



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** IV **Month of publication:** April 2022

DOI: <https://doi.org/10.22214/ijraset.2022.41733>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Stock Market Indices Direction Prediction: Time Series, Macro Economic Factors and Distributed Lag Models

Divyaanshu Das

The International School Bangalore

Abstract: A number of studies have been conducted to model the stock market indices using pure time series models or regression models based on macro economic variables. In this study, instead of focusing on modeling the actual levels of stock market indices I focus on predicting the direction (up/down) as investors who rely on technical analysis are more interested in the direction of stock market index than the actual prediction value. Therefore, in this study I look at best modelling approach for the direction prediction: time series (ARMA) or macro factor models or combination of both (ARDL). My study shows that macro factor models outperform for direction prediction as compared to ARMA or ARDL models. The study was performed on stock market direction prediction of stock indices of three South Asia countries: India, Pakistan and Malaysia. The macro economic factors that are considered for direction prediction are: Inflation, Unemployment and Exchange Rate monthly data from March 2016 to September 2021.

Keywords: stock, equity, prediction, models, direction, ARMA, ARDL, macro

I. INTRODUCTION

In general, developing a predictive model involves: (a) taking the known data (historical time series or observed independent/exogenous data), (b) developing a model based on optimization of some cost/error function and then (c) using that developed model to predict new observed data which occur in the future. There are number techniques to build the prediction models and generally, the best model is the one which minimises the prediction error on test data. However, in some cases the stakeholders are more interested in getting the direction of prediction (up/down) than the actual predicted values. This is particularly true for exotic derivatives trading such as binary option where the traders are more interested in getting the stock market direction prediction accurate than actual values/levels. Therefore, in this study, I want to study which models/techniques such as time series or regression based on macro economic variable or combination of both are best suited for direction prediction. In this study I have chosen stock market indices of three South Asian countries: India, Pakistan and Malaysia monthly data over 5 years from March 2016 to Sep 2021. Using this observed time series data, prediction models are built using: (i) pure time series model (ARMA), (ii) macro economic factors based regression model and (iii) combination of both (ARDL) and these models are used to test the direction prediction. The macro economic variables selected for this study are: Inflation, Exchange Rate, Consumer Pricing Index and Unemployment monthly data downloaded from the Global Economy [1] website. The study shows that the regression model based on macro economic variables outperforms as compared to time series or ARDL when it comes to direction prediction consistently for all three countries. The paper is organised as follows. The section 2 briefly review the current literature survey, section 3 talks about the data, section 4 describes the methodology for modelling the stock market direction prediction, section 5 discusses the results and finally, in section 6 results are concluded along with the next steps.

II. LITERATURE SURVEY

Stock market indices of any country is largely dependent on number of macro economic factors like interest rates, industrial production, output, money supply, political environment, government stability etc. In the literature, I found large number of studies have been conducted in finding the causal relationship between stock price movements and macro economic variables. Among such studies some notable one are [1] [2] [3] [4] and main conclusions from these studies are that the macro economic factors do influence the stock price movements. The literature survey also reveals that causal relationship between stock price movements and macro economic data varies from one country to another country. For example, in the study conducted by Muhammed and Rasheed [5] which tests the relationship between stock returns and exchange rate they conclude that the causal relationship exists only for Bangladesh and Sri Lanka but not for Pakistan and India.

Moreover, there were large number of studies to predict the stock price movements by using time series models like ARMA & ARIMA. One such study conducted by Mondal et al. [6] tried to predict the stock prices of 56 different stocks from seven sectors by using the time series model like ARMA and concluded that ARIMA model gives prediction accuracy of over 85%. In another study conducted by Erfani [7], they have used Stock Price Index (TISP) to find the long .. In the study conducted by Mahdi [8], they have used the simple ARMA model to test the market efficiency and financial stability of the London Stock Exchange and S&P 500 stock indices using monthly and annual data. In their study they concluded that predicting the month stock return outperforms the prediction of annual returns. In addition, their study also reveals that predicting S&P 500 outperforms the London Stock Exchange.

III. METHODOLOGY

A. Approach for Modelling

In this study to predict the direction of stock mark indices of the selected countries, I have applied three different variations of the multivariate linear regression. The variations of these three models are:

1) *Auto Regressive Moving Average Model (ARMA)*: In this variation, the time series is modelled by considering the two lags. That is obser- vation at time t_n is function of $f(t_{n-1}, t_{n-2})$. This can be mathematically expressed as:

$$y(t_n) = \alpha_0 + \alpha_1 \cdot y(t_{n-1}) + \alpha_2 \cdot y(t_{n-2}) + e(t_n) \tag{1}$$

In the above equation, $y(t_n)$ is the observation at time t_n , and $y(t_{n-1}), y(t_{n-2})$ are observations at times t_{n-1} and t_{n-2} respectively. The term $e(t_n)$ represent the error term. The α terms are regression coefficients which are estimated by minimising the least square error between the ac- tual and predicted value. Passive autocorrelation function plots are used to decide the number of lag terms to include in the ARMA model. One such plot is shown in the Figure. 1 for India’s stock and macroeconomic data.

2) *Macroeconomic Factor based Multivariate Linear Regression Model (ARMA)*: In this variation, the stock market indices is modelled as a function of macro economic variables *inflation rate, unemployment rate and exchange rate*. This can be mathematically expressed as:

$$y(t_n) = \alpha_0 + \alpha_1 \cdot I(t_n) + \alpha_2 \cdot U(t_n) + \alpha_3 \cdot ER(t_n) + e(t_n) \tag{2}$$

In the above equation, $y(t_n)$ is the stock market observation at time t_n , $I(t_n)$ is the inflation, $U(t_n)$ is the unemployment rate and $ER(t_n)$ is the exchange rate of the respective countries for 1 USD. The term $e(t_n)$ represent the error term. The α terms are regression coefficients which are estimated by minimising the least square error between the actual and predicted value.

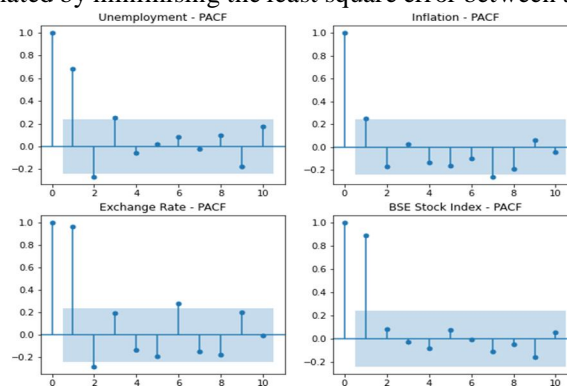


Figure 1: India Time Series Data - PACF Plots.

3) *Auto Regressive Distributed Lag Model (ARDL)*: This variation of the model is combination of 1 and 2 where a regression model is built by combining the lag terms of stock market indices at previous time period (auto regressive) and lag terms of macroeconomic terms (lag terms) at previous time periods. Mathematically, it is expressed as:

$$y(t_n) = \alpha_0 + \alpha_1 \cdot y(t_{n-1}) + \alpha_2 \cdot y(t_{n-2}) + \beta_1 \cdot I(t_n) + \beta_2 \cdot U(t_n) + \beta_3 \cdot ER(t_n) + e(t_n) \tag{3}$$

In the above equation α terms are the regression coefficients of the auto regressive lag terms and β terms are the regression coefficients of the macroeconomic factor lag terms estimated by minimising the least square error between the actual and predicted values approach for counting direction prediction and equations comments on Adj R2

B. Approach for Direction Prediction

The following methodology is adopted to test the model accuracy in terms of stock market directional movement prediction.

- 1) The given time series data of 65 terms are divided into training and test data comprising 59 observation for the model fit and 6 observations for testing.
- 2) Ignoring the actual levels of estimation, a directional predict is treated as correct prediction if both the actual and estimated values at t_n and t_{n-1} has the same sign. That is if $\hat{y}(t_n) - \hat{y}(t_{n-1}) > 0$ and $y(t_n) - y(t_{n-1}) > 0$; or $\hat{y}(t_n) - \hat{y}(t_{n-1}) < 0$ and $y(t_n) - y(t_{n-1}) < 0$ have the same signs then the direction prediction is correct otherwise incorrect. In this equation $\hat{y}(t)$ represents the predicted value from the model and $y(t)$ represents the actual value.

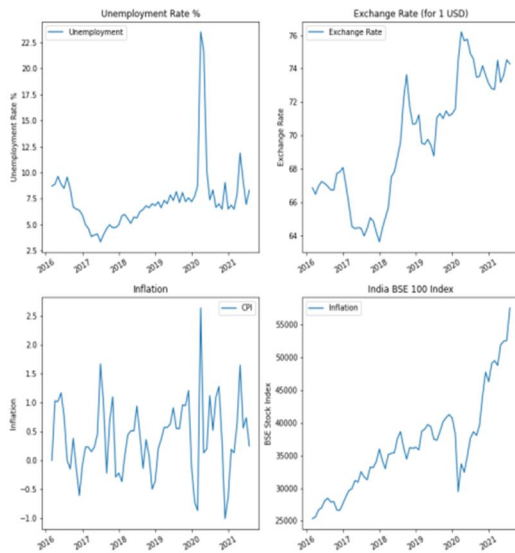


Figure 2: India - Time Series.

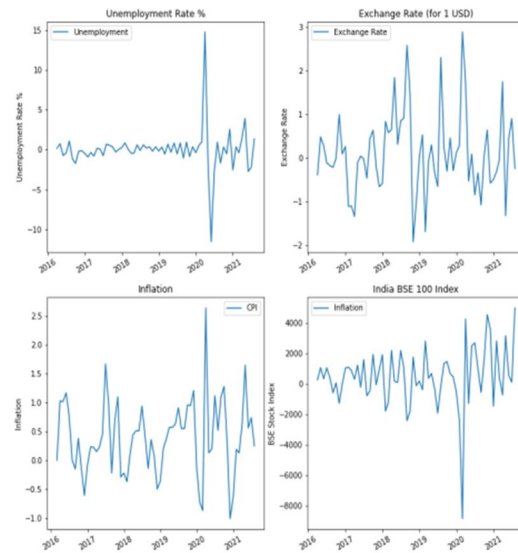


Figure 3: India - Stationary Series.

- 3) The directional prediction is measured over both training and test data for each of the regression model described in the previous section for comparison.

In addition to the direction prediction, R^2 goodness of fit of the regression model is also captured for model comparison.

IV. DATA

The following monthly data covering the period from '01Mar2016' to '01-08- 2021' (a total of 66 monthly observations) are used in this study.

- 1) Monthly stock market indices data of India (BSE 100), Pakistan (KSE 100) and Malaysia (KLCI) biggest stock exchanges downloaded from Yahoo website.
- 2) Monthly macro economic variables Inflation, Exchange Rate, Unemployment Rate and Consumer Price Index downloaded from the GlobalEconomic Data website [10].

The Figures. 14, 4, 6 show that stock market index, unemployment and exchange rate plots have clear trend (either upward or downward) and then applying AD Fuller stationary tests depicts that the data is non-stationary. Therefore, to make the data stationary, I took the first difference of the giventime series as described by the equation below.

$$y(t_n) = y(t_n) - y(t_{n-1}) \tag{4}$$

The plots of the first difference transformed data is shown in the Figure. 16,5, 7 and in these plots we can clearly see that the trend is removed and the

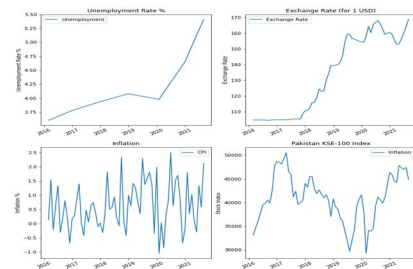


Figure 4: Pakistan - Time Series.

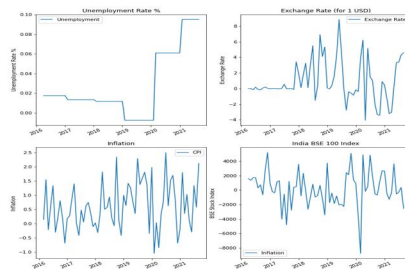


Figure 5: Pakistan - Stationary Series.

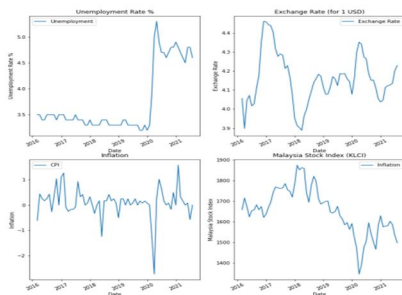


Figure 6: Malaysia - Time Series.

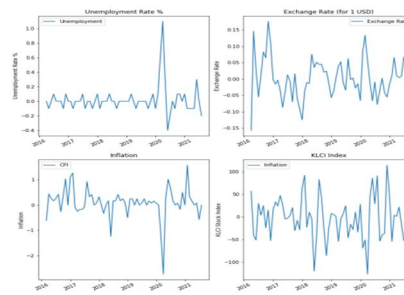


Figure 7: Malaysia - Stationary Series.

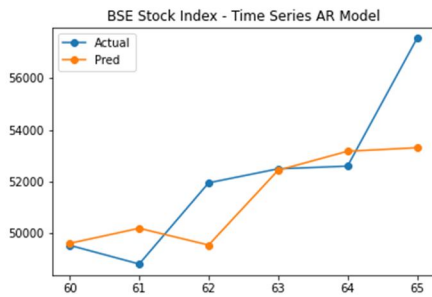


Figure 8: India Stock Prediction ARMA Model..

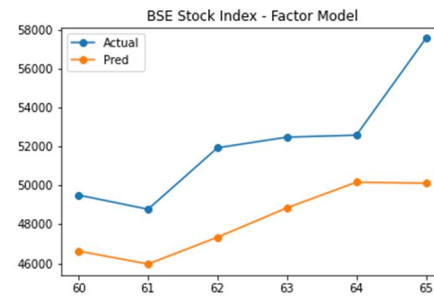


Figure 9: India Stock Prediction Macro Model.

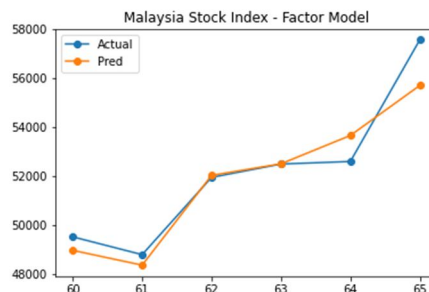


Figure 10: India Stock Prediction -ARDL Model.

Overall mean of the series is around zero. To confirm that the data is truly stationary, the AD Fuller tests are applied on the first difference data and found that the null hypothesis of H_0 the series is not stationary is rejected as the test statistics values are within the 5% critical values.

V. RESULTS

The three linear regression models described in Methodology section were fitted to the stock and macroeconomic data of the three South Asia countries: India, Pakistan and Malaysia. The results of $AdjR^2$ and stock market direction prediction results for each model and country is tabulated below.

Mode	Training Data (Direction Prediction)			Test Data (Direction Prediction)		
	Correct	Incorrect	% Success	Correct	Incorrect	% Success
ARMA	33	26	56%	3	2	60%
Macro	42	17	71%	4	1	80%
ARDL	39	20	66%	5	0	100%

Table 1: India Stock Market Direction Prediction [!h]

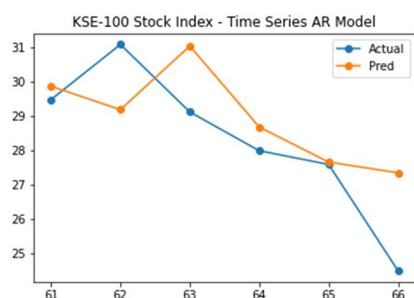


Figure 11: Pakistan Stock Prediction - ARMA Model..

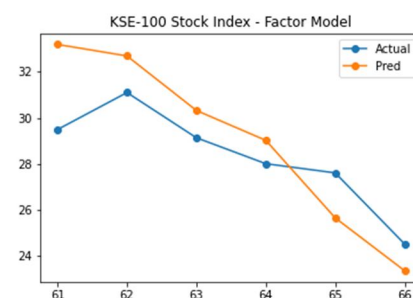


Figure 12: Pakistan Stock Prediction - Macro Model.

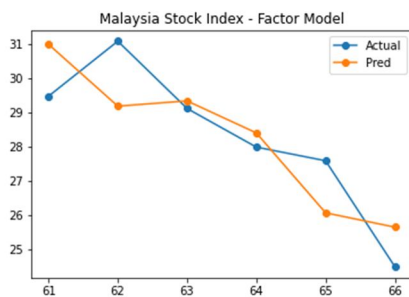


Figure 13: Pakistan Stock Prediction -ARDL Model.

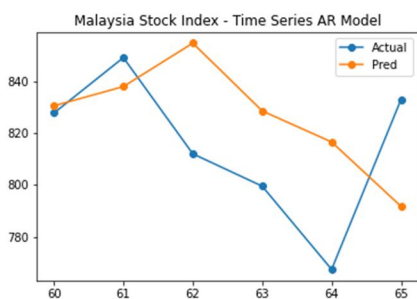


Figure 14: Malaysia Stock Prediction -Figure 15: Malaysia Stock Prediction - ARMA Model..

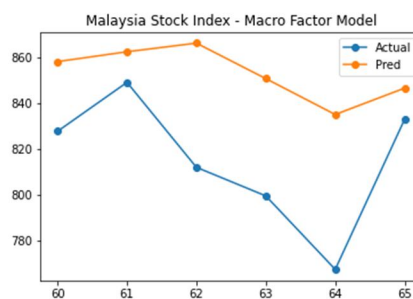


Figure 15: Malaysia Stock Prediction - Macro Model.

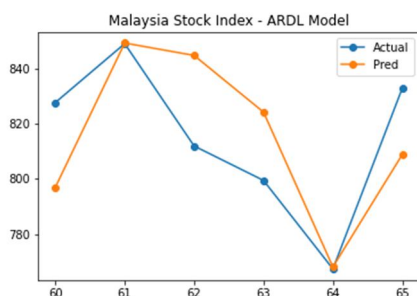


Figure 16: Malaysia Stock Prediction -ARDL Model.

Mode	Training Data (Direction Prediction)			Test Data (Direction Prediction)		
	Correct	Incorrect	% Success	Correct	Incorrect	% Success
ARMA	34	25	58%	3	2	60%
Macro	38	21	64%	4	1	80%
ARDL	36	23	61%	4	1	80%

Table 2: Pakistan Stock Market Direction Prediction

Mode	Training Data (Direction Prediction)			Test Data (Direction Prediction)		
	Correct	Incorrect	% Success	Correct	Incorrect	% Success
ARMA	32	27	60%	3	2	60%
Macro	40	19	64%	4	1	80%
ARDL	39	20	66%	5	0	100%

Table 3: Malaysia Stock Market Direction Prediction

Index	Adj. R^2		
	ARMA	Macro	ARDL
India	0.92	0.83	0.94
Pakistan	0.97	0.91	0.97
Malaysia	0.66	0.41	0.77

Table 4: Goodness of model fit Adj. R^2 results

VI. CONCLUSIONS

The results summarised in Table. 4 shows that Auto Regressive Distributed Lag (ARDL) model outperforms the regression fit consistently as compared to ARMA or the regression model that considers only macroeconomic variables for all three countries stock market indices. This conclusion logically make sense because the ARDL model has all the information about the stock market levels from the previous periods and also the macroeconomic information to predict the stock market levels for the next period. On the other hand, when the regression model was fit purely on macro economic data the goodness of fit measures drops consistently for all three countries stock market indices. From these tables we can conclude that if the focus of the modelling is in predicting the actual levels of stock market indices then best models are ARDL as first choice and then ARMA model as second choice. This also implies that previous levels of stock market indices is very important to predict the levels for the future periods.

The results summarised in Tables. 1, 2, 3 shows that for stock market di- rection prediction, the ARDL models outperforms as compared to regression model that considers only macro economic data or ARMA model.



This conclusion is also consistent because ARDL model has all the information (current levels and macro data) to predict the stock market directional movement for the future periods. From these results we can conclude that if the focus of the modelling is to predict the directional movement the ARDL is the first choice followed by macroeconomic model.

REFERENCES

- [1] Fama E. F. Stock returns, real activity, inflation and money. 71(4):45–565, 1981.
- [2] Fama E. F. Stock returns, expected returns and real activity. 45(4):1089–1108, 1990.
- [3] Fama E. F. and French L. Business conditions and expected prices on stocks and bonds. 25:23–49, 1989.
- [4] Ross. stock market title. 4(2), 2014.
- [5] Naeem Muhammad and Abdul Rasheed. Stock prices and exchange rates: Are they related? evidence from south asian countries. 41(4):535–550, 2002.
- [6] Labani Shit Prapanna Mondal and Saptarsi Goswami. Study of effectiveness of time series modeling (arima) in forecasting stock prices. 4(2), 2014.
- [7] Alireza Erfani and Ahmad Jafari Samimi. Long memory forecasting of stock price index using a fractionally differenced arma model. 5(10):1721–1731, 2009.
- [8] Mohammad Mahdi Rounaghia and Farzaneh Nassir Zadeh. Investigation of market efficiency and financial stability between s&p 500 and london stock exchange: Monthly and yearly forecasting of time series stock returns using arma model. 456(10):10–21, 2016.
- [9] J. M. Smith and A. B. Jones. *Book Title*. Publisher, 7th edition, 2012.
- [10] Neven Valev. *Business Economic Data for 200 Countries*. The Global Economy, <https://www.theglobaleconomy.com/>.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)