Limitation of ARIMA in extremely collapsed market: A proposed method

Conference Paper · February 2019

DOI: 10.1109/ECACE.2019.8679216

2 authors:

Md. Mujibur Rahman Majumder
Feni University
8 PUBLICATIONS **0** CITATIONS
SEE PROFILE

Md. Imran Hossain
Comilla University
14 PUBLICATIONS **10** CITATIONS
SEE PROFILE

2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 7-9 February, 2019

Limitation of ARIMA in extremely collapsed market: A proposed method

Md. Mujibur Rahman Majumder
Department of Information & Communication Technology
Comilla University
Cumilla, Bangladesh
mujib.contact@gmail.com

Md. Imran Hossain
Department of Information & Communication Technology
Comilla University
Cumilla, Bangladesh
sujan_bgd@yahoo.com

***Abstract*— The prediction of equity market is a perplexing task because of puzzling nature of the stock price. A large variety of factors influence the price of stocks that causes the investors in trouble to predict the nature of stock. Researchers proposed various methods to forecast the upcoming price of stocks by figuring out the nature of stock and by computing internal and external factors. Auto Regressive Integrated Moving Average (ARIMA) Box-Jekins method is one of the eminent methods that forecast the future value of stock based on previous time series data. But while experimenting, we have found a crack point where ARIMA showed unstable behavior i.e. it returned 3 types of values (fixed, negative and positive) in prediction. This was after the prices of stock extremely collapsed. To solve this problem, we have proposed a new model that also predicts the future prices from previous prices but obtained greater accuracy than the ARIMA and also solve the negative and fixed value prediction problem occurred in ARIMA.**

Keywords— Time series forecasting; ARIMA; Stock market prediction; Data Mining;

I. INTRODUCTION

A stock market is gathering of vendee and vendor where investors buy and sellout securities such as shares, bonds and other monetary apparatus of public companies.

Stock market are 2 kinds based on the movement of the market-

1. Bull Market
2. Bear Market

The concept of bull and bear market relates to the percentage rise and fall of any financial market. When price of stocks, currencies, bonds and monetary apparatus are rising systematically or prospective to upturning is called *bull market*. It is considered that, if rising of an equity market >20% considered as bull market [1]. From another source, the percentage value of raising is said to 15% [2].

Bull market occurs when supply is less, hereby increases the demand of product. This is also occurred if corporate profit and GDP of a country is increased. One can earn profit from bull market if he/she can predict the bottom and pick point of the bull market and buy and sell accordingly.

If price of stocks, currencies, bonds and monetary apparatus are decreased systematically or prospective to downfall is called *bear market*. It is considered that, if declination of an equity market is <20%, it is called bear market [3]. From another source, the percentage value of falling is said to 15% [2].

Investor sells the stock rather than buy in case of bear market situation. At bear market situation, demand is less for a product but supply is greater. This is also occurred if biz profit is reduced as well as inert economy and employment are downward in a country.

Stock price prediction is the process of ordain the oncoming worth of stocks on an equity exchange. Researchers have proposed various method by analyzing behavior of market, factors influencing the domestic market as well as sentiment of stock from people to predict stock price so that investor profited from this sector. The successful forecasting of a share's oncoming price is profitable business for the investors. However, wrong prediction may cause misfortune in their life.

Autoregressive Integrated Moving Average (ARIMA) is one of the widely used method used by the investor in predicting stock price. However, while research, we have found that ARIMA couldn't predict properly in extremely collapsed market. For a large number of stocks, it predicts negative and fixed values. To solve this problem, we have proposed a new model, which predict future value based on previous sequential data of stock for the extremely collapsed market.

We have collected the equity data for our research from the Dhaka Stock Exchange (DSE), Bangladesh. At the end of 2011, price of stocks in DSE were declining to 90% for most of the stocks. That's why we choose these data for our research. We have verified our proposed model by 20 stocks of DSE and compare the result with ARIMA.

In this paper, we have arranged the related work at chapter II, Theory related to this research is given at chapter III, proposed methodology at chapter IV, result and discussion at chapter V and a few wrapping up comments at chapter VI.

978-1-5386-9111-3/19/$31.00 ©2019 IEEE

II. Related work

ARIMA model has extensively used in forecasting at various discipline of forecasting.

Qonita et al. [4] implemented the ARIMA forecasting model over currency exchange rate data. They forecasted the Indonesian rupiah against the global currency US dollar and they had noticed that ARMA (2,2) gives maximum accuracy which is 98.74%.

Adebiyi et al. [5] used ARIMA model into the stock exchange data and forecast New York stock exchange (NYSE) and Nigeria stock exchange's (NSE) future value. They selected the best model by computing and comparing BIC, adjusted R^2 and S.E of regression to find best fitted model.

Siregar et al. [6] dispensed ARIMA Box-Jekins method in prolepsis the raw material of PP Trilene and PP Tintapro plastic products. They receipt an accuracy of 74% for product PP Trilene and 68% for PP Tintapro products.

Nwanko Steve [7] dispensed ARIMA Box-Jekins method over Nigerain exchange (Naira to Dollar) rate data. He found best accuracy for Box-Jekins AR (1) model.

AK Pinar & Yumusak [8] forecasted the consumption of household natural gas of Turkey by using ARIMA (p,d,q) Box-Jekins method. Authors proposed a merged model over traditional model and showed that Mean Absolute Prediction Error (MAPE) was reduced to 2.2% than traditional non-merged model.

Dong et al. [9] had used time-oriented ARIMA model in detecting the traffic flow of road. Author's experiment showed that the time oriented ARIMA model in preferable than the non-oriented model.

Turno et al. [10] proposed a new model to determine the best fit order of ARIMA. In ARIMA Box-Jekins model, best fitted model was calculated form ACF and PACF value. However, they determined the model based on the over-fitting concept.

Haider [11] forecasted the DSE (Dhaka Stock Exchange) return by using ARIMA. He had used ARIMA model for predicting index series and return series and finds that ARIMA (3,1,2) were best fitted model for index series and ARMA (3,1) for return series data.

III. Background theory

The word ARIMA stands for Autoregressive Integrated Moving Average. This model is also familiar by the name of Box-Jenkins models. This model is the combination of Auto Regressive (AR) and Moving Average (MA) model and it comes from the term ARMA(p,q) model. Autoregressive model of p order is known as AR(p) and moving average model with *lagged forecast error q* is known as MA(q ***)*** [12].

ARIMA also known as ARIMA (p,d,q) equation where d is the differences (non-seasonal) needed for stationarity. The general equation of ARIMA can be written by using (1) which is combination of AR part and MA part.

$$\hat{y}_t = \mu + P_1 y_{t-1} + \cdots + P_p y_{t-p} - \theta_1 \epsilon_{t-1} - \cdots - \theta_q \epsilon_{t-q} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

That means, the predicted value is the weighted sum of one or more recent values and/or recent value of error plus a constant.

ARIMA can be classified based on the value of p,d and q. When the value of p is 1 (that means 1 dependence between the term or simple $\mathrm{Yt} - 1$ in the equation) , then it is called *first order autoregressive model* and denoted by ARIMA (1,0,0). It can be written as -

$$\hat{y}_t = \mu + P_1 y_{t-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

For 2nd order it contains another term, y_{t-2} also [13].
When y is non-stationary then the possible model is called random walk model and often causes long term drift in y. It is also said ARIMA (0,1,0) model and equation can be written as [13]–

$$\hat{y}_t = \mu + y_{t-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

In our research, we have used a common method of selecting ARIMA best fit model which is famous by Hyndman-Khandakar algorithm [14]. Hynman- Khandakar algorithm works by combining the unit root test and best fit model is selected from the concept that containing minimum AICc and BICc value.

IV. Proposed Method

Our working procedure consists of following steps-

1. *Read pre-processed weekly stock data:* In this research, weekly stocks data are used. The property of time series forecasting is that it only predicts over ordered data, thus have to order the dataset and filling out the missing values.

2. *Divide data into train set and test set:* To check the model, dataset divided into 2 sets-
    - Train set
    - Test set

    Train set contains all stock data before market collapsed and 1st collapsed data. And remaining are fallen into test. In research, data from 1st January 2010 to 31th December 2012 have used. Market collapse occurred at December 4, 2011. So train set consists data from January 1,2010 to December 4,2011 (1st collapsed data). Remaining data are fallen into test set.

3. *Compute Absolute Percentage Change (APC) of data of train set:* At this point, Absolute Percentage Change (APC)

have been computed between 2 consecutive close price of train set. 1st computes the difference or Change (C) between 2 consecutive Close Price.

$$C = Close_i - Close_{i-1} \ldots\ldots\ldots\ldots\ldots (4)$$

After that, absolute Percentage Change (APC) is computed by using (5).

$$APC_i = \left| \frac{C}{Close_i} \times 100 \right| \ldots\ldots\ldots (5)$$

4. *Compute Mean Absolute Percentage Change subset (MAPCS) of train set:* At this time, the mean value of all possible subset form from train set and store those value in subset array (X_j) is computed. This can be explained by using (6). Where S is the size of test set. The value of j is ranges from 1 to size(train)-S. We have considered this values in predicting because after a stock market is extremely collapsed, it walks steadily as customer neither make any buy and sell onto market at that time. Thus, we have used the range of Absolute Percentage Change (APC) having minimum mean value.

$$X_j = \frac{\sum_j^{j+S} APC_j}{S} \ldots\ldots\ldots\ldots\ldots (6)$$

5. *Determine minimum MAPCS:* Now the subset which has minimum Mean Absolute Percentage Change Subset (MAPCS) value is determined. And we have calculated the initial value and end value of subset which is denoted by P and Q respectively. This can be done by using (7) and (8). Here, S is the size of test (This is actually the total number of observation we want to predict).

$$P = MINIMUM(X_j) \ldots\ldots\ldots\ldots (7)$$
$$Q = P + S \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (8)$$

6. *Compute Predicted value (p):* The predicted value (p) is computed by multiplying last value of train (1st collapsed close price:$Close_N$) with minimum Mean Absolute Percentage Change Subset (MAPCS) values. This can be written as equation (9).

$$PredVal\ (p) = Close_N \times APC(P:Q) \ldots\ldots (9)$$

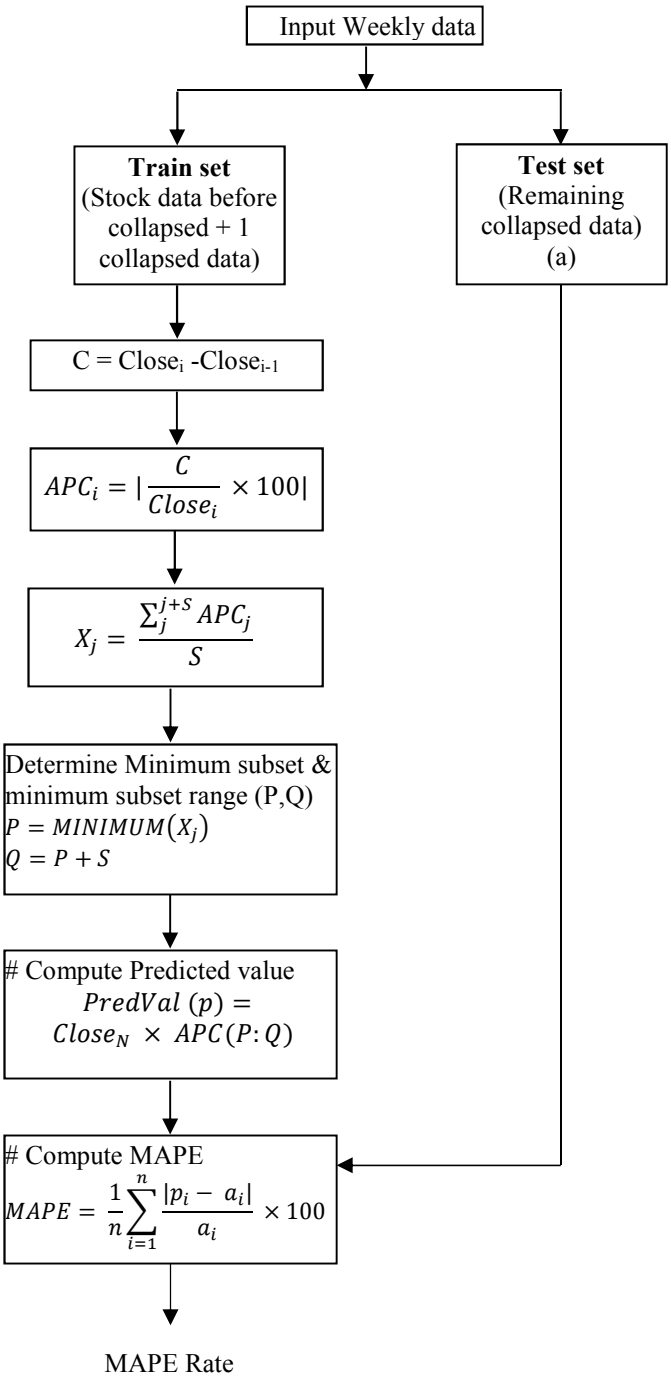

Figure 1: Flow chart of proposed model.

V. RESULTS AND DISCUSSION

To verify this model, Mean Absolute Percentage Error (MAPE) is computed. The equation of MAPE can be written by using (10).

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \frac{|p_i - a_i|}{a_i} \times 100 \ldots\ldots\ldots\ldots\ldots (10)$$

Where p_i is the value of predicted stock price, a_i is the actual close price of stock in day i and n is the total number of days.

We have tested our model for 20 collapse stock of DSE (Dhaka Stock Exchange). And by applying this model,we have predicted next 7 months (30 observation) close price of stock.

It is noticed that best fit model returns fixed value for all 30 observations in 14 stocks, positive (regular) value for 2 stocks and negative value for 4 stocks. But is not accepted that any model returns negative and fixed value as predicted price.

In this study, this proposed model is better than ARIMA (1,0,0) model in all 3 cases (positive, fixed & negative value as predicted value by ARIMA best fit) and better than ARIMA best fit model in 2 cases (positive & negative value as predicted value by ARIMA best fit) but MAPE value of our model is slightly greater an ARIMA best fit in fixed value forecasting. However, this value is also acceptable as prediction accuracy is almost 83%.

The important properties of this model is that there is no chance of this model to predict any negative value. Comparison of this model with ARIMA is depict in Table 1, Table 2 and Table 3.

a. When ARIMA best fit model return fixed value as predicted value

Table 1: Performance analysis when best fit return fixed value

Stock Name	MAPE for ARIMA(1,0,0) model	MAPE for best fitted model	MAPE (Proposed)
Nitol Ins	81.63	6.88	7.98
One Bank	82.06	27.66	28.72
Paramount Ins	14.97	10.15	11.44
Pharma Aids	90.97	29.24	29.29
Reliance Ins	80.95	7.93	8.72
Rupali Bank	93.13	13.06	13.36
Singer BD	80.18	14.18	14.32
Sonargaon	12.6	16.46	17.36
South East	84.67	19.17	19.28
Standard Bank	84.3	16.71	16.95
Tallu Spinn	72.19	23.24	24.79
Titas Gas	77.63	12.82	13.19
Trust Bank	83.09	20.26	21.32
United Ins	80.11	13.51	15.00
MEAN	**72.74**	**16.52**	**17.27**

b. When ARIMA best fitted model return regular(positive) value as predicted value

Table 2: Performance analysis when best fit return positive value

Stock Name	MAPE for ARIMA(1,0,0) model	MAPE for best fitted model	MAPE (Proposed)
Olympic Ins	20.84	20.84	35.68
Renata	89.86	89.86	18.86
MEAN	**55.35**	**55.35**	**27.27**

c. When ARIMA best fitted model return false(negative) value as predicted value

Table 3: Performance analysis when best fit return negative value

Stock Name	MAPE for ARIMA(1,0,0) model	MAPE for best fitted model	MAPE (Proposed)
National Poly	94.58	321.50	17.66
Pragati Life	93.81	448.86	12.58
Progressive Ins	93.4	146.78	21.56
Usmania Glass	94.37	473.16	12.90
MEAN	**94.04**	**347.58**	**16.17**

At table 1, we only keep the MAPE value of stocks which gives fixed prediction value in ARIMA best fit model. From this, it is clear that proposed model is better than ARIMA (1,0,0) as the average value of MAPE for proposed model is 17.27% very much less than ARIMA (1,0,0) which is 72.74%. But slightly greater than ARIMA best fit model which is 16.52%.

At table 2, we only keep the MAPE value of stocks which gives positive prediction value in ARIMA best fit model. From this, it is clear that proposed model is better than ARIMA (1,0,0) as well as ARIMA best fit model. The average MAPE for proposed model is 27.27% and 55.35% for both ARIMA (1,0,0) and best fit model. Both value is same as ARIMA (1,0,0) was the best fit model for those stocks.

At table 3, we only keep the MAPE value of stocks which gives negative prediction value in ARIMA best fit model. From this, it is clear that proposed model is better than both ARIMA (1,0,0) and ARIMA best fit model. The average MAPE for proposed model is calculated to 16.17% and 94.04% for ARIMA (1,0,0) and 347.58% (very poor) for ARIMA best fit model.

VI. Conclusion

From the result, it can be said that, ARIMA model is not work well at extremely collapsed stock market as it predicts fixed and negative price in its best fitted model prediction. But our model is works fine than ARIMA with accuracy of almost 80% at that situation.

However, this model is worked fine when we want to predict price after a market extremely collapsed but

performance may be degraded other than collapsed market. But our model is better than ARIMA as it is not predict fixed and negative value in its prediction.

REFERENCES

[1] https://www.investopedia.com/terms/b/bullmarket.asp

[2] https://investinganswers.com/financial-dictionary/stock-market/bull-market-922

[3] https://www.investopedia.com/terms/b/bearmarket.asp

[4] Adiba Qonita, Annas Gading Pertiwi and Triyanna Widiyaningtyas , "Prediction of Rupiah Against US Dollar by Using ARIMA", 19-21 September 2017, Proc. EECSI 2017, Yogyakarta, Indonesia.

[5] Ayodele A. Adebiyi, Aderemi O. Adewumi, and Charles K. Ayo, "Stock Price Prediction Using the ARIMA Model", 2014, UKSim-AMSS 16th International Conference on Computer Modelling and Simulation.

[6] Baihaqi Siregar, Erna Budhiarti Nababan, Alexander Yap, Ulfi Andayani and Fahmi, "Forecasting of Raw Material Needed for Plastic Products Based in Income Data Using ARIMA Method", 2017, 5th International Conference on Electrical, Electronics and Information Engineering (ICEEIE).

[7] Nwankwo Steve C, "Autoregressive Integrated Moving Average (ARIMA) Model for Exchange Rate (Naira to Dollar)", July 2014, Academic Journal of Interdisciplinary Studies MCSER Publishing, Rome-Italy, Vol 3 No 4, ISSN 2281-3993

[8] Mustafa Akpinar and Nejat Yumusak, "Forecasting Household Natural Gas Consumption with ARIMA Model: A Case Study of Removing Cycle", 2013, 7th International Conference on Application of Information and Communication Technologies. DOI: 10.1109/ICAICT.2013.6722753

[9] Honghui Dong, Limin Jia, Xiaoliang Sun, Chenxi Li and Yong Qin, "Road Traffic Flow Prediction with a Time-Oriented ARIMA Model", 2009, Fifth International Joint Conference on INC, IMS and IDC

[10] Turno, Subanar, Dedi Rosadi and Suhartono, " New Procedure for Determining Order of Subset Autoregressive Integrated Moving Average (ARIMA) Based on Over-fitting Concept", 2012, International Conference on Statistics in Science, Business and Engineering (ICSSBE), DOI: 10.1109/ICSSBE.2012.6396643

[11] ASM Shakil Haider, "Forecasting Dhaka stock exchange (DSE) return: An Autoressive Integrated Moving Average (ARIMA) approach", 2009, North South Business Review, Volume 3, Number 1, ISSN 1991-4938.

[12] Md. Kamruzzaman, Md. Mohsan Khudri, and Md. Matiar Rahman, "Modeling and predicting stock market returns: A case study on dhaka stock exchange of Bangladesh",2017, Dhaka Univ. J. Sci. 65(2): 97-101

[13] https://people.duke.edu/~rnau/411arim.htm

[14] Hyndman, R. J., & Khandakar, Y. "Automatic time series forecasting: The forecast package for R. Journal of Statistical Software", 2008, 27(1), 1–22. https://doi.org/10.18637/jss.v027.i03