Comparative Analysis of Time Series Models for Short-term Price Forecasting of Monetary Commodities

Article *in* Harbin Gongcheng Daxue Xuebao/Journal of Harbin Engineering University · August 2023

3 authors, including:

Aakash Purohit
Jai Narain Vyas University

Shubham Mathur
Jai Narain Vyas University

Comparative Analysis of Time Series Models for Short-term Price Forecasting of Monetary Commodities

Aakash Purohit
purohitaakash27@gmail.com | M.B.M. University, Jodhpur

Shubham Mathur
shubham.mathur341@gmail.com | M.B.M. University, Jodhpur

Nemi Chand Barwar
ncbarwar@gmail.com | M.B.M. University, Jodhpur

Abstract-Time Series Models are used for data that is directly related to time. Unlike regression-based models, these models consider certain time dependencies to provide better forecasting results. Time Series Models are well suited for forecasting stock or monetary commodity prices as these commodities move upward and downward with time. Time Series Models consider seasonal trends and also consider volatility while making predictions on the data. Every monetary commodity has some amount of volatility present in it, especially cryptocurrency which is highly volatile. For these monetary commodities Time Series Models are well suited for future price prediction. Short-term predictions of these monetary commodities can help investors get a general idea about how Gold or Cryptocurrency markets would move in the immediate future. While INR prices in the short term can help determine how it would immediately fare against global currencies. In this research paper, we have compared three different time series models namely ARMA, ARIMA, and Prophet to provide forecasts of three monetary commodities namely INR, Gold, and Cryptocurrency. The results depict a comparison between time series models selected in short-term price prediction of these monetary commodities.

Keywords: Time series models, Cryptocurrency, Price prediction, Prophet, ARIMA, ARMA.

I. Introduction

In simple terms, when any data is directly related to time or it moves according to time then that data is referred to as time series data. Just like regression models are used for economic growth, stocks, price prediction, etc. [1], time series models also can perform these tasks but with better efficiency due to their focus on parameters related to time dependency. These parameters are helpful for their time series analysis. Some of these parameters are-

Trends: the long-term pattern in data that shows the movement of a series to higher or lower values over a long period of time.

Seasonality: the short-term changes that are generally in the form of a repeating frequency with fixed frequency (with respect to time like an hour, week, month, etc).

Cyclicity: the rises or falls in data without any fixed frequency or duration.

Noise: the random variations that do not follow any particular pattern in the time series.

Stationarity: it occurs when the data of the time series does not depend on the time when it is observed.

To predict such time-series data, there are specialized time-series models available. These models utilize the above-mentioned parameters into account to give future predictions. It is generally believed that the data needs to be stationary to get better predictions from the Time

Series Models but there are some models present today that work despite certain outliers.

In our research we have chosen to compare time series models by forecasting the prices of three monetary commodities; Bitcoin cryptocurrency (BTC), Indian Rupee (INR), and Gold. These three are selected because each of them has different parameters and properties concerning time series data, which are discussed in the Proposed Methodology section.

Cryptocurrency is primarily a collection of data that is used as a medium of exchange. Coin ownership is stored in a public ledger using cryptography techniques. This ledger can track every individual transaction. Cryptocurrency does not exist in physical form. It is decentralized hence there is no single authority that controls it. The first cryptocurrency was Bitcoin which was founded in 2008 [2]. Bitcoin remains the best cryptocurrency as it leads the overall cryptocurrency market. Since Bitcoin is the best & the biggest cryptocurrency that exists, all the other cryptocurrencies available on the market have been termed “altcoins”. Cryptocurrencies are being used to buy real products or services but it is primarily used by people to invest like they would do in stocks or other assets.

Just like Cryptocurrencies, other monetary commodities are apt for price prediction as well. Gold is yet another commodity that is traded across the globe. People invest in Gold on both long-term and short-term basis. Despite being a very huge market, Gold prices are considered unpredictable because they depend on a lot of factors. Another commodity is Indian Rupee which directly relates to India’s growth worldwide. If Rupee’s value will increase in comparison to the US Dollar then it’ll mean India will move ahead as a nation since USD is considered a standard for comparison. A comparison among these commodities can also shed some light on how different these commodities are.

The time series models selected are; Prophet, ARIMA, and ARMA. Facebook’s Prophet is considered the state-of-the-art model for time series forecasting due to its ability to avoid some of the limitations associated with other models. These limitations can be related to seasonality, trends, error, etc. [3], while ARIMA and ARMA are the classical statistical models for the same. It is evident that for long-term price forecasting Prophet has outperformed other time series models but during our research, we have found that these classical models outperformed Prophet in short-term price forecasting. We have tried to figure out the possible reasons for these results and presented them in this paper.

II. Literature Survey

Forecasting using time series models, especially statistical models like ARIMA and ARMA has been done quite a lot in the past. Researchers have focused on either one particular model or they have taken into account one monetary commodity like Cryptocurrency or Stocks. Some of the researchers have tried to compare multiple commodities of the same category for instance [4], [5], [6], and [7] have all compared multiple cryptocurrencies using time series models. In [8], the researchers have introduced and outlined the advantages of a state-of-the-art Time Series Model known as Prophet. In [9], the researchers forecasted Bitcoin price using several models including Prophet and ARIMA and they concluded that while both models performed equally well, ARIMA performed fairly better than Prophet. In [10], the researchers forecast the prices of stocks using ARIMA and LSTM. They concluded that ARIMA outperformed LSTM, this further indicates that Time Series Models are better suited for price forecasting given the nature of these models and their ability to handle the data with time dependencies.

Several researchers have tried to predict cryptocurrency or stock prices using other Machine Learning algorithms as well as using modified models whereas our research paper focuses on Time Series Models as they are. In [11], the researchers have tried to make a modified model using Neural Network and Autoregressive properties of Time Series Models to predict Bitcoin prices. In [12], the researchers have predicted future prices of Bitcoin & Ethereum using several traditional Machine Learning algorithms like Naive Bayes, Random Forest, etc. Their research was focused on comparing how Bitcoin is different from other cryptocurrencies such as Ethereum when it comes to future price prediction.

III. Time-Series Models

There are many Time Series models available for the prediction of future price values. As stated before, we have chosen three models for this research.

ARMA: It is one of the most popular time series models available today. It is simply a combination of AR & MA with AR being autoregressive & MA being moving average. It combines the best of both worlds to provide an output that is better than results produced individually by AR & MA models. ARMA cannot stationarize the data itself, so a modified ARMA-GARCH model is used to achieve stationarity, or a more advanced ARIMA model is used [13].

AR or Autoregressive Model uses past observations as inputs to different regression equations to predict the next value. Since the present value is dependent on some previous past values, the value doesn't change too drastically. It uses PACF (Partial Correlation Function) to find a correlation between t & t-k. PACF will determine the order "p" which is primarily the number of significant values available in the given data. PACF will calculate lags of all the previous time values concerning the current time and the significant resultant values will be the order of the AR model. MA or Moving Average model computes errors or residuals of past time series that have an impact on the current time value to compute future values in the series. In very general terms, we can say that it states that the next observation is the mean of every past observation. The unexpected factors at time periods t-1, t-2, t-2, etc., that impact the current time period value t is referred to as "Errors" or "Residuals". MA uses ACF (Auto-correlation Function) to find the correlation between t & t-k as ACF considers all past observations irrespective of their impact on future or present time periods. It takes into account all the lags available in the data. The ACF values that are more significant are then taken into consideration by MA for analysis.

ARIMA: It stands for Auto-Regressive Integrated Moving Averages Model. ARIMA comprises 3 major elements namely p, d & q where p represents several auto-regressive terms in the model, q represents the number of moving average terms and d represents the total number of differences that have been applied for estimating values. When a simple regression is used to estimate the current value of data, a single lag is used. This is not useful when the data contains multiple lags that have to be looked upon for making the prediction. So, ARIMA is used when multiple lags are taken into account. ARIMA model has 2 major types and they are Seasonal ARIMA and Non-Seasonal ARIMA [14]. ARIMA can be expressed by equation (1) [15].

$$Y_t = c + \Sigma(\varphi_i * Y_{t-i}) + \Sigma(\theta_j * \varepsilon_{tj}) + \varepsilon_t \quad (1)$$

where:
Y_t: The time series value at time t,
c: A constant (optional),
φ_i: Coefficients of the autoregressive (AR) terms for lags i,
Y_{t-i}: The time series value at time t-i,
θ_j: Coefficients of the moving average (MA) terms for lags ,j
ε_t: The error term (white noise) at time t, and
ε_{tj}: The error term (white noise) at time t-j

ARIMA is very similar to ARMA but the major difference is in the "I" which stands for integrated. This "I" represents the ability of the ARIMA model to achieve stationarity. This is like a pre-processing phase to stationarize the time series data if needed. The need for stationarity arises since forecasting is difficult because time series do not provide any certainty to predict correct values. The chances of correct prediction can improve greatly if the statistical properties of data remain similar in the future when compared to their past values. This can be achieved by a stationary time series. A stationary time series whose statistical properties such as mean, variance or auto-correlation, etc remain constant over time. The majority of statistical forecasting methods assume that a time series can be rendered approximately stationary after some mathematical transformation. ARIMA does this via "differencing" which is denoted by "d" and is also known as the degree of difference. The factors left after differencing are termed residuals. These residuals may sometimes contain data that can be used to improve forecasting. So, a thorough diagnostic check is performed where all the left residuals are taken into account and it is checked whether the residuals have any correlation with the model. If they don't then the residuals are left alone but if they do then these residuals become part of the stationary data as shown in Figure 1. ARIMA has proven to be very accurate in cryptocurrency price prediction by studies conducted [16].

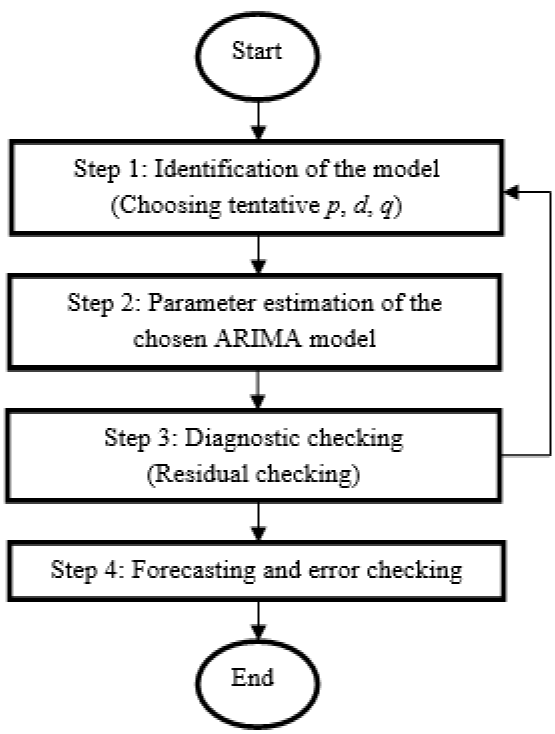

Figure 1: Forecasting model of ARIMA [17]

Prophet: It is open-source software released by Facebook's core data science team. It is a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, & daily seasonality along with holiday effects. Prophet works best with time series that have strong seasonal effects & several seasons of historic data.

The equation for Prophet forecasting is understood by equation (2) [18]:

$$y(t) = g(t) + s(t) + h(t) + et \quad (2)$$

where:

y(t): the additive regression model,

g(t): trend factor,

s(t): seasonality factor,

h(t): holiday factor, and

et: error term

The prophet is used for accurate forecasting and planning, It is very fast as it generates results within seconds. Prophet does not require data pre-processing & works with missing data as well as with several outliers. Users can tweak forecasts by manually adding domain-specific knowledge. Prophet is used in many applications across Facebook for producing reliable forecasts, planning, and goal setting. It has usually outperformed all its competitors in the majority of cases. Prophet can adjust and make changes to its forecast. It can take human interpretation as a parameter to improve the forecast and expand the knowledge domain. Prophet is optimized for business forecasts related to Facebook like time, daily, and weekly observations of history, within a year, trend changes, missing observations, huge outliers & trends that have a non-linear growth curve. Prophet model is well suited for cryptocurrency price prediction because of the nature of cryptocurrencies. Although cryptocurrency data does not have seasonality they are highly volatile and this volatility creates a lot of outliers. Prophet is built for problem statements like these & hence it is apt for volatile data such as cryptocurrency. The basic workflow of Prophet is shown in Figure 2.

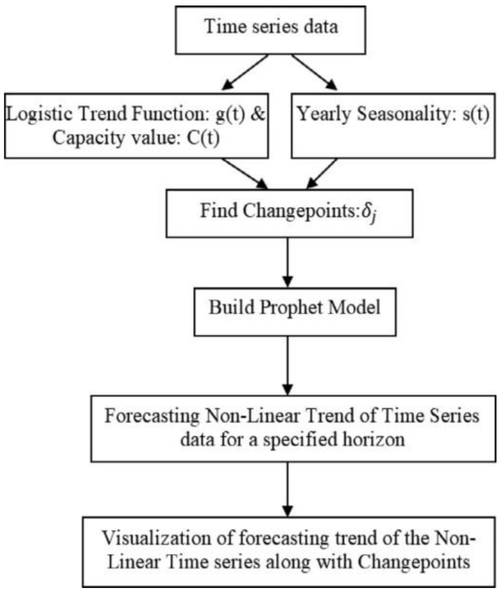

Figure 2: Workflow diagram of Prophet [19]

IV. Proposed Methodology

As shown in Figure 3, we have taken three separate datasets for three different monetary commodities. These datasets contain daily closing prices of Cryptocurrency (Bitcoin), Gold (in INR), and INR value (in comparison to USD) ranging from 2014-2022. We then performed Exploratory Data Analysis (EDA) on these datasets where we took into account several time series factors such as trend, seasonality, cyclicity, and stationarity.

The time series in the cryptocurrency dataset showed an upward trend, it was cyclic, stationary, and non-seasonal. In the case of Gold, there was a downward trend, and the time series was seasonal, cyclic as well as stationary. Lastly, in the case of INR, the time series was cyclic, non-seasonal, and non-stationary, and showed an upward trend. Augmented Dickey-Fuller test was performed on all the datasets which confirmed that all three datasets are non-stationary. The ADF values and p-values suggested that despite being non-stationary, the datasets were not too volatile.

After EDA, the datasets were split to predict prices for the next 60 days. Time series models were then trained and results were evaluated for these datasets. After this, a detailed comparative analysis was performed on the time series models and amongst the monetary commodities as well.

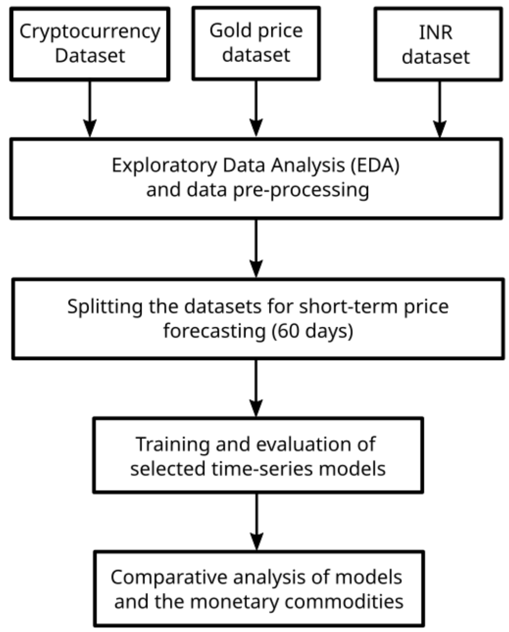

Figure 3: Proposed methodology

V. Experimental Results

Forecasting results are calculated based on certain parameters known as "errors".
They simply showcase the difference between actual values and predicted values. These errors are then used to calculate the accuracy of the models. Errors taken into account for this research paper are:

RMSE (Root Mean Square Error): RMSE provides the measure of the average magnitude of the errors between predicted and actual values. RMSE is one of the most popular metrics to evaluate error percentages in forecasts and predictions. It is calculated by taking the square root of the average of the squared differences between predicted and actual values, hence a lower RMSE value results in a better-performing model [20]. The formula of RMSE is shown in equation (3).

$$RMSE = \sqrt{[\Sigma (y_i - \hat{y}_i)^2 / n]} \quad (3)$$

MAPE(Mean Absolute Percentage Error): MAPE provides the measure of the relative accuracy of predictions or forecasts by calculating the average percentage difference between predicted and actual values. MAPE is expressed as a percentage and a smaller MAPE indicates a more accurate predictive model. The formula of MAPE is shown in equation (4)

$$MAPE = (1 / n)\, \Sigma\, |(y_i - \hat{y}_i) / y_i| \times 100\% \quad (4)$$

Table 1 shows MAPE, RMSE, and Accuracy scores for the time series models used in the research paper.

Table 1: Forecast results

Metrics	Cryptocurrency price forecast			Gold price forecast			INR price forecast		
	Prophet	ARIMA	ARMA	Prophet	ARIMA	ARMA	Prophet	ARIMA	ARMA
RMSE	23108	6126	5390	6	5	3	7	3	7
MAPE	56.73	13.82	11.88	3.91	4.29	1.9	9.44	1.9	6.83
Accuracy	43.27	86.18	88.12	96.09	95.71	98.01	90.56	98.01	93.17

Figure 4 showcases a visualization of accuracy scores for all three-time series models used.

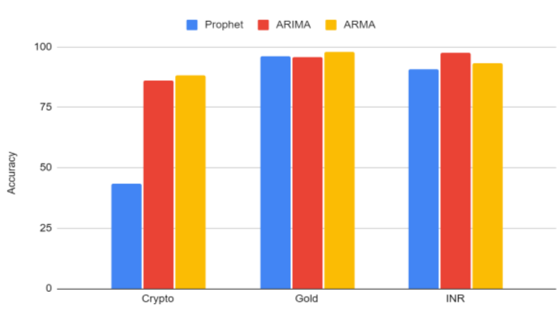

Figure 4: Comparison of forecast accuracies

VI. Conclusion

Although all models performed fairly well, it is evident from the results that ARMA and ARIMA outperformed PROPHET by a fair margin in almost all categories. PROPHET did not perform as well as the other two statistical models in the case of Cryptocurrency while it also did not come on top in the case of Gold and INR. Despite being a state-of-the-art and modern Time Series Model, the results showed that it failed to perform better than traditional Time Series Models such as ARMA and ARIMA. It is known that PROPHET requires a large amount of data to perform well which is not possible in the case of certain monetary commodities such as Cryptocurrencies as it was only introduced in 2008. This is one of the major reasons why it performed poorly in the case of Bitcoin.

ARMA was the overall winner when these models were compared despite being the oldest and most basic of all three models. ARMA outperformed PROPHET and ARIMA in the case of Cryptocurrency and Gold. ARMA performed well in the case of INR price prediction as well. ARIMA provided the best accuracy in the case of INR price prediction. The datasets taken into account were non-stationary but the Dickey-Fuller test performed on them while EDA suggested that they were not too volatile and that may be the reason why ARMA performed better than a theoretically better model like ARIMA or a state-of-the-art model like Prophet. There was no seasonal component in the case of the INR dataset but there was a cyclic component present in it, this factor made it more volatile and that's possibly why ARIMA came on top in the case of INR. Prophet, while performing quite well in the case of INR and GOLD, fell short against ARMA and ARIMA simply because it was overfitting in comparison to the nature of the data.

References:

1. Ray, Susmita. "A quick review of machine learning algorithms." In *2019 International conference on machine learning, big data, cloud and parallel*

computing (COMITCon), pp. 35-39. IEEE, 2019.

2. Ramya, N. "Crypto-Currency Price Prediction using Machine Learning." In *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 1455-1458. IEEE, 2022.
3. Ali, Syed Zishan, Pooja Bhargava, Ranjana Soni, and Shivanshi Tiwari. "A Novel Approach for Analytical Comparison of Cryptocurrencies using Time Series." In *2022 International Conference on Electronics and Renewable Systems (ICEARS)*, pp. 1904-1908. IEEE, 2022.
4. Indulkar, Yash. "Time series analysis of cryptocurrencies using deep learning & fbprophet." In *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, pp. 306-311. IEEE, 2021.
5. Patel, Mohil Maheshkumar, Sudeep Tanwar, Rajesh Gupta, and Neeraj Kumar. "A deep learning-based cryptocurrency price prediction scheme for financial institutions." *Journal of information security and applications* 55 (2020): 102583.
6. Kolla, Venkata Ravi Kiran. "Forecasting the Future of Crypto currency: A Machine Learning Approach for Price Prediction." *International Research Journal of Mathematics, Engineering and IT* 7, no. 12 (2020).
7. Pintelas, Emmanuel, Ioannis E. Livieris, Stavros Stavroyiannis, Theodore Kotsilieris, and Panagiotis Pintelas. "Investigating the problem of cryptocurrency price prediction: a deep learning approach." In *Artificial Intelligence Applications and Innovations: 16th IFIP WG 12.5 International Conference, AIAI 2020, Neos Marmaras, Greece, June 5–7, 2020, Proceedings, Part II 16*, pp. 99-110. Springer International Publishing, 2020.
8. Taylor, Sean J., and Benjamin Letham. "Forecasting at scale." *The American Statistician* 72, no. 1 (2018): 37-45.
9. Tiwari, Raj Gaurang, Ambuj Kumar Agarwal, Rajesh Kumar Kaushal, and Naveen Kumar. "Prophetic analysis of bitcoin price using machine learning approaches." In *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)*, pp. 428-432. IEEE, 2021.
10. Agarwal, Shiva, and Naresh Babu Muppalaneni. "Stock Market Price and Cryptocurrency Price Prediction." In *2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, pp. 1-6. IEEE, 2022.
11. Al-Haija, Qasem Abu. "Time-Series Analysis of Cryptocurrency Price: Bitcoin as a Case Study." In *2022 International Conference on Electrical Engineering, Computer and Information Technology (ICEECIT)*, pp. 49-53. IEEE, 2022.
12. Wimalagunaratne, Minul, and Guhanathan Poravi. "A predictive model for the global cryptocurrency market: A holistic approach to predicting cryptocurrency prices." In *2018 8th International Conference on Intelligent Systems, Modelling and Simulation (ISMS)*, pp. 78-83. IEEE, 2018.
13. Feng, Wenjun, Yiming Wang, and Zhengjun Zhang. "Can cryptocurrencies be a safe haven: a tail risk perspective analysis." *Applied Economics* 50, no. 44 (2018): 4745-4762.
14. Yenidoğan, Işil, Aykut Çayir, Ozan Kozan, Tuğçe Dağ, and Çiğdem Arslan. "Bitcoin forecasting using ARIMA and PROPHET." In *2018 3rd international conference on computer science and engineering (UBMK)*, pp. 621-624. IEEE, 2018.
15. Parab, Leenata Jaymin, and Prashant P. Nitnaware. "Evaluation of Cryptocurrency coins with Machine Learning algorithms and Blockchain Technology." In *2022 IEEE Region 10 Symposium (TENSYMP)*, pp. 1-5. IEEE, 2022.
16. Khedr, Ahmed M., Ifra Arif, Magdi El-Bannany, Saadat M. Alhashmi, and Meenu

Sreedharan. "Cryptocurrency price prediction using traditional statistical and machine-learning techniques: A survey." *Intelligent Systems in Accounting, Finance and Management* 28, no. 1 (2021): 3-34.

17. Abu Bakar, Nashirah, and Sofian Rosbi. "Autoregressive integrated moving average (ARIMA) model for forecasting cryptocurrency exchange rate in high volatility environment: A new insight of bitcoin transaction." *International Journal of Advanced Engineering Research and Science* 4, no. 11 (2017): 130-137.
18. De Leon, Louise Gabriel N., Rafael C. Gomez, Martin Lance G. Tacal, Jonathan V. Taylar, Verlyn V. Nojor, and Alonica R. Villanueva. "Bitcoin Price Forecasting using Time-series Architectures." In *2022 International Conference on ICT for Smart Society (ICISS)*, pp. 1-6. IEEE, 2022.
19. Dash, Sujata, Chinmay Chakraborty, Sourav K. Giri, and Subhendu Kumar Pani. "Intelligent computing on time-series data analysis and prediction of COVID-19 pandemics." *Pattern Recognition Letters* 151 (2021): 69-75.
20. Lyu, Haoran. "Cryptocurrency Price forecasting: A Comparative Study of Machine Learning Model in Short-Term Trading." In *2022 Asia Conference on Algorithms, Computing and Machine Learning (CACML)*, pp. 280-288. IEEE, 2022.