

# Curvilinear and gamma generalized linear models for modeling the dependency of air pollution levels on meteorological conditions in Kathmandu valley

Srijan Lal Shrestha (✉ [srijan\\_shrestha@yahoo.com](mailto:srijan_shrestha@yahoo.com))

Tribhuvan University <https://orcid.org/0000-0002-9390-8215>

---

## Research Article

**Keywords:** Box-Cox transformation, exponential model, gamma generalized linear model, meteorology, ozone air pollution, particulate air pollution

**Posted Date:** January 27th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1299421/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# **Curvilinear and gamma generalized linear models for modeling the dependency of air pollution levels on meteorological conditions in Kathmandu valley**

## **Author**

### **Principal & Corresponding Author**

Srijan Lal Shrestha

Professor

Central Department of Statistics

Tribhuvan University

Kirtipur, Kathmandu, Nepal

Email ID: srijan\_shrestha@yahoo.com; srijan.shrestha@cds.tu.edu.np

ORCID ID: 0000-0002-9390-8215

## **Acknowledgements**

The paper is based upon the findings of the World Health Organization (WHO) study under Urban Health Initiative (UHI) with project title “To provide support on the analysis of air pollution data in Kathmandu valley” The author would like to express deep gratitude to WHO for the initiation of the study and Pierpaolo Mudu, Technical Officer, Air Quality and Health (AQH) for providing an excellent opportunity to work on a WHO study related to air pollution assessment of Kathmandu valley. The author is also thankful to all the data providers including Database Platform, The World Air Quality Project, US Embassy (Kathmandu), International Centre for Integrated Mountain Development (ICIMOD), Open Data Nepal and Department of Hydrology and Meteorology (DHM). Lastly, the author is grateful to all the supporters, and those who have assisted and helped during the study.

## Abstract

Effects of meteorological parameters including confounding effects of seasonality and lag effects on air pollution levels are investigated and quantified for Kathmandu valley, Nepal using daily local temporal data of 2017-2020 available through air pollution monitoring by Department of Environment and US Embassy, Kathmandu, Nepal, and meteorology monitored by Department of Hydrology and Meteorology, Kathmandu, Nepal. Data modeling are performed through Exponential, Box-Cox transformed and Gamma Generalized Linear Models since air pollution levels are found non-normal with the presence of non-constant error variance in linear modeling. Results depict high proportions of observed air pollution variations (79%-85%) explained by the fitted models with around 5% reduction in  $PM_{10}$  and  $PM_1$  levels per  $1^\circ C$  increase in average temperature and significant increase in surface  $O_3$  level (0.177 Box-Cox transformed value) per  $1^\circ C$  increase in average temperature. Similarly, around 0.7% and 2% decrease in  $PM_1$  and  $PM_{10}$  per 1% increase in relative humidity, 0.032 decrease in transformed value of  $PM_{2.5}$  per 1 mm increase in rainfall, and 7.3% decrease in  $PM_{10}$  per 1 m/s increase in wind speed are also detected. Other effects are also quantified in terms of Box-Cox transformed values for statistically significant effects due to relative humidity and wind. In conclusion, meteorological conditions are found significant contributing factors in determining air pollution levels as demonstrated by statistical modeling. On the long run, atmospheric conditions can play vital roles in air pollution situation shifts mainly due to climate change characterized by changes in meteorological values.

**Keywords:** Box-Cox transformation, exponential model, gamma generalized linear model, meteorology, ozone air pollution, particulate air pollution,

## 1 Introduction

Air pollution levels in the ambient air is primarily dependent upon its sources such as vehicular, industrial, domestic fuel combustion, solid waste, etc. Apart from these sources, the temporal variation in air pollution can also be attributed to local topography such as plain terai region, hilly region or mountainous region, weather or meteorological condition assessed by different parameters like temperature, rainfall, humidity and wind speed and direction, regional transport and atmospheric chemistry. Even though many studies have been conducted to assess the dependence of air pollution levels on weather parameters through model building at other parts of the world, such studies conducted in Kathmandu have been very few so far. Because of the lack of adequate number of studies that actually quantified the relationship between air pollution levels with atmospheric conditions based upon local daily data, the present study has been carried out to fulfill the gap which is based upon statistical modeling using local available air pollution and meteorological data. Moreover, air pollution situation in Kathmandu valley has been important environmental and public health concerns to Kathmandu valley inhabitants as shown by many studies conducted previously [2, 3, 5, 8, 14, 18, 19]. According to world air quality report 2019, Nepal is amongst the top 10 air polluted nations in the world, ranked eight as worst polluted as per the population weighted ambient  $PM_{2.5}$  level and Kathmandu ranked sixth worst air polluted among capital cities of the world [7].

Many studies conducted earlier at different places have shown association between air pollution and meteorological parameters. A study conducted at North Chennai, a coastal city in India, during monsoon, post-monsoon, summer, and pre-monsoon seasons for 2010-11 using regression analysis analyzed the influence of temperature and relative humidity on ambient  $SO_2$ ,  $NO_x$ , suspended particulate matter (SPM) and respirable SPM concentrations. The results of the study showed that both  $SO_2$  and  $NO_x$  were negatively correlated in summer and moderately and positively correlated during post-monsoon season with temperature. Weak to moderate correlations existed between the temperature and ambient pollutant concentration during all seasons indicating the influence of inconstant thermal variation in the coastal region. Statistically significant negative correlations were found between humidity and particulates in all the four seasons [9]. Similarly, a study conducted in Dhaka, Bangladesh during 2013-2017 using linear and polynomial regression showed that  $PM_{2.5}$  and  $PM_{10}$  were negatively related with temperature and relative humidity. While relationships of PM and temperature in all other seasons are negative, this research found positive relationship during monsoon season which implied that high temperature with higher humidity during this season contribute to suspension of PM [12]. In January 2013, real-time hourly average concentrations of six pollutants ( $CO$ ,  $NO_2$ ,  $O_3$ ,  $PM_{10}$ ,  $PM_{2.5}$  and  $SO_2$ ) based on data from air quality monitoring stations in major Chinese cities analyzed the air pollution characteristics and their relation to multi-scale meteorological conditions. Meteorological conditions were the primary factor determining day-to-day variations in pollutant concentrations, explaining more than 70% of the variance of daily average pollutant concentrations over China [6]. Moreover, a study conducted in Turkey during 2003-2005 investigated relationship between monitored air pollutant concentrations such as  $SO_2$  and the total suspended particles (TSP) data and meteorological factors such as wind speed, temperature, relative humidity and atmospheric pressure was investigated in months during October-March. According to the results of linear and non-linear regression analysis, it was found that there is a moderate and weak level of relation between the air pollutant concentrations and the meteorological factors [1].

A study conducted in Kano metropolis, Nigeria during 2018 monitored ambient day-time concentration of NO<sub>2</sub>, PM<sub>10</sub>, SO<sub>2</sub>, H<sub>2</sub>S and CO in the months of dry (April) and wet (August) months and corresponding meteorological data were collected from Nigerian Meteorological Agency. Meteorological parameters temperature, relative humidity and precipitation washout or scavenging effect on the pollutants were analyzed quantitatively. The result showed concentration of the pollutants in the atmosphere were lower under condition of increased precipitation, low temperature and increased humidity level compared to that of the dry season [16]. A similar study investigated multi-timescale meteorological effects on the urban air pollution using measurements data of PM<sub>10</sub>, SO<sub>2</sub>, NO<sub>2</sub>, CO, and O<sub>3</sub> and meteorological variables over the period of 1999–2016 in Seoul, South Korea. The long-term air quality data were decomposed into trend-free short-term components and long-term trends and the effects of meteorology and emissions were quantitatively isolated using multiple linear regression with meteorological variables. In terms of short-term variability, inter-correlations among the pollutants and meteorological variables showed warm and stagnant conditions in the migratory high-pressure system are related to high PM<sub>10</sub>, while the strong irradiance and low NO<sub>2</sub> by high winds at the rear of a cyclone are related to the high O<sub>3</sub>. In terms of long-term trends, decrease in PM<sub>10</sub> ( $-1.75 \mu\text{g m}^{-3} \text{yr}^{-1}$ ) and increase in O<sub>3</sub> ( $+0.88 \text{ppb yr}^{-1}$ ) in Seoul were largely contributed by the meteorology-related trends [17].

There have been very few studies that associated air pollution on meteorological parameters in Kathmandu valley. One of the few studies conducted in Kathmandu valley during 2003-2005 found association between meteorological conditions like temperature, rainfall, humidity, atmospheric pressure, wind direction and speed with elemental concentrations of PM<sub>10</sub> in Kathmandu valley. Increase of rainfall, temperature and humidity had negative correlation with average PM<sub>10</sub> concentration in Kathmandu valley ( $r = -0.358$  with rainfall and max. temperature,  $r = -0.539$  with humidity whereas positive correlation with atmospheric pressure ( $r = 0.237$ ) and wind speed ( $r = 0.162$ ) [4]. Similarly, a more recent study conducted by NHRC in 2014/15 associated ambient PM<sub>2.5</sub> with meteorological parameters and found negative association with temperature ( $-0.711$ ), rainfall ( $-0.345$ ) and humidity ( $-0.207$ ) based upon daily average data for one whole year [14].

## **2 Materials and Methods**

### **2.1 Data**

#### **2.1.1 Air pollution**

Air pollution data monitored by Ministry of Population and Environment (MOPE), Department of Environment (DoEnv), Ministry of Forests and Environment and US Embassy was obtained specifically for Kathmandu valley. Data was compiled from websites as follows.

- Daily PM<sub>10</sub> and PM<sub>2.5</sub> data for the years 2017-2020 were compiled from Data Platform of the World Air Quality project <https://aqicn.org/city/kathmandu/> (The World Air Quality Project, 2020).
- Additionally, PM<sub>2.5</sub> and Ozone data for 2017-2020 was compiled from US Embassy website, specifically for US Embassy installed stations (Maharajgunj and Phora Darbar stations). <https://www.airnow.gov/international/us-embassies-and-consulates/> (US Embassies and Consulates, 2020)
- PM<sub>1</sub> data was obtained from Open Data Nepal for 2019 (OPen Data Nepal, 2020).
- Altogether seven stations within Kathmandu valley namely Ratnapark, Shankapark, US Embassy, Phora Darbar, Bhaisipati, Pulchowk and Bhaktapur were incorporated for the analysis. Data for Kirtipur station was unavailable.

#### **2.1.2 Meteorological data**

Daily meteorological data was collected for temperature (maximum and minimum), rainfall, relative humidity, wind speed from Department of Hydrology and Meteorology (DHM), Government of Nepal (GoN), Kathmandu covering three years daily data 2017-early 2020. Data includes eight stations spread over all the three districts of Kathmandu valley mainly for associating air pollution to atmospheric conditions. The stations are Bhaktapur, Nagarkot, Changunarayan, Godavari, Khokana, Khumalatar, Panipokhari and Kathmandu Airport.

### **2.2 Analysis**

Descriptive analysis and subsequent assessment is based upon monthly averages of meteorological parameters and corresponding air pollution averages with graphical representations. Effect quantification of meteorological parameters, seasonality, trend and autoregressive nature of time series variables are explored through statistical models including curvilinear and Gamma generalized linear model (GLM) based upon daily averages.

### **2.3 Models**

Regression models are built to associate air pollution concentration levels on meteorological variables and confounders like seasonality and trend. Additionally, autoregressive term(s) are also explored and added to account the effects of autoregressive effects since the time series data of daily pollution levels can be autoregressive and found true after

computing autocorrelation coefficients at different lags. Various regression models like curvilinear models, Box-Cox transformed models [13] and Gamma GLM are explored for their suitability since pollution concentration levels are found to be highly skewed and non-normally distributed which ruled out suitability of linear regression models. The functional forms of the accounted models are given below.

### 2.3.1 The Exponential Model

$$y = \beta_0 \beta_1^{x_1} \beta_2^{x_2} \dots \beta_k^{x_k} e^\varepsilon$$

where  $y$  is the response variable (Air pollution level),  $\beta_i$  s are unknown parameters,  $x_i$  s are predictors (Meteorological variables and confounders), and  $\varepsilon$  s are residuals. The model can be linearized by log transformation as follows and parameter estimates are found using ordinary least squares.

$$\ln(y) = \ln(\beta_0) + \{\ln(\beta_1)\}x_1 + \{\ln(\beta_2)\}x_2 + \dots + \{\ln(\beta_k)\}x_k + \varepsilon$$

### 2.3.2 Box-Cox Transformed Model

The Box\_Cox transformation is shown below where  $\lambda$  is a constant determined by goodness of fit and model adequacy test results.

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \ln(y), & \lambda = 0 \end{cases}$$

The corresponding model is:  $y^{(\lambda)} = \beta_0 + \beta_1 x_1 \beta_2 x_2 \dots \beta_k + \varepsilon$

Here also  $y$  is the response variable,  $\beta_i$  s are unknown parameters,  $x_i$  s are predictors (Meteorological variables and confounders), and  $\varepsilon$  s are residuals.

### 2.3.3 Gamma Generalized Linear Model

Generalized Linear Model for Gamma distributed dependent variable can be used when the response variable is positive continuous and non-normal positively skewed instead of curvilinear models with variable transformations provided that the model performs relatively superior to diagnostic tests and data transformations are deliberately avoided for difficulty in interpretation. In the model building process, the option of using gamma GLM are also explored since air pollution levels are found non-normal, positively skewed and dependency of error variance on mean is detected even after using curvilinear and Box-Cox transformed models. Additionally, the assumption of homoscedasticity of residuals in linear regression is also relaxed while using gamma GLM [15]. Finally, the model is used if suitable regarding goodness of fit and other major model adequacy tests.

## 3 Results and Discussion

### 3.1 Descriptive analysis

The monthly averages of air pollution levels (PM<sub>10</sub>, PM<sub>2.5</sub> and ozone) and corresponding meteorological parameters in Kathmandu valley during 2017-2020 were assessed. For PM<sub>1</sub> monthly averages were assessed only for 2019 because of data unavailability for the remaining years during 2017-2020 period. The annual averages of temperature, rainfall, humidity and wind are found to be 18.5°C, 1431.5 mm, 77.6% and 1.3 m/s, respectively. The monthly variation shows lowest average temperature (10.6-13.4°C) in winter during when particulate air pollution levels are found to be high, relatively warmer with average ranging between 15.6 to 19.6°C in most spring/autumn months (March, April, October, November) and highest averages ranging between 21.4 to 23.9°C during most of the summer/monsoon months (April-September) during which period the particulate air pollution levels are found to be relatively low which ascertains that temperature and particulate air pollution are inversely related. Similarly, rainfall occurrence is distinctly very high in Monsoon (July-August) with monthly average between 334.4 to 404.2 mm, moderate in May, June and September (116-181 mm) and low in other dry months (1-91 mm) with lowest in November. Considering relative humidity, monthly average shows that the monsoon or around monsoon months (July-September) were the most humid months with average ranging between 80% to 85% and least averages in March and April time (70-72%) with relatively low rainfall during the period. Regarding wind, monthly averages show that the most windy months in Kathmandu valley were from March to June during which period the average wind speed ranged between 1.6-1.8 m/s whereas lowest in winter or around winter time (November-January) with average ranging between 0.7 m/s to 0.95 m/s. Considering particulate air pollution, monthly averages are found to be highest in winter months with approximately 85-99 µg/m<sup>3</sup>, 64-80 µg/m<sup>3</sup> and 36-41 µg/m<sup>3</sup> for

PM<sub>10</sub>, PM<sub>2.5</sub> and PM<sub>1</sub>, respectively. The averages clearly indicate inverse relationship between temperature and particulate air pollution levels. On the contrary, the averages are found to be lowest in summer / monsoon season with approximately 21-23  $\mu\text{g}/\text{m}^3$ , 13-18  $\mu\text{g}/\text{m}^3$  and 5-17  $\mu\text{g}/\text{m}^3$  for PM<sub>10</sub>, PM<sub>2.5</sub> and PM<sub>1</sub>, respectively. However, the pattern of monthly and seasonal variations in Ozone is found to be very different compared to particulate air pollution. The levels are found to be highest in warm temperatures during the months of Spring/Summer (March – June) with values ranging between 67-79  $\mu\text{g}/\text{m}^3$  and lowest during most of the winter time (December-January) with values ranging between 31-33  $\mu\text{g}/\text{m}^3$ .

Table 1 Monthly averages of air pollution and meteorological parameters in Kathmandu valley (2017-2020)

	Temperature (°C)		Rainfall (mm)		Relative Humidity (%)		Wind (m/s)		PM <sub>10</sub> (µg/m <sup>3</sup> )		PM <sub>2.5</sub> (µg/m <sup>3</sup> )		PM <sub>1</sub> * (µg/m <sup>3</sup> )		O <sub>3</sub> (µg/m <sup>3</sup> )	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
January	10.62	1.20	17.6	1.79	74.47	5.31	0.94	0.29	90.82	29.19	80.11	29.34	37.93	14.87	30.97	9.79
February	13.39	1.61	27.2	4.05	73.25	5.46	1.37	0.45	99.23	27.18	62.85	32.17	24.13	10.85	37.82	19.07
March	16.07	2.06	52.4	4.17	69.55	8.53	1.80	0.47	81.41	34.41	52.12	31.60	22.55	11.25	67.74	17.94
April	19.35	1.35	91.1	4.33	71.94	7.97	1.64	0.66	92.42	111.70	47.06	29.79	21.63	13.72	77.12	22.56
May	21.42	1.46	116.1	5.98	74.08	7.53	1.65	0.50	47.02	40.45	37.75	20.42	21.81	12.01	79.48	20.94
June	23.67	0.89	175.3	7.42	79.21	5.14	1.58	0.42	20.86	20.02	21.00	14.37	9.30	5.63	67.34	23.02
July	23.78	0.98	404.2	15.94	85.34	3.81	1.43	0.71	23.03	23.27	13.95	9.70	17.74	23.80	36.78	17.87
August	23.88	0.66	334.4	10.05	84.95	3.23	1.38	0.77	21.43	22.72	13.00	8.38	5.43	3.78	32.58	14.84
September	22.93	1.19	181.4	8.40	84.51	3.12	1.23	0.44	32.29	28.97	18.27	10.43	5.44	3.15	33.38	18.15
October	19.63	2.22	20.8	1.45	79.50	4.12	1.02	0.52	53.07	56.82	31.93	17.97	13.54	8.21	42.12	20.65
November	15.61	1.51	0.9	0.17	77.64	4.22	0.76	0.27	96.94	63.90	54.14	23.80	35.90	15.55	39.99	10.05
December	11.47	1.51	10.3	2.99	76.05	5.12	0.70	0.35	84.66	52.36	64.02	27.98	41.44	10.20	32.89	7.04
Overall	18.51	4.93	1431.5	8.10	77.57	7.53	1.29	0.62	66.72	59.00	42.40	31.17	21.58	16.55	50.38	25.89

\* For PM<sub>1</sub> only data for 2019 is used; SD = Standard deviation

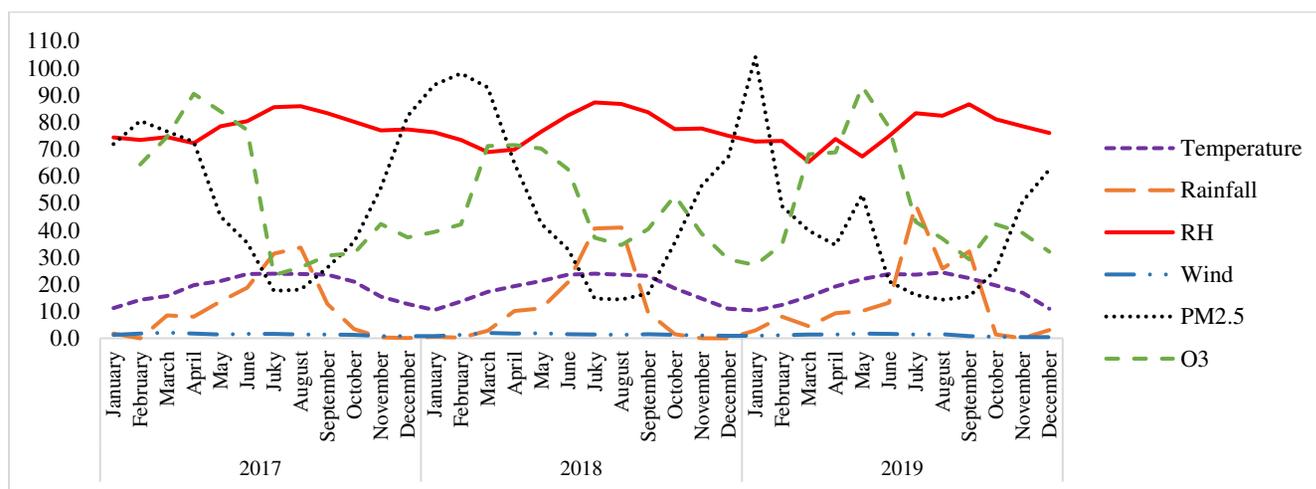


Fig. 1 Monthly statistics of PM<sub>2.5</sub>, O<sub>3</sub> and meteorological parameters; Temperature in °C, Rainfall in mm; RH in %, Wind in m/s; PM<sub>2.5</sub> and O<sub>3</sub> in µg/m<sup>3</sup>

### 3.2 Models

Air pollution level measured by PM<sub>2.5</sub> and PM<sub>10</sub>, PM<sub>1</sub> and O<sub>3</sub> are modeled on different predictors including meteorological parameters, seasonal effects and a daily trend variable. Additionally, lag effects are also explored since data for modeling are essentially time series. Data exploration of air pollution levels showed significant positive skewness for all the parameters considered which suggested suitability of curvilinear, models with transformation, nonlinear models including Gamma generalized linear model (GLM) rather than the linear models since normality assumption of the air pollution levels cannot be accepted for substantially skewed variables. Exponential model with logarithmic transformation, Box-Cox transformed model (response variable with Box-Cox transformation) and gamma GLM was explored for their suitability and the model which is found to be the best among them was chosen for modeling considering various model adequacy tests including goodness of fit assessed by adjusted R<sup>2</sup> or Omnibus test, heteroscedasticity by residual plot, normality by Kolmogorov-Smirnov (KS) test, autocorrelation by plots up to sufficient lag, and multicollinearity by Variance Inflation Factor (VIF). Models with estimated parameters with 95% confidence interval and *p* values are shown in Table 2(A) and Table 2(B). Moreover, the model adequacy test results are also shown. Exponential model, Box-Cox transformed model, and Gamma GLM are found to be suitable for explaining variation of PM<sub>10</sub>, PM<sub>2.5</sub> and Ozone, and PM<sub>1</sub>, respectively. Considering the Box-Cox transformed models, the values of the parameters ( $\lambda$ ) found relatively better for PM<sub>2.5</sub> and O<sub>3</sub> are 0.333 and 0.75, respectively which yielded better goodness of fit and other model adequacy tests.

Table 2 (A) PM<sub>2.5</sub> and PM<sub>10</sub> model estimates

Parameter	PM <sub>2.5</sub>				PM <sub>10</sub>			
	Box-Cox Transformed Model for $\lambda = 0.333$				Exponential Model			
	Coefficient	Sig.	Lower Bound	Upper Bound	Coefficient	Sig.	Lower Bound	Upper Bound
Constant	8.997	0.000	7.963	10.031	5.882	0.000	5.234	6.531
Temperature	-0.083	0.000	-0.108	-0.058	-0.051	0.000	-0.069	-0.034
Rainfall	-0.032	0.000	-0.040	-0.024	-	-	-	-
Relative Humidity	-0.026	0.000	-0.036	-0.016	-0.019	0.000	-0.025	-0.013
Wind	-0.148	0.008	-0.257	-0.038	-0.073	0.095	-0.159	0.013
Trend	-0.001	0.000	-0.001	0.000	0.000	0.017	0.000	0.000
Spring	0.594	0.000	0.377	0.810	0.597	0.000	0.412	0.781
Autumn	0.176	0.075	-0.018	0.370	0.428	0.000	0.294	0.561
Winter	0.278	0.095	-0.048	0.603	0.166	0.139	-0.054	0.386
LAG	0.049	0.000	0.046	0.052	0.006	0.000	0.005	0.007
Adjusted R <sup>2</sup>	0.852				0.791			
Autocorrelation	-0.12 < r < 0.12				-0.08 < r < 0.2			
Constant variance	More or less constant				More or less constant			
Normality	Slightly non-normal				Slightly non-normal			

Table 2 (B) PM<sub>1</sub> and O<sub>3</sub> model estimates

Parameter	PM <sub>1</sub> Gamma GLM				O <sub>3</sub> Box-Cox Transformed Model for $\lambda = 0.75$			
	Coefficient	Sig.	Lower Bound	Upper Bound	Coefficient	Sig.	Lower Bound	Upper Bound
Constant	3.504	.000	2.990	4.019	19.731	.000	17.041	22.420
Temperature	-.046	.000	-.063	-.029	.177	.000	.126	.228
Relative Humidity	-.007	.028	-.013	-.001	-.183	.000	-.219	-.147
Spring					1.716	.000	1.048	2.384
Autumn	-.148	.019	-.272	-.025				
Winter	-.354	.000	-.543	-.166				
LAG 1	.034	.000	.027	.041	.279	.000	.267	.292
LAG 2	.006	.087	-.001	.013				
Omnibus test/R <sup>2</sup>	Significant at 1% level				0.85			
Autocorrelation	-0.06 < r < 0.30				-0.03 < r < 0.1			
Constant variance	Slightly non-constant				Fairly constant			
Normality	Slightly non-Normal				Slightly-normal			
Multicollinearity	Absent (VIF < 5)				Absent (VIF ≤ 2)			

### 3.3 Effects

#### 3.3.1 Temperature

Effects of temperature on pollution level has been found most evident amongst the predictors for all air pollutants and statistical models explored with 1°C increase in average temperature found associated with 5.1% and 4.6% decrease in PM<sub>10</sub> and PM<sub>1</sub> levels, respectively. Similarly, 1°C increase in average temperature is also found associated with 0.083 decrease in Box-Cox transformed unit of PM<sub>2.5</sub> and in contrast, 0.177 increase in Box-Cox transformed unit of Ozone. The negative association between temperature and particulate air pollutants demonstrates cold weather mainly in winter increases particulate air pollution in the ambient air significantly compared to warm atmospheric condition mainly because particulate pollutants are trapped near the ground during colder, calmer months due to temperature inversion. During a temperature inversion, smoke and dust particles are difficult to rise and disperse in the atmosphere which is very much evident in a place like bowl shaped Kathmandu valley characterized by low wind flow. On the contrary, ground level ozone is found relatively higher in warm temperature compared to cold temperature primarily because pollutants emitted by vehicles and industries and other sources chemically react in the presence of sunlight producing Ozone and therefore is most likely to reach unhealthy levels during hot sunny days in urban environments.

#### 3.3.2 Rainfall

Rainfall is found to be statistically associated with PM<sub>2.5</sub> level with increase in 1cm rainfall decreases 0.32 PM<sub>2.5</sub> expressed in Box-Cox transformed value which indicates that rainfall decreases air pollution concentration in the ambient air. Except for PM<sub>2.5</sub>, strikingly rainfall is found statistical insignificant even at 15% level for the rest of the air pollutants considered. This may be primarily due to the effect of multicollinearity among the meteorological parameters when all of them are included in modeling as monsoon season with high rainfall is characterized by warm temperatures and high relative humidity compared to winter season, for instance. Otherwise, if monthly averages of particulate air pollution are assessed then it is found that the pollution levels are least during monsoon time including Kathmandu valley which is a strong evidence that rainfall washes away dust particles from air. Nevertheless, since other major parameters like temperature and relative humidity are found statistically significant in models with particulate air pollution levels, rainfall effect was not found statistically significant.

#### 3.3.3 Relative Humidity

Similar to the temperature effect on air pollution levels, relative humidity is also found to be statistically associated with air pollution levels in all the four considered models. Examining the direction and magnitude of effects, it is found that increase in relative humidity decreases ambient air pollution levels with 1% increase in relative humidity is found to be associated with 1.9% and 0.7% decrease in PM<sub>10</sub> and PM<sub>1</sub> levels, respectively. Similarly, 1% increase in relative humidity is also found associated with 0.026 and 0.183 decrease in Box-Cox transformed unit of PM<sub>2.5</sub> and Ozone, respectively. The reduction in Ozone level associated with increase in relative humidity has also been found in other studies also [10, 11].

#### 3.3.4 Wind

Wind is another important weather parameter that effects air pollution level significantly as shown by many studies (NHRC, 2016). Wind disperses air contaminants away from their source and therefore, generally, higher wind is found associated with lower air pollution concentration. The present study has also found negative association between PM<sub>10</sub> and PM<sub>2.5</sub> with wind speed with 7.3% and 0.148 Box-Cox transformed value of PM<sub>10</sub> and PM<sub>2.5</sub> levels per 1 m/s increase in wind speed, respectively.

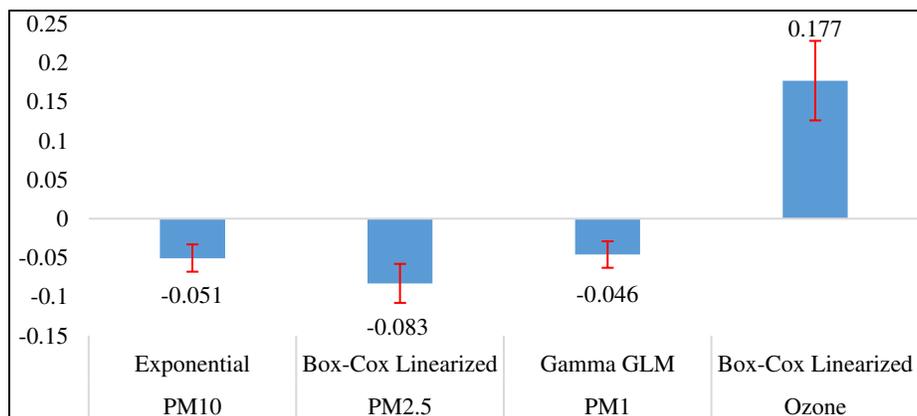


Fig. 2 Change in pollution level per 1°C increase in temperature

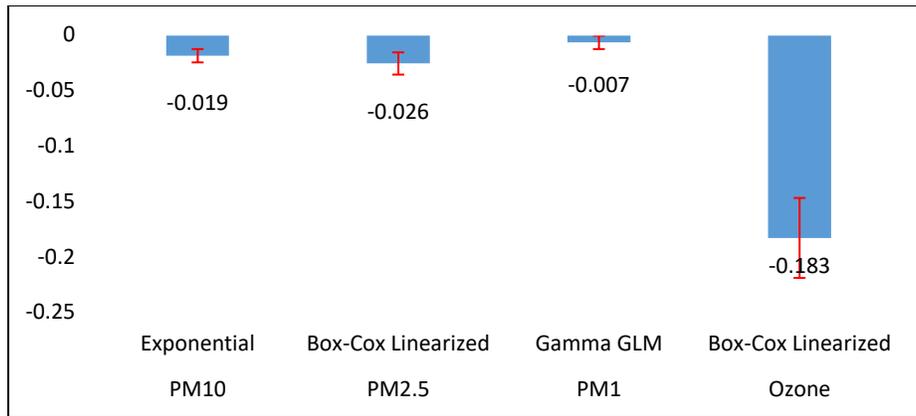


Fig. 3 Change in pollution level per 1% increase in relative humidity

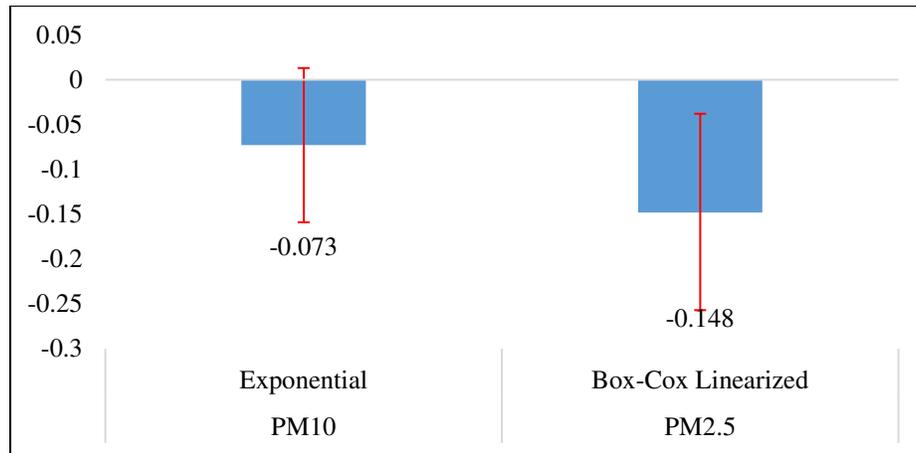


Fig. 4 Change in pollution level per 1 m/s increase in Wind

### 3.3.5 Seasonal effects

Along with atmospheric parameters, seasonality can be major contributing factor on variation in pollution levels. With several seasonal features characterized by joint effects of meteorological parameters, air pollution levels tend to differ significantly for different seasons. Exploration of this through modeling, it is found that Winter season characterized by cool temperature, lower relative humidity and wind flows showed that PM<sub>10</sub> and PM<sub>2.5</sub> increased by 16.6% and 0.278 Box-Cox transformed value compared to Summer. Similarly, Spring season characterized by dry and relatively cool conditions showed that PM<sub>10</sub>, PM<sub>2.5</sub> and O<sub>3</sub> increased by 59.7%, 0.594 and 1.716 Box-Cox transformed values compared to Summer, respectively. Similar results are obtained for Autumn season for PM<sub>10</sub> and PM<sub>2.5</sub> compared to Summer, except for PM<sub>1</sub> where the effect seems to be different which could be due to different season specific joint effect of Autumn compared to Spring for PM<sub>1</sub>. For Ozone, the effect due to Autumn season is found statistically insignificant.

### 3.3.6 Lag effects

Models depending upon time series data are often affected by lagged variables due to which auto-regressive terms may be required to explain variation in the response variable while modeling major predictors. In the present model building process, this has been explored and is found that the first and second lagged values of the responses are found statistically significant in explaining variation of the air pollution levels. All the particulate air pollution parameters including O<sub>3</sub> are found to be affected by their own values at Lag1 and are statistically significant with positive association at 1% level. For PM<sub>1</sub>, additionally Lag2 values is found statistically significant.

### 3.3.7 Model adequacy tests

The goodness of fit, normality, heteroscedasticity, multi-collinearity and autocorrelation checks were performed under model adequacy test requirements of acceptance of the fitted models. Models performed good as regards to goodness of fit test with 79% - 85% of the variance in the response variables explained in curvilinear models or Omnibus Test

resulting highly significant for Gamma GLM. Regarding, normality and heteroscedasticity, residuals are found slightly non-normal and variances fairly constant. In order to obtain normal residuals, other curvilinear/nonlinear models were also explored but residuals were still found non-normal or slightly non-normal. Consequently, the fitted models are accepted and indicate that further researches could be required to achieve relatively more accurate modeling results. Considering autocorrelation, the models are found not much affected by high autocorrelations ( $-1 < r < 0.3$ ) considering high sample size for data modeling.

#### **4 Conclusion**

Though the major sources that govern the overall average of air pollution in a local environment are emissions from anthropogenic sources like vehicular, industrial, domestic, solid waste and others, natural characteristics like local topography, atmospheric conditions (also caused by human activities) also determine air pollution levels and its temporal variations. The present analysis focused on the atmospheric conditions determined by meteorological parameters like temperature, rainfall, relative humidity and wind. Exploring the dependency of temporal air pollution variation through statistical models including curvilinear and nonlinear models namely exponential, Box-Cox transformed, and gamma GLM revealed statistically significant associations between air pollution levels and the meteorological parameters with address of time series affected confounding variables such as seasonality and autoregressive dependence (lagged effect) with high proportions of variance in air pollution levels explained (79% to 85%). With the lack of adequate number of studies that quantified the effects in Kathmandu valley based upon local data, the present analysis would be useful in assessing quantification of meteorological effects on air pollution level. Results showed around 5% reduction in particulate air pollution ( $PM_{10}$  and  $PM_1$ ) per  $1^\circ C$  increase in average temperature and significant increase in surface  $O_3$  air pollution (0.177 Box-Cox transformed value) per  $1^\circ C$  increase in average temperature. Similarly, around 0.7% and 2% decrease in  $PM_1$  and  $PM_{10}$  per 1% increase in relative humidity and 7.3% decrease in  $PM_{10}$  per 1 m/s increase in wind speed are also detected. Other effects are also quantified in terms of Box-Cox transformed values for statistically significant effects due to rainfall, relative humidity and wind. In conclusion, meteorological conditions are significant contributing factors in determining air pollution levels as demonstrated by statistical modeling of local data in Kathmandu valley. On the long run, atmospheric conditions can play vital roles in air pollution situation shifts mainly due to climate change characterized by changes in meteorological parameter values.

#### **Author Declarations**

**Funding:** The paper is based upon the study funded by WHO under urban health initiative (UHI) program implemented in Kathmandu. WHO reference no. of the study is 202576053.

**Conflict of Interests / Competing Interests:** The author declare that there is no conflict of interests. The author has no relevant financial or non-financial interests to disclose. The author has no competing interests to declare that are relevant to the content of this article.

**Ethical approval:** The study does not involve human or animal participant data. Analysis and modeling are based upon secondary data. Available air pollution data of Kathmandu valley was acquired from DataPlatform, World Air Quality project. Upon request, data was sent by the provider through email upon the conditions that the data will only be used for the WHO study and need of acknowledgement of the data provider. Remaining air pollution data was also downloaded from AirNow website of US Embassy and Open Data portal of Nepal, both of which provide freely downloadable data access. Similarly, meteorological data was obtained (purchased) from the Department of Hydrology and Meteorology (DHM), Kathmandu upon formal request and fulfilling conditions required for its use.

#### **Data availability/transparency**

Monthly meteorological and air pollution data are provided in the manuscript itself. Air pollution monitoring data is available data from related websites as mentioned in ethical approval section and results can be verified from the websites. Author is neither entitled nor permitted to supply raw data except its use for analysis of the study.

**Code availability:** Not applicable since raw data is not entitled to be supplied.

**Authors' Contribution:** The paper has sole author. From collection of data, designing the study, analyzing, interpretations, manuscript writing, including others are done by the author. The submission is approved by the author.

## References

1. Akpınar, E. K., Akpınar, S., & Öztıp, H. F. (2009). Statistical analysis of meteorological factors and air pollution at winter months in Elazığ, Turkey. *Journal of Urban and Environmental Engineering*, 3(1), 7-16.
2. Central Bureau of Statistics. (2019). *Environment Statistics of Nepal 2019*. Kathmandu: CBS. <https://cbs.gov.np/environment-statistics-of-nepal-2019/>. Accessed on May 12, 2021.
3. Department of Environment (2017). *Air Quality Management Action Plan for Kathmandu Valley*. Kathmandu: Ministry of Population and Environment. <http://doenv.gov.np/public/uploads/Pdffile/Final%20Report%20on%20AQM%20Action%20Plan%202017-42479.pdf>. Accessed on May 20, 2021.
4. Giri, D., Krishna Murthy, V., & Adhikari, P. R. (2008). The influence of meteorological conditions on PM10 concentrations in Kathmandu Valley. *Int. J. Environ. Res.*, 2(1), 49-60.
5. Gurung, A., & Bell, M. L. (2012). Exposure to airborne particulate matter in Kathmandu Valley, Nepal. *Journal of Exposure Science and Environmental Epidemiology*, 22, 235-242. <https://doi.org/10.1038/jes.2012.14>
6. He, J., Gong, S., Yu, Y., Yu, L., Wu, L., Mao, H., Song, C., Zhao, S., Liu, H., Li, X., & Li, R. (2013). Air pollution characteristics and their relation to meteorological conditions during 2014-2015 in major Chinese cities. *Environmental Pollution*, 223(April 2017), 484-496. <https://doi.org/10.1016/j.envpol.2017.01.050>
7. IQAir. (2019). *2019 World air quality report*. IQ Air. Retrieved June 5, 2021. <https://www.iqair.com/world-most-polluted-cities/world-air-quality-report-2019-en.pdf>
8. Islam, M. R., Jayarathne, T., Simpson, I. J., Werden, B., Maben, J., Gilbert, A., . . . , & Stone, E. A. (2020). Ambient air quality in the Kathmandu Valley, Nepal, during the pre-monsoon: concentrations and sources of particulate matter and trace gases. *Atmos. Chem. and Phys.*, 20, 2927-2951. <https://doi.org/10.5194/acp-20-2927-2020>
9. Jayamurugan, R., Kumaravel, B., Palanivelraja, S., & Chockalingamm M. P. (2013). Influence of temperature, relative humidity and seasonal variability on ambient air quality in a coastal urban area. *International Journal of Atmospheric Sciences*, 2013(264046). <https://doi.org/10.1155/2013/264046>
10. Jia, L., & Xu, Y. (2014). Effects of relative humidity on ozone and secondary organic aerosol formation from the photooxidation of benzene and ethylbenzene. *Aerosol Science and Technology*, 48(1), 1-12. <https://doi.org/10.1080/02786826.2013.847269>
11. Kavassalis, S. C., & Murphy, J. G. (2017). Understanding ozone-meteorology correlations: A role for dry deposition. *Geophysical Research Letters*, 44(6), 2922-2931. <https://doi.org/10.1002/2016GL071791>
12. Kayes, I., Shahriar, S. A., Hasan, K., Akhter, M., Kabir, M. M., & Salam, M. A. (2019). The relationships between meteorological parameters and air pollutants in an urban environment. *Global J. Environ. Sci. Manage*, 5(3): 265-278. <https://doi.org/10.22034/gjesm.2019.03.01>
13. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis*. John Wiley and Sons Inc.
14. Nepal health Research Council. (2016). *Situation Analysis of Ambient Air Pollution and Respiratory Health Effects in Kathmandu Valley*. Kathmandu: NHRC. <http://nhrc.gov.np/wp-content/uploads/2017/06/ambient-air-pollution-book.pdf>. Assessed on March 14, 2021
15. Ng, V., & Cribbie, R. A. (2016). Using the gamma generalized linear model for modeling continuous, skewed and heteroscedastic outcomes in psychology. *Current Psychology*, 36(2), 225-235. <https://doi.org/10.1007/s12144-015-9404-0>
16. Oji, S., & Adamu, H. (2020). Correlation between air pollutants concentration and meteorological factors on seasonal air quality variation. *Journal of Air Pollution and Health*, 5(1), 11-32.
17. Seo, J., Park, D. R., Kim, J. Y., Youn, D., Lim, Y. B., & Kim, Y. (2018). Effects of meteorology and emissions on urban air quality: a quantitative statistical approach to long-term records (1999-2016) in Seoul, South Korea. *Atmos. Chem. Phys.*, 18, 16121-16137. <https://doi.org/10.5194/acp-18-16121-2018>, 2018
18. Shrestha, S. L. (2007). Time series modelling of respiratory hospital admissions and geometrically weighted distributed lag effects from ambient air pollution within Kathmandu valley. *Environ Model Assess*, 12(3): 239-251. <https://doi.org/10.1007/s10666-006-9071-5>
19. Shrestha, S. L. (2012). Particulate air pollution and daily mortality in Kathmandu valley, Nepal: Associations and distributed lag. *The Open Atmospheric Journal*, 6, 62-70. <https://doi.org/10.2174/1874282301206010062>