DOI: 10.21917/ijms.2020.0183

ANALYSIS AND PREDICTION OF THE SPREAD OF COVID-19

F. Ramesh Dhanaseelan[1], A. Oliver Bright[2], M. Jeya Sutha[3] and M.J. Brigit Gilda[4]

[1,3]Department of Computer Applications, St. Xavier's Catholic College of Engineering, India
[2]Department of Business Administration, St. Xavier's Catholic College of Engineering, India
[4]Department of Physics, M.S. University Constituent College of Arts and Science, India

Abstract

Covid-19, the disease caused by a novel Severe Acute Respiratory Syndrome Corona Virus-2 (SARS-CoV-2), is a highly contagious disease. On January 30, 2020 the World Health Organization (WHO) declared the outbreak as a Public Health Emergency of International Concern. As of September 10, 2020; 27891329 laboratory-confirmed and 903991 deaths have been reported globally. Time series models play an important role in predicting the impact of the covid-19 outbreak and taking the necessary measures to respond to this crisis. In this work, Autoregressive Integrated Moving Average (ARIMA) model is employed to forecast the epidemiological trend of COVID-19 prevalence of US, India and Brazil, the most affected countries as of September 10, 2020. The prevalence data of COVID-19 from 31 December 2019 to 10 September 2020 were collected from the WHO website. Different ARIMA models were formed with different ARIMA parameters. ARIMA (1,2,1), ARIMA(2,2,0) and ARIMA(2,2,2) were considered as the best model for forecasting the confirmed cases of US, India and Brazil whereas ARIMA(1,1,1), ARIMA(0,2,1) and ARIMA(2,2,3) were considered as the best model for forecasting the number of death cases in these countries in the next four weeks. Forecasting covid-19 prevalence trend of US, India and Brazil can help health authorities and politics to plan and supply resources effectively as well as make a plan to maintain the stable national economic growth in this terrible situation. It also helpful to the Government and public to take necessary precautions as early as possible.

Keywords:

ARIMA, Time Series, Forecasting, COVID-19, Pandemic

1. INTRODUCTION

In December 2019, there was a pandemic in china which was highly contagious, continuously killing human lives. China was using all sort of internal mechanism and even sought other countries help to bring that epidemic under control. The cause of epidemic was unknown and many experts engaged to find the real name of the disease. Later it was found that it was a typical type of virus and there is no medicine available to cure and contain its spreading. It was identified as corona virus disease, named COVID-19 [10] [21]. Many studies are undertaken to analyse COVID-19 in different angles.

This contagious disease made the whole world to turn their faces and watch astonishingly even though it originated from China [2]. It is considered as one of the worst hit epidemic ever the mankind faced. On 10th September it was reported that 2,78,91,329 were contracted with this disease and 903991 people lost their life around the globe. The World Health Organization exhibited this disease as a "global pandemic". The United States declared a national emergency for their country. The disease started spreading rapidly damaging the health and survival of the people. The whole world realized the danger of losing the life and livelihood. All the countries in the world understood the potential outbreak of this infectious disease which is called as COVID-19. As the social distancing is essential to control the wide spreading, many countries declared the emergency and lockdown. All the establishments were made to close and the distribution of goods and services were awfully affected. It paralyzed the global economy. People are made to be held up in their residences. Many countries announced punishment for violating the social distancing norms. The hospitals and health care centers were flooded with patients contracted of this suspected disease. The positive cases and death toll started shooting up. Hence this study rose up and crucial matter of concern to rejuvenate the health care system so that the confidence of protection against this pandemic will be cultivated in the minds of the people.

In order to analyse and forecast many models and methods may be used statistically to predict the number of cases contracted with this disease which will help to plan the number of additional materials and resources are needed to deal with the outbreak [30]. Estimating the expected burden of disease is essential for public health officials to effectively and timely manage medical care and other resources needed to overcome such disease. It would help to develop the control devices to prevent and control the intensity of the outbreak.

Many researches used different time series models and methods [12] [22] [29], regression models [20], neural network simulation models [15] [18] [28] [16] [13] [26] for the predictions. Such tools were applied to forecast the affected cases of different diseases. Normally the spreading of diseases is due to many factors. There will be many influential factors which will amount the causes of the diseases. Hence general spread of the outbreak is characterized by tendencies and randomness. Therefore the mentioned statistical tool are insufficient to analyse the epidemic randomness and very difficult to generalize.

Many models are used for forecasting based on the trends of the available data. They are used to estimate the outcome of a future period/periods based on the current available values. Many researches have used different forecasting techniques in the field of economics, production, finance and other fields. These models are also used in medical and health care studies also. In the recent studies Autoregressive Integrated Moving Average (ARIMA) model has been used by many researches [4]. ARIMA has been applied in health care field with more accuracy and reliability. It has got a good amount of confidence level and decomposition, exponential smoothing and more reliability. It is one of the best models with more accuracy and simplicity especially in forecasting of seasonal variables. In health care studies ARIMA model has been applied for predicting influenza mortality, malaria incidents, hemorrhagic fever, dengue fever, typhoid fever, tuberculosis, thrombo topenia and many other infectious diseases [9] [6] [8] [14] [28] [25] [31] [7] [17] [5].

Hepatitis E Virus [18], Hantavirus Pulmonary Syndrome [15], Pertussis [27], Hepatitis [24], Hepatitis B Virus [22], Severe

Fever with Thrombocytopenia Syndrome [19], Pertussis [23], Human Brucellosis [26], Pulmonary Tuberculosis [13] and Brucellosis [4] are some of the diseases predicted using ARIMA models.

ARIMA models are emphasizing the temporal dependent formation of a time series, given the changing trends, periodic changes, and random distortions in the time series. It is relatively easy to explain to the end-user since ARIMA methods do not contain much mathematics or statistics. The users can have better idea of how the prediction model could be developed and more reliability in the process of decision making.

As this virus is rapidly spreading in US, India and Brazil as per the statistics given by WHO (https://covid19.who.int/), during 31st December 2019 to 10th September 2020, this period is taken for the study by applying this ARIMA model. This would help predict and estimate the probable cases contracting with this virus and expected depth cases based on the current pattern of spreading.

2. METHODS

2.1 DATA COLLECTION

Since the initial confirmed cases identified vary to countries, the time period chosen for the analysis also varies which are given with their descriptive statistics in Table.1.

Table.1. Descriptive statistics of the dataset used

Type of Dataset	Time Duration considered	Country	Min	Max	Mean	Standard Deviation
Total_ Cases_ Confirmed	21.01.20 - 10.09.20	US	1	6359720	1.9846 e+6	2.0286 e+6
	30.01.20- 10.09.20	India	1	4465863	7.4269 e+5	1.1623 e+6
	26.02.20 - 10.09.20	Brazil	1	4197889	1.2099 e+6	1.367 e+6
Total_ Deaths	01.03.20- 10.09.20	US	1	190815	95045.7	62361.01
	13.03.20- 10.09.20	India	1	75062	19069.6	22108.8
	18.03.20- 10.09.20	Brazil	1	128539	49586.9	43178.8

As of September 10, 2020, the outbreak of COVID-19 had been confirmed in around 215 countries or territories. Fig.1 shows cumulative corona contracted and death cases reported in US, India and Brazil till September 10, 2020. Around 27891329 cases of coronavirus contracted cases and 903991 deaths have been reported worldwide. The cumulative confirmed cases of COVID-19 reported is 6359720 in US, 4465863 in India and 4197889 in Brazil. Meanwhile the apparent mortality rate is 3.06% in Brazil, 3.00% in US and 1.68% in India. Hence, it is essential to analyze the prevailing trend particularly in US, India and Brazil. This would help the countries to take necessary measures to prevent and control the number of persons contracting with this epidemic.

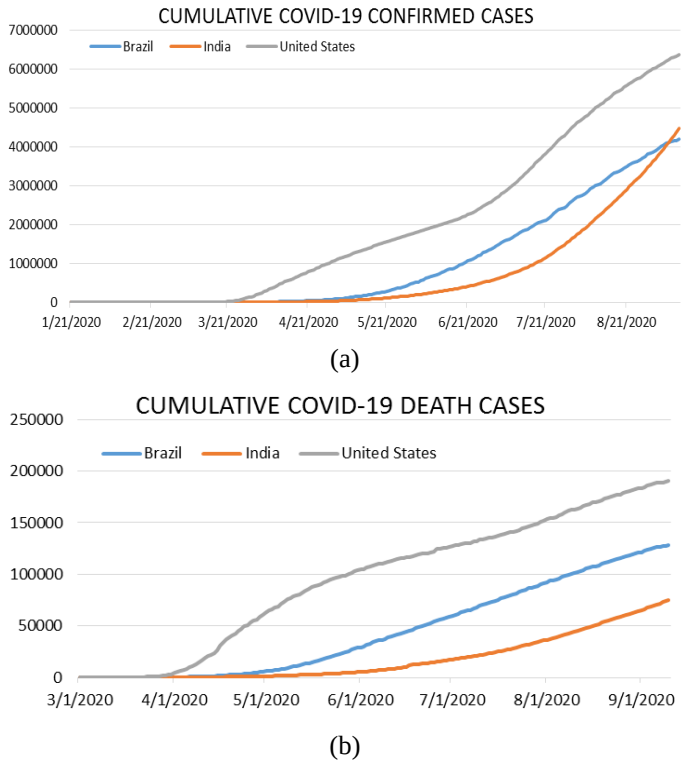


Fig.1. Cumulative COVID-19 cases as on September 10, 2020
(a) Confirmed cases (b) Death cases

2.2 ARIMA MODELS

Time series is a set of data points indexed in time order [3]. Time series analysis is the process of analyzing the time series [1] [11] [14]. It is mostly concerned with identifying three different aspects of the time series which can be used to better clean, understand and forecast the data. It is highly applicable to any kind of data like seasonality, cyclical and fluctuating.

The autoregressive component $AR(p)$ refers the past values for the series Y in the regressive equation, where p denotes the number of lags used. The parameter d in the integrated component $I(d)$ refers the degree of differencing. The moving average component $MA(q)$ refers the error of the model, where q denotes the number of terms included in the model. ARIMA model can be expressed with differencing, autoregressive and moving average component as in Eq.(1):

$$Y_t = c + \emptyset_1 e_{t-1} + \emptyset_2 e_{t-2} + \ldots + \emptyset_q e_{t-q} + e_t \quad (1)$$

2.3 MODEL SELECTION

The accuracy of a model is verified by comparing both the actual values and the predicted values. Mean Absolute Error (*MAE*) and Root Mean Square Error (*RMSE*) are considered as the performance criteria to predict the accuracy of the developed ARIMA models. These are expressed mathematically in Eq.(2) and Eq.(3).

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2} \quad (2)$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n} |e_t| \quad (3)$$

where

n is the number of time points and

e_t is the difference between the estimated and observed values.

Akaike information criteria (AIC) and Bayesian information criterion (BIC) models evaluate and estimate the quality of each model relative to the other. The lower the AIC and BIC, the better the model should be at forecasting. All analyses were done with Python 3.7.4 64-bit software.

3. RESULTS AND DISCUSSION

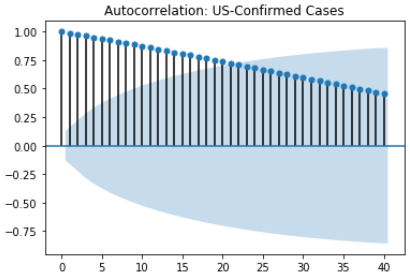
3.1 ARIMA MODEL FOR FORECASTING THE PREVALENCE OF COVID-19 PANDEMIC

The ARIMA model consists of (i) building the model (ii) calculation of parameters (iii) diagnostic checking and (iv) prediction. The first and foremost step is to check whether the dataset is stationary and seasonal. The dataset is called stationary when the statistical properties like mean, variance and autocorrelation are constant over time. Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) show the time series plot in which seasonality and stationary of the dataset can be verified. Estimated autocorrelations and partial autocorrelations for the time series of US, India and Brazil are shown in Fig.2 and Fig.3 respectively.

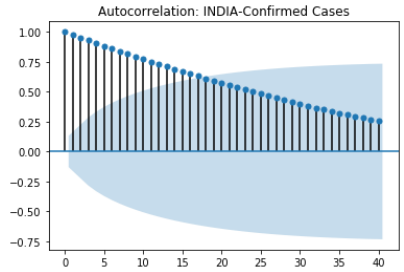
Table.2. Parameters of ARIMA models

Country	Type of Dataset	Best Model	RMSE	AIC	BIC
US	Confirmed_ cases	ARIMA (1,2,1)	6879.28	4623.07	4636.86
	Death_ cases	ARIMA (1,1,1)	374.27	2929.89	2942.94
India	Confirmed_ cases	ARIMA (2,2,0)	3938.69	4099.74	4113.37
	Death_ cases	ARIMA (0,2,1)	97.10	2305.23	2314.81
Brazil	Confirmed_ cases	ARIMA (2,2,2)	11258.43	4052.27	4075.21
	Death_ cases	ARIMA (2,2,3)	192.79	2309.36	2331.51

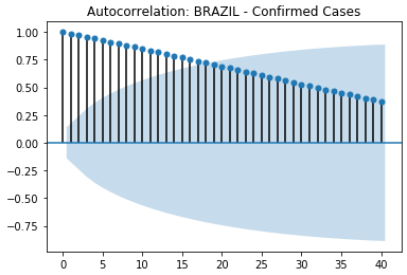
The ACF plots, as in Fig.2, show that the dataset is not stationary, since there is very low decreasing of autocorrelations. Therefore, second-order difference was taken to make all series stationary in the dataset. Different ARIMA models were created and their performance were tested with various statistical tools. Models with least RMSE and MAE values were considered as the best model. Accordingly, ARIMA(1,2,1), ARIMA(2,2,0), ARIMA(2,2,2) models were chosen as the best models for COVID-19 confirmed cases for US, India and Brazil respectively. Similarly, ARIMA(1,1,1), ARIMA(0,2,1) and ARIMA(2,2,3) models were chosen as the best models for COVID-19 death cases for US, India and Brazil respectively. The fitted models for the COVID-19 confirmed cases and death cases for US, India and Brazil are shown in Fig.4. The forecasting values for the COVID-19 confirmed cases and death cases for the next 30 days for US, India and Brazil are shown in Fig.5. As shown in Table 3, US will have 7434912 confirmed cases and 215315 death cases, India may reach 7463000 confirmed cases and 112188 death cases and Brazil will have 5209463 confirmed cases and 152531 death cases in the next 30 days.



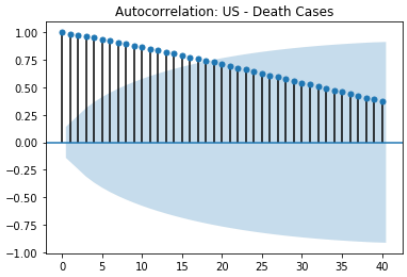
(a)



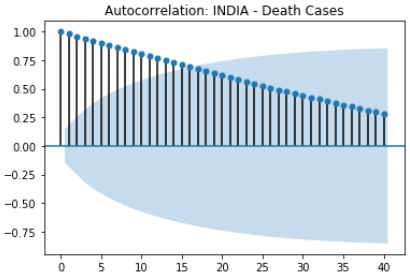
(b)



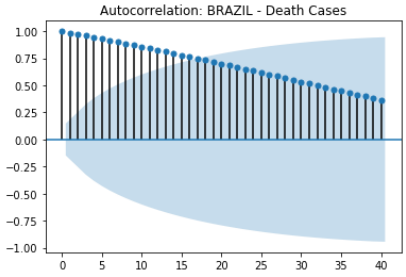
(c)



(d)



(e)



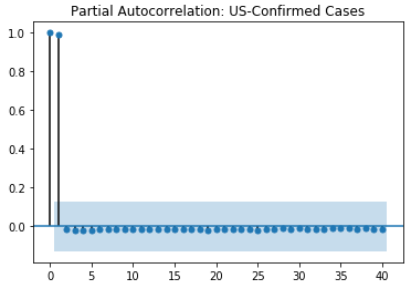
(f)

Fig.2. Estimated autocorrelations for COVID-19 (a-c) Confirmed cases (d-f) Death cases

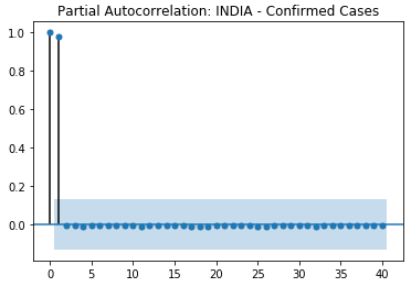
4. DISCUSSION

The COVID-19 affected the goods and operation of all the services. It created a panic among all the employees around the globe. People are held up in their residences. The migrant labors are struggling in all the countries. Ultimately, most of the people are struggling to earn their livelihood. It has been creating lot of challenges in employee engagements, per capita income and national income. Even many countries are wobbling to protect the people from poverty. The distribution of essential services is disturbed. There is lot of hurdles to the people in commuting for meeting basic needs.

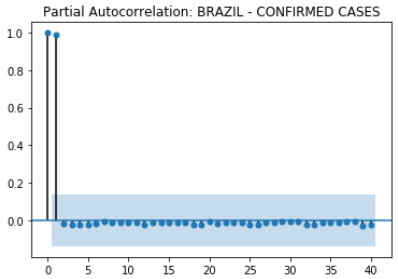
In order to develop an effective healthcare system for the prevention and spreading of this COVID-19 pandemic, effective strategies need to be implemented. Forecasting the epidemiological trend of the prevailing COVID-19 outbreak is crucial for the regulation of production activities, allocation of medical resources and for the national income distribution. Therefore, it is emerged as a need to create a reliable forecasting model that can pave a way to governments a way to decide on microeconomic strategies and allocation of healthcare resources effectively.



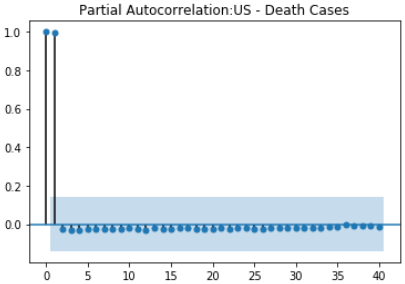
(a)



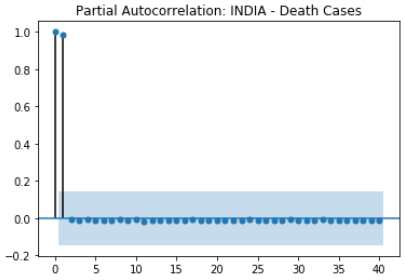
(b)



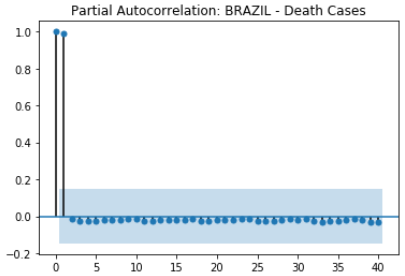
(c)



(d)



(e)



(f)

Fig.3. Estimated partial autocorrelations for COVID-19 ((a)-(c) Confirmed cases (d)-(f) Death cases)

ARIMA model is simple and easy to apply in forecasting such situation [23]. In this work ARIMA model is used for analyzing the current situation of the COVID-19 pandemic in US, India and Brazil. The models present the ongoing trend and extend of the outbreak estimation in the next four weeks. During this COVID-19 pandemic, this work is the first to apply ARIMA models for forecasting COVID-19 pandemic in the next four weeks in US, India and Brazil.

Fig.4. Fitness of the ARIMA model for COVID-19 ((a)-(c) Confirmed cases (d)-(f) Death cases)

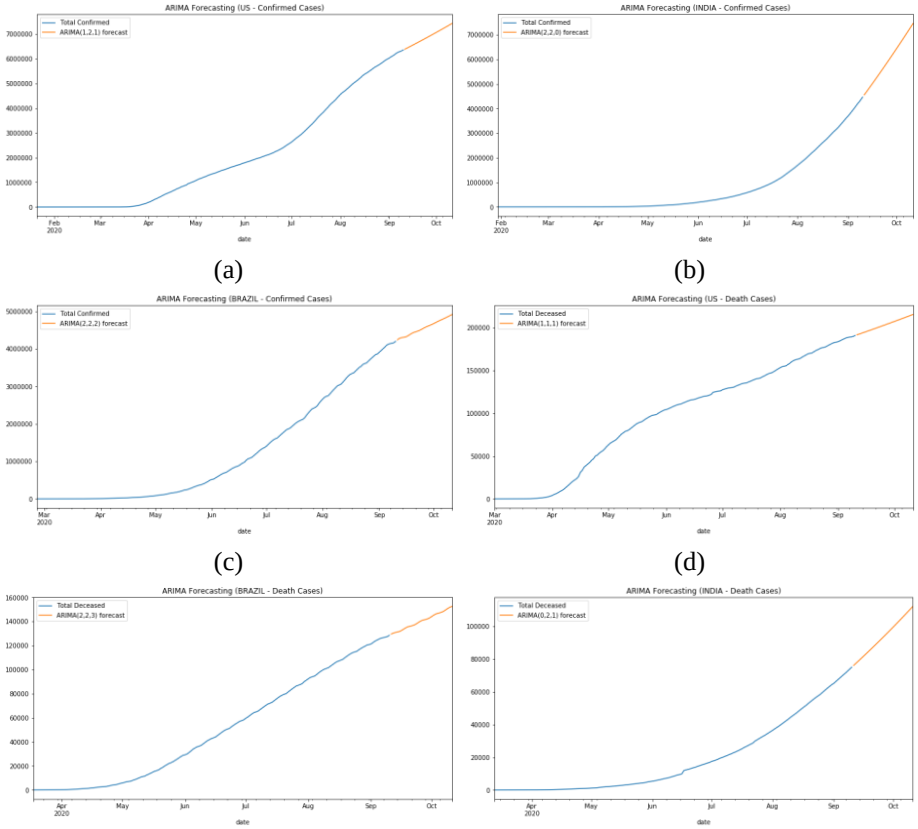


Fig.5. Actual and forecasted COVID-19 cases for the next four weeks ((a)-(c) Confirmed cases (d)-(f) Death cases)

Table.3. Prediction of total confirmed and Death cases of COVID-19 for the next four weeks

Date	US		India		Brazil	
	Confirmed cases	Death cases	Confirmed cases	Death cases	Confirmed cases	Death cases
10.09.2020	6392122	191567	4556426	76165	4241468	129586
11.09.2020	6424777	192323	4645270	77274	4276388	130395
12.09.2020	6457610	193082	4736820	78389	4295268	130892
13.09.2020	6490597	193844	4829575	79511	4302831	131239
14.09.2020	6523729	194609	4921787	80639	4310386	131713
15.09.2020	6557005	195377	5014108	81774	4327069	132501
16.09.2020	6590422	196149	5107223	82915	4354253	133555
17.09.2020	6623981	196922	5200854	84062	4386190	134632
18.09.2020	6657682	197699	5294732	85215	4415210	135474
19.09.2020	6691524	198478	5388974	86375	4437162	136004
20.09.2020	6725508	199260	5483688	87541	4453444	136379
21.09.2020	6759634	200044	5578822	88713	4468983	136878
22.09.2020	6793901	200831	5674335	89892	4488283	137689
23.09.2020	6828309	201620	5770248	91077	4512545	138769
24.09.2020	6862859	202411	5866578	92268	4539530	139876
25.09.2020	6897551	203204	5963315	93466	4565720	140752
26.09.2020	6932384	203999	6060454	94670	4588889	141314
27.09.2020	6967359	204797	6157998	95880	4609313	141719

28.09.2020	7002475	205596	6255949	97096	4629092	142242
29.09.2020	7037733	206397	6354306	98319	4650401	143076
30.09.2020	7073132	207200	6453068	99548	4674044	144181
01.10.2020	7108673	208004	6552235	100784	4699197	145319
02.10.2020	7144356	208811	6651809	102026	4724274	146228
03.10.2020	7180180	209619	6751788	103274	4748123	146823
04.10.2020	7216146	210428	6852172	104529	4770700	147256
05.10.2020	7252253	211239	6952963	105789	4792876	147805
06.10.2020	7288501	212052	7054159	107056	4815683	148662
07.10.2020	7324892	212865	7155761	108330	4839601	149793
08.10.2020	7361423	213681	7257768	109610	4864352	150960
09.10.2020	7398097	214497	7360181	110896	4889232	151903
10.10.2020	7434912	215315	7463000	112188	4913663	152531

In India, number of confirmed cases increases rapidly compared to other two countries. It is expected that India becomes the first position in terms of confirmed cases at the second week of October 2020. Though India has higher increasing rate of confirmed cases, number of death cases are comparatively lower than other countries. As of September 10, 2020, India has the death rate of 1.68 while US and Brazil have 3.00 and 3.06 respectively. It is to be predicted that India will have the death rate 1.50, US will have 2.89 and Brazil will have 3.10 in the next four weeks. It is also estimated that India has an average of 97163 number of confirmed cases and 1203 number of death cases in the next four weeks whereas US has an average of 34832 confirmed cases and 792 death cases and Brazil has 21975 confirmed cases and 763 death cases per day in the next four weeks.

5. CONCLUSION

Time series analysis plays an important role in forecasting the epidemiological trend of the prevailing COVID-19 outbreak. In this work, ARIMA models were applied for estimating the epidemiological trend of the outbreak of this contagious disease on the recently most affected countries US, India and Brazil. This study would help the people to take the preventive and corrective measures to protect them from contracting of covid-19. The outcome of the study based on the pattern growth in covid-19 cases and death during the period will convince the people and authorities to take right decisions in facing such pandemic situations. The estimated results can help health authorities and governments to decide and distribute timely resources effectively in healthcare management. This will result in economic stability and fruitful financial planning. The countries may take required precautionary measures to control the inflation, unemployment and poverty so that country's economic and financial position will be stable. The growth in per capita income national income can be maintained efficiently.

REFERENCES

[1] D. Benvenuto, M. Giovanetti, L. Vassallo, S. Angeletti and M. Ciccozzi, "Application of the ARIMA Model on the COVID- 2019 Epidemic Dataset", *Data in Brief*, Vol. 29, pp. 1-18, 2020.

[2] S. Boccaletti, W. Ditto, G. Mindlin and A. Atangana, "Modeling and Forecasting of Epidemic Spreading: The Case of Covid-19 and Beyond", *Chaos Solitons Fractals*, Vol. 135, pp. 1-18, 2020.

[3] G.E. Box, G.M. Jenkins, G.C. Reinsel and G.M. Ljung, "*Time Series Analysis: Forecasting and Control*", John Wiley and Sons, 2015.

[4] L. Cao, L. Ting, H. Hui, J. Li, X. Yin and J.Wang, "Relationship of Meteorological Factors and Human Brucellosis in Hebei Province", *Science of the Total Environment*, Vol. 703, pp. 1-22. 2020.

[5] Y. Chen, Y. Leng, Y. Lu, L. Wen and Y. Qi, "Epidemiological Features and Time-Series Analysis of Influenza Incidence in Urban and Rural Areas of Shenyang, China, 2010-2018", *Epidemiology and Infection*, Vol. 148, No. 1, pp. 1-25, 2020.

[6] A. Earnest, M.I. Chen, D. Ng and Y.S. Leo, "Using Autoregressive Integrated Moving average (ARIMA) Models to Predict and Monitor the Number of Beds Occupied during a SARS Outbreak in a Tertiary Hospital in Singapore", *BMC Health Services Research*, Vol. 5, pp. 1-8, 2005.

[7] X. Fang, W. Liu, J. Ai and C. Bao, "Forecasting Incidence of Infectious Diarrhea using Randomforest in Jiangsu Province", *BMC Infectious Diseases*, Vol. 20, pp. 1-8, 2020.

[8] J. Gaudart, O. Toure, N. Dessay, A.L. Dicko and O.K. Doumbo, "Modelling Malaria Incidence with Environmental Dependency in a Locality of Sudanese Savannah Area, Mali", *Malaria Journal*, Vol. 8, No. 61, pp. 1-26, 2009.

[9] P. Guan, D.S. Huang and B. Zhou, B. Sen, "Forecasting Model for the Incidence of Hepatitis a based on Artificial Neural Network", *World Journal of Gastroenterology*, Vol. 10, No. 24, pp. 3579-3582, 2004.

[10] W.J. Guan, Z.Y. Ni, Y. Hu and J.X. He, "Clinical Characteristics of Coronavirus Disease 2019 in China", *The New England Journal of Medicine*, Vol. 8, No. 2, pp. 1-18, 2020.

[11] Z. He and H. Tao, "ARIMA Model of Positive-Rate of Influenza Viruses among Children in Wuhan, China: A Nine

Year Retrospective Study", *International Journal of Infectious Diseases Epidemiology*, Vol. 74, pp. 61-70, 2018.

[12] V. Kurbalija, M. Ivanovi, D. Schmidt and C. Hinrichs, "Time-Series Analysis in the Medical Domain: A Study of Tacrolimus Administration and Influence on Kidney Graft Function", *Computers in Biology and Medicine*, Vol. 50, pp. 19-31, 2014.

[13] Q. Liu, Z. Li, Y. Ji and W. Wang, "Forecasting the Seasonality and Trend of Pulmonary Tuberculosis in Jiangsu Province of China using Advanced Statistical Time-Series Analyses", *Infection and Drug Resistance*, Vol. 12, No. 1, pp. 2311-2322, 2019.

[14] Q. Liu, X. Liu, B. Jiang and W. Yang, "Forecasting Incidence of Hemorrhagic Fever with Renal Syndrome in China using ARIMA model", *BMC Infectious Diseases*, Vol. 11, pp. 1-24, 2011.

[15] E.O. Nsoesie, R.J. Beckman and M.V. Marathe, "A Simulation Optimization Approach to Epidemic Forecasting", *PLoS One*, Vol. 8, No. 2, pp. 34-45, 2013.

[16] C. Orbann, L. Sattenspiel and J. Dimka, "Defining Epidemics in Computer Simulation Models: HOW Do Definitions Influence Conclusions?", *Epidemics*, Vol. 19, pp. 24-32, 2017.

[17] S. Polwiang, "The Time Series Seasonal Patterns of Dengue Fever and Associated Weather Variables in Bangkok (2003-2017)", *BMC Infectious Diseases*, Vol. 20, pp. 208-217, 2020.

[18] H. Ren, J. Li, Z.A. Yuan and Y.H. Lu, "The Development of a Combined Mathematical Model to Forecast the Incidence of Hepatitis E in Shanghai, China", *BMC Infectious Diseases*, Vol. 13, pp. 1-6, 2013.

[19] J.M. Sun, L. Lu, K.K. Liu and Q.Y. Liu, "Forecast of Severe Fever with Thrombocytopenia Syndrome Incidence with Meteorological Factors", *Science of the Total Environment*, Vol. 626, pp. 1188-1192, 2018.

[20] M.C. Thomson, A.M. Molesworth, M.H. Djingarey, K.R. Yameogo and L.E. Cuevas, "Potential of Environmental Models to Predict Meningitis Epidemics in Africa", *Tropical Medicine and International Health*, Vol. 11, No. 2, pp. 78-788, 2006.

[21] C. Wang, P.W. Horby and G.F. Gao, "A Novel Coronavirus Outbreak of Global Health Concern", *Lancet*, Vol. 395, No. 1, pp. 470-473, 2020.

[22] Y. Wang, Z. Shen and Y. Jiang, "Comparison of ARIMA and GM (1,1) Models for Prediction of Hepatitis B in China", *PLoS One*, Vol. 13, No. 1, pp. 1-11, 2018.

[23] Y. Wang, C. Xu, Z. Wang and J. Yuan, "Time Series Modeling of Pertussis Incidence in China from 2004 to 2018with A Novel Wavelet based SARIMA-NAR Hybrid Model", *PLoS One*, Vol. 13, No. 2, pp. 1-23, 2018.

[24] W. Wei, J. Jiang, H. Liang, L. Gao and B. Liang, "Application of a Combined Model with Autoregressive Integrated Moving Average (ARIMA) and Generalized Regression Neural Network (GRNN) in Forecasting Hepatitis Incidence in Heng County, China", *PLoS One*, Vol. 11, No. 1, pp. 34-45, 2016.

[25] W. Wu, J. Guo, S. An, P. Guan and P. Ren, "Comparison of Two Hybrid Models for Forecasting the Incidence of Hemorrhagic Fever with Renal Syndrome in Jiangsu Province, China", *PLoS One*, Vol. 10, No. 2, pp. 1-13, 2015.

[26] . Wu, S.Y. An, P. Guan, D.S. Huang and D.S. Zhou, "Time Series Analysis of Human Brucellosis in Mainland China by using Elman and Jordan Recurrent Neural Networks", *BMC Infectious Diseases*, Vol. 19, pp. 1-11, 2019.

[27] Q. Zeng, D. Li, G. Huang, J. Xia and H.J. Wang, "Time Series Analysis of Temporal trends in the pertussis incidence in Mainland China from 2005 to 2016", *Scientific Reports*, Vol. 6, pp. 1-8, 2016.

[28] X. Zhang, Y. Liu, M. Yang, T. Zhang and T. Young, "Comparative Study of Four Time Series Methods in Forecasting Typhoid Fever Incidence in China", *PLoS One*, Vol. 8, No. 2, pp. 1-17, 2013.

[29] L. Zhang, L. Wang, Y. Zheng and X. Zheng, "Time Prediction Models for Echinococcosis based on Gray System Theory and Epidemic Dynamics", *International Journal of Environmental Research and Public Health*, Vol. 14, No. 2, pp. 1-16, 2017.

[30] S. Zhang, M. Diao, W. Yu, L. Pei and D. Chen, "Estimation of the Reproductive Number of Novel Coronavirus (COVID-19) and the Probable Outbreak Size on the Diamond Princess Cruise Ship: A Data-Driven Analysis", *International Journal of Infectious Diseases*, Vol. 93, pp. 201-204, 2020.

[31] Y.L. Zheng, L.P. Zhang and X.L. Zhang, "Forecast Model Analysis for the Morbidity of Tuberculosis in Xinjiang, China", *PLoS One*, Vol. 10, No. 1, pp. 1-13, 2015.