Original Article

COVID-19 Forecasting: A Statistical Approach

Arti Saxena[1], Falak Bhardwaj[2], Vijay Kumar[3]

Abstract

Background: SARS-coronavirus-2 is a new virus infecting people and causing COVID-19 disease. The disease is causing a worldwide pandemic. Although some people never develop any signs or symptoms of disease when they are infected, other people are at very high risk for severe disease and death. ***Objective*:** If we're able to intervene to prevent even some transmission, we can dramatically reduce the number of cases. And this is the public health goal for controlling COVID-19. ***Methods:*** This article initializes an approach for comparatively accurate values prediction of new cases and deaths for a particular day in order to be considered for preventive measures. The three statistical analysis methods considered for forecasting are Fbprophet, Moving average and the Autoregressive Integrated Moving Average algorithm. ***Results:*** The results obtained are in-line with the past and present trend of COVID-19 data collected from WHO website. ***Conclusion:*** The output is satisfactory for further consideration.

Keywords: *SARS-coronavirus-2; COVID-19, Statistical Analysis; Forecast; Fbprophet; Moving Average; Autoregressive Integrated Moving Average.*

Bangladesh Journal of Medical Science, Special Issue on COVID-19. 2021. Page : 85-96
DOI: https://doi.org/10.3329/bjms.v20i5.55401

Introduction

A coronavirus is the virus which causes COVID-19. Corona viruses are a huge, diverse virus community, and in order to be able to see them, a powerful microscope is needed. Coronavirus stands for crown. There are several different types of corona viruses, and a large variety of mammals and birds are infected every year and some also cause mild respiratory disease in humans, hence, corona viruses are not new. The virus that causes COVID-19, is called SARS coronavirus[2.] This virus originated in bats and is infected with this virus all the time. It developed the ability to jump between species and infect people. This is the third coronavirus known since 2002 that has developed. All of these corona viruses have emerged in bats and now attack individuals and can be transmitted from individual to individual. The first virus was named extreme acute respiratory syndrome, which emerged in Guangdong, China in 2002, or SARS coronavirus. The next virus that appeared in the Middle East in 2012 was the Middle Eastern Respiratory Syndrome Coronavirus, or MERS, which now induces human infections and minor outbreaks. More recently, SARS coronavirus 2 appeared in Wuhan, China, at the end of 2019. Since the virus is similar to the virus that caused the first SARS coronavirus, it was named SARS coronavirus[2]. In other cells, viruses have to survive, and so they replicate in those cells and begin to kill other cells in the body. Diabetes, hypertension, any form of lung disease such as asthma or emphysema or COPD, which is a chronic obstructive pulmonary condition, are present medical situations that raise the likelihood of serious illness in COVID-19. There is

1. ArtiSaxena
2. Falak Bhardwaj
3. Vijay Kumar
   FET, Manav Rachna International Institute of Research and Studies Faridabad, Haryana, India

Correspondence: ArtiSaxena, FET, Manav Rachna International Institute of Research and Studies, Faridabad, Haryana, India. Email: artisaxena.fet@mriu.edu.in

also a high chance of relentless COVID-19 illness for patients suffering from heart failure, liver disease and some form of kidney disease. High sensitivity is also seen in individuals who do not have good immune systems. Often, owing to certain drugs for treating some illnesses, such as hormones or those that affect their immune response, few persons have damaged immune systems. If a person is HIV-positive and he/she is on managed HIV infection treatment, so he/she is not at significantly high risk of serious COVID-19 disease. In general, it depends on the welfare of the person before he or she gets disease and access to treatment as well. Death from COVID-19 is rare among young people and stable individuals, but it does happen. Death is very frequent in elderly individuals or adults who have COVID-19 and it rises with age. In the US, 2% to 5% of individuals aged 65 to 75 years are expected to die from COVID-19. For those aged 75 to 85, the risk increased to 4 to 10 percent and was over 10 percent in people aged more than 85 years. Mechanical ventilation is expected of many prisoners with relentless lung disease. This is a ventilator-called breathing system. Ventilator protects lung cases as their body battles infection. Needy patients will be helped by this form of artificial breathing to help their lungs work so that their body can constantly receive the oxygen it requires when working hard to battle this infection. Nowadays, for patients with serious COVID-19 illness, this is the only regimen we have.

"The reproductive number is the number of individuals who would be affected by one infectious person if anyone with whom they have contact is vulnerable to the disease." It is a quantitative indicator of the disease's distribution. In other words, we can see that as a clever way to analyse how easily an illness in the population can spread. The higher reproductive number value correlates to the greater number of individuals that, during every epidemic, would get infection. Each person with COVID-19 can, on average, infect two to three other individuals. If we can rule out just one infection, that means that now each human infects only one person instead of two, so over time we can see a significant decrease in the number of people infected. Therefore, only one person is infected by the first person, and one person is infected by another person, and only one person is infected by that person again, so we end up with only four people who would be confirmed positive in this epidemic, instead of the fifteen that would have taken place if some infection at the initial point had not been ruled out. Thus, we will not interrupt the whole transmission in this manner, but we will surely have a major effect on slowing down the epidemic.

Material and Methods

Forecasting, however, requires ample historical data with no perfect prediction. Forecasts depend upon data reliability and variables of interest for prediction. Ankarali et al.[1] modeled the outbreak with different time series models and also predicted the indicators. They also evaluated the trends and seasonal effects. Petropoulos and Makridakis[3] in their paper introduced an even handed way to predict the continuation of the COVID-19. They assumed that the data used is reliable and also in future will pursue the past model of the disease. They described the timeline of a live forecasting exercise with massive potential implications for planning and decision making and provided objective forecasts for the confirmed cases of COVID-19.Roosa et al.[4] identified the initial cluster of severe pneumonia cases that triggered the COVID-19 epidemic in Wuhan, China in December 2019. They used the models that describe the empirical relationship of phenomena to each other in a way that is consistent with theory but not derived from it and validated these models during previous epidemics to develop the cause and assess short-term forecasts of the cumulative number of confirmed reported cases in Hubei province, the epicenter of the epidemic, and for the overall trajectory in China, excluding the province of Hubei. They collected daily reported cumulative confirmed cases for the 2019-nCoV outbreak for each Chinese province from the National Health Commission of China. Mean estimates and uncertainty bounds for both Hubei and other provinces have remained relatively stable in the last three reporting dates (February 7th – 9th). According to their prediction, epidemic has reached saturation in both Hubei and other provinces. Their decisions recommend that the constraint approach enforced in China was strongly abbreviating the transmission and that the epidemic growth had become slowed in nowadays. For various tasks like goal setting and anomaly detection, a common and obvious technique of data science is Forecasting. It involves serious challenges in resulting a high-quality reliable prediction. Taylor et al.[5] addressed these challenges with a practical way to forecast at scale. They proposed a modular regression model with interpretable parameters that can be intuitively adjusted by analysts with domain knowledge about the time series. Zakariah et al.[6] research paper on Laboratory Diagnostics in COVID-19: What We Know So Far provided a significant direction in this article.

Data Preprocessing

The data is extracted from Microsoft COVID tracker (https://www.bing.com/covid/local/india). The initial attributes were dates, new cases, total cases, new deaths, total deaths. Total cases and total deaths are the cumulative sum of values of cases and deaths respectively for a particular day. There were no significant errors in the data other than missing values. They were replaced with 0, as the ambiguity lied in the initial entry of the data, i.e., early days of COVID-19 when the values were almost zero.

Algorithms

It is a general assumption for all-time series analysis that the data is interpreted as deterministic and stochastic by the sum of two separate components. In this analysis, the very important feature is that the random noise is produced by individual shocks to the system, though it is always violated. And thus, it can be shown that the forecast approach such as exponential smoothing is very promising.

Fbprophet

It is an automatic forecasting procedure [Fig.2] for time series data. The point of attraction in this is the nonlinear trend which is fitted with weekly, yearly, and daily seasonality. The time complexity of this algorithm [Fig.1] is comparatively less than the other forecasting algorithm.

```
confirmed.columns = ['ds','y']
confirmed['ds'] = pd.to_datetime(confirmed['ds'])


m = Prophet(interval_width=0.99)
m.fit(confirmed)
future = m.make_future_dataframe(periods=30)

forecast = m.predict(future)
confirmed_forecast_plot = m.plot(forecast)
```

Fig. 1: *Code Snippet of FBProphet*

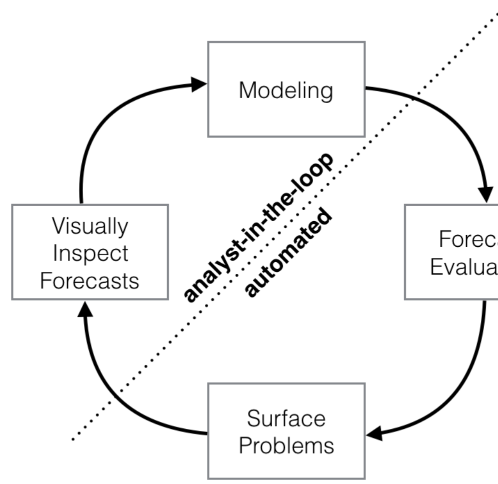

Fig. 2: Flowchart for Fbprophet

ARIMA

It is a common and commonly used statistical technique for forecasting time series. ARIMA is an alias that is used for Integrated Moving Average Auto-Regressive[Fig.4]. For time series forecasting, this is one of the simplest and most powerful machine learning algorithms [Fig.3]. It is a model class that captures in time series data a suite of various normal temporal structures.

$$Y_t = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + .. + \beta_0 Y_0 + \epsilon_t$$

$$Y_{t-1} = \beta_1 Y_{t-2} + \beta_2 Y_{t-3} + .. + \beta_0 Y_0 + \epsilon_{t-1}$$

Fig.3: Equation for ARIMA model.

Flowchart of ARIMA algorithm

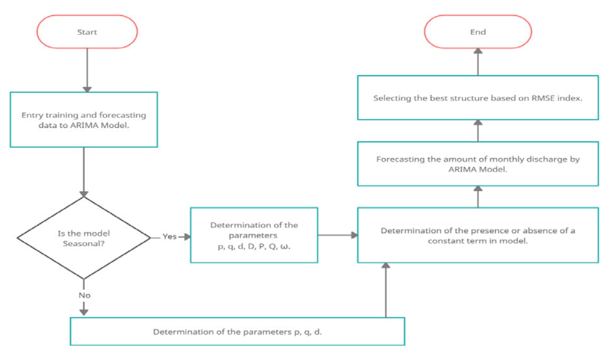

Fig.4: Flowchart for ARIMA

Moving Average

In math, the equation for evaluating data points by generating a sequence of averages of various subsets of the complete data set is a moving average. It is often referred to as a moving mean or rolling mean and is a kind of filter for finite impulse response[Fig.5]. For any stationary time series, it serves as a general model class and states that the non-deterministic weekly stationary time series can be interpreted as the sum of square weights that are the actual past and potential input values.

$$MA_n = \frac{\sum_{i=1}^{n} D_i}{n}$$

Where,

n = number of periods in the moving average

D_i = demand in period i

Fig. 5: Moving Average

Auto-Regression

Autoregression is a model of the time series that uses previous time phase measurements as feedback to a regression equation to estimate the value at the next point of time. It is a very basic principle that can contribute to detailed predictions on a number of issues in a time series.

Here, εis white noise process, β is the weights from values of input [Fig.6].

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + \epsilon_t$$

Fig. 6: Auto-Regression

Autocorrelation Function

Over successive time periods, autocorrelation reflects the degree of resemblance between a given time series and a lagged variant of itself [Fig. 7]. Autocorrelation tests the relationship of a variable's present value to its previous values. Arbitrary snapshots of the process at various points of time and analysing the general behaviour of the series determine the stationary existence of a time series. The initial tests tolerate the actions of the autocorrelation function, promising to classify the deviation from the stationarity is a solid and slowly dying ACF(Autocorrelation Function).

$$\hat{r}_k = \frac{\sum_{t=k+1}^{T}(y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2}$$

Fig. 7: Autocorrelation Function

Partial Autocorrelation Function

In order to get the order of the moving average phase, ACF is an important tool since it is supposed to break off after moving average order, i.e. q. In achieving the order of AR(Auto-regressor), i.e., p, ACF is not promising. As it would be a combination of exponential decay and damped sinusoidal expression. Therefore, it has been shown to be promising to examine this action of the time series PACF. The partial autocorrelation function (PACF) gives the partial association of a stationary time series with its own lagging values in time series analysis, regressing the time series values at all shorter lags [Fig. 8]. This contrasts with the role of autocorrelation, which does not regulate other lags.

$$PACF(y_i, y_{i-k}) = \frac{Covariance(y_i, y_{i-k}|y_{i-1}, y_{i-2}, \dots y_{i-k+1})}{\sigma_{y_i|y_{i-1}, y_{i-2}, \dots y_{i-k+1}} \sigma_{y_{i-k}|y_{i-1}, y_{i-2}, \dots y_{i-k+1}}}$$

Fig. 8: Partial Autocorrelation Function

Results

The forecasting results are as follows:

Fbprophet

The forecast for number of cases and number of death according to Fbprophet can be visualized in the following graph and the tabular data is available in Table 1.1 and 1.2, respectively. (Fig. 9.1, 9.2)

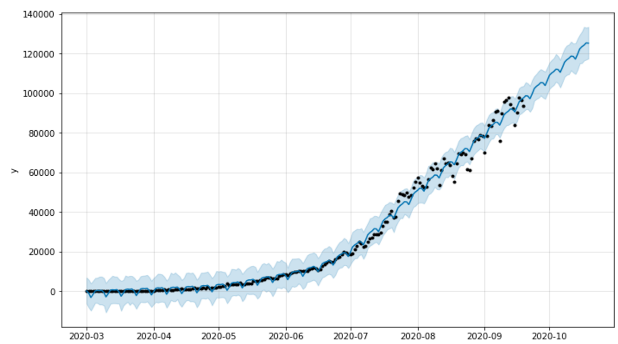

Figure 9.1: The forecast for number of cases.

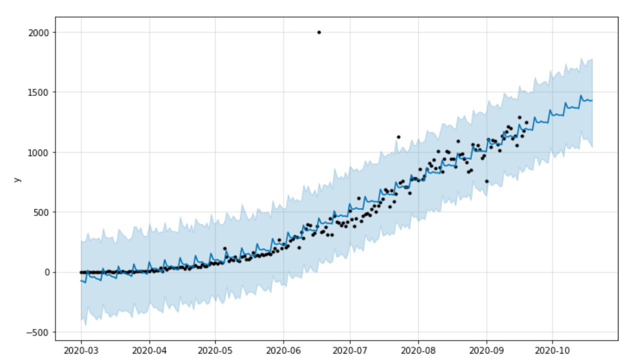

FIGURE 9.2: NUMBER OF DEATHS ACCORDING TO Fbprophet.

Moving Average Function (MA)

The forecast for number of cases and number of deaths according to Moving Average can be visualized in the following graph and the tabular data is available in Table 2.1 and 2.2, respectively. (figure 10)

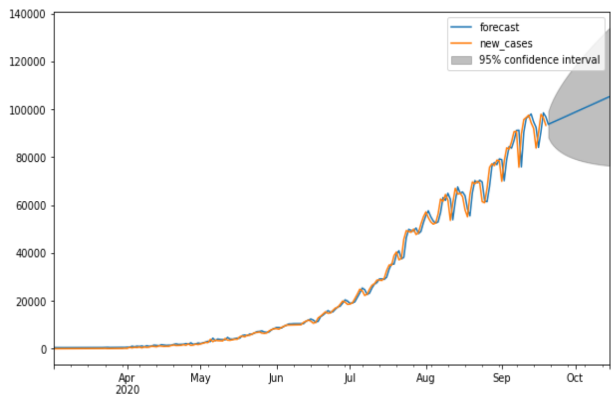

Figure 10.1: The forecast for number of casesaccording to Moving Average.

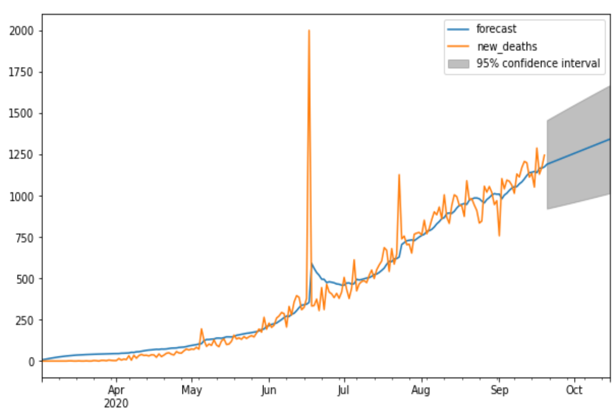

Figure 10.2: The forecast for number of deaths according to Moving Average.

Auto-Regressive Integrated Moving Average (ARIMA)

The forecast for number of cases and number of deaths according to ARIMA can be visualized in the following graph and the tabular data is available in Table 3.1 and 3.2, respectively (fig. 11).

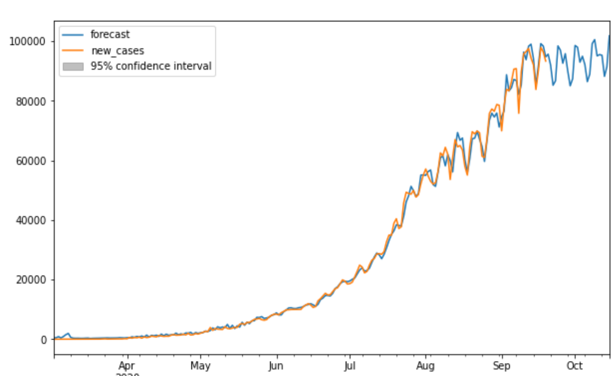

Figure 11.1: The forecast for number of cases according to ARIMA.

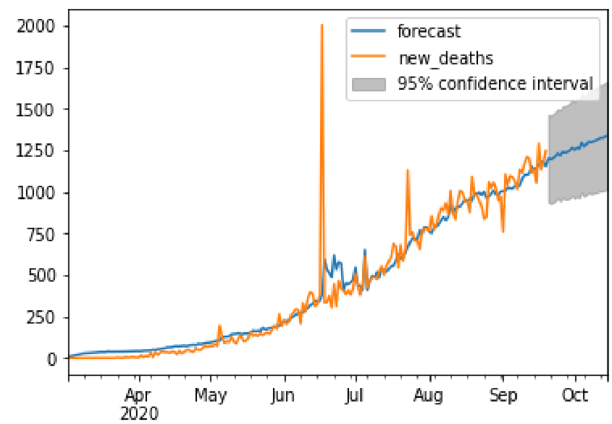

Figure 11.2: The forecast for number of deaths according to ARIMA.

Analysis

Data analysis is an essential part of forecasting and it is an integral part to determine the approach to proceed further with the forecasting. There are several approaches mentioned here, as original approach, logarithmic approach, exponential approach.

The two instances of Data considered are:

1. Number of Cases.
2. Number of Deaths.

The analysis includes:

1. Original Trend.
2. Logarithmic trend.
3. Exponential trend.
4. Moving Average.
5. Shifted Original Trend.
6. Trend Decompose.
7. MA Model.
8. Autocorrelation Function & Partial Autocorrelation Function.
9. ARIMA Model.

Number of Cases

For first instance, the number of cases were considered. It is the number of cases recorded per day in the interval of 01 March 2020 to 19 September 2020.

Original Trend

Here, as the original trend manifests, there was not much variation, shift or alteration in the pattern for a better part of initial two and a half months. The first variation is visible in May and the shift came after June. The high alteration and shift is visible in

Mid-July and hence till end there is a steep and peak variation in the data (fig. 12).

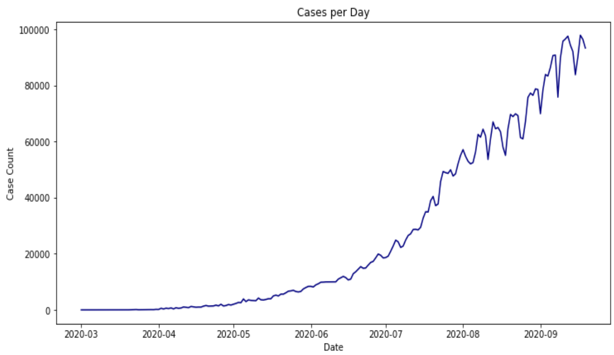

Figure 12.1: Original Trend of COVID-19 cases.

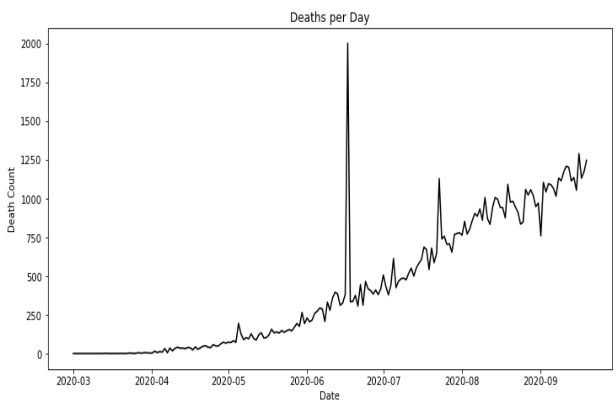

Figure 12.2: Original Trend of COVID-19deaths.

Logarithmic Trend

Logarithmic trend suggests that the high variation starts from the beginning and the trend becomes linear with time which is exactly in contrast with the original trend. Hence, Logarithmic trend was not considered for further steps (fig. 13).

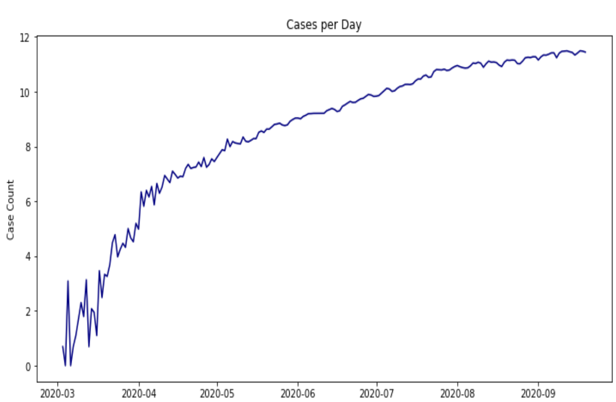

Figure 13.1: Logarithmic Trend of COVID-19 cases.

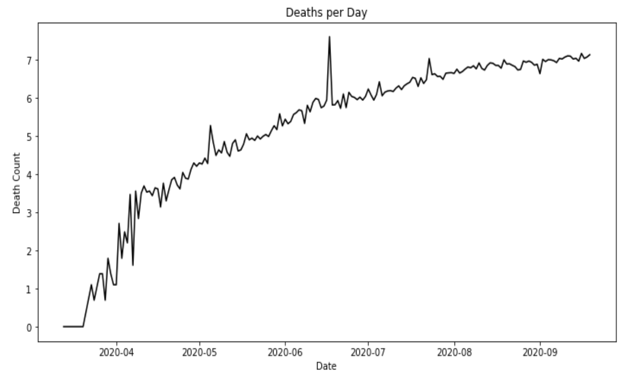

Figure 13.2: Logarithmic Trend of COVID-19 deaths.

Exponential Trend

The exponential trend seems to be stable w.r.t. the original trend. Hence, the exponential trend was also not considered for further analysis (fig. 14).

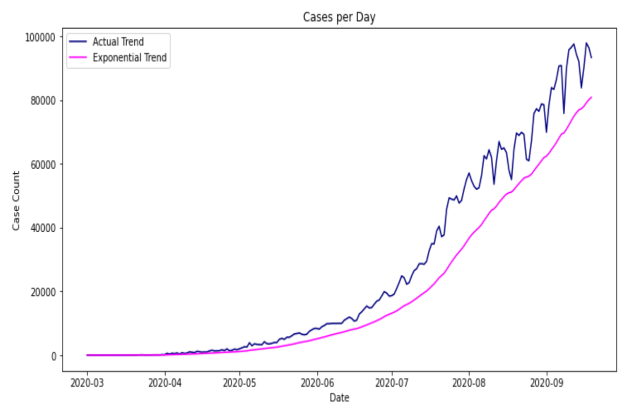

Figure 14.1: Exponential Trend of COVID-19 Cases.

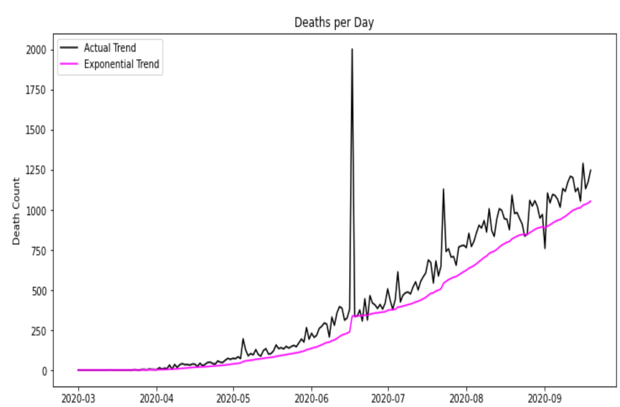

Figure 14.2: Exponential Trend of COVID-19Deaths.

Moving Average

The moving average of the original trend suggests that the data can be analyzed further and there is no irregularity or missing values (fig. 15).

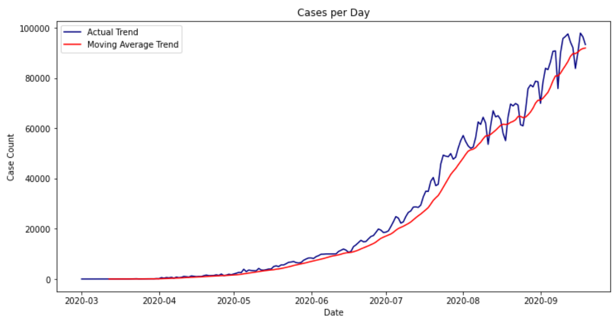

Figure 15.1: Moving Average of COVID-19Cases.

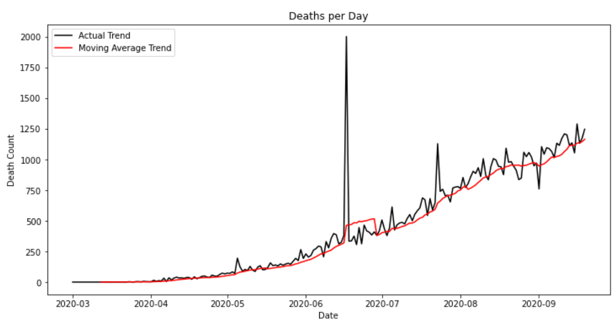

Figure 15.2: Moving Average of COVID-19 Deaths.

Shifted Original Trend

The original trend is shifted to extract the trend of the data irrespective of the count (fig. 16).

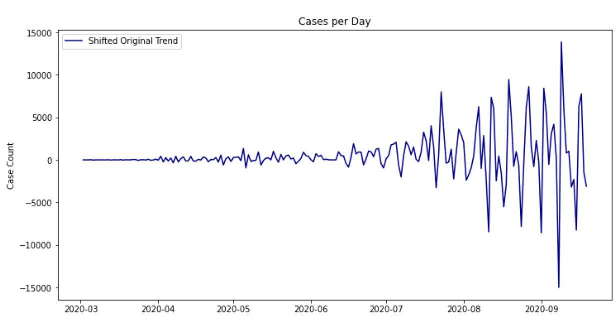

Figure 16.1: Shifted Original Trend of COVID-19 Cases.

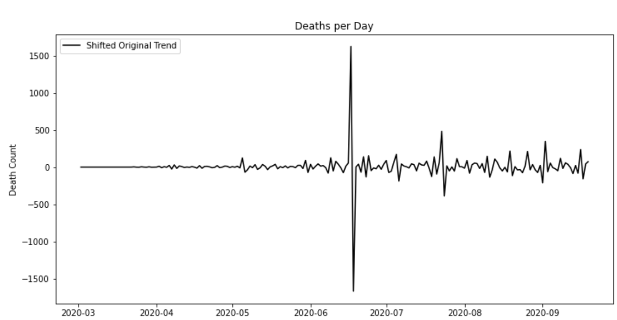

Figure 16.2:Shifted Original Trend of COVID-19 Deaths.

Trend Decompose

Trend decompose is decomposing the different attributes of trend to extract meaning out of the variations in the data. The trend decomposition is applied on the shifted data for better interpretation (figs. 17 and 18). Here, the trend is decomposed in three attributes:

1. General Trend: The increase and decrease in the data irrespective of seasonality and irregular variations.
2. Seasonal Trend: The trend of impact on data due to a particular season or for a certain period of time.
3. Residual Trend: The trend of the irregular values w.r.t. time and its impact on the data.

Figure 17: Trend Decompose for Number of Cases per day.

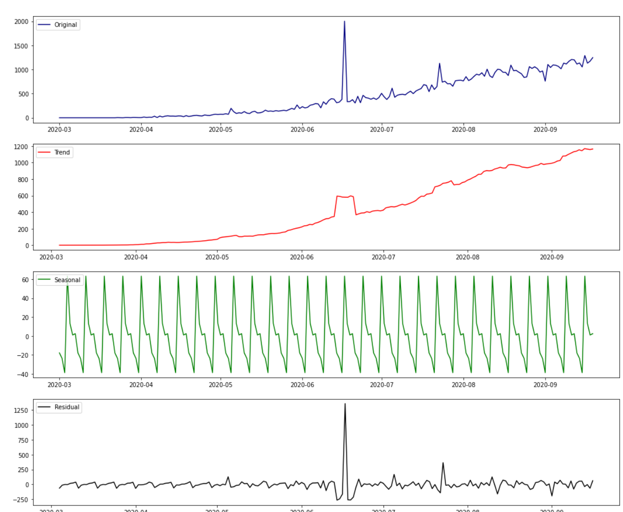

Figure 18: Trend Decompose for Number of Deaths per day.

MA Model:

The Moving Average model is applied by equating the value of p=0 in the ARIMA Model in order to eliminate the autoregressive factor (fig. 19). Here, RSS stands for Residual Sum of Square. It measures the variation of modelling errors. The general formula is:

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

where, y_i – the observed value

$\hat{y}_i$ – the value estimated by the regression line

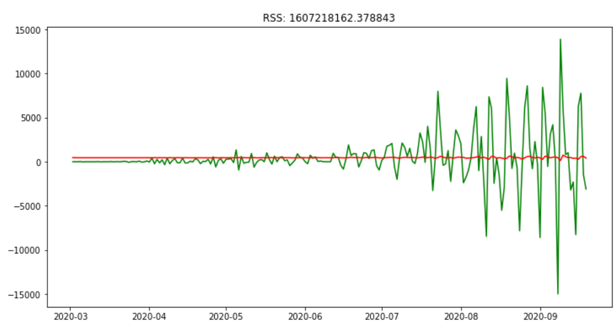

Figure 19.1: The Moving Average model for number of cases.

Figure 19.2: The Moving Average model for number of deaths.

Autocorrelation Function & Partial Autocorrelation Function

ACF(Autocorrelation Function) and PACF(Partial Autocorrelation Function), assist in calculating the order of moving average, p and auto-regressor, q. The *q-value* is the first point of contact. Similarly, the *p-value* is the very first cut point on the graph of PACF, which touches the y-axis at 0, i.e., the *p-value* and *q-value* are the first roots of PACF and ACF, respectively (fig. 20.1, 20.2).

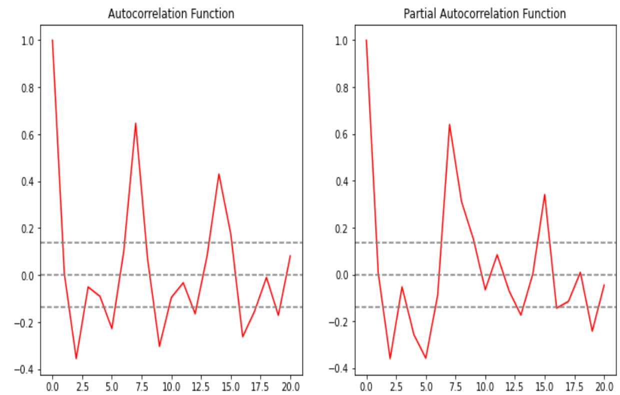

Figure 20.1: Autocorrelation Function & Partial Autocorrelation Function for Number of Cases.

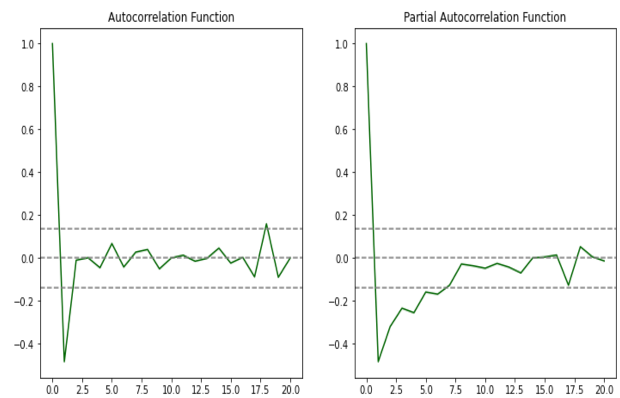

Figure 20.2: Autocorrelation Function & Partial Autocorrelation Function for Number of Deaths.

ARIMA Model

The p and q values generated from ACF and PACF are then combined to fit the model for ARIMA to evaluate the RSS value of the modelled data. The main goal of this operation is to fit the data with least possible RSS value by increasing or decreasing the p or q values (fig. 21).

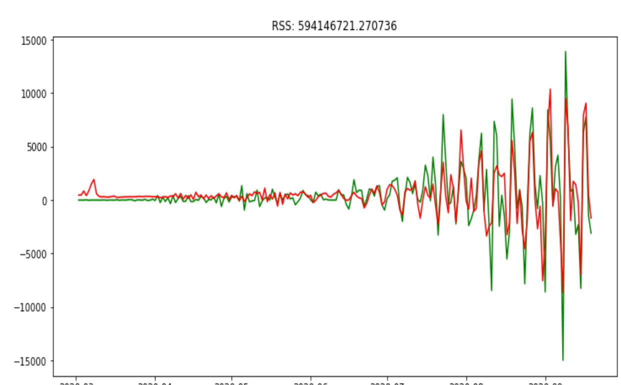

Figure 21.1: ARIMA Model Number of Cases.

Figure 21: ARIMA Model for Number of Deaths.

Discussion

The sole purpose of this article is to describe the COVID-19 situation as well as predicting the future condition and count of the said attributes as per the real-world data declared by WHO for INDIA, to which the research actually succeeded.The following research initiates a concrete explanation of the algorithms used for forecasting the number of cases and number of deaths.

Conclusion

This research concludes that the two main algorithms used i.e., ARIMA (Auto Regressive Integrated Moving Average) and Fbprophet are well developed for predictive analysis of seasonal as well as non-seasonal data as they explicitly capture the seasonality and non-seasonality out of the trend that data follows. The results are satisfactorily promising in view of the recent COVID-19 cases of deaths as well as new cases.

Conflict of Interest

The authors of this article have no conflict of Interest.

Consent for Publication

The authors of this articles are agreeing to give this articlefor publication in your esteemed journal **Recent Patents on Engineering.**

Acknowledgement

This research would like to acknowledge and express gratitude towards WHO (World Health Organization) and Microsoft Bing Blog for providing us with the real time updates related to the COVID-19 disease. From symptoms to case and death count, the research was provided with each and every aspect with accurate values without which it wouldn't be possible.

Availability of Data and Materials

The research is carried out on the basis of studies issued and published by the WHO (World Health Organization) and the data being provided by Microsoft Bing blog which provided real-time data of COVID-19 Case Count and Death Count.

Funding details:

Not Applicable

Appendix

Table1.1: Number of Case FBProphet

Date	Predicted	Lower Bound	Upper Bound
2020-09-20	98,656	91,960	106,027
2020-09-21	98,520	91,955	106,219
2020-09-22	97,138	90,693	103,875
2020-09-23	99,590	93,221	106,916
2020-09-24	102,283	96,027	108,711
2020-09-25	103,432	96,933	110,945
2020-09-26	104,203	96,728	111,382
2020-09-27	105,325	98,743	111,568
2020-09-28	105,190	97,648	112,002
2020-09-29	103,808	96,526	110,397
2020-09-30	106,259	98,445	112,753
2020-10-01	108,953	102,359	116,109
2020-10-02	110,101	103,298	117,130
2020-10-03	110,873	103,412	117,106
2020-10-04	111,995	104,734	119,321
2020-10-05	111,860	104,829	118,810
2020-10-06	110,478	103,179	117,739
2020-10-07	112,929	106,141	119,875
2020-10-08	115,622	108,155	121,981
2020-10-09	116,771	109,864	123,838
2020-10-10	117,542	108,536	125,066
2020-10-11	118,665	111,119	125,840
2020-10-12	118,529	111,194	126,207
2020-10-13	117,148	109,853	124,157
2020-10-14	119,599	111,905	127,125
2020-10-15	122,292	114,424	130,979
2020-10-16	123,441	115,722	131,105
2020-10-17	124,212	116,547	132,311
2020-10-18	125,335	116,880	133,935
2020-10-19	125,199	116,324	134,216

Table 1.2: Number of Deaths FBProphet

Date	Predicted	Lower Bound	Upper Bound
2020-09-20	1,185	865	1,485
2020-09-21	1,187	855	1,504
2020-09-22	1,181	851	1,493
2020-09-23	1,290	964	1,623
2020-09-24	1,246	947	1,576
2020-09-25	1,244	886	1,562
2020-09-26	1,255	941	1,572
2020-09-27	1,245	936	1,587
2020-09-28	1,247	914	1,589
2020-09-29	1,241	909	1,552
2020-09-30	1,350	1,036	1,682
2020-10-01	1,306	989	1,613
2020-10-02	1,304	953	1,633
2020-10-03	1,316	984	1,661
2020-10-04	1,305	995	1,670
2020-10-05	1,307	1,022	1,632
2020-10-06	1,301	959	1,629
2020-10-07	1,410	1,072	1,739
2020-10-08	1,367	1,030	1,663
2020-10-09	1,364	1,029	1,701
2020-10-10	1,376	1,060	1,701
2020-10-11	1,365	1,030	1,687
2020-10-12	1,368	1,042	1,704
2020-10-13	1,361	1,023	1,664
2020-10-14	1,470	1,178	1,784
2020-10-15	1,427	1,106	1,739
2020-10-16	1,425	1,110	1,714
2020-10-17	1,436	1,107	1,762
2020-10-18	1,426	1,073	1,769
2020-10-19	1,428	1,040	1,778

Table 2.1 Number of Cases MA

Date	Predicted	Lower Bound	Upper Bound
2020-09-20	93,725	88,196	99,254
2020-09-21	94,187	86,285	102,089
2020-09-22	94,649	84,938	104,360
2020-09-23	95,111	83,878	106,344
2020-09-24	95,573	83,002	108,145
2020-09-25	96,035	82,254	109,816
2020-09-26	96,497	81,605	111,390
2020-09-27	96,959	81,033	112,886
2020-09-28	97,422	80,524	114,319
2020-09-29	97,884	80,068	115,699
2020-09-30	98,346	79,657	117,034
2020-10-01	98,808	79,285	118,330
2020-10-02	99,270	78,947	119,592
2020-10-03	99,732	78,640	120,824
2020-10-04	100,194	78,359	122,028
2020-10-05	100,656	78,103	123,208
2020-10-06	101,118	77,870	124,366
2020-10-07	101,580	77,656	125,504
2020-10-08	102,042	77,461	126,623
2020-10-09	102,504	77,283	127,725
2020-10-10	102,966	77,121	128,811
2020-10-11	103,428	76,974	129,883
2020-10-12	103,890	76,840	130,940
2020-10-13	104,352	76,720	131,985
2020-10-14	104,815	76,611	133,018
2020-10-15	105,277	76,513	134,040
2020-10-16	105,739	76,427	135,051
2020-10-17	106,201	76,350	136,051
2020-10-18	106,663	76,283	137,043
2020-10-19	107,125	76,225	138,025

Table 2.2 Number of Deaths MA

Date	Predicted	Lower Bound	Upper Bound
2020-09-20	1192	924	1460
2020-09-21	1198	927	1468
2020-09-22	1204	931	1477
2020-09-23	1210	934	1485
2020-09-24	1216	938	1494
2020-09-25	1222	941	1503
2020-09-26	1228	945	1511
2020-09-27	1234	949	1520
2020-09-28	1240	952	1528
2020-09-29	1246	956	1537
2020-09-30	1252	960	1545
2020-10-01	1258	963	1553
2020-10-02	1264	967	1562
2020-10-03	1270	971	1570
2020-10-04	1276	974	1579
2020-10-05	1283	978	1587

Date	Predicted	Lower Bound	Upper Bound
2020-10-06	1289	982	1595
2020-10-07	1295	986	1604
2020-10-08	1301	989	1612
2020-10-09	1307	993	1620
2020-10-10	1313	997	1628
2020-10-11	1319	1001	1637
2020-10-12	1325	1005	1645
2020-10-13	1331	1009	1653
2020-10-14	1337	1013	1661
2020-10-15	1343	1017	1670
2020-10-16	1349	1020	1678
2020-10-17	1355	1024	1686
2020-10-18	1361	1028	1694
2020-10-19	1367	1032	1702

Table 3.1 Number of Cases ARIMA

Date	Predicted
2020-09-20	95,657
2020-09-21	91,996
2020-09-22	85,235
2020-09-23	86,704
2020-09-24	98,425
2020-09-25	96,987
2020-09-26	92,606
2020-09-27	95,818
2020-09-28	90,041
2020-09-29	85,037
2020-09-30	87,398
2020-10-01	98,580
2020-10-02	97,971
2020-10-03	92,966
2020-10-04	94,957
2020-10-05	92,185
2020-10-06	86,428
2020-10-07	88,935
2020-10-08	99,195
2020-10-09	100,542
2020-10-10	95,080
2020-10-11	95,522
2020-10-12	95,270
2020-10-13	88,215
2020-10-14	91,449
2020-10-15	101,808
2020-10-16	103,678
2020-10-17	97,753
2020-10-18	97,279
2020-10-19	98,183

Table 3.2 Number of Deaths ARIMA

Date	Predicted	Lower Bound	Upper Bound
2020-09-20	1,201	939	1,463
2020-09-21	1,194	930	1,459
2020-09-22	1,202	936	1,469
2020-09-23	1,213	946	1,481
2020-09-24	1,230	962	1,499
2020-09-25	1,217	944	1,490
2020-09-26	1,240	964	1,516
2020-09-27	1,236	955	1,516
2020-09-28	1,241	956	1,525
2020-09-29	1,247	961	1,534
2020-09-30	1,265	977	1,553
2020-10-01	1,252	962	1,542
2020-10-02	1,265	973	1,558
2020-10-03	1,259	964	1,554
2020-10-04	1,296	997	1,595
2020-10-05	1,272	971	1,574
2020-10-06	1,285	982	1,588
2020-10-07	1,301	997	1,605
2020-10-08	1,297	986	1,608
2020-10-09	1,303	991	1,615
2020-10-10	1,309	995	1,623
2020-10-11	1,316	1,000	1,632
2020-10-12	1,324	1,006	1,642
2020-10-13	1,324	1,004	1,644
2020-10-14	1,334	1,013	1,656
2020-10-15	1,338	1,014	1,662
2020-10-16	1,343	1,017	1,669
2020-10-17	1,349	1,021	1,676
2020-10-18	1,358	1,028	1,687
2020-10-19	1,359	1,028	1,690

References

1. Ankarali, H., Erarslan, N., &Pasin, O. Modeling and short-term forecasts of indicators for covid-19 outbreak in 25 countries at the end of march. *Bangladesh Journal of Medical Science,* 2020;**19,** 6-20. https://doi.org/10.3329/bjms.v19i0.47611
2. DOUGLAS C. M., CHERYL L. J. and MURAT K. *Introduction to Time Series Analysis and Forecasting*, A JOHN WILEY &. SONS, INC., PUBLICATION. 2008.
3. Petropoulos, F., &Makridakis, S. (2020). Forecasting the novel coronavirus COVID-19. *PloS one*, 2020;**15**(3), e0231236.
4. Roosa, K., Lee, Y., Luo, R., Kirpich, A., Rothenberg, R., Hyman, J. M., ... &Chowell, G. Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. *Infectious Disease Modelling*, 2020; **5**: 256-263.
5. Taylor, S. J., &Letham, B. Forecasting at scale. *The American Statistician*, ***72*(**1), 37-45.
6. Zakariah, S., Tengku Jamaluddin, T. Z., Mahayidin, H., Dahari, Z., & Ibrahim, R. (2020). Laboratory Diagnostics in COVID-19: What We Know So Far. *Bangladesh Journal of Medical Science*, 2018; **19**: S 93-S 100. https://doi.org/10.3329/bjms.v19i0.48199