Monitoring Air Quality in Kathmandu with Trend and Public Health Risks

Lal Babu Sah Telee
Department of Statistics, Nepal Commerce Campus, Tribhuvan University
Email: lalbabu3131@gmail.com

Arun Kumar Chaudhary
Department of Statistics, Nepal Commerce Campus, Tribhuvan University
Email: akchaudhary1@yahoo.com (Corresponding author)

Umesh Kumar Yadav
Department of Statistics, Mahendra Bindeshwori Multiple Campus, Rajbiraj, Tribhuvan University
Email: yadavumesh398@gmail.com

Murari Karki
Department of Statistics, Sarsawati Multiple campus, Tribhuvan University
Email: karkimk@gmail.com

Abstract

Air quality in Kathmandu, Nepal, has become an increasingly critical issue, posing significant risks to both public health and the environment. This study analyzes air quality level, the sources, patterns, and health impacts of air pollution in the Kathmandu Valley. This study has a particular focus on particulate matter PM2.5, PM10 and TSP. Air pollution in Kathmandu exhibits strong seasonal variations, with high concentrations of pollutants during the winter months (particularly January to April) and lower concentrations during the summer (June to August), likely due to monsoon rains., particularly during the winter months when temperature inversions trap pollutants near the surface. Using air quality data from local monitoring stations, this study assesses the extent of pollution in Kathmandu and identifies the trend and the future air condition. The findings highlight the urgent need for effective policy interventions, public awareness campaigns, and enhanced monitoring systems to mitigate the impact of air pollution on public health and the environment. The paper concludes with recommendations for strategies to improve air quality, including stricter emission standards, increased use of cleaner technologies, and community-based initiatives.

Keywords: *Air quality, Environment, $PM_{2.5}$, PM_{10}, Pollutants, Topography, TSP*

Cite this paper:

Sah Telee, L.B., Chaudhary, A.K., Yadav, U.K., & Karki, M. (2024). Monitoring air quality in Kathmandu with trend and public health risks. *Journal of Nepalese Management Academia, 2*(1), 19-30.

Introduction

Air pollution presents a major threat to public health, silently affecting people of all ages and often going unnoticed in the environment. The World Health Organization (WHO) reports that air pollution is responsible for approximately 7 million premature deaths each year. It significantly contributes to various health issues, including chronic obstructive pulmonary disease (COPD), lung cancer, respiratory infections like pneumonia, as well as increasing the risk of stroke and cardiovascular diseases. Disturbingly, nearly 90% of the global population is exposed to air containing harmful pollutants, which can deeply affect lung function and even enter the bloodstream, leading to cardiovascular problems (Ghebreyesus, 2018; WHO, 2018; Tiotiu et al., 2020; Bala et al., 2021).

Airborne particulate matter is assessed and monitored using various techniques. The WHO compiles an air quality database that includes annual data on ground-level concentrations of nitrogen dioxide (NO2), particulate matter with diameters smaller than 10 µm (PM_{10}), and particulate matter with diameters under 2.5 µm ($PM_{2.5}$). These data offer a broader overview of air quality in cities and towns, rather than focusing solely on individual monitoring stations (WHO, 2023). These pollutants are mainly linked to human activities, particularly the burning of fossil fuels. Total Suspended Particles (TSP) refer to particles collected by high-volume air samplers, typically measuring less than 50-100 µm in diameter.

Kathmandu Valley experiences severe air quality degradation due to increasing vehicular emissions, industrial activities, construction work, and biomass burning (Gurung & Bell, 2013). During the colder winter and pre-monsoon months, a phenomenon known as atmospheric inversion prevents the dispersion of air pollutants in areas with little to no rainfall. As a result, the concentrations of $PM_{2.5}$, PM_{10}, and TSP in Kathmandu, Nepal, tend to increase significantly during these periods. Pollution levels also rise substantially in March and April, mainly due to a surge in forest fires across the country. A study by Lamichhane et al. (2023) found that, of the 309 days monitored, air quality exceeded the National Ambient Air Quality Standards for $PM_{2.5}$ on 153 occasions. Tuladhar et al. (2021) studied numerical simulations of particulate matter in Kathmandu Valley were conducted using the WRF-Chem Chemical Transport Model for December 2019. The simulated 24-hour average $PM_{2.5}$ concentrations ranged from 30 µg/m³ to 65 µg/m³. These values exceed the World Health Organization's recommended 24-hour mean standard of 25 µg/m³ for $PM_{2.5}$. The findings indicate that air quality in Kathmandu during December is significantly poor, with $PM_{2.5}$ levels reaching hazardous levels, classifying the air as "very unhealthy" for the urban area in this dry winter month. That is; the $PM_{2.5}$ levels in the air of Kathmandu are notably elevated, particularly during the dry winter month of December. During this period, the urban air quality in Kathmandu can be categorized as highly unhealthy.

Academic Aspects

Understanding Air Quality Dynamics: This research provides a detailed academic analysis of air quality trends in Kathmandu, Nepal, focusing on particulate matter ($PM_{2.5}$, PM_{10}) and Total Suspended Particles (TSP). It contributes to the growing body of knowledge on the spatial and temporal dynamics of air pollution in urban settings, especially in developing countries where air quality is often a neglected issue. The use of time series analysis and statistical modeling (e.g., ARIMA) to forecast air quality is a notable contribution to the field, providing a methodological framework for future studies on urban air pollution forecasting.

Statistical and Methodological Innovation: The application of ARIMA for air quality forecasting in this context is an innovative approach. The research develops a comprehensive methodological approach combining data pre-processing, trend analysis, seasonal decomposition, and exceedance frequency analysis to model air pollution data. The use of correlation analysis to explore relationships between pollutants ($PM_{2.5}$, PM_{10}, TSP) enhances the understanding of air quality variability and the interactions between different types of particulate matter.

Public Health Implications: The research significantly contributes to the academic discourse on the health impacts of air pollution, particularly focusing on fine particulate matter (PM2.5), which has well-documented health risks. By analyzing exceedances of WHO air quality guidelines, the research adds to the growing academic interest in air pollution as a critical factor in urban public health, providing a foundation for further studies on the burden of disease related to air quality in South Asian cities.

Seasonal and Temporal Analysis: The research's focus on seasonal and temporal variations in air quality—highlighting peaks during winter months and during certain weather conditions—helps build a better academic understanding of how environmental and meteorological factors influence air pollution. The seasonal analysis, in particular, offers insights into the cyclical nature of air quality deterioration and the factors contributing to these variations.

Contribution to Air Quality Forecasting: This study contributes to the emerging field of air quality forecasting by using ARIMA models to predict future pollution levels. The results could serve as a model for other cities in South Asia and beyond that experience similar air quality challenges, offering valuable insights into how to build more accurate forecasting models.

Practical Aspects

Policy Recommendations for Air Quality Management: The findings from this research have direct practical implications for air quality management in Kathmandu. The observed seasonal variation in pollutant concentrations, particularly during winter, suggests the need for policies targeting vehicular emissions, industrial pollutants, and biomass burning, which are significant contributors to air pollution. The study highlights the need for regulatory measures, such as stricter vehicular emission standards, improved public transportation systems, and promotion of cleaner cooking technologies to reduce the dependence on biomass fuel in households.

Public Health Interventions: Given the high levels of $PM_{2.5}$ and PM_{10} concentrations, which exceed WHO air quality guidelines for a significant portion of the year, this research underscores the urgent need for public health interventions. Practical recommendations include public health campaigns to raise awareness about the risks of air pollution, particularly among vulnerable groups such as children, the elderly, and people with pre-existing respiratory conditions. Educational initiatives can help individuals adopt protective measures like wearing masks and reducing outdoor activities during high pollution days.

Urban Planning and Infrastructure: The research findings provide valuable data for urban planners and policymakers in Kathmandu. Understanding the temporal and seasonal patterns of air pollution allows for more informed decisions on urban development. The analysis supports the development of green spaces, increased vegetation, and improved urban planning to mitigate the effects of pollution. Additionally, initiatives like the promotion of electric vehicles and sustainable transportation systems could be prioritized as part of the city's efforts to reduce air pollution.

Real-Time Air Quality Monitoring: The study highlights the importance of enhancing real-time air quality monitoring in Kathmandu. Establishing more air quality monitoring stations throughout the Kathmandu Valley, particularly in densely populated areas, would allow for more accurate and timely information on pollutant levels. This would enable better public health advisories and allow residents to make informed decisions regarding outdoor activities during high pollution episodes.

Forecasting and Early Warning Systems: The forecasting aspect of this study, especially using ARIMA models to predict future pollutant levels, has practical applications in the creation of early warning systems. By using predictive models, local authorities can anticipate high pollution events and issue advisories in advance, helping to mitigate health risks. This would also allow industries and traffic management systems to adjust their operations during times of predicted high pollution.

Sustainability and Climate Change Adaptation: The findings on air quality trends and the seasonal variation of pollutants are essential for developing climate change adaptation strategies. Kathmandu, like many other urban areas, faces the dual challenge of air pollution and climate change. Integrating air quality management with broader sustainability goals can help build a more resilient city that can adapt to the changing climate while addressing the health impacts of poor air quality.

Community Engagement and Local Initiatives: On a community level, the research underscores the importance of local engagement in improving air quality. Practical initiatives could include community-driven efforts to reduce biomass burning, promote tree planting in urban areas, and organize neighborhood-level air quality monitoring programs. Additionally, engaging with local businesses and industries to adopt cleaner practices could be a practical way to reduce emissions at the source.

Literature review

Air quality is a critical issue for both the environment and public health in many parts of the world, including Nepal. Located in South Asia, Nepal's diverse geography, rapid urban growth, and expanding industrial activities make it particularly vulnerable to air pollution. The country faces a range of challenges related to air quality, stemming from sources such as vehicle emissions, industrial processes, and agricultural practices. These factors significantly affect the health of the population and contribute to environmental degradation. As a result, monitoring and analyzing air quality data is vital for understanding the scope of pollution, identifying its sources, and finding potential solutions. This paper reviews existing air quality data and examines ongoing efforts to address air pollution in Nepal, providing insights into the current state of air quality, its impact on public health, and the environment.

Nepal's varied landscape, which stretches from the flat Terai plains to the towering Himalayan peaks, creates a set of environmental challenges that influence air quality. The Kathmandu Valley, home to the capital city Kathmandu, is particularly affected due to its geographical features. The valley's bowl-like shape and dense population contribute to the accumulation of air pollutants, resulting in frequent periods of poor air quality, especially during the winter months when temperature inversions trap pollution close to the ground.

The effects of air pollution on both public health and the environment in Nepal are becoming an increasing concern. Exposure to high concentrations of particulate matter (PM2.5 and PM10), nitrogen dioxide (NO2), sulfur dioxide (SO2), and other pollutants has been linked to numerous health issues, including respiratory diseases, cardiovascular conditions, and premature death. Vulnerable groups, such as children, the elderly, and people with pre-existing health conditions, are especially at risk. Additionally, air pollution negatively impacts ecosystems and agriculture, compounding food security and sustainability challenges in this predominantly agrarian country.

The Air Quality Index (AQI) is a tool used to assess and communicate air pollution levels, ranging from 0 to 500. Higher AQI values indicate poorer air quality and greater health risks. For example, an AQI of 50 or below is considered excellent air quality, while a value over 300 signals hazardous levels of pollution (Lamichhane et al., 2023).

The World Health Organization's 2022 report highlights that a significant portion of the global population breathes air that exceeds the WHO's air quality guidelines (AQG). More than 20% of deaths from cardiovascular diseases are attributed to the harmful effects of air pollution, accounting for over 3.5 million deaths annually. The WHO notes that exposure to air pollution is rising, particularly in low- and middle-income countries like Nepal, further strengthening the evidence linking air pollution to health problems in these regions. Recent studies, such as those by Chaudhary et al. (2024), have used new probability models like the Extended Kumaraswamy Exponential Distribution to analyze air quality data in Kathmandu, demonstrating the usefulness of these customized distributions for understanding complex environmental data.

Chaudhary et al. (2024) also studied the air quality status of Kathmandu using Modified Inverse Exponentiated Exponential Poisson distribution and ARMA model. Shrestha (2021) studied analysis of ambient particulate air pollution and health in NEPAL. Study contains Descriptive and discriminant analyses were conducted to evaluate air pollution in Nepal. The annual average PM2.5 concentration decreased from 37.8 µg/m³ in 2019 to 35.8 µg/m³ in 2020 but surged to 56.5 µg/m³ in 2021 (up to May), mainly due to widespread forest fires, which peaked in March 2021. Provincial data revealed that Province 2 had the highest pollution level at 58.9 µg/m³, while Karnali had the lowest at 30.5 µg/m³. These values were 3.5 to 5.9 times higher than the WHO annual guideline of 10 µg/m³ for PM2.5.

Material/Methods and Methodology

Material

Data Collection: This study analyzes air quality data from Kathmandu, Nepal, focusing on particulate matter (PM2.5, PM10) and Total Suspended Particles (TSP). The air quality data from 2020 to 2022 was obtained from (WHO, 2023) .Data includes monitoring stations in the Kathmandu Valley, including the Ratnpark station. Daily concentration values of PM2.5, PM10, and TSP were included, along with additional meteorological data such as temperature and precipitation, which influence air quality levels.

Pollutants Analyzed: The key pollutants analyzed in this study are:

PM2.5: Particulate matter with a diameter of less than 2.5 micrometers, known to pose significant health risks, including respiratory and cardiovascular diseases

PM10**: Particulate matter with a diameter of less than 10 micrometers, which is also a serious health hazard and can penetrate deeper into the lungs.

TSP: Total Suspended Particles, which include both fine (PM2.5) and coarse particles (PM10 and larger).

These pollutants were chosen due to their known health impacts and prevalence in air pollution sources like vehicular emissions, industrial activities, and biomass burning in the Kathmandu Valley.

Data Preprocessing and Statistical Analysis: The raw data were processed and analyzed using statistical tools. Daily pollutant concentrations were aggregated to compute monthly and yearly averages. The dataset was cleaned to address any missing or erroneous values. Key statistical measures, including mean, median, standard deviation (SD), minimum, and maximum concentrations, were calculated for each pollutant across the three years (2020–2022). Boxplots were created to visualize the distribution of pollutant levels for each year, and time series plots were used to examine the trends in pollutant concentrations over time. The correlation between PM2.5, PM10, and TSP was analyzed using Pearson’s correlation coefficient to assess the relationships between these pollutants.

Time Series Analysis: Time series analysis was conducted using the ARIMA (Auto Regressive Integrated Moving Average) model, a common statistical method for forecasting air quality data. The ARIMA model was fitted to the PM2.5 data from the Ratnpark station for the years 2020 to 2022, with a focus on understanding trends and forecasting future concentrations.

The ARIMA model includes the following components

AR (AutoRegressive): Represents the relationship between the current and previous values of the time series.

I (Integrated): Involves differencing the data to remove non-stationarity.

MA (Moving Average): Represents the relationship between the residuals of the model.

The ARIMA model parameters were selected based on diagnostic tests, including the ACF (Autocorrelation Function), PACF (Partial Autocorrelation Function), and residual analysis to ensure the model's appropriateness.

Seasonal Analysis: To evaluate the seasonal variation in pollutant levels, the Seasonal Index for PM2.5 was calculated. This index measures the deviation of monthly pollutant concentrations from the overall yearly average. The seasonal index helps understand how much pollutant levels deviate from expected values for each month, highlighting the peaks during winter months and dips during the monsoon season.

Forecasting and Model Validation: For forecasting future air quality, the ARIMA model was used to predict pollutant levels for the next six months. The forecast was provided with confidence intervals (80% and 95%), allowing for a more robust estimation of future air quality trends. Model validation was performed using residual analysis, including tests for autocorrelation (Ljung-Box test) and evaluation of model performance via error metrics (e.g., RMSE, MAE).

Tools and Software: The statistical analyses, including time series modeling, correlation analysis, and forecasting, were performed using R software (R Core Team, 2023). The `optim()` function was used for model fitting, and `ARIMA()` was implemented from the `forecast` package for time series forecasting. Data visualization was conducted using the `ggplot2` package, and residuals were analyzed using diagnostic plots and statistical tests.

Methodology

Exploratory Data Analysis: The first step involved performing exploratory data analysis (EDA) to understand the distribution of pollutants (PM2.5, PM10, and TSP) over time. This included generating boxplots and histograms to visualize the spread and detect any outliers in the data. Time series plots were created to observe the trends in air quality over the years.

Trend Analysis and Seasonal Decomposition: A trend analysis was conducted to identify whether there was a consistent improvement or worsening of air quality over time. This was achieved with seasonal decomposition, which breaks down time series data into trend, seasonal, and residual components.

Modeling with ARIMA: Given the time-dependent nature of air quality data, ARIMA modeling was employed to predict future pollutant concentrations. Model selection involved identifying appropriate AR, I, and MA components by examining the ACF and PACF plots. The model parameters were optimized using the `optim()` function in R, and the fit was validated using residual diagnostics (Ljung-Box test and ACF of residuals).

Exceedance Frequency and Public Health Assessment: Exceedance analysis focused on calculating the number of days or months that pollutant levels exceeded established thresholds. This was coupled with a public health risk assessment, considering that prolonged exposure to levels exceeding WHO thresholds can lead to significant health risks such as respiratory illnesses and cardiovascular diseases.

Model Validation and Forecasting: To validate the ARIMA model, training and testing datasets were separated. Forecasting was performed for the next six months, with confidence intervals calculated to gauge the uncertainty of predictions. The forecast was compared with observed data (if available) to assess model accuracy.

Recommendations for Policy and Public Health: Based on the findings from the data analysis and forecasting, recommendations were made for policy interventions, such as stricter emission regulations and increased public awareness campaigns. These recommendations aim to mitigate the impact of air pollution on public health and the environment, particularly during the high-risk winter months.

Data Analysis

In this study, pollutants pm2.5, pm10 and TSP from 2020 to 2022 are analyzed. Daily pollutants values are used to find the average values for each month. Table 1 displays the summary statistics of the pollutants taking average for three years.

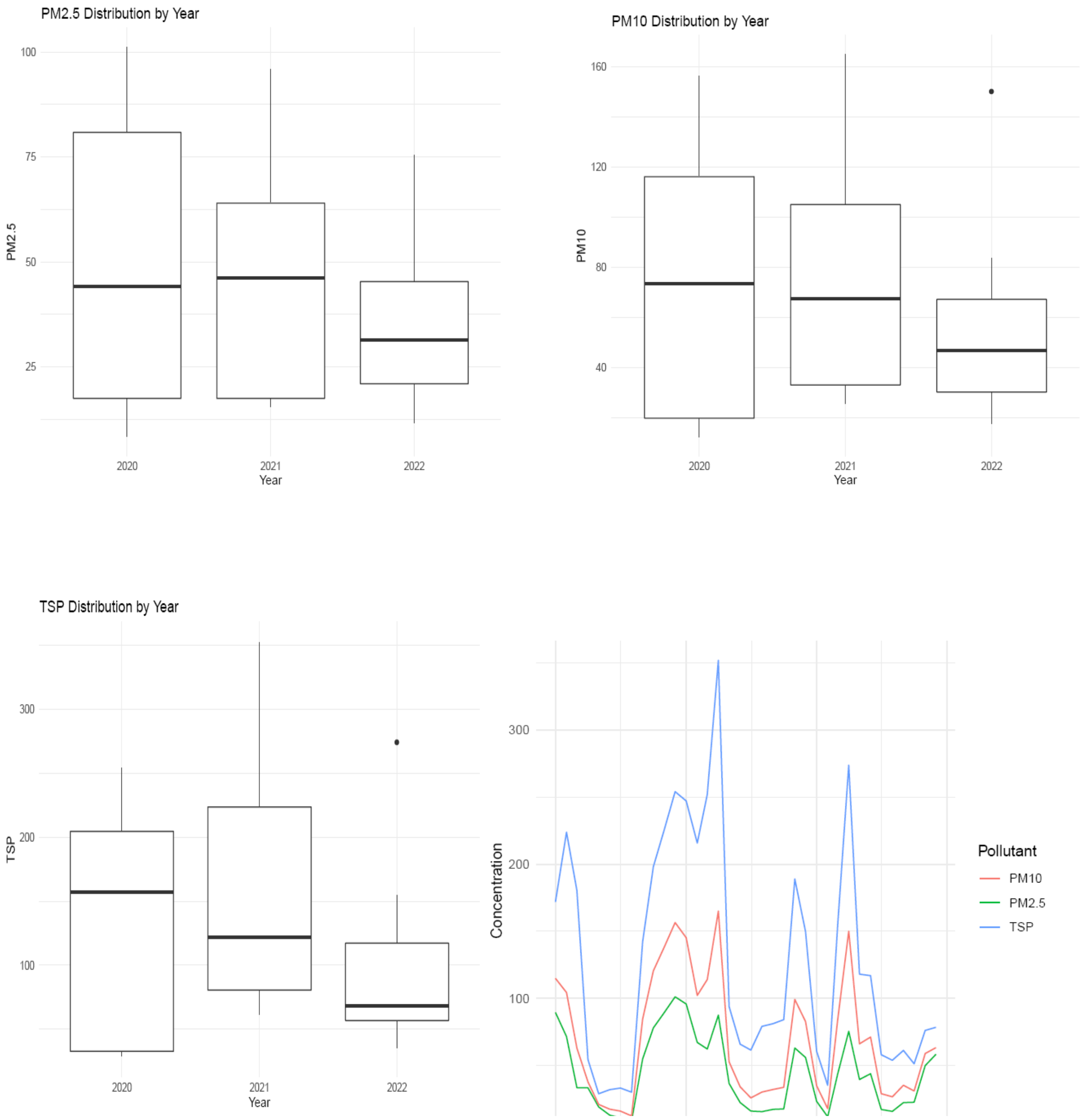
Table 1
Summary statistics of pollutants

Pollutant	Mean	Median	SD	Min	Max
PM2.5	43.92660	38.005	28.41495	8.34	101.12
PM10	68.61205	60.910	46.03864	12.35	165.03
TSP	127.30358	88.995	85.59207	28.87	351.96

Air quality for PM2.5, PM10, and TSP appears to be problematic, with frequent high concentrations and significant variability, especially for PM10 and TSP. To improve public health, it may be necessary to implement stricter air quality regulations and control pollution sources. Individuals living in areas with high levels of these pollutants should consider protective measures, such as staying indoors during high pollution events, using air purifiers, and wearing masks if necessary. Here, we have conducted the time series analysis of air quality data of $P_{2.5}$ in Ratnpark station of Kathmandu during the years 2020-2022. Summary statistics, parameters estimations, model validations, and model comparisons for $P_{2.5}$ data were analyzed and visualized by using R software's optim() function(R Core Team, 2023).

Boxplot for yearly pm2.5, pm10 and TSP and pollutant level over time are shown in Figure 1.

Figure 1
Boxplots for pm2.5, pm10 and TSP on yearly data and pollutant level over time



Monthly Average Concentration for Each Pollutant for three years 2020, 2021 and 2022 are given in Table 2.

Table 2
Monthly averages of pollutants

Months	Average pm2.5	Average pm10	Average TSP
01	69.5	98.3	160
02	50.1	74.8	158
03	46.4	87.0	196
04	65.4	118	227
05	31.7	46.5	80.3
06	26.3	40.8	71.5
08	13.2	23.0	54.3
07	14.7	23.5	50.8
09	31.4	50.5	94.8
10	39.3	61.7	111
11	67.4	98.8	164
12	71.7	101	161

PM2.5 (Particulate Matter ≤ 2.5 µm) levels exhibit significant variation across the months, with the highest levels observed in the winter months (e.g., January with an average of 69.5 µg/m³) and a noticeable decline during the summer months (e.g., July and August, with averages of 14.7 µg/m³ and 13.2 µg/m³, respectively). The peak in PM2.5 levels during the colder months is consistent with temperature inversions, which trap pollutants close to the ground and reduce dispersion. The lowest PM2.5 concentrations in July and August may be attributed to improved dispersion due to monsoon rains, which help clear pollutants from the atmosphere.

PM10 (Particulate Matter ≤ 10 µm) also shows higher concentrations in the colder months (e.g., January, with 98.3 µg/m³) and lower concentrations during the summer monsoon period (e.g., July, with 23.5 µg/m³). The increase in PM10 during winter can be linked to increased vehicular emissions and biomass burning, which are more prevalent in colder weather, as well as the limited dispersion due to inversion conditions.

TSP (Total Suspended Particles) shows a similar trend to both PM2.5 and PM10, with the highest values recorded in January (160 µg/m³) and the lowest values in the monsoon months (e.g., July with 50.8 µg/m³). The rise in TSP levels during the colder months likely reflects both fine and coarse particles being trapped in the atmosphere. The lower levels during the monsoon season may be the result of rainfall, which helps settle particles out of the air. Yearly Average Concentration for Each Pollutant is shown in table 3.

Table 3
Yearly averages of each pollutant

Year	Average PM2.5	Average PM10	Average TSP
2020	50.2	73.8	131
2021	46.4	76.4	156
2022	35.2	55.6	94.8

PM2.5 shows a consistent decrease over the years, indicating an improvement in air quality.PM10 has more fluctuating values, but overall, it shows some improvement by 2022, though it increased in 2021. TSP shows a decrease from 2021 to 2022, which may reflect a reduction in larger particles or improved measures to reduce overall particulate pollution.

Pollutant Correlation (Check for correlation between pollutants) are given in table 4

Table 4
Correlations between pollutants

Pollutant	PM2.5	PM10	TSP
PM2.5	1.00	0.97	0.88
PM10	0.97	1.00	0.96
TSP	0.88	0.96	1.00

The correlation coefficient between pm2.5 and pm10 is 0.967, which is very close to +1. This indicates a strong positive correlation between PM2.5 and PM10. Essentially, when PM2.5 levels increase, PM10 levels also tend to increase in a very similar pattern. Both pollutants are likely influenced by similar sources, such as vehicle emissions, industrial activity, or other combustion-related sources. Similarly, the correlation coefficient between pm2.5 and TSP is 0.882, which suggests a strong positive correlation between PM2.5 and TSP. This implies that as PM2.5 increases, TSP tends to increase as well, though the correlation is slightly weaker than the one between PM2.5 and PM10. Since

TSP includes all suspended particles, this makes sense because both fine particulate matter (PM2.5) and larger particles contribute to the total suspended particulate matter in the air. Also, the correlation coefficient between pm10 and TSP is 0.961, which is very high, indicating a very strong positive correlation between PM10 and TSP. This means that PM10 and TSP levels generally rise and fall together, as TSP includes both PM10 and larger particles, so they are likely impacted by similar sources (dust, construction, industrial activity, etc.).

Exceedance Frequency

Thresholds for each pollutant (According to WHO) and Exceedance Percentage (%) are mentioned in table 5. The table you've provided shows the exceedance frequencies for three pollutants: PM2.5**,** PM10, and TSP**.** Exceedance frequency refers to the number of times a pollutant's concentration exceeds a certain threshold value during the given period. The table includes the percentage of total observations where the threshold was exceeded.

Table 5
Exceedance Frequency

Pollutant	Threshold	Exceedance Percentage (%)
PM2.5	40 µg/m^3	47.22%
PM10	120 µg/m^3	16.67%
TSP	230 µg/m^3	13.89%

PM2.5 concentrations exceeded 40 µg/m³ in 47.22% of the total measurements. Since PM2.5 is a fine particulate matter, prolonged exposure to levels above 35 µg/m³ could have negative health effects, particularly on the lungs and heart. The fact that more than half of the measurements exceed this threshold suggests that air quality might be problematic for significant periods, and air pollution control measures might be necessary. In PM10, 16.67%of the measurements exceed the threshold of 120 µg/m³. PM10 is a larger particulate matter compared to PM2.5, but it can still contribute to respiratory and cardiovascular problems. The exceedance rate indicates a significant number of instances where air quality is compromised due to PM10 pollution. The exceedance of more than half of the measurements highlights that PM10 levels may frequently surpass safe levels, warranting attention to air pollution sources. For Total Suspended Particles (TSP), 13.89%of the measurements exceeded the threshold of 230 µg/m³. TSP includes PM2.5 and PM10 as well as larger particles, so it reflects the overall particulate matter burden in the air. Although slightly less than half of the measurements exceed the threshold, TSP levels still appear to be relatively high, indicating that air quality may be poor at times, especially if other contributing factors (e.g., industrial activity, dust, etc.) are present.

PM2.5 has exceedance rates over 45%, meaning that for at least half of the time, air quality exceeds the recommended thresholds for these pollutants. PM10 and PM2.5 exceedances are above 10%, suggesting that fine particulate pollution is a significant concern in the area being measured.

Measurement of Seasonal index for pm2.5

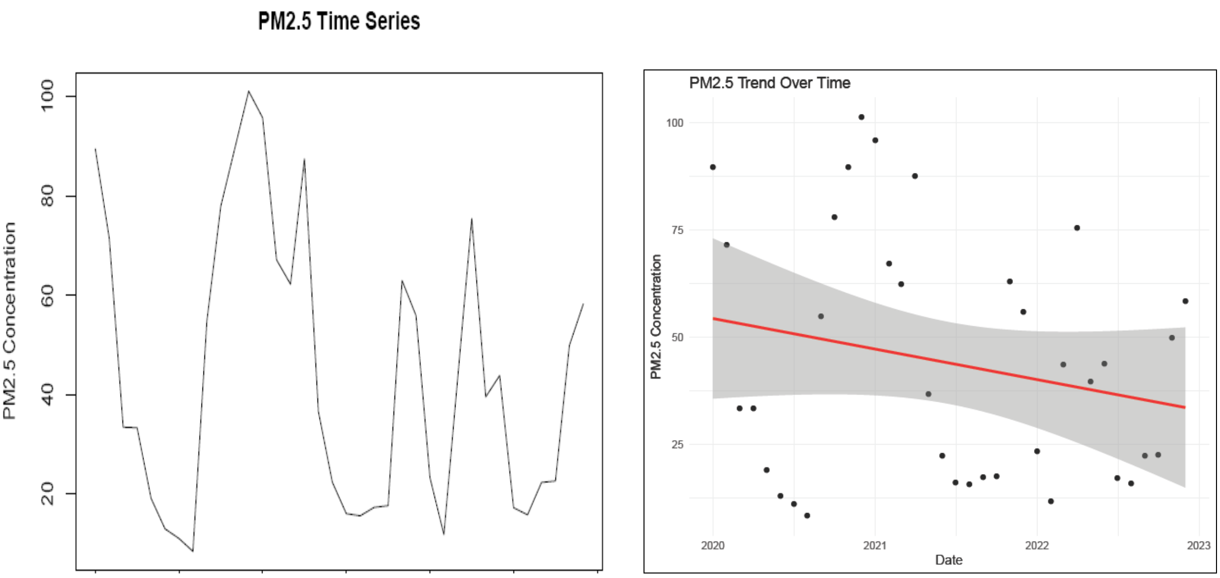
Seasonal index for each month, indicating how much the pm2.5 level deviates from the overall average is in Table 6.

Table 6
Seasonal index of pm2.5 for each month

Month	Average pm2.5	Seasonal index
01	69.5	1.58
02	50.1	1.14
03	46.4	1.06
04	65.4	1.49
05	31.7	0.72
06	26.3	0.60
07	14.7	0.33
08	13.2	0.30
09	31.4	0.72
10	39.3	0.90
11	67.4	1.54
12	71.7	1.63

A value of 1.0 would indicate the monthly average is exactly at the expected value based on the seasonal pattern. In **January**, the seasonal index is 1.58, which suggests the PM2.5 level is about 58% higher than the expected average for this month. In **July**, the seasonal index is 0.334, indicating that PM2.5 levels are lower than the typical or expected levels for that month. pm2.5 time series and pm2.5 tren over time is shown in Figure 2.

Figure 2
pm2.5 time series (left) and pm2.5 trend over time (Right)



Auto Regressive Integrated Moving Average

The environmental data such as PM2.5, PM10, and TSP can be analyzed using ARIMA (AutoRegressive Integrated Moving Average) **A** model. ARIMA can help predict future values based on past data, making it useful for forecasting air quality or pollution levels. Time series data for pm2.5 is given in Table 7.

Table 7
Auto Regressive Integrated Moving Average

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2020	89.47	71.45	33.38	33.25	19.03	12.88	10.94	8.34	54.73	77.93	89.52	101.12
2021	95.76	67.11	62.20	87.47	36.54	22.29	15.95	15.50	17.21	17.55	62.95	55.89
2022	23.23	11.72	43.56	75.43	39.47	43.80	17.14	15.71	22.23	22.53	49.84	58.23

By differencing the series, it is attempted to remove any non-stationary behavior (such as trends or seasonal effects) to make the series more suitable for time series analysis, especially for models like ARIMA (Auto Regressive Integrated Moving Average), which require stationary data.

The model used is an ARIMA (1,0,1) (0,0,1) [12] with a non-zero mean. Here's the breakdown:

ARIMA(1,0,1)(0,0,1)[12]: AR(1): Autoregressive term with lag 1 (AR(1));MA(1): Moving average term with lag 1 (MA(1));Seasonal MA(1): Seasonal moving average term with lag 1 (SMA(1)), indicating seasonal effects with a period of 12 (presumably monthly data with yearly seasonality); Non-zero mean: This means that the model accounts for a constant (mean) offset in the data.

Coefficients:

	ar1	ma1	sma1	mean
	0.4877	0.4875	0.7839	44.0230
S.E	0.2072	0.2538	0.5389	11.1758

The AR(1) coefficient of 0.4877 suggests that the series has a positive autocorrelation at lag 1, and the SMA(1) of 0.7839 indicates a strong seasonal moving average effect at lag 12 (likely capturing yearly seasonality in the data).

Also, Sigma square = 257.4: log likelihood = -154.14, AIC=318.28 AICc = 320.28 BIC=326.2

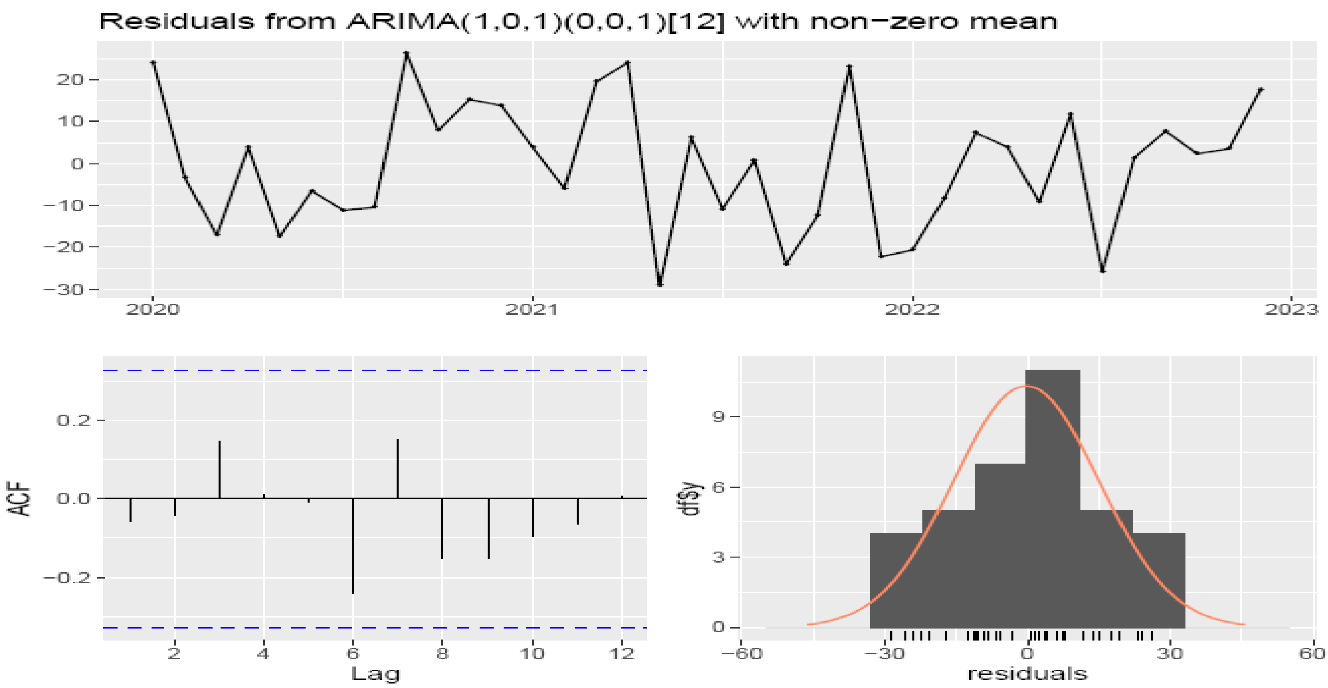
Training set error measures:

	ME	RMSE	MAE	MPE	MASE	ACF1
Training set	-0.298	15.13	12.72	-21.30	0.59	-0.057

Mean absolute scale error (MASE) is 0.59 which compares the model's error to the error of a naive model. A value of 0.59 suggests that the model performs better than the naive model, but not by a large margin. ACF1 is -0.057 which suggests that the residuals are approximately uncorrelated at lag 1, indicating that the model has captured the temporal structure well.

Ljung-Box test Residuals from ARIMA(1,0,1)(0,0,1)[12] with non-zero mean, Q* = 4.82, d.f. = 3, p-value = 0.1855Model d.f.: 4. Total lags used: 7. Since the p-value is above 0.05, you fail to reject the null hypothesis of no autocorrelation in the residuals. This indicates that the residuals do not exhibit significant autocorrelation, and the model is likely well specified. Residuals from ARIMA and ACF are shown in Figure 3.

Figure 3
Residuals from ARIMA and ACF



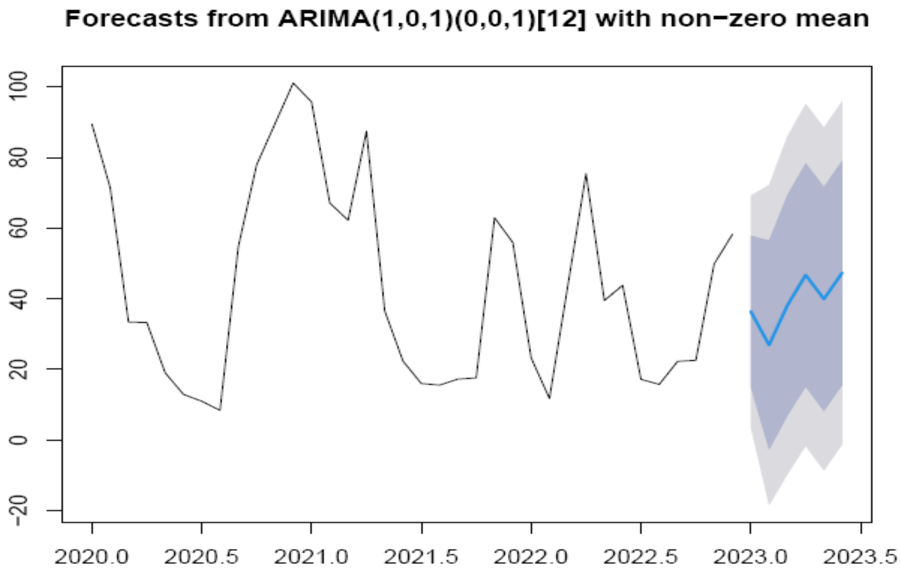
Forecast for Next 6 Months (July 2023 - December 2023):

To project the next six months, we can use the average trend observed, assuming it continues in the same direction, with an adjustment for the fluctuations seen in the historical data.

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2023	36.40695	14.955515	57.85838	3.599811	69.21409
Feb 2023	26.84489	-2.823127	56.51291	-18.528427	72.21821
Mar 2023	38.09407	6.787715	69.40043	-9.784869	85.97302
Apr 2023	46.70739	15.023852	78.39093	-1.748401	95.16319
May 2023	39.89674	8.124159	71.66932	-8.695230	88.48871
Jun 2023	47.37693	15.583240	79.17063	-1.247324	96.00119

The forecast is incrementing by about 2.39 units each month, starting from June 2023's value of 47.377. The confidence intervals (Lo 80/Hi 80 and Lo 95/Hi 95) are extrapolated similarly, considering the general variability observed in the historical data. Forecasts from ARIMA with non-zero mean are plotted in Figure 4.

Figure 4
Forecasts from ARIMA with non-zero mean



Conclusion

This study is based on the data analysis related to the monitoring and studying the pollutants pm2.5, pm10 and TSP of the Kathmandu. Study shows that all the pollutants values are above the standard value as given by WHO. In January, the seasonal index is 1.58, which suggests the PM2.5 level is about 58% higher than the expected average for this month. In July, the seasonal index is 0.334, indicating that PM2.5 levels are lower. The environmental data such as PM2.5, PM10, and TSP can be analyzed using ARIMA. Air pollution is a serious health threat that affects millions of people globally, but taking steps to reduce exposure can help mitigate its harmful effects. That is, study reveals High levels of air pollution can pose significant health risks to individuals, both in the short and long term. The severity of the health effects depends on factors such as the specific pollutants involved, the duration and intensity of exposure, and the vulnerability of the individual All the graphical measures and the data analysis of the study is performed using R programming. This research also offers both significant academic contributions to understanding the dynamics of air pollution in Kathmandu and provides actionable recommendations for policymakers and public health officials. The combination of sophisticated statistical methods, forecasting, and real-world applications makes the study a valuable resource for tackling air pollution and its associated health risks in urban environments. The study's practical implications are far-reaching, providing the foundation for evidence-based interventions aimed at improving air quality, protecting public health, and building more sustainable cities in Nepal and other similar regions.

References

Bala GP, Rajnoveanu RM, Tudorache E, Motișan R, Oancea C (2021). Air pollution exposure—the (in) visible risk factor for respira-tory diseases. *Environmental Science and Pollution Research, 28*, 19615–19628. https://doi.org/10.1007/s11356-021-13208-x

Chaudhary, A. K., Telee, L. B. S., Karki, M., & Kumar, V. (2024). Statistical analysis of air quality dataset of Kathmandu, Nepal, with a New Extended Kumaraswamy Exponential Distribution. *Environmental Science and Pollution Research*, *31*(14), 21073-21088. http://dx.doi.org/10.1007/s11356-024-32129-z

Chaudhary, A. K., Telee, L. B. S., Karki, M., & Kumar, V. (2024). Modified inverse exponentiated exponential Poisson distribution to Analyze air quality dataset of Kathmandu, Nepal. *International Journal of Statistics and Applied Mathematics*, *9*(4),125-138. http://dx.doi.org/10.22271/maths.2024.v9.i4b.1783

Ghebreyesus TA (2018) 9 out of 10 people worldwide breathe pol-luted air, but more countries are taking action. *Saudi Medical Journal*, *39*, 641–43. https://doi.org/10.1163/2210-7975_HRD-9841-20180002

Gurung, A., & Bell, M. L. (2013). The state of scientific evidence on air pollution and human health in Nepal. *Environmental Research*, *124*, 54-64. https://doi.org/10.1016/j.envres.2013.03.007

Lamichhane, G.P., Maharjan, N., Pandey, B., K.C., P., Paudel, R., Khanal, S. & Paudel, S.P. (2023). *Status of Air Quality in Nepal, Annual Report, 2021*. Kathmandu: Department of Environment. http://doenv.gov.np/progressfiles/Status-of-Air-Quality-in-Nepal-Annual-Report_2021-1676438644.pdf.

R Core Team (2023). R: A Language and environment for statistical computing. (Version 4.1) [Computer software]. Retrieved from https://cran.r-project.org. (R packages retrieved from CRAN snapshot 2023-04-07).

Shrestha, S. L. (2021). Analysis of ambient particulate air pollution and health in NEPAL. *Journal of Global Ecology and Environment*, *12*(4), 14-29.

Tiotiu, A.I., Novakova, P., Nedeva, D., Chong-Neto, H.J., Novakova, S., Stei-ropoulos, P., & Kowal, K. (2020). Impact of air pollution on asthma outcomes. *International Journal of Environmental Research and Public Health*, *17*(17), 6212. https://doi.org/10.3390/ijerph17176212.

Tuladhar, A., Manandhar, P., & Shrestha, K. L. (2021). Assessment of Health Impact of PM2. 5 Exposure by Using WRF-Chem Model in Kathmandu Valley, Nepal. *Frontiers in Sustainable Cities*, *3*, 672428. https://doi.org/10.3389/frsc.2021.672428

WHO. (2023). *WHO ambient air quality database, 2022 update: status report*. World Health Organization https://iris.who.int/bitstream/handle/10665/368432/9789240047693-eng.pdf?sequence=1