Statistical Identification and Prediction of Dominant Air Pollutant over Four cities of Rajasthan State: Role of Global Climate Change perceptive

Ranadeep Datta, Bikram Mahapatra and Sachin Sachdeva

ISPS M.Sc. Student Project on

"Statistical Innovations and Strategies for Global Warming Challenges"

Department of Statistics

Central University of Rajasthan

Kishangarh, Rajasthan 305 817

Abstract

According to a report of the World Health Organization (1), four cities of Rajasthan viz. Jaipur, Jodhpur, Udaipur, and Kota are in the world's most polluted list. Air pollution controls are a viable strategy because one of the most significant effects of air pollution is global warming. Air pollutants are a threat to any ecosystem whether it has one dominant species or is highly biodiverse. Global climate changes can result in impacts to air quality. Atmospheric warming associated with dominant air pollutants has the potential to increase ground-level ozone in many regions, which may present challenges for compliance with the ozone standards in the future. In this project, an in-depth study of the temporal heterogeneity of Air Quality Index (AQI) and identification of dominant air pollutants across the four seasons of these four cities. Kruskal-Wallis rank-sum test method was investigated and the dominant air pollutants which are affecting these four cities were identified. In this study, the spatial heterogeneity of dominant air pollutants was analyzed using the Wilcoxon-signed rank test. The monthly forecasting of dominant air pollutants for the duration of one year till July 2020 over the cities have also been examined using time series modelling.

The result shows that AQI in the four sites under study has significant temporal heterogeneity. On the other hand, the spatial study of AQI shows that there was no significant spatial heterogeneity between all the pairs of sites.PM10 and PM2.5 were found to be the dominant pollutants. Results of the forecast show PM10 is mostly above permissible limits at all four study sites. There is no gradual decrease or increase in the mean PM10 concentrations over the period of study except for Jaipur.

Keywords: Forecasting, AQI, Kruskal-Wallis rank Test, SARIMA, Spatial-Temporal prediction, Wilcoxon- signed-rank test.

The motivation of the Project:

In recent years, the global temperature of the earth is increasing. The most important point of worry is the diurnal temperature is increasing in every region of the globe. That increasing temperature may come from different sources. Aside from contributing to limiting global warming, strong reductions in methane, black carbon and ground-level ozone have other key benefits for sustainable development. They can protect health and avoid premature deaths by improving air quality. The air quality is an important factor for increasing the temperature and also to modulate the atmospheric phenomena on a global scale. In a county like India, farmers are mostly dependent upon the precipitation which effecting the economy. To reach the Paris Agreement goal of limiting warming to 1.5 (or even 2) degrees Celsius, rapid reduction of CO_2 emissions is absolutely necessary, but will not in itself be sufficient. The Intergovernmental Panel on Climate Change (IPCC) special report on the impacts of global warming of 1.5 °C stresses that deep reductions in emissions of non-CO_2 climate forcers, particularly the air pollutants methane and black carbon, are also crucial. We are very mucking keen on finding out the dominant air pollutant that affecting the four cities of Rajasthan which is much contributing to the global temperature change. Also, the forecasting of such pollutants for any region is a challenging task.

Objectives of the Project:

Based on the above background the objectives of the project are as follows:

(1) Understand the spatiotemporal behaviour of air pollutants over the four cities of Rajasthan.
(2) Examination of Air Quality Index based on the data of CPCB over the four cities.
(3) Detection of dominant air pollutant of four polluted cities.
(4) Prediction of monthly dominant air pollutant using a sophisticated statistical technique such as Autoregressive Integrated Moving Average (ARIMA) Models in the coming years.
(5) Influence of climate change on fine particulate matter and other air pollutants.

These objectives may give the following deliverables:

(1) Understand the interactions between naturally emitted compounds and man-made pollutants in the atmosphere.
(2) Prediction of Air pollutants for the polluted cities in the context of global climate change.
(3) Identify the co-benefits of reducing air pollutants that also reduce the impacts of climate change.
(4) Develop adaptation and mitigation strategies options to reduce the major air pollutants in the respective cities.

(1) Introduction

Climate change and air pollutants share common sources. The most important factor for air quality is the concentration of pollutants nearer the earth's surface. Emissions of air pollutants and their precursors determine regional air quality and can alter climate.

Poor air quality and climate change are closely linked. Burning fossil fuels releases both air pollutants and greenhouse gases. Thus, reducing air pollution from these sources will help to improve air quality and address climate change at the same time as said by UN Environment climate change specialist NiklasHagelberg - "*When addressing air pollution, we also address a critical and easy-to-implement solution to climate change*".

Many Indian cities (including Capital city New Delhi) are included in the list of most polluted cities in the world (3)About 80 per cent of cities in India violates the prescribed standards of ambient air quality (2)Multiple sources contribute to the problem and hence, sector-specific strategies are required for control of air quality.

Spatial distribution of emissions of particulate matter (PM) shows that emission intensity is highest in the Indo-Gangetic plains as well as in the states of Gujarat, Rajasthan, Tamil Nadu, and Maharashtra. (14)

Figure 1*: Indo-Gangetic plain (Wikipedia)*

Rajasthan is a famous tourist destination for Indians as well as for foreign tourists. The Pink City (Jaipur), Suncity (Jodhpur), LakeCity (Udaipur) are popular tourist destinations, featuring many palaces, and forts. Also, according to a recent report of WHO (2014) Jodhpur, Jaipur, Udaipur, and Kota are the most polluted cities in Rajasthan. Based on these facts, the main goal of this study is to pursue a systematic analyse of the spatiotemporal heterogeneity of air pollution at the zonal spatial scale and the seasonal scale by taking four most polluted cities of Rajasthan.

The dynamics of urban air pollution constitute a broad process, which is influenced by factors including seasonal climate and weather, urban functional zoning, and the spatiotemporal distribution of topography, the sources of pollution. Therefore, understanding the spatial heterogeneity of air pollution distribution, pollution sources, and driving forces are essential to formulate appropriate governing policies and plans for pollution control. Study area, methods, and methodology

(2) Study Area

The four meteorological monitoring sites located nearest to environmental monitoring sites were selected. The names, locations of the environmental monitoring sites, are Jaipur (75.8167480 E, 26.875395 N), Jodhpur (73.037522 E, 26.292011 N), Udaipur (73.697468 E, 24.587656 N), and Kota(75.521256 E, 25.143892 N) and the spatial distribution of these environmental monitoring sites is shown in Figure 1. Other demographic and residential areas information is shown in Table 1. The data under study is collected from CPCB(Daily data from Aug 2017 to July 2019) & RPCB(Monthly data from April 2010 to July 2019)(4).

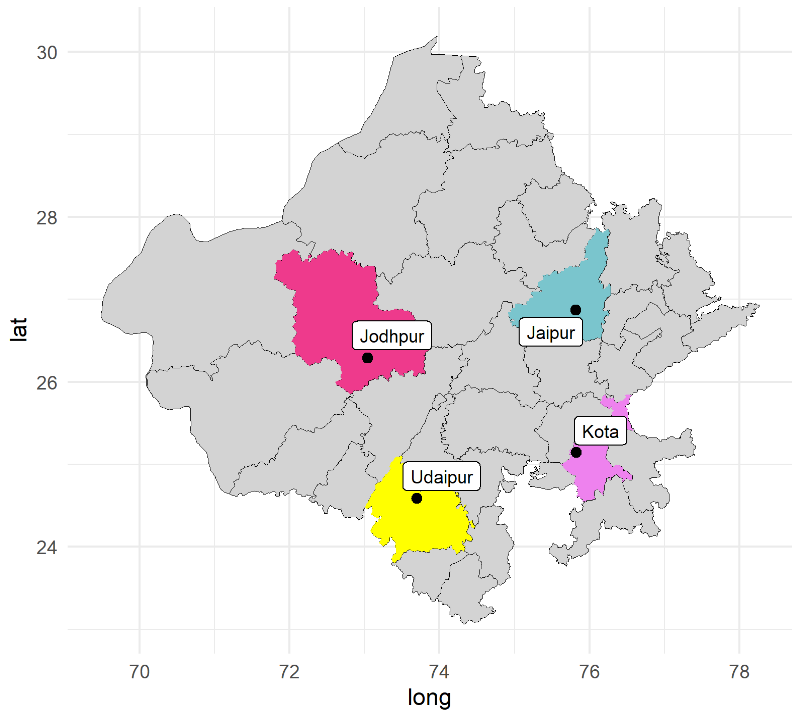

Figure 1: Air Quality Monitoring Stations

Table 1 Demographic information about the four cities

Zones	Population	Residential Area (km^2)
Jaipur	66,26,178	465.09
Jodhpur	36,87,165	214.47
Kota	19,51,014	127.63
Udaipur	30,68,420	109.37

Seasonal wise wind plots of the Jaipur city. Wind prevailing conditions can be analyzed by these Wind-Rose plots:

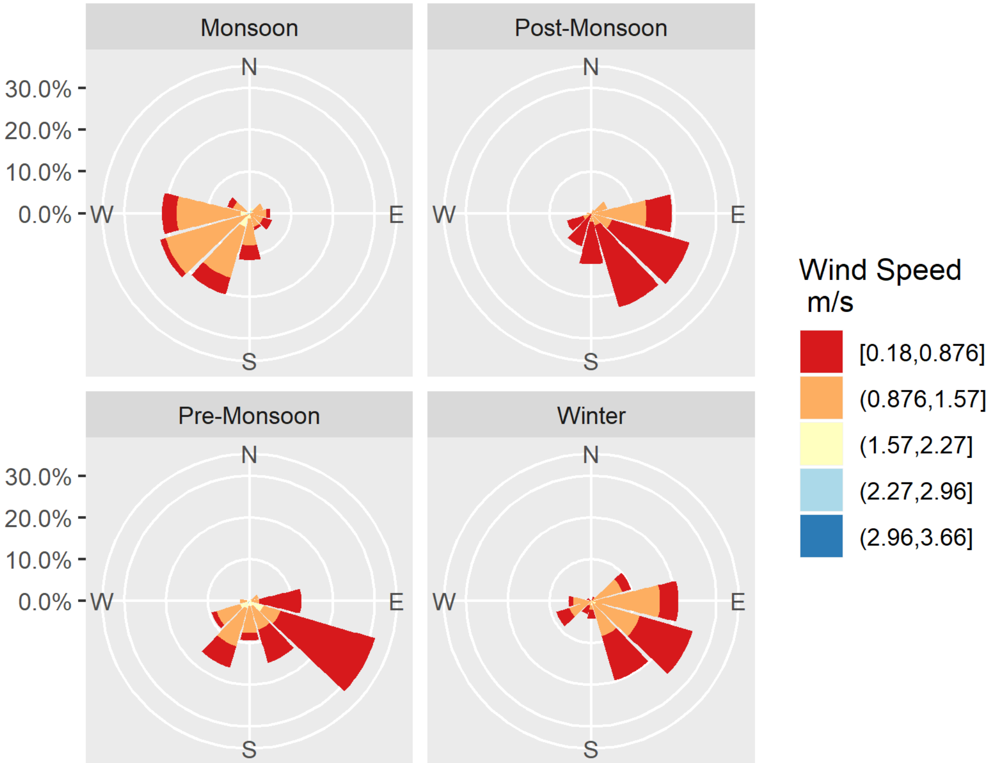

Figure 2 Wind-Rose chart for Jaipur

In Jaipur city prevailing wind condition was South-East (SE) in pre-Monsoon, West (W) in monsoon, SE in post-monsoon and SE in Winter. The wind direction of other cities is given in the following table.

Table 2: *Prevailing Wind Condition*

Zones	Prevailing Wind Condition			
	Pre-Monsoon	**Monsoon**	**Post-Monsoon**	**Winter**
Jaipur	SE	W	SE/E	SE
Jodhpur	SW	SW	E	E
Kota	SE/S	W	SE	SE
Udaipur	SE	W	SE	SE

Table 3: Status of capital NAMP Station (Rajasthan) based on AQI in the year 2018(4)

City	Station code	Location	Area
Jodhpur	274	Regional Office, Jodhpur	Industrial
	273	Sojati Gate	Residential
	376	Mahamandir Police Thana	Residential
	411	Housing Board	Residential
	413	DIC Office	Industrial
	412	Shastri Nagar	Residential
Udaipur	321	Regional Office MIA, Udaipur	Industrial
	320	Ambamata, Udaipur (Chandpur Sattllite Hospital)	Residential
	294	Town Hall, Udaipur	Residential
Kota	17	Regional Office, Kota	Industrial
	325	M/s Samcore Glass Ltd	Industrial
	326	Municipal Corporation Building, Kota	Residential
Jaipur	298	RSPCB Office, JhalanaDoongari	Residential
	410	RIICO Office MIA, Jaipur	Industrial
	296	Ph.D. Office, Ajmeri Gate	Residential
	408	Office of the District Educational Officer, Chandpole	Residential
	409	Regional Office North, RSPCB,6/244 Vidyadhar Nagar	Residential
	297	VKIA, Jaipur (Road no.-6)	Industrial

(3) Methodology

(3.1) Air Quality Index

Daily average concentrations of PM_{10}, O_3, $PM_{2.5}$, SO_2, NO_2, and CO were measured at the four environmental monitoring sites from 24 October 2017 to 31 July 2019. Sub-AQI of each pollutant was calculated using the following model.

$$I_p = \frac{I_{HI} - I_{LO}}{B_{HI} - B_{LO}}\left(C_p - B_{LO}\right) + I_{LO}$$

where,

I_p denotes the sub AQI of the p^{th} pollutant.

C_p denotes the concentration of the p^{th} pollutant.

B_{HI}denotes the breakpoint concentration of the p^{th} pollutant.

B_{HI}denotes the breakpoint concentration greater than or equal to the given concentration.

B_{LO}denotes the breakpoint concentration less than or equal to a given concentration

I_{HI}denotes AQI value corresponding to B_{HI}

I_{LO}denotes AQI value corresponding to B_{LO}

Once the sub-indices are formed, they are aggregated using the formula

$$AQI = \max(I_1, I_2, I_3, \ldots, I_n)$$

Daily AQ scores were sorted into six levels: good, satisfactory, moderately Pullulated, poor, very poor and Severe. Dominant air pollutants were identified according to the AQ level classification criteria specified by the Indian Meteorological Department (IMD).

Table 2 *India AQI Category and Range(6)*

AQI Category	AQI Range	Intensity
Good	0-50	
Satisfactory	50-100	
Moderately polluted	100-200	
Poor	200-300	
Very Poor	300-400	
Severe	400-500	

(3.2) Spatiotemporal Heterogeneity Analysis Methods

The four environmental monitoring sites each represent different districts, so the Spatiotemporalheterogeneity of the distributions of AQ and dominant air pollutants among the sites was studied separately.

(3.2.1) Temporal heterogeneity of air quality

AQ and dominant air pollutants have distinct seasonal characteristics, so the temporal heterogeneity of the distribution of AQ and dominant air pollutants was studied on seasonal scales. The AQI and concentrations of dominant air pollutants can be represented by continuous random variables with unknown probability distributions, and without a paired relationship, so the Kruskal-Wallis rank-sum test method(9) was suitable for testing the differences between the distributions for AQ and the dominant air pollutants on the seasonal scale. The advantages of the Kruskal-Wallis rank-sum test method is that it is a nonparametric test method, multiple group samples can be tested together, and the overall levels of samples can be visually compared with box-plots. The testing process is described below.

Assume that variable x_i, i= 1, 2,…,k denotes a sample of AQI at k sites. First, the null hypothesis H_0 is built as the Equation

$$H_0: F(x_1) = F(x_2) = F(x_3) = F(x_4)$$

$F(x_i)$ denotes the continuous distributions of xi; H_0 denotes AQI at four sites from the same populations.

We rank all $N = \sum_{i=1}^{k} n_i$ observations from smallest to largest, without considering which sample they come from, and assign the smallest observation as rank 1, the next smallest rank 2, …, and the maximum observation as rank N.

Let R_{ij}be the rank of the data pointx_{ij}, and define R_i to be the sum of the ranks in the i^{th}AQI. That is $R_i = \sum_{j=1}^{n_i} R_{ij}$, and then denote each sample mean by $\bar{R}_i = \frac{R_i}{n_i}$.

Let $\bar{R}$ represent the overall mean, and because is$\sum_{i=1}^{k} R_i = \frac{N(N+1)}{2}$,then $\bar{R} = \frac{\sum_{i=1}^{k} R_i}{N} = \frac{N+1}{2}$.

The Kruskal-Wallis test statistic measures the degree to which the actual observed mean ranks $\bar{R}_i$differ from their expected value $(N + 1)/2$. If this difference is large, the null hypothesis H_0 is rejected.

Then the Kruskal-Wallis test statistic is:

$$S_t = \frac{12}{N(N+1)} \sum_{i=1}^{k} n_i(\bar{R}_i - \bar{R})$$

The null hypothesis is tested: when $S_t > \chi^2_{k-1}$ or P ($S_t < \chi^2_{k-1}$) < α, then H_0 is rejected. The null hypothesis H_0 means that there is no difference between the distributions of the AQ and the dominant air pollutants in each of the four seasons. $F(x_1)$ denotes the continuous distributions of x_i; H_0 denotes AQI at four sites from the same populations.

(3.2.2) Spatial heterogeneity of air quality and dominant air pollutant

Spatial heterogeneity of the distributions of AQ and dominant air pollutants among the sites was studied using the Wilcoxon signed-rank test method(10).

The Wilcoxon signed-rank test method was suitable for testing the differences between their distributions since AQI and concentrations of dominant air pollutants can be represented by continuous random variables and there are one-to-one correspondence relations between the four monitoring sites. The advantages of the Wilcoxon signed-rank test method are that it is a nonparametric test method and that it makes full use of the paired difference information in order to reflect the details of the variation between two samples. The testing procedure is described below.

Assume that variable $(x_i, y_i), i = 1, 2, \cdots, n$ denotes the i^{th} AQI values at the same time at two sites. Given a random variable $d_i = x_i - y_i - M_0$, M_0 is the mean of $(x_i - y_i)$ and the median of the population of d_i is M_d. The null hypothesis H_0 means that the AQI and dominant air pollutants are not different between pairs of sites.

$$H_0: M_d = M_0$$

Then, the statistic is constructed as an Equation.

$$Z = \frac{T - n(n+1)/4}{\sqrt{2n(n+1)(2n+1)/48}}$$

$$T = \sum_{i=1}^{n} r(|d_i|) I(d_i > 0)$$

$$I(\rho) = \begin{cases} 1 & \text{if } \rho \text{ is true} \\ 0 & \text{if } \rho \text{ is false} \end{cases}$$

The statistic Z approximately obeys the standard normal distribution, if n >15. Finally, the null hypothesis is tested: When $M_0 < Z_{1-\alpha/2}$ or $M_0 > Z_{\alpha/2}$ or $P(Z_{1-\alpha/2} < M_0 < Z_{\alpha/2}) < a$ then H_0 is rejected, a is significance level, $< Z_{1-\alpha/2}$ and $Z_{\alpha/2}$ are 1 - $a/2$ and $a/2$ quantiles of Z.

(3.3) The relation between the dominant pollutant and other factors

The correlation between dominant air pollutants and their effect factors is complex and not linear, so nonlinear models are more appropriate for finding the relation between the dominant pollutant and other factors. Spearman's Rank Correlation coefficient is used for this.

For a sample of size n, the n pollutant concentrations X_i, Y_i are converted to ranks $r(X_i), r(Y_i)$, and r_s is computed from:

$$r_s = \rho_{r(X_i),r(Y_i)} = \frac{cov(r(X_i), r(Y_i))}{\sigma_{r(X_i)}\sigma_{r(Y_i)}}$$

Where

ρ denotes the usual Pearson correlation coefficient but applied to the pollutant concentrations. $\mathrm{cov}(\mathrm{r}(\mathrm{X_i}), \mathrm{r}(\mathrm{Y_i}))$ is the covariance of the pollutant concentrations. $\sigma_{\mathrm{r}(\mathrm{X_i})}$ and $\sigma_{\mathrm{r}(\mathrm{Y_i})}$ are the standard deviations of the pollutant concentrations.

(3.4) Prediction of dominated air pollutant using Time Series Modeling Procedure

The seasonal ARIMA model is used for trend analysis of air pollutant data of PM_{10} for the period of January 2010 to July 2019 of all four cities

(3.4.1) Autoregressive Integrated Moving Average (ARIMA) Models

The ARIMA model is a generalization of an ARMA model to include the case of non-stationarity as well. In ARIMA models a non-stationary time series is made stationary by applying finite differencing of the data points. The mathematical formulation of the ARIMA (p, d, q) model using lag polynomials is given below

$$\left(1 - \sum_{i=1}^{p} \varphi_i L^i\right)(1-L)^d y_t = \left(1 + \sum_{i=1}^{p} \theta_i L^i\right)\epsilon_t$$

Where L is the lag operator, the φ_i are the parameters of the autoregressive part of the model, the θ_i are the parameters of the moving average part and the ϵ_t are error terms. The error terms ϵ_t are generally assumed to be independent, identically distributed variables sampled from a normal distribution with zero mean.

Here, *p*, *d*, and *q* are integers greater than or equal to zero and refer to the order of the autoregressive, integrated, and moving average parts of the model respectively. The integer *d* controls the level of difference. When *d* = 0, then it reduces to ARMA(*p*, *q*) model. An ARIMA(*p*,0,0) is nothing but the AR(*p*) model and ARIMA (0,0, *q*) is the MA(*q*) model.

(3.4.2) Seasonal Autoregressive Integrated Moving Average (SARIMA) Model

The ARIMA model is for non-seasonal non-stationary data. Box and Jenkins [6] have generalized this model to deal with seasonality. Their proposed model is known as the Seasonal ARIMA (SARIMA) model.

One shorthand notation for the model is ARIMA (*p*, *d*, *q*) × (*P*, *D*,*Q*)*S*, Where *p* = nonseasonal AR(Autoregressive) order, *d* =non-seasonal differencing, *q* = non-seasonal MA (Moving Average) order, *P* = seasonal AR order, *D* = seasonal differencing, *Q* = seasonal MA order, and *S* = time span of repeating seasonal pattern. The mathematical formulation of a SARIMA (p, *d*, *q*) $\times (\mathrm{P}, \mathrm{D}, \mathrm{Q})^{\mathrm{s}}$, model in terms of lag polynomials is given below.

$$\emptyset_{(P)}(\{L\})^{s}\varphi_{P}(L)(1-L)^{d}(1-L^{s})^{d}y_t = \Theta_p(\{L\})^{s}\,\theta_q(L)$$

$$\text{i.e.} \qquad \emptyset_{(P)}(\{L\})^{s}\varphi_{P}(L)z_t = \Theta_p(\{L\})^{s}\,\theta_q(L)$$

Here z_t is the seasonally differenced series.

(3.5) Stationarity Analysis

The stationarity of data is checked by the Augmented Dicky-Fuller test. If data is non-stationary, seasonal differencing of appropriate order is used to remove non-stationarity from the series. A first-order seasonal difference is the difference between an observation and the corresponding observation from the previous year and is calculated as.

$$Z_t = y_t - y_{t-s}$$

For monthly time series *s* = 12 and for quarterly time series *s* = 4.

(3.5.1) Augmented Dicky-Fuller (ADF) test

The Augmented Dickey-Fuller test is a type of statistical test called a unit root test. The intuition behind a unit root test is that it determines how strongly a time series is defined by a trend. It uses an autoregressive model and optimizes an information criterion across multiple different lag values. The test is applied to the model

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \cdots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t,$$

where γ is the parameters of the autoregressive part of the model, α is a constant, β the coefficient on a time trend and p the lag order of the autoregressive process.

The unit root test is then carried out under the null hypothesis $\gamma = 0$, against the alternative hypothesis of $\gamma < 0$, i.e.

H_0: Timeseries is not stationary

H_1: Time series is stationary

The test statistic of the ADF test is given by

$$DF_\tau = \frac{\hat{\gamma}}{SE(\hat{\gamma})}$$

We interpret this result using the p-value from the test. A p-value below a threshold (such as 5% or 1%) suggests we reject the null hypothesis (stationary), otherwise, a p-value above the threshold suggests we fail to reject the null hypothesis (non-stationary).

(4) Results

(4.1) Spatiotemporal Heterogeneity of AQ Distribution

(4.1.1) Temporal Heterogeneity of AQI Distribution

The AQI differences at each site i.e. Jaipur, Jodhpur, Udaipur & Kota were analyzed on a seasonal time-scale. The Kruskal-Wallis rank-sum test method is suitable for testing the differences at a significant level of 5% between the AQI because there is no paired relationship in the seasonal AQI data at each site and the distribution type and parameters for AQI is unknown.

Now here the Calculated Statistics S_t = 464.1856 and at 5% significance level $\chi^2_{0.05,3}$ = 7.815. So, $S_t > \chi^2_{0.05,3}$, then we reject the null hypothesis that is there was significant temporal heterogeneity in terms of the overall level of AQI between the four seasons at each site.

The height of the boxes is smaller in monsoon except for Jodhpur but the outlier values are larger so it can be inferred that AQI varied less in summer and the height of the boxes is larger in winter except in Kota. So, it can be seen that the AQI in monsoon was less varied and lower on average than in the other three seasons, while the AQI in winter was more varied and higher on average than in the other three seasons except in Jodhpur & Kota respectively. It can be inferred that the AQ was best in monsoon and worst in winter at sites Jodhpur & Kota respectively.

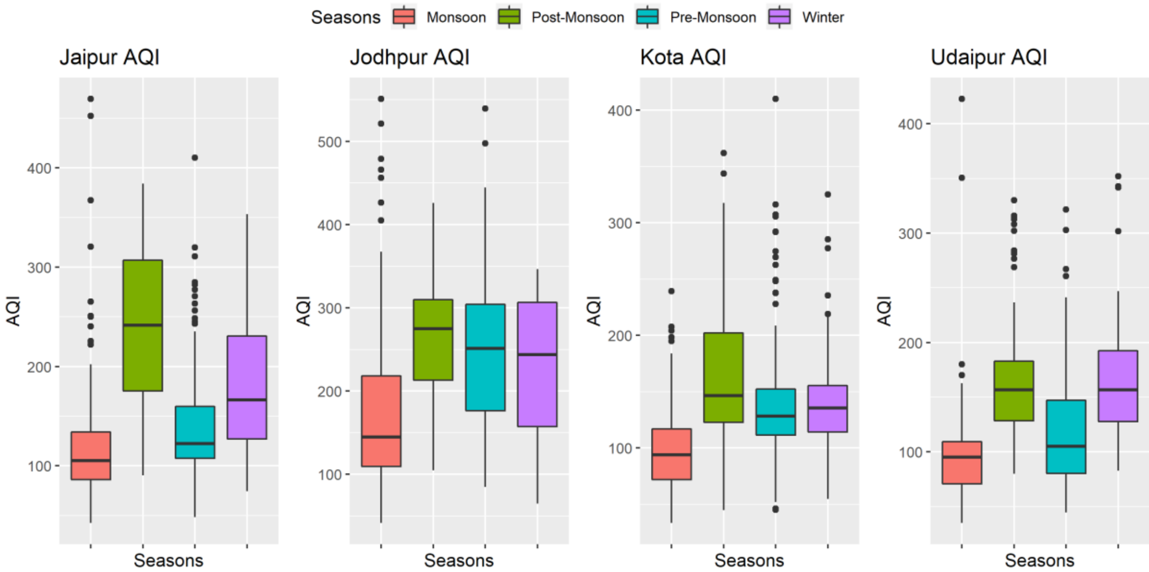

Figure 3 Box-Plot of AQI

(4.1.2) Spatial Heterogeneity of AQI Distribution

The spatial heterogeneity of overall levels of AQ can be reflected by comparing the differences in AQI between the four sites.

In order to further analyze the spatial heterogeneity and characteristics of air pollutant distribution at the four sites, the Wilcoxon signed-rank test method was applied to test the differences in AQI at a significant level of 5% between pairs of sites.

Table 3 Result of Wilcoxon test of AQI differences among the four cities

Pair of Sites	P-Value	Ho
JAIPUR - JODHPUR	0.451029	Accept
JAIPUR - KOTA	0.052623	Accept
JAIPUR - UDAIPUR	0.093542	Accept
JODHPUR - KOTA	0.447255	Accept
JODHPUR - UDAIPUR	0.685478	Accept
KOTA - UDAIPUR	0.840699	Accept

According to the null hypothesis testing method, it can be inferred that the AQ was significantly similar at all pairs of sites. So, it can be inferred that there is no significant spatial heterogeneity between all the pairs of sites.

To further analyze the temporal heterogeneity of air pollutant distribution, the number of days for each level of AQ and the dominant air pollutants in each season were calculated. Figure.shows that AQ level counts at all four sites generally presented significant temporal heterogeneity. The number of days on which AQ ranked as polluted was the largest in Post -Monsoon in Jaipur & Jodhpur. In Monsoon AQ ranked as moderately polluted in each site. But also, in Monsoon AQ is quite satisfactory in Kota & Udaipur than the other two cities. In Winter AQ ranked as moderately polluted in all the cities.

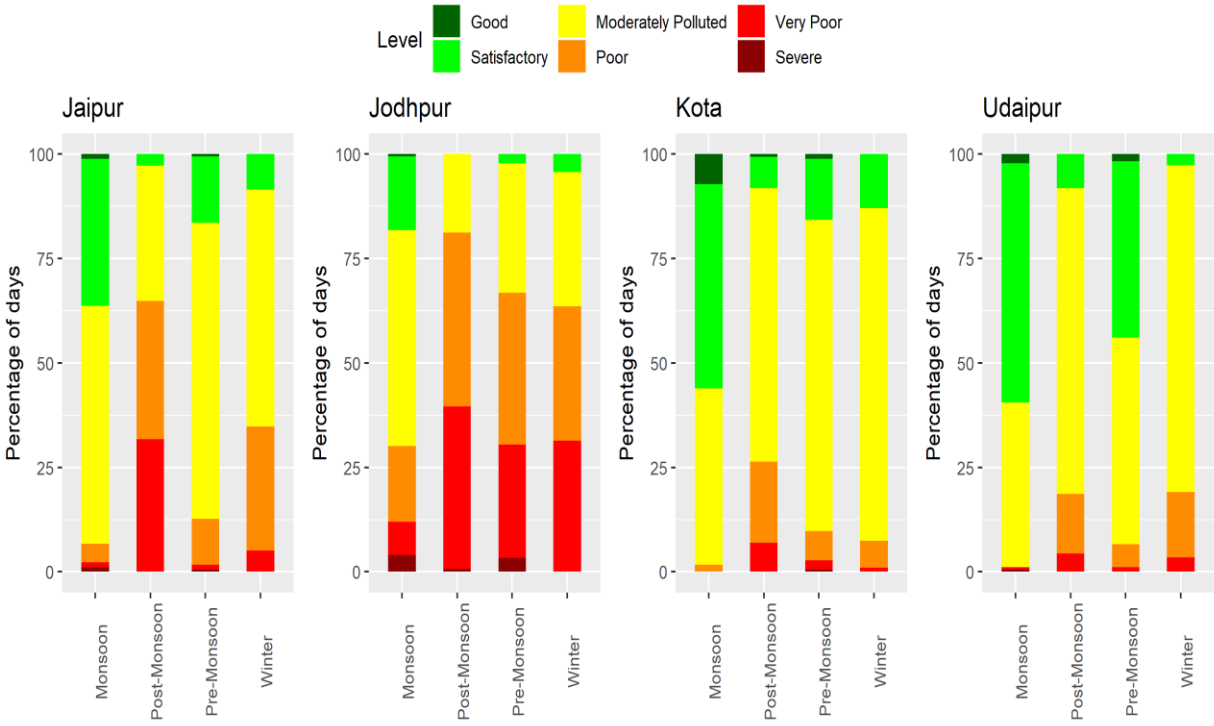

Figure 4 Percentage of number of days of each level of Air Quality (AQ)

(4.2) Dominant pollutant in Each city

First, the heterogeneity of the dominant air pollutants, in terms of the number of days on which each pollutant was dominant, was examined.PM2.5 and PM10 were dominant air pollutants in 2-year span at all 4 cities.

The following figures describe the dominant air pollutant in each season in all four cities. These figures show the percentage of days of dominant air pollutants.

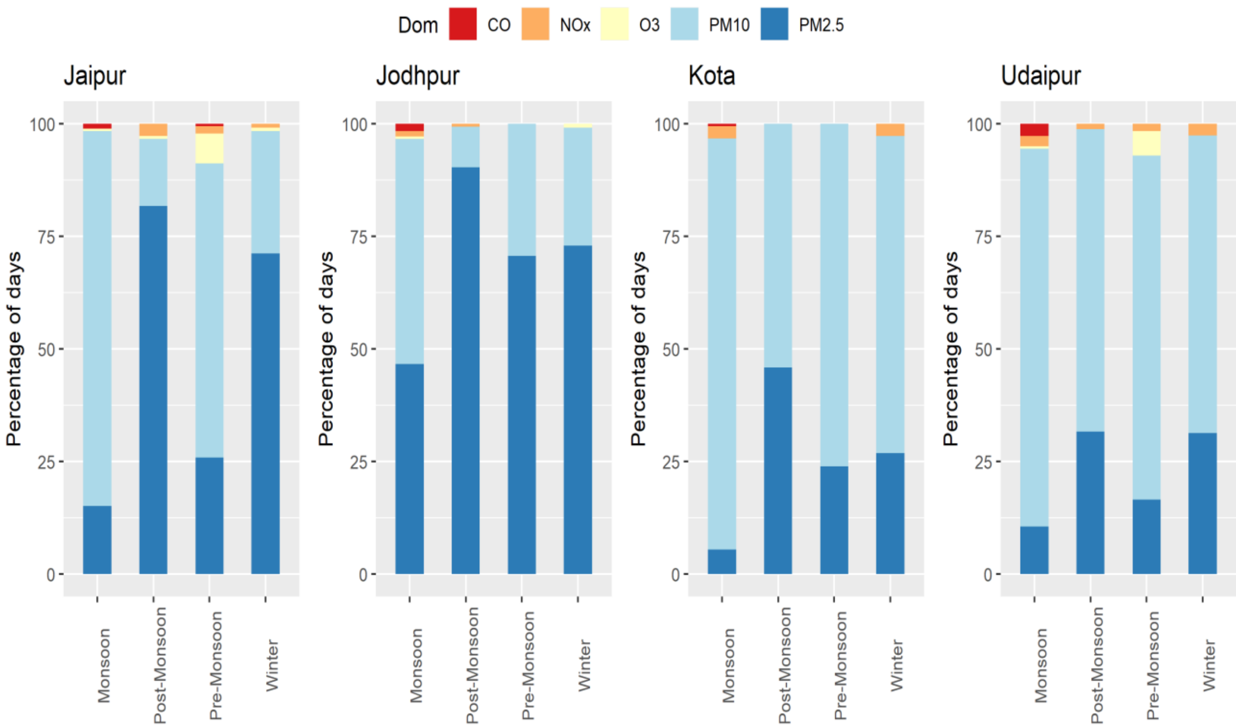

Figure 5 Percentage of number of days of dominant air pollutants

(4.2.1) Spatiotemporal Heterogeneity of Dominant Air Pollutants Distribution

In order to further identify the distribution characteristics of the two dominant air pollutants, the spatial heterogeneity was tested with Daily concentrations of the two dominant air pollutants among the four sites. The Wilcoxon signed-rank test method is suitable for testing these differences because the concentrations of the two dominant air pollutants have a one-to-one correspondence by time among the sites. Therefore, the Wilcoxon signed-rank test method was applied to test the difference at a significant level of 5%.

Table 4 Results of Wilcoxon test of PM10 & PM2.5 differences among the four sites

	Pre-Monsoon		**Monsoon**		**Post Monsoon**		**Winter**	
Jaipur -Jodhpur	PM 2.5		PM 10		PM 2.5		PM 2.5	
	Ho	P - Value	Ho	P - Value	Ho	P - Value	Ho	P - Value
	Reject	0.00366	Reject	0.023087	Reject	0.002817	Reject	.000043

	Pre-Monsoon		**Monsoon**		**Post Monsoon**		**Winter**	
Jodhpur-Udaipur	PM 10		PM 10		PM 2.5		PM 2.5	
	Ho	P - Value	Ho	P - Value	Ho	P - Value	Ho	P - Value
	Reject	0.002077	Reject	0.000058	Accept	0.093542	Reject	0.010772

	Pre-Monsoon		Monsoon		Post Monsoon		Winter	
Udaipur-Kota	PM 10		PM 2.5		PM 10		PM 10	
	Ho	P - Value	Ho	P - Value	Ho	P - Value	Ho	P - Value
	Reject	0.000011	Reject	0.00001	Reject	0.00257	Accept	0.070451

	Pre-Monsoon		Monsoon		Post Monsoon		Winter	
Jaipur - Udaipur	PM 10		PM 10		PM 2.5		PM 2.5	
	Ho	P - Value	Ho	P - Value	Ho	P - Value	Ho	P - Value
	Accept	0.06906	Reject	<0.00001	Accept	0.89562	Accept	0.87335

	Pre-Monsoon		Monsoon		Post Monsoon		Winter	
Kota - Jaipur	PM 10		PM 10		PM 2.5		PM 2.5	
	Ho	P - Value	Ho	P - Value	Ho	P - Value	Ho	P - Value
	Accept	0.0657404	Reject	<0.00001	Reject	0.000098	Accept	0.909555

	Pre-Monsoon		Monsoon		Post Monsoon		Winter	
Kota - Jodhpur	PM 10		PM 10		PM 2.5		PM 2.5	
	Ho	P - Value	Ho	P - Value	Ho	P - Value	Ho	P - Value
	Accept	0.18773	Accept	0.440707	Reject	0.00061	Reject	0.032436

For the Wilcoxon signed-rank test of the concentration of PM10 between the pairs of sites Jodhpur - Udaipur, and Udaipur – Kota in Pre -Monsoon the p-values were smaller than 0.05 Whereas between the pairs of sites like Jaipur – Udaipur, Kota – Jaipur & in Kota – Jodhpur the p- values were greater than 0.05. Thus, it can be deduced that there was significant spatial heterogeneity in the concentrations of PM_{10} only between Jodhpur - Udaipur, Udaipur – Kota on the other hand In the pair of Jaipur – Udaipur, Kota – Jaipur & in Kota – Jodhpur, there was no significant spatial heterogeneity in the concentration of PM 10. In the concentration of $PM_{2.5}$ between the pair of sites Jaipur – Jodhpur, Kota – Jaipur & in Kota – Jodhpur in Post -Monsoon the p- values were smaller than 0.05 whereas in the pairs like Jaipur – Udaipur & Jodhpur – Udaipur the p-values were greater than 0.05. Thus, it can be deduced that there was significant spatial heterogeneity in the concentrations of $PM_{2.5}$ between Jaipur – Jodhpur, Kota – Jaipur & Kota – Jodhpur on the other hand in the pair of Jaipur – Udaipur & Jodhpur – Udaipur, there was no significant spatial heterogeneity. In the concentration of $PM_{2.5}$ the pair of sites Jaipur -Jodhpur, Jodhpur- Udaipur, Kota – Jodhpur in Winter the p-values were smaller than 0.05 whereas in the pairs like Jaipur – Udaipur & Kota – Jaipur the p- values are greater than 0.05. Thus it can be deduced that there was significant spatial heterogeneity in the concentrations of $PM_{2.5}$ between Jaipur -Jodhpur, Jodhpur- Udaipur, Kota – Jodhpur on the other hand in the pair of Jaipur – Udaipur & Kota – Jaipur, there was no significant spatial heterogeneity in Winter. Similarly, the concentration of PM_{10} between the pairs of sites Jaipur -Jodhpur, Jodhpur- Udaipur, Jaipur – Udaipur, Kota – Jaipurin Monsoon the p-values were smaller than 0.05 Whereas in Kota - Jodhpur the *p*-values were greater than 0.05. Thus, it can be deduced that there was significant spatial heterogeneity in the concentrations of PM10 between Jaipur -Jodhpur, Jodhpur- Udaipur, Jaipur – Udaipur, Kota – Jaipur except in Kota - Jodhpur there was no significant spatial heterogeneity in the concentration of PM_{10} in Monsoon.

(4.3) The Effect of the Relevant Factors on the Dominant Air Pollutants

The Spearman rank correlation coefficient between dominant air pollutants and the effect factors at the four sites was calculated. The main driving factors which are affecting PM_{10} and $PM_{2.5}$ were pollutants such as SO_2, NO_2, and CO. Other factors like Temperature, Pressure and wind speed have diverse effects on dominant air pollutants.

Table 5 Correlation between dominant air pollutants and the effect factors at the four cities

	Jaipur		**Jodhpur**		**Kota**		**Udaipur**	
	PM_{10}	**$PM_{2.5}$**	**PM_{10}**	**$PM_{2.5}$**	**PM_{10}**	**$PM_{2.5}$**	**PM_{10}**	**$PM_{2.5}$**
SO_2	0.210	0.259	-0.078	0.015	0.361	0.370	0.133	0.091
NO_2	0.296	0.412	0.141	0.165	0.111	0.348	0.465	0.447
CO	0.228	0.279	0.063	0.131	0.303	0.522	0.199	0.065
Temp	0.047	-0.195	0.085	0.076	-0.110	-0.252	-0.531	-0.507
BP	0.204	0.319	0.017	0.018	0.406	0.171	0.543	0.530
WS	-0.175	-0.266	-0.149	-0.267	-0.293	-0.355	-0.417	-0.385

Variation in SO_2, NO_2 and CO have a positive but weak correlation with a variation of $PM_{2.5}$ and PM_{10} in Jaipur, Kota, and Udaipur except in Jodhpur, SO_2hasa weak negative correlation with PM_{10}.

Variation in temperature has little effect on the variation of $PM_{2.5}$ and PM_{10} at Jaipur, Jodhpur, and Kota except in Udaipur, Variation in temperature is strongly negatively correlated with the variation of $PM_{2.5}$ and PM_{10} at all four cities. Variation in Pressure has little effect on the variation of $PM_{2.5}$ and PM_{10}. Variation in Wind speed is negatively correlated with variation of $PM_{2.5}$ and PM_{10} at all four cities

(4.4) Prediction of dominant air pollutant - PM_{10}

Particulate Matter (PM) can be emitted directly or formed in the atmosphere. The particles released directly to the atmosphere are called "Primary particles", these include dust from roads and black and/or elemental carbon from combustion sources. The particles formed in the atmosphere, "Secondary particles", are formed in the atmosphere from chemical reactions involving primary gaseous emissions like NO_2, SO_2, and NH_3. Particulate matter does not only cause health effects, but it also plays a role in the greenhouse effect and global warming because of its contribution to cloud formation.

The function *auto.arima()* in R was used to fit the best ARIMA Model to our PM10 monthly concentration data for each of the four cities. The best fit for each city was found to be different. The NAAQ standard for PM10 is 100 µg/m^3 annually,represented by a red-dotted line in the figures below, which is usually breached in all four cities at least once in a year. It also observed that PM10 concentration is usually highest during a year at the end of the two halves, i.e., early monsoon and late post-monsoon respectively. This concentration starts to fall during Winter and reaches the lowest point during Monsson, confirming our observations of the previous study on two-year data.

The best fit models for each city found were used to forecast PM10 concentrations for the next 24 months. The black line represents the observed values and the red line shows model fitted values in the figures below. The blue line represents the predicted mean PM10 values and the blue shaded region is the 95% confidence limit.

Jaipur

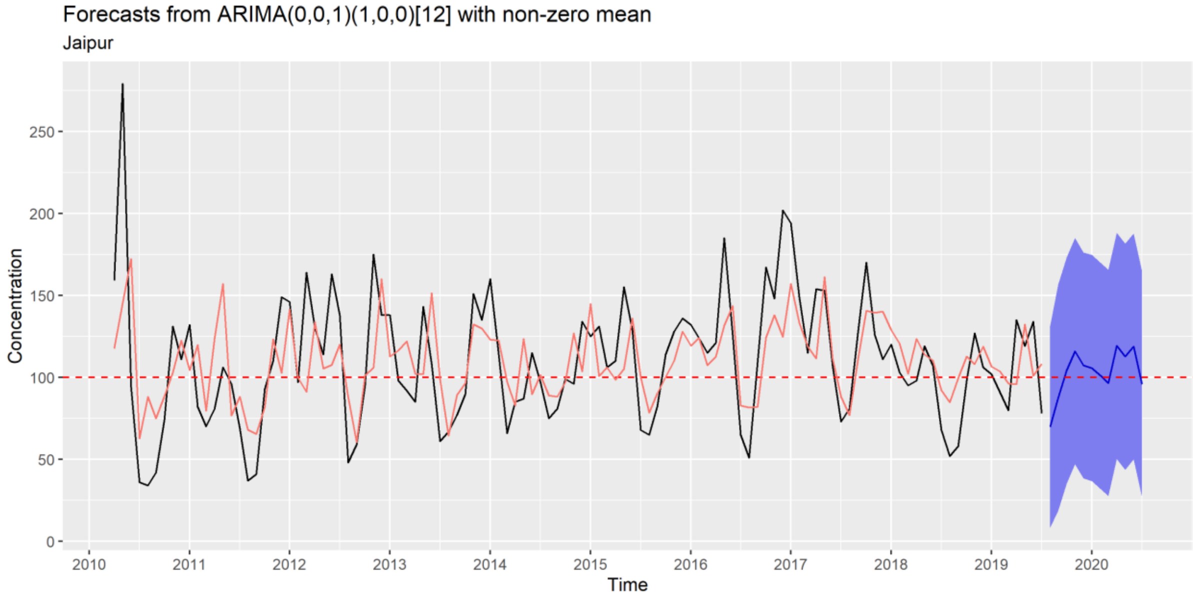

Figure 6 Forecast based on monthly data

ARIMA (0,0,1)(1,0,0)[12]		
	H_0	p-value
Augmented Dickey-Fuller Test	Accept	0.6109

The ARIMA(0,0,1)(1,0,0)[12] model is used for forecasting of mean PM_{10} value in Jaipur. The forecast shows a relative fall in PM_{10} concentration when compared to the corresponding period in previous years.

Jodhpur

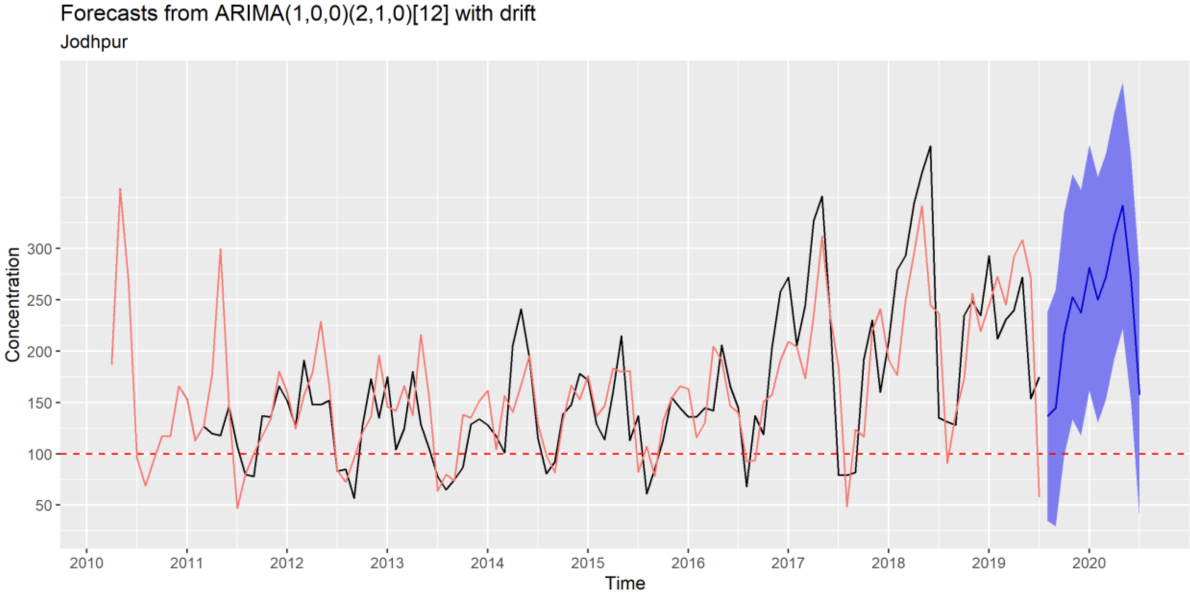

Figure 7 Forecast based on monthly data

ARIMA (1,0,0)(2,1,0)[12]		
	H_0	p-value
Augmented Dickey-Fuller Test	Accept	0.2754

The ARIMA (1,0,0)(2,1,0)[12] model was used for forecasting mean PM_{10} values in Jodhpur. The model fits very well in the early stage, with fitted values overlapping the observed values during the period 2010-11 which shows a good fit. The forecast shows a similar trend of high values at the end of the two halves in a year, in fact, higher than the values in the corresponding period in the previous year.

Kota

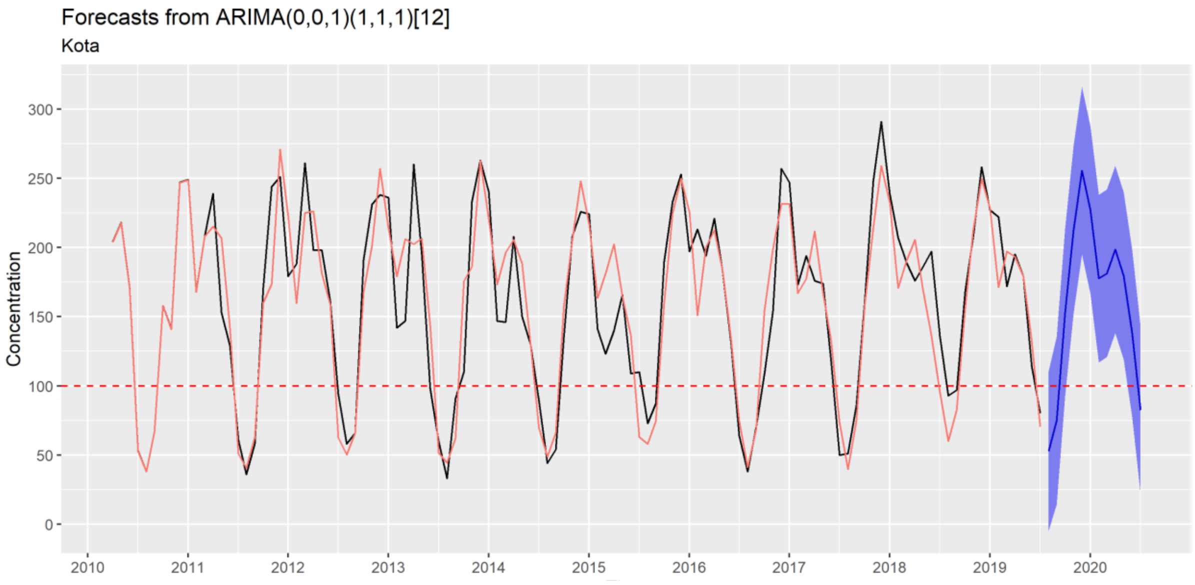

Figure 8 Forecast based on monthly data

ARIMA (0,0,1)(1,1,1)[12]		
	H_0	p-value
Augmented Dickey-Fuller Test	Accept	0.3804

The PM10 values varied very much in Kota, achieving extreme values in almost every year. ARIMA (0,0,1)(1,1,1)[12] model was used to forecast mean PM_{10} values for Kota also the fitted values overlap the observed values at an early stage.

Udaipur

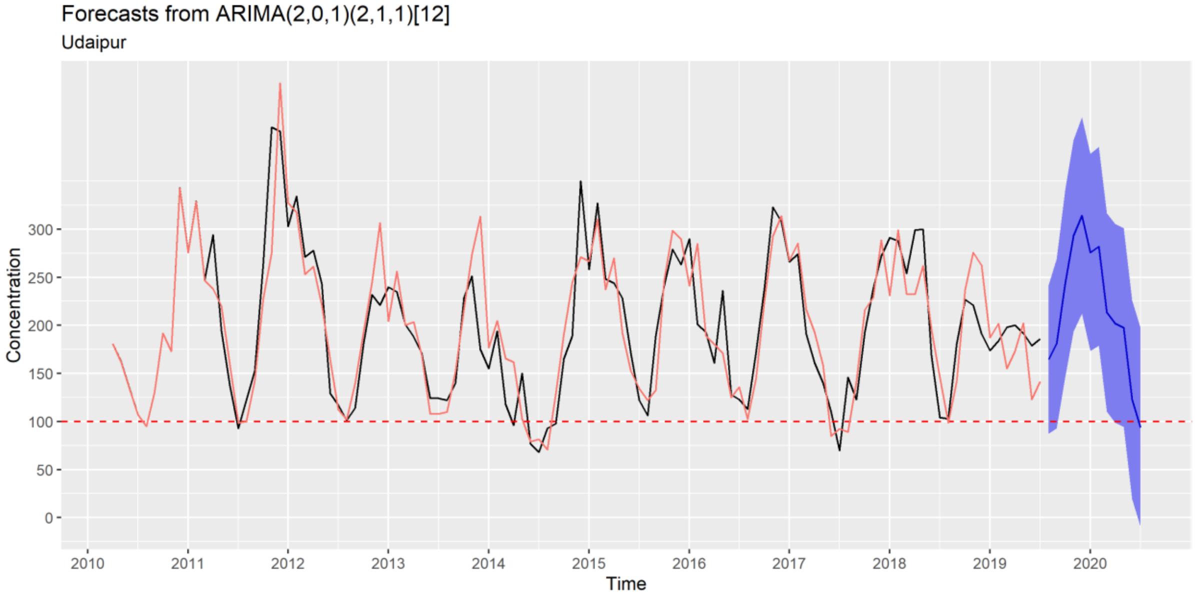

Figure 9 Forecast based on monthly data

ARIMA (2,0,1)(2,1,1)[12]		
	H_0	p-value
Augmented Dickey-Fuller Test	Accept	0.5237

PM_{10} concentrations in Udaipur show downfall when comparing recent years, but the forecast shows otherwise. ARIMA (2,0,1)(2,1,1)[12] model was used to forecast mean PM_{10} values in Udaipur.

Conclusions:

Global warming, as a prime factor of climate change, is caused by a blanket of pollution that traps heat around the earth. This pollution comes from cars, factories, homes, and power plants that burn fossil fuels such as oil, coal, natural gas, and gasoline. Global warming pollution knows no boundaries. The particular matter is a widespread air pollutant, present wherever people live. The major impacts of air pollutants on human health. Since even at relatively low concentrations the burden of air pollution on health is significant, effective management of air quality aiming to achieve WHO AQI levels are necessary to reduce health risks to a minimum. The size of particular matter particles is directly linked to their potential for causing health problems. Also, the second point is the influence of global climate change on particulate matter concentrations was investigated by means of air quality modelling simulations to finding out the hot spot regions for different air pollutants over different regions.

The spatiotemporal heterogeneity of air pollution distribution has been systematically analyzed by using the Kruskal-Wallis rank-sum test and the Wilcoxon signed-rank test. The relation between the dominant pollutant and other pollutant were evaluated using Spearman's rank correlation coefficient.

We found that there was significant temporal heterogeneity in terms of the overall level of AQI between the four seasons at each site. In the temporal view of our study, the number of days on which AQ ranked as polluted was largest in Post -Monsoon in Jaipur & Jodhpur.

On a spatial scale, air quality was relatively better at the Kota and Udaipur sites, and worst at the Jaipur & Jodhpur. The concentration of PM_{10} was highest at Jodhpur; this was mainly because of building dust from rural reconstruction e.g. Sajoti Gate, Mahamandir Police station, Housing Board residential area, industrial dust from steel plants & marble factory from Industrial area like DIC office area. The concentration of $PM_{2.5}$ was highest at Jaipur, this is mainly due to building dust from rural reconstruction e.g. JhalanaDoongari, Ajmeri Gate residential area, industrial dust from steel plants, thermal power plants, etc from an industrial area like RIICO Office MIA, VKIA, Jaipur (

Table).

PM_{10} and $PM_{2.5}$were the dominant air pollutant in all four cities. We tried to make a statistical contribution to global warming by forecasting the concentration of dominant air pollutants (PM_{10}) because particulate matter does not only cause health effects, it also plays a role in the greenhouse effect and global warming because of its contribution to cloud formation or soils. The findings of this study may provide a comprehensive database for framing an appropriate strategy for necessary preventive measures.

This study reveals that the particulate pollutant, PM_{10} is mostly above permissible limits at all four study sites. There is no gradual decrease or increase in the mean PM_{10} concentrations over the period of study except for Jaipur, which shows a slightly decreasing trend.

The methods explored in this study can be applied to the spatiotemporal heterogeneity analysis of pollutant distribution in other polluted regions in India e.g. Delhi, Kanpur, Allahabad, Kolkata, etc.

The major finding of the study is the dominant air pollutants over the four cities such as PM10 & PM2.5. The modelling study for such air pollutant (monthly) for these cities for the year 2020 were detailed explained. In this way, we can find out the hot spot of major air pollutants over the different regions. The modelling study also can be improved and validated by other sophisticated techniques for the robustness of the study for other regions of India. These major air pollutants may cause environmental problems such as acid rain, increasing black carbon causes human health problems, global dimming causes ecological problems such as changes in evaporation and rainfall patterns, droughts and/or increased rainfall cause problems for agriculture. We all need to start putting the planet first if we want to be able to thrive on it. The planet will survive with or without us. If we want it to be a place where we can live, we need to stop pointing fingers and expecting someone else to fix it. This will only work from the bottom up.

Particulate air pollution can be reduced using current technologies.

Interventions resulting in a reduction in the health effects of air pollution range from regulatory measures (stricter air quality standards, limits for emissions from various sources), structural changes (such as reducing energy consumption, especially that based on combustion sources, changing modes of transport, land use planning) as well as behavioural changes by individuals by, for example, using cleaner modes of transport or household energy sources. There are important potential co-benefits of integrating climate change and air pollution management strategies, as evidenced by the importance of the PM indicator and climate change contributor black carbon.

References

(1) https://www.hindustantimes.com/jaipur/4-rajasthan-cities-in-world-s-most-polluted-list-jodhpur-the-worst/story-6VFNZBzmsnqKzG3GmmtZ1L.html

(2) https://app.cpcbccr.com/ccr/#/caaqm-dashboard-all/caaqm-landing/caaqm-comparison-data

(3) Zhen Cheng a , Lina Luo a , Shuxiao Wang b, ∗, Yungang Wang c,i , Sumit Sharma d , HikariShimadera e , Xiaoliang Wang f , Michael Bressi g , Regina Maura de Miranda h , Jingkun Jiang b , Wei Zhou b , Oscar Fajardo b , Naiqiang Yan a , Jiming Hao b.,(2016) Status and characteristics of ambient PM2.5 pollution in global megacities.Environment International

(4) http://rspcbmis.environment.rajasthan.gov.in/NAMP_Report.aspx?HeaderId=4&MenuId=6

(5) Fiore, A.M., Naik, V., Spracklen, D.V., Steiner, A., Unger, N., Prather, M., Bergmann, D., Cameron-Smith, P.J., Cionni, I., Collins, W.J. and Dalsøren, S., (2012). Global air quality and climate. Chemical Society Reviews.

(6) National Air Quality Index, Central Pollution Control Board India.

(7) Mohapatra, K., and Biswal S.K. (2014).Effect of particulate matter (PM) on plants, climate, ecosystem, and human health. *International Journal of Advanced Technology in Engineering and Science*..

(8) Dadhich A. P., Goyal, R., Dadhich, P.N. (2017).Assessment of Spatiotemporal Variations in the air quality of Jaipur city, Rajasthan, India.The*Egyptian Journal of Remote Sensing and Space Sciences.*

(9) Ostertagova, E.; Ostertag, O.; Kovac, J, (2014)Methodology and Application of the Kruskal-Wallis Test.*Appl. Mech. Mater.*

(10) Yeo, I.K. (2017). An algorithm for computing the exact distribution of the Wilcoxon signed-rank statistic. *Journal ofKoreanStatistical Society.*

(11) Naveen V. and Anu, N. (2017).Time Series Analysis to Forecast Air Quality Indices inThiruvananthapuram District, Kerala, India.*International Journal of Engineering Research and Application*

(12) Lei J. and Ling B.(2018). Spatiotemporal characteristics of urban air pollutions and theircausal relationships: Evidence from Beijing and its neighbouring cities. Scientific Reports, *Nature research Journal.*

(13) A policy brief by the energy and Resources Institute (TERI), Feb 2018.

(14) David, L. M., Ravishankara, A. R., Kodros, J. K., Pierce, J. R., Venkataraman, C., &Sadavarte, P. (2019). Premature mortality due to PM2. 5 over India: Effect of atmospheric transport and anthropogenic emissions. *GeoHealth*.