Chapter 5
The Dual Impact of Lockdown on Curbing COVID-19 Spread and Rise of Air Quality Index in India

Swagatam Roy and Ahan Chatterjee

Abstract The number of reported cases in India has been scaling up in geometric progression despite the stringent lockdown norms imposed to keep people indoors since late March. Interestingly, on 31st March there were 1117 affected cases, 33,610 on 30th April and 511,478 till June 26—an unprecedented rise in the numbers. In the present research article, we propose a differential equation-based mathematical model for modeling India’s COVID-19 that incorporates the lockdown effect while looking at the future predictions in terms of the spread and the extent to which lockdown has been effective in India. We have estimated the growth of COVID-19 across India using modified SIR modeling, which is a Compartmental model in Epidemiology. Further, the use of SIQR model to estimate the growth of this disease across the country. Also, a constant factor has been introduced in the model to measure the number of corona-affected patients count due to any accidental mass crowd gathering. Along with that, we analyse the pollution level of India under three conditional scenarios viz. Pre-Lockdown, During Lockdown and After Lockdown. From the epidemiological evidences, it is evident that several pollutants like pm 2.5, NO2, SO2, O3, CO noxious effects of pollution. Here we will analyse the basic contributing factor of pollution and which majorly impacts AQI. We will also visualise the change of AQI in the context of the season or a particular time, i.e. during the festive season and Diwali pollution highly increases and it continues till April. During COVID-19, to avoid the contamination and spreading of the virus, Govt of India declared Lockdown and due to this all the industrial works gets stopped and the reduction of vehicular waste also reduced and thus the concentration of pollutants (μg/m3) decreases immensely. It can be interpreted that due to closure of industries and decreases in the number of vehicles, the concentration of the pollutants decreases thus it can be said that COVID-19 is a blessing to nature. But after reopening, i.e. unlock 1 the concentration, increases rapidly and immensely and from the reports, it is evident that in only in ecological regions, there is an increase of 400%. Thus after
unlock, people are avoiding social gathering maintaining social distance preferring own vehicle rather than the public vehicle. Also, sales of the cycle are increasing promoting greenery and the use of pollution-free vehicles. Is the Nightmare and pandemic situation helps to maintain ecological balance? In this paper we try to analyse these facts keeping different factors into consideration, we will deal with the trend and seasonality of AQI and predict using time series analysis and LSTM. We will build a model which will give a satisfactory output of the quality of air and how the pollutants hamper human health using mathematical models. The novelty in the paper is the comparative study of the models under two scenarios viz., what could have been the figure without lockdown and social distancing and with lock-down and social distancing, along with the AQI Analysis on the same said scenarios. Simultaneously, this is correlated with the predictions for the rise of air quality level.

Keywords LSTM · COVID-19 · SIQR model · AQI

5.1 Introduction

In 1720 Plague, followed by the Cholera outbreak in 1820 and Spanish Flu in 1920; it seems that in every 100 years a pandemic chases the existence of human race and no one has a clue to prevent that. As the famous saying goes,—‘History repeats itself’ and in 2020 we witnessed another pandemic with the name of Novel-Corona Virus (COVID-19). This disease was first identified in December 2019, in the Wuhan province of China and as the fate would have been, it has spread around the globe. The WHO had declared this as a public health emergency on 30 January 2020 and subsequently the first case of COVID-19 in India has been detected on that day itself. The seriousness and complexity of this pandemic could be assessed by the figure of infected people which is close to 1,00,000 as on date. Moreover, the outbreak of this disease has now spread to over 200 countries.

Following the trends in other countries, the Government of India announced a nation-wide lockdown on 24 March 2020 to prevent and slow down the spread of COVID-19 across the length and breadth of India. The lockdown has been imposed to buy time and get things ready to fight the pandemic, i.e. it has allowed the government to follow the rulebook of pandemics, which is to add more hospital beds, to increase the number of test kits, to increase the number of PPE kit for the frontline workers. In India, conditions are very much challenging as there is a high-density population across the country, unavailability of vaccines adding to the woes, thus, making it a herculean task as the numbers go up continuously [1].

The lockdown which has been imposed across the nation completed two months as on 24 May 2020. This lockdown has ‘a pro and a con’, where on the one side the spread of the virus is getting slowed due to implementation of social distancing. But on the other hand, there is a massive recession the economy is going to experience in the coming months. The GDP could possibly go down below the zero benchmark as the Reserve Bank of India has predicted. With regards to India’s employment
figures as ILO (International Labour Organisation) reports, only about 18 per cent belong to the salaried class; so for 467 million belonging to the self-employed class and the non-salaried class (including contractual and non-contractual and especially, the migrant labourers), the prolonged lockdown is becoming far more threatening than the danger of being affected with COVID-19 and many might die from hunger, fatigue, suicides etc.,—a question of life versus livelihood as the possibility of layoff looms large.

Susceptible-Infected-Recovered (SIR) model is an important tool to understand the coronavirus transmission-based statistical simulations. In this work of ours, we have taken the window of daily infected counts from 2 March 2020 to 31 June 2020 and fitted the Susceptible-Infectious-Quarantine-Recovered (SIQR) model into the data. We have introduced a constant factor in this model which will measure the growth of COVID-19 infected cases due to any accidental mass gathering. In addition, we have created a situation-based modeling to create the results. Mainly, two scenarios have been taken into account. In the first scenario, we have assessed the growth of coronavirus without any lockdown and social distancing and in the second case, we have assessed the same with lockdown and social distancing for comparison purposes.

The paper is structured as follows. Section 5.2 contains the Literature Review. In Sect. 5.3, we have briefly described our proposed SIQR model (after the incorporation of a constant) along with the numerical analysis and simulation results regarding the spread of coronavirus in India under two circumstances viz. with lockdown and social distancing and without social distancing and lockdown. Section 5.3, contains the AQI forecasting using Time Series Decomposition and using deep learning models. Section 5.4 presents a correlative study of lockdown and AQI across three scenarios viz. With no such COVID-19 scenario, what would be under lockdown and finally after lockdown scenario?

5.2 SIQR Dynamic Model

The study of COVID-19 spread can be executed by the fundamental rules of SIR modeling. The model was introduced by Kermack and McKendrik [2]. In this modeling approach, they divide the entire population into partial groups and study the contagion and spread of the disease across the groups using the parameter of the rate of change of the size of these groups. For our case, the basic model has been improvised by assumptions to accommodate the spread of the virus and different GDP growth rate simulations by bringing in more parameterised restrictions for better results in the Indian context [3, 4].
5.2.1 Theoretical Framework

In this section, we propose our dynamic model to predict the spread of COVID-19 across the nation. The spread of this virus is following an exponential growth rate path, creating a massacre around the globe. The aim of this section is to forecast the path of daily infected cases and to measure the extent of the epidemic in India [5].

In this paper, we use Susceptible-Infectious-Quarantine-Recovered (SIQR) model. In this model, we divide our entire population of 130 crores in India into four categories or sub-parts. The entire population is assumed to be \( N \) and it has been normalised to one for better assessment. The different categories in which we have divided are as follows—Susceptible \( S \), Infectious \( I \), Quarantine \( Q \), Removed (either recovered or deceased). The total number of active cases is being denoted by \( C \), and it’s the sum of Infectious and Removed, i.e. \( C = I + R \) [6].

The rate of change of these quantities has been shown using differential forms and they are denoted as \( \frac{dS}{dt} \), \( \frac{dI}{dt} \), \( \frac{dQ}{dt} \), \( \frac{dR}{dt} \), respectively. The equations of this model are shown below:

\[
\frac{dS}{dt} = -\frac{\beta_{1} S}{N} I \quad (5.1)
\]

\[
\frac{dI}{dt} = \sigma E - \gamma I \quad (5.2)
\]

\[
\frac{dQ}{dt} = \frac{\beta_{1} S}{N} I - \sigma E \quad (5.3)
\]

\[
\frac{dR}{dt} = \gamma r i \quad (5.4)
\]

\[
\frac{dD}{dt} = d \gamma I - \tau D \quad (5.5)
\]

\[
\frac{dC}{dt} = \sigma E \quad (5.6)
\]

\[
\frac{dC_n}{dt} = N_m \quad (5.7)
\]

where

\( \gamma = \) Infectious period time.
\( \gamma_r = \) Relation between infected population and infected one.
\( \sigma = \) Mean Latent Period.
\( d = \) Proportion of severe cases.
\( \tau = \) Mean duration of public reaction time.
\( N_m = \) Fraction of population, infected due to accidental mass gathering like the Jamat case.
\( \beta_t = \text{Transmission rate.} \)

The transmission rate is governed by the factor \( \alpha \), which is the government policies or measures to curb the spreading pandemic. The function of the transmission rate is represented by Eq. 5.8 [5].

\[
\beta_t = \beta_0(1 - \alpha)
\left(1 - \frac{D}{N}\right)^k
\]

(5.8)

where

- \( \beta_0 = \text{The numerical value is assumed from basic reproduction number.} \)
- \( \alpha = \text{i.e. } R_0 = \frac{\beta_0}{\gamma}. \)
- \( (1 - \frac{D}{N}) = \text{Impact of governmental and individual action through pandemic.} \)
- \( k = \text{Intensity of Self Reaction varied from } [0, 10^5], \text{ increasing over time.} \)
- \( D = \text{Public perception of risk w.r.t death and serious cases varied from } [0, 10^5] \text{ also.} \)

The above said parameters will vary from country to country even from state to state in line with the policies implemented by the governments as protective measures. In our study we have used a step function to define the transmission factor represented in Eq. 5.9 [7]. In this study, we have assumed that the virus is not mutating at a constant rate, thus taking the reaction of the system as constant and exogenous [8].

\[
R_t = \begin{cases} 
R_0 & 0 < t < t_1 \\
R_1 & t > t_1 
\end{cases}
\]

5.2.2 Estimation of Initial Transmission Rate

In India, the policy which has been followed is that the positive tested patient will be transferred to the quarantined facility immediately for a 14-day period. The infected people’s status is soon converted to quarantine status thus there is a change of dynamics in this process [1, 9]. In order to implement this factor in our model, we consider the total and the susceptible population as equivalent, i.e. \( \frac{S}{N} \approx 1. \)

\[
I + (\gamma + \sigma)I - \sigma(\beta - \gamma)I = 0 \quad (5.9)
\]

Integrating Eq. 5.9 we get,

\[
I_t = I_1 e^{-\frac{(\gamma + \sigma + \sqrt{(\gamma + \sigma)^2 + 4\sigma \beta})}{2}} + I_2 e^{-\frac{(\gamma + \sigma - \sqrt{(\gamma + \sigma)^2 + 4\sigma \beta})}{2}} \quad (5.10)
\]
Table 5.1: Parametric values used in our model

| Parameter | Description                        | Value              |
|-----------|------------------------------------|--------------------|
| $N_0$     | Initial population                 | 130 Crore          |
| $S_0$     | Initial susceptible population     | $0.9 \times N_0$   |
| $E_0$     | Exposed population for each infected | $24 I_0$ (assumed) |
| $I_0$     | Initial state of infected person   | 4                  |
| $\alpha$  | Lockdown and other action strength | varied             |
| $k$       | Intensity of people’s reaction     | 1117 (constant)    |
| $\sigma^{-1}$ | Latent period (mean)         | 3 days             |
| $\gamma^{-1}$ | Infectious period (mean)       | 6 days             |
| $d$       | Ratio of severe cases             | 0.26               |
| $\tau^{-1}$ | Duration of public reaction (mean) | 12 days            |

Now, the curve is being fitted to the data available for infected COVID-19 patients. In our model, we have varied the parameters for better visualisation across cases by simulating the same and choosing the best possible outcome.

Table 5.1 shows the parametric values which have been taken into consideration while modeling the SIQR epidemic model based on some recent epidemic and pandemic studies.

5.2.3 Simulation of Mathematical Modeling

In this section, the results of our mathematical modeling have been presented. The parametric values which have been used to assess our model should be treated as an average value for India [10]. At first, we have calculated the growth factor for India across the time duration taken. The growth factor will help us to analyse the transmission rate and $\alpha$ representing government action policies [11].

From Fig. 5.1 we can clearly see that the growth factor (as given below) of India is more than 1.5 (approximately). According to the rule, if the growth factor of a country is more than 1 then it denotes that the disease is spreading at a very rapid pace. (This graph is for scenario II, i.e. with lockdown and social distancing). Based on the growth factor calculated we have approached the transmission rates [12].

\[
\text{Growth Factor} = \frac{\text{Cases on day } t}{\text{Cases on day } t - 1}
\]

The initial value of the transmission rate $\beta_0$ in our model is taken as 0.50. In the first case, we have assumed that there would be no lockdown across the country. In this case, the value of $\alpha$ is taken as 0.8 as there is no government intervention and the
situation is normal [9, 13]. The value of $\alpha$ depends on the strict measures on social distancing and lockdown. In Figs. 5.2 and 5.3 we have shown a graph which will denote the scenario for ‘no lockdown’ in India [14].
From Figs. 5.2, and 5.3 we can see the scenario in India could have been much more devastating if there were no lockdown measures implemented under two circumstantial conditions, namely, with more crowds gathering and low crowd gathering. We can see that infected people count could have been crossed the 1 lakh mark ‘without lockdown and more mass gathering’ (Fig. 5.2) as compared to the ‘low crowd gathering’ (Fig. 5.3). These two cases under ‘without lockdown and social distancing’ could have led to more deaths across the nation much ahead of the expected time as seen from Figs. 5.2 and 5.3. Thus, from this simulation, we are confident that lockdown was needed to limit the spread of COVID-19 [15].

But is lockdown a permanent solution when we look at the fall in growth rate? We analyse this in the coming sections. In the next scene we have simulated the COVID-19 spread across India with social distancing and lockdown. In this case, the $\alpha$ value taken was very small near 0.2 [16, 17]. Lower value indicates there are stronger restrictions of lockdown across the country. In Fig. 5.4 we have presented the scenario with lockdown till October in the graph based on the assumption of taking the initial month to be April 2020. With lockdown measures, the numbers are expected to go down from mid-August onwards [18, 19].

### 5.3 Introduction to Air Pollution

Everything in this world is changing. Man lives within technology and technology is increasing gradually but still in this tech-savvy world the basic essential needs of a human life like air, water remains the same. As we know from the wise old saying that love is in the air but most pathetic thing is that air is polluted. So, it has to be ensured that pollution shouldn’t be the price of development and prosperity and being a responsible citizen, it is our duty as well as responsibility to be a part of the solution to stop the pollution. Basically, pollution is the consequence of the incapability of proper recycling of waste and pollutants. Pollutants are basically a substance that
instigates distasteful effects, or and hampers the natural resources of the environment. It has a long as well as short-term ravaging effect to the environment and hampers as well as destroys human amenities, health, etc. Among them some are biodegradable that means they get decomposed into organic matters by bacteria, fungus, etc., and some are non-biodegradable (like plastic) which means they are incapable or takes a huge amount of time for decomposition and even cannot decompose at all in some cases. These non-biodegradable wastes have more adverse effect. In the highly populated and fast developing countries in the world like India, the upswings and fluctuations in the economy led to environmental threats and the urbanisation as well as industrialisation paved the way to increase air pollution leading to social instability. From Fig. 5.5, a question can be raised that is the global growth-promoting the growth of air pollution?[20–23]

As per reports and the facts published worldwide, more than 2 million premature deaths are the adverse effect of air pollution only and more than 4 million deaths are consequence to the household pollutions only. From the different results, it is evident that 10% of total death in the world is only due to air pollution and so it is becoming a major threat to mankind and India is among the top five countries having the highest pollution. The other countries of southeast like Bangladesh, Pakistan, Nepal are highly polluted. In India, some of the polluted cities according to different reports are New-Delhi, Kanpur, Faridabad, Ghaziabad, Gaya (the report is considered mainly based on concentration on the pollutants. Also South as well as Southeast Asia, are facing a worse conditions. In Fig. 5.2, deaths per day from air pollution of two most densely populated countries in the world, i.e. India and China are taken into account. It is evident that the death rate is increasing. Air pollution can be said as the damaging effect caused to the air due to release of several toxic pollutants which is life-threatening and detrimental to mankind which may cause several chronic diseases and other problems like lung cancer, asthma, anoxaemia, corneal opacity,
irritation, etc. The lungs is the organ which is mainly responsible for the respiration and it is spongy in nature. Due to smoking and exposure to smoke as well as inhaling of toxic gases compelled the cells to divide and this led to the growth of tumours resulting in breathing problems and developing lung cancer and also asthma is also a major consequence which is designated by inflammation of lungs, obstruction in the path as well as high mucous formation rate. Also, environmental hazards like global warming, acid rain, temperature inversion are effects. Due to global warming, there is high climatic distress and rise in sea level has also become a major reason to worry. Acid rain has a high range of destructive effects on the environment as well as on human health. Several respiratory diseases like asthma, bronchitis, pneumonia are some of the harmful effects of it. Also, acid rain disturbs the PH of the water level by making it acidic and thus hampers the aquatic life. Apart from that, it has corrosion effect on marble and its marble statues. Several matrices used for calculation and understanding the pollution are PSI (Pollution Standard Index) which is used for the risk assessment. It is mainly based on the concentration (ug/m3) of pollutants. AQI (Air Quality Index) is the estimation of the condition of air and it is divided into six ranges. AQI is first introduced in 1986. Based on concentration, the calculation of AQI takes place over a particular averaging period. AQI is a Piecewise Linear function which is also known as segmented function of the concentration of the pollutants. In terms of mathematics, concentration can be transformed to AQI as [24, 25]:

\[
AQI = \frac{I_H - I_L}{C_H - C_L} \cdot (C - C_L) + I_L
\]

where \(C\) is the concentration of the pollutants.
Table 5.2 Showing quality of air based on AQI

| AQI value | Condition |
|-----------|-----------|
| 0–50      | Good      |
| 51–100    | Satisfactory |
| 101–200   | Moderate  |
| 201–300   | Poor      |
| 301–400   | Very poor |
| 401–500   | Bad       |

$C_L$, $C_H$ are concentration breakpoint.
$I_H$, $I_L$ are index breakpoint.

There are six ranges based on which the quality of the air is determined. It can be understood further by Table 5.2.

In these upcoming sections, we will try to understand the pollution from different sectors, the relationship among the pollutants and how they are affecting AQI and causing pollution. Later we will go into deep analysis and prediction will be done and we will have a view on how COVID-19 is affecting air pollution and how it has become a blessing in disguise to the environment.

## 5.4 Pollution from Different Sectors

From the smoke and ashes of a small cigarette to smoke of large chimneys of industries, there are pollutants everywhere. In the context of the current scenario, pollution is increasing with the increase in technology. In the next sections, we will have a brief walk through the different pollution sector prevailing in our environment [26].

### From vehicles:

Growth of the economy of a country is paving a way for the growth of different sectors. Many automobile sectors are growing and nowadays preferring more of private or self-owned vehicles over public transport which leads to the growth of sales of car worldwide. The emission of CO (Leading to production of carboxyhaemoglobin), Hydrocarbon, VOC, different oxides of nitrogen ($NO_x$) takes place and apart from that in heavy-duty vehicles, emission of particulate matters, lead particles takes place which is harmful to respiratory tracks [27].

### From Industries:

High growth in technology is developing the industries and thus developing a large amount of pollutants. From the industries and wastes of industries emission of different pollutants like ($NO_x$), sulphur dioxide ($SO_2$), ethylene, formaldehyde, toluene, benzene, etc., takes place. Also smoke from chimneys and wastes from the battery industries, ore and mining industries as well as burning of coal and fossil fuels like natural gas; petroleum is polluting the environment a lot [28–30].
From Household works:

Still today in many houses, cooking is done using the burning of biomass and kerosene which is a high source of pollution. Also, emission of CFC from AC and fridges causes a lot of pollution. These are some of the sources of indoor air pollution. Also RSPM, SPM is one of the key pollutants of household pollution.

5.5 Dataset Preparation

The dataset has been taken from Kaggle, IQAir, air-quality.com, aqicn and other different sources. The AQI value and concentration of other pollutants from June 2018 to June 2020 is taken into account. Doing basic analysis, we can conclude that AQI value changes with the season as during festival time, i.e. from September end to February end. There are numerous number of festivals and lighting of crackers promotes the rise in the value of AQI and this value again gets down from March to August. Also, it is observed that AQI is directly proportional to the concentration of the pollutants but each pollutant has a different contributing factor. This is further elaborated in the next section.[31, 32]

5.6 Understanding Basic Relationship of Pollutants with Pollution

From the data, we are trying to analyse the consequence and affecting the rate of different pollutants and how they are affecting AQI. In Fig. 5.6 the dataset of the pollutants is plotted [33–35].

From the correlation map, it is evident that particulate matters (PM 2.5, PM 10) NOx has a high correlation with AQI value. Top 6 are tabulated in Table 5.3

Fig. 5.6 Different pollutants along with AQI over time
Here, the spearman’s correlation is calculated. From Table 5.2 it is evident that particulate matter has a high positive correlation with AQI means that AQI value is directly proportional to the concentration of pollutants and it has a high impact on AQI value.

5.7 Detailed Analysis

Here, we will be mainly focusing on time series analysis, i.e. using the theoretical approach. In the later part for higher accuracy, we will be focusing on RNN and LSTM based modeling.

5.7.1 Time Series Analysis

Time series is one of the significant analytical tools which is widely used in the analysis of time-dependent data. It is a cluster of points taken at a regular interval of time over a period i.e. in lemons language in a single line, time series represents the points on a graph or listed data collected based on time periods. By doing so we can get data which can be misplaced during the collection of data at random points or times, thus taking elements and collecting data at a particular interval may help in the compilation of the data providing a better view to the problem. It helps to detect how the points are affected due to pacing of time we can make time series modules and develop them as it will make the data we have more explanatory and support the predictions that we can derive by this periodic data and models.

In mathematical term, for a single point it is represented as follows:

\[ y_p(t) = y(t - 1) \]

For multiple points it is represented as:

Table 5.3  Top 6 pollutants affecting AQI

| Name of the particle | Correlation |
|----------------------|-------------|
| Pm 2.5               | 0.965206    |
| Pm 10                | 0.971774    |
| NO                   | 0.815388    |
| NO₂                  | 0.862036    |
| NOₓ                  | 0.891316    |
| SO₂                  | 0.783169    |
\[ y_p(t) = \sum_{i=0}^{i-1} y_{t-i} \]

As we all know that monitoring and measurement of error is important which represented as \((e_t)\) error at a particular time and \(\beta_t\) is the coefficient of the first lag. For moving average model, it is represented as

\[ y = c + \varepsilon_t + \theta_1 \cdot \varepsilon_{t-1} + \theta_2 \cdot \varepsilon_{t-2} + \cdots \]

For weighted average model:

\[ y_p(t) = \sum_{i=0}^{i-1} \beta_i \cdot y_{t-i} \]

The smoothing is required. If \(\alpha\), be smoothing factor then it is represented:

\[ \hat{y}_p(t) = \alpha \cdot y_t + (1 - \alpha) y_p(t-1) \]

There are four parts in which a series need to be decomposed for further analysis:

(a) Trend
(b) Seasonality
(c) Irregularity/noise
(d) cyclic.

Here, we mainly test the trend which is basically the tendency of movement of data and seasonality which is a certain pattern of repetition over time.

For implementing models the first step is to understand the stationary and the difference and steps required for the series to make it stationary. For that Dickey-Fuller’s Test is a unit root test which tests the null hypothesis in the equation where \(\alpha\) is the first lag coefficient:

\[ y(t) = c + \beta t + \alpha \cdot y_{t-1} + \varphi \Delta y_{t-1} + \varepsilon_t \]

In augmented Dickey-Fuller test the equation becomes:

\[ y(t) = c + \beta t + \alpha \cdot y_{t-1} + \theta_1 \Delta y_{t-1} + \theta_2 \Delta y_{t-2} + \cdots + \varepsilon_t \]

Here, the ADF test is performed. The rolling mean along with rolling standard deviation is plotted. Rolling mean is the running average and here window of 12 is taken, i.e. month-wise moving average is taken to make a finite impulse filter and to understand fluctuations and trends. To understand how the original series of data is deviating from average data, we have to take care of the standard deviation and since it is on a running basis so it is called rolling std deviation.
From the graph, it can be said that it has a seasonal factor which can be further understood by plotting the seasonal difference in Fig. 5.7.

We can understand about the seasonal component of the series. Then for stationary check ADF is performed. Then the result of ADF can be seen from Table 5.4.

Here, the obtained p-value is greater than 0.05 significance level and ADF/T stats is greater than all critical values so the null hypothesis is accepted that means the series is not stationary. For the visualisation purpose, if we plot individually the concentration of pollutants we can get even clear about the biasness of the pollutants toward the season. It is evident that apart from particulate matters, the concentration of sulphur and sulphur dioxide increases during the seasonal time. This is because the crackers contain sulphur in it and due to burning sulphur dioxide is formed and thus the concentration increases to a large extent during this season. For further analysis of trend, seasonality and other components, seasonal decomposition is done and the results are plotted in Fig. 5.8 (Table 5.5).

| Name of parameter | Value       |
|-------------------|-------------|
| ADF/T statistic   | −1.37128    |
| P-value           | 0.595969    |
| 1% critical value | −3.440147   |
| 5% critical value | −2.86586    |
| 10% critical value| −2.56905    |
Table 5.5 Result seasonal decomposition

| Date      | AQI Res | AQI seas | AQI trend | Date      | AQI Res | AQI seas | AQI trend |
|-----------|---------|----------|-----------|-----------|---------|----------|-----------|
| 15-10-2018| −12.9174| −3.65398 | 109.5714  | 15-10-2019| 7.628376| 1.514481 | 125.8571  |
| 16-10-2018| −24.9431| 1.514481 | 118.4286  | 16-10-2019| −4.16045| 2.303306 | 130.8571  |
| 17-10-2018| 23.69669| 2.303306 | 141       | 17-10-2019| −37.7216| 0.578711 | 131.1429  |
| 18-10-2018| −8.43585| 0.578711 | 163.8571  | 18-10-2019| 12.48609| 0.228195 | 136.2857  |

5.7.2 Autoregressive Model

Here, the observations of previous steps are taken as input to an equation for forecasting the next step values. Here, the aim is to forecast value for \((t + 1)\) when input is for \((t − 1)\) and \((t − 2)\).

ARIMA (Autoregressive Integrated Moving Average)

When we handle time series the ARMIA model or the autoregressive integrated moving average model becomes a key in forecasting and predicting solutions for problems. It is a class of statistical models. In terms of mathematics, if \(L\) be a lag operator,

\[
1 - \sum_{i=1}^{p-d} \varphi_i L^i = (1 - L)^d
\]
It is the process which expresses this polynomial factorisation property with \( p = p' - d \), and is given by:

\[
(1 - \sum_{i=1}^{p} \varphi_i L^i) \cdot (1 - L)^d \cdot X_t = (1 + \sum_{i=1}^{q} \theta_i L^i) \varepsilon_t
\]

So at a particular case in the ARMA \((p + d, q)\) model, having an autoregressive equation of polynomial consisting d unit-roots the equation becomes:

\[
(1 - \sum_{i=1}^{p} \varphi_i L^i) \cdot (1 - L)^d \cdot X_t = \delta + (1 + \sum_{i=1}^{q} \theta_i L^i) \varepsilon_t
\]

Here, AR means autoregression means the relationship between dependent variables and lags observation is taken into account. Integrating means to understand the difference and implementing it to make series stationary. It is represented as ARIMA \((p, d, q)\) where \( p \) is lag order, \( d \) is the degree of difference and \( q \) is the order of MA. The best two fitted model are with order 001 which can be visualised in Figs. 5.9 and 5.10.

So the best fit for above two comes in \( p = 0 \ q = 1 \) and \( d \) varies.

**SARIMAX (Seasonal Autoregressive Integrated Moving Averages with Exogenous Regressors)**

Here, the seasonal component is mainly focused. But for systematic way PCF and ACF are needed [36].

There is a large spike at lower lag, i.e. high autocorrelation will be a higher effect. Sarimax is applied and the results are seen (Figs. 5.11 and 5.12).
5.8 Artificial Neural Network and LSTM-Based Modeling

Artificial intelligence is the study of intelligent agents. It was coined in 1956 by John McCarthy. Actually, the concept was brought in this area mainly for software which could perform smart computation like humans. The systems capable of smart
and intelligent high-level computations are just an extended version of conventional computing. They are basically roots of the fifth generation. The ANN approach is inspired from the Neurons of the Nervous system of the human brain where there are numerous numbers of interconnected systems along with fast, rapid and parallel execution is going on. ANN is also a data-driven model based on mathematics and the large interconnected architecture. It basically learns from the data by training and giving better results over time and training letting it to think in a rational way to do things with accuracy. The main goal is to understand the complex relationship between the input and the output node. It mainly consisted of three layers: Input, Hidden and Output layers. Neural networks get the knowledge by identifying the patterns as well as insights in data. The MLP is composed of multiple layers of operation nodes that derive the data and are connected to the input and output layers in a directed graph combination. The hidden layers are responsible to perform the various calculations and make predicted outputs. The finer the layers become the more accurate the output is tended to be. The number of layers is not fixed and depends totally on the problem. There are other two terms, i.e. weights and biases. Each node is connected and weight, i.e. coefficients determine the impact of input features which constitute its structure.

RNN is basically a type of Neural Network. Recurrent Networks is able to process a successive recursion with the help of transition function to internal hidden vector state of the input.

It is an MLP where the earlier hidden unit activations have a loop rather feedback loop which goes into the neural network along with the input features. The input at instance \( t \) has some previous data which is at time \( t - 1 \). The cyclic connections are the ultimate power of RNN which makes it more powerful than the normal neural
network to handle sequence data and implement sequence modeling. RNNs had achieved great success in sequence labeling and prediction of time series data.

LSTM is the modified version of RNN which is basically done to overcome the drawbacks of the recurrent network. It has some special blocks known as memory block in the hidden layer which are having a good self-connection network which stores the information of the temporary state and the flow is operated by the input (save vector) and output gates (focus vector). There are mainly three gates input gate, output gate and forget gate.

The equation cell state is:

\[ p_t = f_t \odot p_{t-1} + i_t \odot p'_t \]

Remember vectors are the forget cells. If the output of this gate is 1 then the information is preserved else deleted. In this way it works.

LSTM is widely for predication of Univariate as well as multivariate time series models. In our paper, the observation of AQI over time period of two years from June 2018 to June 2020 is considered and it is modeled (Figs. 5.13 and 5.14).

Here, we used the different optimiser and tested loss which is illustrated in Table 5.6.

### 5.9 Effect of COVID-19 on Air Pollution

COVID-19 is creating a pandemic situation throughout the world breaking the different pillars like the economy of a country. The epidemic broke at a rapid pace creating an exponential growth. It is taking away the jobs and as a result unemployability rate is increasing creating a dangerous economic crisis. It causes fibrosis of
lungs, i.e. lungs tissues is been replaced non-working tissues which sometimes result in death. First, it is present in the upper respiratory tract and it causes the dryness of the throat resulting in cough and then through the respiratory tract it goes to the lungs. Enveloped in the non-separated positive RNA virus belongs to the mother family of Coronaviridae and Nidovirales. Initially, it was assumed that 2019-nCoV is only transmitted from animal–human contact, but later found that human–human transmission is also occurring. The death rate is increasing day-by-day which is a big reason to worry. The number of total affected cases is increasing insignificance [37].

The first case in India was detected on 30th January 2020 and after that affected cases were increasing rapidly. Also, the people affected in a day is increasing exponentially (Fig. 5.15).

As per reports published by different top institutes of India as well as foreign institutes it is evident that this is increasing day-by-day and may reach to peak point between November 2020–February 2021.

It is highly contagious and to avoid spread or rather community spread, the Govt. of India announced lockdown from March end to May end 2020. Due to this the industrial works, offices were stopped and due to this number of vehicles operating on the roads also decreased. So this lockdown became a blessing to the environment
and the pollution amount got reduced. AQI value decreased significantly which can be visualised using Fig. 5.16.

Due to lockdown, the amount vehicles, industries operating were reduced and thus pollutant amount also dropped significantly and the concentration amount is very less than the same time of last year which can be visualised. During early lockdown, AQI
The Dual Impact of Lockdown on Curbing COVID-19 Spread …

Fig. 5.17  Showing AQI value decreasing

Table 5.7  Showing average AQI

| Average AQI before lockdown | Average AQI after lockdown |
|-----------------------------|----------------------------|
| 174.094118                  | 75.648649                  |

value decreases to a high extent resulting in pollution-free environment which is plotted in Fig. 5.17 [38]

After Lockdown we can see that the average AQI is decreasing which can be seen in Fig. 5.17 and Table 5.7. The upper line is for before the lockdown and the lower line is for after lockdown.

We can also see that due to lockdown and due to high spreading of the corona, people are trying to avoid social gathering and it also affects the air quality which can be visualised using Figs. 5.18 and 5.19.

It is predicted that if less amount of festivals and less amount of crackers are there, then AQI value will be highly decreased which is predicted as in Table 5.8.

5.10 Conclusion and Future Scope

In this paper we have tried to find the liaison between lockdown and the AQI Index of the air. At first we have proposed a lockdown model using SIQR modeling, through which we have demonstrated how the COVID-19 spread has been capped using lockdown and consequently we have shown how the AQI Index varied during the lockdown period. We have also presented conditional scenarios where it has been shown that without lockdown how would be the figure of COVID-19 across the nation and subsequently the AQI. Now, currently the Unlock 1, 2 are on the cards to revive
Fig. 5.18  Yearly and monthly box plot for NO₂

Fig. 5.19  Yearly and monthly box plot for PM 2.5

Table 5.8  Prediction of AQI

| Month     | Monthly average AQI |
|-----------|---------------------|
| March     | 52.41               |
| April     | 52.21               |
| May       | 52.15               |
| June      | 51.14               |
| July      | 52.917              |
| August    | 52.854              |
| September | 53.760              |
| October   | 58.765              |
| November  | 59.658              |
| December  | 62.652              |
The down trotting economy the cases per day is increasing in a stipulated manner along with the AQI is also increasing as shown through our analyses. Mother Nature has definitely being healed through this period but with the advent of the people and vehicles on the street it is getting polluted once again but at a much lower rate. As per our predictions we estimate a much lower pollution rate over upcoming Diwali festive season for mainly two reasons viz. One driving factor will be lowering economy and it is expected that people wouldn’t hold much money to spend on the crackers rather than trying to meet their needs. And the other would be, definitely COVID-19, as with the current figures it is pretty evident that it will not end anywhere near November 2020 unless and until the vaccines arrive which is also pretty much difficult on the cards. The paper also shows two-dimensional analysis of the AQI fluctuations, i.e. one is through the time series statistical modeling and the other through the LSTM Based modeling which is heavily used for time series prediction, and both the models are giving similar predictions.

References

1. Roy S (2020) Spread of COVID-19 in India: a simple algebraic study. Available at SSRN 3597973
2. Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. Proc R Soc Lond Ser A, Containing Papers of a Mathematical and Physical Character 115(772):700–721
3. Jha V, Dinesh TA, Nair P (2020) Are we ready for controlling community transmission of COVID 19 in India? Epidemiol Int 5(1):10–13 (E-ISSN: 2455-7048)
4. Mittal S, An exploratory data analysis of COVID-19 in India. https://www.researchgate.net/publication/340965511_An_Exploratory_Data_Analysis_of_COVID-19_in_India
5. Arti MK, Bhatnagar K, Modeling and predictions for COVID 19 spread in India. ResearchGate. https://doi.org/10.13140/RG.2.2.11427.81444
6. El Deeba O, Jallouc M (2020) The dynamics of COVID-19 spread in Lebanon. arXiv, arXiv-2005.
7. Ghosh P, Ghosh R, Chakraborty B (2020) COVID-19 in India: state-wise analysis and prediction. medRxiv
8. Freeman B, Taylor G, Gharabaghi B, Thé J (2017) Forecasting air quality time series using deep learning. J Air Waste Manage Assoc 68. https://doi.org/10.1080/10962247.2018.1459956
9. Roy S, Roy Bhattacharya, K (2020) Spread of COVID-19 in India: a mathematical model. Available at SSRN 3587212
10. Sharma VK, Nigam U (2020) Modelling of Covid-19 cases in India using regression and time series models. medRxiv
11. Cohen AJ, Ross Anderson H, Ostro B et al (2005) The global burden of disease due to outdoor air pollution. J Toxicol Environ Health. Part A 68(13–14):1301–1307. https://doi.org/10.1080/15287390590936166
12. Gupta S, Raghuwanshi GS, Chanda A (2020) Effect of weather on COVID-19 spread in the US: a prediction model for India in 2020. Sci Total Environ 138860
13. Gopal R, Chandrasekar VK, Lakshmanan M (2020) Dynamical modelling and analysis of COVID-19 in India. arXiv:2005.08255
14. Singh BP, Singh G (2020) Modeling tempo of COVID-19 pandemic in India and significance of lockdown. medRxiv
15. Gupta S (2020) Epidemic parameters for COVID-19 in several regions of India. arXiv:2005.08499
16. Ghorani-Azam A, Riahi-Zanjani B, Balali-Mood M (2016) Effects of air pollution on human health and practical measures for prevention in Iran. J Res Med Sci 21:65. Published 2016 Sep 1. https://doi.org/10.4103/1735-1995.189646
17. Sherstinsky A (2020) Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. Physica D 404:132306. https://doi.org/10.1016/j.physd.2019.132306
18. Sharma VK, Nigam U (2020) Modeling and forecasting for Covid-19 growth curve in India. medRxiv
19. Shekhar H (2020) Prediction of spreads of COVID-19 in India from current trend. medRxiv
20. Agatonovic-Kustrin S, Beresford R (2000) Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. J Pharm Biomed Anal 22(5):717–727. https://doi.org/10.1016/s0731-7085(99)00272-1
21. Bergstra A, Brunekeef B, Burdorf A (2018) The effect of industry-related air pollution on lung function and respiratory symptoms in school children. Environ Health A Global Access Sci Source 17:30. https://doi.org/10.1186/s12940-018-0373-2
22. Sinnott RO, Guan Z (2018) Prediction of air pollution through machine learning approaches on the cloud. In: 2018 IEEE/ACM 5th international conference on big data computing applications and technologies (BDCAT). Zurich, pp 51–60. https://doi.org/10.1109/BDCAT.2018.00015.
23. Bhalgat P, Bhoite S, Pitare S (2019) Air quality prediction using machine learning algorithms. Int J Comput Appl Technol Res 8. https://doi.org/10.7753/ICATR0809.1006
24. Carbajal-Hernández JJ (2012) Assessment and prediction of air quality using fuzzy logic and autoregressive models. Atmos Environ 60:37–50
25. Nallakaruppan MK, SurejIlango H (2017) Location aware climate sensing and real time data analysis. In: 2017 world congress on computing and communication technologies (WCCCT). IEEE
26. Li Y, Chen Q, Zhao H, Wang L, Tao R (2015) Variations in pm10, pm2.5 and pm1.0 in an urban area of the Sichuan basin and their relation to meteorological factors. Atmosphere 6(1):150–163
27. Mahajan S, Chen L-J, Tsai T-C (2017) An empirical study of PM2.5 forecasting using neural network. In: IEEE smart world congress, At San Francisco, USA
28. Franklin BA, Brook R, Pope CA (2015) Air pollution and cardiovascular disease. Curr Prob Cardiol 40(5):207–238. ISSN 0146-2806
29. Bansal M, Aggarwal A, Verma T, Sood A (2019) Air quality index prediction of Delhi using LSTM. https://doi.org/10.13140/RG.2.2.26885.70884
30. Chang Y-S, Chiao H-T, Abimannan S, Huang Y-P, Tsai Y-T, Lin K-M (2020) An LSTM-based aggregated model for air pollution forecasting. Atmos Pollut Res 11(8):1451–1463. ISSN 1309-1042. https://doi.org/10.1016/j.apr.2020.05.015
31. Beluvadi S, Rajagopal S, Ranjani R, Mohan R (2020) Air quality forecasting using LSTM RNN and wireless sensor networks. Procedia Comput Sci 170:241–248. https://doi.org/10.1016/j.procs.2020.03.036
32. Gul S, Khan GM (2020) Forecasting hazard level of air pollutants using LSTM’s. In: Maglianiis I, Iliadis L, Pimenidis E (eds) Artificial intelligence applications and innovations. AIAI 2020. IFIP advances in information and communication technology, vol 584. Springer, Cham. https://doi.org/10.1007/978-3-030-49186-4_13
33. Jiao Y, Wang Z, Zhang Y (2019) Prediction of air quality index based on LSTM. In: 2019 IEEE 8th joint international information technology and artificial intelligence conference (ITAIC). Chongqing, China, pp 17–20. https://doi.org/10.1109/ITAIC.2019.8785602
34. Chaudhary V, Deshbhtrarat A, Kumar V, Paul D, Samsung (2018) Time series based LSTM model to predict air pollutant’s concentration for prominent cities in India
35. Kumar A, Goyal P (2013) Forecasting of air quality index in Delhi using neural network based on principal component analysis. Pure Appl Geophys 170:711–722. https://doi.org/10.1007/s00024-012-0583-4
36. Chatterjee A, Mukherjee S (2020) The impact of lockdown on GDP growth & COVID-19 spread: insights from a mathematical simulation exercise for India
37. Chatterjee A (2020) A bi-fold approach to assess the impact of lockdown on suppressing covid-19 spread & small cap stock market mayhem in India. J inform sys & oper manage 14(2)
38. Chatterjee A, Roy S (2021) An analytics overview & LSTM-based predictive modeling of covid-19: a hardheaded look across India. In Mach Intell Soft Comput, Springer, Singapore, pp. 289-307