

Improving Bitcoin price prediction power by time-scale decomposition and GMDH-type neural network: A comparison of different periods and features

Maryam Seifaddini^{†*}, Amir Habibdoust[‡]

†Department of Computer Science, University of Guilan, Rasht, Iran ‡Department of Economics and Accounting, University of Guilan, Rasht, Iran Email(s): m.seifaddini@guilan.ac.ir, amir.habibdoost@gmail.com

Abstract. This paper aims to improve the predictability power of a machine learning method by proposing a two-stage prediction method. We use Group Modeling Data Handling (GMDH)-type neural network method to eliminate the user role in feature selection. To consider recent shocks in Bitcoin market, we consider three periods, before COVID-19, after COVID-19, and after Elon Musk's tweeter activity. Using time-scale analysis, we decomposed the data into different scales. We further investigate the forecasting accuracy across different frequencies. The findings show that in shorter period the first, second and third lag of daily prices and trade volume produce valuable information to predict Bitcoin price while the seven days lag can improve the prediction power over longer period. The results indicate a better performance of the wavelet base GMDH-neural network in comparison with the standard method. This reveals the importance of trade frequencies' impact on the forecasting power of models.

Keywords: Forecasting, machine learning, wavelet, Bitcoin price, GMDH. *AMS Subject Classification 2010*: 34A34, 65L05.

1 Introduction

Cryptocurrencies have attracted much attention among traders and, financial experts since these markets are characterized by high volatility, relatively smaller capitalization, decentralized, and high market data availability [9]. Bitcoin is the first decentralized cryptocurrency that was released in 2008. Bitcoin gradually received more attention, and the price of Bitcoin raised from 457 US dollars in 2014 to around 40000 US dollars in April 2022, with a 773B market cap.

Received: 8 July 2022/ Revised: 27 August 2022/ Accepted: 4 September 2022

DOI: 10.22124/JMM.2022.22638.2003

^{*}Corresponding author.

Bitcoin does not necessarily work such other asset prices such as gold or stock markets, and so on. Klein et al. (2018) found that Bitcoin is not similar in terms of volatility, correlation, and portfolio diversification and "that Bitcoin is no safe haven and offers no hedging capabilities for developed markets" [20]. Recently, studies about Bitcoin have accelerated, specifically in forecasting Bitcoin price and movements. While based on the efficient market hypothesis (EMH) it is impossible to earn abnormal returns because prices reflect all information, there is evidence of the possibility of price prediction to earn returns [11]. On the other hand, the Adaptive Market Hypothesis (AMH) [22], a well-known finance theory, asserts it is possible for one makes a profit in the market.

Due to different trading horizons, financial markets are not homogenous from the decision horizon's point of view. While time series and traditional method neglect this paramount point, time-scale analysis consider the problem by decomposing data into different scales. High volatility and trading at different frequencies make Bitcoin a suitable candidate for a time-scale study. By the time-scale approach, information that is hidden at different frequencies would be uncovered; consequently, it may expand and deepen our understanding of the time horizon and scale level (frequency) role in predicting future prices. In this paper, we aim to predict daily bitcoin prices by a combing wavelet decomposition method and Group Modeling Data Handling (GMDH)-type neural network method. The paper contributes to the related literature by employing an inductive and self-organized approach to the bitcoin market, which uses data at different scales as input. In addition, we consider three periods, before COVID-19, after COVID-19, and after Elon Musk's tweeter activity. The periods include January 1, 2018, to March 10, 2020, January 1, 2018, to January 27, 2021, and January 1, 2018, to January 19, 2022. The first period covers data before the recent pandemic. The second period covers the data during the pandemic; however, the period excludes data after Elon Musk's Twitter activity. Finally, we contribute to the literature by considering the importance of different frequencies in the prediction of Bitcoin prices.

Comparing standard and time-scale based predictions shows that the time-scale based prediction produces significantly more accurate forecasts. We also show that the first, second, and third lag of daily prices and trade volume can predict Bitcoin price in the shorter period, while the seven-day lag can improve the prediction power over a longer period. The results indicate a better performance of the wavelet base GMDH-neural network in comparison with the standard method. This reveals the importance of trade frequencies' impact on the forecasting power of models.

The remainder of the paper is organized as follows. Section 2 presents a concise review of the relevant literature and theoretical background. Section 3 explains the data and method. The empirical results and discussion are presented in Section 4, and Section 5 makes a conclusion.

2 Related literature and theoretical background

Price prediction has been attractive for traders and researchers. The prediction techniques have extended from linear regression to time series methods and recently, machine learning methods have been developed to forecast prices in the financial market. Using machine learning to predict cryptocurrencies markets also is growing. To build a model based on a theoretical foundation, there are two challenging issues. First, are bitcoin prices predictable? Second, which inputs (features) should be selected? The first question has a long history in economics and finance literature. EMH states that it is impossible because the prices reflect all related information. However, EMH is rejected by researchers theoretically and empirically [3, 12]. Alternatively, the Adaptive Market Hypothesis (AMH) states prices in the asset

market are predictable. Findings about Bitcoin market are also mixed. Some studies showed that Bitcoin market does not present an efficient market behavior [29]; however, some studies found weak form of efficiency [2], and others found the AMH hold for Bitcoin market [19]. Hence, our prediction's rational and theoretical foundation lies in the AMH and empirical findings of the predictability of bitcoin prices. The second challenge in supervised prediction is choosing input variables or features based on theoretical background. Input variables (features) can be selected based on two general categories: Fundamental variables (e.g. macroeconomic variable, gold price, and etc.) [5, 8, 32]. For example, [5] found that among macroeconomics variables dollar index has biggest influence on bitcoin price in long run.

The second general category is technical variables (e.g. lagged prices, volume, moving average prices and etc.) [10, 21, 25]. For example, [21] used optimal Functional Link Neural networks to predict daily, weekly, and monthly Bitcoin closing price movements. He also has done a comparative study among FLNs models using and showed the superiority of optimal FLNs models to predict Bitcoin price movements. [10] using multitude of popular technical indicators showed that technical analysis combined with non-linear forecasting models becomes statistically significantly dominant relative to the random walk model on a daily horizon. Their comparative analysis revealed that the random forest model outperforms other models.

In addition, it is possible to use market sentiment [14,30,31] as well as other cryptocurrencies. For example, [31] using news and historical data, showed that their model is able to predict price movements by indicating whether to buy, sell or hold. [30] also employed machine learning tools and available social media data for forecasting the price movement of the Bitcoin, Ethereum, Ripple, and Litecoin cryptocurrency market movements. They found that it is possible to predict Twitter data by itself could be used to predict certain cryptocurrencies. The findings support outperforming of NN to the other models [30].

The empirical results about input selection are mixed; however, [17] comprehensive research showed that technical inputs outperform other features. They also found that predictability of bitcoin increase for longer prediction horizons. We also choose different technical variables to predict bitcoin prices. Some researchers have focused on using the deep learning method to predict the cryptocurrencies prices. For example, [27] employed LSTM and GRU-based hybrid to predict Litecoin and Monero price movements. They show that their method predicts the prices with high accuracy and can be applied in various cryptocurrencies price predictions. Some researchers also compare machine learning methods' performance with the time series method. [15] compared Gated Recurrent Network model (GRU) and LSTM method with ARIMA method in the prediction of bitcoin prices. Their results show that the deep learning models outperform the poorly performing ARIMA prediction.

Based on our knowledge, the number of research, which used wavelet decomposition to predict bitcoin prices are rare. Recently, [26] combined wavelet and long short-term memory networks methods to forecast bitcoin. The findings revealed that Wavelet-LSTM forecast one-step-ahead Bitcoin return in more accurate. Although there are researches that have investigated bitcoin properties by wavelet decomposition [4, 18], the predictability of bitcoin using a decomposition method and a self-organized method is not fully understood yet.

3 Data and method

3.1 Data

We employed the daily price of Bitcoin from YahooFinance. We select three periods based on date of the COVI-19 pandemic announcement and Elon Musk's tweets about bitcoin. The first period ranges from January 1, 2018 to March 10, 2020, which excludes pandemic and Elon Musk's tweet effects. The second period, from January 1, 2018 to January 27, 2021, includes the time of COVID-19 pandemic but excludes Elon Musk's tweet. Finally, the third period ranges from January 1, 2018 to January 19, 2022, which is longer than other periods and includes to mentioned two mentioned periods. To present a better look at the Bitcoin prices, Figure 1 shows the Bitcoin price over the three periods. We choose a different

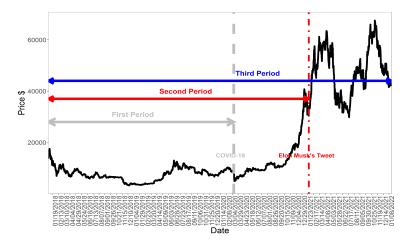


Figure 1: Bitcoin price over three periods.

combinations of lagged price and trade volume variables. For example, model number 1 has five inputs, the first, second, and third lag of daily prices as well as first and second lag of daily trade volume (Table 1). We divided the data into training datasets (80 percent of data points) and testing datasets (20 percent of data points) for in-sample forecasting. All analyses used RStudio (Version 1.4.1717-R 4.0.4) and Matlab 2016 for preparing data and estimating.

3.2 Time-Scale analysis

When frequencies are changing, as what happens in financial markets, standard time series methods can not reveal enough and proper information for analysis. We need a technique to transform data to both frequency and time domain simultaneously in this situation. While Fourier Transform (FT) transforms data to frequency domains, the time scale or wavelet analysis transforms data to both time and frequency domains. Wavelet analysis is based on multiresolutional analysis, which was developed by [23] and [28]. Following Mallat's representation, a time series or a signal X(t) can be written as follows:

$$X(t) \approx s_i(t) + d_i(t) + d_{i-1}(t) + \dots + d_1(t),$$
 (1)

	Model							
	Number		Ir	iput V	ariabl	es		
1,	1	p1	p2	р3	v1	v2	-	-
	2	p 1	p2	p3	p 7	v1	v2	v7
January 2018- March 1 2020	3	p1	p2	p3	v1	v2	v3	-
5 2	4	p5	p7	v5	v7	-	-	-
, 1, - 10,	5	p 1	p2	p3	v1	v2	-	-
January 1, 2018- March 10, 2020	6	p 1	p2	p3	p7	v1	v2	v7
Jan 2 Ma	7	p 1	p2	p3	v1	v2	v3	-
	8	p5	p7	v5	v7	-	-	-
January 1, 2018 to January 19, 2022	9	p 1	p2	p3	v1	v2	-	-
	10	p 1	p2	p3	p7	v1	v2	v7
	11	p 1	p2	p3	v1	v2	v3	-
J Pa	12	p5	p7	v5	v7	-	-	-

Table 1: Input variables (features).

- **p1**: Bitcoin Daily price with 1-day lag, **p2**: Bitcoin Daily price with 2 days lag,
- **p3**: Bitcoin Daily price with 3 days lag, **p5**: Bitcoin Daily price with 5 days lag,
- p7: Bitcoin Daily price with 7 days lag, v1: Bitcoin Daily Trade Volume with 1-day lag,
- v2: Bitcoin Daily Trade Volume with 2 days lag, v3: Bitcoin Daily Trade Volume with 3 day lag,
- v5: Bitcoin Daily Trade Volume with 5 day lag, v7: Bitcoin Daily Trade Volume with 7 day lag.

where, $s_j(t)$ is a trend component (the wavelet "smooth"), $d_j(t)$ are detail components (wavelet "crystals"), and j represents the number of scales. The discrete wavelet transform (DWT) process is summarized in Figure 2^1 . The variable X(t) is passed by two types of filter step by step until we reach an approximated variable that includes just trend information. The two mentioned filters are called low-pass and high-pass filters. Indeed, in each step or frequency, some part of the information. At first step, we can write X(t) at scale 1 (J=1) based on decomposed components, $d_1(t)$ (- the high frequency part) and $S_1(t)$ (- the low frequency part):

$$X(t) \approx s_1(t) + d_1(t). \tag{2}$$

Now, we again write such decomposition for $S_1(t)$

$$s_1(t) \approx s_2(t) + d_2(t). \tag{3}$$

By doing such recursive procedure until reaching our desirable scale J, we have a set of variables which include j detailed or high-frequency components and 1 smoothed component. With such data in hand, we can do prediction or any other analysis at different frequencies and domains at the same time. It should be noted that the signal or time series will be subsampled by 2 at each step, which can be a drawback for DWT because we need dyadic data. To tackle this issue and other disadvantages of DWT method, we used the Maximum Overlap Discrete Wavelet Transform (MODWT) as our method of time-frequency

¹ Borrowed form [7]

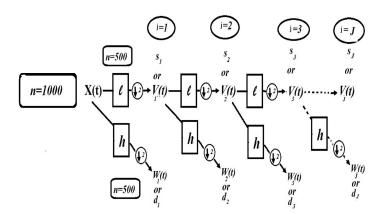


Figure 2: Discrete Wavelet Transform (DWT) process.

decomposition². Furthermore, as mentioned before, the data are filtered out by low pass filters and high pass filters. These filters can be written with different functions. There are different wavelet functions like Morlet, Mexican hat, Daubechies, and so forth. This study used a Daubechies 4-tap (D4) wavelet and decomposed the data into 6 scales.

3.3 modeling and forecasting using GMDH-type neural networks

GMDH algorithm was introduced by [16] to solve complex system problems; specifically, the GMDH neural network has been developed for prediction and forecasting tasks in different disciplines such as statistics, economics, finance, medicine, etc. It is an inductive and self-organized machine learning approach that is suitable for complex model prediction [1]. The central idea is straightforward, but the process is complex and suitable to model complex and nonlinear phenomena and short-term predictions. Supposed one aims to model a complex system, S, with multiple inputs, $X = (x_1, x_2, x_3, ..., x_n)$, and one output, y, as depicted in Figure 3. The inputs and output relationship can generally be written as follows:

$$y = f(x_1, x_2, x_3, \dots, x_n). \tag{4}$$

In which f is an elementary function, y is output, and x_i is input. The goal is to find a proper estimation of

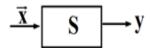


Figure 3: Complex system.

elementary function, f, so that it can predict the output \hat{y} . To do so, the least square method is employed to estimate the coefficient of the elementary function. One of the favorite functions is the complicated

² DWT also have another disadvantages. For detailed information see [6]

discrete form of the Volterra functional series, which is known as the Kolmogorov-Gabor polynomial:

$$y = a_0 + \sum_{i=1}^n a_i x_i + \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n a_{ijk} x_i x_j x_k + \cdots.$$
 (5)

The central idea of GMDH is based on a Divide and Conquer strategy so that the Kolmogorov-Gabor polynomial decomposing to several sub-functions such as (6)

$$Z_{ij} = G(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2.$$
 (6)

It is clear that Z_{ij} has two input variables, and the number of all Z_{ij} is equal $\binom{2}{n}$. In this way, such a partial quadratic description is recursively used to reconstruct the complete Kolmogorv-Gabor polynomial. Eq. (4) can be presented in matrix form

$$Y = xA. (7)$$

In which A is the vector of coefficients in Eq. (6). Multiplying both sides of Eq. (7) by the transpose of a the matrix X gives

$$X^T Y = (X^T X)A. (8)$$

Using the least-squares technique for multivariate regression, the solution of the normal equations will be as follows

$$A = (X^{T}X)^{-1}X^{T}Y. (9)$$

In practice, the mentioned recursive procedure is done as follows. The date is divided into two section, train data, and test date. This procedure is started from the first hidden layer (in which there are $\binom{2}{n}$ neurons, Z_{ij}). The coefficients of Eq. (5) (or Eq. (7) in matrix form) are estimated by the least square

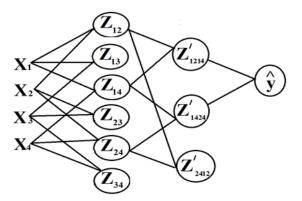


Figure 4: GMDH-type neural network.

method. Now all estimated equations are evaluated by external criteria or thresholds for prediction error such as RMSE³. All estimated equations which are passed the threshold will be saved to construct the next hidden layer, and others are dropped out. The next hidden layer is also constructed by Eq. 6 in which inputs are Z_{ij} , which was produced in the previous step (layer). The estimation and selection⁴ are made again. This process, called the self-organization of models, will end up until just one neuron remains. Figure 4 depicts a schematic GMDH-type neural network.

³ Root Mean Square Error

⁴ Which remember the natural selection in evaluation theory.

3.4 Combining wavelet decomposition and GMDH-type neural network

We predict the prices by doing a two-step process. First, the time series of Bitcoin and all other input variables are decomposed into six scales using Daubechies 4-tap (D4) wavelet. Then, we predict prices at different scales by GMDH-type neural network method and sum up forecasted prices to achieve predicted time series in time-domain environment (Figure 5).

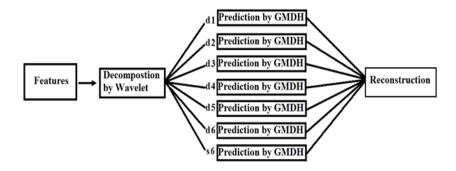


Figure 5: GMDH-type neural network.

4 Results and discussion

It is critical to assess and compare the performance of the forecasting. In this study, the Mean Absolute Error (MAE) (Eq. (11)), and the Root Mean Square Error (RMSE) (Eq. (10)) are employed to providing a fair performance evaluation among the utilized methods. The smaller value for RMSE and MAE indicate the higher accuracy for the model.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (P_t - FP_t)^2},$$
(10)

$$MAE = \frac{1}{n} \sum_{t=1}^{n} |P_t - FP_t|, \tag{11}$$

where P_t , FP_t , and n are the actual values, the predicted value, and the number of forecasted data (or test data), respectively. In addition, to make a better sense of prediction performance, we normalized our metric as (Eqs. (12) and (13)):

$$RMSE = \frac{RMSE}{max_{P_t} - min_{P_t}} \tag{12}$$

$$NMAE = \frac{MAE}{max_{P_c} - min_{P_c}} \tag{13}$$

Since measures are sensitive to outliers, the outlier detection test is performed [13]. Table 2 shows the Grubbs' test results. The results for 24 (GMDH and Wavelet-GMDH) models are presented in Table 3. Here, we explain some part of the results in detail. The RMSE and MAE of model number 1 for GMDH

Table 2: Grubbs'test results.

	Highest value is an outlier			Lowest value is an outlier		
Model Number		statistic		statistic		
Models	G	U	P-value	G	U	P-value
1-GMDH	1.901	0.857	1	2.001	0.979	1
1-WGMDH	1.803	0.983	1	1.913	0.981	1
2-GMDH	1.901	0.981	1	2.031	0.989	1
2-WGMDH	1.812	0.971	1	1.925	0.881	1
3-GMDH	1.901	0.857	1	2.001	0.979	1
3-WGMDH	1.803	0.983	1	1.915	0.980	1
4-GMDH	1.982	0.979	1	2.111	0.976	1
4-WGMDH	1.655	0.985	1	2.497	0.967	1
5-GMDH	1.804	0.983	1	2.002	0.979	1
5-WGMDH	1.882	0.981	1	1.813	0.982	1
6-GMDH	1.791	0.983	1	2.005	0.979	1
6-WGMDH	1.815	0.982	1	1.882	0.981	1
7-GMDH	1.804	0.983	1	2.002	0.979	1
7-WGMDH	1.814	0.982	1	1.882	0.981	1
8-GMDH	1.995	0.979	1	2.104	0.977	1
8-WGMDH	1.559	0.975	1	1.774	0.983	1
9-GMDH	1.859	0.961	1	1.967	0.975	1
9-WGMDH	1.959	0.891	1	2.501	0.903	1
10-GMDH	1.853	0.923	1	1.804	0.880	1
10-WGMDH	1.982	0.879	1	2.101	0.876	1
11-GMDH	1.7655	0.905	1	2.697	0.867	1
11-WGMDH	1.816	0.984	1	1.802	0.781	1
12-GMDH	1.789	0.748	1	2.145	0.964	1
12-WGMDH	1.879	0.857	1	2.732	0.946	1

prediction are 246.61 and 180.6, respectively. However, using Wavelet-GMDH method, the mentioned metrics significantly decrease to 85.05 and 62.45. The normalized metrics are 0.71% (NRMSE), 0.52% (NMAE) for Wavelet-GMDH prediction, 2.07% (NRMSE), and 1.51% (NMAE) for GMDH prediction. The model number 3, which includes first, second and third lag of price and trade volume as input variables, has RMSE value of 83.18 and MAE value of 60.26 for Wavelet-GMDH, 246.77, and 177.45 for GMDH prediction. The NRMSE and NMAE for Wavelet-GMDH is 0.7% and 0.5% and the metrics for GMDH prediction are 2.07% and 1.49%, respectively. In the first period, the model number 3 produces

		Wavelet-GMDH		GMDH		Wavelet-GMDH		GMDH	
	Model Number	RMSE	MAE	RMSE	MAE	NRMSE	NMAE	NRMSE	NMAE
1, 0,	1	85.05	62.45	246.61	180.60	0.71%	0.52%	2.07%	1.51%
	2	160.72	82.75	245.35	178.75	1.35%	0.69%	2.06%	1.50%
January 2018- March 1 2020	3*	83.18	60.26	246.77	177.45	0.70%	0.50%	2.07%	1.49%
ř 2	4	287.15	169.48	593.88	447.21	2.41%	1.42%	4.98%	3.75%
1,	5	122.70	85.93	990.77	548.88	0.48%	0.33%	3.85%	2.13%
January 2018- March 1 2020	6	219.64	113.10	941.75	553.64	0.85%	0.44%	3.66%	2.15%
Januar 2018 March 2020	7*	122.57	85.81	990.77	548.88	0.48%	0.33%	3.85%	2.13%
	8	414.97	281.34	2947.95	1626.04	1.61%	1.09%	11.44%	6.31%
January 1, 2018 to January 19, 2022	9	3448.39	2704.77	17075.62	15556.02	5.36%	4.20%	26.54%	24.18%
	10*	1105.84	837.21	16918.32	15416.53	1.72%	1.30%	26.30%	23.96%
	11	3041.06	2393.47	16413.93	14951.90	4.73%	3.72%	25.52%	23.24%
Ja	12	2873.27	2258.15	53001.19	48319.87	4.47%	3.51%	82.39%	75.11%

Table 3: Prediction accuracy.

Models with lowest NRMSE and NMAE are indicated by *

lowest forecasting error. In addition, seven and five days lags of daily price and trade volume do not improve the accuracy of prediction. However, the third lag of daily trade volume significantly improves the forecasting performance. The results for the second period, which covers the data after COVID-19, but is limited to period Elon Musk' tweet, shows that the model number 7 has lowest RMSE (122.57), MAE (85.81), NRMSE (0.476%), and NMAE (0.333%) values for Wavelet-GMDH prediction and RMSE (990.77), MAE (548.88), NRMSE (3.845%), and NMAE (2.130%) values for GMDH prediction. However, the model number 5 produces errors closest metrics to model number 7.

In the second period, the general message of prediction is same as fist shorter period; the seven and five days lag does not improve the forecasting performance. In longer third period, the models number 10 has higher accuracy rather than other models. The metrics for Wavelet-GMDH prediction are 1105.84 (RMSE), 837.21 (MAE), 1.72% (NRMSE), 1.30% (NMAE).

The prediction accuracy for GMDH prediction are 16918.32 (RMSE), 15416.53 (MAE), 26.30% (NRMSE), and 23.96% (NMAE). It is clear that on our longer period the entering seven days lag of daily trade volume improve the prediction.

Considering normalized metrics, model number 7 produces a more accurate prediction rather all alternative models in all periods, the models number 3 and 6 have second lowest NRMSE and NMAE.

Accuracy improvement

Table 4 presents accuracy improvement which is obtained by applying Wavelet-GMDH model. It is clear that there are significant improvement in all models by employing Wavelet-GMDH method. It should be noted that the metrics are statistically different based on the multivariate version of the Diebold-Mariano test [24] (see Table 5).

	r	
Model Number	Percentage Improvement in RMSE	Percentage improvement in MAE
1	65.5%	65.4%
2	34.5%	53.7%
3	66.3%	66.0%
4	51.6%	62.1%
5	84.3%	87.6%
6	79.6%	76.7%
7	84.4%	87.6%
8	82.7%	85.9%
9	79.8%	82.6%
10	93.5%	94.6%
11	81.5%	84.0%
12	94.6%	95.3%

Table 4: Improvement in accuracy metrics.

Table 5: Test for equal predictive accuracy of forecasting models.

	lag length*	statistic	p-value **
ury 1, 8-	3	38.432	4.921e-07
January 2018- March 1 2020	5	50.352	2.059e-08
nuary 1, 2018- arch 10, 2020	3	99.10	3.961e-08
January 2018- March 1 2020	5	44.302	7.096e-06
nuary 1, 2018- arch 10, 2020	3	101.10	4.001e-09
January 2018- March 1 2020	5	51.516	5.078e-07

^{*}It is assumed that the autocorrelation of loss differentials is essentially zero beyond this lag length.

In the first period, the highest improvement occurs in model number 3 with, 66.3% improvement in RMSE and 66.0% in MAE. The model number 7 in the second period produce highest improvement among other models in the period by 84.4% and 87.6% for RMSE and MAE, respectively. Wavelet-GMDH method produces highest improvement in model number 12 (94.6% for RSME and 95.3% for MAE). The maximum improvement occurs in the third period, which is a longer, and includes data after COVID-19 and Elon Musk's tweets. Based on the results, the Wavelet-Based GMDH-type neural

^{**} Alternative hypothesis: Equal predictive accuracy does not hold.

	NRMSE						
Model Number	d_1	d_2	d_3	d_4	d_5	d_6	<i>s</i> ₆
1	2.2209%	2.3380%	1.4804%	0.8495%	0.5828%	0.2381%	0.0256%
2	2.1867%	2.3091%	1.4629%	2.9860%	2.2262%	0.4131%	0.0490%
3	2.1725%	2.0047%	0.9785%	0.5897%	0.5250%	0.2255%	0.0222%
4	3.1189%	3.3862%	0.9285%	6.2328%	2.6242%	3.9183%	1.2294%
5	2.2311%	1.9694%	0.6633%	0.1557%	0.1573%	0.1323%	0.0238%
6	2.2312%	1.9695%	0.6664%	0.1853%	0.0570%	1.6048%	0.0115%
7	2.2312%	1.9694%	0.6624%	0.1850%	0.1573%	0.1346%	0.0188%
8	3.0378%	3.3513%	4.7963%	3.5265%	1.9387%	1.2349%	0.2602%
9	8.2096%	7.9946%	2.3072%	0.8982%	0.1926%	0.1597%	5.9665%
10	8.1822%	7.9730%	2.3060%	0.7715%	0.1959%	3.4854%	0.5715%
11	8.2096%	7.9946%	2.3072%	0.8982%	0.1926%	0.1597%	5.2162%
12	8.7299%	8.9314%	12.0012%	13.1491%	6.6741%	1.8940%	2.4552%

Table 6: NRMSE at different scales

network produces more reliable forecasting model rather pure GMDH-type neural network in all our three periods. Indeed, employing Wavelet-GMDH-type neural network the forecasting accuracy increase significantly and considerably. The improvement rises as the length of period increases, which reveals the importance of hidden information in different frequencies. The possible explanation for improving the performance of the wavelet GMSH models underlies the decomposition of time series data. Indeed, wavelet deposition makes series stationary at different scales. This would increase the performance of prediction at different scales. Our findings show that first, second, and third lag of daily price along with first, second, and third lag of daily trade volume have more predictive power; however, this is not true for third period. The model with seven days lag has more predictive power in comparison with other alternative models in the period.

Accuracy across scales

We calculate forecasting error in all scales for all models (Tables 6 and 7). In model number 1 lowest NRMSE and NMAE occurs at less noisy data (trend component (s_6)) with value of 0.026% and 0.019%, respectively and the model produces highest error at scales d_1 (2.338% for NRMSE and 1.544% for NMAE) and d_2 (2.338% for NRMSE and 1.544% for NMAE). However, this arrangement is not true for all models. For example, in model number 4 the lowest NRMSE and NMAE (0.9285% and 0.6296%) occur in scale d_3 . The model number 12 has the largest NRMSE and NMAE among all models across scales except scales d_6 and s_6 . These findings emphasize the dissimilar prediction power of input variables on different scales.

Figure 6 depicts a general picture of the metrics for all scales and model. It would be interesting to see how our prediction accuracy change across different scales and models. The plot box method is employed to show the variation of NRMSE and NMAE across scales (1 to 6). The vertical axis shows the

	NMAE						
Model Number	d_1	d_2	d_3	d_4	d_5	d_6	<i>s</i> ₆
1	1.7368%	1.5440%	0.6816%	0.2631%	0.2280%	0.1670%	0.0187%
2	1.7191%	1.5283%	0.6746%	0.5769%	0.5918%	0.1901%	0.0333%
3	1.6893%	1.4168%	0.6491%	0.1994%	0.2008%	0.1580%	0.0156%
4	2.4108%	2.5641%	0.6296%	3.6926%	1.8624%	2.4702%	0.8621%
5	1.5300%	1.3436%	0.4077%	0.0891%	0.0529%	0.0611%	0.0145%
6	1.5301%	1.3444%	0.4104%	0.0985%	0.0315%	0.3358%	0.0063%
7	1.5301%	1.3436%	0.4077%	0.0984%	0.0529%	0.0621%	0.0118%
8	2.1428%	2.3748%	3.3039%	2.3589%	1.1511%	0.6355%	0.1738%
9	6.1336%	6.1123%	1.7897%	0.6820%	0.1203%	0.0934%	4.6079%
10	6.1078%	6.0999%	1.7923%	0.5866%	0.1229%	1.1475%	0.4695%
11	6.1336%	6.1123%	1.7897%	0.6820%	0.1203%	0.0934%	4.0410%
12	6.8952%	7.2333%	9.4280%	9.4291%	4.5848%	1.4557%	2.1163%

Table 7: NMAE at different scales.

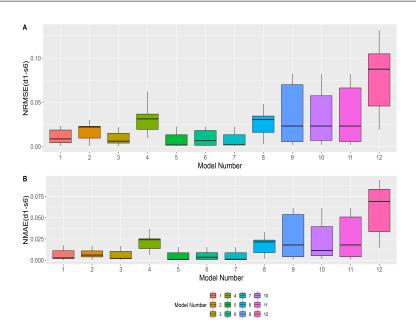


Figure 6: Forecasting error disparities at different scales.

value of NRMSE and NMAE, and the horizontal axis shows the model numbers. In longer period, which includes data after COVID-19 and Elon Musk's tweeter activity, the range of the NRMSE and NMAE values cross scales are wider than other models. However, there is no substantial variation in the first and second periods. In addition, the models with lower NRMSE and NMAE (3, 5, 6, 7, and 10) exhibit lower range in the NRMSE and NMAE values across scales. For example, the range of changes of NRMSE across scales in model 3 is 0.0215, and the value for NMAE is 0.01674. The variance is 0.00006 and

	NRMSE	NMAE		
Model Number	Variance	Range	Variance	Range
1	0.00007	0.02312	0.00004	0.01718
2	0.00010	0.02937	0.00003	0.01686
3	0.00006	0.02150	0.00004	0.01674
4	0.00027	0.05304	0.00010	0.03063
5	0.00008	0.02207	0.00004	0.01516
6	0.00008	0.02220	0.00003	0.01524
7	0.00008	0.02212	0.00004	0.01518
8	0.00020	0.04536	0.00011	0.03130
9	0.00112	0.08050	0.00065	0.06040
10	0.00100	0.07986	0.00059	0.05985
11	0.00108	0.08050	0.00062	0.06040
12	0.00161	0.11255	0.00091	0.07973

Table 8: Variance and range of forecasting error measures across scales.

0.00004 for NRMSE and NMAE, respectively. The range of NRMSe in model number 8 is 0.04536, and the variance is 0.00020. The range of NMAE is 0.03130, and its variance is 0.00011. Model number 12, which belongs to our longer period, has a wider range for NRMSE(0.11255) and NMAE(0.07973). The variances are also larger compared to other models(0.11255 for NRMSE and 0.00091 for NMAE) (See Table 8 for more details).

Figure 7 displays a comparison between the ranges of forecasting errors of the two methods. Wavelet-

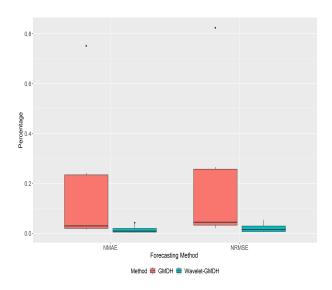


Figure 7: Forecasting error disparities.

GMDH neural network type decreases the range and variation in error measures of proposed models sig-

	NRMSE	NMAE		
	Variance	Range	Variance	Range
Wavelet-GMDH	0.000291	0.048848	0.000191	0.038715
GMDH	0.049063	0.803335	0.04232	0.736254

Table 9: Variance and range of forecasting error measures across models

nificantly. The variance of NRMSE and NMAE values across 12 models for Wavelet-GMDH method are 0.000291 and 0.000191, while the numbers are 0.049063 and 0.04232 for GMDH method, respectively. The range of metrics (0.803335 for NRMS and 0.736254 for NMAE) is also wider for GMDH method in comparison with Wavelet-GMDH method (0.048848 for NRMS and 0.038715 for NMAE) (See Table 9). It is clear that using wavelet decomposition decrease the variation of forecasting error across models.

5 Conclusion

This paper attempts to compare the forecasting power of the wavelet-Based GMDH-type neural network and standard GMDH-type neural network in the prediction of Bitcoin price over three periods. Using daily price and trade volume data, we run 12 models with different features using GMDH type neural network. We decomposed the data into six scales using Daubechies 4-tap (D4) wavelet and used a GMDH-type neural network over the scales. The findings provide practical implications for Bitcoin traders. First, in the short run, the first, second, and third lag of trade volume and price provide valuable information to predict bitcoin price; however, widening the analysis period shows that the seven-day lag of trade volume should be considered a predictor. Second, our findings show that the Wavelet-GMDH-type neural network significantly outperforms the standard GMDH method in forecasting Bitcoin prices. In addition, the results indicate that disparities of forecasting error in the Wavelet-GMDH method across models are lower in the standard GMDH method. The finding supports the importance of hidden information on different scales of Bitcoin prices which are neglected in standard time series.

Acknowledgements

The authors would like to thank the reviewers for their helpful comments and suggestions.

References

- [1] N. Amanifard, N. Nariman-Zadeh, M. Borji, A. Khalkhali, A. Habibdoust, *Modeling and Pareto optimization of heat transfer and flow coefficients in microchannels using GMDH type neural networks and genetic algorithms*, Energy Convers. Manag. **49** (2008) 311–325.
- [2] A.F. Bariviera, *The inefficiency of Bitcoin revisited: A dynamic approach*, Econ. Lett. **161** (2017) 1–4.

- [3] S. Basu, Investment performance of common stocks in relation to their priceearnings ratios: A test of the efficient market hypothesis, J. Finance. **32** (1977) 663-682.
- [4] E. Bouri, S.J.H. Shahzad, D. Roubaud, L. Kristoufek, B. Lucey, *Bitcoin, gold, and commodities as safe havens for stocks: New insight through wavelet analysis*, Q Rev Econ Finance. **77** (2020) 156–164.
- [5] W. Chen, H. Xu, L. Jia, Y. Gao, Machine learning model for Bitcoin exchange rate prediction using economic and technology determinants, IIF. 37 (2021) 28-43.
- [6] P.M. Crowley, A guide to wavelets for economists, J. Econ. Surv. 21 (2007) 207–267.
- [7] P.M. Crowley, A. Habibdoust, *Assessing the exchange rate exposure of US multinationals*, Research Discussion Papers 34/2013, Bank of Finland.
- [8] A.H. Dyhrberg, *Bitcoin, gold and the dollarA GARCH volatility analysis*, Finance Res. Lett. **16** (2016) 85–92.
- [9] M. Ferreira, S. Rodrigues, C.I. Reis, M. Maximiano, *Blockchain: A tale of two applications*, Appl. Sci. **8** (2018) 1506.
- [10] N. Gradojevic, D. Kukolj, R. Adcock, V. Djakovic, *Forecasting Bitcoin with technical analysis: A not-so-random forest?*, Int. J. Forecast. **39** (2023) 1–17.
- [11] J. Green, J.R.M. Hand, X.F. Zhang, *The supraview of return predictive signals*, Rev. Account. Stud. **18** (2013) 692–730.
- [12] S. J. Grossman, J. E. Stiglitz, On the impossibility of informationally efficient markets, AER. 70 (1980) 393–408.
- [13] F. E. Grubbs, Sample criteria for testing outlying observations, Ann. Math. Statist. (1950) 27–58.
- [14] R. Hakim das Neves, *Bitcoin pricing: impact of attractiveness variables*, Financ. Innov. **6** (2020) 1–18.
- [15] I.A. Hashish, F. Forni, G. Andreotti, T. Facchinetti, S. Darjani, A hybrid model for bitcoin prices prediction using hidden Markov models and optimized LSTM networks, IEEE, ETFA. (2019) 721–728.
- [16] A. G. Ivakhnenko, *Polynomial theory of complex systems*, IEEE, Trans. Syst. Man Cybern. Syst. **4** (1971) 364–378.
- [17] P. Jaquart, D. Dann, C. Weinhardt, *Short-term bitcoin market prediction via machine learning*, JFDS. 7 (2021) 45–66.
- [18] S.H. Kang, R. P. McIver, J. A. Hernandez, *Co-movements between Bitcoin and Gold: A wavelet coherence analysis*, Phys. A: Stat. Mech. Appl. **536** (2019).
- [19] S. Khuntia, J.K. Pattanayak, *Adaptive market hypothesis and evolving predictability of bitcoin*, Econ. Lett. **167** (2018) 26–28.

- [20] T. Klein, H.P. Thu, T. Walther, *Bitcoin is not the New GoldA comparison of volatility, correlation, and portfolio performance*, Int. Rev. Financ. Anal. **59** (2018) 105–116.
- [21] I.E. Livieris, N. Kiriakidou, S. Stavroyiannis, P. Pintelas, *An Advanced CNN-LSTM Model for Cryptocurrency Forecasting*, Electronics. **10** (2021) 287.
- [22] A.W. Lo, The adaptive markets hypothesis, J. Portf. Manag. 30 (2004) 15–29.
- [23] S.G. Mallat, *Multiresolution approximations and wavelet orthonormal bases of L(R)*, Trans. Am. Math. Soc. **315** (1989) 69–87.
- [24] R.S. Mariano, D. Preve, *Statistical tests for multiple forecast comparison*, J. Econom. **169** (2012), 123–130.
- [25] S.C. Nayak, *Bitcoin closing price movement prediction with optimal functional link neural networks*, Evol. Intell. **15** (2022) 1825–1839.
- [26] N. Parvini, M. Abdollahi, S. Seifollahi, D. Ahmadian, Forecasting Bitcoin returns with long short-term memory networks and wavelet decomposition: A comparison of several market determinants, Appl. Soft Comput. **121** (2022) 108707.
- [27] M.M. Patel, S. Tanwar, R. Gupta, N. Kumar, *A deep learning-based cryptocurrency price prediction scheme for financial institutions*, JISA. **55** (2020) 102583.
- [28] G. Strang, Wavelets and dilation equations: A brief introduction, SIAM Review. 31 (1989) 614–627
- [29] A. Urquhart, The inefficiency of Bitcoin, Econ. Lett. 148 (2016) 80–82.
- [30] F. Valencia, A. Gmez-Espinosa, B. Valds-Aguirre, *Price movement prediction of cryptocurrencies using sentiment analysis and machine learning*, Entropy. **21** (2019) 589.
- [31] A.D. Vo, Q.P. Nguyen, C.Y. Ock, Sentiment Analysis of News for Effective Cryptocurrency Price Prediction, IJKE. 5 (2019) 47–52.
- [32] Y. Zhu, D. Dickinson, J. Li, *Analysis on the influence factors of Bitcoins price based on VEC model*, FIN. **3** (2017) 1–13.