Comparison of Deterministic, Stochastic, and Mixed Approaches to Cryptocurrency Dynamics Analysis

Victor Dostov[1,2(✉)], Pavel Pimenov[1,2], and Pavel Shoust[1,2]

[1] Federal State Budgetary Educational Institution of Higher Education, Saint-Petersburg State University, 7-9 Universitetskaya Emb., St Petersburg 199034, Russia
dostov@npaed.ru

[2] Russian Electronic Money and Remittance Association, 5/2 Orlikov per, Moscow 107078, Russia

Abstract. Two approaches are most frequently used to predict the development of cryptocurrency market: deterministic and stochastic. The deterministic approach seeks to explain the development of cryptocurrencies through the relationship of several indicators. A stochastic approach (such as ARIMA) seeks to optimize the parameters of a statistical model. This article aims to compare approaches to the assessing cryptocurrencies development using the number of active Bitcoin and Ethereum wallets. For this purpose, a deterministic model based on the Verhulst equation, and a stochastic model based on ARIMA was formulated. The results show that the usage of relative differences wins over absolute ones. At the same time, the predictive value of a purely deterministic model on short segments is not very high, but it has the advantage in analytical form. Further will focus on the combination of a deterministic Bass-type model and statistical methods with stochastic analysis tools.

Keywords: Cryptocurrency dynamics analysis · ARIMA · Deterministic approach · Stochastic approach · Bitcoin · Ethereum

1 Introduction

Over the past five years, cryptocurrencies have become an integral part of the exchange trading, banking, and financial markets. Central banks in many countries are conducting experiments to create national cryptocurrencies-CBDC. Ripple digital currency has become an important factor in the development of the interbank payments market. In such conditions, the analysis of the dynamics of cryptocurrencies development takes on high practical and theoretical significance.

In modern studies of the dynamics of cryptocurrencies, two approaches dominate: deterministic (averaging and inferring the determinant component) [4, 6] and stochastic (using modern regression methods – for example, ARIMA) [15, 18]. Each of these approaches has both undeniable advantages and significant disadvantages. This article

© Springer Nature Switzerland AG 2021
O. Gervasi et al. (Eds.): ICCSA 2021, LNCS 12952, pp. 19–30, 2021.
https://doi.org/10.1007/978-3-030-86973-1_2

aims to conduct a comparative analysis of the application of both approaches to identify these advantages and disadvantages, the level of their reliability and effectiveness in describing the processes in the cryptocurrency market, as well as to consider the possibility and prospects of combined analysis using both stochastic and deterministic models.

The paper will be divided into four parts. In the first section, we look at the main aspects of the concept of "payment system", its classification into C2B and P2P, and the place of cryptocurrencies in this classification. In the second part, we consider the main idea of the deterministic approach, give successful examples of its implementation, and offer our own research on this issue. In the third part, we consider the stochastic approach and examples of its application. We present the authors' approach to predicting the dynamics of the number of wallets in major cryptocurrencies systems using ARIMA methods. Finally, we compare the results of the analysis and draw some conclusions about the choice between approaches, as well as the prospects for applying a mixed strategy.

2 Methodology

For this study, we selected two of the most famous cryptocurrencies – Bitcoin (hereinafter referred to as BTC) and Ethereum (hereinafter referred to as ETH). These cryptocurrencies are perfectly suitable for research purposes, because of a long history of development and a large number of active addresses.

In addition, we believe that these cryptocurrencies can be considered as a "payment system". For the purposes of the article, we will define the payment system as a set of tools, banking procedures and interbank money transfer systems that provide money circulation [8].

Considering Bitcoin and Ethereum as payment systems allows us to apply deterministic approaches to evaluate classical payment systems. This will be demonstrated in the next part of the article. It should be noted that these approaches may differ significantly depending on the payment system in the payment system. There are two models-P2P (payment from person to person) [16] and C2B (client for business) [8]. In C2B payments, a limited number of clients make payments to a company with an infinite capacity [13]. P2P payments are made between individuals within a closed community.

P2P transactions dominate in cryptocurrency schemes, therefore, when considering the degree of accuracy of the analysis, we suggest relying on studies focused on the analysis of P2P payment systems.

3 Deterministic Approach

The generally accepted approach to describing the diffusion of innovations relies on the assumption that it is possible to empirically describe the behavior of complex systems with a small number of parameters [18]. The essence of the approach is the construction of various analytical models, which include the assumption about the mechanisms that determine the behaviour of the system. The models are constructed either by using computer modelling, in which mechanisms determine the rules of model dynamics or

by using differential, less often integral, equations of model dynamics, in which these mechanisms are represented as a mathematical relationship between the states of the system in the present or past.

Often, studies that use a deterministic approach set an ambitious goal of predicting the dynamics of cryptocurrency prices for the needs of technical analysis. For example, Stephen Dipple, Abhishek Choudhary, James Flamino, Boleslaw K. Szymanski & G. Korniss [4] use standard geometric Brownian motion to model cryptocurrency rates. For social media activity and cryptocurrency trading volume, the authors use the Ornstein-Uhlenbeck geometric process. The researchers propose their own model of the correlation between different stochastic variables, which occur through the noise in the corresponding stochastic differential equation.

In our previous works, we have studied the application of a deterministic approach to the study of the behaviour of payment systems [6]. We did not try to create an algorithm for predicting the price, as the high influence of the randomness factor and hype will limit the practical significance of such a model. On the other hand, predicting the number of clients is of more practical value. We assumed that there are payment systems with binary behaviour, where the client's decision to use the system is significantly affected by the presence of other customers in this system.

To predict the dynamics of cryptocurrencies, we used the modified Bass [18] and Verhulst [14] equations. We adapted Bass' idea of interpreting the appearance of a quadratic multiplier because of the simulation effect. In P2P payment systems, you cannot make a transfer if your recipient is not included to the system.

In our calculations, we assumed that when forecasting the P2P market, the relationship between the following parameters should be considered:

- current number of users x.
- the maximum number of users, for example, the entire audience of a given country N. Therefore, the number of potential users not currently participating in the system is N-x.
- audience capture rate, which reflects the probability that a given user will start using the service: a > 0 (the reverse time of the decision) within a given period.
- number of new customers (registrations) R per unit of time.
- fatigue coefficient from system b.
- In addition, we considered the fatigue effect, which means the user can stop using the system at any time, which is not in the original Bass model.

As a result, to predict the P2P market, we proposed the Eq. (1):

$$\frac{dx}{dt} = a(N - x)x - bx \tag{1}$$

The model was used to predict the dynamics of changes in the number of active wallets (addresses) of the most popular cryptocurrencies – Bitcoin [1] and Ethereum [1].

The results of the study are shown in Fig. 1.

The usage of a deterministic approach has shown that due to the complex nature of cryptocurrencies (they are both a means of payment and an investment tool) [10], qualitative and quantitative analysis become more complex. To compensate the investment

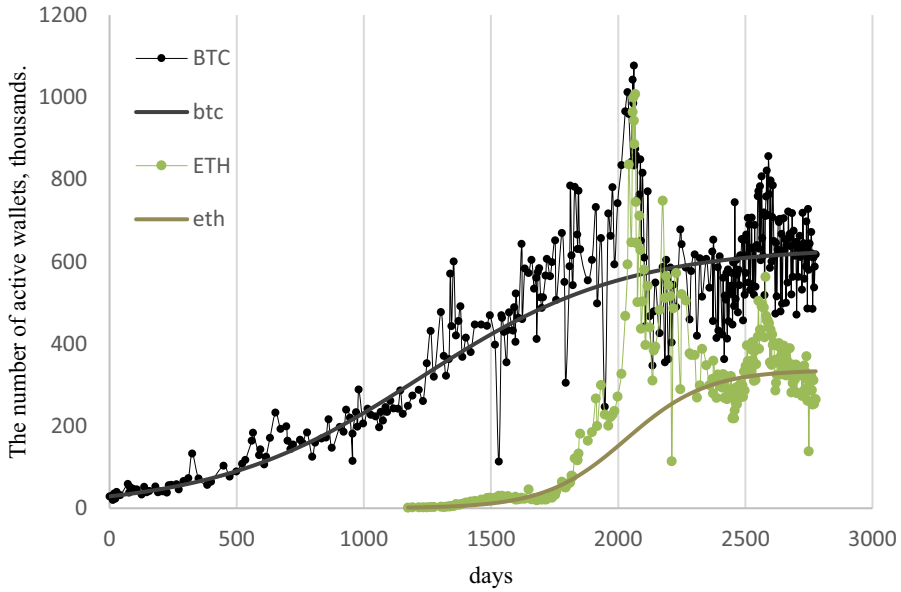

Fig. 1. Statistics of active wallets (addresses) Bitcoin and Ethereum 20.05.2012–27.12.2019

behavior of cryptocurrencies and get the correct overall trend, we used the interpolation method. For the purpose of simplifying the problem, we applied the method of the smallest SPO to the initial sample, summed by years. To do this, we obtained an interpolation with an average deviation of 5% (Fig. 2), introduced mainly by one period of speculative cryptocurrencies of 2017–2018 [17]. We excluded these anomalous points from the STR calculation by further refining the coefficients. Then we superimposed the resulting function on the full source data, as shown in Fig. 2. For this combination, the SPO is about 20%, which is quite expected in case of the initial noise level of the curve.

Similar calculations were carried out for Ethereum (ETH), for which the STR is 27%, also due to one anomalous value (Fig. 2).

Figure 1 and Fig. 2 demonstrate that our equation describes the behavior of both cryptocurrencies well. The exception is the period of 2017–2018, when the growth in the number of wallets was determined not by real transactions, but by speculative behavior. However, we believe that the deterministic approach, if applied to payment systems based on cryptocurrencies, is limited. This is because of two aspects that affect on the behavior of cryptocurrencies. First, the payment functionality, according to our assumption, is quite deterministic. Secondly, the investment component provides high volatility of the cryptocurrency exchange rate in competition with fiat currencies. This leads to a higher noise level of data in comparison with classical payment systems.

As shown in [6], equation coefficients (1) can be considered as constant for a limited period, since the natural development of services and markets in the long term inevitably leads to their change. The rapid growth in the popularity of bitcoin at the beginning of 2021 was because of speculative investments and is not a sign of the inefficiency of the model. During such periods, the model deviations become large, and the quality of the

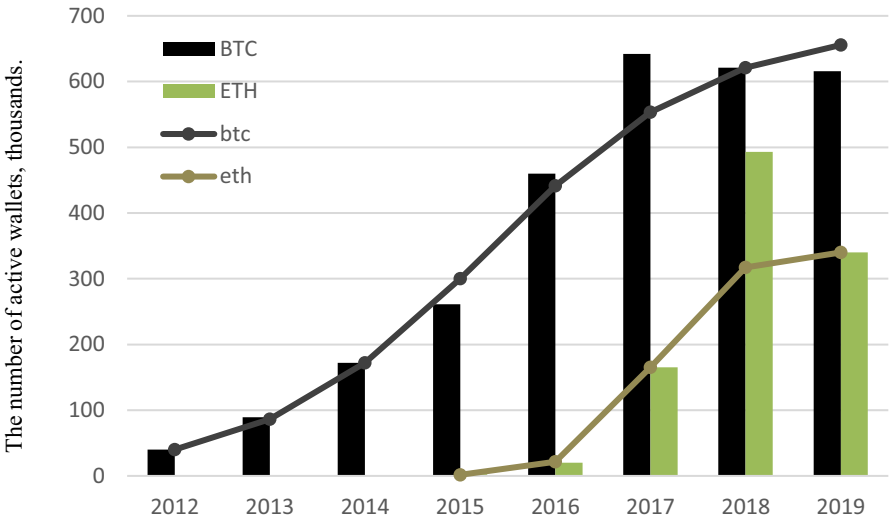

Fig. 2. Statistics of active Bitcoin and Ethereum wallets per day, thousands, averaged over years, actual (histograms) and calculated values.

forecast temporarily decreases. We expect to see a decrease in the exchange rate and a stabilization of user activity, after which the model can be reinterpreted considering new statistics.

4 Stochastic Approach

For the stochastic approach, a set of existing data is considered, and the parameters of the statistical model are optimized, which gives optimal results for the available data and for extrapolation of these data for future periods. The most effective methods at the moment are based on the theory of time series [11] and implemented in the form of algorithms of the ARIMA family [2]. A big advantage of these algorithms is the consideration of random, noise factors in the behaviour of the simulated systems, as well as some secondary systematic, for example, seasonal trends. At the same time, models do not include the available information about the nature of the simulated systems, often have weak explanatory power, require a large amount of data for training, and often poorly describe the behaviour of systems with long-term changing trends.

The application of ARIMA on the example of the cryptocurrency market is studied in detail in many academic works. of Russian authors. In this study, we present the results M. R. Safiullin, A. A. Abdukayev, L. A. Elshin [14] research where the authors suggested that stochastic models in the absence of effective tools for predicting exchange rate fluctuations of cryptocurrencies should show themselves most effectively.

The reliability and adequacy of the results were confirmed by comparing the actual and predictive parameters of the bitcoin exchange rate and based on a high R-square value. The research suggests big promise for using ARIMA in business process modelling based on usage of crypto transactions, including prediction of cryptocurrency price to minimize losses induced by high volatility rates.

However, in certain conditions, ARIMA does not provide an initial increase in efficiency in comparison with other forecasting methods. Sukhanov E. I., Shirnaeva S. Yu., Kozhemyakin D. A. [15] demonstrate the imperfection of forecasting in relation to exchange trading. They developed a special algorithm: decisions about buying or selling cryptocurrencies were made based on the forecast of price changes for each next day using the ARIMA/GARCH model. The results of each trading day were recorded, and eventually, a time series of the profitability of the trading system was obtained. Based on the comparison of statistical indicators of two-time series, conclusions about the feasibility of using this model for trading on the stock exchange were made.

The authors found that the model is well suited for the considered data (there is no autocorrelation in the residuals). However, the trading strategy based on this model did not give good results. The trading strategy was slightly ahead of the benchmark in its indicators, but statistical tests of the strategy and benchmark in case of making profitability did not show significant differences.

Wirawan, I.M., Widiyaningtyas, T., Hasan, M.M. [19] use ARIMA to model Bitcoin prices as well. Research shows that the ARIMA model is feasible to be used as a predictive method of Bitcoin for one to seven days ahead.

In our paper, to obtain an objective comparison of the stochastic and deterministic approaches, we propose to compare the effectiveness of using both approaches. For this purpose, we will analyze the data used in our previous work [6] using the ARIMA method.

In our study, the pmdarima and statsmodels.tsa packages of the Python programming language were used to build ARIMA. The R programming language tools were also used. The usage of the models considered in the work assumed the implementation of five main iterations:

1. Building a time series.
2. Checking the series for stationarity (as a result, the class of the used model is determined-ARMA or ARIMA).
3. The model parameters were selected by using auto-ARIMA. Auto-ARIMA works by conducting differencing tests (i.e., Kwiatkowski–Phillips–Schmidt–Shin, Augmented Dickey-Fuller or Phillips–Perron) to determine the order of differencing, d, and then fitting models within ranges.
4. Assessment of the reliability and adequacy of the constructed model.
5. Development of predictive parameters of the time series under study.

The task of defining criteria for a deterministic model applied to highly noisy data is not trivial, because direct formal comparison of a purely deterministic model with a model that includes random components does not seem correct to us. Considering the question of the significance of the deterministic model for a strongly stochastic process, we take as a basis the approach to the theory of time series described in the article by Kantorovich [11]. In his research, various trends and random components are distinguished in the data sample. Further, as a null hypothesis, we assume that the developed model closely describes the trend, and a noticeable further improvement is possible only taking into account random components that exist beyond this model. Let's also assume that the autoregressive integrated moving average model (ARIMA) is fairly

accurate. Therefore, we will propose the following criterion for the quality of our model as a correct description of the trend:

1. The remainder of the optimal model (relative BTC-btc or absolute (BTC-btc)/btc) give a stationary process.
2. Applying the ARIMA model to the time series obtained at the first step, respectively, gives a model with parameters (p, 0, q), where p and q are the order of the AR - and MA-components of the econometric model.
3. ARIMA estimated on the residuals of the model predicts better than the model applied to the original series.
4. We should make a special condition to implement this approach:
5. A deterministic model is trained by minimizing absolute deviations (or by another classical criterion).
6. The residuals of the deterministic model (relative and absolute) are calculated.
7. The stationarity criterion is applied to the remainders: the Dickey-Fuller test [9] or the CPSS test [12]. If the stationarity hypothesis is accepted, we presume that the model correctly describes the deterministic trend.
8. To determine the predictive strength of a deterministic trend, we apply various ARIMA modifications to the initial data and model residuals and compare the prediction accuracy for both cases.

The results are shown in Table 1. As data for ARIMA, we used 1) the first differences of the original series, 2) the BTC-btc differences, 3) the relative differences of the original data, and 4) the relative differences (BTC-btc)/BTC. The model was trained on 1 365 values, after which a sample of 30 values was predicted.

5 Comparison of Deterministic, Stochastic, and Mixed Approaches

The assumption that the trend is stationary for the BTC-btc differences and ETH-eth differences is confirmed by the behaviour of auto-ARIMA. For both Bitcoin and Ethereum, we observe the smallest prediction error for most of the metrics used in the latest model – ARIMA for relative differences with a deterministic trend. The order of integration of the series is determined based on the KPSS test. The use of the alternative Dickey-Fuller criterion does not change the conclusion in favour of the third model. The residuals from the ARIMA model are white noise, as indicated by the high P-values in the corresponding tests.

It should be noted that at a small interval and under the condition of the relative stability of the source data, the predictive power of ARIMA on real data and ARIMA on the differences between real data and the deterministic model is approximately the same, with a small advantage of the model on the BTC-btc differences. However, because this paper aims to identify trends in the cryptocurrencies' development, and not to create a tool for technical analysis, the most important issue is the functioning of ARIMA over long periods. We expect that in this case, the differences will be more significant. The above-mentioned research by Wirawan, I. M., Widiyaningtyas, T., Hasan, M. M. shows that ARIMA loses its effectiveness on the seventh day of the observations period. For

Table 1. The value of forecast error metrics for different ARIMA models

Bitcoin		Error metrics				P-value of the Box-Pierce and Lewing-Box tests
		RMSE	MSE	MAE	MAPE	
1	ARIMA in differences (4, 1, 2)	0.133	0.018	0.105	0.008	0.122 0.121
2	ARIMA for BTC-btc (4, 1, 2)	0.157	0.025	0.134	1.194	0.126 0.124
3	ARIMA in relative differences (5, 0, 3)	0.163	0.027	0.137	0.990	0.207 0.204
4	ARIMA for (BTC-btc)/btc (1, 0, 0)	0.021	0.0005	0.019	3.107	0.409 0.408
Ethereum		RMSE	MSE	MAE	MAPE	
1	ARIMA in differences (3, 2, 1)	0.292	0.085	0.273	0.021	0.967 0.966
2	ARIMA for ETH-eth (2, 1, 1)	0.299	0.089	0.279	1.015	0.962 0.961
3	ARIMA in relative differences (3, 1, 1)	0.139	0.019	0.104	1.023	0.968 0.968
4	ARIMA for (ETH-eth)/eth (2, 1, 1)	0.024	0.001	0.022	1.035	0.878 0.877

the purposes of our paper, we performed the ARIMA calculation on a smaller amount of data to increase the number of test data. The predicted data for ARIMA based on real data (1, 1, 2) and ARIMA based on BTC-btc differences (2, 0, 2) is shown in Fig. 3.

Both models, after a period of confidently following the real data, demonstrate stationarity in the long term. This situation is typical for ARIMA. Nevertheless, it should be noted that data from the deterministic ARIMA model demonstrates a more adequate trend prediction. MAPE for ARIMA with BTC-btc differences is better than for ARIMA based on real data (0.1609 vs. 0.1716). Thus, the effectiveness of ARIMA with the usage of BTC-btc differences and ETH-eth differences strongly depends on the correctness of the expert assessment of the deterministic model.

Table 1 shows that using the difference with the model gives better results than differentiating. On the other hand, the usage of relative differences wins over absolute ones because of the high volatility of the source data. Therefore, we can conclude that: 1) in the above sense, the deterministic model is reliable and 2) the combination of the deterministic model and the ARIMA methodology is a promising direction for the study of highly volatile payment systems [3]. At the same time, the predictive value of a purely deterministic model on short segments is not very high, but it has the advantage in analytical form. We suppose that other criteria of optimization of parameters of the

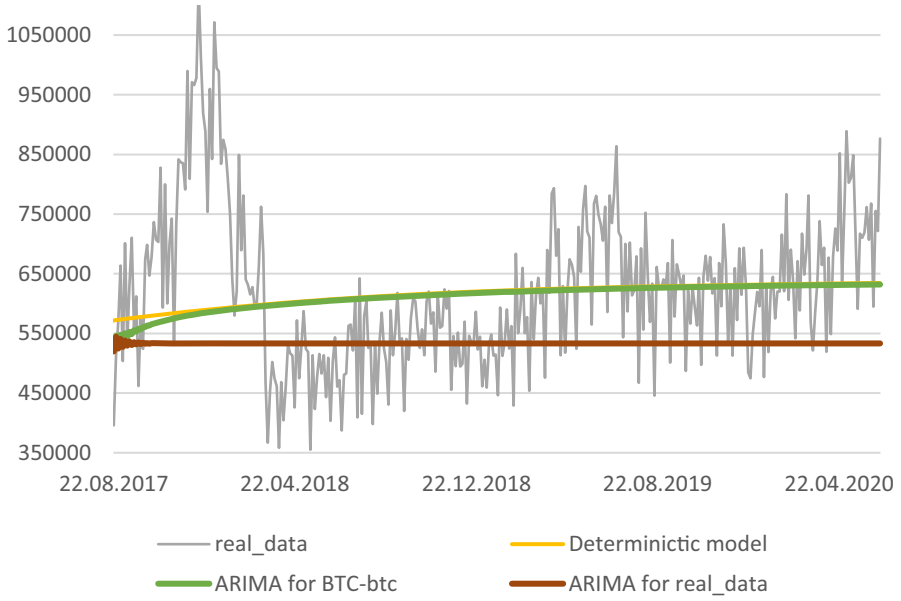

Fig. 3. ARIMA Prediction Charts

deterministic model can be tested to improve the quality of the mixed approach and we are planning to do this in future. We also see that for the BTC model, the parameter $d = 0$, which indicates that our model takes the trend into account qualitatively. For other models, $d = 1$. This may indicate the presence of small unrecorded, for example, seasonal trends, higher stationarity requirements in the ARIMA implementation, or the need for more careful parameters' selection of our model in order to increase stationarity, for example, based on the Dickey-Fuller criterion.

6 Using Mixed Approach to Analyze the Dynamics of Other Cryptocurrencies

It is necessary to note that the effectiveness of using a combination of stochastic and deterministic approaches in configuration of combining ARIMA methods and the Bass equation may differ depending on the nature of the cryptocurrency and the initial idea of their positioning.

Both Bitcoin and Ethereum were initially positioned as an alternative payment system to fiat money. It was assumed that cryptocurrencies would become a faster and safer way to pay for goods and services. Nevertheless, the regulatory barriers of many countries make cryptocurrencies as volatile speculative trading instruments.

Some cryptocurrencies in these conditions have retained the features of payment systems and are used for their intended purpose (for example, Bitcoin and Ethereum). This is reflected in the presence of a characteristic periodization of the growth of the number of wallets, explained by the Bass equation.

On the other hand, if the cryptocurrency was created for the purpose of distribution as a tool for financing a project (for example, ICO product) or as a joke (for example, Dogecoin), this cryptocurrency does not have many payment system features. As an example, let's look at the graph of unique addresses for Dogecoin [5] (Fig, 4).

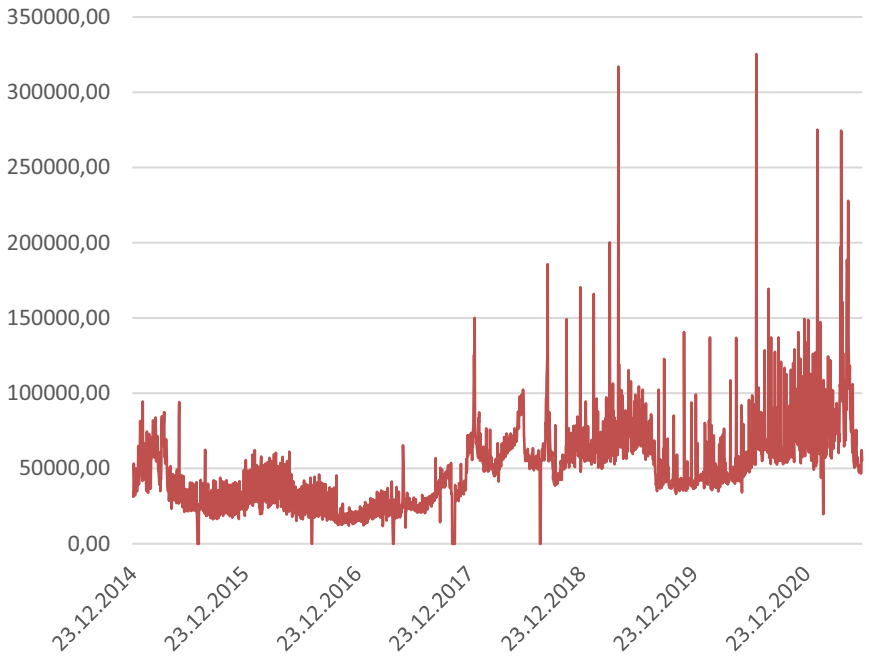

Fig. 4. Number of active Dogecoin wallets

Dogecoin is named after the Internet meme, which demonstrates the frivolity of this coin. It should be noted that the graph of the number of addresses does not have the typical inflection that is observed in the Bitcoin or Ethereum charts. Number of active Dogecoin wallets is feint. This growth is provided not by real demands, but by series of twitters by one of cryptocurrency opinion leaders, E.Mask [7] and therefore are not stochastic in sense of [3] and not deterministic [18]. That's why it can be assumed that the application of the approach described in this paper in case of predicting the number of unique wallets for such cryptocurrency is difficult.

On the other hand, we assume that the model can be quite accurate if it is used to predict the development of Central Bank Digital Currencies (CBDC). Indeed, CBDC is essentially the same cryptocurrency – a distributed ledger, the usage of tokens, a similar level of security. However, unlike existing cryptocurrencies, the price of CBDC and its security will be guaranteed by the state regulator. In fact, CBDC will be a classic payment system. We are confident that for this type of cryptocurrency mixed approach can become an effective tool for predicting scalability. We propose to test this hypothesis when at least one CBDC starts functioning, and the regulator starts to publish statistics on its development.

7 Conclusions

The study shows that the application of both a deterministic and stochastic model to such strongly stochastic systems as cryptocurrencies, in general, has a positive result. It was shown that using the result of the model as a trend can improve the quality of such models forecast.

We have shown the advantages and disadvantages of both a deterministic and stochastic approach to predicting the dynamics of the development of cryptocurrencies. In particular, we found that deterministic models show low efficiency over brief time segments; however, they are sufficiently efficient in general trend forecasts, although the efficiency decays in the period of active growth over a long period of time. On the other hand, stochastic (in particular, ARIMA) forecasting methods show excellent results in short-term periods but cannot effectively predict the behaviour of a P2P system over long distances because of insufficient emphasis on non-numeric factors.

In this regard, the combination of a deterministic Bass-type model and statistical methods with stochastic analysis tools seems to be a prospective direction and will be the subject of further research.

References

1. Bitcoin Active Addresses historical chart. https://bitinfocharts.com/comparison/bitcoin-activeaddresses.html. Accessed 13 Mar 2021
2. Box, G.E.P., Jenkins, G.M.: Time Series Analysis: Forecasting and Control, 575 p. Holden-Day, San Francisco (1970)
3. Christodoulos, C., Michalakelis, C., Varoutas, D.: Forecasting with limited data: combining ARIMA and diffusion models. Technol. Forecast. Soc. Change **77**(4), 558–565 (2010)
4. Dipple, S., Choudhary, A., Flamino, J., Szymanski, B.K., Korniss, G.: Using correlated stochastic differential equations to forecast cryptocurrency rates and social media activities. Appl. Netw. Sci. **5**(1), 1–30 (2020). https://doi.org/10.1007/s41109-020-00259-1
5. Dogecoin Active Addresses historical chart. Homepage https://bitinfocharts.com/comparison/activeaddresses-doge.html
6. Dostov, V., Shoust, P., Popova, E.: Using mathematical models to describe the dynamics of the spread of traditional and cryptocurrency payment systems. In: Misra, S., et al. (eds.) ICCSA 2019. LNCS, vol. 11620, pp. 457–471. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-24296-1_36
7. Elon Musk's Bitcoin Fun Continues. Bloomberg. https://www.bloomberg.com/opinion/articles/2021-05-17/elon-musk-controls-bitcoin-and-dogecoin-prices-with-pure-magic
8. Four Different Types of Services. Homepage https://localfirstbank.com/content/different-types-of-banking-services/. Accessed 10 Mar 2021
9. Fuller W. A.: Introduction to Statistical Time Series. Wiley (2009)
10. Genkin, A., Micheev, A.: Blockchain. How it Works and What Awaits Us Tomorrow, 592 p. Alpina Publisher (2018)
11. Kantorovich, G.G.: Time series analysis. HSE Econ. J. **2020**(1), 85–116 (2006)
12. Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y.: Testing the null hypothesis of stationarity against the alternative of a unit root. J. Econometrics **54**(1–3), 159–178 (1992)
13. Laffont, J.-J.: Regulation and Development, 440 p. Cambridge University Press (2005)

14. Recherches mathématiques sur la loi d'accroissement de la population, dans Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles. 1845. № 18, pp. 1–42
15. Safiullin, M.R., Abdukaeva, A.A., El'shin, L.A.: Methodological approaches to forecasting dynamics of cryptocurrencies exchange rate using stochastic analysis tools (on the example of bitcoin). Finan. Theor. Pract. **22**(4), 38–51 (2018). https://doi.org/10.26794/2587-5671-2018-22-4-38-51
16. Takako, F.-G.: Non-cooperative Game Theory, 260 p. Springer, Japan (2015). https://doi.org/10.1007/978-4-431-55645-9
17. The annual cryptocurrency market review: what's left after the hype. Bitnews Today (2019). https://bitnewstoday.com/news/the-annual-cryptocurrency-market-review-what-s-left-after-the-hype. Accessed 19 Mar 2021
18. The Bass Model Homepage. http://bassbasement.org/BassModel/Default.aspx. Accessed 15 Mar 2021
19. Wirawan, I.M., Widiyaningtyas, T., Hasan, M.M.: Short Term prediction on bitcoin price using ARIMA method. In: International Seminar on Application for Technology of Information and Communication (iSemantic), Semarang, Indonesia, pp. 260–265 (2019). https://doi.org/10.1109/ISEMANTIC.2019.8884257