# Spatio-temporal Analysis of Air Quality Index and Risk Simulator for Health Infrastructure Planning During COVID-19 Pandemic

Sanjana Pai Nagarmat, Saiyed Kashif Shaukat
*Research & Development Centre*
*Hitachi India Pvt. Ltd.*
Bangalore, India
email: {sanjana,saiyed.shaukat} @hitachi.co.in

***Abstract*—The COVID-19 pandemic has thrown light on the need to prioritize and improve the existing health systems globally. Especially in South Asia, the pandemic has left a long-term scar on the economy and people's lives. It has burdened the cities to improve their overall health and medical infrastructure facilities. Though several makeshift solutions developed in due time provided intermediate relief to the cities, a long-term solution to make the city self-sufficient in terms of resources, especially in situations like the pandemic becomes necessary. Thus, a collaborative approach to bring about an overall change in the existing health system is required. In this work, we provide an in-depth Spatio-temporal analysis of the air quality in an Indian smart city - Pune and propose a simulator tool that assists authorities in making informed decisions on adding medical infrastructure facilities. This simulator tool takes in static demographical inputs, user inputs related to the city infrastructure, and forecasted dynamic inputs like critical zones or hotspots along with the overall air quality levels to predict the risk scores of wards. This numerical predicted score is an indicator of the ward level risk. Concerned authorities can use this tool to foresee the ward conditions and understand the impact of the risk score in a ward with incremental variations in the ward facilities during the pandemic. Further, the Spatio-temporal analysis of the Air Quality Index (AQI) shows how certain strategies and guidelines imposed during the pandemic can help bring in significant improvement in the environmental health condition of the city. It provides a detailed summary of the variation in the air quality index and pollutant concentrations in Pune from March 2020 to July 2021. Overall, this work provides a retrospective view on the impact of demographic, medical infrastructure, and environmental parameters on the ward condition and how improvement in these parameters can help cities ameliorate the existing health system in challenging times like the pandemic.**

***Keywords*—*COVID-19; air quality index; risk score.***

## I. Introduction

The research presented in this paper is an extension of our previous work [1]. Coronavirus disease (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus. It eventually started spreading across the globe and soon became a matter of international concern. The World Health Organization (WHO) declared this disease a pandemic in March 2020. Since then, it has caused more than 526 million cases across the globe claiming millions of lives [2].

Like several other countries around the world, India also faced challenges related to the economy, infrastructure, medical facilities, human resources, and environment due to the pandemic. The second wave of COVID-19 in India has had severe consequences in the terms of escalating positive cases, shortage of medical infrastructure facilities reduced medical supplies for treatments, and increased death rates, particularly in the younger population [3]. Scientists and researchers across the globe have been trying to identify the causes for the same. However, the reason for this is still not apparent and is beyond current scientific explanations. Purva et al. [3] believe that India's poor air quality index could be a potential factor as to why the spread of the infection has been severe across the country. Given that 9 of the 15 most polluted cities globally are in India, it could be postulated that the ability of the Indian population to fight against COVID-19 is impaired because people's lungs are severely affected by air pollution. Further, a study by Comunian et al. [4] suggest that an increase in fine particulate matter $PM_{2.5}$ is associated with an increased risk of COVID-19 infection. The atmospheric particulate matter could create a suitable environment for transporting the virus to greater distances. This could induce inflammation in lung cells and exposure to the particulate matter could further increase the susceptibility and severity of the COVID-19 patient symptoms.

In our previous work [1], we studied the effects of COVID-19, particularly in the smart city of Pune, a western Indian state of Maharashtra, and devised a scoring mechanism to dynamically predict the risk levels and the environmental health in the city. This work is an extension of the same. In this paper, we further do an in-depth analysis of the air quality conditions in Pune city during the period of the COVID-19 pandemic and implement a simulator that helps authorities foresee the reduction of risk scores in the city with the addition of medical infrastructure facilities during the pandemic. The timeline of the COVID-19 wave in India during our study can be broadly divided into three phases: Phase 1

(March 2020 to Oct 2020), Maintenance Phase (November 2020 to February 2021), and Phase 2 (March 2021 to June 2021). In our paper, we select the following months in each of the three phases for detailed analysis. Phase 1 (March, April, May 2020), Maintenance Phase (January, February, March 2021), and Phase 2 (May, June, July 2021). According to reports [5], during the second wave, Maharashtra was India's worst-affected state, just as it was in the first wave of COVID-19. A total of 35 percent of cases were reported in the first wave, while 6 percent of cases were registered during the Maintenance Phase. The second wave reported 59 percent cases in Pune. Pune being a smart city has taken several initiatives to fight against this pandemic. Right from setting up a command-and-control center for the supervision and community surveillance of containment zones, tracking the number of positive patients, and planning health care resources, Pune smart city team has been actively providing support for the COVID-19 management. However, due to the unprecedented nature and non-anticipated spread of the virus, the 2nd phase of the pandemic had a devastating effect on the city. In this work, we analyze the ward level risk scores with changes in the ward medical and infrastructure facilities and propose tools, and suggestions to help reduce the risk levels in such situations. The contributions of this work are as follows:

- **Spatio-temporal analysis of Air Quality Index (AQI) during the pandemic**: Descriptive statistics are used to assess the Spatio-temporal variations of various environmental attributes including AQI, sound (dB), ozone (ppb), $NO_2$ (ppb), $SO_2$ (ppb) and particulate matter $PM_{2.5}$ ($\mu g/m^3$) during the 3 phases of the COVID-19 pandemic. Spatial analysis is used to study the variations in attributes across the wards. While temporal analysis is used to study the variations in attributes of the ward over time. Quantification of the pollutant levels across the study period is provided to understand how the lockdown impacted the AQI levels in the city. Suggestive measures to control the degradation of ward environmental health are provided.
- **A simulator tool to analyze the reduction in the ward level COVID-19 risk score with incremental change in the ward medical and infrastructure facilities**: As introduced in our previous work [1], we had developed an aggregated data-rich model that predicts the ward level risk scores. This numerical score was indicative of the overall risk associated with the ward in terms of its medical infrastructural health, demographics, and environmental health. Based on the risk scores, the wards were categorized into risk zones - Severe (100-80), High (80-60), Moderate (60-40), Low (40-20), and Very Low (20-0). The higher the risk value, the greater the COVID-19 risk. In this work, we provide an extension to this model to help simulate and analyze the reduction in the ward level COVID-19 risk score with incremental change in the ward medical and infrastructure facilities. The inputs to the tool include static inputs like the ward demographics (ward population, literate population, children under the age of 6, number of houses in the ward, family size, and working population), dynamic inputs like the number of hotspots, health score of wards along with user inputs like additional hospitals and beds to be added. The dynamic inputs are forecasted based on previous data while the user inputs are directly fed to the model. Training data is prepared to reflect the ward attributes and is used to train the model. Based on the user inputs, required features are forecasted and then fed to the model to get the corresponding risk scores. This simulator tool provides detailed information on the impact of critical attributes including medical infrastructure facilities (hospitals, beds in hospitals), ward demographics (population statistics), and ward environmental health (AQI, population, and tree count) on the ward risk levels. With the help of this tool, concerned authorities can foresee the risk scores for various wards in Pune Municipal Corporation (PMC)-Pune, categorize the wards into risk zones based on the risk scores, and try to understand the impact of incremental improvement in the ward medical infrastructure facilities on the ward risk via simulation.

Overall, the work in this paper provides granular ward level AQI statistics and localized information dynamically considering the change in the attributes. The simulator tool provides adequate information to concerned authorities to plan and further improve the infrastructure and medical facilities in the ward. The simulation helps them understand how improving the ward facilities will impact the risk score and in turn their categorization to risk zone.

The remaining part of the paper is organized as follows. Section II describes the overall methodology, and Section III provides technical information about the data sources, data analysis, and modeling with suitable graphs and plots. Section IV presents the discussion of the results while Section V describes the conclusion of our approach.

## II. Methodology

This section provides details on the study area and briefly explains the contributions of this research.

### A. Study Area

Pune is the seventh most populous city in India and the second-largest city in the state of Maharashtra, with an estimated population of 7.4 million as of 2020 [6]. Pune is also the 101st largest city in the world by population. It is also one of the fastest-growing cities in the Asia-Pacific region. The Mercer 2019 Quality of Living rankings evaluated local living conditions in more than 440 cities around the world where Pune ranked at 143 [7]. PMC is the civic body that governs the inner limits of Pune spread over an area of 331.26 sq. km. Based on the 2011 census, the data from 144 wards of the PMC region is considered in this study [8]. The government of India has started the National Smart Cities Mission program to develop smart cities across the country. The Union Ministry of Urban Development in India is responsible for implementing this mission in collaboration with the state governments of the respective cities. Projects were launched in 20 cities selected in the first batch of the mission through a city challenge competition. Pune was one among them that was shortlisted by the Minister of Urban Development.

### *B. COVID-19 Phase-wise AQI Analysis in Pune*

Air Quality data [9] of Pune city is collected, pre-processed, and analyzed over the study period, Phase 1 (March 2020 to May 2020), Maintenance Phase (January 2021 to March 2021), and Phase 2 (May to July 2021). Ward-wise variations across months and days are studied in detail. We summarize our findings phase-wise and interpret the reasons for the variations in the parameter. This sub-section provides an overview of how the overall AQI has varied in Pune as a result of various precautionary measures that were taken during the study period.

### *C. Simulator tool to analyze the reduction in the ward level COVID-19 risk score with incremental change in the ward medical and infrastructure facilities.*

This tool helps users analyze the variations in the risk score of the ward that is calculated based on essential parameters like the ward demographics, environmental health as well as the current medical and infrastructure facilities. The simulator tool is built on top of the Gradient Tree Boosting-based risk prediction model proposed in our previous work [1]. Based on the current risk score of the ward, users can vary the input features and understand their impact on the risk scores. Improvement in the ward medical facilities will reduce the risk level in the ward and its impact can be easily understood with this simulator. This makes it easier to help city authorities plan and improve the ward medical facilities in challenging times like the pandemic. Appropriate user interface screens are implemented to access the features of this simulator tool.

## III. Implemention and Results

This section provides details on the data sources involved, analysis, and data modeling along with the results and supported user screens.

### *A. COVID-19 Phase-wise AQI Analysis in Pune*

*1) Data Sources:* Air Quality data [9]: The Pune Urban Data Exchange (PUDX) platform provides data related to the air quality sensors installed in several parts of Pune city. 50+ sensors installed across the city actively measure the concentrations of pollutants. The sensors provide AQI readings every 15 minutes. A pipeline is created, and a job is scheduled to collect this data regularly. The collected sensor data is then mapped to the appropriate wards.

*2) Data Analysis: Spatio-temporal analysis of AQI in PMC-Pune:* AQI for the wards in Pune across a period of months is studied in detail. We studied the AQI over the study period, Phase 1 (March 2020 to May 2020), Maintenance Phase (January 2021 to March 2021), and Phase 2 (May to July 2021). According to the study [10], Pune had been ranked 299th on the air pollution index across the globe, 8th in the state, and 74th across the country. However, in 2020, during Phase 1, we saw a significant decrease in the AQI levels across wards in PMC-Pune, especially during April. The average AQI levels saw a decrease of 36% from March to May. This reduction can be majorly attributed to the restriction in activities that were imposed with a nationwide lockdown persisting across these months. Figure 1 shows the average AQI levels in PMC-Pune during this phase. We also studied the variations of AQI at individual ward levels both month-wise and day-wise. Table I below shows the month-wise AQI values in 10 wards during Phase 1 and variation in the AQI between May-20 and March-20.

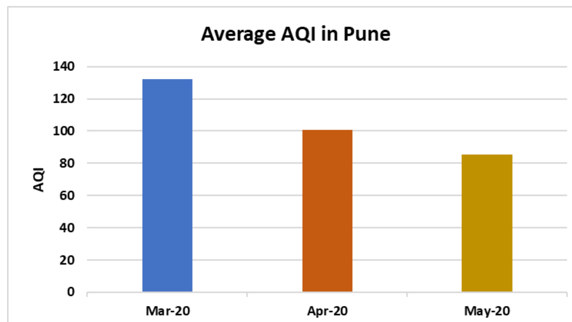


Figure 1. Average AQI in PMC Pune during Phase 1

TABLE I
Ward-wise variation in AQI during Phase 1

| Ward | Mar-20 | Apr-20 | May-20 | Variation(%) |
|---|---|---|---|---|
| AundhGaon | 107.74 | 111.62 | 119.79 | 11 |
| Baner-Balewadi | 49.38 | 33.23 | 30.94 | -37 |
| Hadapsar | 195.61 | 141.25 | 149.66 | -23 |
| Lohagaon | 87.03 | 55.63 | 13.04 | -85 |
| Parvati Darshan | 138.01 | 104.89 | 94.47 | -32 |
| PhuleNagar | 156.04 | 90.73 | 96.99 | -38 |
| S.G.Rugnalaya | 77.43 | 65.32 | 28.49 | -63 |
| S.Mahavidhyalaya | 65.15 | 53.25 | 55.08 | -15 |
| Ved Bhavan | 117.54 | 64.75 | 42.84 | -64 |
| Wadiya College | 124.94 | 98.58 | 100.26 | -20 |

The negative variation indicates improvement in the AQI level while a positive variation shows how it has degraded over time. Most of the wards showed improvements in their air quality index. Some wards showed as little as 2% improvement in their AQI levels while a few others showed 100% improvement. It can be noticed that the AQI levels were relatively lower in April-2020 when strict lockdown rules were imposed across the city. However, by the end of May, the unlock phases had begun. Increased activity during this period in some wards might have led to positive variation being seen in their

AQI levels. This significant decrease in various attributes can be contributed to the suspension of activities and restrictions imposed on citizens during the lockdown. Residential areas like Lohagaon, Kothrud, Baner-Balewadi, and many others as shown in Table I showed a significant reduction in attributes like AQI, sound, and pollutant concentrations with the imposition of lockdown. Figure 2 shows variations in Sound, $NO_2$, $SO_2$, $PM_{2.5}$ and Ozone levels during this phase across Pune. Month-wise and day-wise AQI variations of the ward, Baner-Balewadi during Phase 1 are shown in Figure 3.

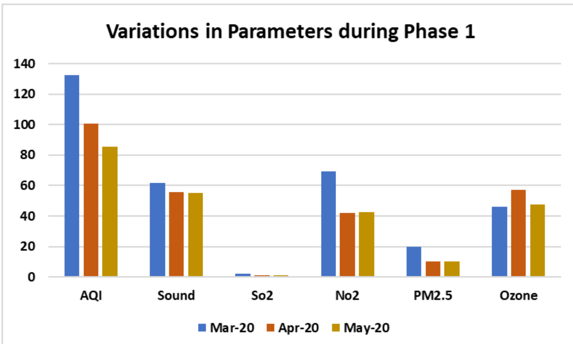


Figure 2. AQI Parameter variation during Phase 1

Overall, during Phase 1 PMC-Pune had

- 36.0% decrease in AQI level
- 11.0% decrease in Sound level
- 58.0% decrease in $SO_2$ level
- 39.0% decrease in $NO_2$ level
- 48.0% decrease in $PM_{2.5}$ level
- 2.0% increase in Ozone level

With the lifting of restrictions, activities across the country slowly began to resume after the first wave of COVID-19. During the Maintenance Phase from November 2020 to January 2020, there were a relatively lesser number of cases seen across the country. By the end of February, the cases saw a surge in number, and people started cautiously following the COVID-19 protocols. Overall, during this phase, the AQI levels decrease by 16% as shown in Figure 4. Figure 5 shows the variation in other AQI parameters.

This decrease in AQI levels was also reflected in some wards during this period. However, the effect of the relaxation of COVID-19 rules and protocols was seen in contributing to increasing the AQI levels in a few more. Table II shows the ward-wise variations across months in this period. Month-wise and day-wise AQI variations of the ward, Baner-Balewadi during the Maintenance Phase are shown in Figure 6.

Overall, during the Maintenance Phase, PMC-Pune had

- 17.0% decrease in AQI level
- 1.0% increase in Sound level
- 36.0% decrease in $SO_2$ level
- 20.0% decrease in $NO_2$ level
- 29.0% decrease in $PM_{2.5}$ level
- 114.0% increase in Ozone level

TABLE II
WARD-WISE VARIATION IN AQI DURING MAINTENANCE PHASE

| Ward | Jan-21 | Feb-21 | Mar-21 | Variation(%) |
|---|---|---|---|---|
| AundhGaon | 93.89 | 88.69 | 81.92 | -13 |
| Baner-Balewadi | 57.43 | 56.52 | 50.24 | -13 |
| Hadapsar | 103.11 | 84.73 | 116.86 | 13 |
| Lohagaon | 98.38 | 84.55 | 83.55 | -15 |
| Parvati Darshan | 139.55 | 101.83 | 104.97 | -25 |
| PhuleNagar | 226.34 | 155.44 | 104.02 | -54 |
| S.G.Rugnalaya | 101.33 | 104.06 | 103.85 | 2 |
| S.Mahavidhyalaya | 108.69 | 119.04 | 135.93 | 25 |
| Ved Bhavan | 79.56 | 68.19 | 62.57 | -21 |
| Wadiya College | 161.18 | 101.89 | 114.05 | -29 |

According to a study, [11] a high $NO_x$ level reacts with Ozone and mops it up. The Ozone that escapes to cleaner areas has no $NO_x$ to further cannibalize it and as a result, Ozone concentration builds up in these areas. This explains the increase in concentrations of Ozone in the atmosphere during this phase.

According to the report, [6], nearly 59% of the cases in Pune were reported during Phase 2. It was noticed that the numbers were increasing rapidly. With changing symptoms unlike those seen in Phase1, people in the age group of 20 to 45 years were mainly affected during this phase. They showed symptoms like low-grade fever and severe breathlessness. During this phase, Maharashtra was one of India's worst-affected states in the country. The Chief Minister of Maharashtra then announced a lockdown starting from the second week of April that lasted till May. We analyzed the AQI levels during this phase as shown in Figure 9 and other parameters as shown in Figure 8.

A few wards saw drastic changes in the AQI levels during this phase. Table III shows the ward-wise variations across months in this period. Month-wise and day-wise AQI variations of the ward, Baner-Balewadi during Phase 2 are shown in Figure 9. Overall, during phase 2, PMC-Pune had

- 6.0% decrease in AQI level
- 1.0% increase in the Sound level
- 21.0% increase in $SO_2$ level
- 9.0% decrease in $NO_2$ level
- 21.0% increase in $PM_{2.5}$ level
- 30.0% decrease in Ozone level

Though levels of AQI, Ozone, and $NO_2$ decreased during this period, there was a significant increase in the $SO_2$ level. With the increase in COVID-19 cases, cities began manufacturing essential medical aids including the PPE kits in large quantities demanding higher energy consumption and in turn leading to higher emissions of $SO_2$ from the power plants. This may have led to increased $SO_2$ levels in the city.

Utilizing a mix of epidemiological data, satellite data, and other monitoring information from around the world, Pozzer et al. [12] estimated that on average, 15% of worldwide deaths from COVID-19 may be linked to chronic exposure to air pollution. They also found out that an increase in exposure to hazardous air pollutants is associated with a 9% increase

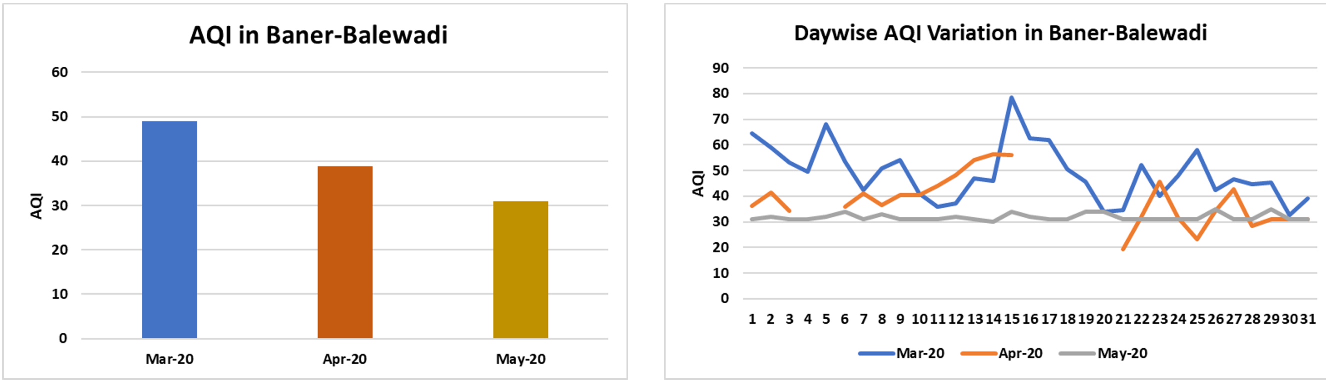


Figure 3. AQI variations in Baner-Balewadi during Phase 1

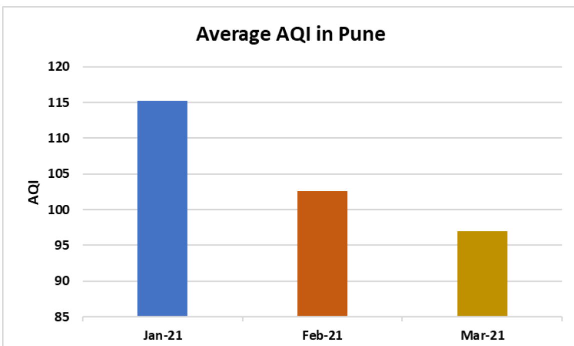


Figure 4. Average AQI in PMC Pune during Maintenance Phase

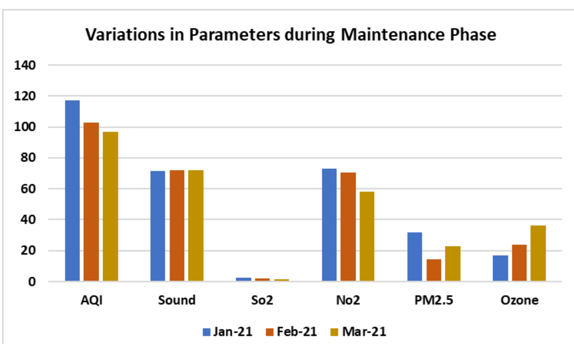


Figure 5. AQI Parameter variation during the Maintenance Phase

TABLE III
WARD-WISE VARIATION IN AQI DURING PHASE 2

| Ward | May-21 | Jun-21 | Jul-21 | Variation(%) |
|---|---|---|---|---|
| AundhGaon | 74.75 | 74.91 | 73.22 | -2 |
| Baner-Balewadi | 22.87 | 39.82 | 36.57 | 60 |
| Lohagaon | 59.89 | 56.18 | 67.29 | 12 |
| Hadapsar | 56.52 | 53.05 | 67.81 | 20 |
| Parvati Darshan | 65.31 | 59.34 | 64.37 | -1 |
| PhuleNagar | 73.29 | 64.31 | 58.39 | -20 |
| S.G.Rugnalaya | 91.34 | 61.79 | 59.83 | -34 |
| S.Mahavidhyalaya | 81.30 | 85.30 | 73.97 | -9 |
| Ved Bhavan | 65.31 | 74.32 | 70.42 | 8 |
| Wadiya College | 69.75 | 65.04 | 62.96 | -34 |

TABLE IV
AQI, COVID-19 CASES AND FATALITY RATE

| | AQI | COVID-19 cases (%) | Fatality(%) |
|---|---|---|---|
| Phase 1 | 102.14 | 35 | 52 |
| Maintenance Phase | 104.98 | 6 | 8 |
| Phase 2 | 70.68 | 59 | 40 |

in death among patients with COVID-19. With this as the reference, we summarized the AQI levels along with the total COVID-19 cases reported (in %) and the fatality rate against the reported cases in Pune in Table IV.

During the Maintenance Phase, there was an increase in the average AQI level and the fatalities that occurred in the reported cases. It may be ascertained that higher concentrations of air pollutants in certain wards may have caused increased respiratory stress, thereby increasing vulnerability to severe illness during Phase 2 of COVID-19. As reported by another study by Wu et al. [13], just a small increase (1 microgram per cubic meter) in long-term average exposure to fine particle pollution is associated with an 11% increase in the COVID-19 death rate for that country. We summarized the average variations of pollutant concentrations in Pune across the study period in Table V. Most of the parameters saw an increase in their concentration levels during the Maintenance phase or the relaxation phase. This increase in levels may have contributed to the surge in cases during Phase 2. Figure 11 shows the variations of pollutants across the study period. From this, we can essentially conclude that the higher the air pollution index and the pollutant concentrations, the more it correlated to poor health outcomes due to COVID-19 in Pune. Figure 10 shows various diseases attributed by air pollution in India in 2019 [14]. It can thus be inferred that with certain restrictions and measures, the concentration of pollutants can be significantly reduced in the city. This will not only help in improving the overall air quality but will also help prevent premature deaths, morbidity issues, and several other serious respiratory diseases in people. It will largely contribute to preventing the aggravation of respiratory stress, especially during challenging

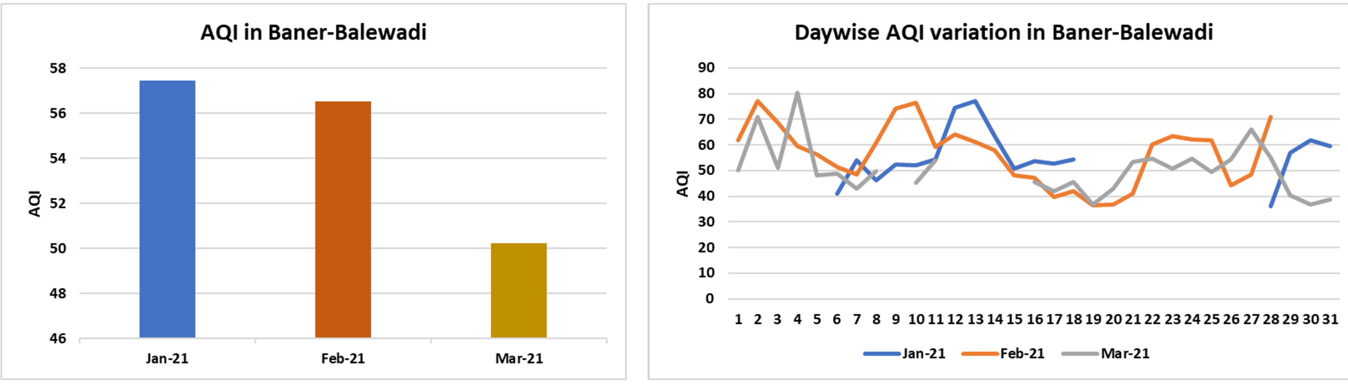


Figure 6. AQI variations in Baner-Balewadi during the Maintenance Phase

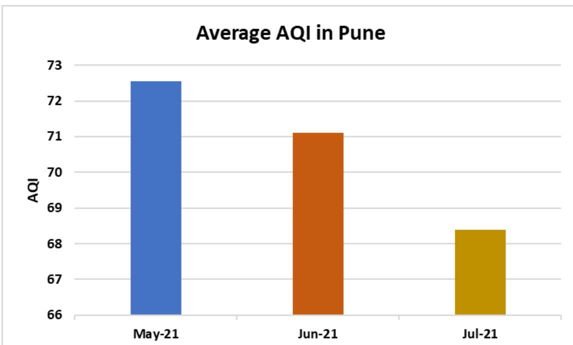


Figure 7. Average AQI in PMC Pune during Phase 2

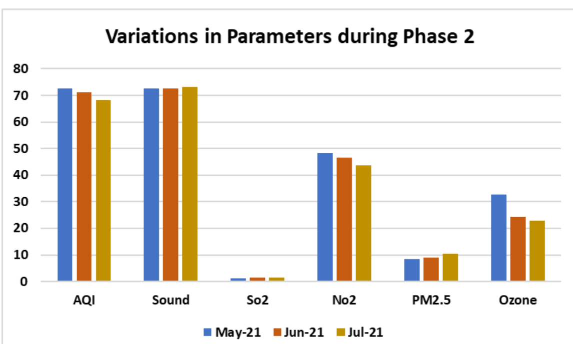


Figure 8. AQI Parameter variation during Phase 2

times like the pandemic.

*B. Simulator tool to analyze the reduction in the ward level COVID-19 risk score with incremental change in the ward medical and infrastructure facilities.*

This tool helps users analyze the variations in the risk score of the ward that is calculated based on essential parameters like the ward demographics, environmental health as well as the current medical and infrastructure facilities. As described in our previous work [1], we had developed a comprehensive risk scoring model that considered thirteen features as described in Table VI to model the risk level of the ward.

The data collected for these features was used to train and test the model. Several machine learning-based prediction algorithms like Linear Regressor, Random Forest, K-Nearest Neighbor, and Gradient Tree Boosting were used to predict the risk scores based on the thirteen input features selected. The predicted risk score values and their variations for the wards across days were studied in detail. The average prediction error rate of models using various machine learning algorithms was compared and the Gradient Tree Boosting algorithm was selected to predict the ward level risk scores. The work in this paper provides an extension to this model. Based on the current risk score of the ward, users can vary the input features like the number of beds(x) and hospitals(y) and understand their impact on the risk scores. Improvement in the ward medical facilities will reduce the risk level in the ward and its impact can be easily understood with this simulator. This makes it easier to help them plan and improve the ward medical facilities in challenging times like the pandemic. The thirteen features that are used to predict the risk score of the ward can be divided into three categories - static, user inputs, and forecasted inputs.

*1) Data inputs:*

- *Static inputs*: Publicly available ward information like - ward area, number of houses, number of literates, number of children below age 6, working population, and average family size in the wards is collected from the data sources provided by the government [15].
- *User inputs*: Additional number of beds and hospitals that are to be added in a ward are taken as input from the user.
- *Forecasted inputs*: Informed estimates on the parameters like the hotspots and AQI of the ward that varies with time are forecasted using Prophet model [16].

*2) Modelling the Simulator:* Appropriate data collection, data preparation, analysis and pre-processing, model training, and testing using thirteen features are done to build the risk model. Data for training the model is collected daily from the previously mentioned sources. First, a pipeline is created, and a job is scheduled for periodic data collection. Data is

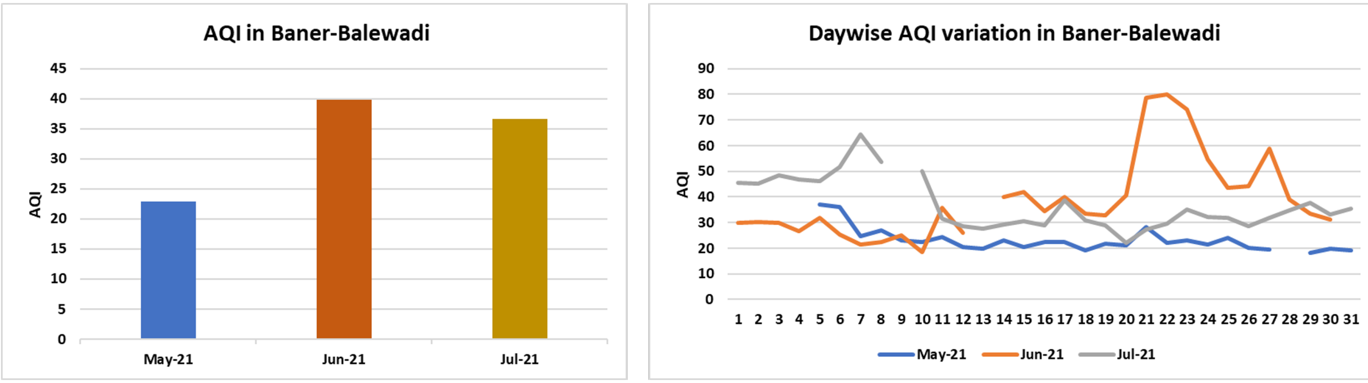


Figure 9. AQI variations in Baner-Balewadi during Phase 2

TABLE V
AQI and its Paramter Variations across the Study Period

| | Mar20 | Apr-20 | May-20 | Jan-21 | Feb-21 | Mar-21 | May-21 | Jun-21 | Jul-21 |
|---|---|---|---|---|---|---|---|---|---|
| AQI | 132.42 | 100.67 | 85.27 | 115.27 | 102.64 | 97.04 | 72.54 | 71.11 | 68.39 |
| Sound | 61.99 | 55.81 | 55.29 | 70.45 | 72.30 | 72.25 | 72.61 | 72.46 | 73.29 |
| $NO_2$ | 69.57 | 41.99 | 42.50 | 73.47 | 69.93 | 58.41 | 48.33 | 46.63 | 43.81 |
| $SO_2$ | 2.08 | 1.13 | 0.88 | 2.48 | 1.73 | 1.65 | 1.21 | 1.39 | 1.47 |
| $PM_{2.5}$ | 19.51 | 10.27 | 10.14 | 32.38 | 15.20 | 22.37 | 8.54 | 9.06 | 10.36 |
| Ozone | 46.30 | 57.25 | 47.43 | 16.16 | 22.74 | 35.26 | 32.62 | 24.23 | 22.86 |

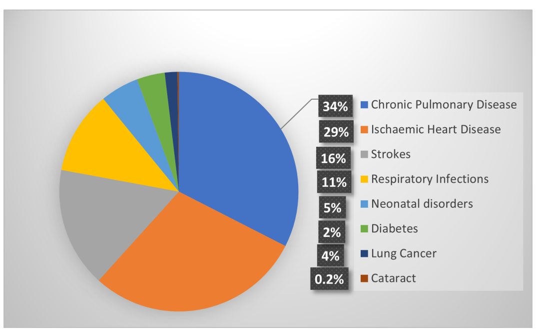


Figure 10. Causes of Deaths attributable to Air Pollution in India in 2019.

TABLE VI
Features used to calculate the COVID-19 risk score

| Feature# | Feature Name |
|---|---|
| 1 | Health Score |
| 2 | Houses |
| 3 | Literate Population |
| 4 | Population under 6 |
| 5 | Family Size |
| 6 | Working Population |
| 7 | Hospitals |
| 8 | Oxygen beds |
| 9 | Beds without oxygen |
| 10 | ICU ventilator beds |
| 11 | ICU beds without ventilator |
| 12 | Additional beds |
| 13 | Hotspots |

collected from all the above-mentioned data sources on a day-to-day basis. This is followed by data pre-processing, data aggregation, and model building. In the pre-processing step, the geo-coordinates and addresses available in the collected data are mapped to appropriate ward IDs using geocoding and mapping techniques. Further, in the data aggregation step, based on ward IDs the data is aggregated at the ward level. Data is cleaned and the missing values are imputed with appropriate techniques. This data is used to train the risk score prediction and simulator model. With the help of the risk score, users (city authorities, planners so on) can categorize the wards into various levels of risk and further select the ward to see how the risk score varies with changes in the current ward medical infrastructure facilities. The risk score initially calculated by the model is used as the reference (R). Using the simulator tool, they can understand and analyze the impact of incremental variations in the ward facilities like the number of beds (x) and hospitals(y) over the risk score during the analysis period. The tool can also help users make decisions on further improving the current ward medical infrastructure facilities. Figure 12 shows the user interface for the simulator tool.

As a first step, a user selects the ward and based on the existing data, the model predicts the risk score (R) of the ward. The initial details are shown in the table on the screen. With this as the reference, the user can fine-tune the input parameters like the number of beds(x1) and hospitals(y1) to be further added under the parameter fine-tuning section to get the new risk score (R1). In the backend, for the selected ward, the simulator tool collects the static data (Features 2,3,4,5,6) from the database. Further, using the previously available data for

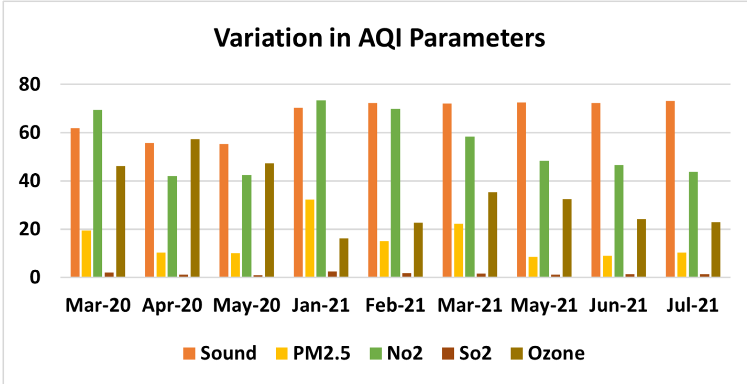


Figure 11. Variations in AQI Paramters

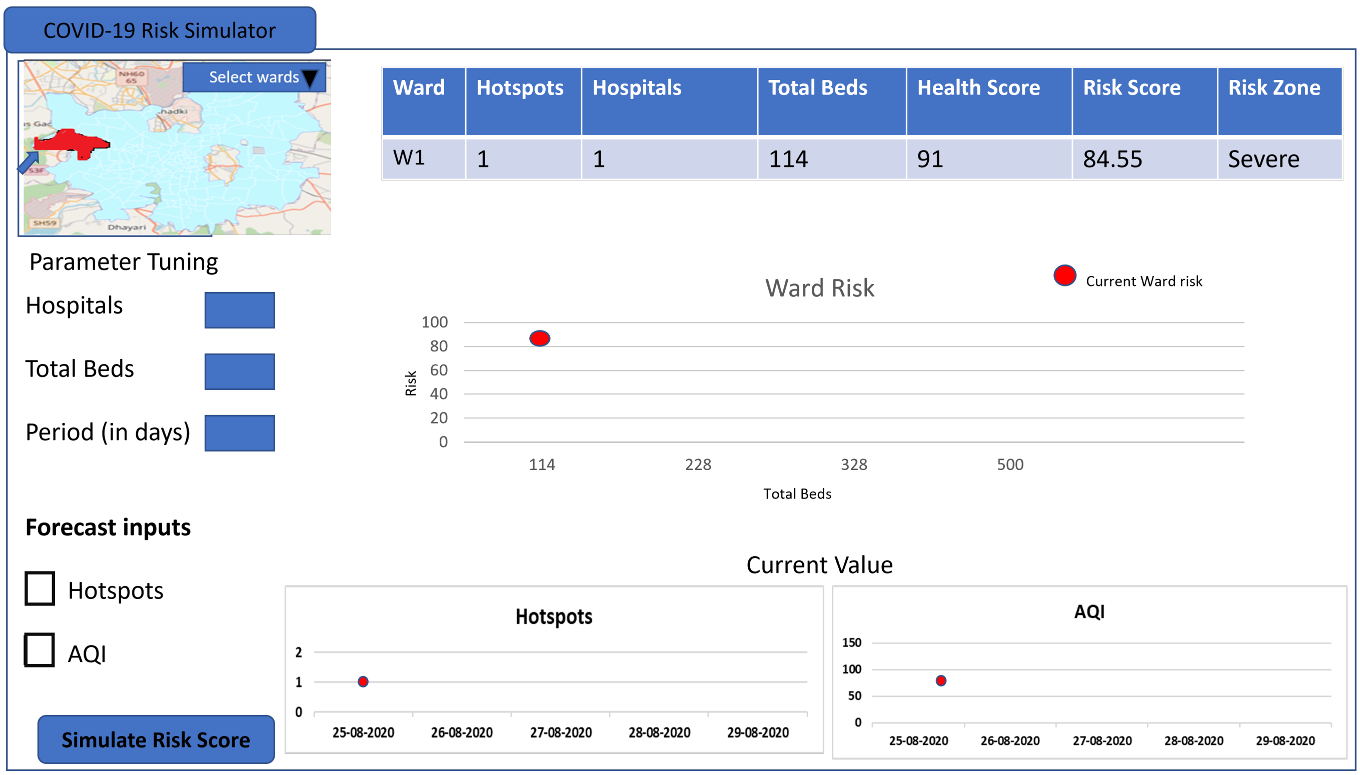


Figure 12. Landing page of the simulator tool

the ward, the tool forecasts the value for the features (Features 1 and 13) for the entered date using the Prophet model. The accuracy of the forecasting model is validated with the actual data for a few wards as shown in Table VII. Using these inputs, the tool predicts the new risk score for the entered inputs. Figure 13 shows the simulated risk scores. Table VIII shows how risk scores for a few wards changed with the change in input features.

From the predicted risk scores, it is evident that the potential risk level of the wards can be reduced by improving the ward medical infrastructure facilities (number of beds, hospitals). In most of the wards, increasing the ward facilities resulted in a change in the ward risk zones. However, it must be noted that no single attribute or feature can individually explain the measure of ward risk. Therefore, using this tool, the user can vary the inputs related to the number of beds and hospitals that can be added to the ward and precisely understand how it impacts the risk score. Users can accordingly increase the value of the input features till the ward risk reduces to an acceptable level. This type of impact analysis in turn assists the authorities to make informed decisions to increase the medical infrastructure facilities in the ward during the pandemic.

## IV. Discussion of Results

In our work, we analyzed the AQI and the ward level COVID-19 risk score for PMC-Pune in detail during the

TABLE VII
ACCURACY OF THE FORECASTING MODEL

| Ward | Date | Forecasted AQI | Actual AQI | Forecasted Hotspot | Actual Hotspot |
|---|---|---|---|---|---|
| 47 | 20-11-2020 | 113.89 | 128.69 | 3 | 3 |
| 47 | 21-11-2020 | 113.11 | 113 | 3 | 3 |
| 47 | 22-11-2020 | 115.29 | 113 | 3 | 3 |
| 47 | 23-11-2020 | 113.44 | 113 | 3 | 3 |
| 47 | 24-11-2020 | 118.50 | 129.5 | 3 | 3 |
| 144 | 20-11-2020 | 135.11 | 146.13 | 2 | 2 |
| 144 | 21-11-2020 | 139.14 | 146 | 2 | 2 |
| 144 | 22-11-2020 | 135.06 | 146 | 2 | 2 |
| 144 | 23-11-2020 | 143.90 | 146 | 2 | 2 |
| 144 | 24-11-2020 | 139.28 | 127.67 | 2 | 2 |

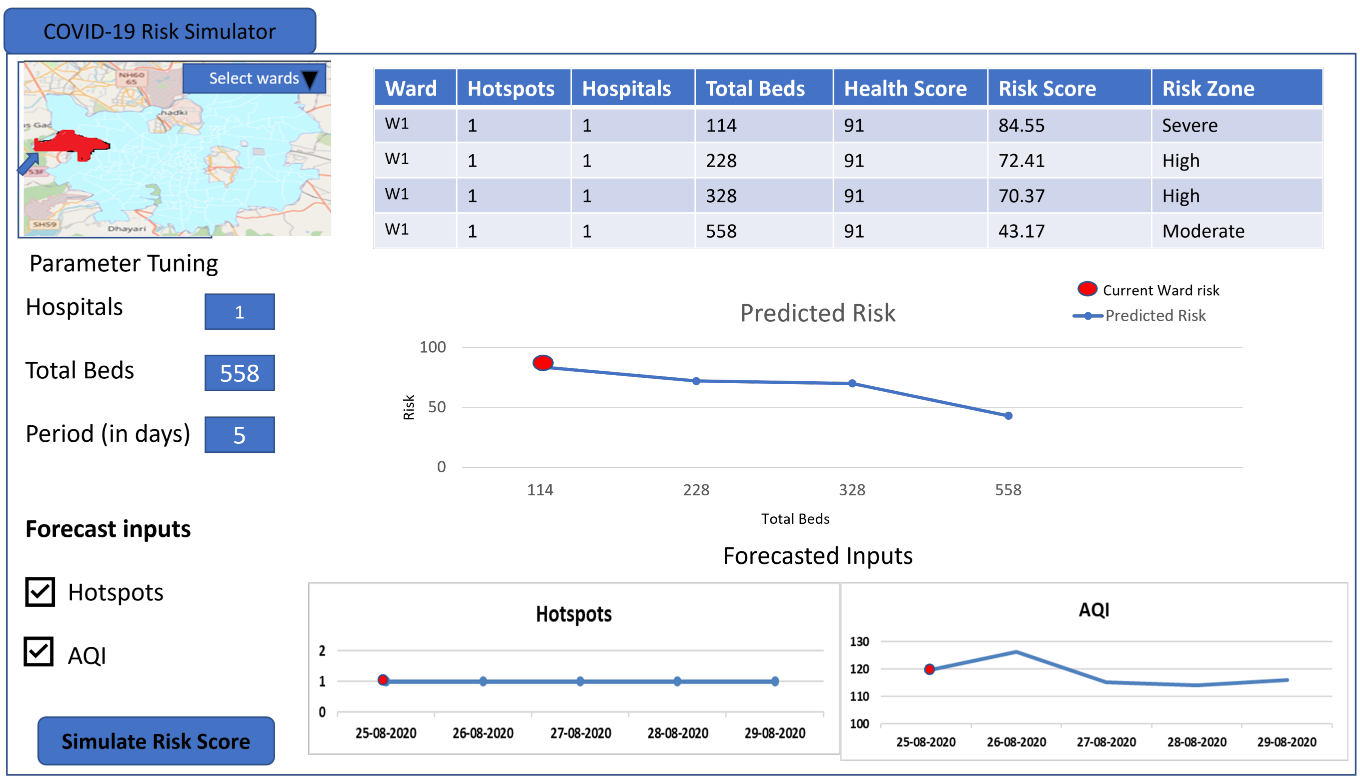


Figure 13. Ward risk analysis with the simulator tool

TABLE VIII
BEDS AND SIMULATED RISK SCORES

| Ward | Population | Health Score | Beds | Risk Score | Risk Zone |
|---|---|---|---|---|---|
| Vishrantwadi | 24678 | 91 | 114 | 84.55 | Severe |
| Vishrantwadi | 24678 | 91 | 228 | 72.41 | High |
| Gokhlenagar | 14614 | 75 | 81 | 64.29 | High |
| Gokhlenagar | 14614 | 75 | 160 | 54.71 | Moderate |
| Bopodi | 15834 | 59 | 60 | 78.26 | High |
| Bopodi | 15834 | 59 | 120 | 64.06 | High |
| Dhanori | 44060 | 91 | 43 | 86.66 | Severe |
| Dhanori | 44060 | 91 | 86 | 80.03 | High |
| Koregaon | 14685 | 82 | 248 | 42.57 | Moderate |
| Koregaon | 14685 | 82 | 496 | 34.21 | Low |

pandemic. AQI information was collected from air sensors deployed across the city and mapped to their respective wards. During Phase 1, 29 of the 34 wards showed improvement in the AQI levels. This improvement was majorly seen due to the strict lockdown rules imposed across the country to limit the movement of the population. While 21 out of 26 wards and 15 out of 31 wards showed improvement in the AQI levels during the Maintenance Phase and Phase 2, respectively. The surge in the AQI levels can be associated with the minimal restrictions during the unlock phases in the country, increased vehicular movements, increased energy consumption, and an increase in industrial operations related to the production of essential supplies. The significant decrease in certain parameters across the months can be attributed to nationwide lockdown, implementation of protocols, and other precautionary measures taken by the people during the pandemic to curb the COVID-19 cases across the country. Overall, AQI varied by 48%, sound by 18%, Ozone by 51%, $NO_2$ by 37%, $SO_2$ by 29% and $PM_{2.5}$ by 47% from March 2020 to July 2021.

Further, the simulator tool introduced in this paper forecasted the dynamic parameters like the AQI levels in the ward with 94.5% accuracy. Along with the user inputs, these forecasted values were used to simulate the risk scores of the wards. From the simulator results, the impact of improving ward-level medical and infrastructural facilities and its association with the ward risk score became evident. This tool not only helps authorities foresee hotspots or AQI levels in the wards over days, and categorize wards based on their risk levels but also helps them make informed decisions on incremental improvements in the ward medical and infrastructure facilities during challenging situations like the pandemic.

## V. Conclusion and Future Work

This paper provides an in-depth Spatio-temporal analysis of the AQI levels of the wards in PMC-Pune. Since our study is done at a granular ward level, with such analysis it becomes easier for authorities and city planners to take action in improving the overall environmental health and living conditions of the city. The study highlights how the air quality and other air pollutants varied in the wards over the months in PMC-Pune during the pandemic. With the imposition of restrictions, the wards saw significant improvements in the AQI. Thus, it advocates the need to improve the overall environmental health conditions of the wards in the city. Further, the simulator tool introduced in this work, helps authorities foresee and understand the impact of the risk score in the ward with incremental variations in the ward facilities during the pandemic. City planners can use this tool to assess and estimate the required increase in the ward medical facilities that would bring in a reduction in the ward risk score. Thus, this tool acts as an aid that assists city planners in their ward planning activities. When the overall condition of the ward improves, its ability to handle situations like the pandemic increases. With better facilities, the living conditions in the ward can improve and the ward can gradually attain self-sufficiency. Overall, this paper is based on the data available in PMC-Pune. However, the same can be extended to other smart cities where similar data is available.

## References

[1] S. P. Nagarmat, S. K. Shaukat, “A Data-Driven Approach for Region-wise Environmental Health and COVID-19 Risk Assessment Scores”, SMART 2021, The Tenth International Conference on Smart Cities, Systems, Devices and Technologies, ISSN 2308-3727, Pages 14-23, https://www.thinkmind.org/articles/smart_2021_1_30_40011.pdf
[2] COVID-19 Pandemic https://en.wikipedia.org/wiki/COVID-19_pandemic 2022.05.27.
[3] P. Asrani, M. S. Eapen, M. I. Hassan, S. S. Sohal, “Implications of the second wave of COVID-19 in India”, Lancet Respir Med. 2021 Sep;9(9):e93-e94. doi:10.1016/S2213-2600(21)00312-X.Epub2021Jun30.PMID:34216547;PMCID:PMC8245060.
[4] S. Comunian, D. Dongo, C. Milani, P. Palestini, ”Air pollution and COVID-19: the role of particulate matter in the spread and increase of COVID-19’s morbidity and mortality”, Int J Environ Res Public Health. 2020;17 DOI:https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7345938/.
[5] COVID-19 2nd wave. https://indianexpress.com/article/cities/pune/pune-covid-second-wave-fatality-rate-7390479/ 2022.05.27.
[6] Pune https://en.wikipedia.org/wiki/Pune 2022.05.27.
[7] Mercer 2019 Ranking https://mobilityexchange.mercer.com/insights/quality-of-living-rankings 2022.05.27.
[8] PMC wards https://pmc.gov.in/mr 2022.05.27.
[9] Pune pollution https://www.hindustantimes.com/cities/pune-s-air-quality-has-improved-in-2019-as-compared-to-2018-says-global-report/story-LlE88cLNm1hyDdklY22XAP.html 2022.05.30.
[10] Pune AQI https://catalogue.iudx.org.in/dataset/datakaveri.org-04a15c9960ffda227e9546f3f46e629e1fe4132b-rs.iudx.org.in-pune-env-aqm 2022.05.27.
[11] Air Pollutants COVID-19. https://indianexpress.com/article/explained/covid-19-lockdown-air-pollution-ozone-6479987/
[12] A. Pozzer et al.,“ Regional and global contributions of air pollution to risk of death from COVID-19”, Cardiovascular Research, Volume 116, Issue 14, 1 December 2020, Pages 2247–2253, https://doi.org/10.1093/cvr/cvaa288.
[13] X. Wu, R. C. Nethery, M. B. Sabath, D. Braun, F. Dominici, “Air pollution and COVID-19 mortality in the United States: Strengths and limitations of an ecological regression analysis”, DOI:10.1126/sciadv.abd4049.
[14] Affects of air pollution https://www.aqi.in/blog/health-and-economic-analysis-of-deaths-due-to-air-pollution-in-india/ 2022.05.27.
[15] City census department. https://www.pmc.gov.in/en/census 2022.05.27.
[16] Prophet forecasting model https://facebook.github.io/prophet/ 2022.05.27.