

Anticipating cryptocurrency prices using machine learning

Laura Alessandretti^a, Abeer ElBahrawy^b, Luca Maria Aiello^c, and
Andrea Baronchelli^{b,d,*}

^aTechnical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

^bCity, University of London, Department of Mathematics, London EC1V 0HB, UK

^cNokia Bell Labs, Cambridge CB3 0FA, UK

^dUCL Centre for Blockchain Technologies, University College London, UK

*Corresponding author: Andrea.Baronchelli.1@city.ac.uk

March 6, 2019

Abstract

Machine learning and AI-assisted trading have attracted growing interest for the past few years. Here, we use this approach to test the hypothesis that the inefficiency of the cryptocurrency market can be exploited to generate abnormal profits. We analyse daily data for 1,681 cryptocurrencies for the period between Nov. 2015 and Apr. 2018. We show that simple trading strategies assisted by state-of-the-art machine learning algorithms outperform standard benchmarks. Our results show that non-trivial, but ultimately simple, algorithmic mechanisms can help anticipate the short-term evolution of the cryptocurrency market.

Introduction

The popularity of cryptocurrencies has skyrocketed in 2017 due to several consecutive months of super-exponential growth of their market capitalisation [1], which peaked at more than \$800 billions in Jan. 2018. Today, there are more than 1,500 actively traded cryptocurrencies. Between 2.9 and 5.8 millions of private as well as institutional investors are in the different transaction networks, according to a recent survey [2], and access to the market has become easier over time. Major cryptocurrencies can be bought using fiat currency in a number of online exchanges (e.g., Binance [3], Upbit [4], Kraken [5], etc) and then be used in their turn to buy less popular cryptocurrencies. The volume of daily exchanges is currently superior to \$15 billions. Since 2017, over 170 hedge funds specialised in cryptocurrencies have emerged and bitcoin futures have been launched to address institutional demand for trading and hedging Bitcoin [6].

The market is diverse and provides investors with many different products. Just to mention a few, Bitcoin was expressly designed as a medium of exchange [7, 8]; Dash offers improved services on top of Bitcoin's feature set, including instantaneous and private transactions [9]; Ethereum is a public, blockchain-based distributed computing platform featuring smart contract (scripting) functionality, and

Ether is a cryptocurrency whose blockchain is generated by the Ethereum platform [10]; Ripple is a real-time gross settlement system (RTGS), currency exchange and remittance network Ripple [11], and IOTA is focused on providing secure communications and payments between agents on the Internet of Things [12].

The emergence of a self-organised market of virtual currencies and/or assets whose value is generated primarily by social consensus [13] has naturally attracted interest from the scientific community [8,14–30]. Recent results have shown that the long-term properties of the cryptocurrency market have remained stable between 2013 and 2017 and are compatible with a scenario in which investors simply sample the market and allocate their money according to the cryptocurrency’s market shares [1]. While this is true on average, various studies have focused on the analysis and forecasting of price fluctuations, using mostly traditional approaches for financial markets analysis and prediction [31–35].

The success of machine learning techniques for stock markets prediction [36–42], suggests that these methods could be effective also in predicting cryptocurrencies prices. However, the application of machine learning algorithms to the cryptocurrency market has been limited so far to the analysis of Bitcoin prices, using random forests [43], Bayesian neural network [44], long short-term memory neural network [45] and other algorithms [32,46]. These studies were able to anticipate, to different degrees, the price fluctuations of Bitcoin, and revealed that best results were achieved by neural network based algorithms. Deep reinforcement learning was showed to beat the uniform buy and hold strategy [47] in predicting the prices of 12 cryptocurrencies over one year period [48].

Other attempts to use machine learning to predict the prices of cryptocurrencies other than Bitcoin come from non-academic sources [49–54]. Most of these analyses focused on a limited number of currencies and did not provide benchmark comparisons for their results.

Here, we test the performance of three models in predicting daily cryptocurrency price for 1,681 currencies. Two of the models are based on gradient boosting decision trees [55] and one is based on long short-term memory (LSTM) recurrent neural networks [56]. In all cases, we build investment portfolios based on the predictions and we compare their performance in terms of return on investment. We find that all of the three models perform better than a baseline ‘simple moving average’ model [57–60] where a currency’s price is predicted as the average price across the preceding days, and that the method based on long short-term memory recurrent neural networks systematically yields the best return on investment.

The article is structured as follows: In section Materials and Methods we describe the data (see Data description and pre-processing), the metrics characterizing cryptocurrencies that are used along the paper (see Metrics), the forecasting algorithms (see Forecasting algorithms), and the evaluation metrics (see Evaluation). In section Results, we present and compare the results obtained with the three forecasting algorithms and the baseline method. In section Conclusion, we conclude and discuss results.

Materials and Methods

Data description and pre-processing

Cryptocurrency data was extracted from the website Coin Market Cap [61], collecting daily data from 300 exchange markets platforms starting in the period between November 11, 2015 and April 24, 2018. The dataset contains the daily price in U.S. dollars, the market capitalisation and the trading volume of 1,681 cryptocurrencies, where the market capitalization is the product between price and circulating supply, and the volume is the number of coins exchanged in a day. The daily price is computed as the

volume weighted average of all prices reported at each market. Fig. 1 shows the number of cryptocurrencies with trading volume larger than V_{min} over time, for different values of V_{min} . In the following sections, we consider that only cryptocurrencies with daily trading volume higher than 10^5 USD can be traded at any given day.

The website lists cryptocurrencies traded on public exchange markets that have existed for more than 30 days and for which an API as well as a public URL showing the total mined supply are available. Information on the market capitalization of cryptocurrencies that are not traded in the 6 hours preceding the weekly release of data is not included on the website. Cryptocurrencies inactive for 7 days are not included in the list released. These measures imply that some cryptocurrencies can disappear from the list to reappear later on. In this case, we consider the price to be the same as before disappearing. However, this choice does not affect results since only in 28 cases the currency has volume higher than 10^5 USD right before disappearing (note that there are 124,328 entries in the dataset with volume larger than 10^5 USD).

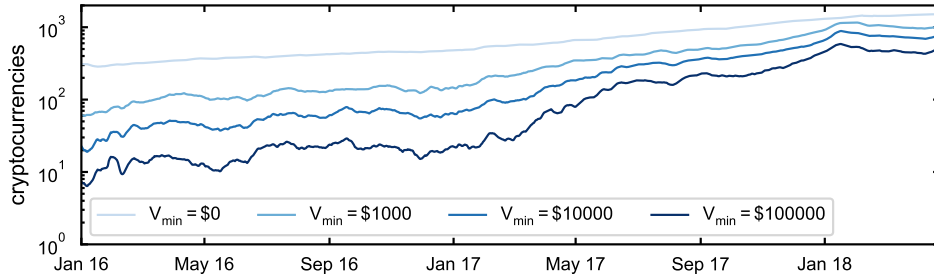


Figure 1: **Number of cryptocurrencies.** The cryptocurrencies with volume higher than V_{min} as a function of time, for different values of V_{min} . For visualization purposes, curves are averaged over a rolling window of 10 days.

Metrics

Cryptocurrencies are characterised over time by several metrics, namely

- Price, The exchange rate, determined by supply and demand dynamics.
- Market capitalization, The product of the circulating supply and the price.
- Market share, The market capitalization of a currency normalized by the total market capitalization.
- Rank, The *rank* of currency based on its market capitalization.
- Volume, Coins traded in the last 24 hours.
- Age, Lifetime of the currency in days.

The profitability of a currency c over time can be quantified through the *return on investment* (ROI), measuring the return of an investment made at day t_i relative to the cost [62]. The index i rolls across days and it is included between 0 and 844, with t_0 = January 1, 2016, and t_{844} = April 24, 2018. Since

we are interested in the short-term performance, we consider the return on investment after 1 day defined as

$$ROI(c, t_i) = \frac{price(c, t_i) - price(c, t_i - 1)}{price(c, t_i - 1)}. \quad (1)$$

In Fig. 2, we show the evolution of the ROI over time for Bitcoin (orange line) and on average for currencies whose volume is larger than $V_{min} = 10^5$ USD at $t_i - 1$ (blue line). In both case, the average return on investment over the period considered is larger than 0, reflecting the overall growth of the market.

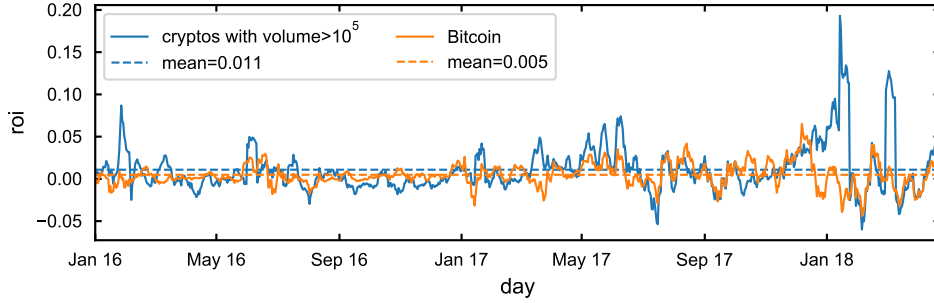


Figure 2: **Return on investment over time.** The daily return on investment for Bitcoin (orange line) and the average for currencies with volume larger than $V_{min} = 10^5$ USD (blue line). Their average value across time (dashed lines) is larger than 0. For visualization purposes, curves are averaged over a rolling window of 10 days.

Forecasting algorithms

We test and compare three supervised methods for short-term price forecasting. The first two methods rely on XGboost [63], an open-source scalable machine learning system for tree boosting used in a number of winning Kaggle solutions (17/29 in 2015) [64]. The third method is based on the long short-term memory (LSTM) algorithm for recurrent neural networks [56] that have demonstrated to achieve state-of-the-art results in time-series forecasting [65].

Method 1: The first method considers one single regression model to describe the change in price of all currencies (see Fig. 3). The model is an ensemble of regression trees built by the XGboost algorithm. The features of the model are characteristics of a currency between time $t_j - w$ and $t_j - 1$ and the target is the ROI of the currency at time t_j , where w is a parameter to be determined. The characteristics considered for each currency are: price, market capitalization, market share, rank, volume and ROI (see equation Eq. (1)). The features for the regression are built across the window between $t_j - w$ and $t_j - 1$ included (see Fig. 3). Specifically, we consider the average, the standard deviation, the median, the last value and the trend (e.g. the difference between last and first value) of the properties listed above. In the training phase, we include all currencies with volume larger than 10^5 USD, and t_j between $t_i - W_{training}$ and t_i . In general, larger training windows do not necessarily lead to better results (see results section), because the market evolves across time. In the prediction phase, we test on the set of existing currencies at day t_i . This procedure is repeated for values of t_i included between January 1, 2016 and April 24, 2018.

Method 2: Also the second method relies on XGboost, but now the algorithm is used to build a different regression model for each currency c_i (see Fig. 4). The features of the model for currency c_i are

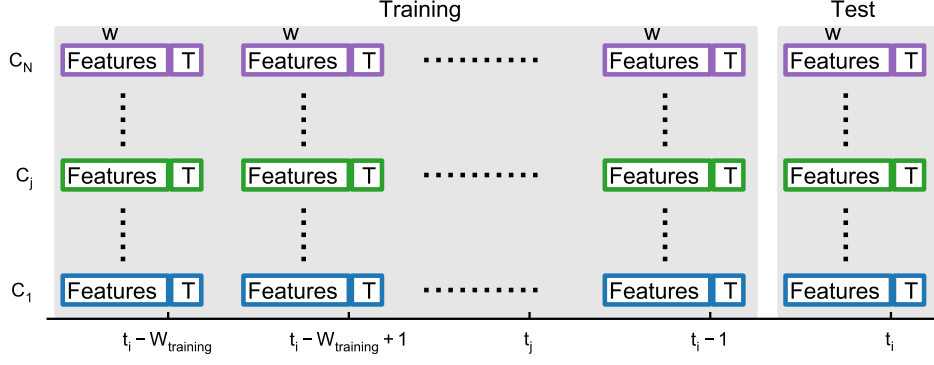


Figure 3: **Schematic description of Method 1.** The training set is composed of features and target (T) pairs, where features are various characteristics of a currency c_i , computed across the w days preceding time t_j and the target T is the price of c_i at t_j . The features-target pairs are computed for all currencies c_i and all values of t_j included between $t_i - W_{\text{training}}$ and $t_i - 1$. The test set includes features-target pairs for all currencies with trading volume larger than 10^5 USD at t_i , where the target is the price at time t_i and features are computed in the w days preceding t_i .

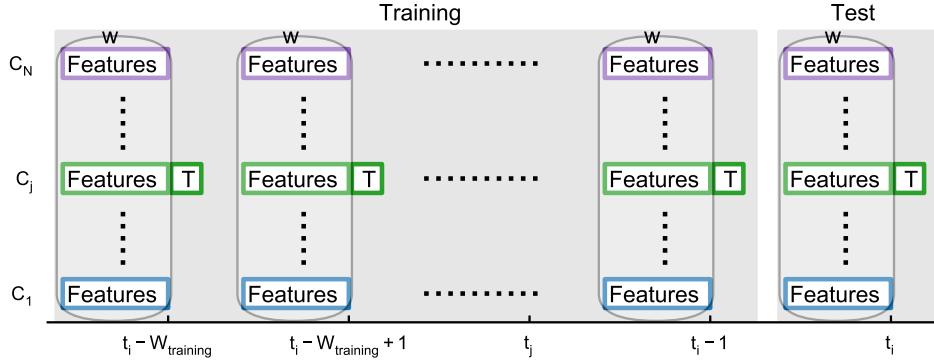


Figure 4: **Schematic description of Method 2.** The training set is composed of features and target (T) pairs, where features are various characteristics of all currencies, computed across the w days preceding time t_j and the target T is the price of c_i at t_j . The features-target pairs include a single currency c_i , for all values of t_j included between $t_i - W_{\text{training}}$ and $t_i - 1$. The test set contains a single features-target pair: the characteristics of all currencies, computed across the w days preceding time t_i and the price of c_i at t_i .

the characteristics of all the currencies in the dataset between $t_j - w$ and $t_j - 1$ included and the target is the ROI of c_i at day t_j (i.e., now the algorithm learns to predict the price of the currency i based on the features of all the currencies in the system between $t_j - w$ and $t_j - 1$). The features of the model are the same used in Method 1 (e.g. the average, standard, deviation, median, last value, difference between last and first value of the following quantities: price, market capitalisation, market share, rank, volume and ROI) across a window of length w . The model for currency c_i is trained with pairs features target between times $t_i - W_{\text{training}}$ and $t_i - 1$. The prediction set include only one pair: the features (computed between $t_i - w$ and $t_i - 1$) and the target (computed at t_i) of currency c_i .

Method 3: The third method is based on Long Short Term Memory networks, a special kind of Recurrent Neural Networks, capable of learning long-term dependencies. As for Method 2, we build a different model for each currency. Each model predicts the ROI of a given currency at day t_i based on the values of the ROI of the same currency between days $t_i - w$ and $t_i - 1$ included.

Baseline method: As baseline method, we adopt the simple moving average strategy (SMA) widely tested and used as a null model in stock market prediction [57–60]. It estimates the price of a currency at day t_i as the average price of the same currency between $t_i - w$ and $t_i - 1$ included.

Evaluation

We compare the performance of various investment portfolios built based on the algorithms predictions. The investment portfolio is built at time $t_i - 1$ by equally splitting an initial capital among the top n currencies predicted with positive return. Hence, the total return at time t_i is:

$$R(t_i) = \frac{1}{n} \sum_{c=1}^n ROI(c, t_i).$$

The portfolios performance is evaluated by computing the Sharpe ratio and the geometric mean return. The Sharpe ratio is defined as:

$$S(t_i) = \frac{\bar{R}}{s_R},$$

where $\bar{R}$ is the average return on investment obtained between times 0 and t_i , and s_R , the corresponding standard deviation.

The geometric mean return is defined as:

$$G(t_i) = \sqrt[t_i]{\prod_{j=1}^{t_i} 1 + R(t_j)},$$

where t_i corresponds to the total number of days considered. The cumulative return obtained at t_i after investing and selling on the following day for the whole period is defined as $G(t_i)^2$.

The number of currencies n to include in a portfolio is chosen at t_i by optimising either the geometric mean $G(t_i - 1)$ (geometric mean optimisation) or the Sharpe ratio $S(t_i - 1)$ (Sharpe ratio optimisation) over the possible choices of n . The same approach is used to choose the parameters of Method 1 (w and $W_{training}$), Method 2 (w and $W_{training}$), and the baseline method (w).

Results

We predict the price of the currencies at day t_i , for all t_i included between Jan, 1st 2016 and Apr 24th, 2018. The analysis considers all currencies whose age is larger than 50 days since their first appearance and whose volume is larger than \$100000. To discount for the effect of the overall market movement (i.e., market growth, for most of the considered period), we consider cryptocurrencies prices expressed in Bitcoin. This implies that Bitcoin is excluded from our analysis.

Parameter setting

First, we choose the parameters for each method. Parameters include the number of currencies n to include the portfolio as well as the parameters specific to each method. In most cases, at each day t_i we choose the parameters that maximise either the geometric mean $G(t_i - 1)$ (geometric mean optimisation) or the Sharpe ratio $S(t_i - 1)$ (Sharpe ratio optimisation) computed between times 0 and t_i .

Baseline strategy: We test the performance of the baseline strategy for choices of window $w \geq 2$ (the minimal requirement for the ROI to be different from 0) and $w < 30$. We find that the value of w maximising the geometric mean return (see Appendix Fig. A1-A) and the Sharpe Ratio (see Appendix Fig. A1-D) fluctuates especially before November 2016 and has median value 4 in both cases. The number of currencies included in the portfolio oscillates between 1 and 11 with median at 3, both for the Sharpe Ratio (see Fig. A1-B) and the geometric mean return (see Fig. A1-E) optimisation.

Method 1: We explore values of the window w in $\{3, 5, 7, 10\}$ days and the training period $W_{training}$ in $\{5, 10, 20\}$ days (see Appendix Fig. A2). We find that the median value of the selected window w across time is 7 for both the Sharpe ratio and the geometric mean optimisation. The median value of $W_{training}$ is 5 under geometric mean optimisation and 10 under Sharpe ratio optimisation. The number of currencies included in the portfolio oscillates between 1 and 43 with median at 15 for the Sharpe Ratio (see Appendix Fig. A2-A) and 9 for the geometric mean return (see Appendix Fig. A2-C) optimisations.

Method 2: We explore values of the window w in $\{3, 5, 7, 10\}$ days and the training period $W_{training}$ in $\{5, 10, 20\}$ days (see Appendix Fig. A3). The median value of the selected window w across time is 3 for both the Sharpe ratio and the geometric mean optimisation. The median value of $W_{training}$ is 10 under geometric mean and Sharpe ratio optimisation. The number of currencies included has median at 17 for the Sharpe Ratio and 7 for the geometric mean optimisation (see Appendix Fig. A3-A and C).

Method 3: The LSTM has three parameters: The number of epochs, or complete passes through the dataset during the training phase; the number of neurons in the neural network, and the length of the window w . These parameters are chosen by optimising the price prediction of three currencies (Bitcoin, Ripple, and Ethereum) that have on average the largest market share across time (excluding Bitcoin Cash that is a fork of Bitcoin). Results (see SI, Fig. A4) reveal that, in the range of parameters explored, the best results are achieved for $w = 50$. Results are not particularly affected by the choice of the number of neurones nor the number of epochs. We choose 1 neuron and 1000 epochs since the larger these two parameters, the larger the computational time. The number of currencies to include in the portfolio is optimised over time by maximising the geometric mean return (see Appendix Fig. A5-A) and the Sharpe ratio (see Appendix Fig. A5-B). In both cases the median number of currencies included is 1.

Cumulative return

In Fig. 5, we show the cumulative return obtained using the 4 methods. The cumulative returns achieved on April, 24 under the Sharpe Ratio optimisation are ~ 65 BTC (Baseline), $\sim 1.1 \cdot 10^3$ BTC (Method 1), ~ 95 BTC (Method 2), $\sim 1.2 \cdot 10^9$ BTC (Method 3). Under geometric mean optimisation we obtain ~ 25 BTC (Baseline), $\sim 19 \cdot 10^3$ BTC (Method 1), ~ 1.25 BTC (Method 2), $\sim 3.6 \cdot 10^8$ BTC (Method 3). The cumulative returns obtained in USD are higher (see Fig. A9). This is expected, since the Bitcoin price has increased during the period considered. While some of these figures appear exaggerated, it is worth noticing that (i) we run a theoretical exercise assuming that the availability of Bitcoin is not limited and (ii) under this assumption the upper bound to our strategy, corresponding to investing every day in the most performing currency results in a total cumulative return of $6 \cdot 10^{123}$ BTC (see Appendix Fig. A6). We consider also the more realistic scenario of investors paying a transaction fee when selling and buying currencies (see Appendix Section A3). In most exchange markets, the fee is typically included between 0.1% and 0.5% of the traded amount [66]. For fees up to 0.2%, all the investment methods presented above lead, on average, to positive returns over the entire period (see Table A1). The best performing method, Method 3, achieves positive gains also when fees up to 1% are considered (see Table A1).

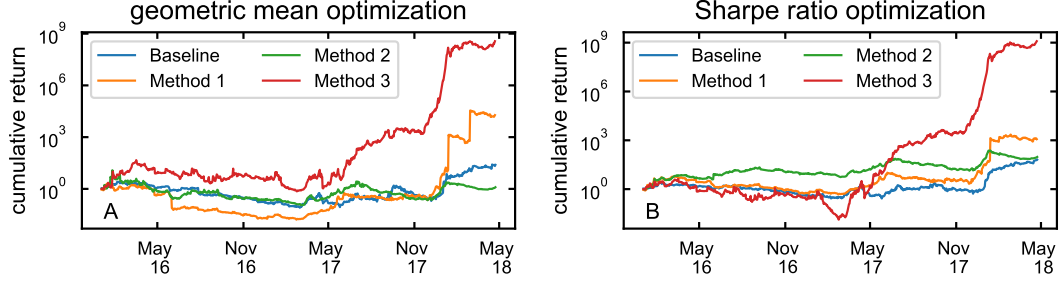


Figure 5: **Cumulative returns.** The cumulative returns obtained under the Sharpe Ratio optimisation (A) and the geometric mean optimisation (B) for the baseline (blue line), Method 1 (orange line), Method 2 (green line) and Method 3 (red line). Analyses are performed considering prices in BTC.

The cumulative return in Fig. 5 is obtained by investing between January 1st, 2016 and April 24th, 2018. We investigate the overall performance of the various methods by looking at the geometric mean return obtained in different periods (see Fig. 6). Results presented in Fig. 6 are obtained under Sharpe ratio optimisation for the baseline (Fig. 6-A), Method 1 (Fig. 6-B), Method 2 (Fig. 6-C), and Method 3 (Fig. 6-D). Note that, while in this case the investment can start after January 1st, 2016, we optimised the parameters by using data from that date on in all cases. Results are considerably better than those achieved using geometric mean return optimisation (see Appendix Fig. A10). Finally, we observe that better performance is achieved when the algorithms consider prices in Bitcoin rather than USD (see Table A2).

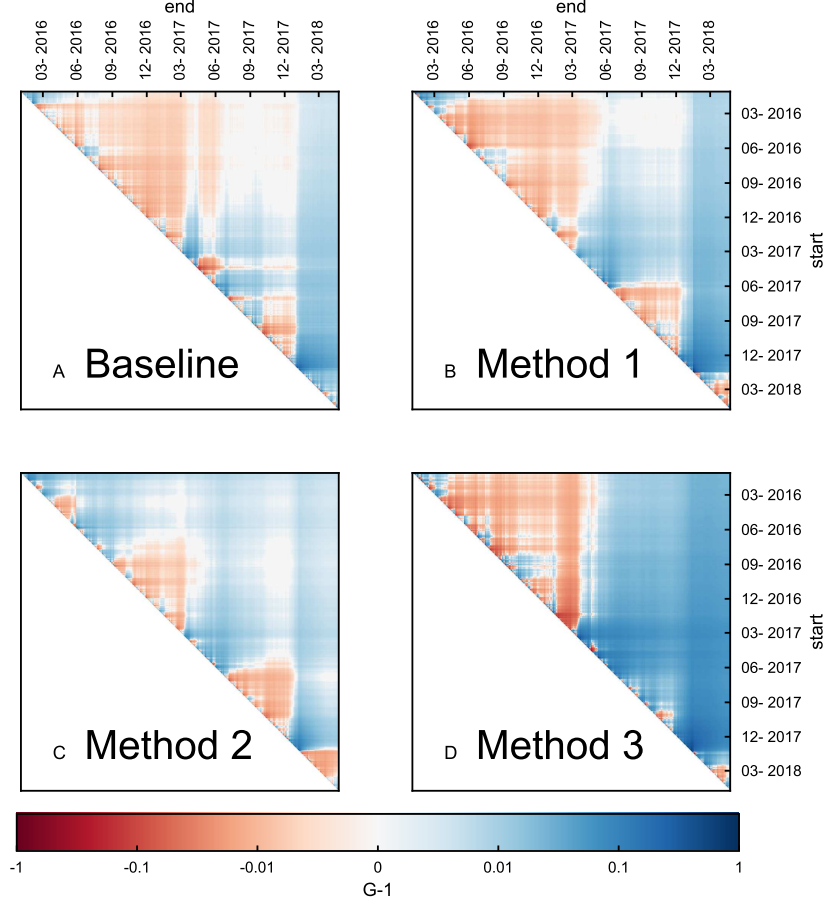


Figure 6: **Geometric mean return obtained within different periods of time.** The geometric mean return computed between time "start" and "end" using the Sharpe ratio optimisation for the baseline (A), Method 1 (B), Method 2 (C) and Method 3 (D). Note that, for visualization purposes, the figure shows the translated geometric mean return G-1. Shades of red refers to negative returns and shades of blue to positive ones (see colour bar).

Feature importance

In Fig. 7, we illustrate the relative importance of the various features in Method 1 and Method 2. For Method 1, we show the average feature importance; For Method 2, we show the average feature importances for two sample currencies: Ethereum and Ripple.

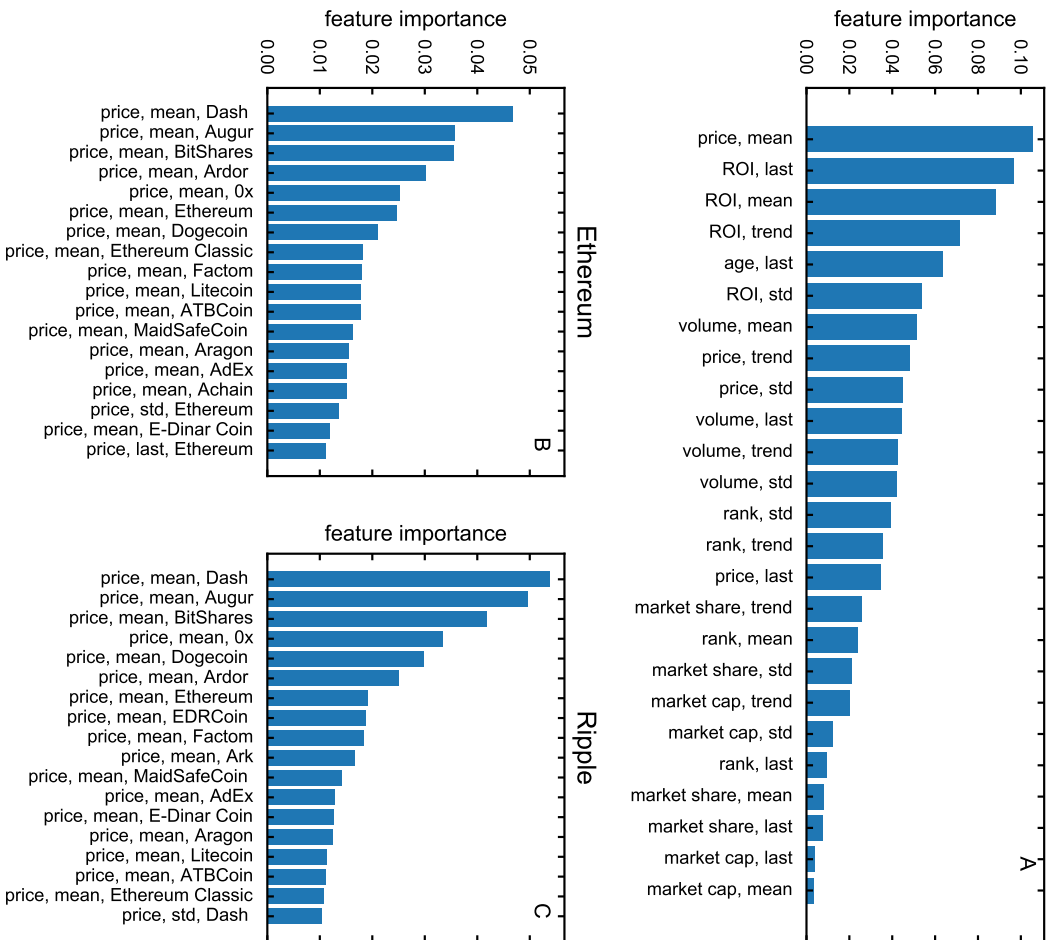


Figure 7: **Feature importance for Methods 1 and 2.** (A) The average importance of each feature for the XGBoost regression model of Method 1. Results are shown for $w = 7$ and $W_{training} = 10$. (B, C) Examples of average feature importance for the XGBoost regression model developed in Method 2. Results are shown for $w = 3$, $W_{training} = 10$, for Ethereum (B) and Ripple (C). For visualization purposes, we show only the top features.

Portfolio composition

The 10 most selected currencies under Sharpe Ratio optimisation are the following:

Baseline: Factom (91 days), E-Dinar Coin (89 days), Ripple (76 days), Ethereum (71 days), Steem (70 days), Lisk (70 days), MaidSafeCoin (69 days), Monero (58 days), BitShares (55 days), EDRCoin (52 days).

Method 1: Ethereum (154 days), Dash (128 days), Monero (111 days), Factom (104 days), Ripple (94 days), Litecoin (93 days), Dogecoin (92 days), Maid Safe Coin (86 days), BitShares (73 days), Tether (59 days)

Method 2: Ethereum (63 days), Monero (61 days), Factom (51 days), Ripple (42 days), Dash (40 days), Maid Safe Coin (40 days), Siacoin (30 days), NEM (26 days), NXT (26 days), Steem (23 days).

Method 3: Factom (48 days), Monero (46 days), Ethereum (39 days), Lisk (36 days), Maid Safe Coin

(32 days), E-Dinar Coin (32 days), BitShares (26 days), B3 Coin (26 days), Dash (25 days), Cryptonite (22 days).

Conclusion

We tested the performance of three forecasting models on daily cryptocurrency prices for 1,681 currencies. Two of them (Method 1 and Method 2) were based on gradient boosting decision trees and one is based on long short-term memory recurrent neural networks (Method 3). In Method 1, the same model was used to predict the return on investment of all currencies; in Method 2, we built a different model for each currency, that uses information on the behaviour of the whole market to make a prediction on that single currency; in Method 3, we used a different model for each currency, where the prediction is based on previous prices of the currency.

We built investment portfolios based on the predictions of the different method and compared their performance with that of a baseline represented by the well known simple moving average strategy. The parameters of each model were optimised for all but Method 3 on a daily basis, based on the outcome of each parameters choice in previous times. We used two evaluation metrics used for parameter optimisation: The geometric mean return and the Sharpe ratio. To discount the effect of the overall market growth, cryptocurrencies prices were expressed in Bitcoin. All strategies, produced profit (expressed in Bitcoin) over the entire considered period and for a large set of shorter trading periods (different combinations of start and end dates for the trading activity), also when transaction fees up to 0.2% are considered.

The three methods performed better than the baseline strategy when the investment strategy was ran over the whole period considered. The optimisation of parameters based on the Sharpe ratio achieved larger returns. Methods based on gradient boosting decision trees (Method 1 and 2) worked best when predictions were based on short-term windows of 5/10 days, suggesting they exploit well mostly short-term dependencies. Instead, LSTM recurrent neural networks worked best when predictions were based on ~ 50 days of data, since they are able to capture also long-term dependencies and are very stable against price volatility. They allowed to make profit also if transaction fees up to 1% are considered. Methods based on gradient boosting decision trees allow to better interpret results. We found that the prices and the returns of a currency in the last few days preceding the prediction were leading factors to anticipate its behaviour. Among the two methods based on random forests, the one considering a different model for each currency performed best (Method 2). Finally, it is worth noting that the three methods proposed perform better when predictions are based on prices in Bitcoin rather than prices in USD. This suggests that forecasting simultaneously the overall cryptocurrency market trend and the developments of individual currencies is more challenging than forecasting the latter alone.

It is important to stress that our study has limitations. First, we did not attempt to exploit the existence of different prices on different exchanges, the consideration of which could open the way to significantly higher returns on investment. Second, we ignored intra-day price fluctuations and considered an average daily price. Finally, and crucially, we run a theoretical test in which the available supply of Bitcoin is unlimited and none of our trades influence the market. Notwithstanding these simplifying assumptions, the methods we presented were systematically and consistently able to identify outperforming currencies. Extending the current analysis by considering these and other elements of the market is a direction for future work.

A different yet promising approach to the study cryptocurrencies consists in quantifying the impact

of public opinion, as measured through social media traces, on the market behaviour, in the same spirit in which this was done for the stock market [67]. While it was shown that social media traces can be also effective predictors of Bitcoin [68–74] and other currencies [75] price fluctuations, our knowledge of their effects on the whole cryptocurrency market remain limited and is an interesting direction for future work.

References

- [1] A. ElBahrawy, L. Alessandretti, A. Kandler, R. Pastor-Satorras, and A. Baronchelli, “Evolutionary dynamics of the cryptocurrency market,” *Royal Society Open Science*, vol. 4, no. 11, 2017. [Online]. Available: <http://rsos.royalsocietypublishing.org/content/4/11/170623>
- [2] G. Hileman and M. Rauchs, *Global Cryptocurrency Benchmarking Study*. Cambridge Centre for Alternative Finance, 2017.
- [3] Binance.com. (2017) Binance. [Online]. Available: <https://www.binance.com/>
- [4] Dunamu. (2017) Upbit. [Online]. Available: <https://upbit.com/>
- [5] P. Inc. (2012) Kraken. [Online]. Available: <https://www.kraken.com/>
- [6] S. Foley, J. Karlsen, and T. J. Putni, “Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies?” *SSRN*, p. <https://ssrn.com/abstract=3102645>, 2018.
- [7] S. Nakamoto, “Bitcoin: A peer-to-peer electronic cash system,” 2008.
- [8] R. Ali, J. Barrdear, R. Clews, and J. Southgate, “The economics of digital currencies,” 2014.
- [9] The Dash Network. (2018) Dash. [Online]. Available: <https://www.dash.org>
- [10] Ethereum Foundation (Stiftung Ethereum). (2018) Ethereum. [Online]. Available: <https://www.ethereum.org/>
- [11] Ripple. (2013) Ripple. [Online]. Available: <https://ripple.com/>
- [12] (2018) Iota. [Online]. Available: <https://iota.org/>
- [13] A. Baronchelli, “The emergence of consensus: a primer,” *Royal Society Open Science*, vol. 5, no. 2, 2018. [Online]. Available: <http://rsos.royalsocietypublishing.org/content/5/2/172189>
- [14] G. P. Dwyer, “The economics of bitcoin and similar private digital currencies,” *Journal of Financial Stability*, vol. 17, pp. 81–91, 2015.
- [15] R. Böhme, N. Christin, B. Edelman, and T. Moore, “Bitcoin: Economics, technology, and governance,” *Journal of Economic Perspectives*, vol. 29, no. 2, pp. 213–38, 2015.
- [16] M. J. Casey and P. Vigna, “Bitcoin and the digital-currency revolution,” *The Wall Street Journal*, vol. 23, 2015.
- [17] S. Trimborn and W. K. Härdle, “Crix an index for blockchain based currencies,” 2016.
- [18] M. Iwamura, Y. Kitamura, and T. Matsumoto, “Is bitcoin the only cryptocurrency in the town? economics of cryptocurrency and friedrich a. hayek,” 2014.

- [19] M. A. Cusumano, “The bitcoin ecosystem,” *Communications of the ACM*, vol. 57, no. 10, pp. 22–24, 2014.
- [20] K. Wu, S. Wheatley, and D. Sornette, “Classification of crypto-coins and tokens from the dynamics of their power law capitalisation distributions,” *arXiv preprint arXiv:1803.03088*, 2018.
- [21] A. Lamarche-Perrin, A. Orléan, and P. Jensen, “Coexistence of several currencies in presence of increasing returns to adoption,” *Physica A: Statistical Mechanics and its Applications*, 2018.
- [22] P. M. Krafft, N. Della Penna, and A. Pentland, “An experimental study of cryptocurrency market dynamics,” *arXiv preprint arXiv:1801.05831*, 2018.
- [23] A. Rogojanu, L. Badea *et al.*, “The issue of competing currencies. case study–bitcoin,” *Theoretical and Applied Economics*, vol. 21, no. 1, pp. 103–114, 2014.
- [24] L. H. White, “The market for cryptocurrencies,” *Cato J.*, vol. 35, p. 383, 2015.
- [25] P. Ceruleo, “Bitcoin: a rival to fiat money or a speculative financial asset?” 2014.
- [26] M. N. Sayed and N. A. Abbas, “Impact of crypto-currency on emerging market focus on gulf countries,” *Life Science Journal*, vol. 15, no. 1, 2018.
- [27] M. A. Javarone and C. S. Wright, “From bitcoin to bitcoin cash: a network analysis,” *arXiv preprint arXiv:1804.02350*, 2018.
- [28] Y. Sovbetov, “Factors influencing cryptocurrency prices: Evidence from bitcoin, ethereum, dash, litcoin, and monero,” 2018.
- [29] F. Parino, L. Gauvin, and M. G. Beiro, “Analysis of the bitcoin blockchain: Socio-economic factors behind the adoption,” *arXiv preprint arXiv:1804.07657*, 2018.
- [30] S. Begušić, Z. Kostanjčar, H. E. Stanley, and B. Podobnik, “Scaling properties of extreme price fluctuations in bitcoin markets,” *arXiv preprint arXiv:1803.08405*, 2018.
- [31] P. Ciaian, M. Rajcaniova, and d. Kancs, “The economics of bitcoin price formation,” *Applied Economics*, vol. 48, no. 19, pp. 1799–1815, 2016.
- [32] T. Guo and N. Antulov-Fantulin, “Predicting short-term bitcoin price fluctuations from buy and sell orders,” *arXiv preprint arXiv:1802.04065*, 2018.
- [33] G. Gajardo, W. D. Kristjanpoller, and M. Minutolo, “Does bitcoin exhibit the same asymmetric multifractal cross-correlations with crude oil, gold and djia as the euro, great british pound and yen?” *Chaos, Solitons & Fractals*, vol. 109, pp. 195–205, 2018.
- [34] N. Gandal and H. Halaburda, “Can we predict the winner in a market with network effects? competition in cryptocurrency market,” *Games*, vol. 7, no. 3, p. 16, 2016.
- [35] H. Elendner, S. Trimborn, B. Ong, T. M. Lee *et al.*, “The cross-section of crypto-currencies as financial assets: An overview,” Sonderforschungsbereich 649, Humboldt University, Berlin, Germany, Tech. Rep., 2016.
- [36] D. Enke and S. Thawornwong, “The use of data mining and neural networks for forecasting stock market returns,” *Expert Systems with applications*, vol. 29, no. 4, pp. 927–940, 2005.

- [37] W. Huang, Y. Nakamori, and S.-Y. Wang, "Forecasting stock market movement direction with support vector machine," *Computers & Operations Research*, vol. 32, no. 10, pp. 2513–2522, 2005.
- [38] P. Ou and H. Wang, "Prediction of stock market index movement by ten data mining techniques," *Modern Applied Science*, vol. 3, no. 12, p. 28, 2009.
- [39] M. Gavrilov, D. Anguelov, P. Indyk, and R. Motwani, "Mining the stock market (extended abstract): which measure is best?" in *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2000, pp. 487–496.
- [40] K. S. Kannan, P. S. Sekar, M. M. Sathik, and P. Arumugam, "Financial stock market forecast using data mining techniques," in *Proceedings of the International Multiconference of Engineers and computer scientists*, vol. 1, 2010, p. 4.
- [41] A. F. Sheta, S. E. M. Ahmed, and H. Faris, "A comparison between regression, artificial neural networks and support vector machines for predicting stock market index," *Soft Computing*, vol. 7, p. 8, 2015.
- [42] P.-C. Chang, C.-H. Liu, C.-Y. Fan, J.-L. Lin, and C.-M. Lai, "An ensemble of neural networks for stock trading decision making," *Emerging Intelligent Computing Technology and Applications. With Aspects of Artificial Intelligence*, pp. 1–10, 2009.
- [43] I. Madan, S. Saluja, and A. Zhao, "Automated bitcoin trading via machine learning algorithms," 2015.
- [44] H. Jang and J. Lee, "An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information," *IEEE Access*, vol. 6, pp. 5427–5437, 2018.
- [45] S. McNally, "Predicting the price of bitcoin using machine learning," Ph.D. dissertation, Dublin, National College of Ireland, 2016.
- [46] K. HEGAZY and S. MUMFORD, "Comparitive automated bitcoin trading strategies."
- [47] A. G. Shilling, "Market timing: Better than a buy-and-hold strategy," *Financial Analysts Journal*, vol. 48, no. 2, pp. 46–50, 1992.
- [48] Z. Jiang and J. Liang, "Cryptocurrency portfolio management with deep reinforcement learning," *arXiv preprint arXiv:1612.01277*, 2016.
- [49] "Traderobot package," <https://github.com/owocki/pytrader>, 2017, accessed: 5 December 2017.
- [50] "Cryptobot package," <https://github.com/AdeelMufti/CryptoBot>, 2017, accessed: 5 December 2017.
- [51] "Cryptocurrencytrader package," <https://github.com/llens/CryptoCurrencyTrader>, 2017, accessed: 5 December 2017.
- [52] "Btc trading package," <https://github.com/bukosabino/btctrading>, 2017, accessed: 5 December 2017.
- [53] "Bitpredict package," <https://github.com/cbyn/bitpredict>, 2017, accessed: 5 December 2017.
- [54] "Bitcoin bubble burst," <https://bitcoinbubbleburst.github.io/>, 2017, accessed: 5 December 2017.

- [55] J. H. Friedman, “Stochastic gradient boosting,” *Computational Statistics & Data Analysis*, vol. 38, no. 4, pp. 367–378, 2002.
- [56] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [57] W. Brock, J. Lakonishok, and B. LeBaron, “Simple technical trading rules and the stochastic properties of stock returns,” *The Journal of finance*, vol. 47, no. 5, pp. 1731–1764, 1992.
- [58] T. Kilgallen, “Testing the simple moving average across commodities, global stock indices, and currencies,” *The Journal of Wealth Management*, vol. 15, no. 1, p. 82, 2012.
- [59] B. LeBaron, “The stability of moving average technical trading rules on the dow jones index,” *Deriv. Use Trad. Regul.*, vol. 5, pp. 324–338, 2000.
- [60] C. A. Ellis and S. A. Parbery, “Is smarter better? a comparison of adaptive, and simple moving average trading strategies,” *Research in International Business and Finance*, vol. 19, no. 3, pp. 399–411, 2005.
- [61] “coinmarketcap.com,” <http://coinmarketcap.com/>, 2013, accessed: 30 May 2017.
- [62] G. T. Friedlob and F. J. Plewa Jr, *Understanding return on investment*. John Wiley & Sons, 1996.
- [63] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. ACM, 2016, pp. 785–794.
- [64] K. Inc. (2018) Kaggle. [Online]. Available: <https://www.kaggle.com/>
- [65] Z. C. Lipton, J. Berkowitz, and C. Elkan, “A critical review of recurrent neural networks for sequence learning,” *arXiv preprint arXiv:1506.00019*, 2015.
- [66] B. Wiki. (2017) Comparison of exchanges. [Online]. Available: https://en.bitcoin.it/wiki/Comparison_of_exchanges
- [67] H. S. Moat, C. Curme, A. Avakian, D. Y. Kenett, H. E. Stanley, and T. Preis, “Quantifying wikipedia usage patterns before stock market moves,” *Scientific reports*, vol. 3, p. 1801, 2013.
- [68] D. Kondor, I. Csabai, J. Szüle, M. Pósfai, and G. Vattay, “Inferring the interplay between network structure and market effects in bitcoin,” *New Journal of Physics*, vol. 16, no. 12, p. 125003, 2014.
- [69] L. Kristoufek, “What are the main drivers of the bitcoin price? evidence from wavelet coherence analysis,” *PloS one*, vol. 10, no. 4, p. e0123923, 2015.
- [70] D. Garcia and F. Schweitzer, “Social signals and algorithmic trading of bitcoin,” *Royal Society Open Science*, vol. 2, no. 9, 2015. [Online]. Available: <http://rsos.royalsocietypublishing.org/content/2/9/150288>
- [71] L. Kristoufek, “Bitcoin meets google trends and wikipedia: Quantifying the relationship between phenomena of the internet era,” *Scientific reports*, vol. 3, p. 3415, 2013.
- [72] A. Urquhart, “What causes the attention of bitcoin?” *Economics Letters*, vol. 166, pp. 40 – 44, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S016517651830065X>

- [73] D. Garcia, C. J. Tessone, P. Mavrodiev, and N. Perony, “The digital traces of bubbles: feedback cycles between socio-economic signals in the bitcoin economy,” *Journal of the Royal Society Interface*, vol. 11, no. 99, p. 20140623, 2014.
- [74] S. Wang and J.-P. Vergne, “Buzz factor or innovation potential: What explains cryptocurrencies returns?” *PloS one*, vol. 12, no. 1, p. e0169556, 2017.
- [75] T. R. Li, A. S. Chamrajnagar, X. R. Fong, N. R. Rizik, and F. Fu, “Sentiment-based prediction of alternative cryptocurrency price fluctuations using gradient boosting tree model,” *arXiv preprint arXiv:1805.00558*, 2018.

Appendix

A1 Parameter optimisation

In Fig. A1, we show the optimisation of the parameters w (A,C) and n (B,D) for the baseline strategy. In Fig. A2, we show the optimisation of the parameters w (A,D), $W_{training}$ (B,E), and n (C,F) for Method 1. In Fig. A3, we show the optimisation of the parameters w (A,D), $W_{training}$ (B,E), and n (C,F) for Method 2. In Fig. A4, we show the median squared error obtained under different training window choices (A), number of epochs (B) and number of neurons (C), for Ethereum, Bitcoin and Ripple. In Fig. A5, we show the optimisation of the parameter n (C,F) for Method 3.

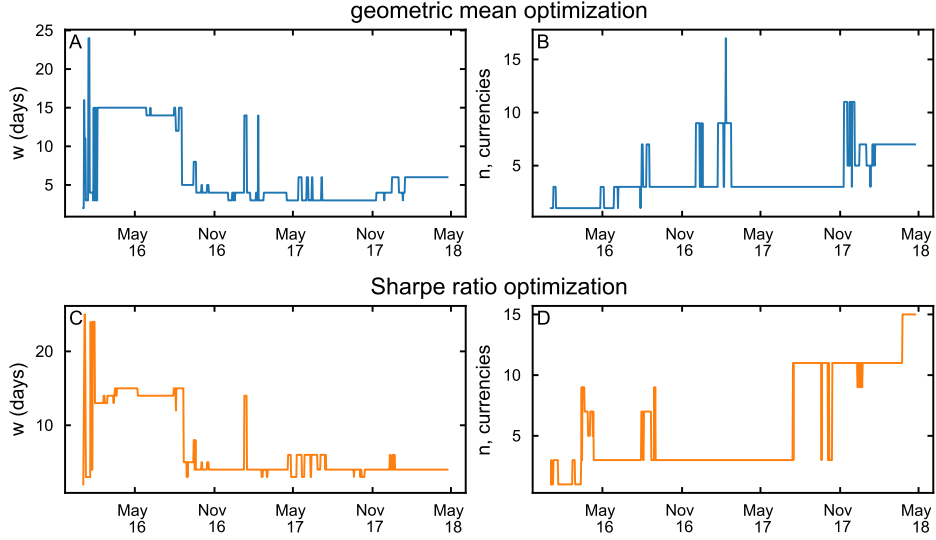


Figure A1: **Baseline strategy: parameters optimisation.** The sliding window w (A,C) and the number of currencies n (B,D) chosen over time under the geometric mean (A,B) and the Sharpe Ratio optimisation (C,D). Analyses are performed considering prices in BTC.

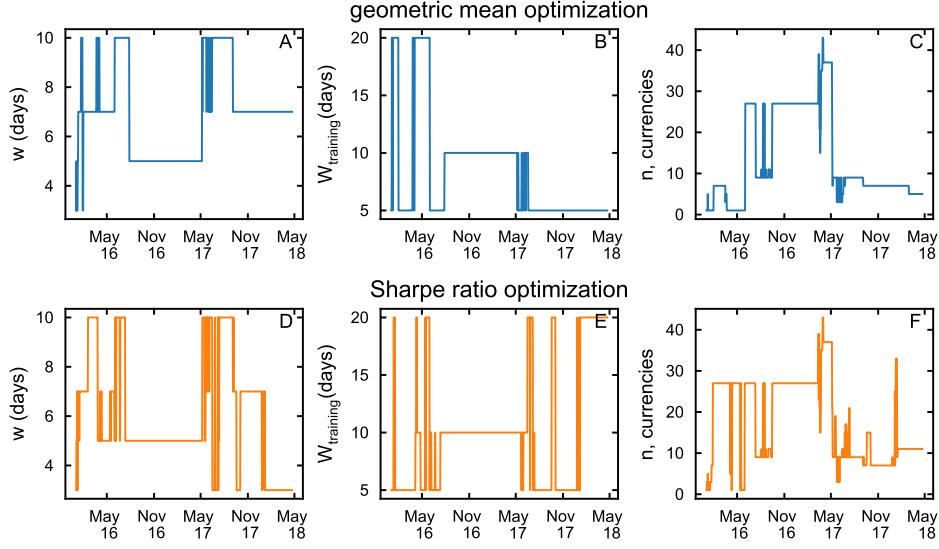


Figure A2: **Method 1: Parameters optimisation.** The sliding window w (A,D), the training window $W_{training}$ (B,E) and the number of currencies n (C,F) chosen over time under the geometric mean (A,B,C) and the Sharpe Ratio optimisation (D,E,F). Analyses are performed considering prices in BTC.

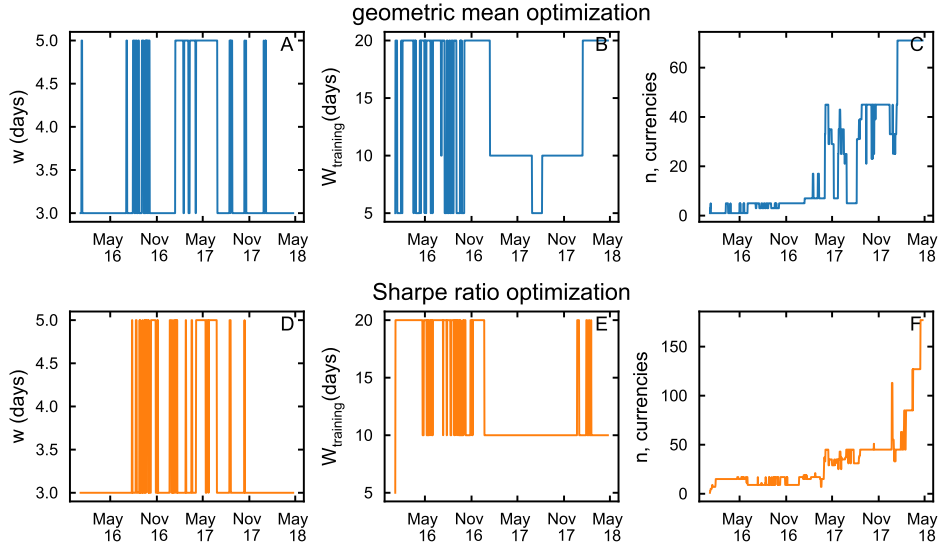


Figure A3: **Method 2: Parameters optimisation.** The sliding window w (A,D), the training window $W_{training}$ (B,E) and the number of currencies n (C,F) chosen over time under the geometric mean (A,B,C) and the Sharpe Ratio optimisation (D,E,F). Analyses are performed considering prices in BTC.

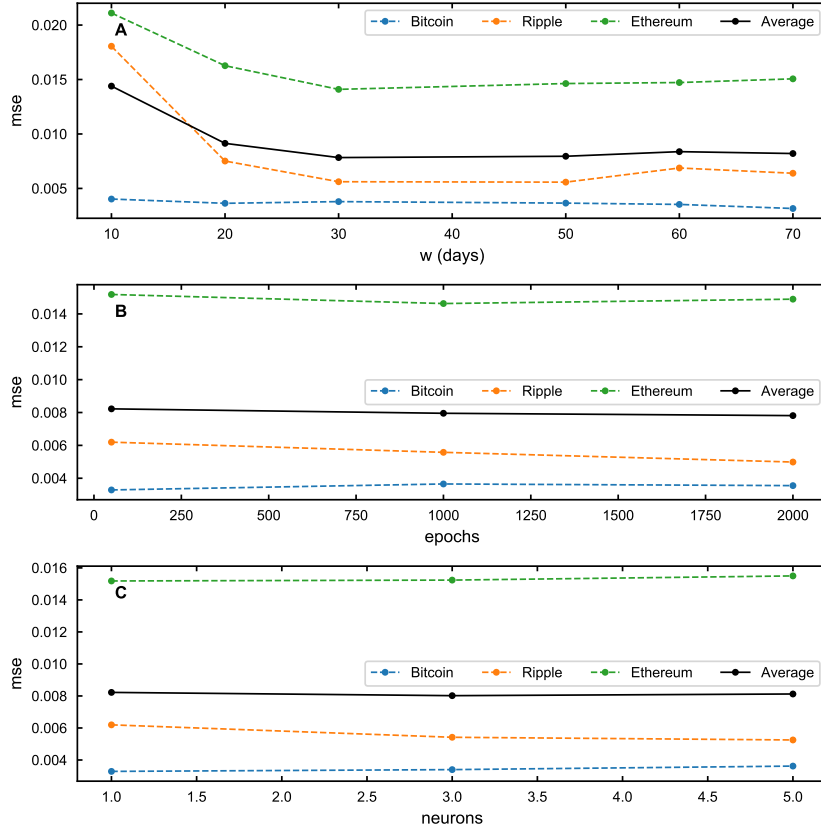


Figure A4: **Method 3: Parameters optimisation.** The median squared error of the ROI as a function of the window size (A), the number of epochs (B) and the number of neurons (C). Results are shown considering prices in Bitcoin.

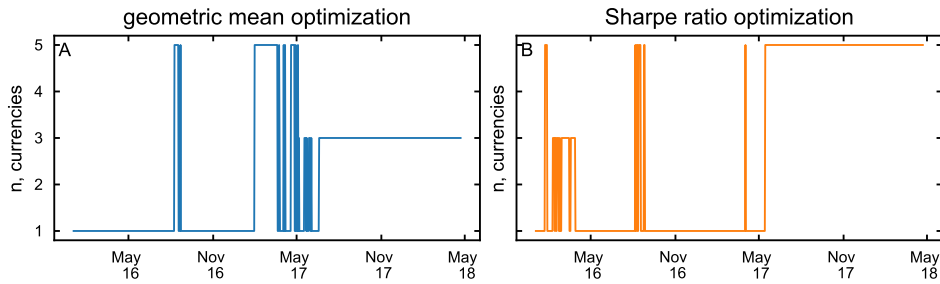


Figure A5: **Method 3: Parameters optimisation.** The number of currencies n chosen over time under the geometric mean (A) and the Sharpe Ratio optimisation (B). Analyses are performed considering prices in BTC.

A2 Return under full knowledge of the market evolution.

In Fig. A6, we show the cumulative return obtained by investing every day in the top currency, supposing one knows the prices of currencies on the following day.

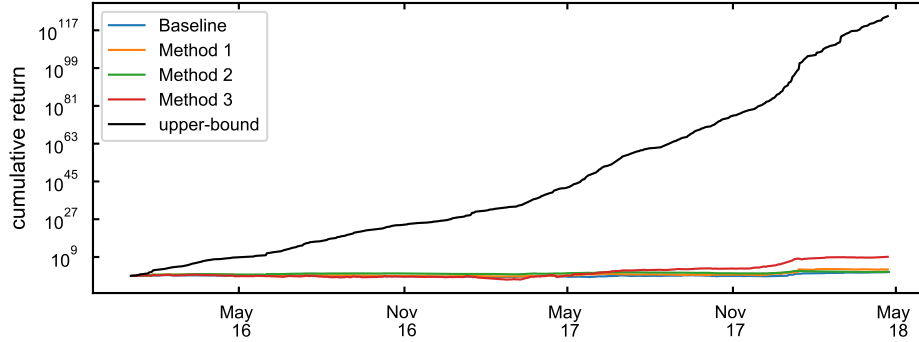


Figure A6: **Upper-bound for the cumulative return.** The cumulative return obtained by investing every day in the currency with highest return on the following day (black line). The cumulative return obtained with the baseline (blue line), Method 1 (orange line), Method 2 (green line), and Method 3 (red line). Results are shown in Bitcoin.

A3 Return obtained paying transaction fees.

In this section, we present the results obtained including transaction fees between 0.1% and 1% [66]. In general, one can not trade a given currency with any given other. Hence, we consider that each day we trade twice: We sell altcoins to buy Bitcoin, and we buy new altcoins using Bitcoin. The mean return obtained between Jan. 2016 and Apr. 2018 is larger than 1 for all methods, for fees up to 0.2% (see Table A1). In this period, Method 3 achieves positive returns for fees up to 1%. The returns obtained with a 0.1% (see Fig. A7) and 0.2% (see Fig. A8) fee during arbitrary periods confirm that, in general, one obtains positive gains with our methods if fees are small enough.

	no fee	0.1%	0.2%	0.3%	0.5%	1%
Baseline	1.005	1.003	1.001	0.999	0.995	0.985
Method 1	1.008	1.006	1.004	1.002	0.998	0.988
Method 2	1.005	1.003	1.001	0.999	0.995	0.985
Method 3	1.025	1.023	1.021	1.019	1.015	1.005

Table A1: **Daily geometric mean return for different transaction fees.** Results are obtained considering the period between Jan. 2016 and Apr. 2018.

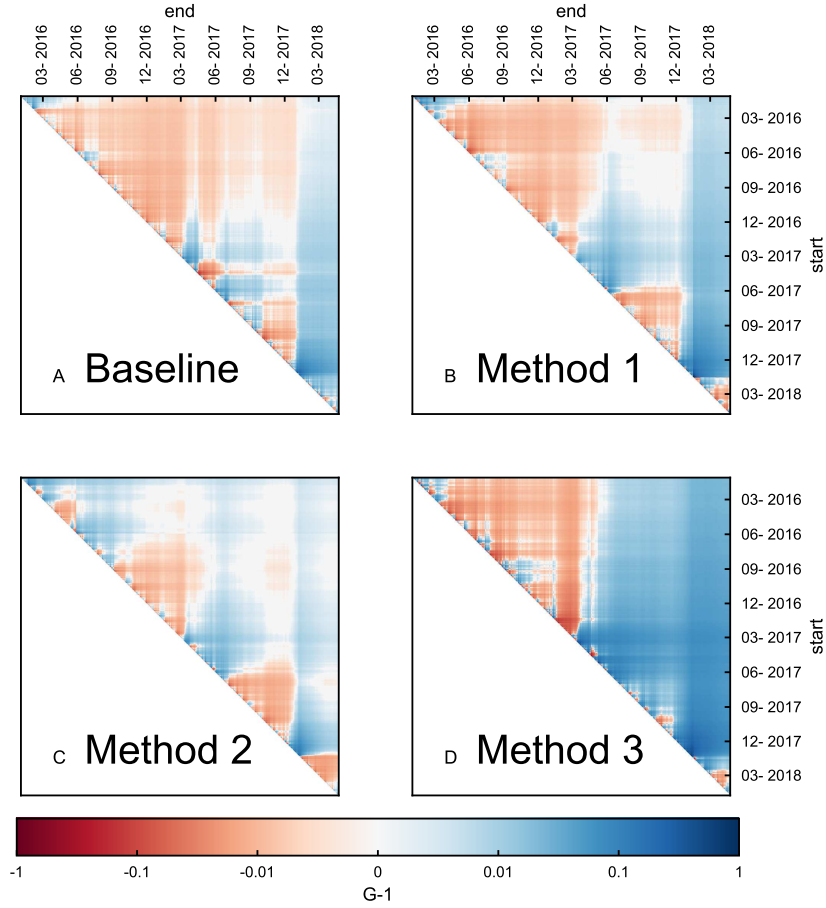


Figure A7: **Daily geometric mean return obtained under transaction fees of 0.1%.** The geometric mean return computed between time "start" and "end" using the Sharpe ratio optimisation for the baseline (A), Method 1 (B), Method 2 (C) and Method 3 (D). Note that, for visualization purposes, the figure shows the translated geometric mean return $G-1$. Shades of red refers to negative returns and shades of blue to positive ones (see colour bar).

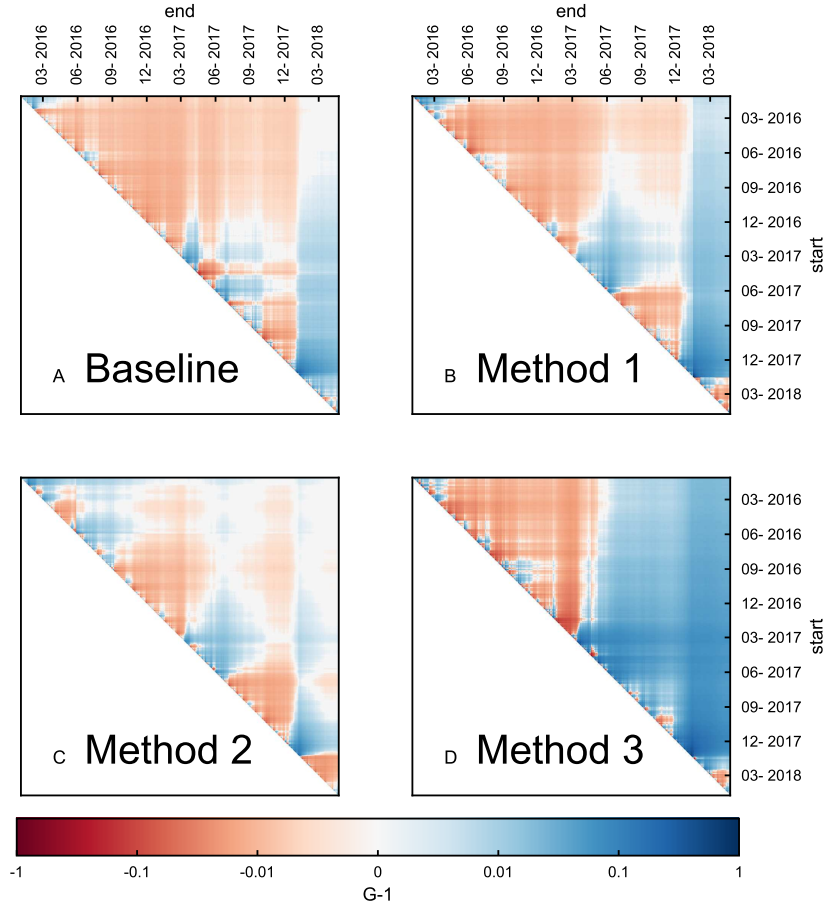


Figure A8: **Daily geometric mean return obtained under transaction fees of 0.2%.** The geometric mean return computed between time "start" and "end" using the Sharpe ratio optimisation for the baseline (A), Method 1 (B), Method 2 (C) and Method 3 (D). Note that, for visualization purposes, the figure shows the translated geometric mean return $G-1$. Shades of red refers to negative returns and shades of blue to positive ones (see colour bar).

A4 Results in USD

In this section, we show results obtained considering prices in USD. The price of Bitcoin in USD has considerably increased in the period considered. Hence, gains in USD (Fig. A9) are higher than those in Bitcoin (Fig. 5). Note that, in Fig. A9, we have made predictions and computed portfolios considering prices in Bitcoin. Then, gains have been converted to USD (without transaction fees). In Table A2, we show instead the gains obtained running predictions considering directly all prices in USD. We find that, in most cases, better results are obtained from prices in BTC.

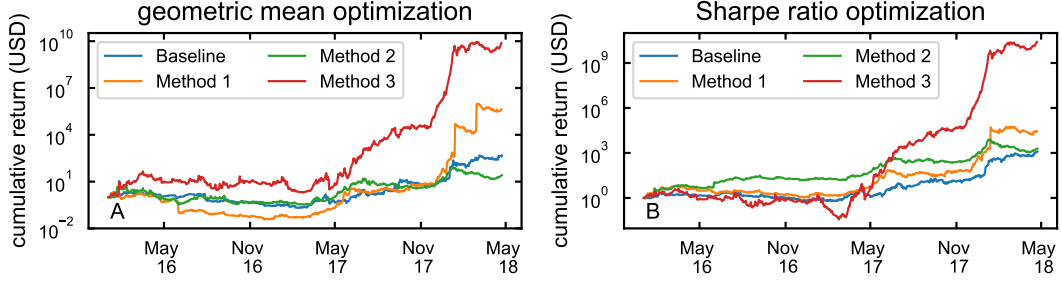


Figure A9: **Cumulative returns in USD.** The cumulative returns obtained under the Sharpe Ratio optimisation (A) and the geometric mean optimisation (B) for the baseline (blue line), Method 1 (orange line), Method 2 (green line) and Method 3 (red line). Analyses are performed considering prices in BTC.

	Geometric mean in USD (from BTC prices)	Geometric mean in USD (from USD prices)
Baseline	1.0086	1.0141
Method1	1.0121	1.0085
Method2	1.0091	1.0086
Method3	1.0289	1.0134

Table A2: **Geometric mean returns in USD.** Results are obtained for the various methods by running the algorithms considering prices in BTC (left column) and USD (right column).

A5 Geometric mean optimisation

In Fig. A10, we show the geometric mean return obtained by between two arbitrary points in time under geometric mean return optimisation for the baseline (Fig. A10-A), Method 1 (Fig. A10-B), Method 2 (Fig. A10-C), and Method 3 (Fig. A10-D).

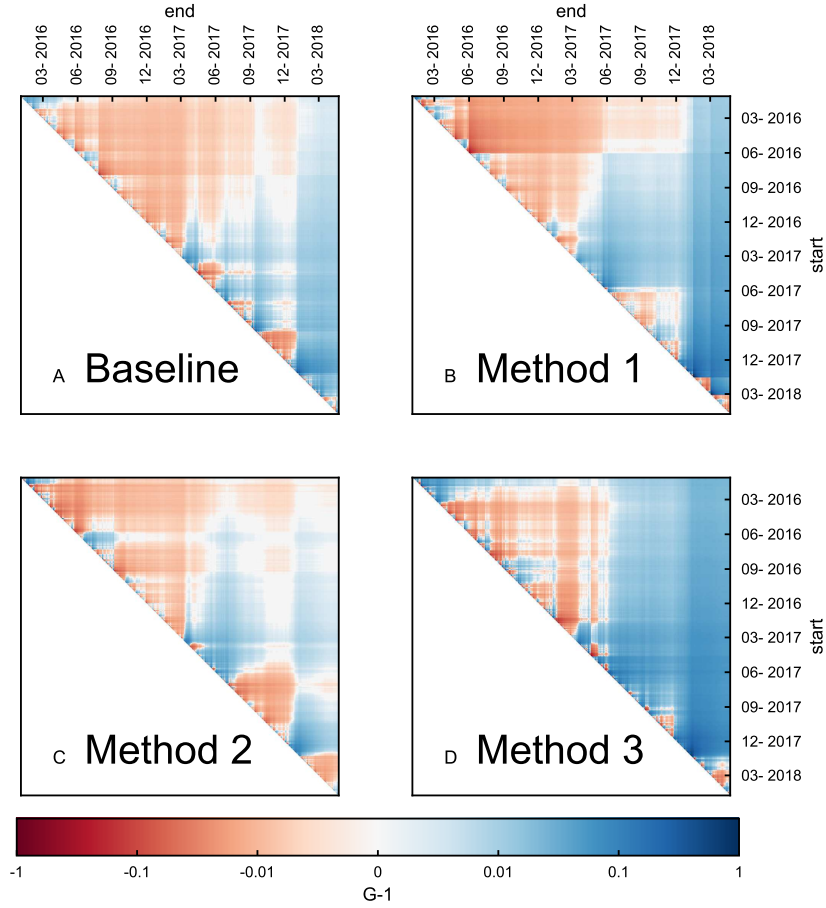


Figure A10: **Geometric mean return obtained within different periods of time.** The geometric mean return computed between time "start" and "end" using the Sharpe ratio optimisation for the baseline (A), Method 1 (B), Method 2 (C) and Method 3 (D). Note that, for visualization purposes, the figure shows the translated geometric mean return $G-1$. Shades of red refers to negative returns and shades of blue to positive ones (see colour bar).