ADITUM

Aditum Journal of Clinical and Biomedical Research

Open Access

Research Article

Forecasting The Prevalence Of Covid-19 In Maharashtra, Delhi, And Kerala Using An Arima Model

Strong P Marbaniang

Department of Public Health and Mortality Studies International Institute for Population Sciences, Mumbai, Maharashtra, India.

Article Info

Received: April 22, 2021
Accepted: April 30, 2021
Published: May 17, 2021

***Corresponding author:** Strong P Marbaniang, Department of Public Health and Mortality Studies International Institute for Population Sciences, Mumbai, Maharashtra, India.

Citation: Strong P Marbaniang. (2021) "Forecasting the prevalence of COVID-19 in Maharashtra, Delhi, and Kerala using an ARIMA model", Aditum Journal of Clinical and Biomedical Research, 2(1); DOI: http;//doi.org/04.2021/1.1028.

Abstract:

Aims:
As the whole world was preparing to welcome the year 2020, a new deadly virus, COVID-19, was reported in the Wuhan city of China in late December 2019. By June 28, 2020, approximately 10 million cases and 0.50 million deaths had been reported globally. There is an urgent need to predict the COVID-19 prevalence to control the spread of the virus.

Methods:
Time-series analyses can help understand the impact of the COVID-19 epidemic and take appropriate measures to curb the spread of the disease. In this study, an ARIMA model was developed to predict the trend of COVID-19 prevalence in the states of Maharashtra, Delhi and Kerala.

Results:
The prevalence of COVID-19 from 16 March 2020 to 27 June 2020 was collected from the website of Covid19india. Several ARIMA models were generated along with the performance measures. ARIMA (1,3,1), ARIMA (2,3,2), and ARIMA (2,3,1) with the MAPE (3.4, 13.61 and 2.01) for Maharashtra, Delhi, and Kerala were selected as the best fit models respectively. The findings show that over the next 30 days, the total number of confirmed COVID-19 cases may increase to 0.439 million in Maharashtra, 0.21 million in Delhi, and 12,113 in Kerala.

Conclusion:
The results of this study can throw light on the intensity of the epidemic in the future and will help the government administrations in Maharashtra, Delhi, and Kerala to formulate effective measures and policy interventions to curb the virus in the coming days.

Keywords: COVID-19, Pandemic, ARIMA, Time series, Forecasting, India

Introduction:

COVID-19, a global pandemic, is an emerging disease that spreads from human to human and is responsible for infecting millions and killing thousands of people since the first reported fatal cases in late 2019. COVID-19 belongs to the family of zoonotic coronaviruses such as the Severe Acute Respiratory Syndrome coronavirus (SARS-Cov) and the Middle East Respiratory Syndrome (MERS-Cov) that have their origin in bats, mice, and domestic animals. The virus first emerged in Wuhan, the capital city of China's Hubei province, in late December 2019. In just a few months, the virus spread rapidly across the world, reaching a total of approximately 4.7 million confirmed cases and 315,496 deaths as of 18 May 2020.[1] The first case of COVID-19 in India was reported in Kerala on the 30th of January 2020, with origins in China.[2] By 17 May 2020, India had registered a total of 95,698 confirmed cases and 30,24 deaths.[3]

As of today, the disease has spread all over the world. The number of confirmed COVID-19 cases vary due to the differences in the testing and disease surveillance capacities across the countries and regions. Since there is no valid treatment method and prevention for this virus yet, effective planning and proper implementation of the health infrastructure and services is the only way to control the spread of the virus. For this reason, accurate forecasting of future total confirmed cases plays a vital role in managing the health system and allows the decision-makers to develop a strategic plan and interventions to avoid a possible epidemic. Also, such estimates help in guiding the intensity and types of interventions needed to lessen the outbreak [.4,5] To estimate the number of additional manpower and resources needed to control the outbreak, a mathematical and statistical modeling tool is required that can be used for making short- and long-term disease forecasting.

In the last few years, studies have used different statistical methods such as multivariate linear regression, [6] simulation-optimization approach, [7] generalized growth model and generalized logistic model, [8] holt method, [9] and grey model [10] to forecast epidemic cases. These statistical models, however, are inadequate for analyzing the influence of randomness on the epidemic outbreak. Random factors play an important role in the spread of a disease as Nakamura and Martinez have described in their study.[11]

Autoregressive integrated moving average (ARIMA) models are the most commonly used prediction models and are considered to be the best [12] for predicting epidemic diseases, such as malaria, [13] tuberculosis,[14] measles,[15] and influenza.[16] An ARIMA model is commonly used for predicting the time series data of infectious diseases, especially for series that have a cyclical or repeating pattern. Mostly, it deals with non-stationary time series in order to capture the linear trend of an epidemic or a disease, and it mainly predicts a future time series value by considering the previous time series values and the lagged forecast error.

In recent studies, different models have been used to predict the prevalence, incidence, and mortality rate of COVID-19. Perone (2020) used an ARIMA model and predicted that Italy would reach the inflection point in terms of cumulative cases during the months of April and May.[17] Zhao et al. (2020) applied the Metropolis-Hastings algorithm and predicted the effects of three epidemic intervention scenarios, that is, suppression, mitigation, and mildness in controlling the spread of COVID-19 in African countries. [18] Wang et al. (2020) used the SEIR model and virus reproduction rate R to predict the number of infectious cases in Wuhan, China.[4]

With the rising number of COVID-19 cases every day, there is a lot of stress on the administration and the health care system in India for accommodating patients with symptoms of COVID-19. Hence, the prediction of the estimated new cases in the coming days will help the health administration make adequate arrangements with ample time.

This paper aims to forecast the prevalence of COVID-19 cases in Maharashtra, Delhi, and Kerala. The COVID-19 data corresponds to the period between 16 March 2020 and 27 June 2020. The best fit ARIMA model was used to estimate the prevalence of COVID-19 cases for a period of 30 days. In addition to highlighting the characteristics of the epidemic and the behavior of its spread, this study also provides the health authorities crucial information about the intensity of the epidemic at peak times using ARIMA model. These models can help predict the health infrastructure and materials the patients will need in the future.

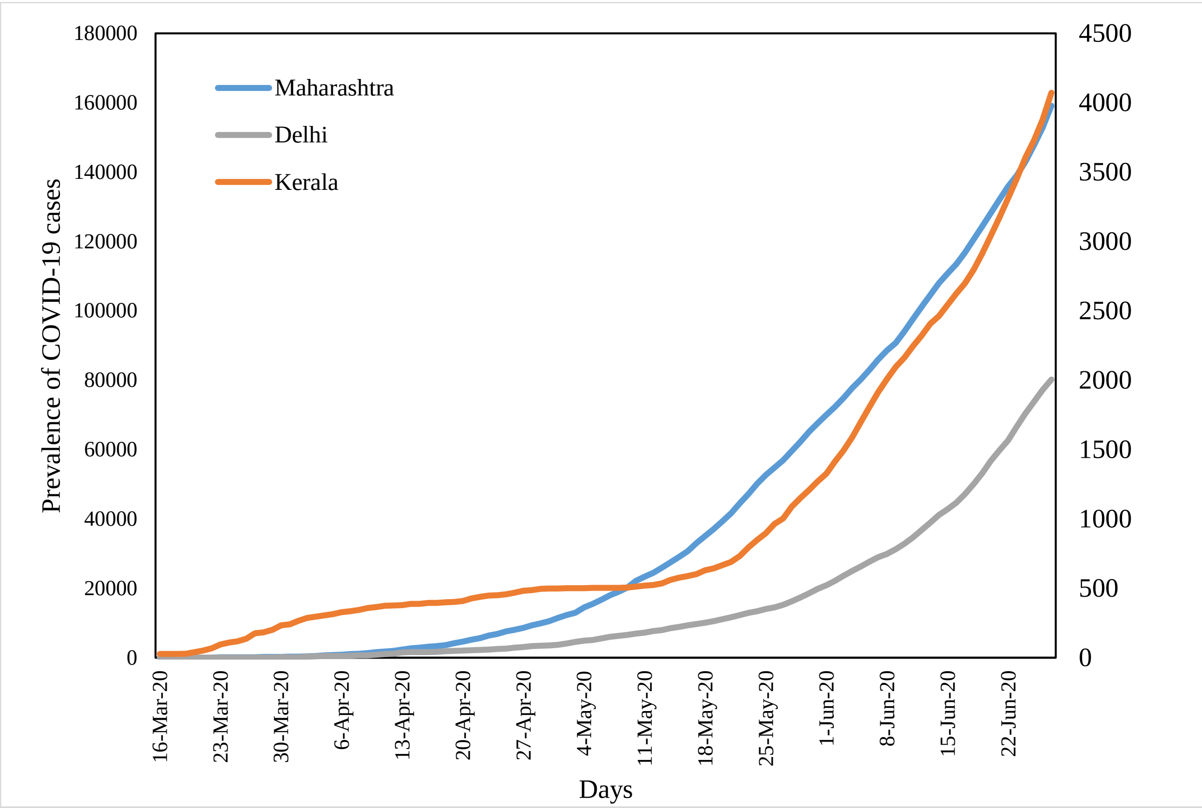


Figure: Represent the Prevalence of COVID-19 cases in Maharashtra, Delhi, and Kerala

Methods:

Data Source:

For the validation and analysis of the proposed study, the prevalence of COVID-19 cases was taken from the website www.covid19india.org and Microsoft Excel was used to build a time-series database. The minimum sample size required for time series forecasting is 30 observations. [19] Hence in this study, 104 time-series observations between 16 March 2020 and 27 June 2020 were used to predict the prevalence of COVID-19 cases over the next 30 days with a 95% confidence interval limit. All analyses were performed using Statgraphics Centurion XVII.II software, with p-value<0.05 as the statistical level of significance.

Statistical Analysis:

ARIMA Model:

A time series is a sequence of observations, each one being recorded at a specific time; it may be measured continuously or discretely.[19] The main aim of a time series is to study past observations and develop an appropriate model to forecast future values. The ARIMA model, first introduced by Box and Jenkins in the 1970s, is the most used time series model if the data show no seasonality pattern. The ARIMA model – generally represented as ARIMA(p,d,q) – is an extension of autoregressive AR(p), moving average MA(q), and ARMA(p,q) models. [16] The letters p, d, and q correspond to order of autoregression, degree of difference, and order of moving average respectively. [19] In an AR(p) model, the current time series value Y_t is expressed as a linear combination of p past observations Y_{t-1} $Y_{t-2}\ldots\ldots Y_{t-p}$ and a random error ε_t, together with a constant term. Similarly, in an MA(q) model, the current time series value Y_t uses past q error terms ε_{t-1} $\varepsilon_{t-2}\ldots\ldots\varepsilon_t$ as the explanatory variables. The general formula of AR(p) and MA(q) models can be expressed as in Eq (1) and (2) respectively.

$$Y_t = C + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \ldots\ldots + \emptyset_p Y_{t-p} + \varepsilon_t \quad (1)$$

$$Y_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots\ldots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (2)$$

Here $\emptyset_i$ (i=1, 2...p) and θ_j (j=1, 2...q) are the autoregressive and moving average parameters respectively. Y_t is the observed value at time t and ε_t the random error (or random shock) at time t. C is the constant term, and μ is the mean of the series. The random shock is assumed to be a white noise process, that is, a sequence of independent and identically distributed (i.i.d) random variables with mean zero and a constant variance σ^2.[19]

The ARMA(p,q) model is a combination of AR(p) and MA(q) models in which the current time series value Y_t is defined linearly in terms of its past p observations as well as the current and past q random shock, together with a constant term. The general formula of an ARMA(p,q) model can be expressed as in Eq (3).

$$Y_t = C + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \ldots\ldots + \emptyset_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots\ldots + \theta_q \varepsilon_{t-q} \quad (3)$$

Where C is a constant and ε_{t-k} (k=1, 2…q) are the values of the previous random shock. Time series analysis requires a stationary time series, that is, the series shows no fluctuation or periodicity with time.[12] In an ARIMA model, a non-stationary time series is made stationary by applying finite differencing to the time series. The differenced stationary time series can be modeled as an ARIMA model to perform an ARIMA forecasting. [16]

Best fit model selection:

Once a model is generated, it is necessary to test the goodness of the model fit before forecasting future values. The accuracy of the model can be determined by comparing the actual values with the predicted values. In this study, we used three performance measures, namely Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE), to test the forecasting accuracy of a particular model. Mathematically, these measures are expressed as in Eq (4), (5), and (6).

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|e_t| \quad (4)$$

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{e_t}{y_t}\right| \times 100 \quad (5)$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2} \quad (6)$$

Where y_t is the actual value at time t, and e_t is the difference between the actual and the predicted values? Also, n is the number of time points. Lower MAE, MAPE, and RMSE values indicate a model that best fits the data. [20]

Steps involved in ARIMA modeling:

Four critical steps are involved in the ARIMA modeling, namely, identification, estimation, diagnostic checking, and forecasting. The first step is to check the seasonality and stationarity of the time series data by drawing a time series plot of the observed series with the corresponding time. A time series is considered as stationary if a shift in time doesn't cause a change in the shape of the distribution, that is, the statistical properties such as mean, variance, and autocorrelation are constant over time. The stationarity of time-series data is important as it helps develop powerful techniques to forecast future values.[21] The second step is to construct the autocorrelation (ACF) and the partial autocorrelation (PACF) plots of the stationary time series to determine the order of the AR and MA processes. The ACF is the correlation between the observation at time t and the observation at a different time lag, while PACF is the amount of correlation between the current observation at time t and the observation at lag k that is not explained by the correlation at all lower-order lags (that is, lag<k).[21] The third step involves estimating the parameters of the best fit model, which is done using the performance measure criteria. The ACF plot of residuals, as well as the Box-Pierce test of white noise, were determined to evaluate the model goodness of fit. The fourth step involves forecasting future values using a good fit model.

Results:

Prevalence and incidence of COVID-19:

Cases	States	Mean	St. Dev	Minimum	Maximum	Skewness	Kurtosis
Prevalence	Maharashtra	39834	46296	38	159133	1.02	-0.21
	Kerala	998	1026	27	4072	1.41	0.97
	Delhi	15061	20259	7	80188	1.62	1.87
Incidence	Maharashtra	1530	1433	3	6368	0.75	-0.06
	Kerala	39	45	0	195	1.26	0.81
	Delhi	771	1015	0	3947	1.61	1.72

Table 1: Descriptive statistics of the Prevalence and Incidence of

Covid-19 in Maharashtra, Delhi, Kerala, and IndiaThe descriptive
statistics of the prevalence and incidence of COVID-19 in Maharashtra, Delhi, and Kerala are given in Table 1. As seen in

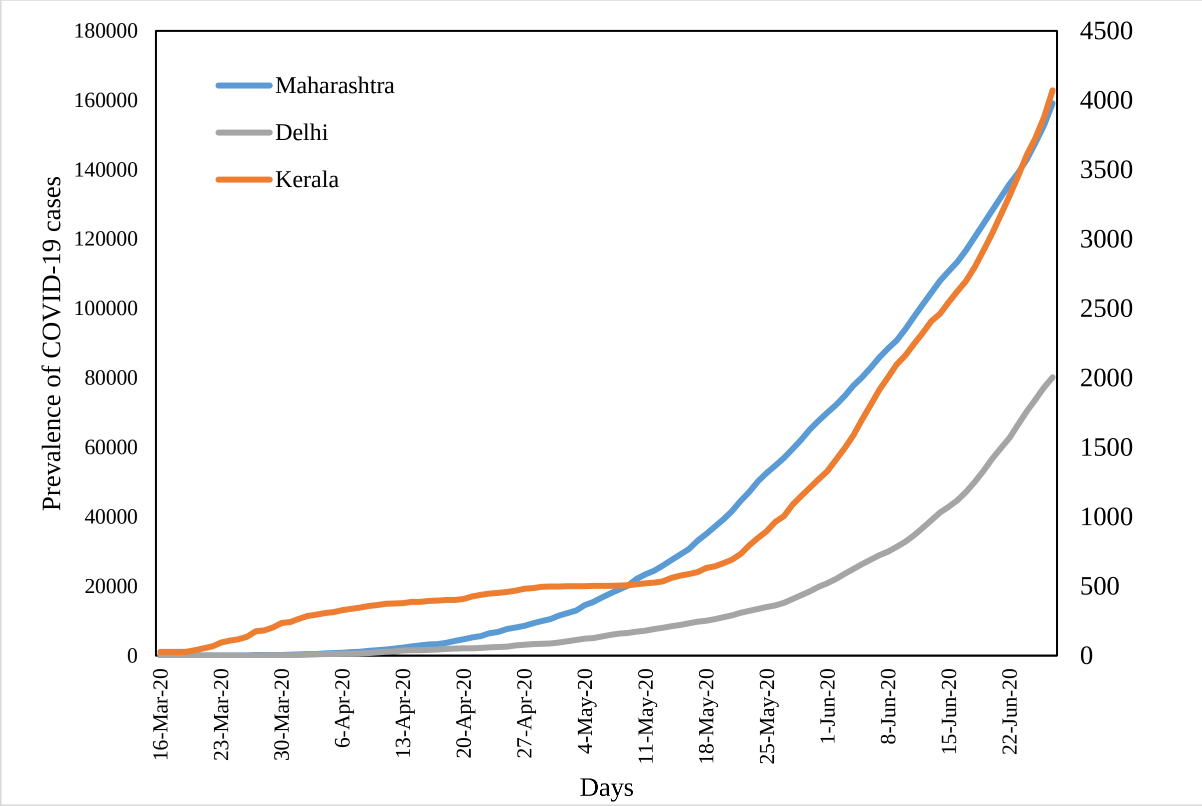
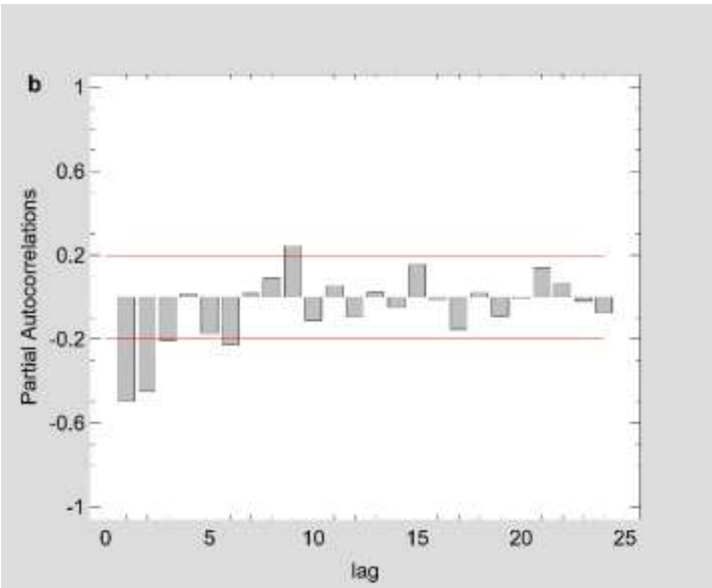
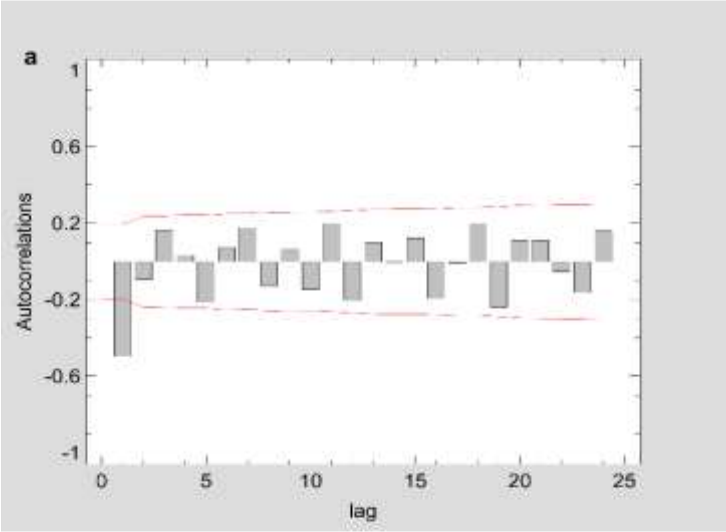
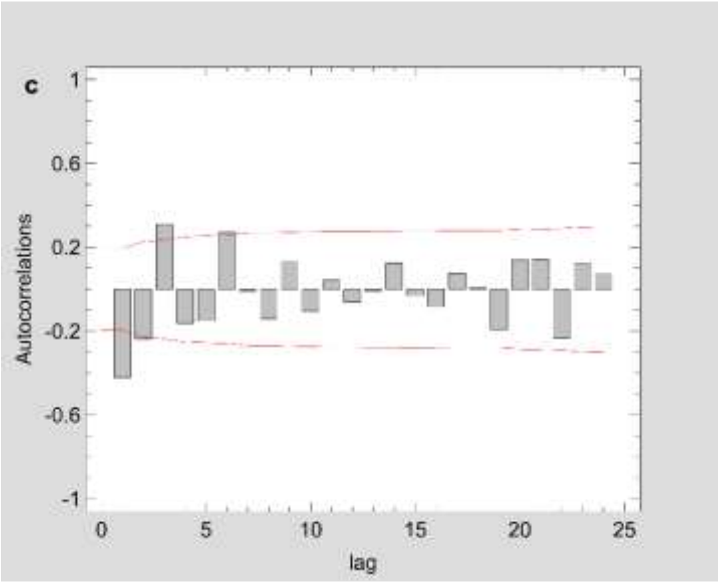


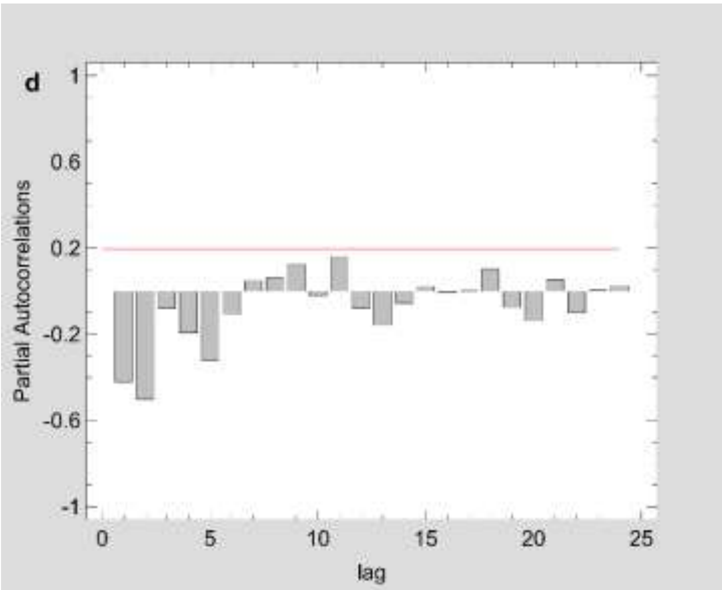
Figure 1: Represent the Prevalence of COVID-19 cases in Maharashtra, Delhi, and Kerala

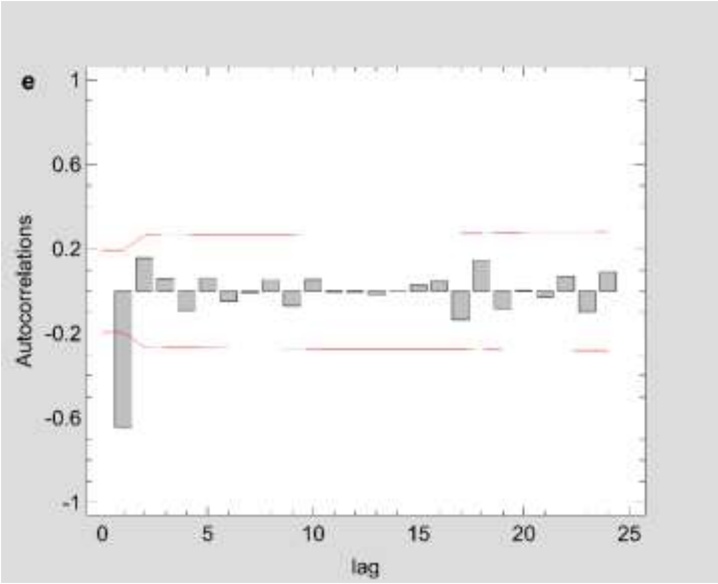
Figure 1 the prevalence curve of COVID-19 has been growing at a steep rate, with the states of Maharashtra and Delhi following a similar trend. With an average of 1530 new cases per day Maharashtra was one of the most hard-hit states of India. In the national capital Delhi, the first case of COVID-19 was reported on the 2nd of March, and since then, the number of confirmed cases has climbed to about 80,188 cases.




a
Autocorrelations
1
0.6
0.2
-0.2
-0.6
-1
0
5
10
15
20
25
lag


c
Autocorrelations
1
0.6
0.2
-0.2
-0.6
-1
0
5
10
15
20
25
lag


d
Partial Autocorrelations
1
0.6
0.2
-0.2
-0.6
-1
0
5
10
15
20
25
lag



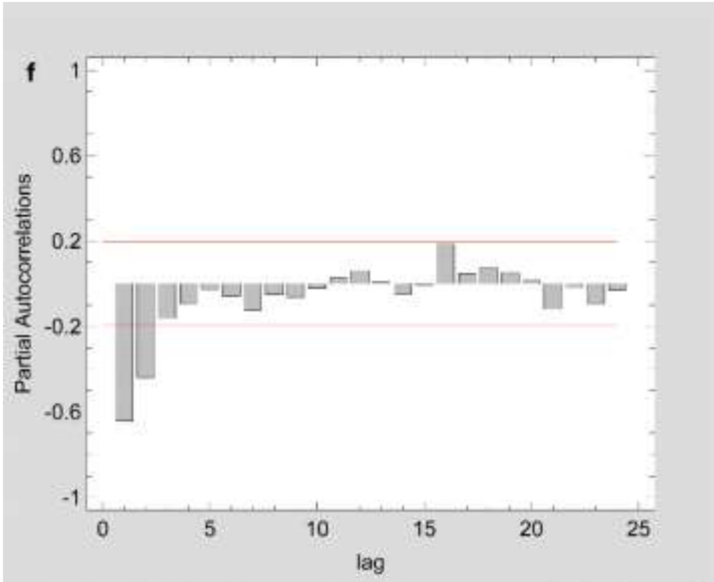


Figure 2: The estimated ACF and PACF plot to predict the trend of Covid-19 prevalence for (a-b) Maharashtra (c-d) Delhi and (e-f) Kerala.

Forecasting the prevalence of COVID-19 pandemic using the time series ARIMA model:

States	Model	RMSE	MAE	MAPE
Maharashtra	ARIMA (1,3,1)	347.193	238.554	3.400
	ARIMA (2,3,1)	344.650	234.582	3.756
	ARIMA (3,3,1)	346.459	232.809	3.102
Delhi	ARIMA (2,3,1)	212.726	135.370	13.534
	ARIMA (2,3,2)	212.033	135.133	13.615
	ARIMA (2,3,3)	213.276	135.087	13.945
Kerala	ARIMA (1,2,1)	12.766	8.878	1.987
	ARIMA (1,3,1)	12.868	8.849	2.032
	ARIMA (2,3,1)	12.684	8.708	2.011

Table 2: Comparison of ARIMA models performance measures.

States	Best fit Model	Parameters	Coefficient	S. E	t-statistic	p-Value
Maharashtra	ARIMA (1,3,1)	$\emptyset_1$	-0.182	0.107	-1.709	<0.10
		θ_1	0.966	0.008	115.853	<0.01
Delhi	ARIMA (2,3,2)	$\emptyset_1$	-0.575	0.263	-2.183	<0.05
		$\emptyset_2$	-0.306	0.103	-2.970	<0.01
		θ_1	-0.478	0.270	1.772	<0.1
		θ_2	0.472	0.830	1.800	<0.1
Kerala	ARIMA (2,3,1)	$\emptyset_1$	-0.500	0.109	-4.586	<0.01
		$\emptyset_2$	-0.223	0.111	-2.005	<0.05
		θ_1	0.962	0.008	113.25	<0.01

Table 3: Parameters of best fit ARIMA models.

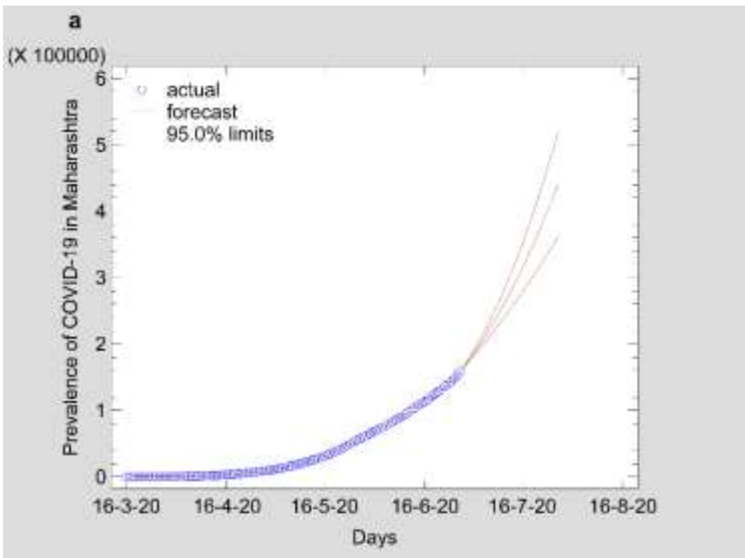
Figure show that the prevalence of COVID-19 in this study shows no seasonal pattern, which is also supported by the autocorrelation plot of the cumulative COVID-19 cases for Maharashtra, Delhi, Kerala (see Appendix). The two lines on the graph indicate the lower and upper limits of the 95% confidence interval. These lines help identify the presence of non-zero autocorrelation. The ACF plot confirms that the prevalence of COVID-19 is not stationary as the autocorrelation is seen to reduce slightly with increasing lag (see Appendix). The first- and second-order differencing were taken to stabilize the mean of COVID-19 prevalence for Maharashtra, Delhi, Kerala. After the third-order differencing, the series became stationary, and the parameters of the ARIMA model were determined according to the ACF and PACF plots as shown in **Error! Reference source not found.**. All the analyses w ere performed on the transformed prevalence of COVID-19. The ARIMA model with the lowest MAPE and with most statistically significant parameters was selected as the best model for forecasting. ARIMA (1,3,1), ARIMA (2,3,2), and ARIMA (2,3,1) were selected as the best fit models for Maharashtra, Delhi, and Kerala. With the $MAPE_{Maharashtra}$= 3.4, $MAPE_{Delhi}$= 13.615, and $MAPE_{Kerala}$= 2.011, the models fitted the prevalence of COVID-19 very well (**Error! Reference source not found.** and Table **2**). All the estimated parameters of the best fit models are presented in Table 3. The fitted and predicted total confirmed COVID-19 cases are presented in **Error! Reference source not f ound.** and **Error! Reference source not found.**. For the next 30 days, the total number of confirmed COVID-19 cases is estimated to be from 3,60,888 to 5,18,538 in Maharashtra, 1,63,017 to 2,52,929 in Delhi, 10,023 to 14,204 in Kerala.

Date	Maharashtra			Delhi			Kerala		
	ARIMA (1,3,1)			ARIMA (2,3,2)			ARIMA (2,3,1)		
	Forecast	Lower limit	Upper Limit	Forecast	Lower limit	Upper Limit	Forecast	Lower limit	Upper Limit
28-Jun-20	165473	164784	166162	83236	82815	83657	4248	4222	4273
29-Jun-20	172036	170586	173486	86517	85596	87437	4432	4386	4478
30-Jun-20	178780	176386	181175	89767	88315	91218	4626	4556	4697
01-Jul-20	185716	182217	189214	93098	91006	95191	4823	4723	4923
02-Jul-20	192843	188090	197595	96529	93706	99352	5026	4892	5159
03-Jul-20	200164	194014	206313	100016	96389	103643	5235	5065	5405
04-Jul-20	207681	199997	215365	103580	99069	108091	5449	5239	5660
05-Jul-20	215397	206043	224751	107224	101753	112695	5669	5415	5924
06-Jul-20	223313	212157	234470	110941	104435	117446	5896	5594	6197
07-Jul-20	231432	218344	244521	114735	107122	122349	6128	5775	6480
08-Jul-20	239756	224606	254907	118608	109814	127402	6366	5959	6773
09-Jul-20	248288	230946	265629	122559	112512	132606	6610	6146	7074

10-Jul-20	257028	237368	276688	126589	115218	137961	6860	6335	7385
11-Jul-20	265980	243873	288086	130700	117932	143468	7117	6527	7706
12-Jul-20	275145	250465	299825	134891	120655	149127	7379	6722	8037
13-Jul-20	284526	257144	311907	139164	123388	154939	7648	6920	8377
14-Jul-20	294124	263914	324335	143519	126132	160905	7924	7121	8727
15-Jul-20	303943	270775	337111	147957	128887	167026	8206	7325	9087
16-Jul-20	313984	277731	350237	152479	131654	173303	8494	7532	9456
17-Jul-20	324249	284782	363717	157086	134434	179737	8789	7742	9836
18-Jul-20	334741	291930	377552	161778	137227	186329	9091	7955	10226
19-Jul-20	345461	299176	391746	166556	140033	193079	9399	8171	10626
20-Jul-20	356412	306523	406301	171421	142853	199989	9714	8391	11036
21-Jul-20	367596	313972	421220	176373	145687	207060	10036	8614	11457
22-Jul-20	379015	321525	436505	181415	148536	214293	10364	8840	11888
23-Jul-20	390672	329182	452161	186545	151401	221689	10700	9070	12330
24-Jul-20	402567	336946	468189	191765	154280	229250	11043	9303	12782
25-Jul-20	414705	344817	484593	197076	157176	236976	11392	9539	13245
26-Jul-20	427086	352797	501374	202478	160088	244869	11749	9779	13719
27-Jul-20	439713	360888	518538	207973	163017	252929	12113	10023	14204

Table 4: Prediction of the total confirmed Covid-19 cases for the next 30 days according to ARIMA models with 95% confidence interval

be black and white print



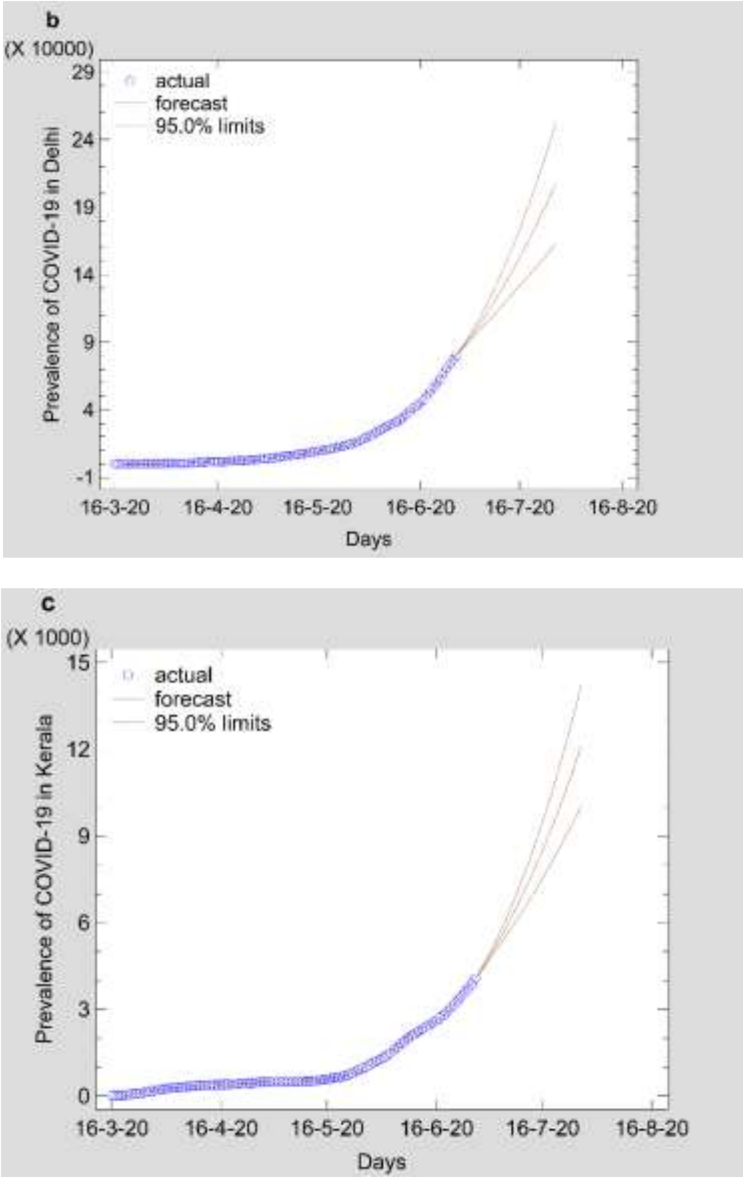


Figure 3: Time series forecasting plot estimated from ARIMA best-fit model for (a) Maharashtra (b) Delhi and (c) Kerala.
Note: Fig 3 should be black and white print

Discussion:

In an effort to slow down the spread of COVID-19, the Indian government took strong measures by announcing a countrywide lockdown on 24 March 2020 as the number of confirmed positive cases were increasing in the country. Estimating the prevalence and intensity of an epidemic is crucial for allocating medical and health resources, production of activities, and even the economic situation of the country. Hence developing a forecasting model that accurately predicts the future intensity of an epidemic can help the government administrators and decision-makers prepare the manpower and medical supplies required during an outbreak. In this study, the ongoing trend and the intensity of the COVID-19 pandemic were estimated using the ARIMA time series model. The ARIMA model is one of the best models and has been extensively employed to predict the incidence of contagious diseases. [22] To the best of our knowledge, this is the first study in India to apply the ARIMA model to estimate the prevalence of COVID-19 in India major states.

India has reported a lower COVID-19 death rate as compared to countries like China, United Kingdom, Italy, Spain, and the United States. [23] However, the total confirmed COVID-19 cases in most of the Indian states show no sign of a downward trend. At the time of writing this article, India had 5,38,529 positive confirmed cases [3] and was expected to overtake Russia total COVID-19 cases. Minhas (2020), in his study, points out that India is another potential epicenter of the global COVID-19 pandemic due to human overpopulation and unhygienic living conditions. [24] Containing the spread of the virus among the economically disadvantaged people, who may not be able to self-isolate, is a challenge. In Maharashtra, the number of daily new cases since March 16 has grown exponentially and crossed the 1000-cases-per-day mark on May 6. Mumbai, the state capital of Maharashtra and also India's financial capital, has been the worst hit city by COVID-19, having recorded 15,750 total cases accounting for 20 percent of all positive COVID-19 cases in India.[25] Kerala reported the first case of COVID-19 in India; however, over a period of one month, the daily new confirmed

cases significantly reduced to zero for five consecutive days.[3] Delhi, the national capital, reported 3947 COVID-19 cases in a single day on June 23, the highest jump so far. With the lockdown curbs being relaxed after May 17, the number of new cases may increase further.[26] This pattern will burden the health system to its maximum capacity. As a result, if adequate measures to contain the spread are not appropriately enforced, and social distancing is not maintained, the number of cases is not expected to plateau any time soon.
An epidemic is a numbers game and as far as numbers are concerned, India has a handful of them. With no valid medical treatment and preventive measures for this virus to date, forecasting the prevalence of the disease is a vital strategy to strengthen the surveillance and allocate health resources accordingly. The results of the study will help health authorities and health care management plan the necessary supply resources, which include medical staff, medical equipment, intensive care facilities, hospital beds, and other healthcare facilities. This will make the epidemic controllable and bring it within the domain of the available healthcare resources in India.

Acknowledgments:

Conflict of interest statement: The authors declare that they have no conflicts of interest.

Role of funding source: The research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethical Approval: Non sought

References:

1. Hopkins J. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. 2020.
2. PIB.No title. 2020.
3. Covid19. Covid19 India. 2020. Available from: https://www.covid19india.org/.
4. Wang H, Wang Z, Dong Y, Chang R, Xu C, Yu X et al. Phase-adjusted estimation of the number of Coronavirus Disease 2019 cases in Wuhan, China. Cell Discovery. 2020; 6(1): 4–11.
5. Zhang S, Diao MY, Yu W, Pei L, Lin Z, Chen D. 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. 2020; 93: 201–204.
6. Thomson MC, Molesworth AM, Djingarey MH, Yameogo KR, Belanger F, Cuevas LE. Potential of environmental models to predict meningitis epidemics in Africa. Tropical Medicine and International Health. 2006; 11(6): 781–788.
7. Nsoesie EO, Beckman RJ, Shashaani S, Nagaraj KS, Marathe MV. A Simulation Optimization Approach to Epidemic Forecasting. PLoS ONE. 2013; 8(6).
8. Chowell G, Luo R, Sun K, Roosa K, Tariq A, Viboud C. Real-time forecasting of epidemic trajectories using computational dynamic ensembles. Epidemics. 2020; 30 (August 2019).
9. Myrzakerimova AB, Shaizat M, Duisebekova KS, Nurmaganbetova MO. Forecasting risk of diseases in Kazakhstan with using mapping technique based on 9 years statistics. Procedia Computer Science. 2020; 170: 75–81.
10. Zhang L, Wang L, Zheng Y, Wang K, Zhang X, Zheng Y. Time prediction models for echinococcosis based on gray system theory and epidemic dynamics. International Journal of Environmental Research and Public Health. 2017; 14(3).
11. Nakamura GM, Martinez AS. Hamiltonian dynamics of the SIS epidemic model with stochastic fluctuations. Scientific Reports. 2019; 9(1): 1–9.
12. Wang Y, Shen Z, Jiang Y. Comparison of ARIMA and GM(1,1) models for prediction of hepatitis B in China. PLoS ONE. 2018; 13(9): 1–11.
13. Anokye R, Acheampong E, Owusu I, Isaac-Obeng E (2018). Time series analysis of malaria in Kumasi: Using ARIMA models to forecast future incidence. Cogent Social Sciences. 2018; 4(1): 1–13.
14. Zheng YL, Zhang LP, Zhang XL, Wang K, Zheng YJ (2015). Forecast model analysis for the morbidity of tuberculosis in Xinjiang, China. PLoS ONE. 2015; 10(3): 1–13.
15. Sharmin S, Rayhan I. Modelling of infectious diseases for providing signal of epidemics: A measles case study in Bangladesh. Journal of Health, Population and Nutrition. 2011; 29(6): 567–573.
16. He Z, Tao H. Epidemiology and ARIMA model of positive rate of influenza viruses among children in Wuhan, China: A nine-year retrospective study. International Journal of Infectious Diseases. 2018; 74: 61–70.
17. Perone G. An ARIMA Model to Forecast the Spread of COVID-2019 Epidemic in Italy. SSRN Electronic Journal. 2020; (April).
18. Zhao Z, Li X, Liu F, Zhu G, Ma C, Wang L. Prediction of the COVID-19 spread in African countries and implications for prevention and controls: A case study in South Africa, Egypt, Algeria, Nigeria, Senegal and Kenya. Science of The Total Environment. 2020; 729, 138959.
19. Yaffee RA, McGee M. Introduction To Time Series Analysis and Forecasting with application of SAS and SPSS (1st ed.). Academic Press. 2000
20. Tseng YJ, Shih YL. Developing epidemic forecasting models to assist disease surveillance for influenza with electronic health records. International Journal of Computers and Applications. 2019; 0(0): 1–6.
21. Brockwell PJ, Davis RA. Introduction to Time Series and Forecasting (Second). Springer. 2001.
22. Wang Y, Xu C, Wang Z, Zhang S, Zhu Y, Yuan J. Time series modeling of pertussis incidence in China from 2004 to 2018 with a novel wavelet based SARIMA-NAR hybrid model. PLoS ONE. 2018a; 13(12): 1–23.
23. Ghosal S, Sengupta S, Majumder M, Sinha B. Prediction of the number of deaths in India due to SARS-CoV-2 at 5–6 weeks. Diabetes and Metabolic Syndrome: Clinical Research and Reviews. 2020; 14(4): 311–315.
24. Minhas S. Could India be the origin of next COVID-19 like epidemic? Science of the Total Environment. 2020; 728.
25. Dutta PK (2020). Covid-19: Maharashtra has 33%, Mumbai 20% of all-India cases. Where they failed. 2020.
26. Dutt A (2020). Biggest single-day spike as Delhi Covid-19 cases cross 8,000. 2020.