government policy-makers and regulators. Specifically, if non-linear speculative trends exist at the market microstructure level of the Bitcoin fluctuations, policy prescriptions to remedy the situation may be necessary. However, the complexity of the Nakano et al. (2018)'s model precludes such lessons because it combines trend (i.e., exponential moving average), trend confirmation (i.e., stochastics), volume (i.e., on-balance-volume) and momentum (i.e., moving average convergence-divergence) technical predictors in a single model. This impedes a sensible economic interpretation as it is difficult to understand what is driving their results. In order to identify the nature of Bitcoin "market regimes", it would be necessary to quantify the relative importance of technical trading ANN inputs over sequential time periods.

The current paper also complements the contributions of Lahmiri and Bekiros (2019) and Stavroyiannis et al. (2019) that primarily analyze high-frequency time-series properties of Bitcoin. The former paper relies on an (autoregressive) deep learning neural network model to forecast returns on Bitcoin, Digital Cash, and Ripple, while the latter paper reveals consistent multifractal properties of Bitcoin prices. In contrast to our work, these recent additions to the literature do not consider the role of technical trading indicators and market volatility in forecasting cryptocurrencies, nor attempt to uncover the point and density predictors of BTC/USD returns.

3. Data

The data set was collected from a Bloomberg Terminal and it spans the period from July 19, 2010 to March 5, 2018. The primary time series is the daily BTC/USD exchange rate. Since Bitcoin's inception in January 2009 its value has skyrocketed, visible in Figure 1. The aggressive appreciation over a short time is reflected in the daily volatility of the BTC/USD (Figure 2), and Bitcoin has become a very popular, yet equally risky, investment option.

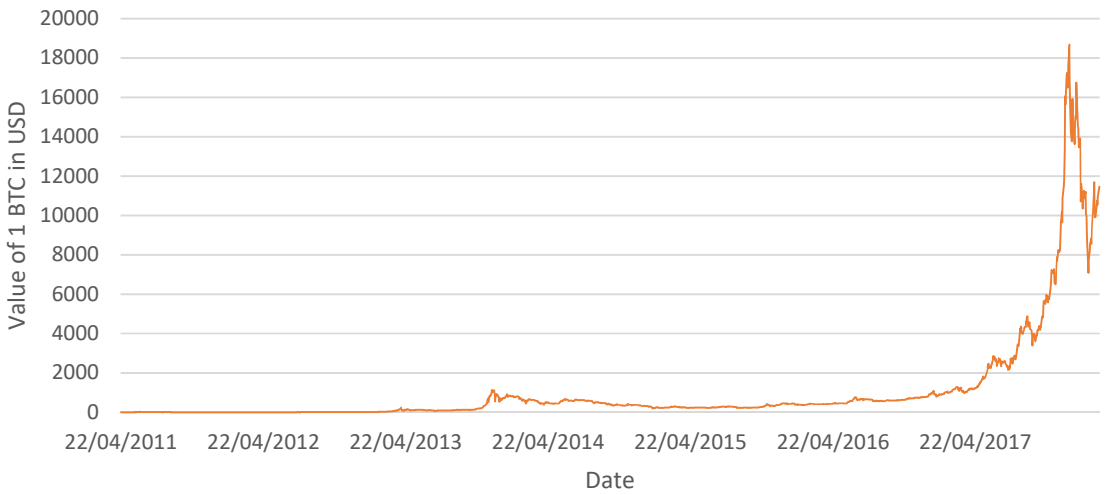


Figure 1: Daily price of BTC/USD

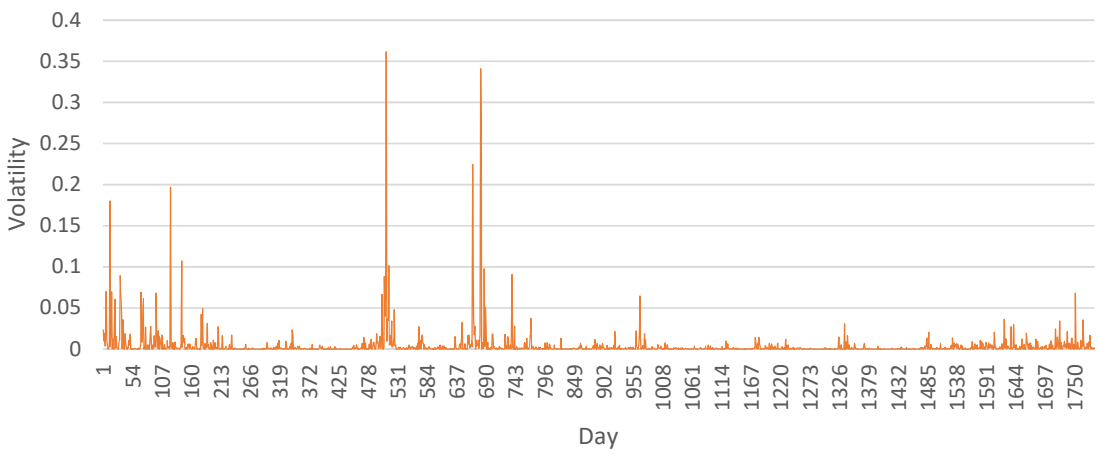


Figure 2: Daily volatility of the BTC/USD exchange rate. Notes: Day 1 refers to April 22, 2011 and the sample ends on March 2, 2018 (day 1784). Subsample 1 runs from day 1 to day 445, subsample 2 from day 446 to day 891, subsample 3 from day 892 to day 1337, and subsample 4 from day 1338 to day 1784.

Preparing the data for use in the ANN model required creating 50- and 200-day moving averages of the BTC/USD daily returns, and then using these indicators to create buy/sell signals by subtracting the price from the moving average at time *t*. These simple buy/sell signals were included as inputs along with log daily returns, as well as their lags of 1, 2, 3, and 4 days. After the adjustments, we were left with 1,784 useable observations beginning on April 22, 2011 and running until March 2, 2018. We also produced four subsamples of the time series to test market efficiency and the robustness of our results over time. The first subsample runs from April 22/2011 – January 10/2013 (445 observations), the second from January 11/2013 – Sep 30/2014 (445 observations), the third from October 1/2014 – June 15/2016 (445 observations), and the fourth from June 16/2016 – Mar 2/2018 (446 observations). We also included the CBOE Volatility Index (VIX) in a model to test whether the presence of a market volatility measure affects the neural network's performance. We provide the summary statistics of data in the Appendix.

4. Methodology

Our goal was to follow a similar methodology to the one outlined in Gençay (1998). The daily spot exchange series for BTC/USD will be denoted by $p_t, t = 1, 2, 3, \dots, T$, and the return series as $r_t =$

$\log(p_t/p_{t-1})$. To calculate the moving averages, we denote the length of the average as *n*, and let m_t^n be the moving average at time *t*, where

$$m_t^n = \left(\frac{1}{n}\right)\sum_{i=0}^{n-1} p_{t-i} \tag{1}$$

From the moving averages we create buy/sell signals, s_t, where

$$s_t^{n^1,n^2} = m_t^{n^1} - m_t^{n^2} \tag{2}$$

and we use [1,50] and [1,200] day moving averages for $[n^1, n^2]$. When s_t is negative (positive) it represents a sell (buy) signal.

The random walk is simply

$$r_t = \alpha + \epsilon_t \quad \epsilon_t \sim ID(0, \sigma^2) \tag{3}$$

and is our benchmark model.

The comprehensive feedforward ANN is constructed as

$$r_t = \alpha_0 + \sum_{j=1}^{d} \beta_j G\left(\alpha_j + \sum_{i=1}^{p} \gamma_{ij} r_{t-i} + \sum_{i=1}^{p} \gamma_{ij} s_t^{n^1,n^2}\right) + \epsilon_t \quad \epsilon_t \sim ID(0, \sigma^2) \tag{4}$$

where *G* is the activation function (sigmoid or tan-sigmoid) with the Levenberg-Marquardt optimization method, chosen for its efficiency (Wilamowski and Yu, 2010), and we have *d* hidden layers. Overfitting is prevented by early stopping, i.e., stopping the training process when the validation set error starts to increase. For the whole sample and each subsample, the data were divided into three parts: estimation data (70%), validation data (15%) and testing (out-of-sample) data (15%). We also modified the comprehensive model to target individual effects of returns in (5), 50-day signal, and 200-day signal in (6):

$$r_t = \alpha_0 + \sum_{j=1}^{d} \beta_j G\left(\alpha_j + \sum_{i=1}^{p} \gamma_{ij} r_{t-i}\right) + \epsilon_t \quad \epsilon_t \sim ID(0, \sigma^2) \tag{5}$$

$$r_t = \alpha_0 + \sum_{j=1}^{d} \beta_j G\left(\alpha_j + \sum_{i=1}^{p} \gamma_{ij} s_t^{n^1,n^2}\right) + \epsilon_t \quad \epsilon_t \sim ID(0, \sigma^2) \tag{6}$$

We first construct the log returns, moving averages, buy/sell signals, and then lag returns and signals four times. VIX is a simple log return when it is added to the inputs of the network.

5. Results

5.1. Feedforward ANN

Table 1 compares six different models applied to the full data range. The comprehensive model considers all 12 inputs (four lags of returns, four lags of [1,50] trading signals, and four lags of [1,200] trading signals) with 12 hidden layer neurons and is tested again with 6 hidden layer neurons. The three remaining models isolate only returns, 50-day signals, or 200-day signals. The subsample periods are shown in Tables 2-5 and they do not consider the 6 hidden layer size model, nor the VIX model.

We can see in Table 1 that the comprehensive model with 12 neurons produced the best predictive performance, outperforming the random walk by about 5.46%. Table 1 shows that the isolated returns model predicted the correct direction over 65% of the time, while the comprehensive model (with the lowest MSPE ratio) predicted direction with nearly 63% accuracy. These success rates are higher than those recorded for foreign exchange rate series (Gençay, 1999).

Model	*Comprehensive*	*Comprehensive (6 neurons)*	*VIX*	*Returns Only*	*50-Day Signal Only*	*200-Day Signal Only*
MSPE of ANN	0.00337	0.00339	0.00342	0.00348	0.00351	0.00339
Random Walk MSPE	0.00357	0.00357	0.00357	0.00357	0.00357	0.00357
Ratio	**0.94537**	0.95041	0.95927	0.97613	0.98473	0.95058
Sign	0.62921	0.58427	0.61049	**0.65169**	0.59551	0.6367
Hidden Layer Size	12	6	12	12	12	12

Table 1: Comparison of ANN predictability following standard set-up with entire data range, 1,784 observations, allocated 70% training, 15% validation, 15% testing (last 267 days). Best model ratio and best sign prediction are in bold.

It is important to note that the VIX input was unable to improve performance over the comprehensive model. The 200-day signal is about 4.42% better than the random walk, compared to 1.53% for the 50-day signal, and 2.39% for returns only. The six hidden layer size model did not drastically alter the results. Next, we will observe the performance of the competing models across four subsamples.

Model	*Comprehensive*	*Returns Only*	*50-Day Signal Only*	*200-Day Signal Only*
MSPE of ANN	0.00047	0.00047	0.00049	0.00049
Random Walk MSPE	0.00052	0.00052	0.00052	0.00052
Ratio	0.91947	**0.9069**	0.95572	0.94784
Sign	0.56897	**0.60345**	0.58621	0.60345
Hidden Layer Size	12	4	4	4

Table 2: Comparison of ANNs within subsample 1 (April 22, 2011 – January 10, 2013), 445 observations, allocated 70% training, 15% validation, 15% testing (last 58 days). Best model ratio and best sign prediction are in bold.

Table 2 is the first subsample from April 22, 2011 to January 10, 2013. The best-performing ANN was returns-only, outperforming the random walk by 9.31%. The comprehensive model in subsample 1 outperforms the random walk by about 8.05%, while the 50- and 200-day signals outperform by 4.43% and 5.22%, respectively. Referring to Figure 2 we can observe that subsample 1 had high daily volatility. This subsample finds itself between subsample 2's extreme volatility and the average return volatility over the entire dataset. The summary statistics in Table 7 tell us that returns had a daily standard deviation of 7.52%, higher than the full dataset return standard deviation of 6.5% shown in Table 6, and lower than the 8.65% of the second subsample. Subsample 1 also has the lowest volatility of the 50 and 200-day moving average signals. The isolated returns model also provides the best sign prediction with over 60% accuracy. This result is consistent in Table 3 with subsample 2.

The second subsample, as mentioned above, has highly volatile returns, but lower moving average signal volatility than the full dataset as we observe in Table 8. The results from subsample 2, shown in Table 3, present the first failure of the comprehensive model to outperform the random walk.

The 50-day signal outperforms the random walk by approximately 10.84%, suggesting that the short-term trend is most successful during periods of higher return volatility. The 200-day signal outperforms the random walk by 6.93%, and the past returns by only 1.21%.

Model	*Comprehensive*	*Returns Only*	*50-Day Signal Only*	*200-Day Signal Only*
MSPE of ANN	0.00181	0.00113	0.00102	0.00107
Random Walk MSPE	0.00115	0.00115	0.00115	0.00115
Ratio	1.58172	0.98795	**0.89162**	0.93067
Sign	0.55172	**0.60345**	0.56897	0.53448
Hidden Layer Size	12	4	4	4

Table 3: Comparison of ANNs within subsample 2 (January 11, 2013 – September 30, 2014), 445 observations, allocated 70% training, 15% validation, 15% testing (last 58 days). Best model ratio and best sign prediction are in bold.

From subsample 2 to subsample 3 we see a drastic change in the daily return volatility and ANN's predictive performance. Subsample 3 has the lowest return standard deviation of 3.64% and lower signal volatility than subsample 2, but higher than that of subsample 1, as shown in Table 9. Figure 2 suggests that this is the steadiest period of Bitcoin data we have. This period also saw a marked increase in Bitcoin trading. The ANN results in Table 4 are striking as the 200-day signal outperforms the random walk by 43.55%. The 50-day signal also saw an unprecedented improvement of 21.02%. The comprehensive and returns-only ANNs also outperform the random walk by 6.71% and 3.67%, respectively.

Subsample 3 strays from the sign prediction results of the first two subsamples and the full data set. The comprehensive model reaches nearly 59% successful sign prediction. Again, however, the best returns forecast was the isolated 200-day signal which offers only 53% sign accuracy.

Model	*Comprehensive*	*Returns Only*	*50-Day Signal Only*	*200-Day Signal Only*
MSPE of ANN	0.00098	0.00102	0.00083	0.0006
Random Walk MSPE	0.00106	0.00106	0.00106	0.00106
Ratio	0.93291	0.96334	0.7898	**0.5645**

Sign	**0.58621**	0.51724	0.55172	0.53448
Hidden Layer Size	12	4	4	4

Table 4: Comparison of ANNs within subsample 3 (October 1, 2014 – June 15, 2016), 445 observations, allocated 70% training, 15% validation, 15% testing (last 58 days). Best model ratio is in bold Best model ratio and best sign prediction are in bold.

The fourth and final subsample has the highest signal volatility of any sample, and relatively low return volatility as shown in Table 10. Table 5 presents the results of the ANN models for subsample 4. The comprehensive model improves upon the random walk by 10.87%. The isolated returns model improves by about 0.1%. The isolated signals fail to dominate the random walk for the first time. Like in subsample 3, the comprehensive model provides the best sign prediction for subsample 4 with about 59%.

Model	*Comprehensive*	*Returns Only*	*50-Day Signal Only*	*200-Day Signal Only*
MSPE of ANN	0.0053	0.00594	0.00599	0.00606
Random Walk MSPE	0.00594	0.00594	0.00594	0.00594
Ratio	**0.89127**	0.99898	1.00837	1.0192
Sign	**0.59322**	0.55932	0.42373	0.52542
Hidden Layer Size	12	4	4	4

Table 5: Comparison of ANNs within subsample 4 (June 16, 2016 – March 2, 2018), 446 observations, allocated 70% training, 15% validation, 15% testing (last 59 days). Best model ratio and best sign prediction are in bold.

Overall, feedforward ANNs improve prediction performance over the random walk; however, this may not hold true over all periods. Regarding sign predictions, it appears the buy-sell signals overall do the poorest job at forecasting direction, while the isolated returns model and the comprehensive do the best.

5.2. Comparative forecast performance

This subsection performs a comparative predictive analysis that aims at establishing how the proposed feedforward ANN model fares against a variety of parametric and non-parametric alternatives. All the predictive models are tested on the most recent 15% of the data (267 daily observations) that is held out-of-sample, while the first 70% of the sample is used for estimation (training) and the subsequent 15% of the sample is used as validation data. In addition to the comprehensive ANN (Equation 4) and the random

walk models that were utilized in the previous subsection, the following predictive models will be considered: 1) autoregressive integrated moving average with exogenous variables – ARIMAX (2,1,1); 2) generalized autoregressive conditional heteroscedasticity in mean – GARCH-M (1,1); 3) linear regression; 4) quantile regression ANN; 5) recurrent ANN; 6) kernel regression. All the competing models are based on the 12 inputs that were used in the comprehensive ANN model (Equation 4). The statistical significance of the difference in the performance of alternative models is tested by using the Diebold-Mariano (DM) test (Diebold and Mariano, 1995). The null hypothesis is that there is no difference in the MSPE of two alternative models (the random walk and the other models).

	Statistic	
	Sign (%)	*MSPE (DM)*
Comprehensive ANN	63%	0.0033 (-2.33*)
RW	-	0.0036
ARIMAX (2,1,1)	60%	0.0193 (4.54)
GARCH-M (1,1)	62%	0.0223 (4.78)
Linear regression	59%	0.1080 (6.82)
Quantile regression ANN	62%	0.0283 (3.15)
Recurrent ANN	62%	0.0034 (-1.71**)
Kernel regression	60%	0.0164 (5.15)

Table 6: Comparison of predictive models for the entire data range, 1,784 observations, allocated 70% training, 15% validation, 15% testing (last 267 days). The Sign and MSPE statistics are used for the out-of-sample (testing) data. The models that are considered are 1) comprehensive ANN model (Equation 4), 2) the random walk (RW) model, 3) ARIMA (2,1,1) model, 4) GARCH-M (1,1) model, 5) linear regression, 6) quantile regression ANN model, 7) recurrent ANN model, and 8) kernel regression model. The Diebold-Mariano (DM) test statistics are reported in the parentheses below MSPEs, where applicable. The critical values are +/- 1.64 and +/- 2.58 for a confidence level of 90% and 99%, respectively. (*) and (**) indicates the DM statistic is significant at 1% and 10% significance level, respectively.

The MSPE's and PERC's of the predictive models are presented in Table 6. Both statistical measures of forecast accuracy are reported in levels. The results in Table 6 indicate that the out-of-sample forecast performance of the linear estimators (GARCH-M (1,1), ARIMAX (2,1,1) and linear regression) do not outperform the random walk model. The DM statistics of the MSPE's of all three linear models are greater than 2.58, which shows statistically significant forecast improvements (at the 1% significance

level) of the random walk model over the linear models. Of the three alternative non-linear models evaluated, only the recurrent ANN model offers statistically significant forecast accuracy improvements over the random walk model. The quantile regression ANN (Taylor, 2000), for any quantile considered (Table 6 reports the MSPE for the 50th percentile), and kernel regression models are markedly inferior in terms of their out-of-sample forecast accuracy to the competing ANN models as well as to the random walk model. We conjecture that the observed forecast improvements of a feedforward ANN lie in its ability to process 12 inputs, whereas higher dimensional data plague the usefulness of kernel regression models. Similarly, training quantile regression ANNs with as many as 12 predictors may lead to overfitting.

5.3. Density forecasting

In addition to the Bitcoin point forecasting that measures the expected (central) tendency of the model output, this subsection explores the precision (uncertainty) around the forecasted values generated by the prediction models. This uncertainty is typically gauged by using density forecasts. In what follows, we will evaluate density forecasts from probability integral transforms (PITs) which are defined as cumulative probabilities (measured by density forecasts) at the realized value of the output. More formally, the PIT of a time series (y_t) is defined as follows (Diebold et al., 1998):

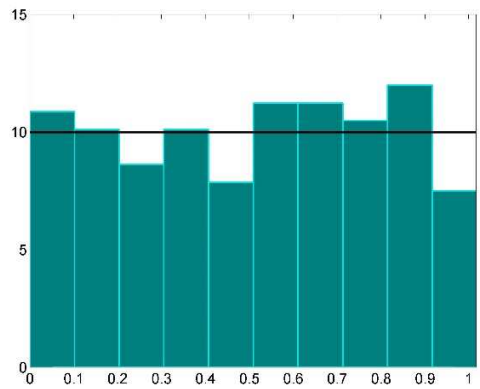
$$PIT_t = \int_{-\infty}^{y_t} p_t(u)du = P_t(y_t) \qquad (7)$$

where P_t is the cumulative distribution function of the density forecasts p_t.

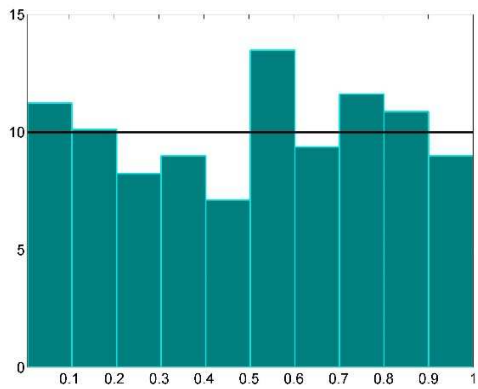
Diebold et al. (1998) show that the PIT is uniform (U[0,1]), independent and identically distributed if the density forecast is correctly specified. In line with Bekiros and Paccagnini (2014), we test the uniformity property of our density forecasts by plotting the empirical distribution of the PIT (i.e., its histogram). We construct the histograms with ten bins and observe the contribution of each bin. Uniform density forecasts would have histograms that resemble equally-sized rectangles, i.e., each bin would contribute 10% of the probability mass. The histograms of the PITs associated with one-day-ahead

density forecasts for the four competing predictive models are displayed in the panels A, B, C and D of Figure 3. The height of each bin reflects the fraction of observations that fell within different deciles of the density forecasts. Hence, the correct specification of density forecasts contains 10% in each bin, as represented by the horizontal solid line.

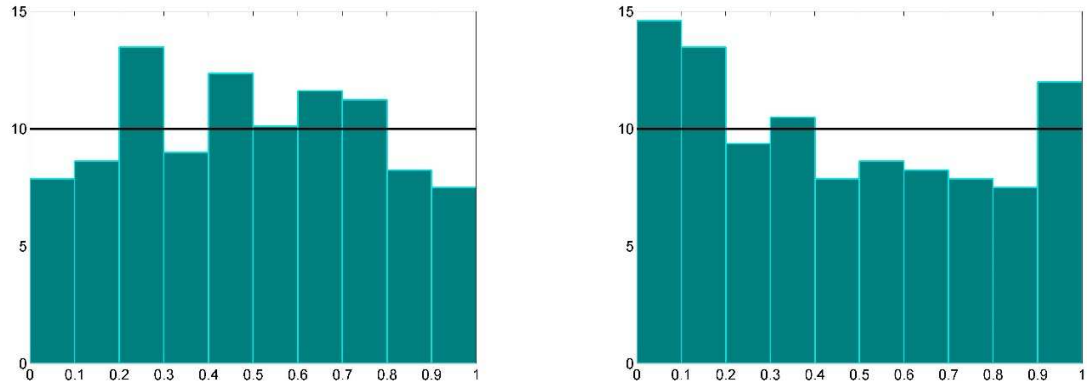
Panel A of Figure 3 suggests that the PITs for the comprehensive ANN model are closest to a uniform distribution, thus, indicating the most accurate density forecasts. Also, the PITs generated by the recurrent ANN model have an unconditional distribution that is approximately uniform, although the histogram exhibits a minor peak in the middle (Panel B, Figure 3). The PIT histogram in Panel C of Figure 3 is not correctly specified because it displays a hump in the middle. This can be interpreted as a slight tendency of the quantile regression ANN model to under-predict the Bitcoin returns which results in the underestimation of uncertainty by the density forecasts. Finally, the ARIMAX model (Panel D, Figure 3) forms a "U" shaped PIT histogram that arises from over-confident predictions. In all, we conclude that the comprehensive ANN model with technical indicators represents the optimal method for both point and density one-day-ahead forecasting of Bitcoin returns.



Panel A



Panel B



Panel C ***Panel D***

Figure 3: Histograms of the PITs for the competing predictive models. Notes: The focus is on the one-day-ahead density forecasts of the BTC/USD exchange rate returns on the out-of-sample part of the data (last 267 days). Panel A: Comprehensive ANN model; Panel B: Recurrent ANN model; Panel C: Quantile regression ANN model; Panel D: ARIMAX model. The horizontal solid line (10% probability mass in each bin) represents a well-specified model. The histograms resemble a uniform distribution when the heights of all bins are at 10.

6. Conclusions

The results in this paper suggest that BTC/USD returns can be forecasted with feedforward ANNs using past returns and simple technical trading rules. ANN can improve point and density forecasting over the range of competing parametric and non-parametric models. There is, however, variability in the predictive performance of the ANNs over time, possibly due to the volatile nature of Bitcoin. It can be concluded that the dynamics of the daily BTC/USD returns exhibits episodes of local trends that potentially arise from the speculative nature of cryptocurrency trading.

Further research should expand on the inputs used, perhaps considering volume, number of daily blocks added to the chain, or various other measurements. As well, similar studies should be done with a larger dataset and subsamples, and should investigate the profitability of other non-linear, non-parametric models. In terms of extending the model, a large variety of high-frequency technical trading inputs could be fed into a deep learning ANN where the relative importance of ANN inputs would be tracked over

time. This would enable us to differentiate distinct “market regimes” in Bitcoin fluctuations and explain market forces behind Bitcoin volatility. Finally, other significant cryptocurrencies could also be studied.

Appendix

	Log Return	*50-Day Signal*	*200-Day Signal*
Mean	0.00513	143.28726	510.90161
Std. Deviation	0.06498	966.20554	1645.01526
Sample Variance	0.00422	933553.14983	2706075.19711
Kurtosis	17.00856	42.71676	28.38834
Skewness	-0.38037	4.20856	4.98645

Table 7: Summary statistics for overall dataset, April 22, 2011 – March 2, 2018

	Log Return	*50-Day Signal*	*200-Day Signal*
Mean	0.00552	0.67935	2.38929
Std. Deviation	0.07520	3.40990	4.83931
Sample Variance	0.00566	11.62741	23.41895
Kurtosis	7.94823	12.36359	3.86129
Skewness	0.05044	2.29904	1.38454

Table 8: Summary statistics for Subsample 1, April 22, 2011 – January 10, 2013

	Log Return	*50-Day Signal*	*200-Day Signal*
Mean	0.00737	25.32859	119.29650
Std. Deviation	0.08650	142.94177	206.38868
Sample Variance	0.00748	20432.35006	42596.28660
Kurtosis	15.53803	10.31652	3.15659
Skewness	-0.65036	2.70574	1.76560

Table 9: Summary statistics for Subsample 2, January 11, 2013 – September 30, 2014

	Log Return	*50-Day Signal*	*200-Day Signal*
Mean	0.00137	1.43376	-25.25517
Std. Deviation	0.03643	50.95643	123.16758
Sample Variance	0.00133	2596.55810	15170.25367
Kurtosis	8.77809	1.41577	-1.18461
Skewness	-0.55987	0.15418	-0.01474

Table 10: Summary statistics for Subsample 3, October 1, 2014 – June 15, 2016

	Log Return	*50-Day Signal*	*200-Day Signal*
Mean	0.00616	551.03600	1912.60464
Std. Deviation	0.04969	1868.91181	2830.51856
Sample Variance	0.00247	3492831.34772	8011835.29887
Kurtosis	3.67048	8.44078	5.26176
Skewness	-0.35876	1.67345	2.32334

Table 11: Summary statistics for Subsample 4, June 16, 2016 – March 2, 2018

References

Balcilar, M., Bouri, E., Gupta, R., & Rouband, D., 2017. Can volume predict Bitcoin returns and volatility? A quantiles-based approach. Economic Modelling 64, 74–81.

Bariviera, A.F., Basgall, M.J., Hasperué, W., & Naiouf, M., 2017. Some stylized facts of the Bitcoin market. Physica A 484, 82–90.

Bekiros, S. & Paccagnini, A., 2014. Estimating point and density forecasts for the US economy with a factor-augmented vector autoregressive DSGE model. Studies in Nonlinear Dynamics and Econometrics 19(2), 107–136.

Brock, W.A., Lakonishok, J., & LeBaron, B., 1992. Simple technical trading rules and the stochastic properties of stock returns. Journal of Finance 47, 1731–1764.

Diebold, F.X., & Mariano, R.S., 1995. Comparing predictive accuracy. Journal of Business & Economic Statistics 13, 253–263.

Diebold, F.X., Gunther, T.A., & Tay, A.S., 1998. Evaluating density forecasts with applications to financial risk management. International Economic Review 39(4), 863–883.

Fama, E.F., 1970. Efficient capital markets: A review of theory and empirical work. The Journal of Finance 25, 383–417.

Gallant, A.R., & White, H., 1988. There exists a neural network that does not make avoidable mistakes. Proceedings of the Second Annual IEEE Conference on Neural Networks, San Diego, CA. IEEE Press, New York, pp I.657 – I.664.

Garcia, D., & Schweitzer, F., 2015. Social signals and algorithmic trading of Bitcoin. Royal Society Open Science 2, http://dx.doi.org/10.1098/rsos.150288.

Gençay, R., & Stengos, T., 1998. Moving average rules, volume and the predictability of security returns with feedforward networks. Journal of Forecasting 17, 401–414.

Gençay, R., 1998. The predictability of security returns with simple technical trading rules. Journal of Empirical Finance 5, 347–359.

Gençay, R., 1999. Linear, non-linear and essential foreign exchange rate prediction with simple technical trading rules. Journal of International Economics 47, 91–107.

Gradojevic N., & Yang J., 2006. Non-linear, non-parametric, non-fundamental exchange rate forecasting. Journal of Forecasting 25(4), 227–245.

Gradojevic N., 2007. Non-linear, hybrid exchange rate modelling and trading profitability in the foreign exchange market. Journal of Economic Dynamics & Control 31(2), 557–574.

Gradojevic N., & Gençay, R., 2013. Fuzzy logic, trading uncertainty and technical trading. Journal of Banking & Finance 37(2), 578–586.

Hornik, K., Stinchcombe, M., & White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Networks 2, 359–366.

Jang, H., & Lee, J., 2018. An empirical study on modeling and prediction of Bitcoin prices with Bayesian neural networks based on blockchain information. IEEE Access 6, 5427–5437.

Karalevicius, V., Degrande, N., & De Weerdt, J., 2018. Using sentiment analysis to predict interday Bitcoin price movements. The Journal of Risk Economics 19, 56–75.

Kristjanpoller, W., & Minutolo, M.C., 2018. A hybrid volatility forecasting framework integrating GARCH, artificial neural network, technical analysis and principal components analysis. Expert Systems With Applications 109, 1–11.

Kuan, C.-M., & Liu, T., 1995. Forecasting exchange rates using feedforward and recurrent neural networks. Journal of Applied Econometrics 10, 347–364.

Kuan, C.-M., & White, H., 1994. Artificial neural networks: An econometric perspective. Econometric Reviews 13, 1–91.

Lahmiri, S. & Bekiros, S., 2019. Cryptocurrency forecasting with deep learning chaotic neural networks. Chaos, Solitons & Fractals 118, 35–40.

Lo, A.W., & MacKinlay, A.C., 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. Review of Financial Studies 1, 41–66.

Nakano, M., Takahashi, A. & Takahashi, S., 2018. Bitcoin technical trading with artificial neural network. Physica A 510, 587–609.

Stavroyiannis, S., Babalos, V., Bekiros, S., Lahmiri, S. & Salah Uddin, G., 2019. The high frequency multifractal properties of Bitcoin. Physica A 520, 62–71.

Taylor, J.W., 2000. A quantile regression neural network approach to estimating the conditional density of multiperiod returns. Journal of Forecasting 19(4), 299–311.

Wilamowski, B.M., & Yu, H., 2010. Improved Computation for Levenberg-Marquardt Training. IEEE Transactions on Neural Networks 21, 930–937.