Joint Mid-Term Forecast of Cryptocurrencies in Technics of Inductive Modelling (on Example of XRP, Waves, ETH)

Bulat Bayburin
Moscow Institute of Physics and Technology, Moscow, Russia
bulatbaiburin@gmail.com

Pavel Mogilev
Technological University of Dublin, Dublin, Ireland
nismorus@gmail.com

Mikhail Alexandrov
Russian Presidential Academy of National Economy and Public Administration, Moscow, Russia
Autonomous University of Barcelona, Barcelona, Spain
malexandrov@mail.ru

John Cardiff
Technological University of Dublin, Dublin, Ireland
john.cardiff@tudublin.ie

Olexiy Koshulko
Glushkov Institute of Cybernetics
Kyev, Ukraine
koshulko@gmail.com

Abstract — **Low accuracy of mid-term value forecasts is an open problem that impedes cryptocurrencies promotion as a new world's markets legit financial instrument. One of the factors is the lack of reliable predictive models including all or the majority of significant objective and subjective factors. In this paper, we propose to partially compensate it by taking into account mutual relations of cryptocurrencies and using the noise-immune algorithms of GMDH technology. Their object of consideration is the dynamics of cryptocurrencies XPR, Waves and ETH for 2 full years from March 2019 till February 2021 with the step of 1 week. We consider forecasting with next different options: 2 periods of cryptocurrency behaviour (calm, crisis), 2 forecast horizons (week, month), 3 models (autoregression, regression and hybrid), 2 basic algorithms from the GMDH Shell tool (combinatorial, neural network). We also study how preliminary smoothing of cryptocurrencies' dynamics affects to forecasting error. The popular Holt-Winters method is used as a baseline. The results prove to be promising, which allows us to propose this the rather simple approach for experts of cryptocurrency market.**

I. Introduction

A. Motivation

In spite of this skepticism related to crypto-market in total, cryptocurrencies remain attractive. Indeed, crypto-assets today are not only a way to make quick money in trading but also a good fundraising opportunity for startups as like as an instrument of making investment process easier for many companies [1]. It is especially important in the epoch of the Covid-19 pandemic, when many sectors of the world economy are paralyzed and investors begin to look for alternative directions for investments and new forms for their rapid implementation. The principal obstacle that inhibits the development of cryptocurrency market is not a high market volatility rate but inability to provide a satisfactory forecast of their dynamics.

There are two reasons for this inaccurate forecast:

- strong dependence of cryptocurrencies' dynamics on various subjective factors as attitude of authorities, interest of business community, confidence of population, which may not be easily taken into account in a formal model;
- imperfection of traditional algorithms for forecasting time series due to its low noise immunity.

These circumstances determined the content of our research and our proposals. Namely:

- we try to compensate the lack of models including the mentioned factors using mutual ties between cryptocurrencies;
- we tune and test well-known GMDH-based algorithms having high noise immunity.

B. Related work

At the moment the Weiss Agency is the unique reliable institution which evaluates behavior of cryptocurrencies assigning appropriate ratings for them [2,3]. Internet also contains resources which forecast different cryptocurrencies based on expert assessments. For example, here [4-6] one can find quantitative forecasts for the cryptocurrencies XRP, Waves and ETH considered in this paper. But these sources usually hide exact algorithms used in their calculations.

On a stable period of cryptocurrency behavior these quantitative assessments prove to be enough good for short-term forecasts (from several days to a week) and worse for mid-term forecasts (from several weeks to a month). Meanwhile for the mid-term forecasts one often prefers qualitative assessments.

The situation cardinally changes, when we deal with forecasts related to an unstable period. Here, quantitative assessments often prove to be unsatisfactory for short-term forecasts and users implement here qualitative assessments as well. However, for mid-term forecasts even these qualitative assessments become ‘dangerous’ due to a very high volatility that took place since October 2020.

In the paper [7] authors study possibilities of GMDH-based algorithms to provide satisfactory mid-term forecast for one given cryptocurrency Waves. They consider both calm and crisis periods covering 3 years 2017-2019 and test all four algorithms from GMDH Shell tool for week and month forecasts.

The authors of [8] show how to use queries to search services of Yandex and Google to improve the quality of mid-term forecast of the cryptocurrencies XRP, Waves and ETH. Such queries are considered as some reflection of subjective factors defining dynamics of these cryptocurrencies. The period of consideration is 2 full years during 2019-2021 including calm and crisis periods. For modeling two GMDH-based algorithms from the mentioned GMDH Shell tool are used.

One should note that financial segment of Internet is full of different opinions and controversial information about cryptocurrencies that looks as an information warfare and may be a subject of separate investigation [9], [10].

C. Problem setting

In this paper, we test four hypotheses:

1) Mutual ties between cryptocurrencies take into account an additional external information that can improve forecasts in comparison with autoregression model.
2) GMDH-based algorithms with their high noise-immunity are better suited to dynamics of time series that can improve forecast in comparison with traditional algorithms based on exponential smoothing.
3) Preliminary smoothing of dynamics for cryptocurrency to be forecasted can essentially improve the results and cubic root is the best one between other procedures under consideration.
4) The mentioned improvements refer to all cryptocurrencies.

To check the first hypothesis, we build regression and hybrid models, where the regressors are other cryptocurrencies with respect to the cryptocurrency under consideration. The forecasts are compared with forecasts on autoregressive model.

To check the second hypothesis, we test two typical GMDH-based algorithms on autoregressive model and compare their forecasts with forecasts of widely known Holt-Winters method.

To check the third hypothesis, we test several options of smoothing on autoregression model with cryptocurrency XRP whose behavior is the most volatile.

To check the forth hypothesis, we consider possibilities of GMDH-based algorithms to forecast cryptocurrencies’ dynamics of XRP, Waves, and ETH. Our options here are: calm and crisis periods, week and month forecast horizons, with preliminary smoothing and without this procedure.

One should do the following two important notes:

- The notions ‘mid-term’ and ‘short-term’ are shortened more in crisis periods than in periods of usual volatility.
- Both mentioned research [7], [8] and this study are preliminary works to show advantages of using certain external information and inductive modeling. Such studies are a preliminary for current research in this area.

II. CRYPTOCURRENCIES

The source of information about cryptocurrency dynamics are weekly data of XRP/Euro, Waves/Euro, and ETH/Euro rates taken from the resources [11, 12]. Each value is the average value over the course of a week. The datasets cover 2 full years. It is equal to 52 x 2 = 104 weeks from March 1, 2019 to February 28, 2021. The so-called stable (calm) period lasts the first 87 weeks and the so-called unstable (crisis) period lasts the last 17 weeks. Fig. 1-3 illustrate the dynamics of all 3 cryptocurrencies for the whole period. Table I presents average and relative average values of each cryptocurrencies for each period. Relative average values refer to the averages for the calm period. Figure 4 illustrates these values.

TABLE I. AVERAGE AND RELATIVE AVERAGE OF DATASETS

Crypto	*Calm*	*Crisis*
XRP	0,64 / 1,00	0,72 / 1,13
Waves	1,14 / 1,00	1,86 / 1,63
ETH	6,00 / 1.00	9,40 / 1,57

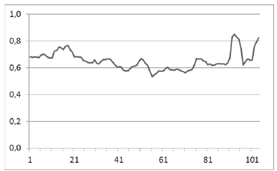

Fig.1. XPR-dynamics

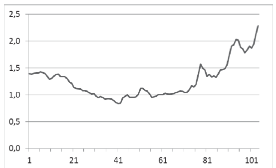

Fig.2. Waves-dynamics

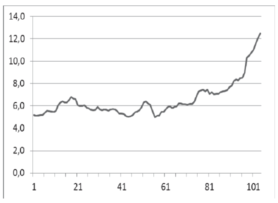

Fig.3. ETH-dynamics

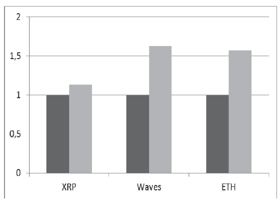

Fig.4. Relative average values

III. MODELS AND METHODS

A. Autoregression, regression and hybrid models

To check the hypothesis about contribution of mutual ties of cryptocurrencies to price formation of these cryptocurrencies we consider 3 models: autoregression, regression and hybrid.

- The popularity of autoregression is easily explained by the fact that such a model allows to ignore any factors outside the time series itself. Here it is assumed that these factors can manifest themselves in the pre-history of cryptocurrencies’ behavior.

- The rRegression model uses normalized values of cryptocurrencies as input variables. In this paper we name cryptocurrencies in regression equation as ‘regressors’.
- The hybrid model is a composition of autoregression and regression models. That is, it includes pre-history of all cryptocurrencies and regressors.

All these models were tested with GMDH-based algorithms for forecasts of cryptocurrencies XPR, Waves and ETH.

B. GMDH-based algorithms from GMDH Shell

Group Method of Data Handling (GMDH) is a technology of machine learning (ML) for creating models with increased noise immunity. The ideas and perspectives of GMDH are presented in many publications; see, for example, [11]. Theoretical bases of GMDH are described the most completely in [12]. Numerous applications of GMDH are reflected in papers and books, which can be downloaded from the resource [13].

In our experiments we use algorithms from the well-known platform GMDH Shell [14]. This platform offers several algorithms for forecast of time series. But in our research, we use only two algorithms:

1) Combi algorithm, which observes all models on each complexity level.

2) Neuro algorithm, which selects the best models on each complexity level and forms a new set of models from the chosen ones based on the principles of evolution.

It should be noted that the complexity of the model is defined by the model’s parameter quantity that needs to be assessed.

In a certain sense these two algorithms are the antipodes having in view their principles of forming new models when moving from one step of the algorithm to the next one.

The models’ quality is evaluated with the procedure of 2-fold cross validation, which corresponds to the symmetric criterion of regularity [11]. Errors are measured by common Mean Absolute Percentage Error (MAPE) being independent on absolute values of errors.

IV. EXPERIMENTS

A. Plan of experiments

In the experiments with GMDH-based algorithms we use options presented in Table II. The term ‘preprocessing’ means here preliminary transformation of cryptocurrencies’ dynamics value series. Totally, we completed 104 experiments with GMDH-based algorithms.

TABLE II. OPTIONS OF EXPERIMENTS WITH GMDH-BASED ALGORITHMS

Options	*Number*	*Contents*
Cryptocurrencies	3	XRP, Waves, ETH
Models	3	autoregression, regression, hybrid
Methods	2	combi, neuro
Preprocessing *)	4	absence, cubic and square roots, log
Periods of modeling	2	calm, crisis
Horizons of forecast	2	week, month

*) Testing only with cryptocurrency XRP

Baseline in our research was determined by the results of forecast with widely known Holt-Winters method [15]. In the experiments with this method, we use options presented in Table III. Totally, we completed 24 experiments with Holt-Winters method.

TABLE III. OPTIONS OF EXPERIMENTS WITH HOLT-WINTERS METHOD

Options	*Number*	*Contents*
Cryptocurrencies	3	XRP, Waves, ETH
Models	1	Autoregression
Preprocessing	2	absence, cubic root
Periods of modeling	2	calm, crisis
Horizons of forecast	2	week, month

B. Forecast with data smoothing

It is well known that in many cases preliminary smoothing of time series can improve (sometimes significantly) results of forecasts. Speaking of “smoothing” we mean transformation of individual values without any averaging. The latter allows to return to source data after forecast using the inverse transformation.

In this research we study effect of transformation with square root, cubic root, and logarithm in the form of log(1+x). All the experiments were completed only with one cryptocurrency XRP using autoregression model: we consider XRP as a representative of all cryptocurrencies, autoregression model allows to see effect of transformation without contribution of any regressors. Tables IV and V show results of forecast that refer to the calm and crisis periods respectively. Figure 8 illustrates the contents of Tables IV and V for algorithm Neuro.

TABLE IV. ERRORS (MAPE, %%) OF MODELING WITH AND WITHOUT TRANSFORMATION ON CALM PERIOD

Transformations	*Week*		*Month*	
	Combi	*Neuro*	*Combi*	*Neuro*
absence	5,6	5,5	9,1	8,9
square root	2,8	2,7	4,4	4,3
cubic root	**1,9**	**1,8**	**3,0**	**2,6**
logarithm	4,8	4,7	8,2	8,0

TABLE V. ERRORS (MAPE, %%) OF MODELING WITH AND WITHOUT TRANSFORMATION ON CRISIS PERIOD

Transformations	*Week*		*Month*	
	Combi	*Neuro*	*Combi*	*Neuro*
absence	20,4	12,9	16,3	12,6
square root	9,1	5,6	9,1	3,7
cubic root	**5,9**	**3,5**	**5,5**	**2,6**
logarithm	16,5	12,4	15,7	9,9

One can see that in all cases:

- transformation with cubic root proves to be the best one, it allows to reduce errors by 3-4 times in comparison with the absence of transformation;
- transformation with logarithm shows results being close to those without transformation;
- transformation with square root shows intermediate results between results with cubic root and logarithm;
- algorithm Neuro demonstrates better results than that of algorithm Combi.

All the other results below refer to experiments with preprocessing in the form of cubic root.

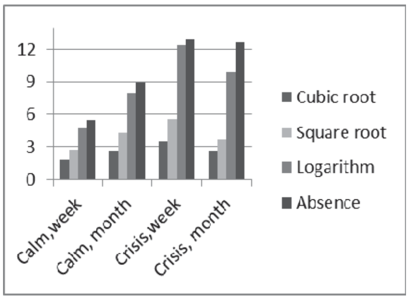

Fig.8. MAPE %%, algorithm Neuro with different transformations on autoregression

C. Modeling with different models

In these series of experiments, we compare autoregression, regression, and hybrid models. Tables VI-VIII show values of MAPE for cryptocurrencies XRP, Waves, and ETH respectively. To make interpretation of the Tables clearer (more representative) we illustrate their contents on Figures 9-11 on the example of algorithm Neuro. In total Neuro and Combi demonstrate close results but Neuro exceeds Combi in the majority of cases.

TABLE VI. ERRORS (MAPE, %%) OF MODELING, XRP

Options	*Autoregression*		*Regression*		*Hybrid*	
	Combi	*Neuro*	*Combi*	*Neuro*	*Combi*	*Neuro*
Calm,week	1,9	1,8	1,1	**0,9**	1,8	1,9
Calm,month	3,0	2,6	3,5	3,8	**1,6**	1,7
Crisis,week	5,9	**3,5**	4,3	3,9	5,3	4,2
Crisis,month	5,5	**2,6**	4,7	3,1	6,2	5,5

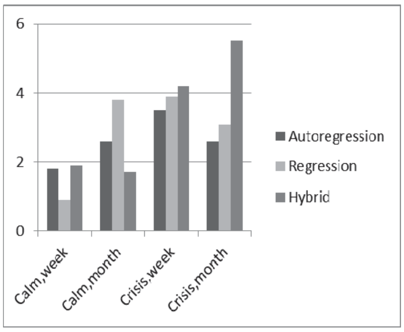

Fig. 9. MAPE %%, modeling with cryptocurrency XRP and algorithm Neuro

Results of modeling with cryptocurrency XRP show:

- usage of regressors improves results for calm period;
- usage of regressors worsens results for crisis period;
- in total forecasts for calm period are better than that for crisis period (it is the expected result).

Notes:

- regressors are used both in regression model and in hybrid model;
- regression is better in some cases than hybrid model and vice versa.

Results of modeling with cryptocurrency Waves show:

- usage of regressors improves results for all cases except the month forecast on crisis period;
- there is no any regularity mentioned in the quality of forecasts on calm and crisis periods.

Here we can do the same notes as those above for XRP.

TABLE VII. ERRORS (MAPE, %%) OF MODELING, Waves

Options	*Autoregression*		*Regression*		*Hybrid*	
	Combi	*Neuro*	*Combi*	*Neuro*	*Combi*	*Neuro*
Calm,week	2,7	2,8	3,2	**2,2**	3,0	2,6
Calm,month	3,3	3,4	3,4	4,3	**2,0**	2,1
Crisis,week	1,9	3,0	2,5	**1,4**	2,7	1,7
Crisis,month	3,5	**1,8**	2,6	2,7	3,8	2,5

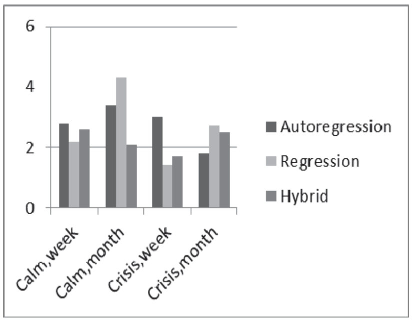

Fig.10. MAPE %%, modeling with cryptocurrency Waves and algorithm Neuro

TABLE VIII. ERRORS (MAPE, %%) OF MODELING, ETH

Options	*Autoregression*		*Regression*		*Hybrid*	
	Combi	*Neuro*	*Combi*	*Neuro*	*Combi*	*Neuro*
Calm,week	2,3	2,1	0,7	**0,5**	1,0	1,1
Calm,month	3,4	3,5	0,9	**0,7**	2,8	3,3
Crisis,week	2,3	2,7	1,0	**0,6**	1,1	1,5
Crisis,month	2,4	2,4	4,0	2,3	2,7	**1,9**

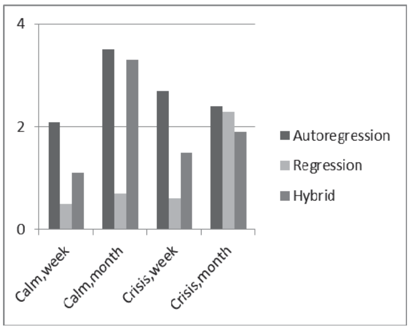

Fig.11. MAPE %%, modeling with cryptocurrency ETH and algorithm Neuro

Results of modeling with cryptocurrency ETH show:

- in all cases usage of regressors improves results;
- there is no any regularity mentioned in the quality of forecasts on calm and crisis periods.

Here we can do the same notes as those above for XRP.

D. Modeling with different algorithms

In these series of experiments, we compare GMDH-based algorithms and Holt-Winters (H-W) method with each other. Tables IX and X contain the results of experiments with cryptocurrency XRP.

It is easy to see that:

- in all cases GMDH-based algorithms show essentially better results than that of Holt-Winters method;
- there are several cases, where Combi proves to be better than Neuro and vice versa.

The similar advantage and variety of results may be showed for the cryptocurrencies Waves and ETH.

TABLE IX. ERRORS (MAPE, %%) OF MODELING XPR, WEEK FORECAST

Models	Calm			Crisis		
	Combi	*Neuro*	*H-W*	*Combi*	*Neuro*	*H-W*
autoregression	1,9	1,8	**2,2**	5,9	3,5	**7,3**
regression	1,1	0,9	-	4,3	3,9	-
hybrid	1,8	1,9	-	5,3	4,2	-

TABLE X. ERRORS (MAPE, %%) OF MODELING XPR, MONTH FORECAST

Models	Calm			Crisis		
	Combi	***Neuro***	***H-W***	***Combi***	***Neuro***	***H-W***
autoregression	3,0	2,6	**4,3**	5,5	2,6	**13,9**
regression	3,5	3,8	-	4,7	3,1	-
Hybrid	1,6	1,7	-	6,2	5,5	-

We summarized all the results (the current ones and the previous ones) concerning cryptocurrencies XRP, Waves, and ETH in Table XI. It shows number of cases, where algorithm Combi is better than algorithm Neuro and vice versa for different periods and horizons of forecast. If the results differ .by less than 1% then such results are considered as equal.

TABLE XI. CASES OF ADVANTAGE OF COMBI AND NEURO, MODELING ALL CRYPTOCURRENCIES

Methods	Calm		Crisis	
	Week	***Month***	***Week***	***Month***
Combi	-	2	3	-
Neuro	1	-	4	6
Equality	8	7	2	3
Total	9	9	9	9

Table XI shows:

- both algorithms demonstrate almost equal results in more than 50% of cases;
- Neuro has certain advantages over Combi on the crisis period.

V. CONCLUSION

A. Results

We completed 3 series of experiments with GMDH-based algorithms. The results of these experiments are:

- first series of experiments shows that cubic root proves the best smoothing procedure;
- second series of experiments shows that regressors always improve results on the calm period but they worsen results on the crisis period for XRP and Waves;
- third series of experiments shows that GMDH-based algorithms demonstrate practically equal results in more than 50% of cases; Neuro has certain advantages over Combi on crisis period; in all cases GMDH-based algorithms are essentially better than the popular Holt-Winters method being the baseline.

B. Future work

In future we intend to use:

- additional external information concerning dynamics of the world economy and finance;
- qualitative scales for cryptocurrencies' values;
- non-linear regression for building prognostic models [16].

REFERENCES

[1] Wikipedia: Blockchain technology, Web: // https://en.wikipedia.org/wiki/Blockchain

[2] M.D. Weiss, J.M. Villaverde, "Dark Shadows with a Bright Future", Report, 2019, Web:// https://weisscryptocurrencyratings.com/reports/WCY/sq/crypto-web-free-report/

[3] Internet resource: Weiss Crypto Agency, Web: // https://weisscrypto.com/

[4] Predictor of XRP, Web: // https:// coinpredictor.io/ripple

[5] Predictor of Waves, Web: // https:// coinpredictor.io/waves

[6] Predictor of ETH, Web:// https:// coinpredictor.io/ ethereum-classic

[7] P. Mogilev, A. Boldyreva, M. Alexandrov, J. Cardiff, "GMDH-based models for mid-term forecast of cryptocurrencies (on example of Waves)", In: Proc. Intern. Conf. CSIT-2020 (Workshop IWIM-2020), Lviv, Ukraine, 2020, pp. 13-16, Web: // https://ieeexplore.ieee.org/abstract/document/9321873

[8] Boldyreva, B. Bayburin, M. Alexandrov, O. Koshulko, P. Mogilev, "Search queries to Internet as a tool for mid-term forecast of cryptocurrencies (XRP, Waves, ETH) on GMDH-based models", In: Proc. Intern. Conf. CSIT-2021 (Workshop IWIM-2021), Lviv, Ukraine, 2021, pp. 375-379, Web: // https://ieeexplore.ieee.org/document/9648740

[9] A. Petrov, O. Proncheva, "Modeling propaganda battle: decision-making, homophily, and echo chambers", In: Artificial Intelligence and Natural Language (AINL-2018), Springer, CCIS book series, vol 930, 2018, pp. 197-209. DOI: 10.1007/978-3-030-01204-5_19; Web:// https://link.springer.com/chapter/10.1007%2F978-3-030-01204-5_19

[10] Petrov, O. Proncheva, "Propaganda battle with two-component agenda," In: Proc. MACSPro (Workshop 2019), Austria, CEUR, vol. 478, 2019, pp. 28–38; Web:// http://ceur-ws.org/vol-2478/

[11] V. Stepashko, "Developments and prospects of GMDH-based inductive modeling," Advances in Intell. Systems and Computing II; Springer, AISC book series, vol. 689, 2017, pp. 346–360101.

[12] V. Stepashko, "Method of critical variances as analytical tool of theory of inductive modeling", J. Autom. Inform. Sci., vol. 40, no. 3, 2008, pp. 4–22

[13] GMDH Resource in IRTC ITS NAS of Ukraine; Web: // http//mgua.irtc.org.ua/

[14] Resource: GMDH Shell, Web: // http// www.gmdhshell.com

[15] C. Holt, "Forecasting trends and seasonals by exponentially weighted averages," Intern. J. of Forecasting, vol. 20, no. 1, 2004, pp. 5–10.

[16] V. Strijov, G. Weber "Nonlinear regression model generation using hyperparameter optimization", Computers and Mathematics with Applications, 60(4), 2010, pp. 981-988.