# A differential evolution-based regression framework for forecasting Bitcoin price

R. K. Jana[1] · Indranil Ghosh[2] · Debojyoti Das[3]

## Abstract

This research proposes a differential evolution-based regression framework for forecasting one day ahead price of Bitcoin. The maximal overlap discrete wavelet transformation first decomposes the original series into granular linear and nonlinear components. We then fit polynomial regression with interaction (PRI) and support vector regression (SVR) on linear and nonlinear components and obtain component-wise projections. The sum of these projections constitutes the final forecast. For accurate predictions, the PRI coefficients and tuning of the hyperparameters of SVR must be precisely estimated. Differential evolution, a metaheuristic optimization technique, helps to achieve these goals. We compare the forecast accuracy of the proposed regression framework with six advanced predictive modeling algorithms- multilayer perceptron neural network, random forest, adaptive neural fuzzy inference system, standalone SVR, multiple adaptive regression spline, and least absolute shrinkage and selection operator. Finally, we perform the numerical experimentation based on—(1) the daily closing prices of Bitcoin for January 10, 2013, to February 23, 2019, and (2) randomly generated surrogate time series through Monte Carlo analysis. The forecast accuracy of the proposed framework is higher than the other predictive modeling algorithms.

✉ R. K. Jana
   rkjana1@gmail.com

   Indranil Ghosh
   fri.indra@gmail.com

   Debojyoti Das
   debojyoti.das@iimb.ac.in

[1] Operations and Quantitative Methods Area, Indian Institute of Management Raipur, Atal Nagar, Raipur, CG 493661, India

[2] Department of Operations Management and IT, Calcutta Business School, Kolkata, WB 743503, India

[3] Finance and Accounting Area, Indian Institute of Management Bangalore, Bengaluru, Karnataka 560076, India

# 1 Introduction

Cryptocurrencies are well known for their hedging, portfolio diversification, and risk mitigation capabilities across the literature (Bouoiyour et al. 2019; Bouri et al. 2018; Koutmos 2019). Bitcoin, the most popular cryptocurrency, has drawn crucial market players' attention at different hierarchies for accomplishing these activities. Simultaneously, the literature predominantly reports Bitcoin's return behavior as chaotic and multifractal (Al-Yahaee et al. 2018; Filho et al. 2018) and more volatile than other financial assets (Chaim and Laurini 2018).

Several researchers report the critical properties of Bitcoin and its interplay with other assets (Tiwari et al. 2018; Wang et al. 2019; Horra et al. 2019; Poyser 2019; Zhang et al. 2020). The assessment of the price dynamics of Bitcoin remains very challenging due to the high degree of nonstationary, nonlinear, volatile behaviors and the influence of various uncontrollable factors (Matta et al. 2016; Cretarola and Figà-Talamanca 2019). The conventional predictive models like moving average (MA), Autoregressive (AR), autoregressive integrated moving average (ARIMA) fail to produce accurate forecasts (Ghosh et al. 2019). Besides, these models suffer from several shortcomings like the strict requirement of the stationary of datasets, homoscedasticity of errors, etc. (Filho et al. 2018). Researchers propose machine learning, computational intelligence, deep learning, and hybrid models (Ghosh et al. 2017, 2019; Kumar and Ravi 2017) to alleviate these limitations.

This work presents an efficient regression-based predictive modeling framework for forecasting a one-day ahead price of Bitcoin. An accurate estimation of a one-day ahead price forecast may help the daily traders make better investment decisions and more profit. Support vector regression (SVR) and polynomial regression with interactions (PRI) help to learn the inherent governing patterns on the decomposed time series. The maximal overlapping discrete wavelet transformation (MODWT) decomposes the original time series. We use polynomial regression with interactions on linear sub-components and SVR on nonlinear sub-components to generate forecasts. Tuning of hyperparameters of SVR is necessary to ensure superior forecast accuracy from the two regression models. Also, we need to estimate the PRI coefficients. DE helps us to achieve these two tasks. We combine the two regression models' forecasts to generate the final prediction and carry out the proposed approach's comparative performance evaluation with MLP, RF, ANFIS, SVR, MARS, and LASSO, considering the original and five distinct surrogate time series simulated from the original Bitcoin price data.

This research's significant contribution is the proposition of a novel predictive modeling approach by combining the optimization and regression frameworks in a granular set up for time series forecasting. The proposed approach consists of four distinct components—MODWT for granular decomposition, SVR and PRI for learning pattern, and DE for hyperparameter tuning of SVR and the PRI model's estimation. It is different from the existing wavelet decomposition-based approaches in terms of using two distinct predictive modeling algorithms on appropriate granular components (Ghosh et al. 2017a, 2019). This approach is useful for predictive modeling of time series displaying an excess degree of volatility like the Bitcoin for making a more accurate forecast.

The rest of the paper's organization is as follows: Sect. 2 presents the previous research. Section 3 describes data and critical statistical properties of Bitcoin prices. Section 4 presents the research methodology, while Sect. 5 reports the results obtained from the proposed framework. Section 6 concludes the paper.

## 2 Previous research

This section clusters the pertinent work to inspect Bitcoin's temporal characteristics, interaction dynamics with other assets, and predictive modeling. Several researchers study the temporal characteristics of Bitcoin (Bariviera 2017; Nadarajah and Chu 2017; Takaishi 2018). A section of literature proves that the Bitcoin market is inefficient (Urquhart 2016; Kristoufek et al. 2018; Tiwari et al. 2018). Researchers examine the same market's multifractal nature through multifractal detrended fluctuation analysis (Filho et al. 2018; Kristjanpoller et al. 2020; Takaishi 2018). The findings suggest that the degree of multifractal nature is more substantial in Bitcoin than several other stock market indices. A copula-based Granger causality test explores the causal interplay between Bitcoin's attention manifested through google trends and Bitcoin returns from January 1, 2013, to December to December 31, 2017 (Dastgir et al. 2019). The study reveals the existence of bidirectional causal nexus in the left and right tail of the distribution.

The interaction between Bitcoin returns and economic policy uncertainty (EPU) explores the hedging capacity of Bitcoin (Demir et al. 2018; Paule Vianez et al. 2020). Wang et al. (2019) measure the risk spillover effect from EPU to Bitcoin and selected equity market uncertainty index, US EPU index, and implied volatility index as proxies for EPU. The results indicate the presence of negligible spillover from EPU to Bitcoin. So, Bitcoin may act as a safe asset for investment under EPU shocks. The Bayesian structural time series approach inspects the temporal dynamics and dependence structure of Bitcoin prices on other assets (Poyser 2019). The findings reveal both positive and negative association structures with different foreign exchange rates, gold prices, and stock indices. The LASSO technique identifies EPU, gold returns, and search intensity as the significant drivers of Bitcoin returns (Panagiotidis et al. 2018).

The literature reports various artificial intelligence and predictive modeling approaches for forecasting future movements of financial assets. For example, a combination of radial basis function neural network (RBFN), K-means clustering, and artificial fish swarm algorithm (AFSA) estimates the future movements of the Shanghai stock exchange (Shen et al. 2011). The model performs better than five other approaches. Similarly, a combination of discrete wavelet transforms (DWT), MARS, and SVR predicts stock market returns (Kao et al. 2013). This approach outperforms DWT-MARS, DWT-SVR, ARIMA, standalone SVR, and ANFIS. Similar other hybrid models exist for stock market prediction (Oztekin et al. 2016; Wang et al. 2016; Chen and Hao 2017). These models also perform better compared to the traditional and some other hybrid forecasting approaches. Unfortunately, they are not used to predict the future prices of Bitcoin.

Various techniques like neural network (Jang and Lee 2017), neuro-fuzzy (Atsalakis 2019), machine learning (Mallqui and Fernandes 2019), deep learning (McNally et al. 2018; Wu et al. 2018), and deep neural network (Nakano et al. 2018), deep learning chaotic neural networks (Lahmiri and Bekiros 2019) predict the directional predictive modeling of daily prices of Bitcoin. These techniques produce superior predictive performances over their traditional counterparts and are more useful for nonlinear and chaotic financial markets. The literature reveals that Bitcoin is very closely associated with a myriad of global financial assets. The multifractal interactions and spillovers of Bitcoin with these assets make the temporal dynamics very complicated. It is challenging to identify explanatory variables beforehand due to these associations. The degree of (in)efficiency owing to the Bitcoin bubble and other extreme events further complicates

the predictive analysis task. Therefore, developing a robust predictive framework for predicting future prices of Bitcoin can be of paramount importance for practical implications.

Time-series decomposition-based predictive approaches that utilize lagged information [滞后信息] as explanatory features for estimating financial assets' future movements emerge to be successful. However, it is rare to find a granular framework-based predictive model driven by lagged information for forecasting Bitcoin prices. The present work contributes to the existing literature by proposing a granular predictive architecture that can recognize Bitcoin price movements' intricate patterns. Wavelet decomposition is an instrumental paradigm for this purpose. SVR works as the learning algorithm. Deployment of DE for optimizing SVR parameters augments the quality of the final forecasts. Wavelet decompositions generate granular linear and nonlinear subseries. Pattern recognition algorithms apply to these subseries for generating forecasts. The use of SVR on all the components may result in overfitting. Thus, we apply metaheuristic driven SVR to the nonlinear component and PRI model to the linear component to reduce this effect. The DE algorithm selects the PRI parameters. [here it explains the method of research]

## 3 Data description

We compile daily closing prices of Bitcoin for January 10, 2013, to February 23, 2019 for conducting the predictive exercise. During late 2017, Bitcoin prevails to be the most dominant cryptocurrency with more than 42% market share capitalization. As a result, we observe a steady upward growth of Bitcoin price during this period. Due to heavy criticism and the Bitcoin bubble, the growth eventually declines afterward. So, we consider the study period consisting of significantly normal, bullish, and bearish phases. Figure 1 portrays the temporal evolutionary dynamics of Bitcoin prices during the considered period. Table 1 reports the
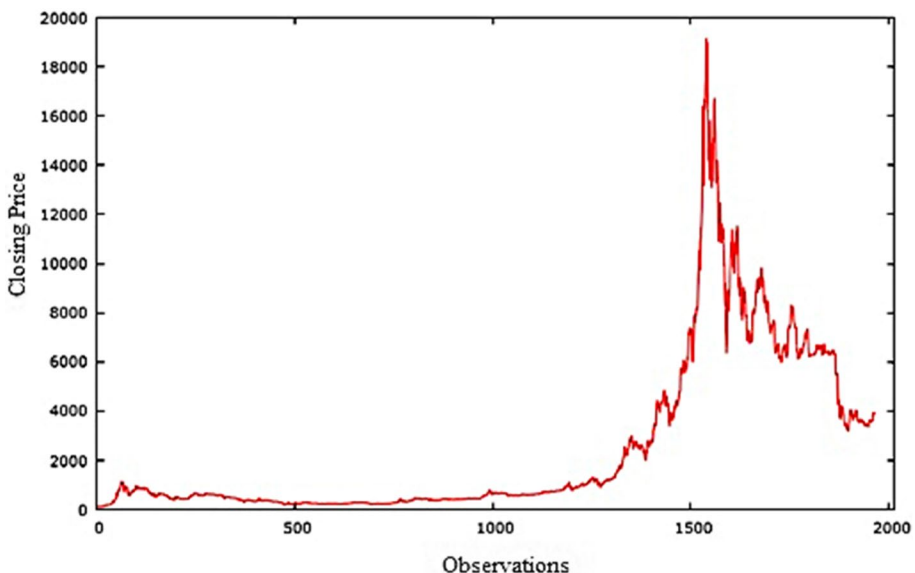
[展示折线图进行统计分析的说明]



**Fig. 1** Temporal evolutionary pattern of Bitcoin

results of the statistical tests and additional properties to comprehend the underlying dataset's empirical properties.

Financial time series exhibits nonstationary, nonparametric, and chaotic traces (Avdoulas et al. 2018; Guerard et al. 2020). Mere visualization of the temporal evolutionary patterns may not capture such properties distinctly. Throughout the literature, researchers rely upon statistical and econometric tests for decoding such properties (Das et al. 2018; Ghosh et al. 2020; Jana et al. 2019; Boukhatem et al. 2020). The present work follows the same principles for comprehending the fundamental temporal properties of Bitcoin prices.

Outcomes of Jarque–Bera and Frosini tests indicate that daily closing prices of Bitcoin over the considered period do not abide by the normal distribution. Test statistics of Philip-Perron and Augmented Dickey-Fuller tests emerge insignificant, indicating unit-roots' presence in Bitcoin price dynamics. So, the time series under consideration is nonstationary. Therefore, the chosen time series's complex evolutionary pattern implies that the traditional econometric models for forecasting are not appropriate for deriving forecasts. We perform Terasvirta's and White's NN tests to further evaluate the behavioral patterns for gaining more in-depth insights. The test statistic values of both these tests are significant. So, there exists a strong nonlinear trait in the evolutionary patterns of Bitcoin prices. The Hurst exponent identifies the randomness and persistence of the time series. The Hurst exponent value of 0.89404 ($> 0.5$) implies the presence of fractional Brownian motion with a persistent trend (Ghosh et al. 2017). The persistent trend further indicates the presence of volatility clustering. Thus, the immediate past information on Bitcoin price does play a significant role in explaining future behavior. The present study exploits this property by setting one, two, three, four, and five day back Bitcoin prices as independent variables to predict a one-day ahead price. The partial autocorrelation plots shown in Sect. 5 validates the selection of five lagged variables.

## 4 Research methodology

The proposed optimization-based regression approach has four distinct components—MODWT, SVR, PRI, and DE. As the closing prices of Bitcoin are persistent, we select the explanatory variables based on the lagged information. This section elucidates the details of individual components.

### 4.1 MODWT

MODWT decomposes the original Bitcoin prices into granular subcomponents. It determines the respective wavelet coefficients accounting for different time horizons. Daubechies least asymmetric filter of length eight (la8) helps in the decomposition process. The technique has additional advantages over the conventional 'Haar' wavelet filters in smoother operations (Gencay et al. 2002) and the generation of unrelated coefficients (Cornish et al. 2006). MODWT segregates the Bitcoin price time series into time and frequency space by expressing the sequence as projections of the father ($\phi$) and mother ($\psi$) wavelets representing approximation and units, respectively. Mean figures of approximation and detail components are one and zero, as shown in the following equations:

$$\int \phi(t)dt = 1 \qquad (1)$$

**Table 1** Empirical properties of Bitcoin price

| Jarque–Bera Test | Frosini Test | Philip-Perron Test | Augmented Dickey-Fuller Test | Terasvirta's Test | White's NN Test | Hurst Exponent |
|---|---|---|---|---|---|---|
| 2031.86*** | 5.5217*** | − 9.5062# | − 2.5873# | 33.7355*** | 23.2801*** | 0.89404 |

#Not Significant, ***Significant at 1% level of significance

$$\int \psi(t)dt = 0 \tag{2}$$

We express the long-scale linear component through the father wavelet, whereas the mother wavelet represents the linear component's nonlinear and volatile deviations. Mathematically, we write them as:

$$\phi_{(j,k)} = 2^{-j/2}\phi\left(\frac{t - 2^j k}{2^j}\right) \tag{3}$$

$$\psi_{(j,k)} = 2^{-j/2}\psi\left(\frac{t - 2^j k}{2^j}\right) \tag{4}$$

From the original time series, we can approximate the coefficients as follows:

$$s_{(j,k)} = \int f(t)\phi_{(j,k)} \tag{5}$$

$$d_{(j,k)} = \int f(t)\psi_{(j,k)} \tag{6}$$

The combination of wavelet functions formulates the original time series $f(t)$ as follows:

$$f(t) = \sum_k s_{J,k}\phi_{J,k}(t) + \sum_k d_{J,k}\psi_{J,k}(t) + \cdots + \sum_k d_{j,k}\psi_{j,k}(t) + \cdots + \sum_k d_{1,k}\psi_{1,k}(t) \tag{7}$$

We write Eq. (7) as:

$$f(t) = S_J + D_J + D_{J-1} + \cdots + D_j + \cdots + D_1 \tag{8}$$

where

$$S_J = \sum_k s_{j,k}\phi_{j,k}(t) \tag{9}$$

$$D_J = \sum_k d_{j,k}\psi_{j,k}(t), j = 1, 2, \ldots, J \tag{10}$$

The orthogonal constituents $[S_J, D_J, D_{J-1}, \ldots, D_1]$ are the outcomes of the decomposition process of the original time series $f(t)$. $D_J$ denotes the wavelet detail at $J^{th}$ level. $S_J$ defines the sum of deviations at the detail scales. In the proposed predictive analytics exercise, we carry out six decomposition levels that resulted in one approximation or linear component and six detail components accounting for high frequency or nonlinearity traits of the time series. Subsequently, predictive algorithms applied to the respective components yield predictions. For the MODWT decomposition.

The time series consisting of the closing prices of Bitcoin along with explanatory features become inputs for the decomposition exercise. The outputs comprise of granular decomposed series of Bitcoin prices and other explanatory features.

## 4.2 Support vector regression (SVR)

SVR (Vapnik, 1995) is an extension of the SVM algorithm applicable for performing nonlinear regression. A nonlinear mapping function $\theta(x)$ projects the input $x_i \in R^p$ to a feature space with a high dimension and establishes a linear relationship between the target and predictor constructs. Mathematically, we represent this as follows:

$$f(x) = w^T \theta(x) + b \tag{11}$$

where $w$ and $b$ are the weight and bias parameters, respectively, account for the decision boundary.

The goal is to estimate the parameters that eventually result in at most $\varepsilon$ deviation from the target for all training samples. This set up of SVR is known as $\varepsilon$-SVR.

$$\left| y_i - f(x_i) \right|_\varepsilon = \begin{cases} 0 & if \left| y_i - f(x_i) \right| \le \varepsilon \\ \left| y_i - f(x_i) \right| - \varepsilon & Otherwise \end{cases} \tag{12}$$

where $y_i$ denotes the target variable. We estimate $f(x)$ by minimizing the following quadratic optimization problem through determining the parameters $w$ and $b$ using slack constructs, $\xi_i, \xi_i^*$:

$$Minimize : \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} (\xi_i + \xi_i^*) \tag{13}$$

subject to

$$\begin{cases} w^T x_i - y_i \le \varepsilon + \xi_i \\ y_i - w^T x_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0, i = 1, \dots, n \end{cases} \tag{14}$$

where $C$ is the regularization parameter represents the trade-off between the fitness of regression and resultant error in training. We write the dual problem with the final solution using Eq. (11) as follows:

$$f(x) = \sum_{i=1}^{n} \left( \alpha_i - \alpha_i^* \right) k(x_i, x_i^*) + b \tag{15}$$

where $k(x_i, x_i^*)$ is a kernel function, $\alpha_i$ and $\alpha_i^*$ are Lagrange multipliers. We use the Radial basis kernel (RBF) defined by $k(x_i, x_j) = \exp\left(-\|x_i - x_j\|^2 / 2\sigma^2\right)$ where $\sigma$ denotes the width of the kernel. The three critical parameters $C$, $\sigma$, and $\varepsilon$ that drive the learning process and affect the prediction performance. The DE provides their optimum values. The SVR-DE learning model applied to the nonlinear decomposed components obtained through MODWT performs the predictive analytics task.

多项式回归

## 4.3 Polynomial regression with interaction (PRI)

这个地方其实是比较重要的，好像是比较传统的统计方法

PRI is an extension of simple linear regression. It fits a polynomial to establish the relationship between the target and explanatory variables. Usually, the ordinary least square (OLS) technique estimates the regression coefficients, excluding the predictors' interactions. These two

issues lead to the underfitting of the regression curve when time-series movements are highly <mark>nonstationary and nonlinear</mark>, as in Bitcoin prices. We use a third-order quadratic regression model with one level of interaction with five lagged variables as explanatory variables. The polynomial of order three with one level of interaction is as follows:

$$Y_t^{est} = \alpha + \sum_{i=1}^{3} \beta_i Y_{t-1}^i + \sum_{j=1}^{3} \beta_{j+3} Y_{t-2}^j + \sum_{k=1}^{3} \beta_{k+6} Y_{t-3}^k + \sum_{l=1}^{3} \beta_{l+9} Y_{t-4}^l + \sum_{m=1}^{3} \beta_{m+12} Y_{t-5}^m$$

(five lagged variables as explanatory variables)

$$+ \sum_{n=1}^{4} \delta_n Y_{t-1} Y_{t-n-1} + \sum_{o=1}^{3} \delta_{o+4} Y_{t-2} Y_{t-n-2} + \sum_{p=1}^{2} \delta_{p+7} Y_{t-3} Y_{t-p-3} + \sum_{q=1}^{1} \delta_{q+9} Y_{t-4} Y_{t-n-4}$$

(quadratic regression 二次项解释项)

(16)

where $Y_t^{est}$ represents the estimated Bitcoin price at time t, $Y_{t-i}$ accounts for the previous information of Bitcoin price at $i(i = 1, 2, \ldots, 5)$ period back, $\alpha$, $\beta$, and $\delta$ are the respective coefficients of polynomial and interaction terms. 残差吗?

The DE estimates the regression coefficients $\alpha$, $\beta_i$ and $\delta_i$ given in the Eq. (16) as follows:

$$\begin{cases} Find : \alpha, \beta_i, \delta_i \\ Argmin : \sum \left( Y_t^{act} - Y_t^{est} \right)^2 \end{cases}$$

(17)

where $Y_t^{act}$ represents actual Bitcoin price at time $t$.

## 4.4 Differential evolution

The DE is a widely used metaheuristic for solving complex optimization problems in the continuous domain (Storn and Price 1997). It is a population-based searching algorithm that traverses the search space intelligently to obtain local or near-optimal solutions. The following are the detailed procedures of individual phases of the DE.

### 4.4.1 Initialization

The initial population of the DE comprises of $P$ vectors $\{\bar{x}_{i,G}, i = 1, 2, \ldots, P\}$, where G denotes the generation number. We randomly generate the initial population from a uniform distribution specified lower and upper bounds. The possible solutions can be represented as a D-dimensional vector as follows:

$$\bar{x}_{i,G} = \left( x_{i,1,G}, x_{i,2,G}, \ldots, x_{i,D,G} \right)$$

(18)

The selection, recombination, and mutation operations subsequently modify the initial population.

### 4.4.2 Mutation

In traditional DE, three vectors estimate the mutation vector $\left( \bar{v} \right)$. Mathematically, we may represent it as follows:

$$\bar{v}_{i,G} = \bar{x}_{r_1,G} + F * \left( \bar{x}_{r_2,G} - \bar{x}_{r_3,G} \right)$$

(19)

where $\bar{v}_{i,G}$ is the donor vector, and $F$ is the scaling parameter that varies between 0 and 2.

### 4.4.3 Recombination

In this phase, the components of a donor vector generate a trial offspring vector as follows:

$$u_{i,j,G} = \begin{cases} v_{i,j,G}, & if\ rand_j \leq CR\ or\ j = rn(j) \\ x_{i,j,G}, & otherwise \end{cases} \tag{20}$$

where $rand_j(j = 1,2,\ldots,D)$ denotes random numbers between 0 and 1; $rn(j)$ denotes a randomly selected index from $1,2,\ldots,D$ and $CR$ is the crossover parameter between 0 and 1.

### 4.4.4 Selection

The DE uses a greedy approach for selecting candidate vectors for the subsequent generations. It compares trial and target vectors and makes the final selection as follows:

$$\overline{x}_{i,G+1} = \begin{cases} \overline{u}_{i,G}, & if\ f\left(\overline{u}_{i,G}\right) < f\left(\overline{x}_{i,G}\right) \\ \overline{x}_{i,G}, & otherwise \end{cases} \tag{21}$$

The number of generations can be set depending upon the problem's nature or fixed until the convergence in solutions is noticed.

The DE tunes the hyper-parameters of the SVR model. It randomly initializes the values of different parameters as input and traverses the search space intelligently to yield optimal input parameters. In the PRI model, the DE randomly initializes the value of the coefficients. After completion of the search process, it provides the optimal values of the same as the output.

The combination of SVR-DE and PRI-DE predictions generate the final forecast of the next-day Bitcoin price. The SVR-DE approach yields granular forecasts $\widehat{Y}_t^{d_i}(i = 1,2,\ldots,6)$ from disentangled nonlinear components provided by MODWT. The PRI-DE approach produces granular forecasts $\widehat{Y}_t^{s_6}$ from the approximation component fetched by MODWT. We write the final forecast ($\widehat{Y}_t^{Final}$) as follows:

$$\widehat{Y}_t^{Final} = \widehat{Y}_t^{d_1} + \widehat{Y}_t^{d_2} + \widehat{Y}_t^{d_3} + \widehat{Y}_t^{d_4} + \widehat{Y}_t^{d_5} + \widehat{Y}_t^{d_6} + \widehat{Y}_t^{s_6} \tag{22}$$

## 4.5 Predictive performance indicators

Mean squared error (MSE) measures the fitness of the SVR-DE model as follows:

$$MSE = \frac{1}{N} \sum_{t=1}^{N} \left(\widehat{Y}_t - Y_t\right)^2 \tag{23}$$

where, $\widehat{Y}_t$ and $Y_t$ are predicted and actual figures. The PRI-DE model considers Eq. (17) as the fitness function.

*Nash–sutcliffe efficiency (NSE)* NSE measures the residual variance's relative strength resulting from a predictive model with the original variance of the dataset.

$$NSE = 1 - \frac{\sum_{t=1}^{N}\left\{\widehat{Y}_t - Y_t\right\}^2}{\sum_{i=1}^{N}\left\{Y_t - \bar{Y}_t\right\}^2} \qquad (24)$$

The NSE values lie between $(-\infty)$ to 1. The NSE value close to 1 implies excellent predictive performance by the respective model.

*Index of agreement (IA)* IA is a measure that considers the magnitude of error generating from a model to assess the predictive capability.

$$IA = 1 - \frac{\sum_{t=1}^{N}(\widehat{Y}_t - Y_t)^2}{\sum_{i=1}^{N}\left\{\left|\widehat{Y}_t - \bar{Y}_t\right| + |Y_t - \bar{Y}_t|\right\}^2} \qquad (25)$$

Like NSE, the IA figure should ideally be close to 1 for efficient forecasting.

*Theil Index (TI)* We can compute the index as follows:

$$TI = \frac{\left[\frac{1}{N}\sum_{t=1}^{N}\left(\widehat{Y}_t - Y_t\right)^2\right]^{1/2}}{\left[\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{Y}_t\right)^2\right]^{1/2} + \left[\frac{1}{N}\sum_{i=1}^{N}\left(Y_t\right)^2\right]^{1/2}} \qquad (26)$$

Unlike NSE and IA, TI values need to be close to 0 to imply a predictive model's effectiveness.

*Directional predictive accuracy (DA)* For calculating the DA, we use the following relationships:

$$DA = \frac{1}{N}\sum_{t=1}^{N}D_t, D_t = \begin{cases} 1, & (Y_{t+1} - Y_t)\left(\widehat{Y}_{t+1} - Y_t\right) \geq 0 \\ 0, & Otherwise \end{cases} \qquad (27)$$

The DA values near 1 indicate high accuracy of directional prediction obtained by the model, while DA figures near 0 imply the opposite scenario. As the present paper deals with regression, DA's computation eventually evaluates the presented forecasting framework's directional predictive capacity. Models having high DA values are attractive for traders for profitable trading.

Apart from self-assessment of the proposed forecasting approach's predictive performance, the paper also carries out a comparative study with six advanced predictive modeling algorithms: MLP, ANFIS, RF, SVR, MARS LASSO to justify the usage of the prediction framework. We conduct the Diebold-Mariano Statistical test for inspecting the equal predictive ability to draw inference on comparative assessment.

*Diebold-Mariano (DM) test* The DM test assesses the predictive accuracy of multiple forecasting models. We use the mean square prediction error (MSPE) as the loss function in the DM test. The DM statistic (DMS) has the following form:

$$DMS = \frac{\bar{D}}{\sqrt{V_{\bar{D}}/M}} \qquad (28)$$

where

$$\bar{D} = \frac{1}{M} \sum_{t=1}^{M} d_t \tag{29}$$

$$d_t = \frac{1}{M} \sum_{t=1}^{M} \left( \left( Y_t - \widehat{Y}_t^{test} \right)^2 - \left( Y_t - \widehat{Y}_t^{bench} \right)^2 \right) \tag{30}$$

$$V_{\bar{D}} = \gamma_0 + 2 \sum_{q=1}^{\infty} \gamma_q \tag{31}$$

$$\gamma_q = cov\left( d_t, d_{t-q} \right) \tag{32}$$

$\widehat{Y}_t^{test}$ and $\widehat{Y}_t^{bench}$ are predicted values achieved by test and benchmark models. Figure 2 displays the proposed research methodology.

The decomposition-based predictive model utilizes two learning methods—SVR-DE and PRI-DE. The DE tunes the hyper-parameters of the SVR model and estimates the coefficients of the PRI model.

## 5 Results and discussions

We evaluate the proposed approach's performance based on the original Bitcoin price series and five surrogate time series randomly generated through the Monte-Carlo technique from the original Bitcoin price series. The surrogate series are more chaotic and random. Therefore, the performance evaluation based on these series is essential to ascertain the robustness and capability of the proposed forecasting approach. The parameter settings remain constant for both cases.

### 5.1 Performance evaluation based on the original time series

The proposed framework utilizes one-day, two-day, three-day, four-day, and five-day back lagged closing prices of Bitcoin as explanatory features. The partial autocorrelation function (PACF) plot (Fig. 3) justifies the selection.

The first local minima of the PACF plot occurs at lag 5. Subsequently, it exhibits gradual upward and downward movements. The estimated Hurst exponent presented in Table 1 indicates that the series abide by the persistence trend. Therefore, we select the lag at which the PACF value initially reaches close to 0 as the number of lagged variables to form the independent feature set. Then, MODWT applies to the Bitcoin price dataset and explanatory features for six levels of decomposition. Figure 4 displays the original and decomposed series of Bitcoin prices.

This study utilizes the methodology presented by Bou-Hamad and Jamali (2020) to evaluate in-sample (training) and out-of-sample (validation) forecasting performances. We conduct static one-step ahead and dynamic multiple step-ahead forecasting schemes. For both cases, the validation period is the last 252 observations that reflect the movements of Bitcoin prices of the previous year of the considered sample. A rolling window of 747 observations in the static forecasting scheme estimates the one-day ahead closing price.
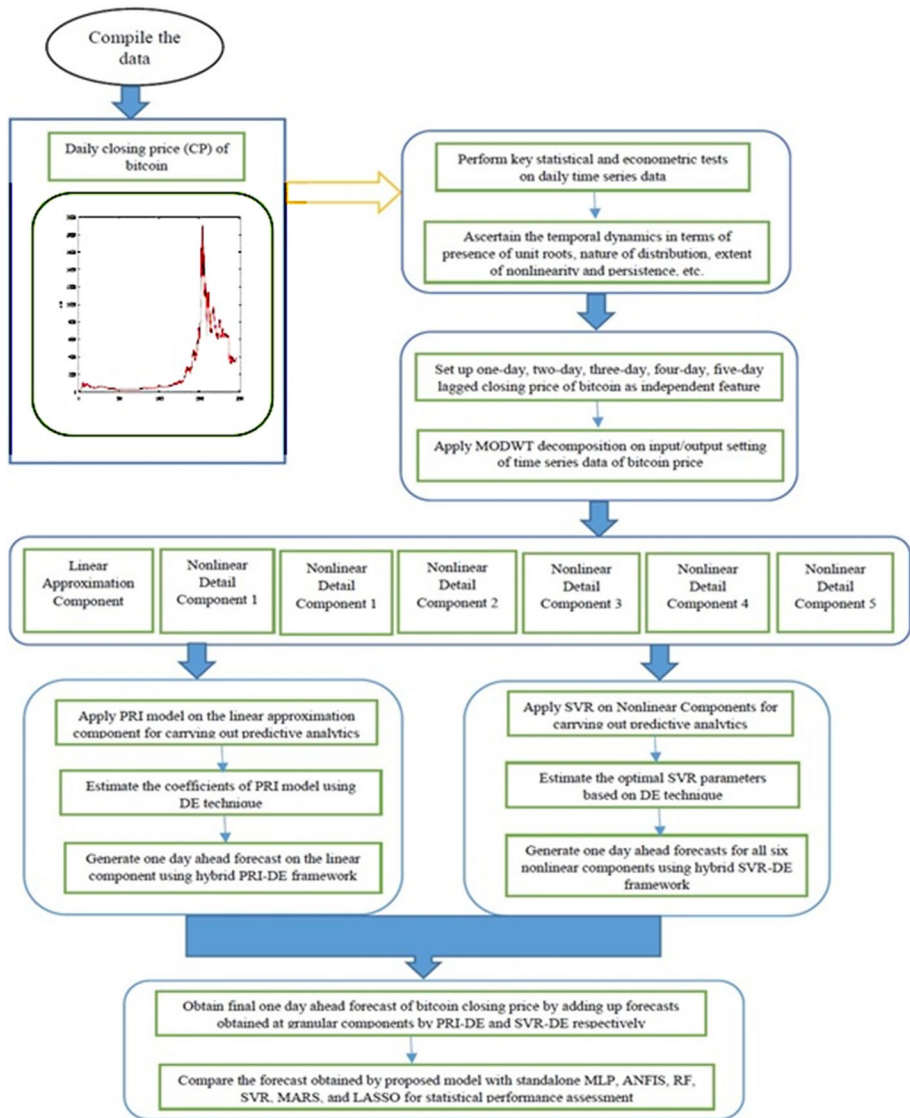
**Fig. 2** Proposed research methodology

The considered sample has a total of 1959 observations. To predict the price of 748th observation, samples spanning from observations 1 to 747 constitute the training period's window. The rolling window omits the oldest information and adds the latest observation while traversing from training to validation segments. Training samples 3 to 749 generates the forecast for 750th observation. Similarly, training samples 13 to 759 generates the forecast for 760th observation in the one-step-ahead forecasting framework.

In a multi-step-ahead scheme, the predictive framework uses the entire in-sample segment for training. Then, we partition the decomposed dataset randomly into training

**Fig. 3** PACF plot
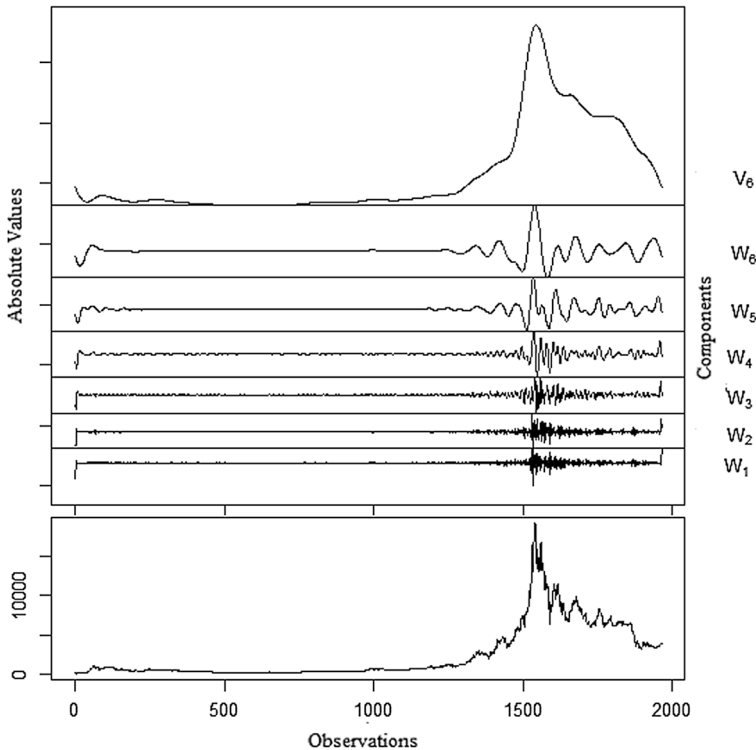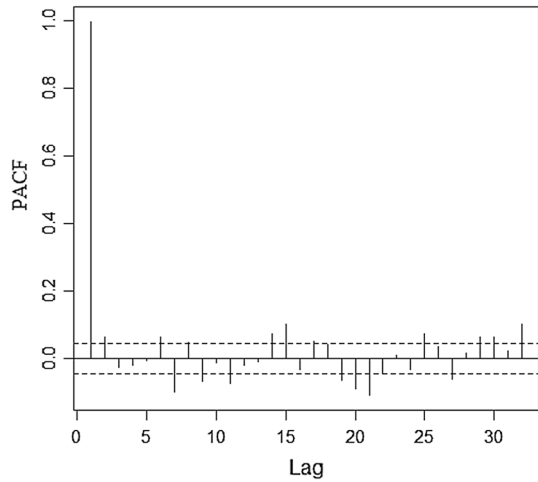




**Fig. 4** MODWT decomposition of Bitcoin price

(80%) and test (20%) data. The SVR-DE and PRI-DE generate forecasts from the respective components. The DE algorithm estimates the SVR hyperparameters. The number of iterations and the population size are 200 and 20, respectively. Figure 5 depicts the
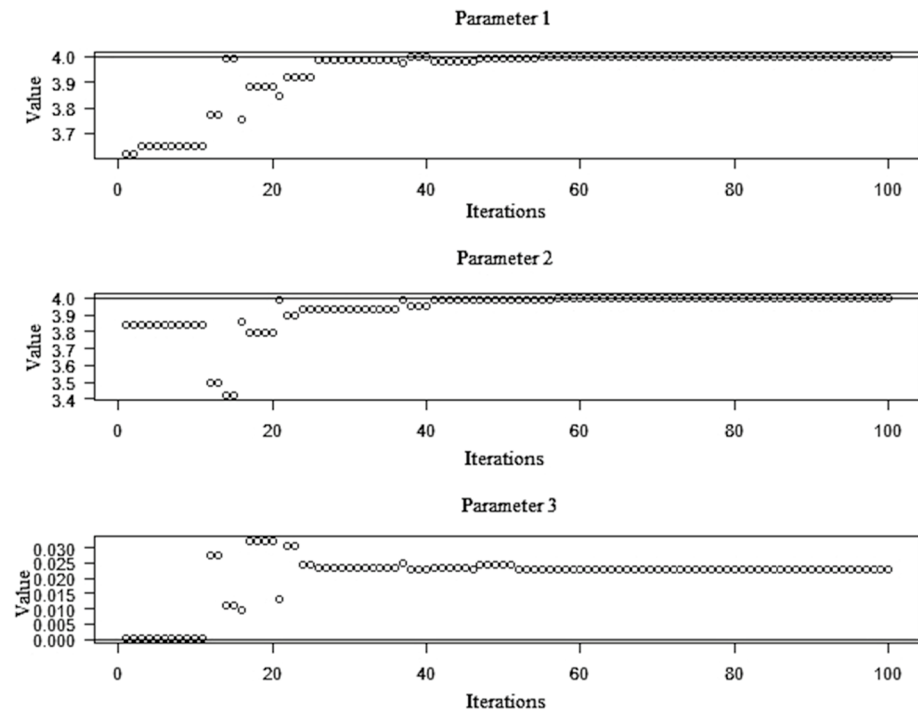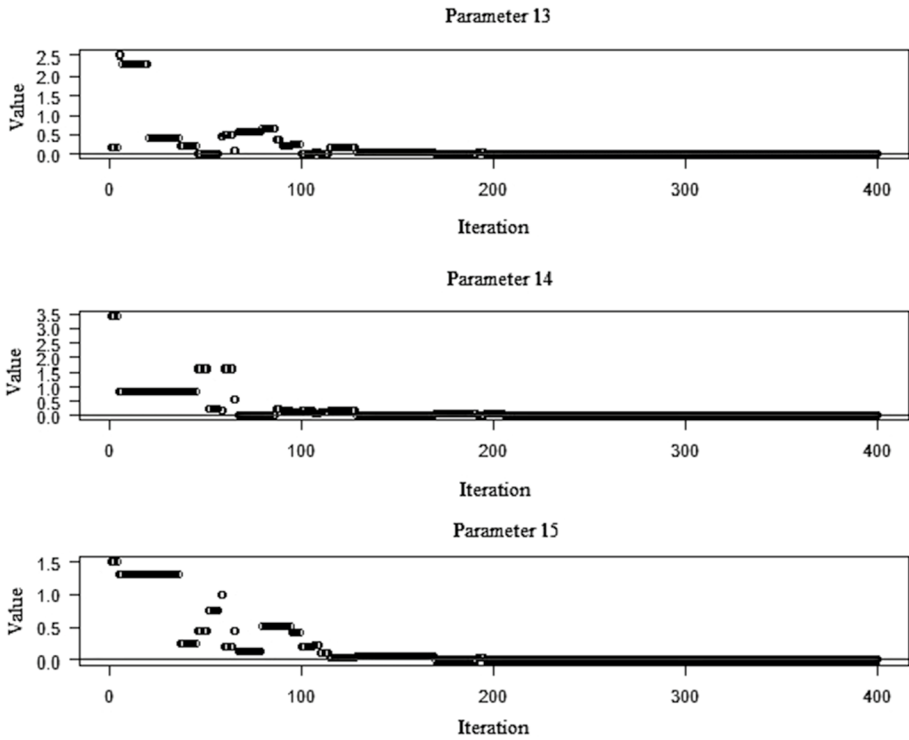
**Fig. 5** Iteration wise performance of the DE in SVR-DE model

performance of the DE in estimating the three parameters $C$, $\sigma$, and $\varepsilon$. We notice that the DE algorithm converges while determining the three parameters. The DE also estimates the coefficients of the PRI-DE model. As the number of parameters is significantly higher in the PRI model, the number of iterations and population size is 400 and 80, respectively. Figure 6 displays a part of the parameter estimation process betraying iteration wise details for a few selected parameters.

The presence of convergence is apparent. After completing parameter estimation and component-wise predictions, we compute the final one day ahead forecast of Bitcoin price using Eq. (20). Table 2 reports four performance indicators' estimated values on out-of-sample data points in static and dynamic setups.

The NSE and IA values of out-of-sample segments in static and dynamic cases are substantially high and close to 1 and show the supremacy of the predictive model. The TI values are low as well, which is ideal for a predictive model. Lastly, the DA statistic is close to 1 for the training and test dataset, which signifies the precise one-day ahead of Bitcoin price predictions. The model has noteworthy performance in estimating the change of price direction of Bitcoin. Hence, the forecasting framework can effectively predict the rise and fall movements, enabling investors to exploit Bitcoin for profitable trading in both short and long terms. The performance of static forecasting is marginally better than that of dynamic forecasting. Thus, the accuracy of the immediate future forecast is comparatively more than the long run futuristic projections. Figure 7 exhibits the actual versus predicted and fitment of actual and forecasted price on out-of-sample data.

**Fig. 6** Iteration wise performance of the DE in PRI-DE model

**Table 2** Forecasting performance of the proposed framework

| Series | IA | TI | DA | NSE |
|---|---|---|---|---|
| *Static forecasting* | | | | |
| Bitcoin price | 0.999703 | 0.027424 | 0.989514 | 0.998841 |
| *Dynamic forecasting* | | | | |
| Bitcoin price | 0.999542 | 0.028139 | 0.988408 | 0.998593 |

Both the figures further validate the claims inferred from the values of four performance indicators. The original price and predicted price display very little discriminability between the actual and forecasted. Table 3 summarizes the key parameters of six competing models.

We generate forecasts from all the competing models in a univariate setup by keeping the same training and test dataset. Subsequently, we calculate the four-performance metrics for comparison. Table 4 furnishes the comparative assessment findings of the proposed approach with the six competing algorithms on test segments of static and dynamic forecasting settings.

The proposed model outperforms the rest of the models as NSE, IA, and DA figures are higher, while the TI is lower than their counterparts. The competing models' forecasts are above the satisfactory level as the NSE, IA, and DA values are higher than 0.9, and the TI value is considerably low. The performance of competing models degrades in the
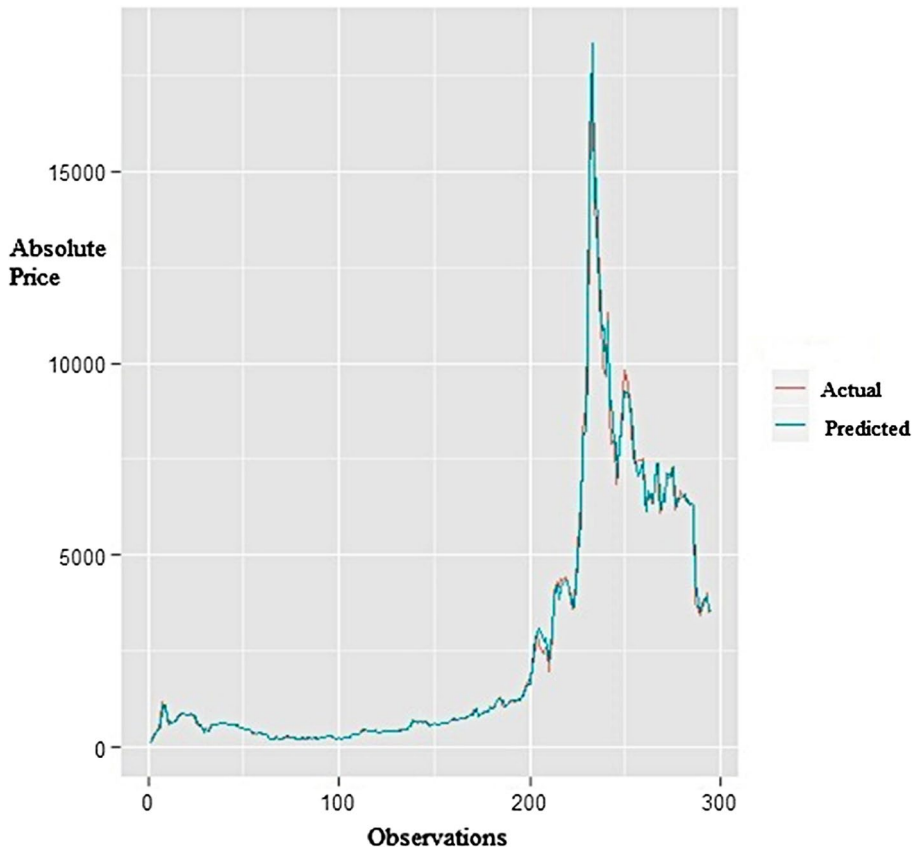
**Fig. 7** Actual and predicted Bitcoin price visualization

**Table 3** Parameters of competing models

| Model | Key parameters |
|---|---|
| MLP | Learning rate = 0.7, momentum = 0.9, number of hidden nodes = 20 |
| ANFIS | Membership function = Gaussian |
| Rejected | |
| RF | Number of trees = 500, features for branching = 2 |
| SVR | $C$=3.5, $\sigma$=3.2, $\varepsilon$=0.1 |
| MARS | Degrees of interaction = 1, generalized cross validation penalty per knot = 2 |
| LASSO | Lambda (regularization) = 0.3 |

multi-step ahead dynamic forecasting. The DM derives the statistical significance and gains more in-depth insights into the competing models' performance. Table 5 shows the outcome of the test.

The DM test is sensitive to the order of constituents of a pair. A positively significant test statistic value implies the statistical superiority of the second model over the first

**Table 4** Forecasting performance evaluation

| Series | Proposed | MLP | ANFIS | RF | SVR | MARS | LASSO |
|---|---|---|---|---|---|---|---|
| *Static forecasting* | | | | | | | |
| IA | 0.999703 | 0.991793 | 0.992093 | 0.992957 | 0.992777 | 0.988731 | 0.990769 |
| TI | 0.027424 | 0.055312 | 0.044386 | 0.039149 | 0.038274 | 0.048718 | 0.047633 |
| DA | 0.989514 | 0.971166 | 0.973414 | 0.980486 | 0.981639 | 0.962394 | 0.967938 |
| NSE | 0.998841 | 0.998020 | 0.994762 | 0.995152 | 0.994953 | 0.987845 | 0.983132 |
| *Dynamic forecasting* | | | | | | | |
| IA | 0.999542 | 0.991766 | 0.992069 | 0.992944 | 0.992759 | 0.988716 | 0.990734 |
| TI | 0.028139 | 0.055327 | 0.044405 | 0.039162 | 0.038283 | 0.048728 | 0.047669 |
| DA | 0.988408 | 0.971154 | 0.973398 | 0.980474 | 0.981627 | 0.962367 | 0.967917 |
| NSE | 0.998593 | 0.998007 | 0.994741 | 0.995127 | 0.994938 | 0.987823 | 0.983113 |

**Table 5** The DM test results

| Models | MLP (1) | ANFIS (1) | RF(1) | SVR (1) | MARS (1) | LASSO (1) | Proposed (1) |
|---|---|---|---|---|---|---|---|
| *Static forecasting* | | | | | | | |
| MLP (2) | – | | | | | | |
| ANFIS (2) | 0.217# | – | | | | | |
| RF (2) | 0.209# | 0.216# | – | | | | |
| SVR (2) | 0.211# | 0.208# | 0.223# | – | | | |
| MARS (2) | − 3.8245*** | − 3.8663*** | − 3.8792*** | − 3.8804*** | – | | |
| LASSO (2) | − 3.8587*** | − 3.8935*** | − 3.8956*** | − 3.8967*** | 0.231# | – | |
| Pro-posed (2) | 4.8429*** | 4.8378*** | 4.8249*** | 4.8227*** | 5.2461*** | 5.2583*** | – |
| *Dynamic forecasting* | | | | | | | |
| MLP (2) | – | | | | | | |
| ANFIS (2) | 0.198# | – | | | | | |
| RF (2) | 0.237# | 0.207# | – | | | | |
| SVR (2) | 0.215# | 0.224# | 0.229# | – | | | |
| MARS (2) | − 3.8369*** | − 3.8681*** | − 3.8783*** | − 3.8804*** | – | | |
| LASSO (2) | − 3.8548*** | − 3.8890*** | − 3.8988*** | − 3.8967*** | 0.231# | – | |
| Pro-posed (2) | 4.8461*** | 4.8451*** | 4.8316*** | 4.8289 *** | 5.2461*** | 5.2583*** | – |

***Significant at 1% level, # Not significant

model. A negatively significant test statistic value implies the statistical superiority of the first model over the second model. Finally, an insignificant test statistic value implies that the performance of the models is significant different. A positive significant test statistic associated with the proposed model against all other competing models signifies the proposed model's statistical superiority over the others. From the significance levels, we conclude that all other models outperform MARS and LASSO. There is no significant difference between the predictive performance of MLP, ANFIS, RF, and SVR. Thus, this test offers the statistical validation of the efficacy and effectiveness of the proposed predictive model. The proposed framework is statistically superior to all the six competing models for forecasting one-step ahead and multi-step ahead Bitcoin prices.

## 5.2 Performance evaluation based on the surrogate time series

We evaluate the efficacy of the surrogate time series to validate and justify the proposed framework's deployment. The present research utilizes *'tseriesEntropy'* package of *'R'* platform for simulating the surrogate series using Sieve Bootsrap (Buhlmann, 1997). The autoregressive (AR) model resamples the residuals with the replacement for enabling simulation. Using the *'surrogate.AR'* function, we generate five such series from the original Bitcoin price series as depicted in Fig. 8. The figure shows that the generated five series are comparatively more volatile than the original one.

As the cryptocurrency market is highly uncertain and chaotic due to various factors, it becomes necessary to examine the proposed predictive model's performance. Tables 6 and 7 show the outcome of the proposed framework's predictive performance and competitive models on static and dynamic forecasting environments on surrogate datasets.

The performance indicator values imply that the predictive performance is more than satisfactory. Also, the performance is not as good as that of the original Bitcoin price prediction summarized in Sect. 5.1. The validation process applies to a more random and volatile series than the original time series. So, this is an expected outcome. The findings suggest that the proposed predictive framework can predict the future values of cryptocurrencies that are comparatively less stable than Bitcoin. Table 8 reports the results of the DM test on the static and dynamic forecasting setups.

Table 8 shows that the predictive performance is like that of the original time series. So, the proposed predictive framework is beneficial in an increasingly volatile and random environment. The overall findings of the predictive modeling bring out several interesting results. Both one-step ahead and multi-step ahead forecasting assessments indicate the proposed model's efficiency for estimating Bitcoin prices. Static one-step ahead forecasting performance is marginally better than its counterparts. It is comparatively easier to predict the future figures in the short duration with better accuracy than in the long term. One-step ahead forecasting estimates the closing price of Bitcoin on the next day. It will be useful for daily trading exercise. In multi-step ahead dynamic forecasting, we estimate the Bitcoin price for almost one year. The overall forecast accuracy would naturally be lower in the dynamic case. However, the values of performance indicators suggest that the obtained predictions are very accurate. This will help in long-term portfolio realignment, hedging, etc.

Kristjanpoller et al. (2020) and Zhang et al. (2020) study the dynamic interplay of Bitcoin with other assets. Paule-Vianez et al. (2020) report the safe heaven properties of Bitcoin. Kristoufek et al. (2018) evaluate the inefficiency of the Bitcoin market. However, these works barely throw any light on developing implementable practical
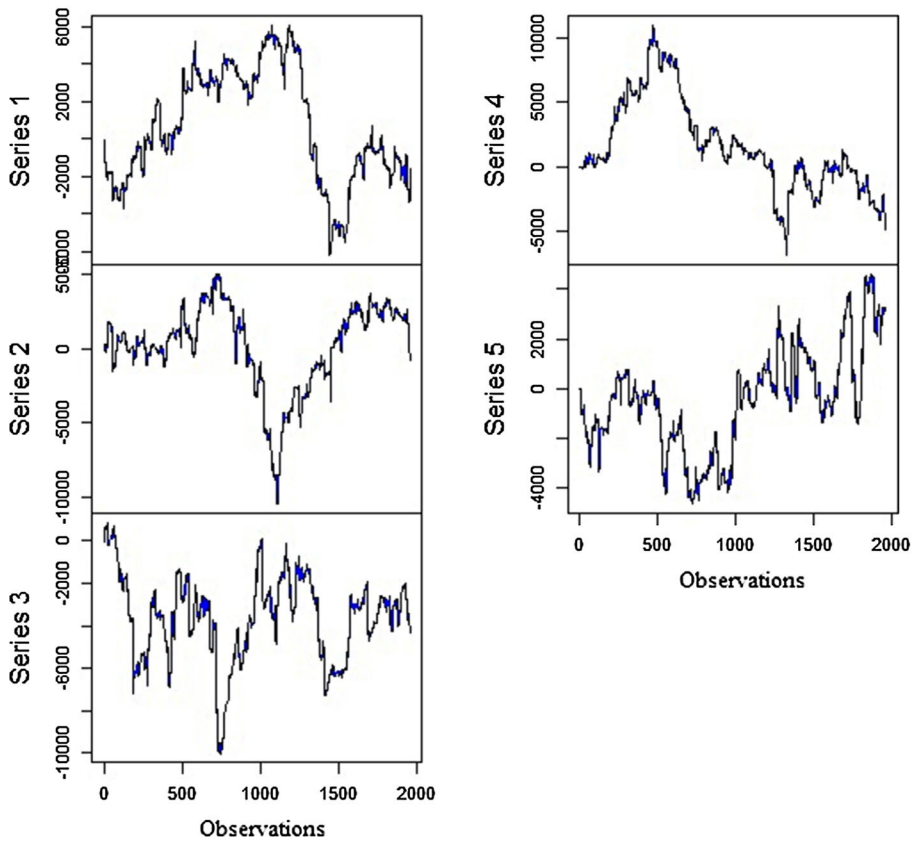
**Fig. 8** Simulated surrogate series

predictive architectures capable of predicting short and long-run movements. Interestingly, the literature manifests the multifractal nature of interplay (Kristjanpoller et al. 2020) and sensitivity to volatility contagion (Zhang et al. 2020) that triggers doubt on the predictability of the Bitcoin price with a high degree of precision in the long-range scale. The proposed approach shows that we can predict the Bitcoin price with utmost precision using lagged information of past prices in static and dynamic environments accounting for short and long duration forecasting.

Thus, despite the uncertain nature of Bitcoin prices (Paule Vianez et al. 2020), the proposed approach can model its evolutionary patterns in an efficacious way. Kristoufek et al. (2018) find the inefficiency structure of Bitcoin in the short run. However, the outcomes of the proposed predictive exercises justified that the inefficiency prevailed for long run time horizons. Predictive accuracy of surrogate series rationalizes proposed architecture's effectiveness in modeling Bitcoin price subject to extreme external shocks. As a future endeavor and in connection with the existing literature, the framework can examine Bitcoin price movements during unprecedented events like Covid-19 pandemic with a lagged based architecture.

**Table 6** Performance on surrogate series in static forecasting

| Series | IA | TI | DA | NSE |
|---|---|---|---|---|
| *Series one* | | | | |
| Proposed | 0.97985 | 0.035044 | 0.96535 | 0.95878 |
| MLP | 0.97082 | 0.061623 | 0.95237 | 0.95532 |
| ANFIS | 0.97166 | 0.062382 | 0.94961 | 0.95549 |
| RF | 0.97231 | 0.059831 | 0.95837 | 0.95731 |
| SVR | 0.97175 | 0.063208 | 0.95318 | 0.96108 |
| MARS | 0.96898 | 0.066013 | 0.94227 | 0.95116 |
| LASSO | 0.96926 | 0.065744 | 0.94359 | 0.95184 |
| *Series two* | | | | |
| Proposed | 0.96389 | 0.037921 | 0.93982 | 0.94337 |
| MLP | 0.95443 | 0.064318 | 0.92096 | 0.93910 |
| ANFIS | 0.95418 | 0.063248 | 0.92262 | 0.93886 |
| RF | 0.95562 | 0.062925 | 0.92126 | 0.93954 |
| SVR | 0.95380 | 0.062475 | 0.92278 | 0.93902 |
| MARS | 0.95085 | 0.065149 | 0.91867 | 0.93626 |
| LASSO | 0.95177 | 0.065022 | 0.91920 | 0.93581 |
| *Series three* | | | | |
| Proposed | 0.94214 | 0.040185 | 0.91887 | 0.91684 |
| MLP | 0.92838 | 0.077623 | 0.90472 | 0.91123 |
| ANFIS | 0.92769 | 0.078014 | 0.90526 | 0.91158 |
| RF | 0.92857 | 0.076352 | 0.90516 | 0.91209 |
| SVR | 0.92711 | 0.077484 | 0.90569 | 0.91202 |
| MARS | 0.92430 | 0.080169 | 0.90291 | 0.90943 |
| LASSO | 0.92461 | 0.080074 | 0.90327 | 0.90959 |
| *Series four* | | | | |
| Proposed | 0.95638 | 0.038772 | 0.90697 | 0.93058 |
| MLP | 0.94344 | 0.069585 | 0.88701 | 0.92155 |
| ANFIS | 0.94275 | 0.069192 | 0.89141 | 0.92119 |
| RF | 0.94393 | 0.068713 | 0.89176 | 0.92208 |
| SVR | 0.94346 | 0.069626 | 0.88637 | 0.92089 |
| MARS | 0.94057 | 0.071233 | 0.88205 | 0.91865 |
| LASSO | 0.94096 | 0.070876 | 0.88289 | 0.91887 |
| *Series five* | | | | |
| Proposed | 0.94196 | 0.039561 | 0.90875 | 0.92658 |
| MLP | 0.92768 | 0.072736 | 0.89912 | 0.91771 |
| ANFIS | 0.92689 | 0.072584 | 0.89873 | 0.91733 |
| RF | 0.92830 | 0.072371 | 0.89966 | 0.91808 |
| SVR | 0.92789 | 0.072905 | 0.89832 | 0.91756 |
| MARS | 0.92494 | 0.073688 | 0.89513 | 0.91412 |
| LASSO | 0.92463 | 0.073641 | 0.89545 | 0.91397 |

**Table 7** Performance on surrogate series in dynamic forecasting

| Series | IA | TI | DA | NSE |
|---|---|---|---|---|
| *Series one* | | | | |
| Proposed | 0.96944 | 0.035759 | 0.96251 | 0.95129 |
| MLP | 0.95893 | 0.062069 | 0.94965 | 0.93986 |
| ANFIS | 0.95814 | 0.061940 | 0.94910 | 0.93967 |
| RF | 0.95917 | 0.060238 | 0.95103 | 0.94223 |
| SVR | 0.95785 | 0.062124 | 0.94815 | 0.94136 |
| MARS | 0.95109 | 0.063496 | 0.94223 | 0.93254 |
| LASSO | 0.95074 | 0.063455 | 0.94178 | 0.93297 |
| *Series two* | | | | |
| Proposed | 0.95031 | 0.038045 | 0.93724 | 0.94106 |
| MLP | 0.96389 | 0.066703 | 0.92761 | 0.92996 |
| ANFIS | 0.94214 | 0.066816 | 0.92691 | 0.92961 |
| RF | 0.95638 | 0.066234 | 0.92778 | 0.93207 |
| SVR | 0.94196 | 0.066589 | 0.92765 | 0.93128 |
| MARS | 0.94196 | 0.068117 | 0.92542 | 0.92766 |
| LASSO | 0.94196 | 0.068077 | 0.92580 | 0.92797 |
| *Series three* | | | | |
| Proposed | 0.92970 | 0.041296 | 0.90855 | 0.91522 |
| MLP | 0.91821 | 0.080927 | 0.90875 | 0.92658 |
| ANFIS | 0.91845 | 0.080918 | 0.90875 | 0.92658 |
| RF | 0.91990 | 0.080335 | 0.90875 | 0.92658 |
| SVR | 0.91873 | 0.080891 | 0.90875 | 0.92658 |
| MARS | 0.91606 | 0.082216 | 0.90875 | 0.92658 |
| LASSO | 0.91575 | 0.081995 | 0.90875 | 0.92658 |
| *Series four* | | | | |
| Proposed | 0.94373 | 0.038837 | 0.91037 | 0.92916 |
| MLP | 0.93236 | 0.070862 | 0.90082 | 0.92018 |
| ANFIS | 0.93277 | 0.070814 | 0.90055 | 0.92033 |
| RF | 0.93344 | 0.070633 | 0.90104 | 0.92105 |
| SVR | 0.93258 | 0.070705 | 0.90029 | 0.92046 |
| MARS | 0.92850 | 0.071276 | 0.89565 | 0.91914 |
| LASSO | 0.92911 | 0.070975 | 0.89533 | 0.91827 |
| *Series five* | | | | |
| Proposed | 0.93106 | 0.039840 | 0.91673 | 0.92347 |
| MLP | 0.92135 | 0.073084 | 0.90347 | 0.92441 |
| ANFIS | 0.92119 | 0.073077 | 0.90304 | 0.92429 |
| RF | 0.92202 | 0.072918 | 0.90395 | 0.92467 |
| SVR | 0.92168 | 0.073109 | 0.90374 | 0.92541 |
| MARS | 0.91825 | 0.073863 | 0.89950 | 0.92118 |
| LASSO | 0.91849 | 0.073839 | 0.89943 | 0.92093 |

**Table 8** The DM test of predictive performance on the surrogate series

| Models | MLP (1) | ANFIS (1) | RF(1) | SVR (1) | MARS (1) | LASSO (1) | Proposed (1) |
|---|---|---|---|---|---|---|---|
| *Static forecasting* | | | | | | | |
| MLP (2) | – | | | | | | |
| ANFIS (2) | 0.244# | – | | | | | |
| RF (2) | 0.189# | 0.226# | – | | | | |
| SVR (2) | 0.234# | 0.239# | 0.213# | – | | | |
| MARS (2) | − 3.7683*** | − 3.7580*** | − 3.8698*** | − 3.8753*** | – | | |
| LASSO (2) | − 3.8491*** | − 3.8881*** | − 3.8479*** | − 3.8811*** | 0.254# | – | |
| Proposed (2) | 5.3628*** | 5.2869*** | 5.2263*** | 5.3259*** | 5.6780*** | 5.7569*** | – |
| *Dynamic forecasting* | | | | | | | |
| MLP (2) | – | | | | | | |
| ANFIS (2) | 0.214# | – | | | | | |
| RF (2) | 0.225# | 0.185# | – | | | | |
| SVR (2) | 0.241# | 0.228# | 0.242# | – | | | |
| MARS (2) | − 3.8316*** | − 3.8579*** | − 3.8856*** | − 3.88923*** | – | | |
| LASSO (2) | − 3.8641*** | − 3.8958*** | − 3.8911*** | − 3.9104*** | 0.239# | – | |
| Proposed (2) | 5.6673*** | 5.6389*** | 5.5427*** | 5.6160*** | 5.8906*** | 6.1425*** | – |

***Significant at 1% level, # Not significant

# 6 Conclusions

This research proposes a novel DE-based regression framework for forecasting one day ahead price of Bitcoin. It contributes to the forecasting literature by combining a metaheuristic search technique with regression models in a granular setup. The utilization of the DE for optimizing the learning algorithms' performance assists in augmenting the forecast quality. The proposed approach emerges as statistically superior to all other competing models on static and dynamic forecasting situations. LASSO and MARS produce comparatively less accurate forecasts. There is no statistical evidence to discriminate RF, MLP, ANFIS, and SVR's predictive performance. The performance of static one-step ahead is comparatively better than the multi-step ahead dynamic forecasting. The proposed model yields accurate forecasts on the surrogate time series and indicates its efficiency in recognizing intricate patterns. It is useful for estimating future figures of highly volatile cryptocurrencies and other chaotic financial assets.

The diversification benefits of Bitcoin are well known. It has the capability of providing investment security in challenging times. These properties of Bitcoin can help in the precise estimation of its movements in short and long-run horizons. The proposed framework is highly efficient in accomplishing this task. As a result, it will help investors in risk mitigation by estimating the future prices of Bitcoin. The efficacy of the proposed framework on the surrogate series demonstrates the predictive framework's capability to secure noteworthy returns in the crisis period.

This research's scope is limited to examining the predictability of the cryptocurrency having the largest market share. The proposed framework can assess the predictability and efficacy of other cryptocurrencies to infer the market's behavioral characteristics at a holistic level. This work provides the opportunity to offer new forecasting frameworks by integrating appropriate alternate components. We may consider a new forecasting framework by combining the ensemble empirical mode decomposition in place of MODWT. Alternate

regression approaches like regression neural network and quantile regression neural network may also work in this setting. Alternate metaheuristics search techniques like PSO, artificial bee colony, and brainstorm optimization may replace the DE. A comparative performance assessment of such predictive frameworks for forecasting highly volatile and chaotic time series may be an interesting research topic.

# References

Al-Yahaee, K. H., Mensi, W., & Yoon, S. M. (2018). Efficiency, multifractality, and the long-memory property of the Bitcoin market: A comparative analysis with stock, currency, and gold markets. *Finance Research Letters, 27,* 228–234.

Atsalakis, G. S., Atsalaki, I. G., Pasiouras, F., & Zopounidis, C. (2019). Bitcoin price forecasting with neuro-fuzzy techniques. *European Journal of Operations Research, 276,* 770–780.

Avdoulas, C., Bekiros, S., & Boubaker, S. (2018). Evolutionary-based return forecasting with nonlinear STAR models: evidence from the Eurozone peripheral stock markets. *Annals of Operations Research, 262,* 307–333.

Bariviera, A. F. (2017). The inefficiency of Bitcoin revisited: A dynamic approach. *Economics Letters, 150,* 6–9.

Bou-Hamad, I., & Jamali, I. (2020). Forecasting financial time-series using data mining models: A simulation study. *Research in International Business and Finance,.* https://doi.org/10.1016/j.ribaf.2019.101072

Boukhatem, J., Ftiti, Z., & Sahut, J. M. (2020). Bond market and macroeconomic stability in East Asia: A nonlinear causality analysis. *Annals of Operations Research.* https://doi.org/10.1007/s10479-020-03519-6

Bouri, E., Gupta, R., Lau, C. K. M., Roubaud, D., & Wang, S. (2018). Bitcoin and global financial stress: A copula-based approach to dependence and causality in the quantiles. *The Quarterly Review of Economics and Finance, 69,* 297–307.

Bouoiyour, J., Selmi, R., & Wohar, M. E. (2019). Bitcoin: competitor or complement to gold? *Economics Bulletin, 39,* 186–191.

Buhlmann, P. (1997). Sieve bootstrap for time series. *Bernoulli, 3,* 123–148.

Chaim, P., & Laurini, M. P. (2018). Is Bitcoin a bubble? *Physica A: Statistical Mechanics and its Applications, 517,* 222–232.

Chen, Y., & Hao, Y. (2017). A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction. *Expert Systems with Applications, 80,* 340–355.

Cornish, C. R., Bretherton, C. S., & Percival, D. B. (2006). Maximal overlap wavelet statistical analysis with application to atmospheric turbulence. *Boundary-Layer Meteorology, 119,* 339–374.

Cretarola, A., & Figà-Talamanca, G. (2019). Detecting bubbles in Bitcoin price dynamics via market exuberance. *Annals of Operations Research.* https://doi.org/10.1007/s10479-019-03321-z

Das, D., Bhowmik, P., & Jana, R. K. (2018). A multiscale analysis of stock return co-movements and spillovers: Evidence from Pacific developed markets. *Physica A: Statistical Mechanics and its Applications, 502,* 379–393.

Dastgir, S., Demir, E., Downing, G., Gozgor, G., & Lau, C. K. M. (2019). The causal relationship between Bitcoin attention and Bitcoin returns: Evidence from the Copula-based Granger causality test. *Finance Research Letters, 28,* 160–164.

Demir, E., Gozgor, G., Lau, C. K. M., & Vigne, S. A. (2018). Does economic policy uncertainty predict the Bitcoin returns? An empirical investigation. *Finance Research Letters, 26,* 145–149.

Filho, A. C. D. S., Maganini, N. D., & Almeida, E. F. D. (2018). Multifractal analysis of Bitcoin market. *Physica A: Statistical Mechanics and its Applications, 512,* 954–867.

Gencay, R., Selcuk, F., & Whitcher, B. (2002). *An introduction to wavelets and other filtering methods in finance and economics.* Academic Press.

Ghosh, I., Sanyal, M. K., & Jana, R. K. (2017). Fractal inspection and machine learning-based predictive modelling framework for financial markets. *Arabian Journal for Science and Engineering, 43,* 4237–4287.

Ghosh, I., Sanyal, M. K. and Jana, R. K. (2017a). Analysis of causal interactions and predictive modelling of financial markets using econometric methods, maximal overlap discrete wavelet transformation and machine learning: A study in Asian context. In: Shankar B., Ghosh K., Mandal D., Ray S., Zhang D., Pal S. (eds) Pattern Recognition and Machine Intelligence. PReMI 2017. Lecture Notes in Computer Science, vol 10597. Springer, Cham.

Ghosh, I., Jana, R. K., & Sanyal, M. K. (2019). Analysis of causal interactions and predictive modelling of financial markets using econometric methods, maximal overlap discrete wavelet transformation and machine learning. *Applied Soft Computing, 82,* 105553.

Ghosh, I,Sanyal, M. K. & Jana, R. K. (2020). Co-movement and dynamic correlation of financial and energy markets: An integrated framework of nonlinear dynamics, wavelet analysis and DCC-GARCH. Computational Economics. https://doi.org/10.1007/s10614-019-09965-0.

Guerard, J. B., Xu, G., & Markowitz, H. (2020). A further analysis of robust regression modeling and data mining corrections testing in global stocks. *Annals of Operations Research.* https://doi.org/10.1007/s10479-020-03521-y

Horra, L. P., & d. l., Fuente, G. d. l. and Perote, K. . (2019). The drivers of Bitcoin demand: A short and long-run analysis. *International Review of Financial Analysis, 62,* 21–34.

Jana, R. K., Tiwari, A. K., & Hammoudeh, S. (2019). The inefficiency of Litecoin: A dynamic analysis. *Journal of Quantitative Economics, 17*(2), 447–457.

Jang, H., & Lee, J. (2017). An empirical study on modeling and prediction of bitcoin prices with Bayesian neural networks based on blockchain information. *IEEE Access, 6,* 5427–5437.

Kao, L. J., Chiu, C. C., Lu, C. J., & Chang, C. H. (2013). A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting. *Decision Support Systems, 54,* 1228–1244.

Koutmos, D. (2019). Market risk and Bitcoin returns. *Annals of Operations Research, 294,* 453–477.

Kristoufek, L. (2018). On Bitcoin markets (in)efficiency and its evolution. *Physica A: Statistical Mechanics and its Applications, 503,* 257–262.

Kristjanpoller, W., Bouri, E., & Takaishi, T. (2020). Cryptocurrencies and equity funds: Evidence from an asymmetric multifractal analysis. *Physica A: Statistical Mechanics and its Applications, 545,* 123711.

Kumar, P. D., & Ravi, V. (2017). Forecasting financial time series volatility using particle swarm optimization trained quantile regression neural network. Applied Soft Computing, 35–52.

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

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

Matta, M., Lunesu, I., & Marchesi, M. (2016). Is Bitcoin's market predictable? Analysis of web search and social media. In: Fred A., Dietz J., Aveiro D., Liu K., Filipe J. (eds) Knowledge Discovery, Knowledge Engineering and Knowledge Management. IC3K 2015. Communications in Computer and Information Science, vol. 631. Springer, Cham.

McNally, S., Roche, J., & Caton, S. (2018). Predicting the price of bitcoin using machine learning. In *2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)* (pp. 339–343). IEEE.

Nadarajah, S., & Chu, J. (2017). On the inefficiency of Bitcoin. *Economics Letters, 150,* 6–9.

Nakano, M., Takahashi, A., & Takahashi, S. (2018). Bitcoin technical trading with artificial neural network. *Physica A: Statistical Mechanics and its Applications, 510,* 587–609.

Oztekin, A., Kizilaslan, R., Freund, S., & Iseri, A. (2016). A data analytic approach to forecasting daily stock returns in an emerging market. *European Journal of Operational Research, 253,* 697–710.

Panagiotidis, T., Stengos, T., & Vravosinos, O. (2018). On the determinants of bitcoin returns: A LASSO approach. *Finance Research Letters, 27,* 235–240.

Paule-Vianez, P.-R., & C. and Gomez-Martinez, R. . (2020). Economic policy uncertainty and Bitcoin. Is Bitcoin a safe-haven asset? European Journal of Management and Business Economics. *European Journal of Management and Business Economics, 29,* 347–363.

Poyser, O. (2019). Exploring the dynamics of Bitcoin's price: a Bayesian structural time series approach. *Eurasian Economic Review, 9,* 29–60.

Shen, W., Guo, X., Wu, C., & Wu, D. (2011). Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm. *Knowledge-Based Systems, 24,* 378–385.

Storn, R., & Price, K. (1997). Differential evolution: A simple and efficient adaptive scheme for global optimization over continuous spaces. *Journal of Global Optimization, 11,* 341–359.

Takaishi, T. (2018). Statistical properties and multifractality of Bitcoin. *Physica A: Statistical Mechanics and its Applications, 506,* 507–519.

Tiwari, A. K., Jana, R. K., Das, D., & Roubaud, D. (2018). Informational efficiency of Bitcoin: An extension. *Economics Letters, 163,* 106–109.

Urquhart, A. (2016). The inefficiency of Bitcoin. *Economics Letters, 148,* 80–82.

Vapnik, V. (1995). *The nature of statistical learning theory* (2nd ed.). Springer.

Wang, J., Hou, R., Wang, C., & Shrn, L. (2016). Improved v -Support vector regression model based on variable selection and brainstorm optimization for stock price forecasting. *Applied Soft Computing, 49,* 164–178.

Wu, C. H., Lu, C. C., Ma, Y. F., & Lu, R. S. (2018). A new forecasting framework for bitcoin price with LSTM. In *2018* IEEE International Conference on Data Mining Workshops (ICDMW) (pp. 168–175), IEEE.

Wang, G. J., Xie, C., Wen, D., & Zhao, L. (2019). When Bitcoin meets economic policy uncertainty (EPU): Measuring risk spillover effect from EPU to Bitcoin. *Finance Research Letters,.* https://doi.org/10.1016/j.frl.2018.12.028

Zhang, Y. J., Bouri, E., Gupta, R., & Ma, S. J. (2020). Risk spillover between Bitcoin and conventional financial markets: An expectile-based approach. *The North American Journal of Economics and Finance*. https://doi.org/10.1016/j.najef.2020.101296

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.