Journal of
Risk and Financial Management

Article

Forecasting the Price of the Cryptocurrency Using Linear and Nonlinear Error Correction Model

Jong-Min Kim [1,*], Chanho Cho [2] and Chulhee Jun [3]

[1] Statistics Discipline, Division of Science and Mathematics, University of Minnesota-Morris, Morris, MN 56267, USA
[2] School of Business and Natural Science, Black Hills State University, Spearfish, SD 57783, USA; chanho.cho@bhsu.edu
[3] Department of Finance, Bloomsburg University of Pennsylvania, Bloomsburg, PA 17815, USA; cjun@bloomu.edu
* Correspondence: jongmink@morris.umn.edu

Abstract: We employed linear and nonlinear error correction models (ECMs) to predict the log returns of Bitcoin (BTC). The linear ECM is the best model for predicting BTC compared to the neural network and autoregressive models in terms of RMSE, MAE, and MAPE. Using a linear ECM, we are able to understand how BTC is affected by other coins. In addition, we performed Granger-causality tests on fourteen cryptocurrencies.

Keywords: cryptocurrencies; Bitcoin; error correction model; Granger causality

check for updates

Citation: Kim, Jong-Min, Chanho Cho, and Chulhee Jun. 2022. Forecasting the Price of the Cryptocurrency Using Linear and Nonlinear Error Correction Model. *Journal of Risk and Financial Management* 15: 74. https://doi.org/10.3390/jrfm15020074

Academic Editors: Aviral Kumar Tiwari and Mark Harris

Received: 30 December 2021
Accepted: 7 February 2022
Published: 10 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The Coronavirus Disease 2019 (COVID-19) pandemic made the investment environment more uncertain. Since then, it has been very difficult for investors to forecast financial markets because of the uncertainty of this pandemic. Bitcoin, a decentralized currency in the cryptocurrency market, has emerged a popular investment asset and is often referred to as the currency of the future. Recently, various error correction models have been applied to the cryptocurrency market. Using an ECM, Liang (2021) found the result that the relationship between Bitcoin rate of return and the relevant indicators to measure monetary function was not significant, which rejects the original assumption that Bitcoin can assume monetary function, indicating that Bitcoin does not have the ability and potential to assume monetary function. Haffar and Fur (2021) analyzed the impact of shocks in the financial markets of emerging and developed countries on the price of Bitcoin using a structural vector error correction model. Keilbar and Zhang (2021) analyzed the role of cointegration relationships within a large system of cryptocurrencies using a vector error correction model (VECM) framework. Szetela et al. (2021) verified the existence of short-term and long-term relationships between the strength of a trend and the volume in bullish and bearish cryptocurrency markets through the application of a VECM to Bitcoin daily data. Giudici and Pagnottoni (2020) investigated return connectedness across eight of the major exchanges of Bitcoin, both from a static and a dynamic viewpoint by employing an extension of the order-invariant forecast error variance decomposition proposed by Diebold and Yilmaz (2012) to a generalized vector error correction framework. Using a time-varying VEC model, Chang and Shi (2020) examined the dynamic information shares of the top four cryptocurrencies: Bitcoin, Ethereum, Ripple, and Litecoin. Kapar and Olmo (2020) proposed an empirical model for analyzing the dynamics of Bitcoin prices by considering a VEC model over two overlapping periods: 2010–2017 and 2010–2019. These findings provided empirical evidence on the presence of a correction in the price of Bitcoin during the period 2018–2019 uncorrelated to market fundamentals. Ibrahim et al. (2020) forecasted the Bitcoin closing price using vector autoregression (VAR) and Bayesian vector autoregression (BVAR) prediction models. Experimental results showed that the VAR models achieved better performance compared to the traditional autoregression models and the BVAR models. Hakim das Neves (2020) studied the relationship between the price of virtual currency, the price of Bitcoin, and the number of Google searches that used the terms bitcoin, bitcoin crash, and crisis between December 2012 and February 2018 by using an error correction model. Goczek and Skliarov (2019) aimed to determine what drives the price of Bitcoin and analyzed a large set of data by using VEC models augmented by factors representing unobservable economic forces. Goczek and Skliarov (2019) also found that the main factor driving the price of Bitcoin is its popularity. Wang et al. (2016) performed a cointegration analysis and used a VEC model to demonstrate that there is a relationship between price of Bitcoin and some variables, including stock price index, the price of oil, and the daily trading volume of Bitcoin. The short-run analysis of Wang et al. (2016) revealed that price of oil and Bitcoin trading volume had little influence on the price of Bitcoin, whereas the stock price index had a relatively larger influence on the price of Bitcoin. Georgoula et al. (2015) used a VECM to investigate the existence of long-term relationships between cointegrated variables. Georgoula et al. (2015) revealed that the price of Bitcoin was positively associated with the number of Bitcoins and negatively associated with the S&P 500 stock market index.

Even though Bitcoin is the most dominant cryptocurrency, and it was found to influence other cryptocurrencies by Kwapień et al. (2021), our goal in this research was to build a model to predict Bitcoin (BTC) log-return prices based on other cryptocurrencies' prices, because BTC has strong correlations with other major cryptocurrencies, such as Ethereum (ETH) and Binance Coin (BNB). Currently, Miller and Kim (2021) have applied several deep learning time-series models to predict BTC log-return prices, but there is no standard guideline for selecting the correct deep learning tools, which requires knowledge of the topology, training method, and other parameters. Therefore, we still need a prediction model from which researchers can make statistical inferences on cryptocurrency price data. We propose linear and nonlinear ECM prediction models compared with the current available univariate time-series models, including the neural network time-series model.

This paper is organized as follows. Section 2 presents the summary and graphical data analysis for the top fourteen cryptocurrencies. Section 3 gives an overview of the econometrical models used in this study. The illustrated comparison study for the proposed methods is performed in terms of the measures of errors is in Section 4, with the conclusion presented Section 5.

2. Description of Data

The cryptocurrency data used in this study were obtained from crypto2 R package. The variables for each of the cryptocurrency datasets before manipulation were low, open, time, high, volume from, volume to, conversion type, conversion symbol, and close. The data period studied for each of the fourteen cryptocurrencies was from 1 January 2019 to 27 August 2021. Each variable is calculated as $\log\left(\frac{P_t}{P_{t-1}}\right)$. Table 1 shows the 14 cryptocurrencies that are used in this paper.

Table 2 shows the summary statistics for each of the cryptocurrency datasets. In terms of the median for log returns, Bitcoin (BTC), Ethereum (ETH), Cardano (ADA), Binance Coin (BNB), XRP, Tether (USDC), Bitcoin Cash (BCH), Litecoin (LTC), Chainlink (LINK), Ethereum Classic (ETC), and Stellar (XLM) have positive values. ADA has the highest median log returns, but DOGE has the highest mean log returns among 14 cryptocurrencies. The values of kurtosis in the log returns of all cryptocurrencies in Table 2 are greater than 3, meaning heavy tails compared to normal distribution. BTC, ETH, ADA, BNB, XRP, BCH, LTC, LINK, and ETC are left skewed while USDT, DOGE, USDC, LUNA, and XLM are right skewed.

Table 1. Variable definitions.

Variable	Name of Cryptocurrency
BTC	Bitcoin
ETH	Ethereum
ADA	Cardano
BNB	Binance Coin
USDT	Tether
XRP	XRP (Ripple)
DOGE	Dogecoin
USDC	USD Coin
LUNA	Terra
BCH	Bitcoin Cash
LTC	Litecoin
LINK	Chainlink
ETC	Ethereum Classic
XLM	Stellar

Table 2. Summary Statistics.

This Table Reports Summary Statistics of the Major Variables Used in Our Analysis.								
Variables	**N**	**Mean**	**Q1**	**Median**	**Q3**	**Max**	**Skewness**	**Kurtosis**
BTC	968	0.260	−1.388	0.174	1.911	17.182	−1.525	24.157
ETH	968	0.315	−1.856	0.279	2.861	23.070	−1.553	19.843
ADA	968	0.431	−2.531	0.333	3.074	27.944	−0.350	10.430
BNB	968	0.453	−1.943	0.197	3.135	52.922	−0.193	21.296
USDT	968	−0.002	−0.118	−0.006	0.112	5.339	0.300	57.251
XRP	968	0.119	−1.912	0.019	1.984	44.476	−0.099	19.238
DOGE	968	0.497	−1.980	−0.035	1.730	151.638	6.138	101.913
USDC	968	−0.002	−0.094	0.001	0.098	4.244	0.876	28.523
LUNA	968	0.159	−5.947	−0.254	5.168	236.929	0.245	18.264
BCH	968	0.135	−2.261	0.026	2.787	42.081	−0.557	18.252
LTC	968	0.172	−2.455	0.131	2.668	26.873	−0.990	13.818
LINK	968	0.456	−3.185	0.216	3.863	48.062	−0.404	13.746
ETC	968	0.255	−2.003	0.211	2.337	35.247	−0.075	14.694
XLM	968	0.113	−2.510	0.061	2.353	55.918	0.884	18.391

Table 3 shows Pearson and Kendall correlations between the price log returns of cryptocurrencies. We can notice that (BTC, ETH), (BTC, XRP), (BTC, BCH), (BTC, LTC), (ETH, ADA), (ETH, BNB), (ETH, XRP), (ETH, BCH), (ETH, LTC), (ETH, ETC), (ETH, XLM), (ADA, XRP), (ADA, BCH), (ADA, LTC), (ADA, XLM), (XRP, BCH), (XRP, LTC), (XRP, ETC), (XRP, XLM), (BCH, LTC), (BCH, ETC), (BCH, XLM), (LTC, ETC), and (LTC, XLM) have high Pearson and Kendall correlations, greater than or equal to 0.50.

Table 4 shows the augmented Dickey–Fuller (ADF) unit root tests for the log returns of 14 cryptocurrencies. The *p*-values in Table 4 are smaller than the significance level 0.05. This means that the log returns of the 14 cryptocurrencies are stationary time-series data. With the tsm and vars R packages, we performed an alternative unit root test when the data include a structural break in Table 4 and we also rejected the null of unit root with constant and trend at the 5% significance level. We also performed principal component analysis by using “prcomp” command in stats R package. Figure 1 shows the 2D and 3D graphical presentations by using the leading two and three principal components in principal component analysis. From Figure 1, LUNA, USDT, and USDC have different locations compared to other cryptocurrencies, including BTC and ETH.

Table 3. Correlation Tables. Data range is from 1/1/2019 to 8/27/2021. Each variable is calculated as $\log(P_t/P_{t-1})$ and expressed as a percentage.

Panel A: Pearson Correlation Matrix													
	BTC	**ETH**	**ADA**	**BNB**	**USDT**	**XRP**	**DOGE**	**USDC**	**LUNA**	**BCH**	**LTC**	**LINK**	**ETC**
ETH	0.82												
ADA	0.67	0.76											
BNB	0.66	0.70	0.62										
USDT	−0.07	−0.11	−0.08	−0.10									
XRP	0.57	0.63	0.60	0.54	−0.06								
DOGE	0.40	0.38	0.37	0.28	−0.04	0.32							
USDC	−0.10	−0.10	−0.08	−0.09	0.66	−0.05	−0.03						
LUNA	0.13	0.09	0.08	0.09	−0.05	0.05	0.04	−0.04					
BCH	0.77	0.81	0.70	0.63	−0.09	0.65	0.43	−0.10	0.11				
LTC	0.79	0.84	0.73	0.68	−0.07	0.66	0.42	−0.07	0.10	0.85			
LINK	0.57	0.68	0.60	0.56	−0.06	0.53	0.34	−0.08	0.05	0.60	0.63		
ETC	0.63	0.72	0.64	0.57	−0.10	0.58	0.45	−0.07	0.04	0.78	0.75	0.55	
XLM	0.62	0.69	0.74	0.57	−0.08	0.73	0.39	−0.09	0.06	0.68	0.68	0.58	0.64
Panel B: Kendall Correlation Matrix													
	BTC	**ETH**	**ADA**	**BNB**	**USDT**	**XRP**	**DOGE**	**USDC**	**LUNA**	**BCH**	**LTC**	**LINK**	**ETC**
ETH	0.61												
ADA	0.48	0.57											
BNB	0.46	0.51	0.46										
USDT	0.06	0.05	0.05	0.02									
XRP	0.50	0.57	0.54	0.44	0.05								
DOGE	0.47	0.46	0.42	0.38	0.07	0.45							
USDC	−0.02	−0.01	0.00	−0.03	0.42	0.00	0.04						
LUNA	0.22	0.17	0.12	0.17	−0.04	0.15	0.10	−0.03					
BCH	0.58	0.62	0.53	0.47	0.05	0.55	0.46	−0.01	0.17				
LTC	0.58	0.63	0.55	0.49	0.05	0.56	0.46	0.00	0.17	0.66			
LINK	0.38	0.48	0.44	0.38	0.04	0.41	0.35	0.00	0.10	0.44	0.44		
ETC	0.48	0.55	0.49	0.43	0.04	0.50	0.45	0.01	0.14	0.59	0.56	0.40	
XLM	0.45	0.53	0.57	0.43	0.04	0.58	0.41	−0.02	0.15	0.51	0.51	0.43	0.49

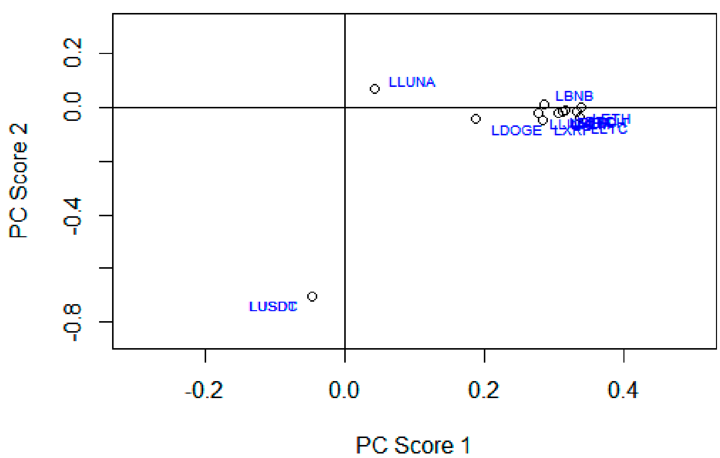

Figure 1. *Cont.*

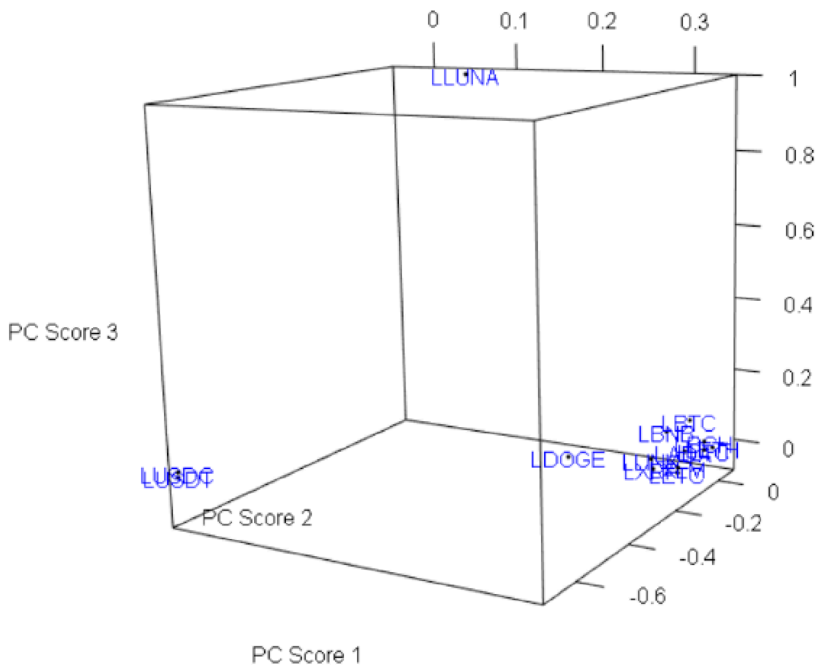

Figure 1. Graphical display by principal component analysis.

Table 4. Augmented ADF and alternative unit root tests. Data range is from 1/1/2019 to 8/27/2021. Each variable is calculated as $\log(\mathrm{P_t}/\mathrm{P_{t-1}})$ and expressed as a percentage.

Cryptocurrency	Augmented ADF	*p*-Value	Alternative Unit Root Testing under Structural Breaks
BTC	−9.5408	0.01	Reject null of unit root at 5%—with constant & trend
ETH	−9.2633	0.01	Reject null of unit root at 5%—with constant & trend
ADA	−8.8816	0.01	Reject null of unit root at 5%—with constant & trend
BNB	−8.0394	0.01	Reject null of unit root at 5%—with constant & trend
USDT	−14.345	0.01	Reject null of unit root at 5%—with constant & trend
XRP	−9.1692	0.01	Reject null of unit root at 5%—with constant & trend
DOGE	−9.4223	0.01	Reject null of unit root at 5%—with constant & trend
USDC	−12.381	0.01	Reject null of unit root at 5%—with constant & trend
LUNA	−12.579	0.01	Reject null of unit root at 5%—with constant & trend
BCH	−9.5168	0.01	Reject null of unit root at 5%—with constant & trend
LTC	−9.7936	0.01	Reject null of unit root at 5%—with constant & trend
LINK	−9.5374	0.01	Reject null of unit root at 5%—with constant & trend
ETC	−9.0545	0.01	Reject null of unit root at 5%—with constant & trend
XLM	−9.6799	0.01	Reject null of unit root at 5%—with constant & trend

3. Econometrical Methods

In this section, we briefly define the econometric methods that we used in this paper. We first want to look at the causality in mean by using the linear Granger causality (Granger 1969) in a vector autoregressive (VAR) system to explore informational linkages between pairs of markets. Given any pair of stationary data (X_t and Y_t), variable X_t Granger-causes Y_t linearly, provided that lags of X_t offer a significant information for explaining the current values of Y_t. The bivariate Granger causality is specified in a VAR system as follows:

$$X_t = \varphi_1 + \sum_{i=1}^{k} a_{1i} X_{t-i} + \sum_{i=1}^{k} b_{1j} Y_{t-j} + v_{1t} \tag{1}$$

and

$$Y_t = \varphi_2 + \sum_{i=1}^{k} a_{2i} X_{t-i} + \sum_{i=1}^{k} b_{2j} Y_{t-j} + v_{2t} \tag{2}$$

where φ_1 and φ_2 are the constant terms of the system of equations; a and b denote estimated coefficients; k is the optimal lag length based on the Akaike information criterion (AIC); and v_{1t} and v_{2t} represent residuals from the VAR model. The general format of an error correction model (ECM) is:

$$\Delta y_t = \beta_0 + \beta_1 \Delta x_{i,\,t} + \cdots + \beta_i \Delta x_{i,\,t} + \gamma(y_{t-1} - (\alpha_1 x_{1,\,t-1} + \cdots + \alpha_i x_{i,\,t-1}))$$

The ECM function of the R Package 'ecm' in Bansal (2021) modifies the equation to the following:

$$\Delta y_t = \beta_0 + \beta_1 \Delta x_{i,\,t} + \cdots + \beta_i \Delta x_{i,\,t} + \gamma y_{t-1} + \gamma_1 x_{1,\,t-1} + \cdots + \gamma_i x_{i,\,t-1}$$

where $\gamma_i = -\gamma\alpha_i$, so it can be modeled as a simpler ordinary least squares (OLS) function using R's lm function.

By default, R's base 'lm' is used to fit the model. However, researchers can opt to use 'earth', which uses Jerome Friedman's multivariate adaptive regression splines (MARS) to build a nonlinear regression model, which transforms each continuous variable into piecewise linear hinge functions. This allows for non-linear features in both the transient and equilibrium terms. ECM models are used for time-series data.

To forecast the log-returns of BTC, we used the Hyndman et al. (2021) "forecast" R package for employing univariate time-series models, such as autoregressive integrated moving average (ARIMA) model, exponential smoothing state space (ETS) model, autoregressive fractional integrated moving average (ARFIMA) model, BATS model (exponential smoothing state space model with Box-Cox transformation, ARMA errors, trend and seasonal components), TBATS, which is a modification of BATS that allows for multiple non-integer seasonality cycles, and neural network autoregressive (NNAR) model, which is a feed-forward neural network with a single hidden layer and lagged inputs for forecasting univariate time series. We also used a hybrid univariate time-series model through the hybridModel function in "forecastHybrid" R package from Shaub and Ellis (2020). The hybridModel function fits multiple individual model specifications to allow the easy creation of ensemble forecasts. With our data, the automated selected model from hybridModel function is (ETS, NNAR, THETAM, TBATS). THETAM fits an exponential smoothing state space model with an artificial neural network to the target variable, having first performed classic multiplicative seasonal adjustment. These two "forecast" and "forecastHybrid" R packages automatically select the best model in each time-series model based on the AIC model selection method.

4. Data Analysis

In this Section, we look at Granger-causality test and perform a comparison of forecasting methods. Firstly, Table 5 shows the Granger-causality test result, which shows that the Granger cause variables with lag 1 order to BTC are ADA, DOGE, ETC, and XLM, the Granger cause variables with lag 1 order to ETH are BTC, ADA, DOGE, ETC, and XLM, and the Granger cause variables with lag 1 order to XRP are BTC and XLM at the 5% significance level.

To forecast the log returns of BTC, we divide the data into 80% of the total observations (968), which are the training data (774 observations), and 20% of the total observations (968), which are the test data (194 observations). To compare the accuracy of the univariate time series models, we employ three measures.

Root mean square (prediction) error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{n}} \tag{3}$$

Mean absolute error deviation (MAE):

$$MAE = \frac{\sum_{t=1}^{n} |y_t - \hat{y}_t|}{n} \tag{4}$$

Weighted Mean absolute percentage error (WMAPE):

$$WMAPE = \frac{\sum_{t=1}^{n} |y_t - \hat{y}_t|}{\sum_{t=1}^{n} |y_t|} \tag{5}$$

The metric errors such as the RMSE, MAE, and WMAPE are used to analyze the performance of the methods. MAE is not sensitive to outliers as they are weighted less than the other observations when comparing actual and predicted values. RMSE takes bias and variance into account, but normalizes the units. Model 1 is a linear ECM model of BTC with 13 other cryptocurrencies; the summary of estimates is shown in Table 6. The R-squared of model 1 was 0.883. Model 2 is a nonlinear MARS-based ECM model of BTC with 13 other cryptocurrencies; the summary of estimates is shown in Table 7. The R-squared of nonlinear ECM (model 2) was 0.888. Table 8 shows the measures of accuracy of forecasting BTC. Among the eight different univariate time series models, model 1 has the smallest values of accuracy in terms of RMSE, MAE, and WMAPE. Thus, our proposed ECM prediction to BTC price log returns performed better than other univariate time-series models, including the neural network time-series model. Therefore, our ECM prediction model can help cryptocurrency market investors to identify threats to capital and earnings well from an uncertain and unexpected financial volatility. Financial policy committees in each country can reduce the difficulty of making future financial decisions with predictable cryptocurrency price log-returns information from our prediction model.

Table 5. Granger Causality Test.

This Table Presents the Correlation Matrix of the Granger Causality. ∗ Indicates Significance at the 5% Levels. Rows Are Granger-Cause with Lag 1 Order and Columns are Granger Effect.				
	BTC	**ETH**	**ADA**	**XRP**
BTC		0.006 **	0.701	0.040 *
ETH	0.768		0.755	0.121
ADA	0.003 **	0.006 **		0.068
BNB	0.370	0.351	0.792	0.378
USDT	0.407	0.475	0.520	0.871
XRP	0.085	0.052	0.627	
DOGE	0.003 **	0.001 **	0.069	0.278
USDC	0.229	0.482	0.455	0.827
LUNA	0.113	0.138	0.156	0.216
BCH	0.655	0.633	0.526	0.666
LTC	0.466	0.037 *	0.199	0.155
LINK	0.220	0.079	0.086	0.110
ETC	0.001 ***	0.001 **	0.116	0.128
XLM	0.000 ***	0.008 **	0.910	0.003 **

Note: * 0.10, ** 0.05, and *** 0.01 significance levels.

Table 6. Summary for model 1 with training data.

	Estimate	**Std. Error**	***t* Value**	***p*-Value**
(Intercept)	0.128	0.073	1.756	0.080
deltaLETH	0.379	0.035	10.960	0.000
deltaLADA	−0.048	0.025	−1.937	0.053
deltaLBNB	0.081	0.022	3.585	0.000
deltaLUSDT	0.577	0.220	2.622	0.009
deltaLXRP	−0.016	0.019	−0.811	0.418

Table 6. *Cont.*

	Estimate	Std. Error	*t* Value	*p*-Value
deltaLDOGE	0.043	0.010	4.405	0.000
deltaLUSDC	−0.458	0.215	−2.135	0.033
deltaLLUNA	0.005	0.002	2.316	0.021
deltaLBCH	0.156	0.027	5.736	0.000
deltaLLTC	0.145	0.031	4.701	0.000
deltaLLINK	−0.013	0.014	−0.932	0.352
deltaLETC	−0.024	0.025	−0.975	0.330
deltaLXLM	0.005	0.022	0.216	0.829
LETHLag1	0.311	0.049	6.395	0.000
LADALag1	−0.019	0.036	−0.525	0.600
LBNBLag1	0.079	0.030	2.586	0.010
LUSDTLag1	0.879	0.360	2.442	0.015
LXRPLag1	−0.034	0.026	−1.285	0.199
LDOGELag1	0.034	0.014	2.484	0.013
LUSDCLag1	−0.639	0.347	−1.841	0.066
LLUNALag1	0.006	0.003	1.781	0.075
LBCHLag1	0.157	0.038	4.124	0.000
LLTCLag1	0.139	0.042	3.344	0.001
LLINKLag1	−0.034	0.020	−1.714	0.087
LETCLag1	−0.010	0.034	−0.293	0.770
LXLMLag1	0.043	0.030	1.426	0.154
LBTCLag1	−0.964	0.037	−26.056	0.000

Note: delta in the table is integrated to the first order, I(1). Residual standard error: 1.988 on 745 degrees of freedom. Multiple R-squared: 0.883; adjusted R-squared: 0.879; F-statistic: 209 on 27 and 745 DF; *p*-value: < 0.000.

Table 7. Summary for model 2 with training data.

	Estimate of Coefficient
(Intercept)	−1.871
yLag1	−0.951
h(−15.1888 − deltaLETH)	−0.776
h(deltaLETH − 15.1888)	0.313
h(−6.33698 − deltaLDOGE)	−0.032
h(deltaLDOGE − 6.33698)	0.039
h(−4.69091 − deltaLBCH)	−0.159
h(deltaLBCH − 4.69091)	0.133
h(2.99004 − deltaLLTC)	−0.146
h(deltaLLTC − 2.99004)	0.221
h(−11.788 − deltaLETC)	0.131
h(2.75879 − LETHLag1)	−0.211
h(LETHLag1 − 2.75879)	0.299
h(LADALag1 − 14.5369)	0.434
h(16.1265 − LDOGELag1)	−0.079
h(12.7908 − LBCHLag1)	−0.150
h(−3.63013 − LLTCLag1)	−0.268
h(LLTCLag1 − 3.63013)	0.146

Note: Selected 18 of 22 terms, and 11 of 27 predictors. Termination conditions: RSq changed by less than 0.001 at 22 terms. Number of terms at each degree of interaction: 1 17 (additive model). GCV = 3.999; RSS = 2817.814; GR-Squared = 0.878; R-Squared = 0.888.

Table 8. Measures of accuracy of forecasting BTC.

	Model 1	Model 2	ARIMA	ETS	ARFIMA	BATS	NNAR	Hybrid
RMSE	2.531	2.728	4.295	4.294	4.296	4.372	4.292	4.316
MAE	1.845	1.926	3.220	3.220	3.222	3.309	3.218	3.245
WMAPE	0.574	0.600	1.002	1.002	1.003	1.030	1.002	1.010

5. Conclusions

By using linear ECMs and nonlinear ECMs, comprising six different univariate time series models such as neural network and autoregressive models to predict the price log returns of cryptocurrencies based on their previous values and relationships with each other, a better understanding can be achieved on whether they can be used to predict the log returns of BTC. We found that the linear ECM was the best model compared to other machine learning univariate time-series models. We can use linear ECM for predicting future log-return prices of each cryptocurrency with highly correlated cryptocurrencies.

Author Contributions: Conceptualization, J.-M.K.; methodology, C.J. and J.-M.K.; software, C.C. and J.-M.K.; validation, C.C. and J.-M.K.; formal analysis, C.C., C.J. and J.-M.K.; investigation, C.J.; resources, C.J. and J.-M.K.; data curation, C.J., and J.-M.K.; writing—original draft preparation, C.C., C.J., and J.-M.K.; writing—review and editing, C.C. and J.-M.K.; visualization, C.C., and J.-M.K.; supervision, C.C. and J.-M.K.; project administration, J.-M.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

Bansal, Gaurav. 2021. *Build Error Correction Models. R Package 'ecm'*. Vienna: Foundation for Statistical Computing.

Chang, Le, and Yanlin Shi. 2020. Does Bitcoin dominate the price discovery of the Cryptocurrencies market? A time-varying information share analysis. *Operations Research Letters* 48: 641–45. [CrossRef]

Diebold, Francis X., and Kamil Yilmaz. 2012. Better to give than to receive predictive directional measurement of volatility spillovers. *International Journal of Forecasting* 28: 57–66. [CrossRef]

Georgoula, Ifigeneia, Demitrios Pournarakis, Christos Bilanakos, Dionisios Sotiropoulos, and George M. Giaglis. 2015. Using Time-Series and Sentiment Analysis to Detect the Determinants of Bitcoin Prices. Paper presented at the 2015 Mediterranean Conference on Information Systems, Samos, Greece, October 3–5.

Giudici, Paolo, and Paolo Pagnottoni. 2020. Vector error correction models to measure connectedness of Bitcoin exchange markets. *Applied Stochastic Models in Business and Industry* 36: 95–109. [CrossRef]

Goczek, Łukasz, and Ivan Skliarov. 2019. What drives the Bitcoin price? A factor augmented error correction mechanism investigation. *Applied Economics* 51: 6393–410. [CrossRef]

Granger, Clive W. J. 1969. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica* 37: 424–38. [CrossRef]

Haffar, Adlane, and Eric Le Fur. 2021. Structural vector error correction modelling of Bitcoin price. *The Quarterly Review of Economics and Finance* 80: 170–78. [CrossRef]

Hakim das Neves, Rodrigo. 2020. Bitcoin pricing: Impact of attractiveness variables. *Financial Innovation* 6: 1–18. [CrossRef]

Hyndman, Rob, George Athanasopoulos, Chrisoph Bergmeir, Gabriel Caceres, Leanne Chhay, Mitchell O'Hara-Wild, Fotios Petropoulos, Slava Razbash, Earo Wang, Farah Yasmeen, and et al. 2021. *Forecast: Forecasting Functions for Time Series and Linear Models*. Vienna: Foundation for Statistical Computing.

Ibrahim, Ahmed, Rasha Kashef, Menglu Li, Esteban Valencia, and Eric Huang. 2020. Bitcoin Network Mechanics: Forecasting the BTC Closing Price Using Vector Auto-Regression Models Based on Endogenous and Exogenous Feature Variables. *Journal of Risk and Financial Management* 13: 189. [CrossRef]

Kapar, Burcu, and Jose Olmo. 2020. Analysis of Bitcoin prices using market and sentiment variables. *The World Economy* 44: 45–63. [CrossRef]

Keilbar, Georg, and Yanfen Zhang. 2021. On cointegration and cryptocurrency dynamics. *Digital Finance* 3: 1–23. [CrossRef]

Kwapień, Jarosław, Marcin Wątorek, and Stanisław Drożdż. 2021. Cryptocurrency Market Consolidation in 2020–2021. *Entropy* 23: 1674. [CrossRef] [PubMed]

Liang, Huan. 2021. Application of Error Correction model with monetary function factors in the return of bitcoin. *Journal of Physics: Conference Series* 1941: 012058. [CrossRef]

Miller, Dante, and Jong-Min Kim. 2021. Univariate and Multivariate Machine Learning Forecasting Models on the Price Returns of Cryptocurrencies. *Journal of Risk and Financial Management* 14: 486. [CrossRef]

Shaub, David, and Peter Ellis. 2020. *Convenient Functions for Ensemble Time Series Forecasts. R Package 'forecastHybrid'*. Vienna: Foundation for Statistical Computing.

Szetela, Beata, Grzegorz Mentel, Yuriy Bilan, and Urszula Mentel. 2021. The relationship between trend and volume on the bitcoin market. *Eurasian Economic Review* 11: 25–42. [CrossRef]

Wang, Junpeng, Yubo Xue, and Minghao Liu. 2016. An Analysis of Bitcoin Price Based on VEC Model. Paper presented at the 2016 International Conference on Economics and Management Innovations, Beijing, China, July 9–10.