JOURNAL OF CRITICAL REVIEWS

ISSN- 2394-5125 VOL 07, ISSUE 12, 2020

Forecasting of COVID-19 Cases in India with ARIMA Models

Dr. Sachi N. Mehta
Assistant Professor
Government Arts and Commerce College, Ahwa, Dang
Dr. Manish R. Patel
Assistant Professor, Department of Economics, VNSGU, Surat.

Abstract
A worldwide danger has been identified as Coronavirus Disease 2019 (COVID-2019), and many research have been carried out on different mathematical models to estimate how likely an epidemic could evolve. These mathematical models are susceptible to possible prejudices, depending on many variables and analyses. Here, we attempt to forecast the spread of COVID-2019 a simple econometric model. (Christian Democrats, 2020) This research was based on confirmation case reports from COVID-19 received daily from the www.MoHFH.gov.in official website between 6 March 2020 and 30 April 2020. The ARIMA model has been utilised for prediction of trends in confirmed cases. The outcome of the Dickey-Fuller increased unit root test indicates that the verified COVID-19 instances are stationary when the initial difference is taken into account. The ARIMA Automatic Prediction indicates that ARIMA was for India (4,2,0). This implies that there are 4 AR delays and 2 MA lags (Moving Average). The findings indicate that throughout this predicted time the instances of Confirm COVID-19 will continue to rise. We may estimate that will perhaps rise by 173 percent on 31 May 2020. The tendency for confirmed COVID-19 cases will deteriorate if the administration takes severe measures such as lockdown, curfew and quarantine. It will reach about 90,000 by the end of May.
Key words: coronavirus, auto-regression, moving average, prediction, moving average,

1. Introduction

COVID-19 is a new, newly identified Coronavirus infectious illness. Before the epidemic started in Wuhan, China, in December 2019, the novel virus and illness had not been recognised. Therefore, the new Coronavirus was named. In 2019, it was discovered. COVID-19 are viruses that belong to a wide family and cause sickness from the common cold to more serious illnesses. Coronavirus illness comprises severe ARS (SARS-CoV) and North East Core (MMERS-CoV) and was called the 2019-nCoV (Dehesh et al, 2020) Virus responsible for this disease was declared by the World Health Organisation or W.H.O. as the name of this disease and the 'COVID-19' virus. The W.H.O designated COVID-19 as a worldwide health emergency in

ISSN- 2394-5125 VOL 07, ISSUE 12, 2020

accordance with its transmission rate. Recent surveys by the Center for Disease Control or C.D.C. show that COWID-19 has, to date, impacted numerous nations worldwide. Everywhere throughout the globe. In the 30th of April 2020 there was confirmed a total of 33,18,739 cases including 2,34,264 deaths. This illness's prevalence is unknown since the prevalence of the condition is so dynamic at now. Epidemiological monitoring and detection capability for suspected cases are evidently different across nations. Many nations worldwide have also reported various instances of COVID-19 infections. Where there is no special therapy for the disease, disease prevention and medical services prepare for it is extremely essential.

On 30 January 2020 in India the first instance of the coronavirus pandemic was recorded in China. As of 30 April2020, total of cases, 34.866 recovery and 1154 deaths in the nation have been verified by the Ministry of Health and Family Welfare. Experts believe that there may be considerably greater number of illnesses, because India is among the lowest test rates worldwide. The infection rate in India with COVID-19 in the worst afflicted nations is estimated to be considerably lower. On 22 March 2020, at Prime Minister Narendra Modi's request, India conducted a 14-hour voluntary public curfew. In 75 areas where COVID instances occurred, as well as in all major cities, the authorities followed up with lock-downs. On 24 March, the Prime Minister also announced 21 days of national lock-up impacting the whole population of 1.3 billion people in India. On April 14, the current national lockout was extended by the Prime Minister until May 3. In its study, based on data from 73 nations, the Oxford Covid-19 government response tracker states that the Indian government has responded to the epidemic more rigorously than any other country. It highlighted quick action by the Government, emergency policymaking and healthcare urgencies, budgetary measures, vaccine research funding and an aggressive reaction, and gave India a "100" mark in terms of stringency.

The treatment system may be able to provide services for new patients by modelling and future prediction of the daily confirmed case number. Models of statistical prediction may be used to anticipate and manage the worldwide pandemic danger. The Auto-Regressive integrated moving average (ARIMA) model may usefully be used in this research to forecast COVID-2019 confirmed cases. This model has more ability in the prediction of natural calamities than other prediction models like the Wavelet Neural Network (WNN) and the SVM.

2. Methods:

2.1 Source of data

Between 6 March and 23 April 2020, confirmed daily COVID-2019 cases were gathered via the official website of the Ministry of Health and Family Welfare (MoHFW), and a time series data base was constructed using Microsoft Excel.

2.2 Models for ARIMA

ARIMA denotes an integrated autoregressive moving average – also sometimes referred to as a Box-Jenkins A series of timelines that are examined through ARIMA should be evaluated if they are stationary or non-stationary. If additional processes are not to be used, time-series data should be stationary in order to execute ARIMA.

ISSN- 2394-5125 VOL 07, ISSUE 12, 2020

2.2.1 Test of Stationarity

Stationarity implies that the characteristics of a time series do not rely on the time of observation of the series. In addition, time series data variation across time should always be consistent. If data in a time series are not stationary, the technique of difference may be used to stabilise existing data. Instead of time series data itself, the time series difference is utilised. In the following equations the first and second differential technique presented may be seen:

$$D\ t = xt - xt\text{-}1$$

$$D\ t = (xt - xt\text{-}1) - (xt\text{-}1 - xt\text{-}2)$$

In addition, the components of ARIMA need be grasped in order to understand them. There are two components ARIMA has for moving average and autoregressive models.

2.2.2 Autoregressive Models (AR)

Any observation X(t) can be explained by some function of its preceding observation X(t-1) plus the error variable, E (t). Thus, having X(t-1) and all the other necessary constants and figures derived from time series, it is possible to forecast the X(t) value. The following equation is a formula for the AR time series analysis.

"X(t) = A (1) * X(t-1) + E(t)"

The equation demonstrates an AR-process with one lag, such that the time series is only analysed on the basis of the preceding series. It's an AR (1) procedure, therefore. As indicated below would be an AR (2) procedure with a second lag:

"X(t) = A (1) * X(t-1) + A (2) * X(t-2) + E(t)"

2.2.3 Moving Average (MA)

ARIMA's second component is the movement of the average models. The distinction between the autoregressive model and the average moving model is that the moving average model focuses more on error constant of past observations or sometimes referred to as earlier lags. In the Moving Average models, for example, the error constants such as E(t-1), E(t-2), E(t-2) are evaluated in place of X(t-1), X(t-2) X(t-3). The following is an example of a moving average equation:

"X(t) = - B (1) * E(t-1) + E(t)"

B (1) is an MA of order 1 as the preceding equation indicates and is multiplied by the lag one term of error. Consequently, model X(t), the expected future value, is always subject to the error term according to the moving averages. An equation of two lags would be using the same logic:

$$X(t) = - B(2) * E(t-2) – B(1) * E(t-1) + E(t)$$

2.2.4 ARIMA

Now that both models are understood, proceed to the ARIMA (Box – Jenkins) mixed model. In order to provide a precise prediction, ARIMA combines both equations. The example of an ARIMA model might be the ARIMA model order of (1,1,1) describing the AR-1-order term, the first time-series differences examined and the first-order MA-Term. Briefly, the criteria and demands for the ARIMA model are stationarity in the first place, then application of the difference, unless there is a stable, secondly constant variance across time.

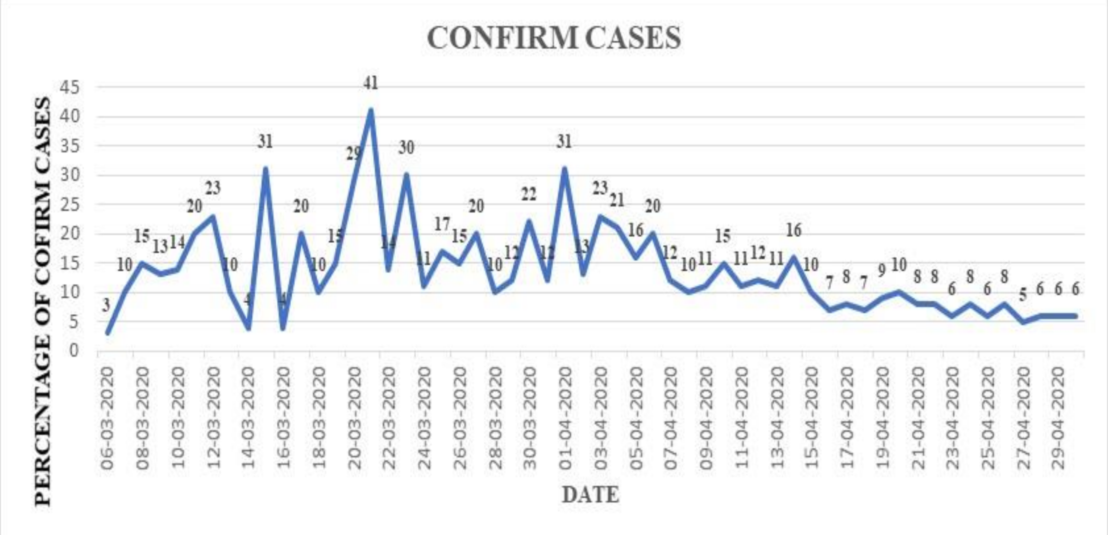
ISSN- 2394-5125 VOL 07, ISSUE 12, 2020

3. RESULTS-:

3.1 Confirm COVID-19 Case Trends percentage.

The percentage trends of COVID-19 are highly fluctuating in Figure I. In supplication, it was only raised by 3 percent than day after day. It rose by 41 percent on 22 March 2020 and this is the greatest growth throughout the whole time. The trends in the COVID-19 confirmation % have risen after 22 March 2020, but decreased. The results of a government effort and last week's lock-out have been raised by 6 percent on average to 7 percent, which shows the government's plan has been somewhat successful.

Figure I: Percentage Trends to Confirm COVID-19 Cases.



3.2 Stationary test result:

The sequence of integration of a time series variable is critical. To identify the sequence of integration of the series, we first conduct unit root tests in levels and initial differences. We use the Augmented Dickey-Fuller (ADF) test to test integration order.

Table I shows that the computed ADF statistics for level confirm COVID–19 instances less than the critical levels, indicating that the variable is not level stationery. To corroborate COVID–19 instances, ADF statistics indicate first-difference stationery.

Table: I Result of Unit Root Test Using Augmented Dickey-Fuller Test

Variable	At Level		At First Difference		Conclusion
	ADF	Prob.	ADF	Prob.	
CONFIRM COVID - 19 CASES	-0.9470	0.3011	-7.691	0.0000	I (1)

3.3 Result of Automatic ARIMA Forecasting

Table II shows the ARIMA technique of 4 lags and 2 lags. This implies there are 4 AR lags and 2 MA lags (Moving Average). The AIC (Akaike information criteria) score is 1.501725. This is the best outcome of all the combinations. As seen in table III.

Table: II Automatic ARIMA Forecasting Summary

Automatic ARIMA Forecasting Selected dependent variable: DLOG(CONFIRM_CASES) Date: 05/01/20 Time: 13:20 Sample: 3/06/2020 4/30/2020 Included observations: 55 Forecast length: 31
Number of estimated ARMA models: 25 Number of non-converged estimations: 0 Selected ARMA model: (4,2)(0,0) AIC value: 1.50172504603

Figure II: Forecast Comparison Graph

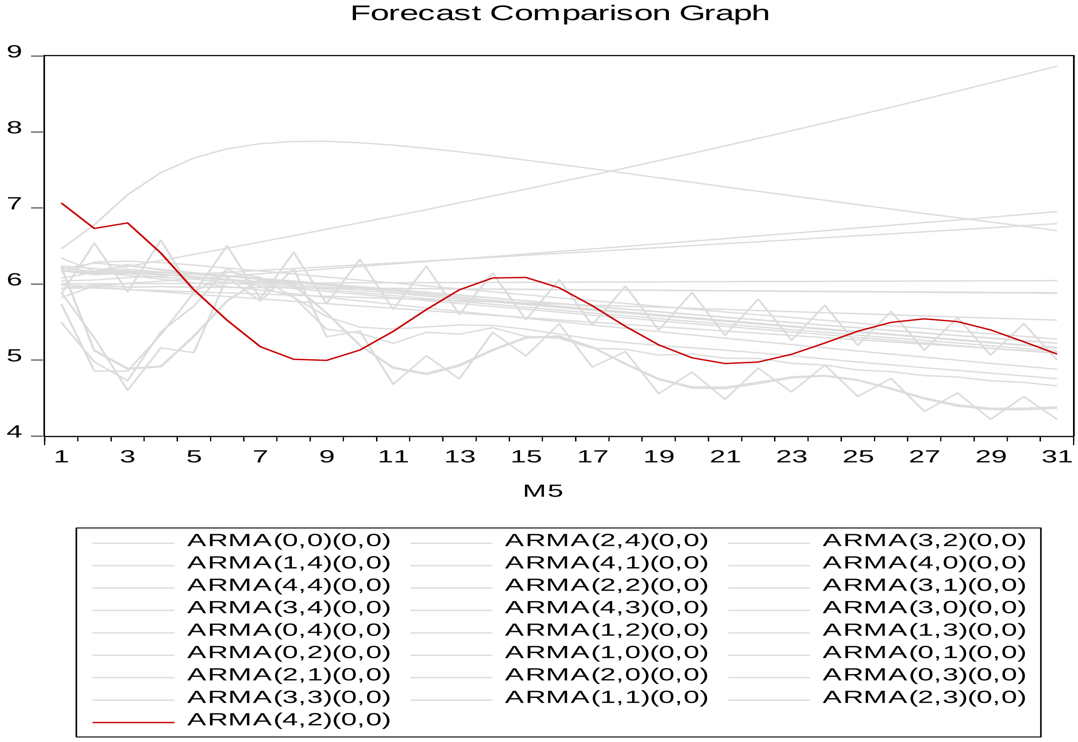


Figure II compares ARIMA lags. The red line depicts the selected lags.

3.4 Data forecasting

Table III shows that the Confirm COVID-19 instances will now be 31 days (from May 1st, 2020 to May 31th 2020). Based on the projected results, it will continue to rise. On May 31, 2020, it may rise to 173 percent. Figure III depicts this.

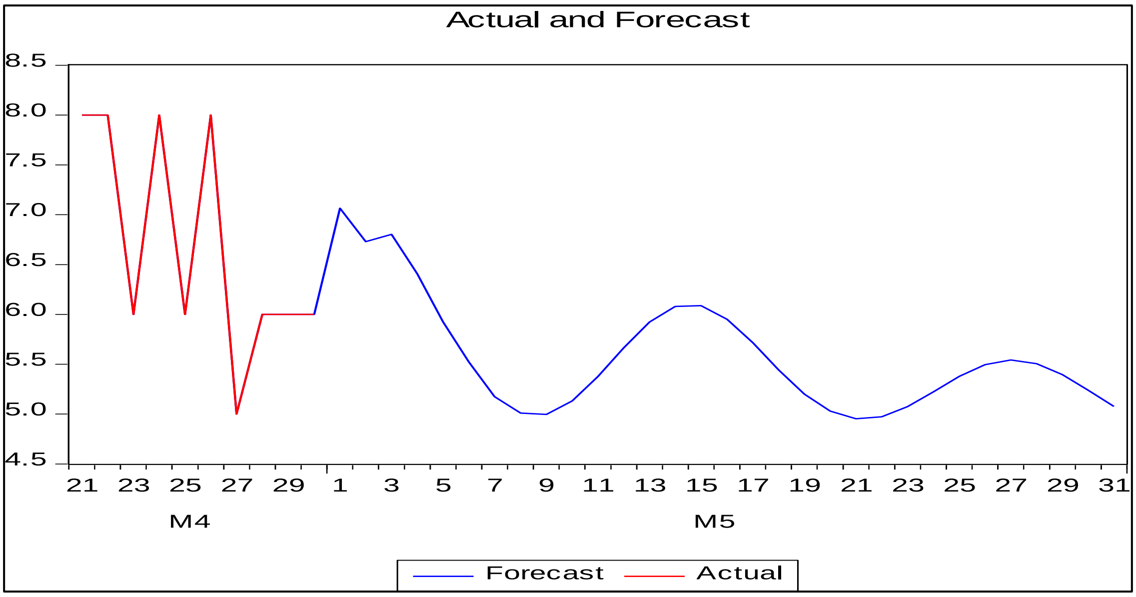
Table: III COVID-19 Confirm Cases forecast from 20-04-2020 to 30-04-2020

DATE	Forecast Confirm COVID-19 Cases	DATE	Forecast Confirm COVID-19 Cases
01-05-2020	7%	17-05-2020	6%
02-05-2020	7%	18-05-2020	5%
03-05-2020	7%	19-05-2020	5%
04-05-2020	6%	20-05-2020	5%

ISSN- 2394-5125 VOL 07, ISSUE 12, 2020

05-05-2020	6%	21-05-2020	5%
06-05-2020	6%	22-05-2020	5%
07-05-2020	5%	23-05-2020	5%
08-05-2020	5%	24-05-2020	5%
09-05-2020	5%	25-05-2020	5%
10-05-2020	5%	26-05-2020	5%
11-05-2020	5%	27-05-2020	6%
12-05-2020	6%	28-05-2020	6%
13-05-2020	6%	29-05-2020	5%
14-05-2020	6%	30-05-2020	5%
15-05-2020	6%	31-05-2020	5%
16-05-2020	6%	TOTAL	173%

Figure: 3 Forecast Comparison Graph



4. Conclusion:

COVID-19 is an infection caused by a newly identified Coronavirus. No one knew about this novel virus and illness until an epidemic in Wuhan, China, in 2009. After then, numerous instances from all around the globe were reported to authorities. India is also one of the nations impacted country. According to this research, in India, COVID-19 is expected to spread in the next 31 days (from May 1st, 2020 to May 31st, 2020). From March 6th, 2020 to April 30th, 2020, daily verified COVID-19 instances were gathered from the MoHFW's official website. The trend of confirmed cases was predicted using the ARIMA model. We utilised E-Views 10 to predict COVID-19 in India. The Confirm COVID-19 instances will be extended to 31 days (from May 1th 2020 to May 31st, 2020). During the projected period, it will continue to grow. On May 31, 2020, it may rise to 173 percent. As a result, the number of confirmed COVID-19 cases will decrease if the government imposes severe measures such as lockdown, curfew It will be about 90,000 by the end of May.

This research is helpful since it forecasts the COVID-2019 pandemic, making it a reliable and objective tool for infection control monitoring. All institutions engaged in public health and infection control may profit from this research since they can use the model to predict the COVID-2019 pandemic on a daily basis and with confidence. This research provides a simple method to assess the COVID-2019 transmission patterns and determine if the infection control or quarantine strategy is effective.

The prediction may, however, influence certain research constraints. Some key independent variables (number of daily corona tests) may be omitted from the model, reducing its practical relevance. Although additional data is required, these models may assist forecast future confirmed cases assuming the viral transmission does not alter dramatically. Because of its rarity, this virus has the potential to spread quickly and severely This ability may influence all forecasts, but to our knowledge, this model is the best at the moment.

SELECTED REFERENCES:

Damodar G. Time Series Econometrics: Forecasting. In: Aric Bright, editor. Basic Econometrics, 4th ed. New York: Tata McGraw Hill; 2003.p.835-65.

Francis X. Unit Roots, Stochastic Trends, ARIMA Forecasting Models, and Smoothing. In: Dennis H and Jennifer B. Elements of Forecasting, 3rd ed. New Delhi: Thomson South-western; 2007.p.331-78.

Jeffrey W. Advanced Time Series Topics: In Joe Sabatino, editor. Introductory Econometrics a Modern Approach 5th ed. Mason OH: South-Western Cengage Learning; 2013.p.632-71.

Joseph T Wu, Kathy Leung, Gabriel M Leung. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modeling study. Lancet 2020;395:689-97.

Joseph T, Kathy Leung, Mary Bushman, Nishant Kishore, Rene Niehus, Pablo M, Benjamin J, Marc Lipsitch, Gabriel M. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. J. Clin.Med 2020;26:506-10

Joseph T, Kathy Leung, Mary Bushman, Nishant Kishore, Rene Niehus, Pablo M, Benjamin J, Marc Lipsitch, Gabriel M. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. Nat Med 2020;26:506-10.

Wang W, Tang J, Wei F. Updated understanding of the outbreak of 2019 novel coronavirus (2019- nCoV) in Wuhan, China. J Med Virol 2020;92:441-7.

Ministry of Health and Family Welfare Government of India (MoHFW). Covid-19 India,[updated 23 April 2020].Available from: https://www.covid19india.org/