Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Diabetes & Metabolic Syndrome: Clinical Research & Reviews 14 (2020) 1419–1427

Contents lists available at ScienceDirect

Diabetes & Metabolic Syndrome: Clinical Research & Reviews

journal homepage: www.elsevier.com/locate/dsx

ARIMA modelling & forecasting of COVID-19 in top five affected countries

Alok Kumar Sahai [a,*], Namita Rath [a], Vishal Sood [a], Manvendra Pratap Singh [b]

[a] Sri Sri University, Cuttack, India
[b] National Institute of Technology, Rourkela, India

ARTICLE INFO

Article history:
Received 14 July 2020
Received in revised form
21 July 2020
Accepted 22 July 2020

Keywords:
COVID-19
SARV-2 Cov
Pandemic
ARIMA
Forecasting

ABSTRACT

Background and aims: In a little over six months, the Corona virus epidemic has affected over ten million and killed over half a million people worldwide as on June 30, 2020. With no vaccine in sight, the spread of the virus is likely to continue unabated. This article aims to analyze the time series data for top five countries affected by the COVID-19 for forecasting the spread of the epidemic.
Material and methods: Daily time series data from 15th February to June 30, 2020 of total infected cases from the top five countries namely US, Brazil, India, Russia and Spain were collected from the online database. ARIMA model specifications were estimated using Hannan and Rissanen algorithm. Out of sample forecast for the next 77 days was computed using the ARIMA models.
Results: Forecast for the first 18 days of July was compared with the actual data and the forecast accuracy was using MAD and MAPE were found within acceptable agreement. The graphic plots of forecast data suggest that While Russia and Spain have reached the inflexion point in the spread of epidemic, the US, Brazil and India are still experiencing an exponential curve.
Conclusion: Our analysis shows that India and Brazil will hit 1.38 million and 2.47 million mark while the US will reach the 4.29 million mark by 31st July. With no effective cure available at the moment, this forecast will help the governments to be better prepared to combat the epidemic by ramping up their healthcare facilities.

The Corona pandemic which originated in Wuhan, China in December 2019 has spread out to the whole world and in six months has caused unprecedented havoc. This extremely virulent strain of corona virus is highly contagious and has already affected over 10,101,998 cases worldwide and has claimed 501,644 lives within seven months [1]. The earlier instances of corona virus namely SARS and MERS were not as contagious and persistent as the 2019-nCov or COVID-19 as it has come to be known. The confusion and lack of transparency in the initial stages of the outbreak only worsened the situation and today 185 countries are suffering from the virus with no cure in sight. The virus in the current form is highly contagious and causes death due to respiratory failure. Due to the differences in epidemiological conditions and testing facilities the spread of the virus has been varied in countries. The worst affected are developed countries like Spain, Italy, France, Germany and the US. Today, US tops the list followed by Brazil, Russia, India and Spain respectively [1].

In India, the first case of COVID19 was reported on January 30, 2020 and the spread in India was extremely slow. As the severity of the viral infection became known the Government of India resorted to a complete lockdown to contain the spread of the virus. The first lockdown was announced on 25th March which was extended gradually till the end of May. Owing to the all-round collapse of the industry and the miseries of the daily wage earners and migrant labours, the government decided to lift the lockdown in a phased manner from June 2, 2020. Migrant labours from the two hotspots of New Delhi and Mumbai migrated to their home states and this large scale export of corona virus resulted in the explosion of the number of cases. Slowly India has entered the top ten countries affected by COVID-109 and today is the third most badly affected country in the world [1].

Despite the claims of the government of increased medical and testing facilities, the number of affected cases is not flattening or abating. The number of new patients every day is reaching 20,000 per day and many concerns are looming over the spread of COVID-19. How many people will be infected tomorrow? How many

* Corresponding author.
E-mail address: alok.s@srisriuniversity.edu.in (A.K. Sahai).

https://doi.org/10.1016/j.dsx.2020.07.042
1871-4021/

deaths will happen tomorrow? When will the infection curve reach inflexion or get flattened? How many people will be affected during the peak period of the outbreak? Are there mathematical models available to answer these questions? Under the circumstances, it is very important to estimate the spread of COVID-19 so that the policymakers, medical personnel and the general public could be better prepared to deal with the emergency.

In this paper, we have employed Auto Regressive Integrated Moving Average (ARIMA) model to predict the incidence and spread of the COVID-19 in India, Russia, Brazil, Spain and the US as the five most badly hit countries [1]. As compared to other econometric models ARIMA models have been used with success in the prediction of several diseases [2–7].

1. Literature review

The past two decades have seen research focused on statistical issues pertaining to a prospective detection of outbreaks of infectious diseases. The challenges arise in early detection and possible evolution of the epidemic for taking the appropriate preventive measures. The rapid growth in this area is called biosurveillance [8,9].

An early model of regression method of outbreak detection was presented by Shewhart [10]. Assuming a normally distributed incidence of infected cases the regression tested for exceeding the mean by a certain multiple of the standard deviation. However, with epidemics, the normal distribution is no longer a valid distribution and most epidemics show an exponential distribution or a highly skewed bell curve [11]. used a simple regression model which computer the expected number of cases at month t calculated as the mean count over t-1, t and t+1 months over a specified number of years. Regression models are used to detect the onset of influenza epidemics [12,13]. When data frequency is not much the normal errors regression model are inadequate and Poisson regression models have been used [14,15]. Unlike the parametric regression models described so far semi parametric models can be used to create a baseline model as used in monitoring the mortality and other related effects. A smoothing method to obtain baseline and standard deviations while working with Salmonella outbreaks was used by Ref. [16,17]. Most regression-based models define a mean at time t and issue an alarm at t if the observed value lies above a certain threshold predetermined by the sample statistics and the quantiles of a suitable normal or Poisson distribution [18]. described non-thresholding regression methods which test the hypothesis that a given value y_t at time t belongs to the same distribution as the baseline distribution.

The regression techniques do not capture the correlation structure of the data. Time series methods offer this advantage over the regression methods. Syndromic and laboratory data collected with daily or weekly frequency are generally autocorrelated with some lags. They may further exhibit correlations associated with the seasonal patterns in the data arising out of weekly or yearly seasonality. Failure to account properly for the autocorrelation in the time series data leads to misspecified models and incorrect forecasts. The Box Jenkins model is designed to take care of the autocorrelation of times series into account.

With outbreak surveillance, the trend is best estimated through a relatively simple procedure. A Serfling model [19] based on trigonometric functions may be used to estimate the trend and seasonal components for time series data with regular seasonality. Simple exponential smoothing [20,21] and Holt- Winters procedure are employed in surveillance studies. Simple exponential smoothing makes predictions by taking a weighted average of past observations, the weights decreasing the farther we go in past with the higher weightage on the more recent data. The Hold-Winters procedure is a variant of simple exponential soothing which allows for local trend and seasonality. This method has been used with success in many surveillance studies and has done better than other forecasting methods [22].

Auto Regressive Integrated Moving Average (ARIMA) models [23] have been widely used for detecting outbreaks of infectious diseases [24–27]. Stationarity of the time series is a prerequisite for fitting an ARIMA model. An investigation of ARIMA modelling showed that it was unable to model eight out of 17 syndromic time series resulting from sparse data [28]. However, for the series which were successfully modelled, one step ahead forecasts were highly acceptable and forecasts up to 3 years in future were obtained by continuously updated models. The traditional ARIMA models require a fairly large number of parameters for the auto correlation to be detected. Further, a model for one syndrome or outbreak cannot be automatically applied to another and the model has to be identified each time. For shorter lengths of time series data, it is prudent to use a hierarchical time series model. It is claimed that the hierarchical times series model can detect outbreaks faster than the lab based exceedance system [29].

The ARIMA model has seen widespread usage in the study of infectious diseases for several time series events. These include leptospirosis and its relationship with rainfall and temperature [5] and the relationship of suicide cases with changes in national

Table 1
ARIMA model specifications.

	INDIA	BRAZIL	RUSSIA	SPAIN	US
MODEL	(4,2,4)	(3,1,2)	(3,0,0)	(4,2,4)	(1,2,1)
AR1	1.375893	1.38224	1.821833	−0.1032	0.224382
Std Err	0.13892	0.232872	0.080456	0.24914	0.199694
T	9.90423	5.93562	22.64392	−0.41423	1.12363
P	0.00000	0.00000	0.00000	0.67941	0.26321
AR2	−1.02318	−0.83395	−0.647	0.265827	
Std Err	0.112545	0.328354	0.161136	0.201803	
T	−9.09125	−2.53978	−4.01521	1.31726	
P	0.00000	0.01227	0.00010	0.19014	
AR3	1.369517	0.448294	−0.17526	−0.43591	
Std Err	0.105984	0.12919	0.080759	0.125818	
T	12.92191	3.47003	−2.17015	−3.46459	
P	0.00000	0.00071	0.03177	0.00073	
AR4	−0.72424			−0.63051	
Std Err	0.132004			0.160281	
T	−5.48652			−3.93377	
P	0.00000			0.00014	
MA1	1.794408	0.824621		0.380611	0.594318
Std Err		0.24318			0.165616
T		3.39099			3.58854
P		0.00092			0.00047
MA2	−1.50841	−0.11077		0.150048	
Std Err		0.23289			
T		−0.47564			
P		0.63513			
MA3	1.591911			−0.37507	
Std Err					
T					
P					
MA4	−0.89826			−0.4913	
Std Err					
T					
P					
AIC	2128.399	2689.998	2077.294	2179.333	2624.271
SBC	2154.546	2707.474	2088.974	2205.481	2632.987
Log Likelihood	−1055.2	−1339	−1034.65	−1080.67	−1309.14

Source: Authors' own calculation.

Table 2
Forecast data for COVID-19 outbreak in India.

DATE	LOWER CI	FORECAST	UPPER CI	STD_ERR
January 07, 2020	603,907	605,084	606,262	600.9606
February 07, 2020	621,640	623,844	626,048	1124.4701
March 07, 2020	639,485	642,773	646,062	1677.9432
April 07, 2020	658,599	663,129	667,658	2311.0731
May 07, 2020	678,031	683,792	689,552	2939.1669
June 07, 2020	697,137	704,039	710,940	3521.313
July 07, 2020	717,145	725,228	733,311	4123.9509
August 07, 2020	738,198	747,529	756,861	4760.9582
September 07, 2020	759,060	769,604	780,147	5379.2975
October 07, 2020	780,073	791,821	803,570	5994.242
November 07, 2020	802,265	815,308	828,351	6654.942
December 07, 2020	824,890	839,280	853,670	7341.9044
07/13/2020	847,242	862,980	878,719	8029.8984
07/14/2020	870,285	887,448	904,610	8756.6787
07/15/2020	894,290	912,992	931,694	9542.059
07/16/2020	918,216	938,511	958,807	10355.0318
07/17/2020	942,191	964,142	986,093	11199.8207
07/18/2020	967,129	990,871	1,014,612	12113.2121
07/19/2020	992,534	1,018,184	1,043,834	13087.0884
07/20/2020	1,017,713	1,045,348	1,072,982	14099.6111
07/21/2020	1,043,389	1,073,132	1,102,875	15175.4833
07/22/2020	1,069,915	1,101,928	1,133,942	16333.7611
07/23/2020	1,096,454	1,130,855	1,165,256	17551.9068
07/24/2020	1,122,982	1,159,885	1,196,787	18828.2916
07/25/2020	1,150,284	1,189,858	1,229,433	20191.4058
07/26/2020	1,178,063	1,220,473	1,262,884	21638.3565
07/27/2020	1,205,677	1,251,050	1,296,423	23149.8286
07/28/2020	1,233,645	1,282,138	1,330,631	24741.8074
07/29/2020	1,262,358	1,314,163	1,365,968	26431.5544
07/30/2020	1,291,161	1,346,437	1,401,714	28202.6936
07/31/2020	1,319,928	1,378,826	1,437,724	30050.3959
January 08, 2020	1,349,317	1,412,029	1,474,740	31996.3832
February 08, 2020	1,379,180	1,445,899	1,512,619	34041.2004
March 08, 2020	1,408,942	1,479,831	1,550,720	36168.6273
April 08, 2020	1,438,956	1,514,196	1,589,436	38388.456
May 08, 2020	1,469,620	1,549,422	1,629,224	40715.8197
June 08, 2020	1,500,436	1,584,987	1,669,539	43139.4772
July 08, 2020	1,531,212	1,620,691	1,710,170	45653.2561
DATE	LOWER CI	FORECAST	UPPER CI	STD_ERR
August 08, 2020	1,562,487	1,657,101	1,751,715	48273.3948
September 08, 2020	1,594,221	1,694,184	1,794,147	51002.6718
October 08, 2020	1,625,914	1,731,415	1,836,916	53828.1115
November 08, 2020	1,657,789	1,769,027	1,880,266	56755.2362
December 08, 2020	1,690,226	1,807,426	1,924,625	59796.9572
08/13/2020	1,722,857	1,846,229	1,969,601	62946.0768
08/14/2020	1,755,460	1,885,202	2,014,944	66196.2032
08/15/2020	1,788,461	1,924,794	2,061,128	69559.2399
08/16/2020	1,821,897	1,965,051	2,108,205	73039.2166
08/17/2020	1,855,344	2,005,528	2,155,713	76626.1288
08/18/2020	1,888,925	2,046,354	2,203,784	80322.6473
08/19/2020	1,922,987	2,087,897	2,252,808	84139.5271
08/20/2020	1,957,271	2,129,890	2,302,509	88072.5738
08/21/2020	1,991,544	2,172,088	2,352,631	92115.7886
08/22/2020	2,026,135	2,214,836	2,403,537	96277.7173
08/23/2020	2,061,130	2,258,229	2,455,329	100562.9208
08/24/2020	2,096,178	2,301,903	2,507,629	104963.9251
08/25/2020	2,131,328	2,345,908	2,560,488	109481.5652
08/26/2020	2,166,887	2,390,567	2,614,248	114124.6156
08/27/2020	2,202,684	2,435,706	2,668,728	118891.0175
08/28/2020	2,238,488	2,481,084	2,723,679	123775.4657
08/29/2020	2,274,547	2,526,959	2,779,371	128783.9514
08/30/2020	2,310,975	2,573,455	2,835,936	133921.0867
08/31/2020	2,347,488	2,620,279	2,893,070	139181.4645
January 09, 2020	2,384,084	2,667,426	2,950,768	144564.7979
February 09, 2020	2,421,026	2,715,174	3,009,322	150078.1141
March 09, 2020	2,458,210	2,763,417	3,068,624	155720.8153
April 09, 2020	2,495,420	2,811,932	3,128,443	161488.4038
May 09, 2020	2,532,836	2,860,904	3,188,973	167384.9613
June 09, 2020	2,570,583	2,910,470	3,250,357	173414.8553
July 09, 2020	2,608,441	2,960,401	3,312,360	179574.3196
August 09, 2020	2,646,370	3,010,654	3,374,938	185862.4365
September 09, 2020	2,684,590	3,061,462	3,438,333	192284.7388
October 09, 2020	2,723,048	3,112,770	3,502,493	198841.583
November 09, 2020	2,761,551	3,164,381	3,567,211	205529.2858
December 09, 2020	2,800,221	3,216,420	3,632,619	212350.5262
09/13/2020	2,839,186	3,269,024	3,698,863	219309.2637
09/14/2020	2,878,278	3,322,020	3,765,762	226403.0034
09/15/2020	2,917,435	3,375,343	3,833,250	233630.5279

Source: Authors' own computation.

Table 3
Forecast and 95% confidence interval for COVID outbreak in Brazil.

DATE	LOWER CI	FORECAST	UPPER CI	STD_ERR
January 07, 2020	1,439,607	1,448,644	1,457,681	4610.585
February 07, 2020	1,467,503	1,484,229	1,500,956	8534.161
March 07, 2020	1,494,870	1,517,010	1,539,150	11296.27
April 07, 2020	1,523,982	1,550,697	1,577,411	13630.03
May 07, 2020	1,554,011	1,585,926	1,617,841	16283.47
June 07, 2020	1,583,200	1,621,272	1,659,344	19424.84
July 07, 2020	1,611,204	1,655,902	1,700,600	22805.37
August 07, 2020	1,638,688	1,690,134	1,741,580	26248.49
September 07, 2020	1,666,099	1,724,467	1,782,836	29780.16
October 07, 2020	1,693,346	1,758,950	1,824,554	33,472
November 07, 2020	1,720,204	1,793,378	1,866,551	37334.16
December 07, 2020	1,746,635	1,827,649	1,908,663	41334.32
07/13/2020	1,772,737	1,861,818	1,950,898	45,450
07/14/2020	1,798,577	1,895,950	1,993,322	49680.85
07/15/2020	1,824,148	1,930,047	2,035,947	54031.33
07/16/2020	1,849,425	1,964,081	2,078,736	58498.89
07/17/2020	1,874,411	1,998,039	2,121,667	63076.58
07/18/2020	1,899,125	2,031,931	2,164,736	67759.09
07/19/2020	1,923,581	2,065,764	2,207,948	72543.85
07/20/2020	1,947,782	2,099,540	2,251,298	77428.9
07/21/2020	1,971,730	2,133,254	2,294,777	82411.53
07/22/2020	1,995,430	2,166,904	2,338,379	87488.69
07/23/2020	2,018,887	2,200,493	2,382,099	92657.68
07/24/2020	2,042,109	2,234,022	2,425,934	97916.23
07/25/2020	2,065,099	2,267,490	2,469,880	103262.2
07/26/2020	2,087,862	2,300,897	2,513,933	108693.6
07/27/2020	2,110,400	2,334,244	2,558,088	114208.2
07/28/2020	2,132,719	2,367,531	2,602,343	119804.1
07/29/2020	2,154,822	2,400,757	2,646,693	125479.6
07/30/2020	2,176,713	2,433,924	2,691,136	131232.8
07/31/2020	2,198,395	2,467,032	2,735,669	137062.1
January 08, 2020	2,219,872	2,500,080	2,780,288	142965.9
February 08, 2020	2,241,146	2,533,068	2,824,990	148942.6
March 08, 2020	2,262,222	2,565,998	2,869,774	154990.8
April 08, 2020	2,283,101	2,598,869	2,914,636	161108.9
May 08, 2020	2,303,787	2,631,681	2,959,574	167295.8
June 08, 2020	2,324,283	2,664,434	3,004,586	173549.9
July 08, 2020	2,344,591	2,697,130	3,049,669	179870.1
DATE	LOWER CI	FORECAST	UPPER CI	STD_ERR
August 08, 2020	2,364,714	2,729,767	3,094,820	186255.1
September 08, 2020	2,384,654	2,762,346	3,140,039	192703.7
October 08, 2020	2,404,414	2,794,868	3,185,322	199214.7
November 08, 2020	2,423,997	2,827,332	3,230,667	205787.1
December 08, 2020	2,443,404	2,859,739	3,276,074	212419.8
08/13/2020	2,462,637	2,892,088	3,321,539	219111.7
08/14/2020	2,481,700	2,924,381	3,367,062	225861.7
08/15/2020	2,500,594	2,956,617	3,412,639	232668.9
08/16/2020	2,519,321	2,988,796	3,458,270	239532.4
08/17/2020	2,537,883	3,020,918	3,503,954	246451.1
08/18/2020	2,556,282	3,052,985	3,549,687	253424.2
08/19/2020	2,574,521	3,084,995	3,595,469	260450.9
08/20/2020	2,592,600	3,116,950	3,641,299	267530.1
08/21/2020	2,610,522	3,148,848	3,687,174	274661.2
08/22/2020	2,628,289	3,180,691	3,733,094	281843.2
08/23/2020	2,645,902	3,212,479	3,779,057	289075.4
08/24/2020	2,663,363	3,244,212	3,825,061	296,357
08/25/2020	2,680,673	3,275,889	3,871,106	303687.3
08/26/2020	2,697,835	3,307,512	3,917,189	311065.4
08/27/2020	2,714,850	3,339,080	3,963,311	318490.7
08/28/2020	2,731,719	3,370,594	4,009,469	325962.5
08/29/2020	2,748,444	3,402,053	4,055,662	333,480
08/30/2020	2,765,027	3,433,458	4,101,890	341042.7

(*continued on next page*)

Table 3 (*continued*)

DATE	LOWER CI	FORECAST	UPPER CI	STD_ERR
08/31/2020	2,781,469	3,464,810	4,148,151	348649.7
January 09, 2020	2,797,771	3,496,107	4,194,443	356300.6
February 09, 2020	2,813,935	3,527,351	4,240,767	363994.6
March 09, 2020	2,829,962	3,558,541	4,287,121	371731.1
April 09, 2020	2,845,853	3,589,678	4,333,503	379509.5
May 09, 2020	2,861,611	3,620,762	4,379,914	387329.3
June 09, 2020	2,877,236	3,651,793	4,426,351	395189.8
July 09, 2020	2,892,729	3,682,772	4,472,814	403090.4
August 09, 2020	2,908,092	3,713,697	4,519,303	411030.7
September 09, 2020	2,923,326	3,744,571	4,565,815	419,010
October 09, 2020	2,938,433	3,775,392	4,612,351	427027.9
November 09, 2020	2,953,412	3,806,161	4,658,909	435083.7
December 09, 2020	2,968,267	3,836,878	4,705,489	443,177
09/13/2020	2,982,997	3,867,543	4,752,089	451307.2
09/14/2020	2,997,604	3,898,157	4,798,709	459473.9
09/15/2020	3,012,090	3,928,719	4,845,348	467676.6

Source: Authors' own computation.

Table 4
Forecast and 95% confidence interval for COVID-19 outbreak in Russia.

DATE	LOWER CI	FORECAST	UPPER CI	STD_ERR3
January 07, 2020	653,535	654,393	655,251	437.7359
February 07, 2020	659,024	660,807	662,590	909.7199
March 07, 2020	664,181	667,085	669,990	1481.792
April 07, 2020	669,040	673,227	677,413	2135.867
May 07, 2020	673,619	679,229	684,840	2862.586
June 07, 2020	677,928	685,091	692,254	3654.641
July 07, 2020	681,978	690,810	699,642	4506.393
August 07, 2020	685,775	696,385	706,994	5413.246
September 07, 2020	689,326	701,813	714,301	6371.346
October 07, 2020	692,635	707,095	721,554	7377.383
November 07, 2020	695,707	712,226	728,746	8428.454
December 07, 2020	698,545	717,208	735,870	9521.978
07/13/2020	701,152	722,037	742,921	10655.63
07/14/2020	703,531	726,712	749,893	11827.27
07/15/2020	705,684	731,232	756,781	13034.97
07/16/2020	707,615	735,597	763,579	14276.89
07/17/2020	709,323	739,804	770,284	15551.35
07/18/2020	710,813	743,852	776,890	16856.75
07/19/2020	712,085	747,740	783,395	18191.59
07/20/2020	713,142	751,468	789,794	19554.42
07/21/2020	713,985	755,034	796,083	20943.87
07/22/2020	714,615	758,438	802,260	22358.65
07/23/2020	715,035	761,677	808,320	23797.49
07/24/2020	715,246	764,753	814,260	25259.17
07/25/2020	715,249	767,663	820,078	26742.52
07/26/2020	715,046	770,408	825,770	28246.4
07/27/2020	714,639	772,986	831,334	29769.71
07/28/2020	714,029	775,398	836,767	31311.38
07/29/2020	713,217	777,642	842,066	32870.36
07/30/2020	712,205	779,718	847,230	34445.63
07/31/2020	710,996	781,625	852,255	36036.18
January 08, 2020	709,589	783,365	857,140	37641.05
February 08, 2020	707,988	784,935	861,882	39259.28
March 08, 2020	706,193	786,336	866,479	40889.93
April 08, 2020	704,207	787,568	870,929	42532.08
May 08, 2020	702,030	788,631	875,232	44184.84
June 08, 2020	699,665	789,524	879,384	45847.3
July 08, 2020	697,114	790,249	883,384	47518.61
DATE	LOWER CI	FORECAST	UPPER CI	STD_ERR
August 08, 2020	694,378	790,804	887,230	49197.9
September 08, 2020	691,459	791,190	890,922	50884.34
October 08, 2020	688,359	791,408	894,457	52577.09
November 08, 2020	685,079	791,457	897,835	54275.33
December 08, 2020	681,623	791,338	901,053	55978.26
08/13/2020	677,991	791,051	904,112	57685.09
08/14/2020	674,185	790,597	907,010	59395.04
08/15/2020	670,209	789,977	909,745	61107.34
08/16/2020	666,063	789,190	912,318	62821.23
08/17/2020	661,750	788,238	914,726	64535.97
08/18/2020	657,272	787,121	916,970	66250.81
08/19/2020	652,631	785,840	919,049	67965.05

Table 4 (*continued*)

DATE	LOWER CI	FORECAST	UPPER CI	STD_ERR3
08/20/2020	647,830	784,396	920,963	69677.95
08/21/2020	642,870	782,790	922,710	71388.83
08/22/2020	637,755	781,022	924,290	73096.99
08/23/2020	632,485	779,094	925,703	74801.74
08/24/2020	627,065	777,007	926,949	76502.43
08/25/2020	621,495	774,761	928,027	78198.38
08/26/2020	615,778	772,358	928,937	79888.95
08/27/2020	609,918	769,799	929,680	81573.5
08/28/2020	603,916	767,086	930,255	83251.4
08/29/2020	597,775	764,219	930,663	84922.04
08/30/2020	591,497	761,200	930,903	86584.81
08/31/2020	585,085	758,030	930,976	88239.11
January 09, 2020	578,542	754,712	930,882	89884.37
February 09, 2020	571,870	751,246	930,621	91519.99
March 09, 2020	565,072	747,633	930,195	93145.44
April 09, 2020	558,150	743,877	929,603	94760.14
May 09, 2020	551,108	739,977	928,846	96363.57
June 09, 2020	543,948	735,937	927,925	97955.19
July 09, 2020	536,673	731,757	926,841	99534.49
August 09, 2020	529,285	727,439	925,593	101,101
September 09, 2020	521,788	722,986	924,184	102654.1
October 09, 2020	514,184	718,399	922,615	104193.4
November 09, 2020	506,476	713,680	920,885	105718.5
December 09, 2020	498,667	708,832	918,996	107228.8
09/13/2020	490,760	703,855	916,950	108,724
09/14/2020	482,758	698,753	914,748	110203.5
09/15/2020	474,664	693,527	912,390	111,667

Source: Authors' own computation.

alcohol policies [30] among others. Time series modelling of infectious disease specially COVID-19 has been reported by several researchers [4,7,31–38].

2. Methodology

COVID-19 daily data of all reported cases were taken from the Worldometers website (worldometers.info/coronavirus/#countries). Data for India was of primary interest but data for the other two countries above and below India in the severity of epidemic were also studied to have a comparison of the epidemic and also investigate the onset of flattening of the curve. Daily data from 15 February to June 30, 2020 was collected and analysed separately for each country. We used data 30th June for modelling and then 77 days out of sample forecast was done based on the ARIMA models fitted to the data. Actual data from 1st to 7th was used to compute the accuracy and forecast error.

2.1. Box Jenkins procedure

Box and Jenkins (1971) popularised a method which combines both autoregressive (AR) and moving average (MA) models. An ARMA (p,q) model is a combination of AR(p) and MA(q) models and is best used for univariate time series modelling. In AR(p) model the future value of a variable is assumed to be dependent upon a linear combination of p past observations and a random error term. Mathematically and AR(p) model can be expressed as follows-

$$Y_t = c + \phi_1 y_{t-1} + \phi_{2yt-2} + \phi_3 y_{t-3} + \phi_4 y_{t-4} + \ldots . + \phi_p y_{t-p} + \varepsilon_t$$

Yt and ε_t are the actual value and the error terms at time period t, ϕ_i (i = 1,2,3,4 ….) are model parameters and c is a constant. Integer p is known as the order of the model. Unlike AR(p) model an MA(q) model uses past errors as explanatory variables. The MA(q) model is given below-

$$Y_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \theta_4 \varepsilon_{t-4} + \ldots . + \theta_p \varepsilon_{t-q} + \varepsilon_t$$

Table 5
Forecast and 95% confidence intervals for COVID-19 outbreak in Spain.

DATE	LOWER CI	FORECAST	UPPER CI	STD_ERR
January 07, 2020	295,067	296,504	297,942	733.5741
February 07, 2020	294,084	296,695	299,307	1332.366
March 07, 2020	292,753	296,853	300,952	2091.702
April 07, 2020	291,496	297,115	302,735	2867.293
May 07, 2020	290,095	297,437	304,779	3745.853
June 07, 2020	288,393	297,774	307,155	4786.492
July 07, 2020	286,419	298,102	309,786	5960.999
August 07, 2020	284,043	298,345	312,648	7297.349
September 07, 2020	281,496	298,554	315,611	8702.891
October 07, 2020	278,836	298,739	318,642	10154.9
November 07, 2020	276,180	298,962	321,744	11623.88
December 07, 2020	273,539	299,246	324,954	13116.37
07/13/2020	270,811	299,568	328,326	14672.39
07/14/2020	267,937	299,903	331,869	16309.53
07/15/2020	264,825	300,198	335,571	18047.73
07/16/2020	261,506	300,447	339,388	19868.05
07/17/2020	258,046	300,664	343,282	21744.3
07/18/2020	254,525	300,883	347,241	23652.35
07/19/2020	251,003	301,140	351,277	25580.76
07/20/2020	247,457	301,439	355,421	27542.25
07/21/2020	243,841	301,766	359,691	29554.04
07/22/2020	240,084	302,085	364,086	31633.81
07/23/2020	236,156	302,372	368,587	33784.09
07/24/2020	232,078	302,623	373,169	35993.12
07/25/2020	227,903	302,859	377,814	38243.43
07/26/2020	223,688	303,107	382,526	40520.45
07/27/2020	219,457	303,387	387,318	42822.67
07/28/2020	215,190	303,700	392,209	45158.72
07/29/2020	210,843	304,024	397,204	47541.89
07/30/2020	206,374	304,335	402,296	49980.92
07/31/2020	201,770	304,619	407,467	52474.53
January 08, 2020	197,055	304,879	412,703	55013.3
February 08, 2020	192,270	305,134	417,999	57585.04
March 08, 2020	187,451	305,406	423,362	60182.65
April 08, 2020	182,605	305,705	428,805	62807.22
May 08, 2020	177,713	306,024	434,336	65466.29
June 08, 2020	172,739	306,347	439,955	68168.54
July 08, 2020	167,658	306,654	445,651	70917.92
DATE	LOWER CI	FORECAST	UPPER CI	STD_ERR
August 08, 2020	162,468	306,940	451,413	73711.6
September 08, 2020	157,194	307,213	457,232	76541.85
October 08, 2020	151,865	307,488	463,111	79400.95
November 08, 2020	146,502	307,779	469,056	82285.62
December 08, 2020	141,106	308,091	475,077	85198.3
08/13/2020	135,655	308,416	481,177	88144.96
08/14/2020	130,124	308,737	487,351	91130.88
08/15/2020	124,500	309,045	493,590	94157.34
08/16/2020	118,788	309,338	499,887	97220.94
08/17/2020	113,008	309,624	506,239	100315.8
08/18/2020	107,183	309,916	512,649	103437.3
08/19/2020	101,322	310,224	519,125	106584.3
08/20/2020	95,421	310,546	525,671	109759.7
08/21/2020	89,462	310,874	532,287	112967.7
08/22/2020	83,427	311,198	538,968	116211.4
08/23/2020	77,312	311,509	545,706	119490.5
08/24/2020	71,123	311,810	552,497	122801.8
08/25/2020	64,879	312,110	559,342	126,141
08/26/2020	58,592	312,418	566,244	129505.5
08/27/2020	52,269	312,739	573,209	132895.5
08/28/2020	45,900	313,070	580,239	136313.5
08/29/2020	39,473	313,402	587,332	139762.5
08/30/2020	32,975	313,729	594,482	143244.1
08/31/2020	26,407	314,046	601,684	146,757
January 09, 2020	19,778	314,358	608,938	150298.7
February 09, 2020	13,099	314,671	616,243	153866.1
March 09, 2020	6381	314,993	623,605	157,458
April 09, 2020	−377	315,325	631,026	161075.2
May 09, 2020	−7182	315,663	638,507	164719.7
June 09, 2020	−14045	316,000	646,045	168393.4
July 09, 2020	−20972	316,331	653,635	172096.9
August 09, 2020	−27961	316,657	661,275	175828.7
September 09, 2020	−35004	316,979	668,963	179586.8
October 09, 2020	−42091	317,306	676,703	183369.2
November 09, 2020	−49217	317640	684,497	187175.5
December 09, 2020	−56385	317,981	692,347	191006.6
09/13/2020	−63601	318,326	700,252	194864.2
09/14/2020	−70873	318,668	708,210	198749.2
09/15/2020	−78203	319,006	716,216	202661.6

Source: Authors' own computation.

Here μ is the mean of the series, θj(j = 1,2,3 … q) are model parameters and is the order of the model. Mathematically an ARMA (p,q) model is represented as follows-

$$Y_t = c + \mu + \phi_1 y_{t-1} + \phi_{2}y_{t-2} + \phi_3 y_{t-3} + \phi_4 y_{t-4} + \ldots . + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \theta_4 \varepsilon_{t-4} + \ldots . + \theta_p \varepsilon_{t-q} + \varepsilon_t$$

The AR and MA can only be applied to a univariate stationary times series. To test the stationarity of a times series we need to test for the presence of unit root. If the series is not stationary in level, we need to differentiate it d (d = 1,2,3 …) times to make it stationary. Such a time series model is called an ARIMA (p,d,q) model.

2.2. ARIMA modelling steps

1. The first step is to check for the stationarity of the times series. This can be done by graphically plotting the series or conducting Augmented Dicky Fuller Test (ADF).
2. Identification of the model. Graphically the AR and MA terms can be deduced from the Autocorrelation function (ACF) and partial autocorrelation function (PACF) plots.
3. ARIMA parameters are estimated by least square method. EVIEWS 8 and JMulti software were used. While EVIEWS required naming of the model(p,d,q) based on ACF and PACF plots, JMulti does the model specification automatically using the Hannan Rissanen model selection algorithm(1982). The best model is selected on the basis of AIC values.
4. The residual analysis is done.
5. Out of sample forecast is carried out based on data from February 15, 2020 to June 30, 2020. A 77 days forward forecast upto September 15, 2020 is done based on the model.
6. The procedure is repeated for the US, Brazil, Russia and Spain to check the model specification and forecasting accuracy for the five most severely affected countries.

3. Results

The first step was to test for unit root in all the five time series. A visual examination of the data plot suggested that the series were exponentially rising and were non stationary. Other than Russian time series of COVID incidence, all other series had to be differentiated. Augmented Dickey Fuller test was conducted to establish that Russian series was stationary in level while Brazil was integrated in the first order and the remaining three series namely India, Spain and US were integrated in second order. The model specification determined by Hannan Rissanen algorithm [29] was India (4,2,4), Brazil (3,1,2), Russia (3,0,0), Spain (4,2,4) and US (1,2,1) respectively. The residuals of the ARIMA series were plotted and found to be stationary.

The ARIMA models were then used to forecast the out of sample COVID outbreak for 77 days up to September 15, 2020. The forecast values for each country are presented in Tables 1–5. The graphical

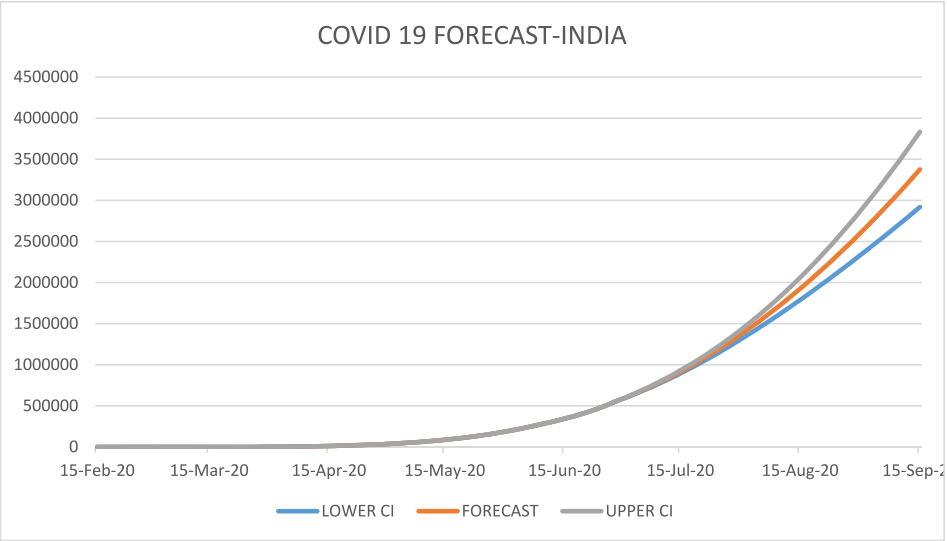

Fig. 1. Covid-19 forecast plot for India.

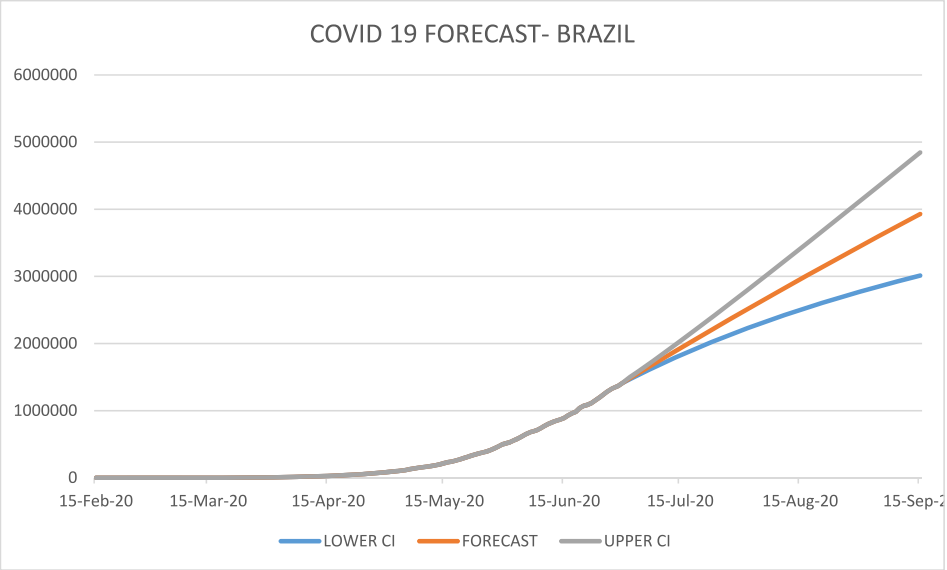

Fig. 2. Covid 19 forecast plot for Brazil.

plots with 95% confidence intervals are presented in Figs. 1–5. We compared the actual data from 1st July to 18th July and checked the forecast efficiency using mean absolute deviation (MAD) and the mean absolute percentage error (MAPE). The MAD was lowest for Spain followed by Russia whereas India, Brazil and US exhibited increasing absolute deviations indicating that actual forecasts lean towards the upper bound of the forecast. In other words, the forecast indicated worsening situation and steepening of the case graph for India, Brazil and US in the days to come. A better measure of the forecast efficiency is the mean absolute percentage error (MAPE) which converts the absolute deviations as percentage of actual numbers. Percentage numbers are easily compared to have a relative estimate of the severity of the spread across the countries under consideration. MAPE for India, Brazil and US were 3.701%,1.844% and 2.885% respectively. It was lowest for Russia and Spain at 1.090% and 0.832% indicating a very tight forecast accuracy. The smaller numbers for Russia and Spain further indicate that the forecast is following the linear trend established by the past data. Spain has even dropped out of the top five countries in the world. Even though the MAPE numbers for US, India and Brazil are all less than 4.0%, the relatively larger numbers indicate a trend which is steepening and leaning towards the upper bound of the forecast. The MAPE numbers validate the accuracy of the forecast. The results are presented in Table 6.

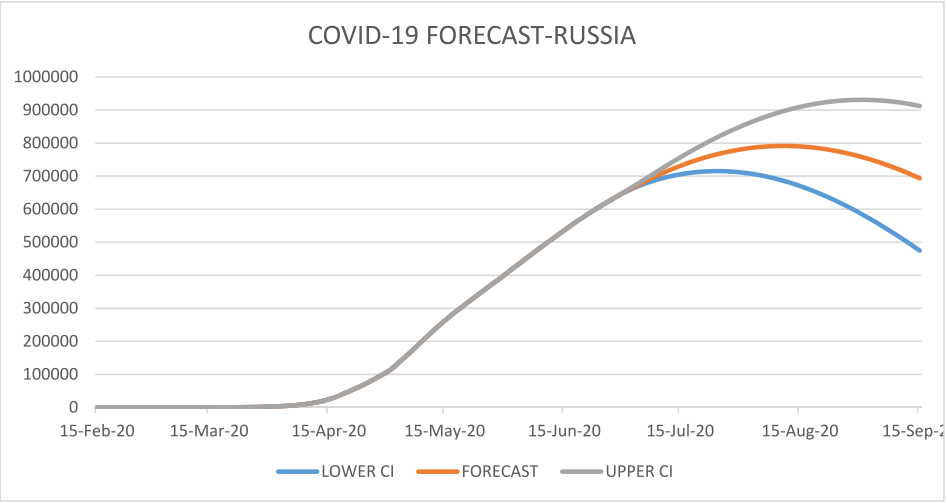

Fig. 3. Covid-19 forecast plot for Russia.

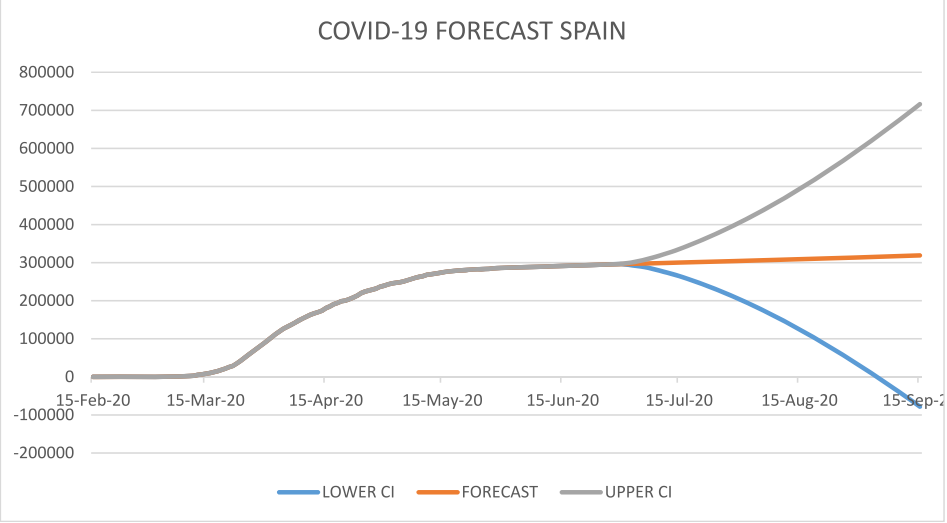

Fig. 4. Covid-19 forecast plot for Spain.

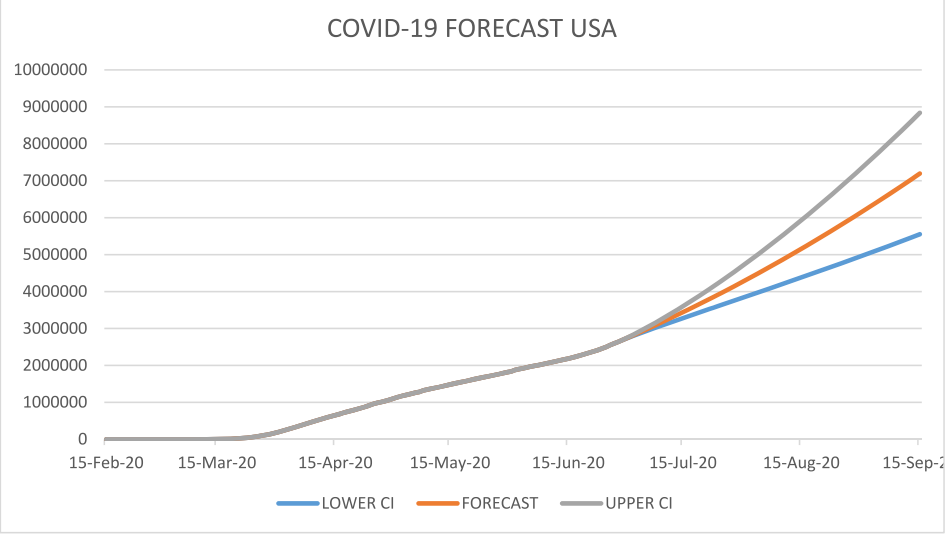

Fig. 5. Covid-19 forecast plot for US.

Table 6
Forecast and 95% confidence intervals for COVID-19 outbreak in US.

DATE	LOWER CI	FORECAST	UPPER CI	STD_ERR
January 07, 2020	2,765,075	2,772,875	2,780,674	3979.266
February 07, 2020	2,803,614	2,818,529	2,833,443	7609.775
March 07, 2020	2,841,873	2,864,474	2,887,074	11530.92
April 07, 2020	2,879,825	2,910,744	2,941,664	15775.66
May 07, 2020	2,917,466	2,957,348	2,997,230	20348.26
June 07, 2020	2,954,817	3,004,288	3,053,759	25240.75
July 07, 2020	2,991,899	3,051,562	3,111,226	30440.89
August 07, 2020	3,028,741	3,099,173	3,169,605	35935.59
September 07, 2020	3,065,365	3,147,119	3,228,874	41712.28
October 07, 2020	3,101,794	3,195,401	3,289,008	47759.35
November 07, 2020	3,138,051	3,244,019	3,349,986	54066.19
December 07, 2020	3,174,153	3,292,972	3,411,791	60623.19
07/13/2020	3,210,117	3,342,261	3,474,405	67421.61
07/14/2020	3,245,959	3,391,885	3,537,812	74453.5
07/15/2020	3,281,694	3,441,846	3,601,998	81711.6
07/16/2020	3,317,334	3,492,142	3,666,950	89189.26
07/17/2020	3,352,892	3,542,774	3,732,656	96880.35
07/18/2020	3,388,377	3,593,741	3,799,104	104779.2
07/19/2020	3,423,802	3,645,044	3,866,286	112880.6
07/20/2020	3,459,175	3,696,683	3,934,191	121179.7
07/21/2020	3,494,505	3,748,657	4,002,809	129671.8
07/22/2020	3,529,801	3,800,967	4,072,134	138352.9
07/23/2020	3,565,070	3,853,613	4,142,157	147218.8
07/24/2020	3,600,319	3,906,595	4,212,870	156265.9
07/25/2020	3,635,557	3,959,912	4,284,267	165490.5
07/26/2020	3,670,788	4,013,565	4,356,342	174889.5
07/27/2020	3,706,020	4,067,554	4,429,087	184459.5
07/28/2020	3,741,257	4,121,878	4,502,498	194197.7
07/29/2020	3,776,507	4,176,538	4,576,569	204101.1
07/30/2020	3,811,774	4,231,533	4,651,293	214167.2
07/31/2020	3,847,062	4,286,865	4,726,667	224393.2
January 08, 2020	3,882,378	4,342,532	4,802,686	234776.8
February 08, 2020	3,917,725	4,398,535	4,879,344	245315.6
March 08, 2020	3,953,108	4,454,873	4,956,638	256007.4
April 08, 2020	3,988,531	4,511,547	5,034,563	266849.9
May 08, 2020	4,023,998	4,568,557	5,113,116	277841.2
June 08, 2020	4,059,513	4,625,902	5,192,291	288979.3
July 08, 2020	4,095,081	4,683,584	5,272,087	300262.2
DATE	LOWER CI	FORECAST	UPPER CI	STD_ERR
August 08, 2020	4,130,703	4,741,600	5,352,498	311688.2
September 08, 2020	4,166,384	4,799,953	5,433,522	323255.3
October 08, 2020	4,202,128	4,858,641	5,515,155	334962.1
November 08, 2020	4,237,936	4,917,665	5,597,394	346806.8
December 08, 2020	4,273,814	4,977,025	5,680,236	358787.8
08/13/2020	4,309,763	5,036,720	5,763,678	370903.6
08/14/2020	4,345,786	5,096,751	5,847,717	383152.7
08/15/2020	4,381,886	5,157,118	5,932,350	395533.7
08/16/2020	4,418,067	5,217,820	6,017,574	408045.1
08/17/2020	4,454,329	5,278,858	6,103,387	420685.8
08/18/2020	4,490,677	5,340,232	6,189,787	433454.2
08/19/2020	4,527,113	5,401,942	6,276,770	446349.3
08/20/2020	4,563,639	5,463,987	6,364,335	459369.7
08/21/2020	4,600,257	5,526,368	6,452,478	472514.2
08/22/2020	4,636,969	5,589,084	6,541,199	485781.8
08/23/2020	4,673,779	5,652,136	6,630,494	499171.2
08/24/2020	4,710,687	5,715,524	6,720,361	512681.4
08/25/2020	4,747,697	5,779,248	6,810,799	526311.3
08/26/2020	4,784,809	5,843,307	6,901,805	540059.8
08/27/2020	4,822,027	5,907,702	6,993,377	553925.9
08/28/2020	4,859,352	5,972,433	7,085,513	567908.8
08/29/2020	4,896,786	6,037,499	7,178,212	582007.3
08/30/2020	4,934,330	6,102,901	7,271,472	596220.5
08/31/2020	4,971,988	6,168,639	7,365,290	610547.6
January 09, 2020	5,009,759	6,234,712	7,459,665	624987.6
February 09, 2020	5,047,647	6,301,121	7,554,596	639539.7
March 09, 2020	5,085,652	6,367,866	7,650,080	654202.9
April 09, 2020	5,123,777	6,434,947	7,746,116	668976.5
May 09, 2020	5,162,022	6,502,363	7,842,703	683859.7
June 09, 2020	5,200,391	6,570,115	7,939,838	698851.6
July 09, 2020	5,238,883	6,638,202	8,037,521	713951.5
August 09, 2020	5,277,501	6,706,625	8,135,750	729158.5
September 09, 2020	5,316,246	6,775,384	8,234,523	744,472
October 09, 2020	5,355,119	6,844,479	8,333,838	759891.3
November 09, 2020	5,394,123	6,913,909	8,433,696	775415.5
December 09, 2020	5,433,258	6,983,675	8,534,093	791,044
09/13/2020	5,472,525	7,053,777	8,635,029	806,776
09/14/2020	5,511,926	7,124,214	8,736,502	822,611
09/15/2020	5,551,463	7,194,987	8,838,512	838548.3

Source: Authors' own computation.

The graphs show that for the US, Brazil and India the situation does not seem to be coming under control. For Russia and Spain, the situation is seemingly under control and it can be said that the epidemic has reached the inflexion point. (see Table 7)

4. Discussion

India had controlled the spread of the pandemic very successfully until the May 31, 2020. Once the lockdown was lifted the migrant labourers and moving out from the hotspots of Delhi and Mumbai resulted in the explosion of the pandemic. The viral explosion that resulted from the lifting of lockdown has seen India break into the top ten affected countries. By the end of June India already touched third place after US and Brazil. The data for Spain showed a flattening of the curve while Russia showed a clear inflexion point and a downward trend. At the time of writing Spain has been pushed down by Peru and Chile.

At the current rate, we estimate India and Brazil to touch 1.38 million and 2.47 million mark respectively while the US is expected to touch 4.29 million mark by the end of July 2020. This modelling is expected to better prepare these countries for the burgeoning demand for healthcare facilities.

Though the results of the forecast were very agreeable, the ARIMA models suffer serious limitations in forecasting, characteristic of the time series models. Regression models take into account the causal variables but ARIMA models have found widespread and successful application in disease outbreak modelling. ARIMA forecast, built on the autoregressive nature of the time series coupled with corrective incremental adjustments, essentially, predicts a linear pattern and fails to predict a series with turning points. We have forecasted the COVID incidence up to September 15, 2020 assuming that no vaccine or other cure would be found by then. The exponentially rising graph of total cases indicates a possible community spread. Any successful medical intervention would, however, change the forecast significantly. Further, even without a vaccine, Russia and Spain have shown slowdown and flattening in the growth curves. If and when that will happen in case of the US, Brazil and India cannot be said based on this forecast. The ARIMA model does not help in predicting the onset of flattening of the pandemic cases.

5. Conclusion

ARIMA modelling of daily reported cases of COVID-19, in the top five countries showed a good forecast as measured by MAD and MAPE. The forecast could be used by the concerned governments to better manage and ramp up their healthcare preparedness for the pandemic.

Table 7
Forecast accuracy with mean absolute deviation (MAD) and mean absolute percentage error (MAPE).

DATE	INDIA		BRAZIL		RUSSIA		SPAIN		US	
	ACTUAL	FORECAST	ACTUAL	FORECAST	ACTUAL	FORECAST	ACTUAL	FORECAST	ACTUAL	FORECAST
1-Jul-20	605,220	605,084	1,453,369	1,448,644	654,405	654,393	296,739	296,504	2,778,452	2,772,875
2-Jul-20	627,168	623,844	1,543,341	1,484,229	661,165	660,807	297,183	296,695	2,835,684	2,818,529
3-Jul-20	649,889	642,773	1,543,341	1,517,010	667,883	667,085	297,625	296,853	2,890,588	2,864,474
4-Jul-20	673,904	663,129	1,578,376	1,550,697	674,904	673,227	297,625	297,115	2,935,770	2,910,744
5-Jul-20	697,836	683,792	1,604,585	1,585,926	681,261	679,229	297,625	297,437	2,982,928	2,957,348
6-Jul-20	720,346	704,039	1,626,071	1,621,272	687,862	685,091	298,869	297,774	3,040,833	3,004,288
7-Jul-20	743,481	725,228	1,674,655	1,655,902	694,230	690,810	299,210	298,102	3,097,084	3,051,562
8-Jul-20	769,052	747,529	1,716,196	1,690,134	700,792	696,385	299,593	298,345	3,163,318	3,099,173
9-Jul-20	794,842	769,604	1,759,103	1,724,467	707,301	701,813	300,136	298,554	3,224,892	3,147,119
10-Jul-20	822,603	791,821	1,804,338	1,758,950	713,936	707,095	300,988	298,739	3,297,170	3,195,401
11-Jul-20	850,358	815,308	1,840,812	1,793,378	720,547	712,226	301,670	298,962	3,359,174	3,244,019
12-Jul-20	879,466	839,280	1,866,176	1,827,649	727,162	717,208	302,352	299,246	3,417,795	3,292,972
13-Jul-20	907,645	862,980	1,887,959	1,861,818	733,699	722,037	303,033	299,568	3,483,584	3,342,261
14-Jul-20	937,487	887,448	1,931,204	1,895,950	739,947	726,712	303,699	299,903	3,549,632	3,391,885
15-Jul-20	970,169	912,992	1,970,909	1,930,047	746,797	731,232	304,574	300,198	3,621,637	3,441,846
16-Jul-20	1,005,637	938,511	2,014,738	1,964,081	752,797	735,597	305,935	300,447	3,695,025	3,492,142
17-Jul-20	1,040,457	964,142	2,048,697	1,998,039	759,203	739,804	307,335	300,664	3,770,012	3,542,774
18-Jul-20	1,077,864	990,871	2,075,246	2,031,931	765,437	743,852	307,335	300,883	3,833,271	3,593,741
MAD		33,614		33,277		8040.4		2530		100,761
MAPE		3.701%		1.844%		1.090%		0.832%		2.885%

Declaration of competing interest

On behalf of my co-authors, I, Dr. Alok Kumar Sahai, confirm that none of the authors have any conflict of interest to report.

References

[1] 10101998 Coronavirus Update (Live). Cases and 501644 deaths from COVID-19 virus pandemic - worldometer. Available at: https://www.worldometers.info/coronavirus/. [Accessed 28 June 2020].
[2] Alsudani RSA, Liu JC. The use of some of the information criterion in determining the best model for forecasting of thalassemia cases depending on Iraqi patient data using ARIMA model. J Appl Math Phys 2017;5:667–79. https://doi.org/10.4236/jamp.2017.53056.
[3] De P, Sahu D, Pandey A, Gulati BK, Chandhiok N, Shukla AK, et al. Post millennium development goals prospect on child mortality in India: an analysis using autoregressive integrated moving averages (ARIMA) model. Health 2016;8:1845–72. https://doi.org/10.4236/health.2016.815176.
[4] Wang Li Y, Peng B, Zhou R, Zhan CY, Liu Z. Mathematical modeling and epidemic prediction of COVID-19 and its significance to epidemic prevention and control measures. Ann Infect Dis Epidemiol 2020;5(1):1052. 2020.
[5] Chadsuthi S, Modchang C, Lenbury Y, Iamsirithaworn S, Triampo W. Modeling seasonal leptospirosis transmission and its association with rainfall and temperature in Thailand using time-series and ARIMAX analyses. Asian Pac J Trop Med 2012;5(7):539–46.
[6] Unkel S, Farrington CP, Garthwaite PH, Robertson C, Andrews N. Statistical methods for the prospective detection of infectious disease outbreaks: a review. J Roy Stat Soc 2011;175(1):49–82.
[7] Kotwal A, Yadav AK, Yadav J, Kotwal J, Khune S. Predictive models of COVID-19 in India: a rapid review. Med J Armed Forces India Online Pub 2020. https://doi.org/10.1016/j.mjafi.2020.06.001.
[8] Shmueli G, Burkom H. Statistical challenges facing early outbreak detection in biosurveillance. Technometrics 2010;52:39–51.
[9] Rolka HR. Preface. Statist. Med. 2011;30:401–2.
[10] Shewhart WA. Economic control of quality of manufactured product. Princeton: Van Nostrand Reinhold; 1931.
[11] Stroup D, Wharton M, Kafadar K, Dean A. Evaluation of a method for detecting aberrations in public health surveillance data. Am J Epidemiol 1993;137:373–80.
[12] Costagliola D, Flahault A, Galinec D, Garnerin P, Menares J, Valleron A-J. When is the epidemic warning cut-off point exceeded? Eur J Epidemiol 1994;10:475–6.
[13] Pelat C, Boëlle P-Y, Cowling BJ, Carrat F, Flahault A, Ansart S, Valleron A-J. Online detection and quantification of epidemics. BMC Med. Informat. Decsn Makng 2007;7. article 29.
[14] Parker RA. Analysis of surveillance data with Poisson regression: a case study. Stat Med 1989;8:285–94.
[15] Jackson MJ, Baer A, Painter I, Duchin J. A simulation study comparing aberration detection algorithms for syndromic surveillance. BMC Med. Informat. Decsn Makng 2007;7. article 6.
[16] Dominici F, Sheppard L, Clyde M. Health effects of air pollution: a statistical review. Int Stat Rev 2003;71:243–76.
[17] Stern L, Lightfoot D. Automated outbreak detection: a quantitative retrospective analysis. Epidemiol Infect 1999;122:103–10.
[18] Parker RA. Analysis of surveillance data with Poisson regression: a case study. Stat Med 1989;8:285–94.
[19] Serfling R. Methods for current statistical analysis of excess pneumonia-influenza deaths. Publ. Health Rep 1963;78:494–506.
[20] Healy MJR. A simple method for monitoring routine statistics. Statistician 1983;32:347–9.
[21] Ngo L, Tager IB, Hadley D. Application of exponential smoothing for nosocomial infection surveillance. Am J Epidemiol 1996;143:637–47.
[22] Chatfield C, Yar M. Holt-Winters forecasting: some practical issues. Statistician 1988;37:129–40.
[23] Box GEP, Jenkins GM. Time series analysis: forecasting and control. San-Francisco: Holden-Day; 1970.
[24] Choi K. An evaluation of influenza mortality surveillance, 1962-1979: 1, Time series forecasts of expected pneumonia and influenza deaths. Am J Epidemiol 1981;113:215–26.
[25] Helfenstein U. Box-Jenkins modelling of some viral infectious diseases. Stat Med 1986;5:37–47.
[26] Reis BY, Mandl KD. Time series modeling for syndromic surveillance. BMC Med. Informat. Decsn Makng 2003;3. article 2.
[27] Williamson G, Weatherby Hudson G. A monitoring system for detecting aberrations in public health surveillance reports. Stat Med 1999;18:3283–98.
[28] Heisterkamp SH, Dekkers ALM, Heijne JCM. Automated detection of infectious disease outbreaks: hierarchical time series models. Stat Med 2006;25:4179–96.
[29] Hannan EJ, Rissanen J. Recursive estimation of mixed autoregressive-moving average order. Biometrika 1982;69:81–94.
[30] Pridemore WA, Snowden AJ. Reduction in suicide mortality following a new national alcohol policy in Slovenia: an interrupted time-series analysis. Am J Publ Health 2009;99(5):915–20.
[31] 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. Diabet. Metabol. Syndr.: Clin Res Rev 2020;14(4):311–5. 2020.
[32] Roy A, Kar S. Nature of transmission of Covid19 in India. MedRxiv 2020:57–66. https://doi.org/10.1101/2020.04.14.20065821.
[33] Tiwari A. Modelling and analysis of COVID-19 epidemic in India. MedRxiv 2020:77–89. https://doi.org/10.1101/2020.04.12.20062794.
[34] Perone G. An ARIMA model to forecast the spread of COVID-2019 epidemic in Italy. SSRN Electron J 2020. https://doi.org/10.2139/ssrn.3564865. April.
[35] Modeling and predictions for COVID 19 spread in India. Retrieved from, https://www.researchgate.net/publication/340362418_Modeling_and_Predictions_for_COVID_19_Spread_in_India. [Accessed 26 June 2020].
[36] 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. https://doi.org/10.1371/journal.pone.0201987. 2018.
[37] Bhola J, Venkateswaran VR, Koul M. Corona epidemic in Indian context: predictive mathematical modelling. MedRxiv 2020. https://doi.org/10.1101/2020.04.03.20047175.
[38] He Z, Tao H. Epidemiology and ARIMA model of positive-rate of influenza viruses among children in Wuhan, China: a nine-year retrospective study. Int J Infect Dis 2018;74:61–70.