

Article

Impact of Green Energy Transportation Systems on Urban Air Quality: A Predictive Analysis Using Spatiotemporal Deep Learning Techniques

Rafia Mumtaz ^{1,*}, Arslan Amin ¹, Muhammad Ajmal Khan ¹, Muhammad Daud Abdullah Asif ¹, Zahid Anwar ² and Muhammad Jawad Bashir ¹

¹ School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan; aamin.ms19seecs@seecs.edu.pk (A.A.); ajmal.khan@seecs.edu.pk (M.A.K.); daud.abdullah@seecs.edu.pk (M.D.A.A.); mbashir.msds20seecs@seecs.edu.pk (M.J.B.)

² Department of Computer Science, The Sheila and Robert Challey Institute for Global Innovation and Growth, North Dakota State University (NDSU), Fargo, ND 58102, USA; zahid.anwar@ndsu.edu

* Correspondence: rafia.mumtaz@seecs.edu.pk

Abstract: Transitioning to green energy transport systems, notably electric vehicles, is crucial to both combat climate change and enhance urban air quality in developing nations. Urban air quality is pivotal, given its impact on health, necessitating accurate pollutant forecasting and emission reduction strategies to ensure overall well-being. This study forecasts the influence of green energy transport systems on the air quality in Lahore and Islamabad, Pakistan, while noting the projected surge in electric vehicle adoption from less than 1% to 10% within three years. Predicting the impact of this change involves analyzing data before, during, and after the COVID-19 pandemic. The lockdown led to minimal fossil fuel vehicle usage, resembling a green energy transportation scenario. The novelty of this work is twofold. Firstly, remote sensing data from the Sentinel-5P satellite were utilized to predict air quality index (AQI) trends before, during, and after COVID-19. Secondly, deep learning models, including long short-term memory (LSTM) and bidirectional LSTM, and machine learning models, including decision tree and random forest regression, were utilized to forecast the levels of NO₂, SO₂, and CO in the atmosphere. Our results demonstrate that implementing green energy transportation systems in urban centers of developing countries can enhance air quality by approximately 98%. Notably, the bidirectional LSTM model outperformed others in predicting NO₂ and SO₂ concentrations, while the LSTM model excelled in forecasting CO concentration. These results offer valuable insights into predicting air pollution levels and guiding green energy policies to mitigate the adverse health effects of air pollution.

Keywords: remotesensing; deep learning; urban air pollution; green energy; Sentinel-5P



Citation: Mumtaz, R.; Amin, A.; Khan, M.A.; Asif, M.D.A.; Anwar, Z.; Bashir, M.J. Impact of Green Energy Transportation Systems on Urban Air Quality: A Predictive Analysis Using Spatiotemporal Deep Learning Techniques. *Energies* **2023**, *16*, 6087. <https://doi.org/10.3390/en16166087>

Academic Editor: Boris Igor Palella

Received: 11 July 2023

Revised: 6 August 2023

Accepted: 16 August 2023

Published: 21 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The rapid urbanization and industrialization over the past seven decades have led to significant air pollution in large cities. Consequently, the air quality in urban centers has severely declined, posing risks to both human health and the environment [1,2]. Unfortunately, there is a lack of spatiotemporal air quality data for populated areas, hindering data-driven interventions to address environmental deterioration [1]. Regular air quality monitoring is essential to devise suitable strategies to prevent its negative effects on human health and the ecosystem of the area of interest [3,4]. Moreover, these monitoring methods can help track the immediate effects associated with the shift toward sustainable energy transportation systems. The detection and monitoring of trace gases using remote sensing data from satellites offer numerous advantages, such as global coverage for extended periods, enabling researchers to examine the concentration of trace gases over a wider area and

map their distribution [5]. Additionally, precise measurements of trace gases at multiple locations help identify sources and sinks, allowing for reasonable budgets to be generated. However, ongoing urbanization and industrialization have complicated the monitoring and control of air quality, particularly in rapidly developing nations, such as China and India. Despite suffering from poor air quality, these nations continue to produce synthetic gases to meet industrial growth without fully understanding the adverse environmental effects [1,6]. The high levels of air pollution in certain regions of Asia, such as South Asia and East Asia, have been associated with higher incidences of respiratory, mental, and other health issues [7–10]. It is estimated that Asia alone accounts for nearly 6.7 million premature deaths annually to poor air quality [11].

Besides India and China, Pakistan is also suffering from high air pollution levels owing to significant population and economic growth. The largest and fastest-growing sources of air pollution in Pakistan over the past decade have been the automotive and industrial sectors. During the period 2001–2013, the number of vehicles in Pakistan increased by 130% [12,13]. The city of Lahore alone accounts for 23–26% of extra carbon monoxide (CO) emissions due to an inadequate and inefficient mass-transit system [14].

In Lahore and Islamabad, emissions from vehicles significantly contribute to the deteriorating air quality, highlighting the urgent need for interventions. The air pollution crisis in Lahore is worsened by the involvement of 40% of the city's 7 million registered vehicles, which emit higher than permissible levels of hazardous air pollutants and contribute to smog-related issues. The situation is exacerbated by traffic congestion and the operation of heavy transport vehicles without road-worthiness certification [15]. It underscores the critical need for a transition to green energy in the transportation sector. Leveraging green transportation systems could substantially reduce air pollution and improve public health. Green transportation, which includes electric vehicles, hybrid cars, biofuels, and effective public transit systems, could substantially reduce air pollution and improve public health. It also helps combat climate change by reducing emissions, conserving energy, and promoting efficient resource use [16–18]. Pakistan was among the top 10 nations most hit by extreme weather events from 1991 to 2010 [19]. Since 2010, Pakistan has experienced numerous instances of intense heatwaves, torrential rains, and widespread floods. Hence, it is important to explore the chemical composition of the atmosphere over Pakistan by monitoring chemically active trace gases for understanding their impact on surface air temperature, heat waves, and climate change.

Atmospheric pollution is mainly caused by higher concentrations of various trace gas species including CO and oxides of nitrogen (NO_x) and sulfur (SO_x). The primary emissions from anthropogenic sources are the trace gases such as CO, nitrogen dioxide (NO_2), and sulfur dioxide (SO_2). CO is a hazardous air pollutant that negatively impacts air quality and poses risks to all forms of life. While present in trace amounts, it can severely impair oxygen supply in the body, leading to severe health problems which include drowsiness and irritation in the eyes [20]. The main sources of CO include vehicular emissions, fossil fuel combustion, industry, home heating, and vegetation burning, as well as natural sources like forest fires and volcanoes [2]. NO_2 , generated from burning fossil fuels in transportation, industry, and power generation, is another hazardous gas contributing to air pollution. Exposure to NO_2 can cause respiratory symptoms, reduced lung function, and increased cardiovascular risks, and has led to millions of premature deaths globally [21,22]. Similarly, SO_2 , generated by natural and human activities such as volcanic eruptions, fossil fuel burning, and industrial operations, directly affects air quality and poses risks to human health, ecosystems, and the environment. Pakistan's heavy reliance on coal and industrial activities has resulted in high SO_2 emissions, exceeding WHO standards [23]. During the COVID-19 pandemic, many cities experienced lockdown measures, resulting in reduced travel, cutting pollution, fuel consumption, and emissions. Post-pandemic, promoting sustainable options like cycling, electric vehicles, and public transport is crucial for climate mitigation. Cities adopt low emissions zones, shared mobility, and innovative transport for efficient, eco-friendly systems [24,25]. This situation closely resembled a green energy

transportation scenario, providing valuable insights into the potential improvements in air quality. Analyzing data from different time periods, including before, during, and after the pandemic, is essential to assess the potential impact of adopting green transportation systems [26]. The Pakistan Environmental Protection Agency (Pak-EPA) is attempting to analyze the concentration of NO₂ in a few Pakistani cities, along with other air quality examinations, but frequent updates are needed to investigate its influence on climate change. By leveraging data from the Sentinel-5P satellite, which measures air pollutants such as NO₂, CO, and SO₂, researchers can obtain frequent, accurate, and comprehensive information on the levels and distribution of these pollutants in urban areas [27]. This research enables the evaluation of the effectiveness of green transportation systems in reducing air pollution and its subsequent positive impact on public health. By accurately predicting the impact of transitioning to green transportation systems, policymakers can make informed decisions to prioritize sustainable transportation solutions and create healthier and more livable cities for everyone. The main contributions of this study are summarized as follows:

- Exploratory data analysis (EDA) is conducted on Sentinel-5p data to analyze the effects of green energy transportation on AQI trends in Lahore and Islamabad.
- Machine learning (ML) and deep learning (DL) models are created to forecast future air pollution levels to provide actionable insights and trends for policymakers to mitigate the harmful effects of air pollution.
- Comparative analysis of the traditional LSTM and bidirectional LSTM model is performed to predict concentrations of CO, NO₂, and SO₂. The bidirectional LSTM model provided an improvement of 10% over the traditional LSTM.

2. Related Work

To effectively address the issue of air pollution and assess the usage of green energy transportation in evaluating air quality, a comprehensive review of the relevant literature was conducted. Monitoring air quality is vital for a sustainable environment, achieved through various methods such as active/passive gas sampling, automatic point monitoring, photochemical/optical sensors, remote optical sensing, and imagery data. These approaches provide a holistic understanding of pollution, enabling precise assessments and targeted interventions. Combined with deep learning, these techniques offer a detailed air quality view, helping policymakers in developing effective pollution control strategies.

In traditional approaches, the active and passive sampling methods involve collecting samples of gases and vapors using pumps, sorbent tubes, or diffusion techniques [28,29]. The other approach that was utilized by the US Environmental Protection Agency was automatic point monitoring to detect and calculate the concentration of selected gases [30]. It provides continuous measurements and real-time data availability, which helps to identify pollution hotspots and develop mitigation strategies.

Traditional air quality monitoring methods have limitations. Active sampling is accurate but expensive, slow, and limited. Passive sampling is less sensitive, delayed, and prone to interference. Automatic point monitoring is costly, fixed, and has technical problems. Despite their usefulness, these methods should be combined with others for a complete understanding of air quality.

Apart from traditional methods, sensor-based systems, like photochemical and optical sensor systems, use light-sensitive sensors to detect pollutants in the air, offering mobility and simultaneous measurement of multiple pollutants [30]. This is especially useful for urban areas with diverse pollution sources. Another sensor-based approach is remote optical monitoring, which employs electromagnetic spectrum measurements to determine pollutant concentrations in real-time [31]. Space-based sensors also utilize image-based monitoring with aerosol optical thickness for assessing air pollutants, using various methods based on the application and available resources [32].

Air quality monitoring using Internet of Things (IoT) sensors allows real-time monitoring of air quality parameters [33]. The Atmospheric Air Surveil System (AASS) is a

transportable prototype that uses IoT sensors to monitor parameters like CO and CO₂ in outdoor environments. The AASS system utilizes microcontrollers, gas sensors, and GPS to measure gas concentrations and transmit the processed data to a Data Acquisition unit via MQTT and cloud services. The data are then stored in a remote server, which can be accessed remotely. This cost-effective AASS system offers real-time air quality data for analysis and decision-making.

The aforementioned techniques provide precise air quality measurements at a specific site, but they are restricted by spatial and temporal constraints. To address this, remote sensing techniques have emerged for broader regional and global air quality monitoring. These methods encompass satellite-based sensing, airborne measurements, and mobile ground-based monitoring [34]. Optical, radar, and LiDAR satellites offer high spatial and temporal resolutions, and advanced satellite-based technologies have the potential to provide highly accurate and comprehensive data than traditional ground-based monitoring methods [35–37].

Recent improvements in satellite and aerial remote sensing technology have made it possible to collect precise data on air pollution across vast areas [38–40]. This aids in precise air quality mapping and trend tracking. Deep learning and machine learning analyze these data for real-time monitoring and prediction; this is crucial for public health in urban areas [38]. These techniques excel due to their capacity to efficiently manage diverse data [39].

In recent years, there has been a growing interest in using machine learning and deep learning techniques for air quality prediction and estimation. Lin et al. used a random forest regression model to forecast PM_{2.5} and nitrate levels based on road site data [41]. The model showed strong predictive accuracy, gauged by the R-squared value. However, precision depends on data quality and site conditions, potentially limiting applicability to diverse locations.

Shafi et al. [42] utilized K-means clustering to detect abrupt changes in air quality. The method successfully grouped data into clusters based on similarity, detecting notable changes linked to weather and human activities. This highlights the K-means clustering promise in crafting early warning systems to predict air quality shifts. These techniques provide prompt action to counter the adverse effects of pollution on health and the environment.

Choi et al. [43] employed affordable sensors and machine learning to monitor Seoul's air quality for urban planning. Their model effectively predicted pollutants, like PM_{2.5} and NO₂, using sensor data. The study underscores the value of budget-friendly sensor-based monitoring and machine learning for the swift identification of pollution areas, providing proactive solutions in air quality management and urban planning.

Li et al. [44] used a machine learning model to assess the impact of clean air actions in improving air quality in Beijing on the basis of data from 2008 to 2017. The findings revealed substantial decreases in pollutants including PM_{2.5}, SO₂, and NO₂ due to these actions. The study underscores the actions' efficacy while underscoring the necessity for ongoing endeavors to sustain and enhance air quality. Moreover, it showcases machine learning's utility in gauging the impact of environmental policies on air pollution.

Huang et al. [45] developed an accurate PM_{2.5} concentration prediction model using remote sensing data and machine learning algorithms. The random forest algorithm performed the best with an R-squared value of 0.80, RMSE of 6.62, and MAE of 4.58. In another study, Banerjee et al. [46] investigated the potential relationship between air pollution, economic growth, and COVID-19 mortality rates in India using machine learning techniques. The study concluded that air pollution levels and economic growth were significant predictors of COVID-19 mortality rates in India. Specifically, a 10 µg/m³ increase in PM_{2.5} concentrations was associated with a 9.4% rise in COVID-19 deaths, while a 1% increase in gross domestic product (GDP) was linked to a 5.5% decrease in COVID-19 deaths.

Cosemans et al. [47] compared the performance of three machine learning algorithms in predicting air pollutant concentrations at different locations across Europe. Random forest and support vector regression outperformed both linear regression and regularization. Researchers have also proposed a deep learning-based model based on air quality and meteorological data to accurately identify the major sources of air pollution [45,48], which can help policymakers take targeted actions to reduce emissions. Zhang et al. and Zhou et al. [49,50] have developed deep learning-based approaches that utilize satellite remote sensing data to identify the sources of particulate matter pollution with high accuracy.

Besides monitoring air quality, researchers have also attempted to estimate the concentration of pollutants and predict air quality based on measured data. Kow et al. [51] proposed a new approach for air quality estimation using image data and deep learning neural networks, achieving high accuracy in predicting AQI values in real time. Similarly, Sharma et al. [52] reported a novel technique for forecasting PM10 concentrations in the most polluted hotspots in Australia using satellite data and deep learning methods, achieving high accuracy with a mean absolute error of less than 10. Another study by Kurnaz et al. [53] predicted the concentrations of two air pollutants, SO₂ and PM10, in the city of Sakarya in Turkey, with high accuracy. Similarly, Mao et al. [54] have reported a deep learning method for predicting air quality. In another study, the researchers proposed an effective convolutional neural network (CNN) for visual understanding of transboundary air pollution based on Himawari-8 satellite images [55]. The CNN-based model was shown to accurately identify and classify different types of pollutants.

This [56] study presents a novel deep predictive model for accurately predicting spatiotemporal PM2.5 in Los Angeles County using meteorological data, wildfire data, remote-sensing satellite imagery, and ground-based sensor data. The model employs a graph convolutional network (GCN) and a convolutional long short-term memory (ConvLSTM) to learn and predict spatiotemporal correlations in air pollution data. The model achieves state-of-the-art accuracy in predicting hourly PM2.5 at seven sensor locations in Los Angeles County. The root mean square error (RMSE) and normalized root mean square error (NRMSE) decrease over time with later frames, but this is expected as the nature of PM2.5 results in concentrations 24 h in the future being more correlated with 24 h in the past as compared to concentrations 48 h in the future.

Das et al. [57] compared the performance of MLP, RNN, and LSTM models in predicting air pollutants such as PM10 and SO₂. The evaluation metrics used were MSE, RMSE, MAE, and R2. The LSTM model outperformed the MLP and RNN models in terms of accuracy. The study also compared the performance of the proposed model with existing studies in the literature and found that the LSTM model predicted PM10 and SO₂ pollutants with high accuracy. The study provides valuable insights into the use of deep learning models for air pollutant prediction.

In [58], multiple techniques for forecasting air pollution levels using statistical and deep learning methods were used. The data were used from government-built air pollution monitoring stations in Kolkata and evaluated the performance of different models based on two performance indicators, RMSE and MAE. It is observed that Holt–Winter-based forecasting models outperform for PM2.5, PM10, and SO₂ time series, while deep learning-based models, such as ConvLSTM and Bi-LSTM dominate for NO₂ time series data.

Shin et al. [59] present a study on the use of an FCN-based deep learning regression model for real-time indoor air quality monitoring. The dataset is preprocessed to reduce skewness and convert the raw 1D dataset into 2D image input/output datasets, after which the model is trained with various hyperparameters. The results show a decrease in the average prediction error for the MAE and RMSE compared with a deep neural network model.

LSTM and BiLSTM networks excel in air quality forecasting by capturing sequential dependencies, handling missing data, and modeling complex temporal relationships. They retain crucial information from past observations, considering weather and pollution

factors, and enhance prediction by incorporating future insights. Optimizing these models requires experimentation, considering data quality, features, and architecture [60].

Machine learning and deep learning offer advantages over traditional methods. They handle large, irregular data, learn intricate patterns, and leverage remote sensing for precise pollution source detection. These models inform policies, aid urban planning, and offer cost-effective data-driven solutions for air quality management [61].

Tables 1 and 2 summarize the performances of various statistical machine learning and deep learning models used for predicting air quality.

Table 1. An overview of air quality prediction studies using machine learning.

| Study | Model | Advantages | Disadvantages | Results |
|-----------------------------|-------------------|--|--|---|
| Choi et al. [43] | LR and RF | Maps urban air quality using mobile sampling with low-cost sensors | Limited to a specific region and type of pollution | 86.1–90.6% Mean absolute error (MAE) |
| Cosemans et al. [47] | Linear regression | Develops Europe-wide spatial models of fine particles and nitrogen dioxide | Limited to Europe and specific pollutants | 70.2–79.8% Coefficient of determination (R2) |
| G.-Y. Lin et al. [41] | RF and XGBoost | Uses machine learning to predict PM2.5 and nitrate concentrations | Limited to a road site station | 0.84–0.88 R2 |
| J. Shafi and A. Waheed [42] | k-means | Detects abrupt changes in air quality using k-means clustering | Limited to detecting abrupt changes, not predicting concentrations | 82–96% F1-score |
| Jin et al. [35] | Random forest | Accurately predicted PM2.5 and nitrate concentrations using long-term data | Only focused on one station | R2 = 0.819 (PM2.5) R2 = 0.812 (nitrate) |

Table 2. An overview of air quality prediction studies using deep learning.

| Study | Model | Advantages | Disadvantages | Results |
|------------------------|--------------|---|--|----------------------------|
| Huang et al. [45] | CNN-LSTM | Predicts air pollutants using meteorological data and satellite-based images | Requires extensive data and computing resources | 85.5–87.2% RMSE |
| Liu et al. [33] | LSTM and MLP | LSTM Outperformed MLP | Limited data | R2 = 0.897, MAE = 14.41 |
| Muthukumar et al. [56] | LSTM and RNN | Comparison of the performance of different deep learning models for forecasting | The performances have not been investigated for all pollutants except PM10 | R2 = 0.86, RMSE = 10.27 |

Table 2. Cont.

| Study | Model | Advantages | Disadvantages | Results |
|---------------------|----------------------|--|--|-------------------------------|
| Middya, et al. [58] | ConvLSTM and Bi-LSTM | Comprehensive analysis, optimal model building, and seasonal variation detection | Focuses only on air pollutant-specific optimal model building and does not cover other aspects of air pollution. | RMSE = 30% MAE = 22.5% |
| Shin et al. [59] | FCN model | The dataset is acquired through computational simulation under various indoor geometrical conditions | Does not provide a detailed analysis of the computational resources required for the FCN-based model | MAE = 43.14% RMSE = 34.77% |

3. Methodology

The study was conducted in two major cities of Pakistan—Lahore and Islamabad. The dataset for the study was based on atmospheric monitoring data collected by the Sentinel-5P satellite from 2018–2021. The dataset was preprocessed, including the conversion of L2 to L3 products, filtering for the study areas, interpolation, and outlier removal. The data were converted from mole/m² to the AQI standard unit. An exploratory data analysis (EDA) was performed to analyze the AQI trends before, after, and during COVID-19 in both cities. Two forecasting models were trained to predict future trends to support data-driven policy interventions for improving AQI. Figure 1 illustrates the methodology followed in this study.

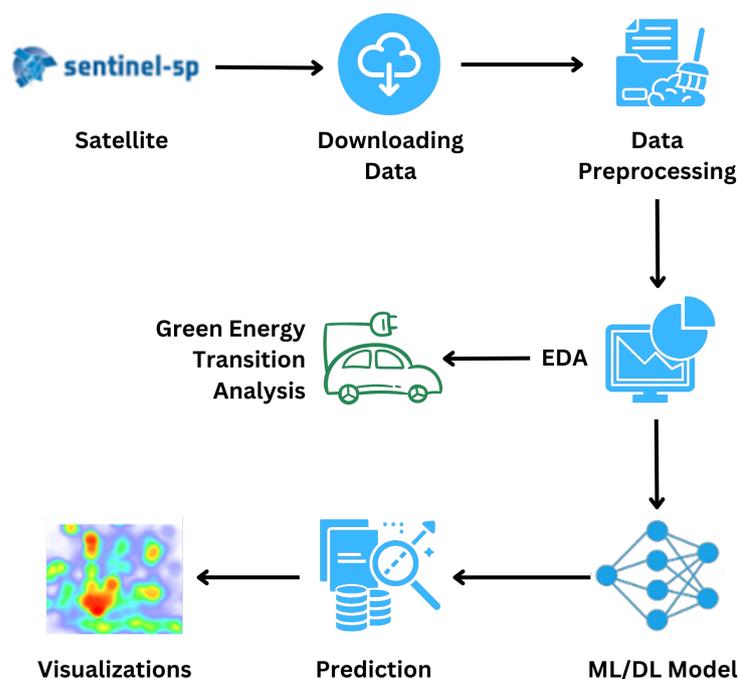


Figure 1. Overview of the proposed methodology.

3.1. Study Area

Air pollution is a serious problem for major population centers of Pakistan as it has been ranked third among the countries with the highest levels of air pollution [62]. Lahore and Islamabad, shown in Figure 2, are two major cities and neither is immune from the curse of environmental pollution. Both cities are renowned for their cultural and historical significance but they also suffer from air quality issues. Lahore is the second largest city and the provincial capital of Punjab with a population of over 11 million people growing at an annual rate of 3% since 1998, resulting in substantial urbanization and a growing reliance on transportation [63]. This trend has led to significant problems with road congestion and increased emissions in the area. According to the annual global survey conducted by IQAir, a Swiss manufacturer of air purifiers, the city of Lahore experienced a significant rise in its air pollution levels in 2022. The city has jumped more than 10 places to become the world's most polluted city. IQAir measures air quality by assessing the concentration of harmful PM_{2.5} particles, which can damage the lungs. Lahore's air quality deteriorated from 86.5 micrograms of PM_{2.5} particles per cubic meter in 2021 to an alarming level of 97.4 micrograms per cubic meter in 2022.

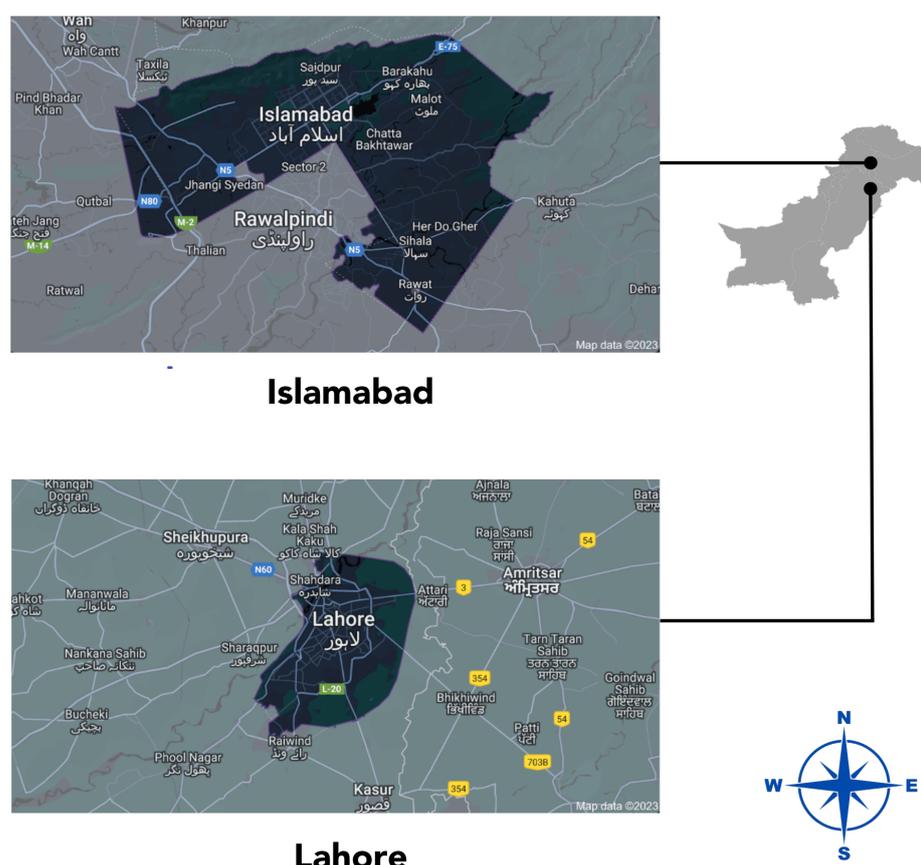


Figure 2. Area of study in Pakistan.

The primary sources of pollution in Lahore comprise transportation, industries, agriculture (through crop residue burning), open waste burning, and inefficient fuel consumption in the commercial and domestic sectors. Air pollution in Lahore is predominantly caused by the transportation sector, accounting for a staggering 83% of total pollution. This sector alone is responsible for 127 Gg of emissions. The majority of these emissions, amounting to 104.76 Gg, are produced by two-stroke vehicles like motorcycles, scooters, and auto-rickshaws. Motorcars, jeeps, and wagons contribute a further 16.34 Gg to the total emissions. The primary pollutant emitted in Lahore is carbon monoxide, resulting from

the incomplete combustion of fuels in mobile engines and other processes, as illustrated in Figure 3.

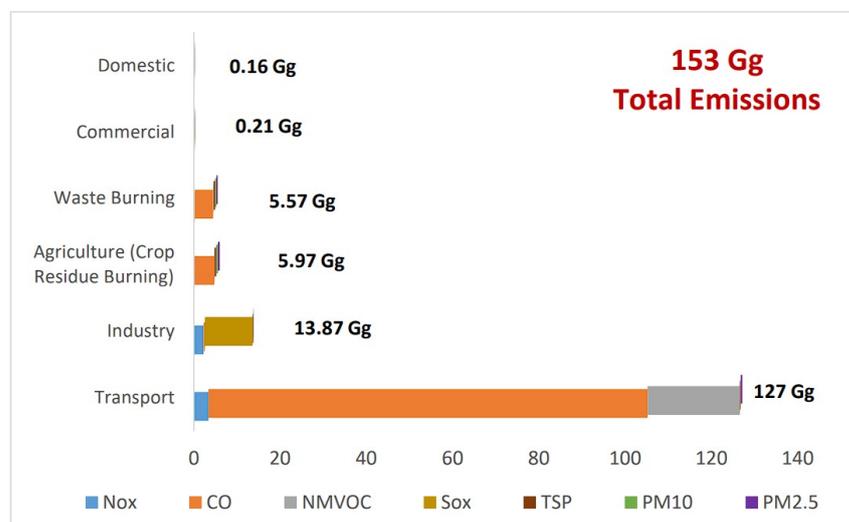


Figure 3. Air pollution sources in Lahore [63].

Non-methane volatile organic compounds (NMVOCs) and nitrogen oxides (NOx) are secondary major pollutants, largely emitted from the transport and industrial sectors. Particulate matter, including total suspended particulates, PM2.5, and PM10, are emitted in lower concentrations. Apart from transportation, emissions from the industrial (9%), domestic (0.11%), and commercial (0.14%) sectors also contribute to the overall pollution levels in Lahore. These sectors primarily use inefficient fuels, such as coal and diesel oil, leading to emissions of pollutants. Additionally, the common practice of burning crop residues (3.9%) and waste (3.6%) in the outskirts of Lahore also contributes significantly to the city's pollution. The resulting pollution levels in Lahore far exceed the recommended limits, leading to a surge in respiratory ailments among the population. It has been estimated that if air quality guidelines were adhered to, residents could potentially increase their life expectancy by an average of 6.8 years [63].

Islamabad, the capital of Pakistan, is home to over 1.7 million people, with an average growth rate of 3.7%. This has resulted in rapid urbanization, causing an increase in transportation [64]. While its air quality is generally better than in Lahore, it still faces pollution challenges. In 2022, it was reported as unhealthy, with the average level of hazardous air pollutant PM2.5 measured at 49.33 micrograms per cubic meter, exceeding the permissible limit of 35 micrograms per cubic meter [65]. Vehicular emissions are identified as the primary cause of particle pollution in Islamabad, leading to levels as high as 41.63 micrograms per cubic meter [66]. Astonishingly, these emissions contribute to a substantial 43% of the country's overall air pollution. The usage of non-compliant diesel fuel, containing hazardous sulfur dioxide, exacerbates the problem. It is crucial to address this issue promptly by implementing stricter regulations on vehicle emissions, promoting cleaner fuels, and ensuring compliance with environmental standards. Taking these measures will help improve air quality and safeguard public health in Islamabad [67]. Emissions in both cities are primarily attributable to transportation activities, which led to the selection of these urban areas for an analysis of the trends in the AQI during the COVID-19 pandemic. Concurrently, a prediction model was also developed. The COVID-19 period, marked by frequent lockdowns, saw a significant reduction in intracity transportation. Our study was devised to evaluate the impact of this reduction on air quality. The change in AQI trends captured by the prediction model provides policymakers with a measure of the effectiveness of their transition toward green energy policies.

3.2. Data Acquisition and Preprocessing

There are various datasets available that provide information on air quality related to CO, NO₂, and SO₂. One of the most commonly used datasets is the one provided by the World AQI project, which collects and aggregates air quality data from different sources worldwide. The World AQI project provides hourly data on various air pollutants and aggregates these data into an overall air quality index that can be used to compare air quality between different cities or regions. While the World AQI project provides a valuable source of information on air quality, certain potential limitations must be considered when using these data. The AQI data are compiled from various sources and are susceptible to gaps, in terms of geographic coverage, data quality, lack of detail, and time lag. The AQI provides a broad overview of air quality but typically lacks the level of detail needed for more localized or in-depth analysis. Additionally, there could be a time difference between the measurement of pollutants and their inclusion in the AQI data. Therefore, it is important to be aware of these limitations and consider using additional sources of data and information to supplement and verify the findings. Other datasets that provide information on air quality related to these pollutants include those provided by national or regional air quality monitoring networks. For example, in the United States, the Environmental Protection Agency (EPA) provides data on air quality through its Air Quality System (AQS), covering thousands of monitoring sites across the country. The measured data provided information about various pollutants, including CO, NO₂, and SO₂. In addition to these datasets, there are also satellite-based datasets that contain information on atmospheric pollutants such as NO₂ and SO₂ on a global scale. For example, the Sentinel-5P mission, which is part of the European Space Agency's Copernicus program, provides high-resolution data on common atmospheric pollutants, including CO, NO₂, SO₂, and O₃. This dataset can be used to monitor air quality on a global scale. Overall, these datasets play an important role in helping scientists, policymakers, and the public to understand and address air quality issues arising from hazardous pollutants like CO, NO₂, and SO₂.

To collect the data for our study, the Python API was used to query the Sentinel-5P database, which contains atmospheric data collected by satellite. Incorporating Lahore and Islamabad as pivotal points in the utilization of Sentinel-5P's data can produce critical knowledge. These cities in Pakistan confront significant air pollution levels, partly due to industrial operations, traffic exhaust, and agricultural burning in surrounding locations. Therefore, they provide ideal scenarios for exploiting Sentinel-5P's data collection capabilities in the environmental and atmospheric monitoring sector. The robustness and reliability of Sentinel-5P's data, collected through the Tropospheric Monitoring Instrument (TROPOMI), are key attributes. TROPOMI's ability to monitor gases, such as CO, NO₂, and SO₂, has proven consistently accurate, making the data highly reliable. Furthermore, the global scientific community, environmental agencies, and government bodies place considerable reliance on the data produced by Sentinel-5P. This extensive data collection can significantly support local and national policymakers in making well-informed decisions about environmental policies and mitigation strategies, underscoring the instrument's significant reliance and relevance. The API was programmed to retrieve data for the specific region of interest, which was defined by a GeoJSON file of Pakistan. GeoJSON is a file format used to represent geographical data and is commonly employed in mapping applications. By providing the GeoJSON file, the Python API was able to extract the data for the cities of Lahore and Islamabad. The satellite collected data daily basis from 2017 to 2021 for the three pollutants of interest, NO₂, CO, and SO₂. However, monitoring data for April 2018 onward is made public, and the same is used in this study. The data were downloaded in 'netcdf' format, which is a standard format used for storing, manipulating, and analysis of scientific data. The size of the dataset for SO₂, NO₂, and CO was 1651.15 GB, 771.26 GB, and 274.23 GB, respectively. The following data preprocessing steps were taken:

- **Conversion from Level-2(L2) to Level-3(L3) Products:** L2 products are the minimally processed or unprocessed data that a satellite sensor has collected. These deliverables often include measurements of particular variables with great spatial and temporal

resolution, such as atmospheric composition. However, the accuracy and interpretation of the measurements may be impacted by noise, artifacts, or anomalies in the L2 data. The L2 data are transformed into L3 products using the HARP Python package to address these limitations. Aggregating the L2 data over more expansive spatial and temporal scales is implemented to reduce data noise and improve measurement accuracy. The conversion of the L2 to L3 product aggregation procedure enables a more thorough understanding of the atmospheric composition in the studied area. L3 products offer a broader viewpoint and capture the qualities of the variables of interest that have been averaged or aggregated. These tools help carry out national or international assessments, capture trends, and explore long-term patterns.

- **Filtering and conversion to CSV:** The L3 data were filtered separately for Lahore and Islamabad for each pollutant to isolate data relevant to the study area and eliminate extraneous data points. The data were then converted from netcdf files to CSV files to simplify data manipulation and facilitate further analysis.
- **Checking for null values:** The CSV files were examined for null values to identify missing data points. It was observed that these missing values were clustered in specific areas, suggesting that interpolation could be employed to estimate the missing values.
- **Interpolation:** Linear interpolation was performed to estimate the missing values based on neighboring data points. This technique is commonly used to impute missing data in scientific datasets and results in the creation of a complete dataset.
- **Outlier removal:** To ensure data quality, any outliers were removed by utilizing GeoJSON files containing the geographical boundaries of Lahore and Islamabad. This step filtered out data points located outside the city boundaries, improving the accuracy and reliability of the dataset.
- **Duplicate values:** To address the issue of duplicate values, the pandas library provides two key functions: duplicated() and drop_duplicates(). The duplicated() function was employed to identify duplicate values in a DataFrame, while the drop_duplicates() function was used to eliminate those duplicates.
- **Conversion to AQI standard unit:** The initial gas concentrations, measured in units of moles per square meter (mol/m^2), were converted to air quality index (AQI) standard units. The mass concentrations of each gas in micrograms per cubic meter ($\mu\text{g m}^{-3}$) were calculated using the molecular weight and molar volume of air. The AQI values for each gas concentration were determined by comparing these to the relevant AQI standards.

After performing the pre-processing steps on Sentinel-5p data, the clean data statistics are shown in Table 3.

Table 3. Statistics of air quality after pre-processing Sentinel-5p data.

| Pollutant | Lahore | | | Islamabad | | |
|-------------------|-----------|--------|--------|-----------|--------|--------|
| | Count | Mean | Std | Count | Mean | Std |
| 1 NO ₂ | 3,068,481 | 304.17 | 352.2 | 1,608,122 | 242.09 | 260.27 |
| 2 SO ₂ | 3,035,501 | 266.5 | 245.67 | 1,574,120 | 328.85 | 394.20 |
| 3 CO | 3,036,501 | 319.84 | 126.07 | 1,583,159 | 220.8 | 89.16 |

It is important to note that the specific AQI standards may vary depending on the location and regulations governing air quality monitoring.

3.3. Training of Machine/Deep Learning Models

Two machine learning models, random forest, and decision tree, as well as two deep learning models, LSTM and bidirectional LSTM, are utilized to predict air quality. The LSTM model, being a recurrent neural network (RNN) variant, is particularly well-suited

for modeling sequential data, making it an optimal choice for analyzing time-series data. On the other hand, random forest and decision tree models are well-known for their ability to handle structured data and make accurate predictions in various domains. Each model was trained using six input features (“latitude”, “longitude”, “year”, “month”, “day”, and “hour”) and one output label (“respective gas”). Initially, the input features consisted of longitude, latitude, and timestamp. However, to enhance the feature set and improve the accuracy of the model, feature engineering was applied to the timestamp. This resulted in the generation of four new features, namely year, month, day, and hour. These additional features are crucial for both increasing the accuracy of the model and performing Exploratory Data Analysis. The dataset is split into 80% for training and 20% for testing with random shuffling. This approach ensures that the model learns from a variety of data points during training, improving its ability to generalize and make accurate predictions on unseen data. Random shuffling also helps in assessing the model’s performance by providing an unbiased representation of the dataset for evaluation.

3.3.1. Decision Tree

The decision tree model is an effective machine-learning approach that enables the division of the feature space into different and independent regions. It can effectively represent non-linear correlations between these variables by using the predictor variables. Decision tree models do have the propensity to overfit the data, which means they could become overly specialized to the training dataset and have trouble generalizing successfully to new data. To mitigate overfitting, k-folds cross-validation is used. Additionally, the criterion parameter is set to mean squared error (MSE), which guides the decision tree’s construction by minimizing the squared differences between predicted and actual values. Also, the splitter parameter is set to “best”, which determines the best possible split point at each node based on the chosen criterion. This encourages the model to make more informed and accurate decisions during the tree-building process.

3.3.2. Random Forest

The random forest model is an ensemble model that combines multiple decision trees to reduce overfitting, as shown in Figure 4. It can improve the performance of the decision tree model by reducing its variance. We also performed k-fold cross-validation for each model to obtain a more reliable estimate of the model’s performance. We used five-fold, shuffled the data, and used MSE as a criterion. The mean of the MSE scores across all folds was utilized to obtain the cross-validation MSE.

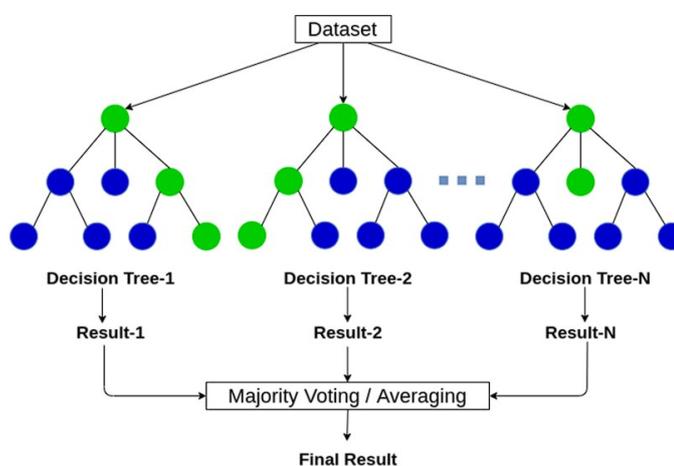


Figure 4. Architecture of the random forest model [68].

3.3.3. Long Short-Term Memory (LSTM) Model

The LSTM regression model consists of one LSTM layer with 50 units and a dense output layer with one unit (Figure 5). The LSTM layer uses the ReLU activation function and has an input shape of (1, number of features). The output layer has no activation function and one unit. The model is compiled using the Adam optimizer and the mean squared error loss function. The mean absolute error is used as a metric to evaluate the model's performance during training. The LSTM layer has several parameters that can be adjusted to optimize the model's performance. The dropout and recurrent dropout parameters are used to prevent overfitting by randomly dropping out some of the LSTM layer's output values during training. The return sequences and return state parameters can be used to return the LSTM layer's output sequences and final state, respectively. The LSTM model is trained for 100 epochs with a batch size of 32. During training, the model's performance is evaluated on a validation set, and the MSE and mean absolute error (MAE) are calculated for the testing set after training.

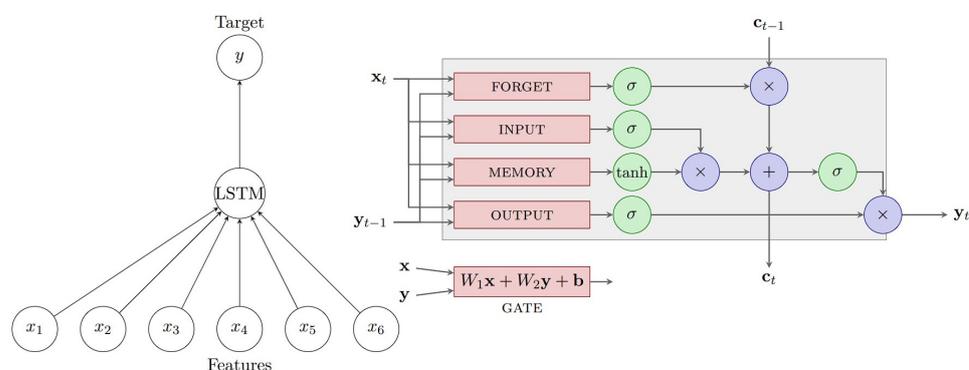


Figure 5. An overview of LSTM architecture.

3.3.4. Bidirectional LSTM

The bidirectional LSTM (Figures 6 and 7) regression model consists of one bidirectional LSTM layer with 50 units and a dense output layer with one unit. Like the LSTM model, the bidirectional LSTM layer uses the ReLU activation function and has an input shape of (1, number of features). The output layer has no activation function and one unit. The model is compiled using the Adam optimizer and the MSE loss function, with the mean absolute error used as a metric to evaluate the model's performance during training. The bidirectional LSTM layer processes the input sequence in both forward and backward directions, allowing the model to take into account both past and future information. This improves the model's performance compared to the LSTM model, especially for time-series data with long-term dependencies. The bidirectional LSTM layer has the same parameters as the LSTM layer, including the dropout and recurrent dropout, return sequences, and return state. The model is trained for 100 epochs with a batch size of 32, and the MSE and MAE are calculated for the testing set after training.

To summarize, this section has discussed the study area, dataset, and analytical techniques used to predict air quality in Lahore and Islamabad based on remote sensing data from the Sentinel-5P satellite. The prediction models employed include machine learning and deep learning techniques such as random forest, decision tree, LSTM, and bidirectional LSTM. The models are trained on preprocessed data and evaluated to predict air quality parameters for the cities of interest. When it comes to air quality forecasting using LSTM, it is essential to address the time-consuming calculations and stability concerns to achieve efficient and accurate predictions. To tackle these challenges and optimize the forecasting process, several strategies can be implemented. Firstly, reducing the number of LSTM layers or units can significantly improve computational efficiency without sacrificing forecasting performance. By optimizing the architecture and finding the right balance between complexity and accuracy, training and inference times can be reduced. In

addition to reducing complexity, regularization techniques play a vital role in stabilizing LSTM models. Applying dropout or recurrent dropout to the LSTM layers helps prevent overfitting and enhances the generalization capability of the model. This ensures that the LSTM network learns meaningful patterns from the air quality data and produces reliable forecasts. To further improve stability, incorporating batch normalization into the LSTM layers is beneficial. By normalizing the activations within each layer, batch normalization helps stabilize the training process, leading to faster convergence and better overall model stability. Addressing gradient explosion or vanishing is essential for LSTM models in air quality forecasting. Implementing gradient clipping techniques prevents the gradients from becoming too large or too small during backpropagation. This regularization technique ensures stable updates to the LSTM parameters, enabling more accurate and reliable predictions. Considering the nature of air quality forecasting, which often involves long sequences, truncated backpropagation through time (BPTT) can be employed. By breaking down the input sequences into smaller subsequences, the memory requirements and computation times are reduced. Although some long-term dependencies may be sacrificed, the trade-off allows for stable and efficient training of LSTM models. Optimizing hardware and software resources is also crucial for efficient air quality forecasting. Leveraging hardware accelerators, such as GPUs or TPUs, can significantly speed up the calculations involved in LSTM training and inference. Additionally, using optimized software frameworks like TensorFlow or PyTorch allows for efficient utilization of parallel processing capabilities and optimized implementations, further enhancing performance. By implementing these strategies specifically in the context of air quality forecasting, researchers and practitioners can effectively address the challenges of time-consuming calculations and stability concerns associated with LSTM models. This leads to more efficient training and inference, improved stability, and reliable forecasts, ultimately aiding in better decision-making and management of air quality.

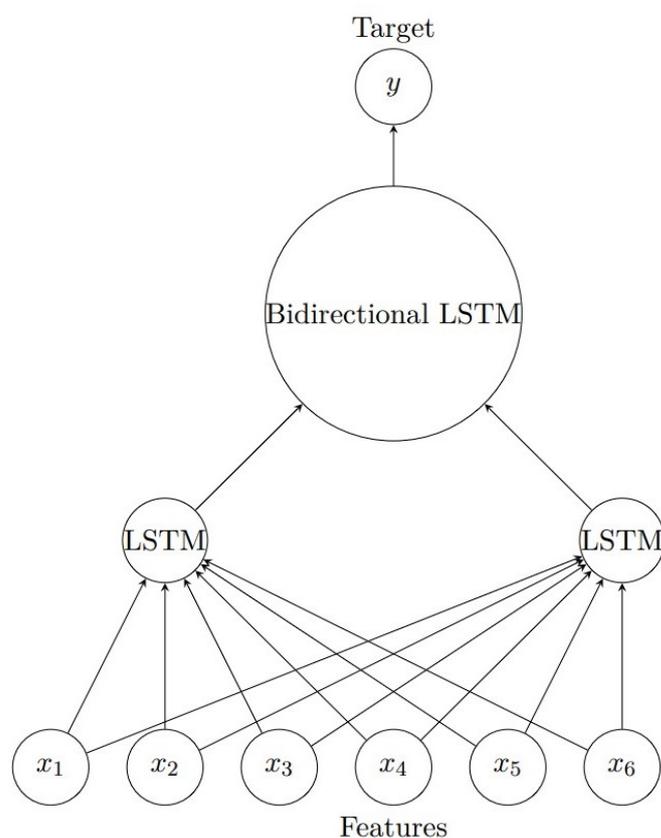


Figure 6. An overview of bidirectional LSTM architecture.

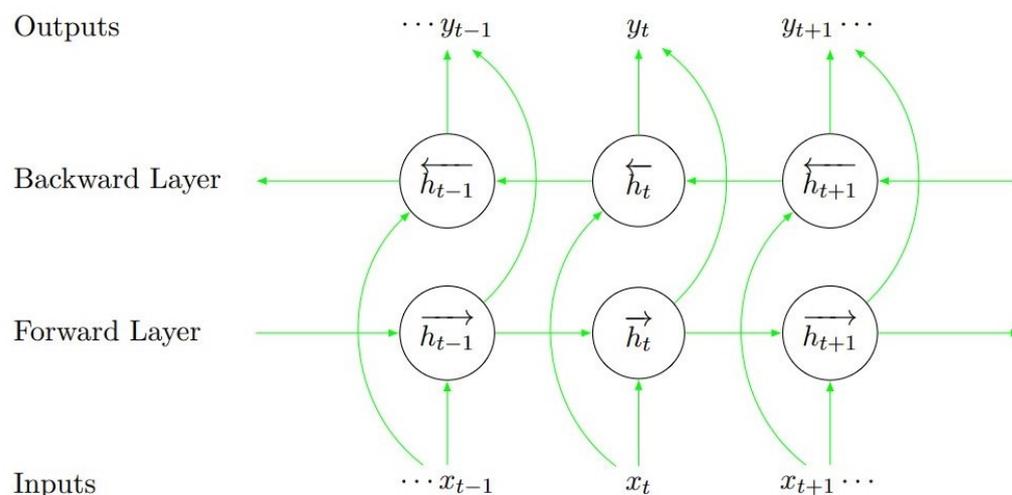


Figure 7. Structure of bidirectional LSTM.

4. Results

The section presents the results of the study on air quality monitoring using machine learning and deep learning techniques. The effectiveness of different models to predict air quality levels is evaluated using commonly employed metrics such as MSE and MAE. The results are then discussed in detail, including exploratory data analysis of seasonal trends in air quality and the impact of COVID-19 on air pollution levels in Lahore and Islamabad.

4.1. Predicting Air Quality Levels Using Machine Learning and Deep Learning

MSE and MAE are used as evaluation metrics to evaluate the machine/deep learning models. MSE measures the average squared difference between the predicted values and the true values of a set of samples. MAE measures the average absolute difference between the predicted values and the true values of a set of samples. The results of the decision tree and random forest are computed for all three hazardous gases but for discussion purposes, only the results for CO are shown in Figures 8 and 9, where it can be seen that both models are overfitting. Overfitting is an undesirable phenomenon in machine learning when a model is too complex, leading to high variance and low bias. In other words, the model fits the training data too closely for training data, and as a result, it fails to generalize well to new data. To overcome this problem the K-fold cross-validation is applied during the model training. K-fold cross-validation is a widely used technique in machine learning that helps to address the issue of overfitting. By dividing the data into K subsets or folds, the model is trained on a different subset each time, allowing for more robust and accurate estimates of the model's performance on new data. This technique can also help to identify potential sources of bias or variability in the data and model, enabling more informed model selection and parameter tuning. The results are calculated for all the gasses and but for the representation of results, only SO₂ results are shown in Figures 10 and 11, where it can be seen that the problem of overfitting is eliminated.

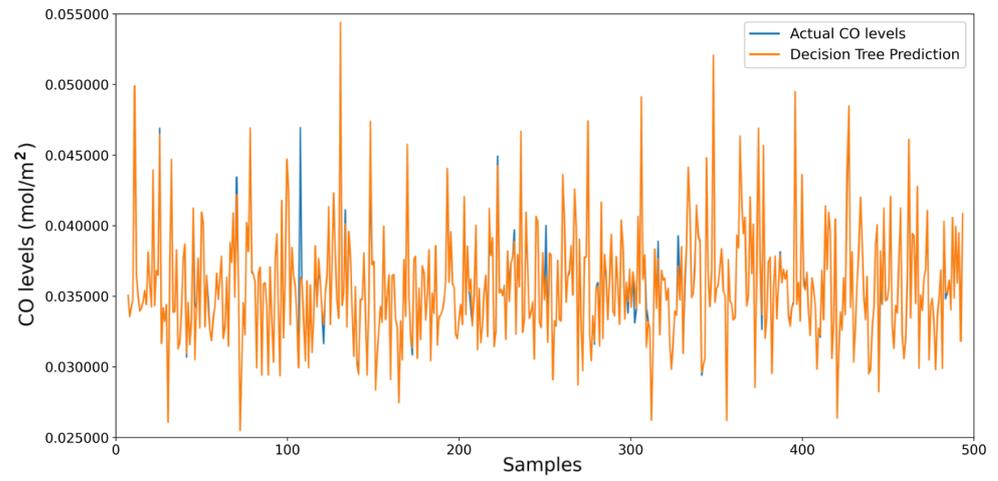


Figure 8. CO decision tree results on 500 samples.

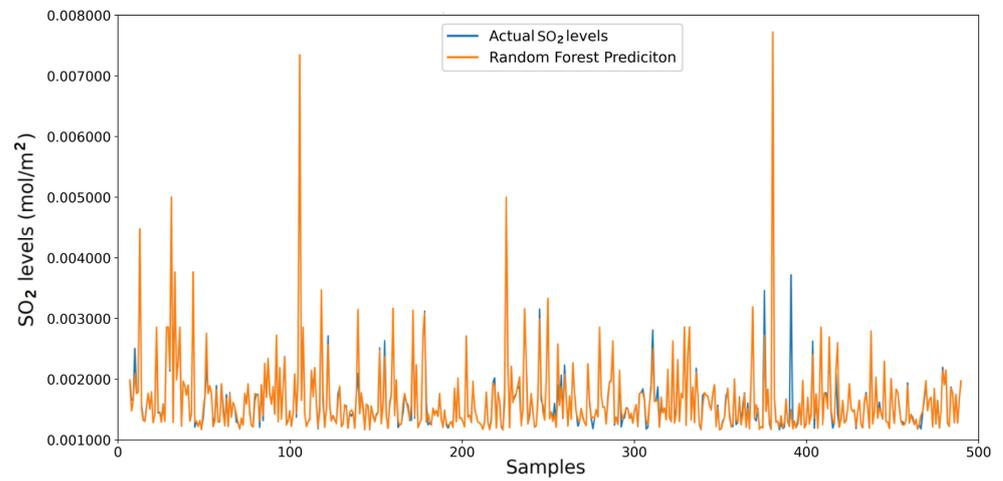


Figure 9. SO₂ random forest results on 500 samples.

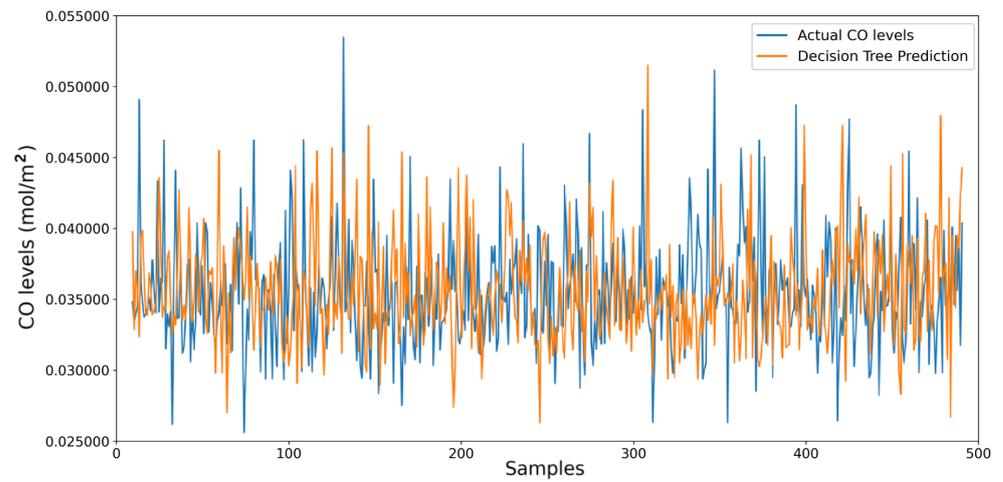


Figure 10. CO K-fold decision tree results on 500 samples.

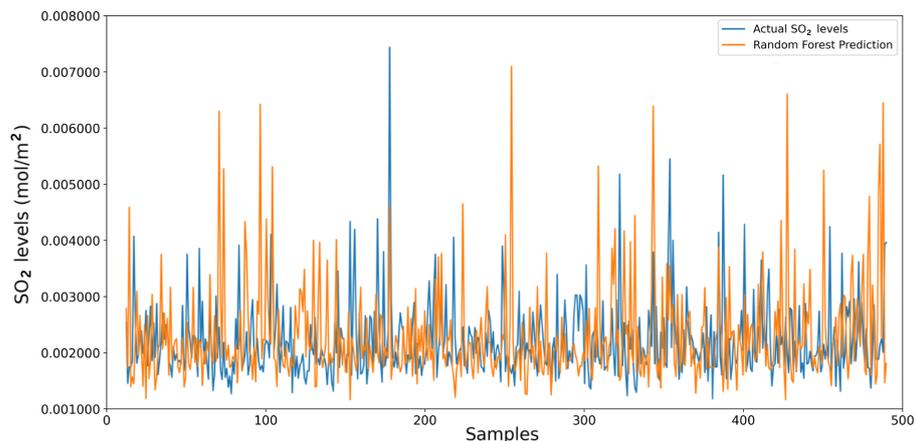


Figure 11. SO₂ K-fold random forest results on 500 samples.

The LSTM and bidirectional LSTM models are utilized for forecasting the concentrations of NO₂, SO₂, and CO gases. However, in terms of presenting the results, the discussion primarily focuses on the prediction of NO₂, as illustrated in Figures 12 and 13, respectively. It is evident from Figure 12 that the LSTM model provides accurate prediction and adequately captures the underlying patterns and trends in the data.

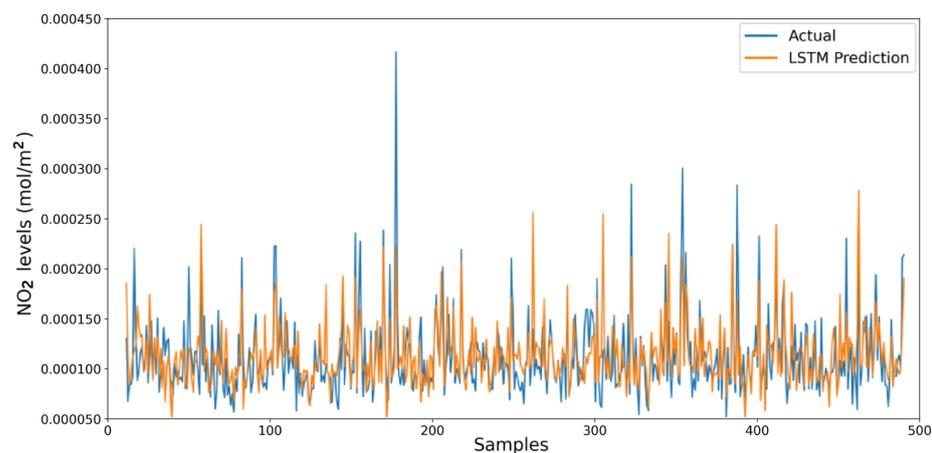


Figure 12. NO₂ LSTM prediction on 500 samples.

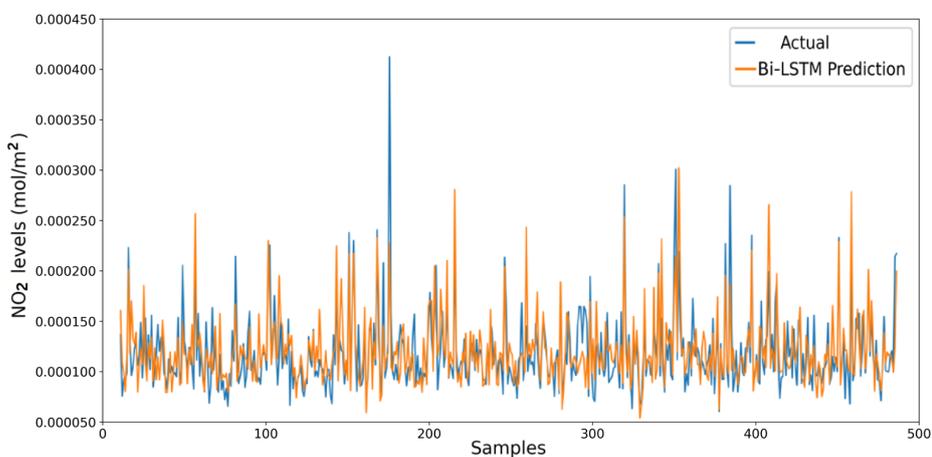


Figure 13. NO₂ bidirectional LSTM prediction on 500 samples.

The bidirectional LSTM results, shown in Figure 13, confirm that the model is effective in accurately predicting air quality levels. The predicted values closely match the actual values, indicating that the model has successfully captured the underlying patterns and trends in the data. Moreover, the results obtained from bidirectional LSTM are improved as compared to the results obtained from the standard LSTM model, indicating that bidirectional LSTM can improve the accuracy of air quality predictions. The previously presented figures display generalized learned patterns derived from a randomly shuffled dataset. In contrast, Figure 14 represents the results of the LSTM model trained on a timestamp-ordered, non-shuffled dataset. This methodological distinction enhances the accuracy of predictions as it respects the temporal continuity of the data, thereby providing a more precise forecast.

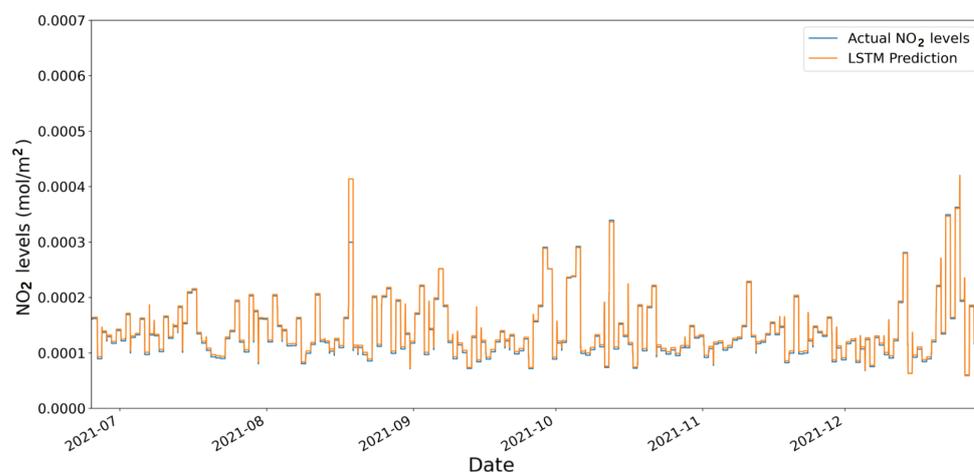


Figure 14. Prediction of NO₂ levels using LSTM in the time period from 1 July 2021 to 31 December 2021.

Evaluation Metrics

The study used both machine learning and deep learning techniques for air quality prediction. The machine learning models included decision tree and random forest, while the deep learning models included LSTM and bidirectional LSTM. These models were trained and tested on the pre-processed dataset using evaluation metrics, such as MSE and MAE.

The results of the study using deep learning models are given in Tables 4 and 5. The results indicate that the bidirectional LSTM outperformed the machine learning models in predicting air pollution levels. The bidirectional LSTM model achieved, respectively, an MSE of 0.41, 0.38, and 0.34 for NO₂, SO₂, and CO. As regards the MAE performance metric, values of 0.41, 0.38, and 0.34 were obtained for NO₂, SO₂, and CO, respectively. This indicates the high degree of accuracy of said bidirectional LSTM model in predicting air pollution levels. The LSTM model also performed well, achieving an MSE of 0.50, 0.44, and 0.47; and an MAE of 0.48, 0.40, and 0.52 for NO₂, SO₂, and CO, respectively. In contrast, the machine learning models, specifically the decision tree and random forest models, provided reasonable predictions of air pollution levels but were not as accurate as the deep learning models.

Table 4. Comparison of LSTM and bidirectional LSTM models for predicting air pollutant concentrations in Lahore.

| Pollutant | LSTM | | Bidirectional LSTM | |
|-------------------|------|------|--------------------|------|
| | MSE | MAE | MSE | MAE |
| 1 NO ₂ | 0.55 | 0.49 | 0.44 | 0.43 |
| 2 SO ₂ | 0.66 | 0.54 | 0.61 | 0.52 |
| 3 CO | 0.34 | 0.41 | 0.26 | 0.37 |

Table 5. Comparison of LSTM and bidirectional LSTM models for predicting air pollutant concentrations in Islamabad.

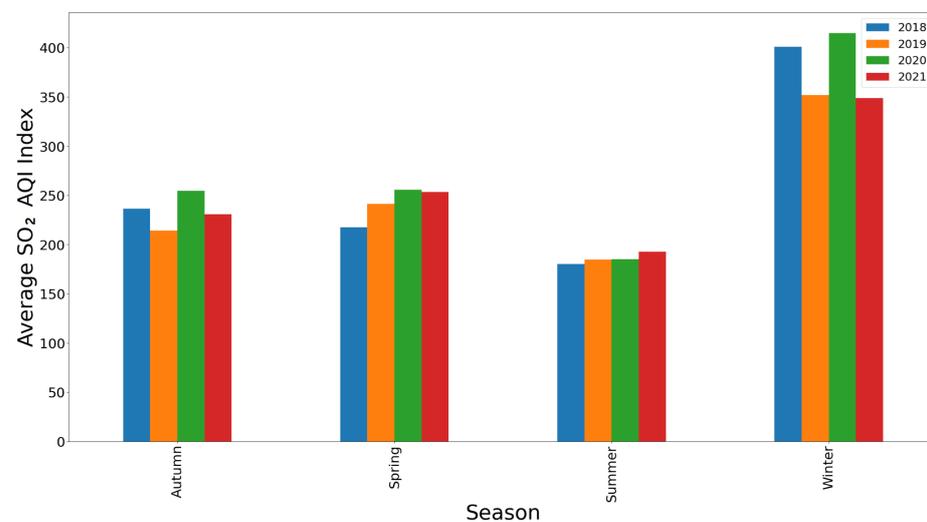
| Pollutant | LSTM | | Bidirectional LSTM | |
|-------------------|------|------|--------------------|------|
| | MSE | MAE | MSE | MAE |
| 1 NO ₂ | 0.50 | 0.48 | 0.41 | 0.43 |
| 2 SO ₂ | 0.44 | 0.40 | 0.38 | 0.39 |
| 3 CO | 0.47 | 0.52 | 0.34 | 0.45 |

4.2. Exploratory Data Analysis (EDA)

After preprocessing the data, the EDA analysis is performed to summarize the main characteristics of the dataset and gain a better understanding of the patterns and relationships within the data.

4.2.1. Seasonal Comparison of AQI Trends

The calculated AQI levels of SO₂ for Lahore and Islamabad are shown in Figures 15 and 16, respectively. Similar patterns were found for seasonal AQI trends of NO₂ and CO. The AQI levels vary throughout the year due to various factors, such as weather conditions, seasonal activities, and changes in human behavior. In Lahore, higher levels of NO₂, SO₂, and CO were observed during the winter months (November to February) due to increased usage of heating appliances and more significant vehicular traffic. During the summer months (June to August), AQI levels were lower due to increased wind speeds and warmer temperatures, which can help to disperse the pollutants.

**Figure 15.** Lahore SO₂ seasonal AQI trends.

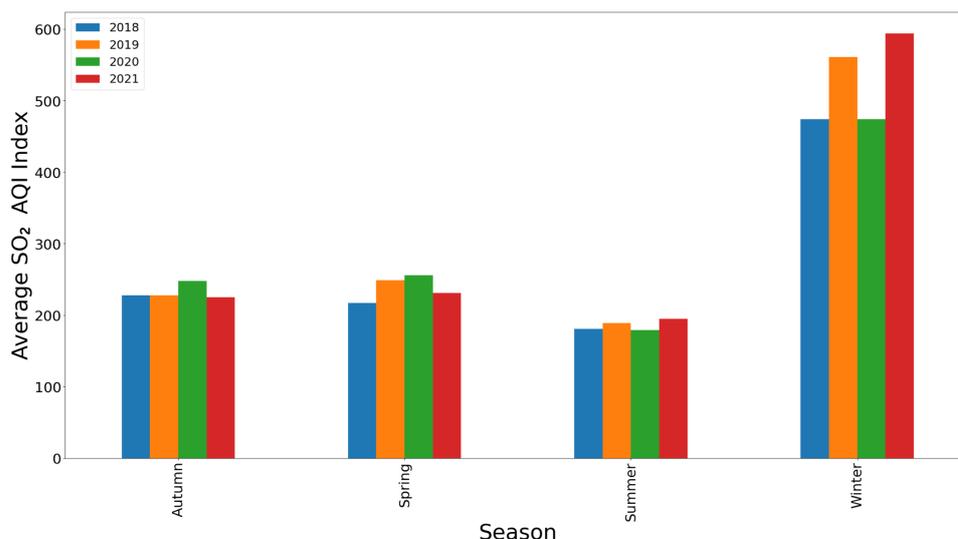


Figure 16. Islamabad SO₂ seasonal AQI trends.

4.2.2. AQI Trends before and after COVID-19 in Lahore and Islamabad

The COVID-19 pandemic had a significant impact on air quality in both cities. The AQI results based on NO₂ levels in Lahore and Islamabad are shown in Figures 17 and 18), respectively. Similar results were obtained for SO₂ and CO pollutants. Before the pandemic, the AQI levels in Lahore and Islamabad were significantly higher, with raised levels of SO₂, NO₂, and CO. However, during the COVID-19 lockdowns, the AQI levels dropped significantly in both cities, with a considerable reduction in SO₂, NO₂, and CO levels.

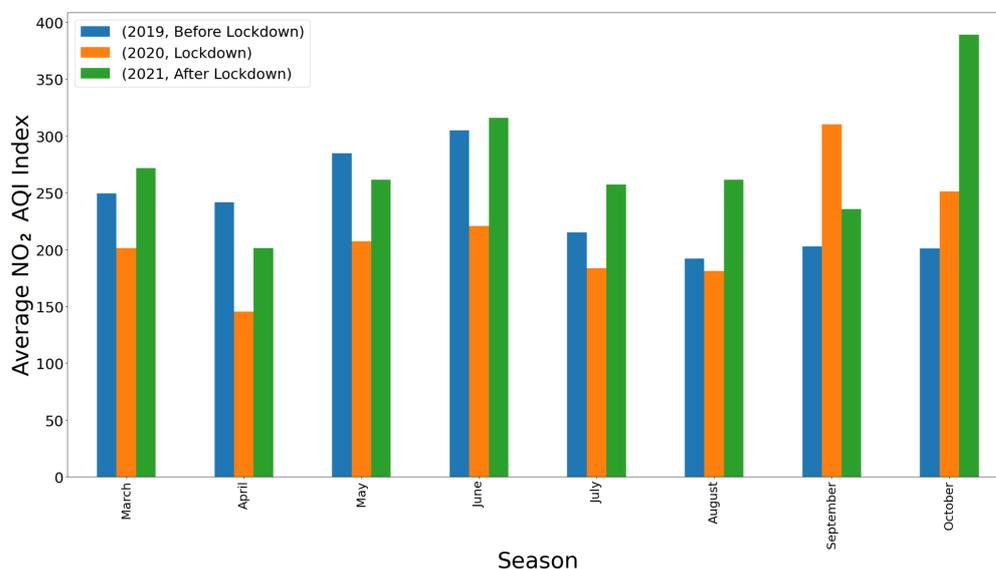


Figure 17. Lahore NO₂—after and before the COVID-19 AQI trend.

In Lahore, the SO₂ levels dropped by 66%, NO₂ levels dropped by 64%, and CO levels dropped by 43% during the lockdown period. Similarly, in Islamabad, the SO₂ levels dropped by 64%, NO₂ levels dropped by 57%, and CO levels dropped by 51% during the lockdown period. After the lockdowns were lifted, the AQI levels started to increase again, with SO₂, NO₂, and CO levels returning to their pre-pandemic levels by the end of 2020.

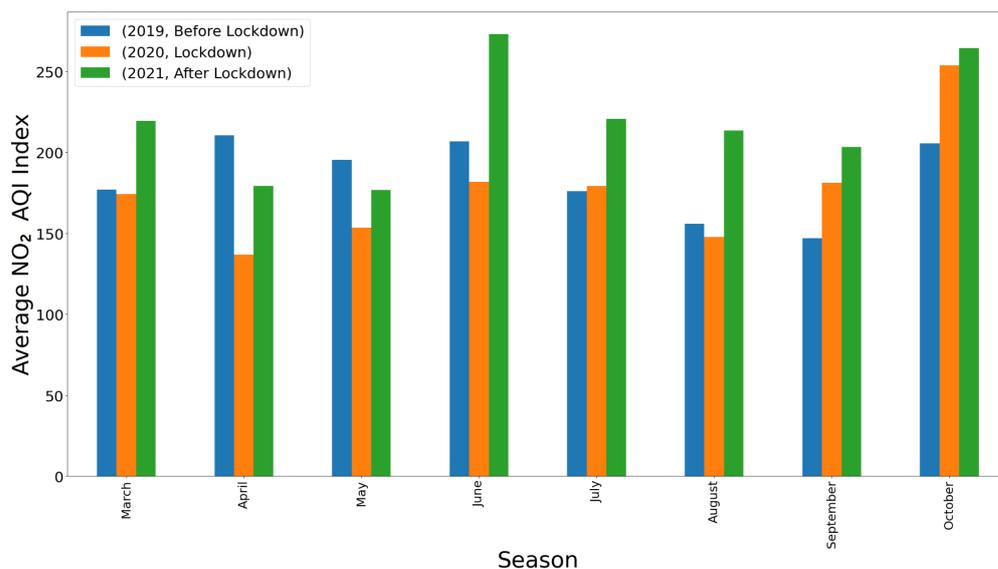


Figure 18. Islamabad NO₂—after and before the COVID-19 AQI trend.

4.3. Heatmap Visualization—A Conclusive Overview

In our air quality monitoring study, we utilized heatmap visualization to effectively represent the spatial distribution of pollutant concentrations in Lahore and Islamabad. This technique facilitates the identification of pollution hotspots and temporal patterns underlying the raw data of the Sentinel-5P satellite that has been processed through deep learning algorithms. We chose Tableau as our data visualization tool for generating heatmaps due to its user-friendly interface, flexibility, and powerful mapping capabilities. The heatmap visualizations are created by assigning a color gradient to the gas concentration values, making it easy to distinguish between areas with high and low concentrations of the pollutant. The AQI chart, depicted in Figure 19, illustrates the air quality index values offering a visual representation of air pollution levels. The heatmaps of the NO₂ gas for Lahore before, during, and after the COVID-19 pandemic are shown in Figures 20, 21, and 22, respectively. Similarly, Figures 23–25 show NO₂ pollution patterns for Islamabad. Heatmaps generated for other pollutants (SO₂ and CO) showed similar patterns for both cities.

| Air Quality Index (AQI) | Index Value | Colors |
|-------------------------|-------------|--------|
| Good | 0-50 | Green |
| Satisfactory | 50-100 | Yellow |
| Moderate | 101-150 | Orange |
| Unhealthy | 151-200 | Red |
| Very Unhealthy | 201-300 | Purple |
| Hazardous | 300-500 | Maroon |

Figure 19. Air quality index chart [69].

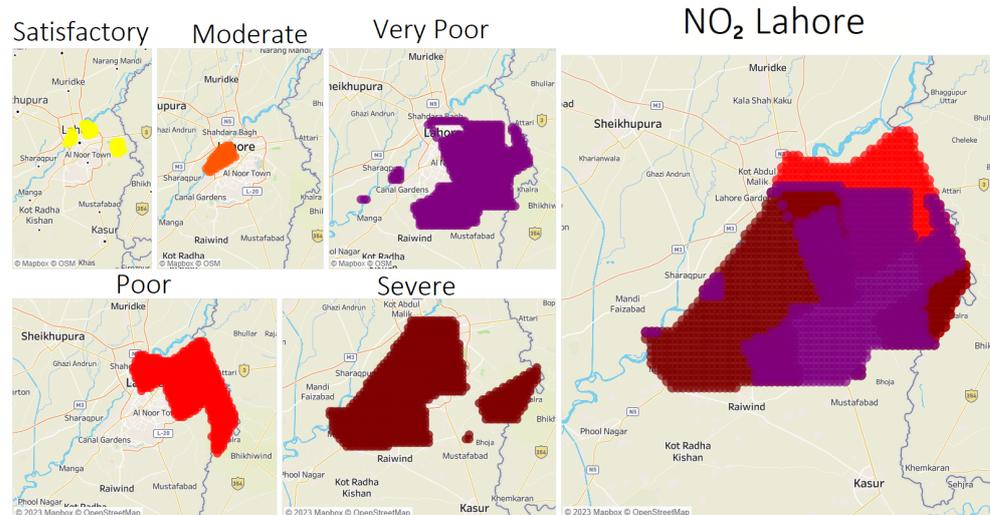


Figure 20. NO₂ trend (Lahore)—before COVID-19.

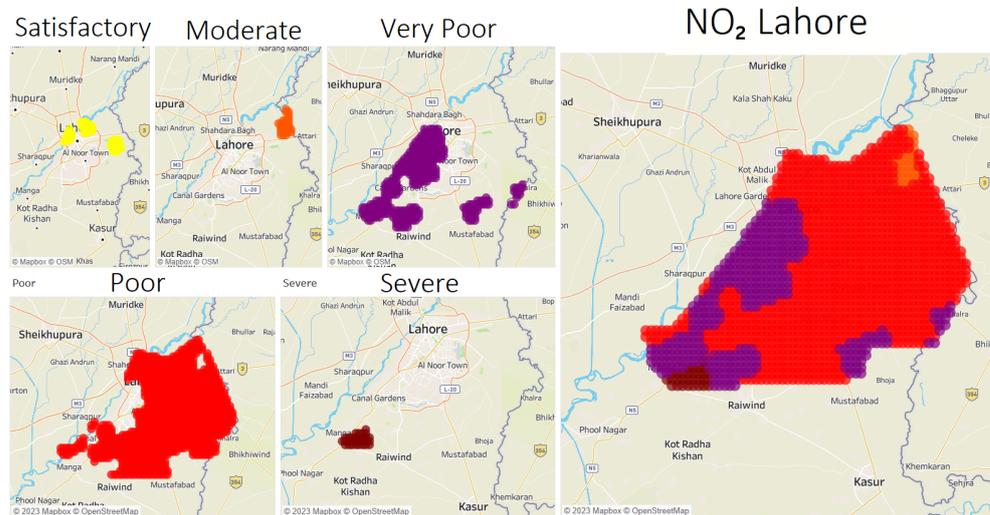


Figure 21. NO₂ trend (Lahore)—during COVID-19.

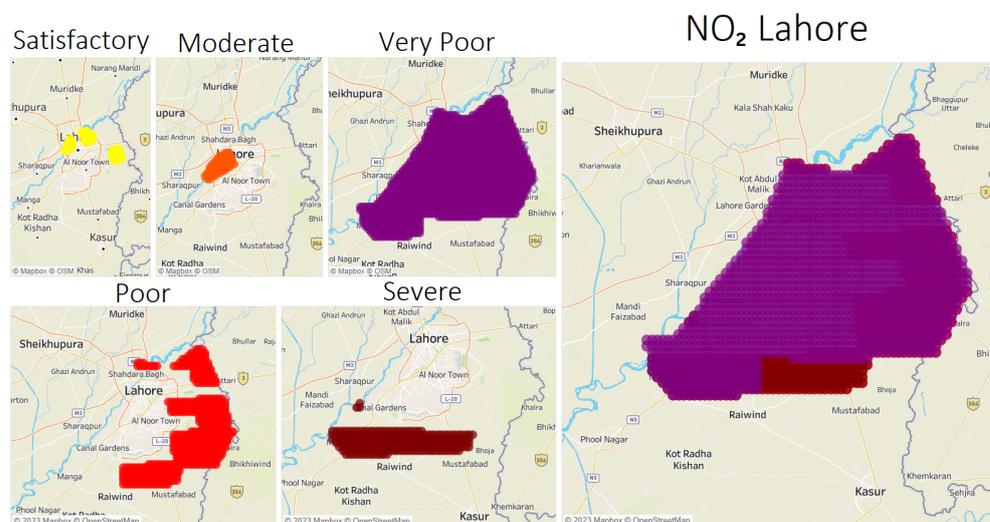


Figure 22. NO₂ trend (Lahore)—after COVID-19.

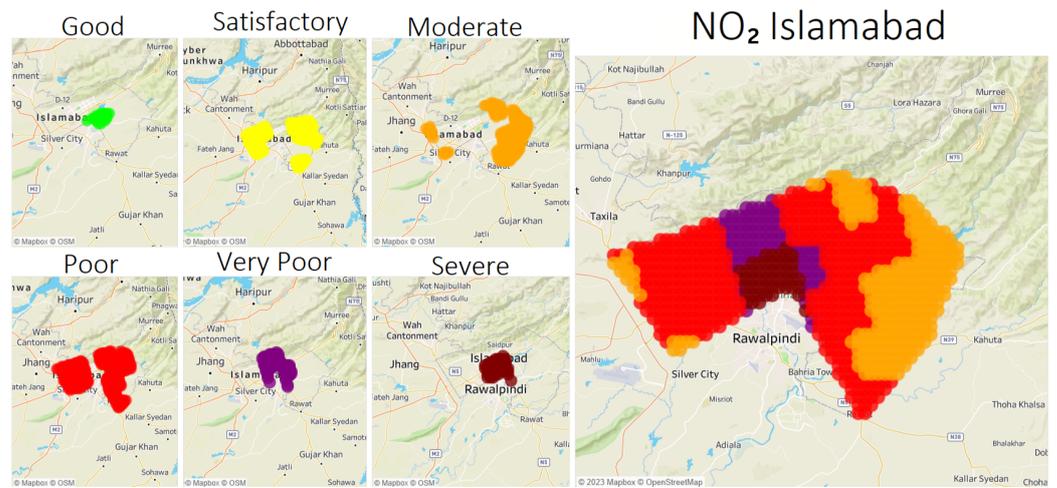


Figure 23. NO₂ trend (Islamabad)—before COVID-19.

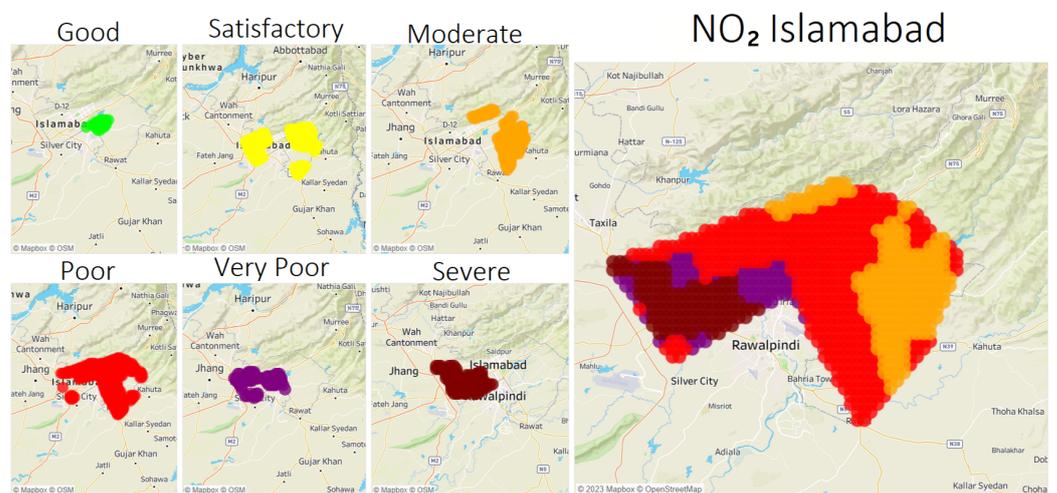


Figure 24. NO₂ trend (Islamabad)—during COVID-19.

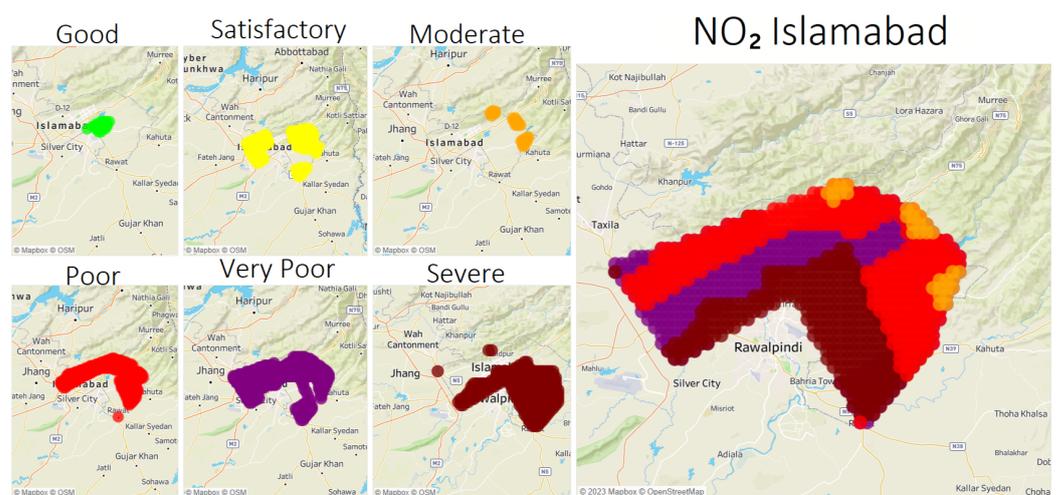


Figure 25. NO₂ trend (Islamabad)—after COVID-19.

In Lahore, a distinct variation in air quality visualizations can be observed as demonstrated by the AQI Index, which transitioned from “Very Poor” and “Severe” categories (pre-COVID-19—Figure 20) to predominantly “Poor” (during COVID-19—Figure 21). This change underscores the improvement in air quality during the COVID-19 pandemic owing to less traffic and industrial activity. During the post-pandemic period (Figure 22), the AQI reverted to “Severe” and “Very Poor” classifications, which suggests a significant influence of human activities on the city’s air quality. Consequently, there is an urgent need for policy development aimed at reducing industrial emissions and traffic related to air pollution.

A comparable trend was observed in Islamabad’s air quality, where a slight improvement was noted within the “Severe” category during the COVID-19 pandemic (Figures 23 and 24). This subtle improvement may be attributed to the implementation of targeted lockdown measures. However, during the post-pandemic (Figure 25), the AQI index values not only reverted but also deteriorated further. This observation emphasizes the necessity to maintain stringent emission standards for both industrial and transportation sectors, particularly in light of increasing population growth. Based on these observations, some action points for policy-making are discussed in Section 5.

The exploratory data analysis presented in this section provides valuable insights into the seasonal trends in air quality and the impact of external factors such as COVID-19 on air pollution levels. The study confirms a significant decrease in air pollution levels during the COVID-19 lockdown, highlighting the potential of reducing emissions to improve air quality. Overall, the findings of the study demonstrate the potential of advanced techniques such as machine learning and deep learning for more accurate and efficient air quality monitoring. The evaluation metrics used in this study, such as MSE and MAE, provide a quantitative assessment of the model’s accuracy, making it easier to compare the effectiveness of different air quality monitoring approaches. The use of deep learning models, specifically bidirectional LSTM, can provide highly accurate predictions of air pollution levels, enabling effective air quality management strategies. The study’s findings can assist policymakers and stakeholders in developing strategies to reduce air pollution and improve public health.

5. Discussion

Policy development plays a vital role in addressing the issue of traffic-related air pollution, which is crucial for the long-term improvement of air quality in urban areas. Pollution data representation (Figures 20–25) can be an effective tool to quickly grasp the extent of the problem and initiate meaningful interventions. Figure 26 visually depicts the noticeable impact of lockdown measures on air quality improvement in both cities. The population of Islamabad is approximately 15.45% of the population of Lahore. As expected, due to its larger population Lahore exhibits a worse average AQI than Islamabad. When examining the average AQI in the Figure, Lahore exhibits a value of 333.5, while Islamabad demonstrates a lower average AQI of 241. Table 6 displays the lockdown time periods in Pakistan. From Figure 26 it is evident that the implementation of lockdown measures led to an improvement in air quality. Indirectly this can be interpreted as potential benefits in terms of improved AQI if cities adopt green transportation systems. The observed improvement in the AQI during the COVID-19 lockdowns confirms the substantial contribution of the transportation system to emissions. This finding underscores the importance of addressing transportation-related factors to improve air quality. By utilizing our predictive model, policymakers can assess the effectiveness of their initiatives in enhancing the AQI in both cities.

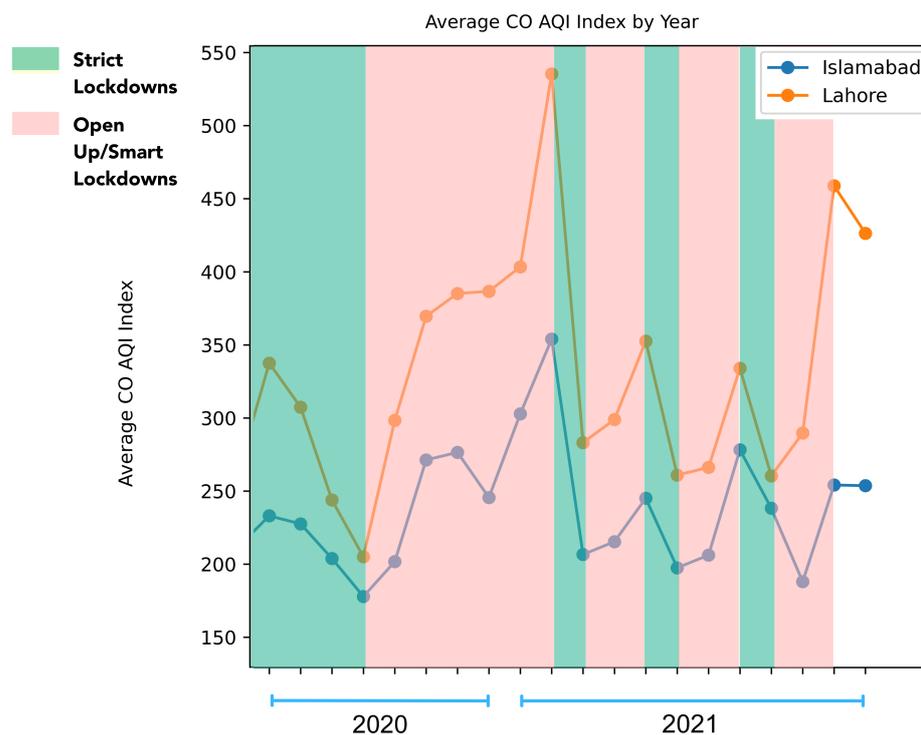


Figure 26. Air quality improvement and decrement analysis during COVID-19 lockdowns and open-ups.

Table 6. Lockdowns Dates in 2020 and 2021.

| Year | Dates | Details |
|------|-------------------------|---|
| 2020 | 23 March–14 April | Nationwide lockdown with the closure of non-essential businesses and public places, ban on public gatherings and flights [70]. |
| 2020 | 16 June–1 July | Smart lockdown in COVID-19 hotspots in seven cities, with restriction of movement and essential services only [71]. |
| 2020 | 20 November–31 December | New restrictions with a ban on indoor weddings and dining, closure of cinemas, theaters, and shrines, and a limit on public gatherings and office staff [70]. |
| 2021 | 15 March–11 April | Spring break for schools in seven cities, and ban on all types of gatherings [70]. |
| 2021 | 8 May–15 May | Complete lockdown in Punjab with the closure of markets, malls, transport, and tourist spots, and only grocery stores, pharmacies, and vaccination centers open [70]. |
| 2021 | 30 July–9 August | Partial lockdown in Sindh due to the Delta variant, affecting intercity transport and flights between Karachi, Lahore, and Islamabad [70]. |
| 2021 | 3 August–12 September | Eased restrictions in 15 districts of Punjab and Islamabad, allowing indoor dining, cinemas, theaters, contact sports, festivals, and full office staff for vaccinated people [72]. |

To formulate effective policies for addressing traffic-related air pollution, several potential approaches can be considered:

1. **Promoting public transportation:** Encouraging the use of public transportation through the development of efficient, accessible, and affordable transit systems can

- help reduce the number of private vehicles on the road, subsequently diminishing traffic-related pollution.
2. **Implementing vehicle emission standards:** Introducing stringent vehicle emission standards can promote cleaner and more efficient technologies in the automotive sector. Regular inspection and maintenance programs can ensure vehicles on the road comply with these regulations.
 3. **Traffic management:** Developing intelligent traffic management systems can help reduce congestion and optimize the flow of traffic, thereby minimizing emissions from idling vehicles.
 4. **Encouraging non-motorized transport:** Investing in infrastructure for cycling and walking can promote healthier and more sustainable modes of transportation, resulting in reduced vehicular emissions.
 5. **Electric vehicle adoption:** Encouraging the adoption of electric vehicles (EVs) through incentives, infrastructure development, and public awareness campaigns can help replace conventional vehicles, leading to a reduction in traffic-generated air pollution.
 6. **Carpooling and ridesharing:** Promoting carpooling and ridesharing options can help reduce the number of single-occupancy vehicles on the road, resulting in lower overall emissions.
 7. **Urban planning and land use policies:** Integrating air quality considerations into urban planning and land use policies can help reduce exposure to traffic-related air pollution by concentrating development around public transportation hubs and creating buffer zones between major roadways and residential areas.
 8. **Transitioning to cleaner energy sources:** Encouraging industries to shift from fossil fuels to cleaner energy sources, such as renewable sources, can significantly reduce air pollution generated by industrial processes.
 9. **Waste management and pollution control:** Implementing effective waste management practices and pollution control technologies can help industries minimize the release of pollutants into the air.
 10. **Industrial zoning and land use policies:** Integrating air quality considerations into industrial zoning and land use policies can help reduce exposure to industrial air pollution by creating buffer zones between industrial facilities and residential areas. By implementing these policy measures and fostering a collaborative approach among government agencies, industry stakeholders, and the general public, traffic-related air pollution can be significantly reduced, leading to improved air quality and public health.

6. Conclusions and Future Work

This research endeavor focused on investigating the potential of green energy transportation systems to significantly enhance air quality in the urban areas of Islamabad and Lahore. To accomplish this, a thorough exploratory data analysis was conducted to assess the feasibility of implementing such systems. Additionally, predictive models were trained and validated to accurately forecast the trends in AQI. Remote sensing data from Sentinel-5P were utilized and machine learning and deep learning models were deployed, such as decision trees, random forests, LSTM, and bidirectional LSTM, to predict pollutant levels. The models exhibited high efficacy, with the trained LSTM model achieving an MSE of 0.50, 0.44, and 0.47 for NO₂, SO₂, and CO, respectively, in Islamabad. The MSE results improved with the trained Bi-LSTM model to 0.41, 0.38, and 0.34 for the same pollutants. In Lahore, the LSTM model produced an MSE of 0.55, 0.66, and 0.34, while the Bi-LSTM model achieved 0.44, 0.61, and 0.26. The findings present substantial evidence that transitioning to green transportation could significantly lessen urban air pollution. Consequently, this underlines the urgent need for a policy shift toward sustainable transportation. The developed predictive models can help policymakers understand the potential impacts of green energy transition efforts on air quality. Nonetheless, it is essential to combine the trained models with other metrics, such as renewable energy usage and specific pollutant

reductions, given the multi-factorial nature of AQI and the varying reliability of predictive models. In the future, the integration of data from various sources will be explored, such as moderate resolution imaging spectroradiometer (MODIS) or cloud–aerosol lidar and infrared pathfinder satellite observation (CALIPSO) satellites, along with existing on-ground monitoring devices. This could generate a more diverse dataset, potentially leading to improved air quality forecasting and a broader understanding of air quality trends. The inclusion of other air pollutants—like ground ozone and particulate matter into predictive models—will further widen the scope of air quality analysis. This comprehensive approach is vital for improving data quality and achieving a holistic understanding of atmospheric conditions. To facilitate this, machine learning models will need to be fine-tuned with a diverse array of parameters that influence atmospheric processes. These models could incorporate features representing influential factors, like El Niño or the Schwabe cycle. Furthermore, the deployment of upscaling or downscaling techniques will play a crucial role in mitigating disparities in spatial resolution among different datasets. Striking a balance between preserving fine-grained details and adjusting resolution will be key in enabling localized predictions. Additionally, developing reporting and monitoring solutions for relevant government bodies and environmental agencies based on the trained models influence decisions around green energy resource management. A geographical expansion of the analysis to other major cities of Pakistan may provide a more holistic view of the country’s air quality dynamics and regional variations. This comprehensive approach will better illustrate the immediate and long-term benefits of transitioning to green energy transportation systems.

Author Contributions: Conceptualization, R.M., A.A. and Z.A.; methodology, R.M., A.A., M.D.A.A. and M.J.B.; software, M.J.B.; validation, R.M., A.A., Z.A., M.A.K. and M.J.B.; investigation, R.M., A.A., Z.A., M.D.A.A. and M.J.B.; writing original draft preparation, R.M., A.A., M.A.K. and M.J.B.; writing—review and editing, R.M., A.A., Z.A., M.A.K. and M.D.A.A.; visualization, R.M., A.A., M.A.K. and M.D.A.A.; supervision, R.M.; data curation, A.A., Z.A. and M.A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data may be requested by reaching out to authors through email.

Acknowledgments: This research work was conducted at the NUST Coventry Internet of Things Lab (NCIL), NUST-SECS, Islamabad, Pakistan, in collaboration with North Dakota State University, USA.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|-----------------|--------------------------------------|
| WHO | World Health Organization |
| NO ₂ | nitrogen dioxide |
| CO | carbon monoxide |
| SO ₂ | sulfur dioxide |
| LSTM | long short-term memory |
| Bi-LSTM | bidirectional long short-term memory |
| RF | random forest |
| DT | decision tree |
| AQI | air quality index |
| MSE | mean square error |
| MAE | mean absolute error |

References

- Power, M.; Cascio, K.; Adgate, A. Air pollution and cardiovascular disease: A window of opportunity. *Curr. Opin. Cardiol.* **2018**, *33*, 578–584.
- Khokhar, M.; Abid, M.; Zahid, M. Air pollution and its impact on agriculture: A review. *Environ. Sci. Pollut. Res.* **2016**, *23*, 1703–1715.
- Chakraborty, P.; Tandon, N.; Bajpai, R. Monitoring of air quality in an urban area of India using lichens. *Environ. Monit. Assess.* **2000**, *64*, 513–525.
- Moretti, M.; Becagli, S.; Cappelletti, F. A multi-disciplinary study of air quality in Florence, Italy. *Atmos. Environ.* **2010**, *44*, 2701–2711.
- Jung, J.; Lee, S.J.; Lee, S. Air Quality Monitoring System Based on Wireless Sensor Network and Its Application to Smart Home. *Sensors* **2019**, *19*, 698.
- Fang, Y.; Chan, C.K. Air Pollution in Mega Cities in China. *Atmos. Environ.* **2015**, *100*, 33–37.
- Ali, M.; Athar, M.A.; Ashfaq, M.; Tariq, M.A. Air Pollution and Human Health: A Review. *Environ. Pollut.* **2015**, *207*, 427–438.
- Khokhar, M.F.; Khokhar, M.I.; Khokhar, M.B.; Ahmad, N.; Akram, M.A.; Iqbal, T.; Javed, M.T. Gaseous and Particulate Pollutants in Ambient Air of Lahore City. *J. Environ. Health Sci. Eng.* **2015**, *13*, 1–8.
- Khokhar, M.B.; Khokhar, M.I.; Ahmad, N.; Khokhar, M.F.; Akram, M.A.; Javed, M.T. Status of Ambient Air Quality in Major Cities of Pakistan: A Review. *Environ. Sci. Pollut. Res.* **2015**, *22*, 18306–18318.
- Shabbir, R.; Siddique, N.A.; Baig, J.A. Air Quality Index of Urban Areas: A Case Study of Lahore, Pakistan. *J. Environ. Public Health* **2015**, 597051.
- World Health Organization. Ambient (Outdoor) Air Quality and Health. Available online: [https://www.who.int/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health) (accessed on 28 May 2023).
- ESoP. *ESoP 2013–2014: Environmental Monitoring Report*; Pakistan Environmental Protection Agency: Islamabad, Pakistan, 2014.
- Ali, S.; Kausar, A.; Malik, R.N. Air quality assessment in Lahore, Pakistan. *Air Qual. Atmos. Health* **2015**, *8*, 11–20.
- Aziz, S.M.; Bajwa, N.; Naeem, A.; Colbeck, I. Modelling carbon monoxide concentrations in Lahore, Pakistan. *Int. J. Environ. Technol. Manag.* **2008**, *8*, 48–59.
- Dawn. 40 pc of 7 m Registered Vehicles in Lahore Contributing to Smog, 2022. Dawn, 15 August 2023. Available online: <https://www.dawn.com/news/1724428/40pc-of-7m-registered-vehicles-in-lahore-contributing-to-smog> (accessed on 22 May 2023).
- Inglesi-Lotz, R.; Dogan, E.; Nel, J.; Tzeremes, P. Connectedness and spillovers in the innovation network of green transportation. *Energy Policy* **2023**, *180*, 113686. [[CrossRef](#)]
- Simoes, V.; Pereira, L.; Dias, A. Enhancing Sustainable Business Models for Green Transportation. *Sustainability* **2023**, *15*, 7272. [[CrossRef](#)]
- Anjum, M.S.; Ali, S.M.; ud din, M.I.; Subhani, M.A.; Anwar, M.N.; Nizami, A.S.; Ashraf, U.; Khokhar, M.F. An Emerged Challenge of Air Pollution and Ever-Increasing Particulate Matter in Pakistan; A Critical Review. *J. Hazard. Mater.* **2021**, *402*, 123943. [[CrossRef](#)] [[PubMed](#)]
- Germanwatch. *Global Climate Risk Index 2012*; Briefing Paper; Germanwatch: Berling, Germany, 2011.
- Derikvand, A.; Taherkhani, A.; Hassanvand, M.S.; Naddafi, K.; Nabizadeh, R.; Shamsipour, M.; Niazi, S.; Heidari, M.; Mokammel, A.; Faridi, S. Indoor Air Quality in the Most Crowded Public Places of Tehran: An Inhalation Health Risk Assessment. *Atmosphere* **2023**, *14*, 1080. [[CrossRef](#)]
- Iqbal, S.; Maqbool, F.; Malik, S.; Awan, M. The Assessment of Air Quality Index of Lahore city for PM10 and NO₂. *J. Biodivers. Environ. Sci.* **2018**, *12*, 74–85.
- Mebrahtu, T.F.; Santorelli, G.; Yang, T.C.; Wright, J.; Tate, J.; McEachan, R.R. The effects of exposure to NO₂, PM_{2.5} and PM₁₀ on health service attendances with respiratory illnesses: A time-series analysis. *Environ. Pollut.* **2023**, *333*, 122123. [[CrossRef](#)]
- Khan, M.F.; Latif, M.T.; Sawalha, M.; Abdellatif, E. The impact of industrial emissions on the air quality of Karachi, Pakistan. *Environ. Monit. Assess.* **2016**, *188*, 676.
- Nundy, S.; Ghosh, A.; Mesloub, A.; Albaqawy, G.A.; Alnaim, M.M. Impact of COVID-19 pandemic on socio-economic, energy-environment and transport sector globally and sustainable development goal (SDG). *J. Clean. Prod.* **2021**, *312*, 127705. [[CrossRef](#)]
- Vichova, K.; Veselik, P.; Heinzova, R.; Dvoracek, R. Road Transport and Its Impact on Air Pollution during the COVID-19 Pandemic. *Sustainability* **2021**, *13*, 11803. [[CrossRef](#)]
- Venter, Z.S.; Aunan, K.; Chowdhury, S.; Lelieveld, J. Air pollution declines during COVID-19 lockdowns mitigate the global health burden. *Environ. Res.* **2021**, *192*, 110403. [[CrossRef](#)]
- Kazemi Garajeh, M.; Laneve, G.; Rezaei, H.; Sadeghnejad, M.; Mohamadzadeh, N.; Salmani, B. Monitoring Trends of CO, NO₂, SO₂, and O₃ Pollutants Using Time-Series Sentinel-5 Images Based on Google Earth Engine. *Pollutants* **2023**, *3*, 255–279. [[CrossRef](#)]
- Zhang, X.; Zhang, X.; Huang, Y.; Chen, J.; Wang, Y. Air pollution monitoring: A review of recent developments. *Environ. Sci. Pollut. Res. Int.* **2021**, *28*, 30594–30609.
- Wang, J.; Lei, Y.; Chen, Y.; Wu, Y.; Ge, X.; Shen, F.; Zhang, J.; Ye, J.; Nie, D.; Zhao, X.; et al. Comparison of active and passive sampling methods for air pollutants in urban environments. *Environ. Sci. Pollut. Res. Int.* **2020**, *27*, 33173–33183.
- U.S. Environmental Protection Agency. *Air Quality Monitoring*; U.S. Environmental Protection Agency: Washington, DC, USA, 2021.
- Wang, W.; Li, J.; Chen, X. Long-path remote sensing of greenhouse gases and air quality: A review. *Atmos. Meas. Tech.* **2020**, *13*, 2705–2734.
- Yang, Y.; Liu, Q. Remote Sensing of Atmospheric Aerosols: Techniques and Applications. *Remote Sens.* **2020**, *12*, 1418.

33. Liu, H.; Li, S.; Li, J.; Li, S.; Li, J. Internet of Things-based air quality monitoring systems: A review. *Environ. Pollut.* **2020**, *265*, 114862. [CrossRef]
34. Lee, H.C.; Lee, J.; Kim, Y.J. Satellite-based remote sensing for air quality monitoring: Recent advances and future prospects. *J. Environ. Manag.* **2021**, *292*, 112782.
35. Jin, J.; Li, X.; Li, L. Advances in satellite remote sensing of environmental variables for epidemiological applications. *Environ. Res.* **2019**, *170*, 186–197.
36. NASA. Ice, Cloud and Land Elevation Satellite-2 (ICESat-2), n.d. Available online: <https://www.nasa.gov/content/goddard/icesat-2> (accessed on 2 March 2023).
37. Siddans, R.; Boesch, H.; Noel, S.; Anand, J.; Bousseres, N.; Bovensmann, H.; Brinksma, E.; Butz, A.; Chimot, J.; Dehn, A.; et al. Validation of the Sentinel-5 Precursor Tropomi level 2 geophysical retrievals. *Atmos. Meas. Tech.* **2019**, *12*, 5263–5350.
38. Liu, Y.; Liu, Y.; Zhang, Q.; Huang, J. Air quality monitoring using deep learning and internet of things. *IEEE Access* **2018**, *6*, 20407–20418.
39. Zhao, Q.; Liu, Y.; Li, X.; Zhang, Y.; He, J. Remote sensing and machine learning techniques for air quality monitoring: A review. *Int. J. Environ. Res. Public Health* **2019**, *16*, 909.
40. Zeng, Q.; Zhou, X.; Wang, J. Air quality monitoring using deep learning and satellite remote sensing. *Environ. Sci. Pollut. Res.* **2020**, *27*, 30250–30264.
41. Lin, G.Y.; Chen, H.W.; Chen, B.J.; Chen, S.C. A machine learning model for predicting PM_{2.5} and nitrate concentrations based on long-term water-soluble inorganic salts datasets at a road site station. *Chemosphere* **2021**, *289*, 133123. [CrossRef] [PubMed]
42. Shafi, J.; Waheed, A. K-Means Clustering Analysing Abrupt Changes in Air Quality. In Proceedings of the 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 5–7 November 2020; pp. 26–30.
43. Choi, S.; Lee, Y.; Kim, E.; Jeong, I.; Kim, H.; Kim, K.H. Mapping urban air quality using mobile sampling with low-cost sensors and machine learning in Seoul, South Korea. *Environ. Pollut.* **2021**, *275*, 116586.
44. Li, R.; Liu, M.; Wang, C.; Dong, L.; Song, Y.; Zhang, J. Assessing the impact of clean air action on air quality trends in Beijing using a machine learning technique. *Atmos. Environ.* **2020**, *231*, 117463.
45. Huang, W.; Lu, Y.; Liu, J.; Ma, Q.; Xu, Q.; Li, R.; Wang, H. Deep learning for predicting air pollutant concentrations using meteorological data and satellite-based remote sensing images. *Remote Sens.* **2020**, *12*, 477.
46. Banerjee, I.; Mondal, I.; Mondal, P.; Saha, I. Pollution, economic growth, and COVID-19 deaths in India: A machine learning evidence. *Environ. Sci. Pollut. Res.* **2021**, *28*, 32789–32799.
47. Cosemans, G.; Janssen, S.; Panis, L.I.; Mishra, V. A Comparison of Linear Regression, Regularization, and Machine Learning Algorithms to Develop Europe-wide Spatial Models of Fine Particles and Nitrogen Dioxide. *Atmosphere* **2021**, *12*, 798.
48. Zhan, Y.; Li, X.; Li, G.; Li, Z. Deep learning for source identification of ambient air pollutants using air quality and meteorological data. *Environ. Sci. Pollut. Res.* **2021**, *28*, 14380–14391.
49. Zhang, J.; Yang, W.; Wang, R.; Zhu, Y.; Liu, Y.; Zheng, Y. Deep Learning-based Air Quality Prediction using Spatiotemporal Correlations. *Int. J. Environ. Res. Public Health* **2019**, *16*, 2510.
50. Zhou, C.; Liu, Y.; Zhang, X.; Li, X. A deep learning-based approach to identify the sources of particulate matter pollution using satellite remote sensing. *Int. J. Environ. Res. Public Health* **2019**, *16*, 1846.
51. Kow, P.Y.; Hsia, I.W.; Chang, L.C.; Chang, F.J. Real-time image-based air quality estimation by deep learning neural networks. *Atmos. Environ.* **2020**, *226*, 117360. [CrossRef] [PubMed]
52. Sharma, E.; Deo, R.C.; Soar, J.; Prasad, R.; Parisi, A.V.; Raj, N. Novel hybrid deep learning model for satellite-based PM₁₀ forecasting in the most polluted Australian hotspots. *Environ. Pollut.* **2021**, *269*, 116198. [CrossRef]
53. Kurnaz, G.; Demir, A.S. Prediction of SO₂ and PM₁₀ air pollutants using a deep learning-based recurrent neural network: Case of industrial city Sakarya. *Environ. Pollut.* **2021**, *272*, 116380. [CrossRef]
54. Mao, W.; Wang, W.; Jiao, L.; Zhao, S.; Liu, A. Modeling air quality prediction using a deep learning approach: Method optimization and evaluation. *Sustain. Cities Soc.* **2021**, *65*, 102567. [CrossRef]
55. Lin, F.; Gao, C.; Yamada, K.D. An Effective Convolutional Neural Network for Visualized Understanding Transboundary Air Pollution Based on Himawari-8 Satellite Images. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]
56. Muthukumar, P.; Nagrecha, K.; Comer, D.; Calvert, C.F.; Amini, N.; Holm, J.; Pourhomayoun, M. PM_{2.5} Air Pollution Prediction through Deep Learning Using Multisource Meteorological, Wildfire, and Heat Data. *Atmosphere* **2022**, *13*, 822. [CrossRef]
57. Das, B.; Dursun, Ö.O.; Toraman, S. Prediction of air pollutants for air quality using deep learning methods in a metropolitan city. *Urban Clim.* **2022**, *46*, 101291. [CrossRef]
58. Middy, A.I.; Roy, S. Pollutant specific optimal deep learning and statistical model building for air quality forecasting. *Environ. Pollut.* **2022**, *301*, 118972. [CrossRef] [PubMed]
59. Shin, S.; Baek, K.; So, H. Rapid monitoring of indoor air quality for efficient HVAC systems using fully convolutional network deep learning model. *Build. Environ.* **2023**, *234*, 110191. [CrossRef]
60. Handhayani, T.; Lewenusa, I.; Herwindiati, D.E.; Hendryli, J. A Comparison of LSTM and BiLSTM for Forecasting the Air Pollution Index and Meteorological Conditions in Jakarta. In Proceedings of the 2022 5th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 8–9 December 2022; pp. 334–339. [CrossRef]

61. Yang, J.; Ismail, A.W. Air Quality Forecasting Using Deep Learning and Transfer Learning: A survey. In Proceedings of the 2022 IEEE Global Conference on Computing, Power and Communication Technologies (GlobConPT), New Delhi, India, 23–25 September 2022; pp. 1–6. [CrossRef]
62. Al Jazeera. Lahore Most Polluted City in Pakistan, Third among Countries: Survey. *Al Jazeera*, 14 March 2023. Available online: <https://www.aljazeera.com/news/2023/3/14/lahore-most-polluted-city-pakistan-third-among-countries-survey> (accessed on 28 May 2023).
63. The Urban Unit of the Planning and Development Department of Punjab. Emission Inventory of Lahore 2023. 2023. Available online: <https://www.urbanunit.gov.pk/Download/publications/Files/8/2023/Emission%20Inventory%20of%20Lahore%202023.pdf> (accessed on 16 June 2023).
64. Wasif Ali, N.; Amir, S.; Iqbal, K.; Shah, A.; Saqib, Z.; Akhtar, N.; Ullah, W.; Tariq, M. Analysis of Land Surface Temperature Dynamics in Islamabad by Using MODIS Remote Sensing Data. *Sustainability* **2022**, *14*, 9894. [CrossRef]
65. Daily Times. Air Pollution High in Capital Amid Soaring Heat, Vehicular Emissions. *Daily Times*, 16 June 2022. Available online: <https://dailytimes.com.pk/953048/air-pollution-high-in-capital-amid-soaring-heat-vehicular-emissions/> (accessed on 28 May 2023).
66. Ministry of Climate Change. *Government of Pakistan*; Ministry of Climate Change: Islamabad, Pakistan, 2021; Chapter 5. Available online: <https://mocc.gov.pk/SiteImage/Misc/files/Chapter-05.pdf> (accessed on 15 June 2023).
67. Graana. Vehicular Emissions Cause Highest Particle Pollution in Islamabad. *Graana Blog*, 17 September 2019. Available online: <https://www.graana.com/blog/vehicular-emissions-cause-highest-particle-pollution-in-islamabad/> (accessed on 15 June 2023).
68. Sharma, A. Decision Tree vs Random Forest Algorithm: Which is the Real Winner? *Analytics Vidhya*, 12 May 2020.
69. AirNow. AQI Basics. Available online: <https://www.airnow.gov/aqi/aqi-basics/> (accessed on 8 March 2023)
70. Wikipedia. COVID-19 Pandemic in Pakistan. 2021. Available online: https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Pakistan (accessed on 16 June 2023)
71. Abbas, W. COV-19: Pakistan Begins Lockdowns; Lahore Epicentre. *Khaleej Times*, 22 March 2021. Available online: <https://www.khaleejtimes.com/coronavirus/covid-19-pakistan-begins-lockdowns-lahore-epicentre> (accessed on 16 June 2023).
72. Geo New. Punjab Eases Lockdown Restrictions in 15 Districts till September 15. *Geo News*, 1 September 2021. Available online: <https://www.geo.tv/latest/368290-punjab-eases-lockdown-restrictions-in-15-districts-till-september-15th> (accessed on 16 June 2023).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.