LINKING AIR QUALITY TO METEOROLOGY: A MULTILINEAR REGRESSION APPROACH

Sadia Afrin*[1] and Mohammad Maksimul Islam[1]

[1]*Assistant Professor, Bangladesh University of Engineering and Technology, Bangladesh, e-mail: sadiaafrin9888@gmail.com*

***Corresponding Author**

ABSTRACT

Dhaka, the capital of Bangladesh, is one of the most polluted cities in the world in terms of air quality. Particulate matter (PM) concentrations of the city frequently exceed the national air quality standards. Although meteorological parameters play a vital role in air quality through downwind transportation and formation of the secondary PM, few studies have been conducted so far to explore the association between meteorology and air quality. In this paper, we apply a multiple linear regression approach to build a statistical model with PM as response variable, and daily mean temperature, precipitation, relative humidity, solar radiation, wind speed, and wind direction components as potential predictors. For analysis, we employ 24-hour average fine ($PM_{2.5}$) and coarse ($PM_{2.5-10}$) particulate matter concentration and meteorological data for the year 2002 to 2004 as this work expands the work by Islam, Afrin, Ahmed, & Ali . The result depicts that meteorological variables can explain 57% and 35% variability in the 24-hour average $PM_{2.5}$ and $PM_{2.5-10}$ concentration respectively when only the direct influence of the parameters is considered. After the inclusion of interaction terms among the parameters, the $PM_{2.5}$ model performances improve by 12% although there is no improvement for coarser fraction. The relative humidity is the most dominating factor explaining 72% of the total variability explained by the $PM_{2.5-10}$ model. Under a humid environment, there is a reduction of coarser particles as the settling and wet deposition of the particles is fostered by high moisture content. Oppositely for $PM_{2.5,}$ temperature, wind speed, and wind direction are the most influential parameters. They altogether explain 94% of the total variability explained by the $PM_{2.5}$ model. Similar to relative humidity, temperature has an inverse relationship with both PM fractions due to the radiative cooling by particles and also for the shutdown of specific PM sources during the non-winter period. Overall, the outcome of this study provides deeper insight regarding the influence of meteorology and their interaction on PM concentrations. The statistical approach developed in the paper is powerful to develop PM forecasting model for predicting the next day PM given the forecasted meteorological data.

Keywords: *Particulate matter, Meteorology, Multivariate, Regression, Variability.*

1. INTRODUCTION

According to the 'Global Burden of Disease' study, air pollution is the dominant driver of global mortality compared to other different types of pollution including water, soil, chemical, and metals, occupational (Cohen et al., 2017). Every year, outdoor and indoor air pollution attributes to around 8 million premature deaths globally (WHO, 2016) and for Bangladesh this number is around 15000 (Mahmood, 2011). Among different pollutants, particulate matter with aerodynamic diameter less than 10 µm (PM_{10}) are more strongly associated with negative health outcomes. A good number of studies have provided evidence of strong positive association between PM concentrations and all-cause mortality, respiratory morbidity, possibly cardiovascular mortality, morbidity (Bae & Hong, 2018; Zheng, Pozzer, Cao, & Lelieveld, 2015). In Bangladesh, concentrations of these particulate matters frequently exceed the Bangladesh national ambient air quality standard (NAAQS) and U.S. NAAQS, especially during winter months (Hossain & Easa, 2012; Rahman, Mahamud, & Thurston, 2019). During this period, PM concentration remains 5 to 6 times higher than the non-winter monsoon concentration. Specifically, PM is the most concerning air pollutant for Dhaka city (Majumder, Sihabut, & Saroar, 2019; Rahman et al., 2019), which is the capital of Bangladesh. According to WHO (2016), Dhaka city is the third most polluted city in the world in terms of poor air quality. It is also one of the most densely populated cities in the world. The current population of Dhaka is around 20 million. This huge population of the city is constantly exposed to this high concentration of PM mostly due to unplanned urbanization, industrialization, and rapid population growth. A recent study by Tasmin et al. (2019) found association of Short-term exposure to PM with worse lung function of school children in Dhaka experiencing high pollution. Surrounding brick kilns operation, Fossil fuel combustion, industrial operation, vehicular emission, re-suspension of dust from unpaved roads and soil, metal smelter, fugitive Pb, Zn sources are the primary source of particulate matter in this city (Begum, Nasiruddin, Randal, Sivertsen, & Hopke, 2014; Islam et al., 2015).

Given the poor air quality of the city and elevated level of particulate matter concentration in the ambient air, a detailed understanding of the dynamics and drivers of air pollution has become a critical research area to explore. However, limited studies have been conducted so far addressing this issue due to the lack of concern and data availability as the city does not have any well-distributed network of air monitoring. Among the existing literature, studies mostly focused on trend analysis (B. A. Begum & Hopke, 2018; Rahman et al., 2019), source apportionment (Begum et al., 2014; Salam, Hossain, Siddique, & Shafiqul Alam, 2008) and seasonal variability assessment of the air pollutants (Islam & Afrin, 2014; Islam et al., 2015). Although the metrological parameter has been proved to be a significant driver of air pollution (Leung et al., 2017; Tai, Mickley, & Jacob, 2010) very few studies have explored the potential impact of meteorology on the air quality of Dhaka city (Afrin, Islam, Ahmed, & Ali, 2014; Islam et al., 2015; Islam, Saroar, & Ahmed, 2018; Kayes et al., 2019). Kayes et al. (2019) applied multiple linear and non-linear regression models to explore role of meteorological parameters on both particulate and gaseous air pollutants. However, this study was limited to only three meteorological variables – temperature, relative humidity, and rainfall. Islam et al. (2015) determined bivariate 'Pearson correlation coefficient' between meteorological parameters and air pollutants at the Shangshad Bhaban site. This study considered solar radiation and wind speed in addition to the other variables considered by Kayes et al. (2019). In addition, Islam et al. (2018) assessed the effect of meteorological parameters on seasonal variation of particulate matter on a different site located at Darus-Salam, Dhaka using multilinear regression and cross-correlation approach between PM_{10} and $PM_{2.5}$.

None of the previous studies considered the role of wind speed in their regression models. However, the literature suggests a strong association of wind direction with high concentrations of PM components (Leung et al., 2017; Tai et al., 2010). In this paper, we included wind speed and direction components in our regression model, which is an extension of our previous analysis (Islam et al., 2015), which was based on bivariate correlation analysis. In this new work, we employed stepwise multivariate regression analysis to explain the influence of meteorological parameters on the overall

PM concentration of Dhaka city in order to give a better insight regarding the drivers of air pollution. The unique feature of the stepwise regression method keeps only the significant predictors in the final model. Our model also considers the interaction between meteorological variables to assess the combined impact of different types of meteorological variables. Additionally, we quantified the relative importance of the different significant meteorological parameters affecting the 24-hour variation of the ambient particulate matter concentration.

2. METHODOLOGY

2.1 Study Area and Data Source

Dhaka is located in the center of the country. This city has a tropical wet and dry climate with an annual mean temperature of 25°C. Compared to the other major cities of the county, air quality of Dhaka city is worse. Hence, this paper focuses on the air quality of Dhaka city. We collected 24-hour average PM concentration data from the 'Shangshad Bhaban' Continuous Air Monitoring Station (CAMS). This is an urban air monitoring site located close to the heavily trafficked Rokeya Sharani (~150 meters) and Manik Mia Avenue (~300 meters) and started operation in 2002. This site measured the 24-hour average $PM_{2.5}$ (particles with aerodynamic diameter < 2.5 µm) and PM_{10} (particles with aerodynamic diameter < 10 µm) concentrations using high volume PM samplers for the period April 2002 to May 2004. After that period, PM concentrations were reported as monthly average for this site. Hence, we consider the 24-hour average particle concentration data from April 2002 to May 2004 in our analysis. To keep consistency with our previous analysis paper (Islam et al., 2015), we split the total PM_{10} concentrations into two size fractions, the finer fraction with $PM_{2.5}$ and the coarser fraction named $PM_{2.5\text{-}10}$ consisting particles' size ranging in between 2.5 µm and 10 µm. In order to assess the influence of meteorological parameters over air quality we consider the concurrent daily average atmospheric temperature, rainfall, relative humidity, and solar radiation recorded at the same site. For daily wind speed and prevailing wind direction, we used the data recorded by Bangladesh Meteorological Department (BMD) as 'Shangshad Bhaban' CAMS did not report daily wind speed and direction for that period. On the other hand, BMD monitors daily data from all meteorological parameters since 1953 and located within 4km from the CAMS site. The meteorological variables included in this study are listed in Table 1.

Table 1: List of meteorological parameters

Parameter	Abbreviation	Unit
Temperature	T	°C
Rainfall	R	cm
relative humidity	RH	%
solar radiation	SR	W/m^2
wind speed	WS	m/s
North-South component of wind direction	CosWD	-
East-West component of wind direction	SinWD	-

2.2 Statistical Analysis

This paper used the stepwise regression method to develop generalized linear models considering multiple variables for the monitoring site. We use the 24-hour average particulate matter concentration (both $PM_{2.5}$ and $PM_{2.5\text{-}10}$) as response variables and all the listed parameters in Table 1 as predictive variables. In this stepwise regression approach, the choice of predictive variables was carried out considering both forward and backward selections of variables. The final model only retains the significant predictors based on the p-value for the F or chi-squared test of the change in the deviance by adding or removing the term. In our analysis, we define a significance level of $P < 0.05$ and also the upper and lower model type. To consider the independent impacts by the individual predictors we set the upper model types as 'linear', while set the upper model type as 'interaction' to incorporate the interaction between meteorological variables. We used the adjusted R^2 value instead of the ordinary R^2 value to explain the model's predictive power, as adjusted R^2 gives an unbiased estimate of the population R^2. We simulate two sets of stepwise generalized linear models by including and excluding the interaction terms among the predictor variables. For each case, we regenerated the model for both $PM_{2.5}$ and $PM_{2.5\text{-}10}$. The schematic diagrams of the two sets of models considering all probable predictors are shown in equation (1) and equation (2).

$$PM = \beta 0 + \sum_1^k \beta_k X_k \qquad (1)$$

$$PM = \beta 0 + \sum_1^k \beta_k X_k + inetraction\ terms \qquad (2)$$

Where, PM could be either 24-hour average $PM_{2.5}$ or $PM_{2.5\text{-}10}$ during the study period considered, β0 is the intercept term or β1 to β7 are the regression coefficients for each of the seven independent meteorological variables. Finally, we applied dominance analysis (Azen & Budescu, 2003) to assess the relative importance of different predictors in a regression model. Following this approach, we also quantified the shared variance by individual predictors to the total variance or explaining power of a model.

3. RESULTS AND DISCUSSION

Considering only the linear terms of all meteorological parameters as per equation (1), the stepwise regression model can explain 57% and 35% variability of the 24-hour average $PM_{2.5}$ and $PM_{2.5\text{-}10}$ concentrations of the ambient air respectively (Figure 1). In this case of excluding interaction terms, temperature, relative humidity, and wind parameters are the most significant predictors for the $PM_{2.5}$ model. In contrast, the wind parameter does not have any significant influence in predicting the $PM_{2.5\text{-}10}$ concentrations during the study period considered. The estimated intercept, regression coefficients and other model parameters for both $PM_{2.5}$ and $PM_{2.5\text{-}10}$ models are presented in Table 2 and Table 3 respectively.

Table 2: Significant parameters of the $PM_{2.5}$ regression model considering only the linear terms of all meteorological parameters as predictor variables

	Estimate	**Standard Error**	**t-Statistics**	**P-Value**
Intercept	416.86	34.10	12.23	2.2E-28
T	-8.39	0.75	-11.26	6.0E-25
RH	-1.20	0.30	-3.95	9.7E-05
WS	-11.41	3.24	-3.52	4.9E-04
SinWD	-10.08	5.13	-1.96	5.0E-02
CosWD	11.98	4.17	2.87	4.4E-03

Table 3: Summary statistics/ Parameters of the $PM_{2.5\text{-}10}$ regression model considering only the linear terms of all meteorological parameters as predictor variables

	Estimate	**Standard Error**	**t-Statistics**	**P-Value**
Intercept	298.80	18.06	16.55	8.3E-45
T	-2.51	0.38	-6.62	1.6E-10
RH	-2.27	0.20	-11.57	4.5E-26

As shown in Table 2, temperature, relative humidity, wind speed has a negative association with 24-hour average $PM_{2.5}$ concentration. Although wind speed coming from the typical wind direction for a specific season showed weak correlation with particle concentrations (Islam et al., 2015), their inclusion in the regression model along with wind direction components increases the explaining power of the $PM_{2.5}$ regression model. These findings also highlight the potential role of wind direction as a model predictor, which has not been previously explored by previous studies. The role of temperature, relative humidity obtained in this study is in agreement with the results obtained by the previous literature (Islam et al., 2015; Islam et al., 2018; Kayes et al., 2019). Likewise the negative association of wind speed obtained in this study, another study (Islam et al., 2018) based on the data of Darus Salam CAMS also found strong negative correlation with wind speed.

The positive association with the North-South component and negative association with the east-west component suggests that $PM_{2.5}$ concentrations increase when winds come from north and west directions. This makes a sense as wind is transporting particles from the brick-kiln clusters located in the north-west of the city (Islam & Afrin, 2015). Overall, from the t-statistics, it can be said that temperature is the most influencing meteorological parameter among the all other predictors. For the $PM_{2.5\text{-}10}$ regression model, only temperature, relative humidity is the significant predictors (Table 3). Opposite to $PM_{2.5}$ model, relative humidity has a higher impact on the total variability of the $PM_{2.5\text{-}10}$ model than temperature. With the increase in moisture content the coarser particles become more susceptible to the different physical processes within the atmosphere. Consequently due to settling and wet deposition, the particle concentration reduces.

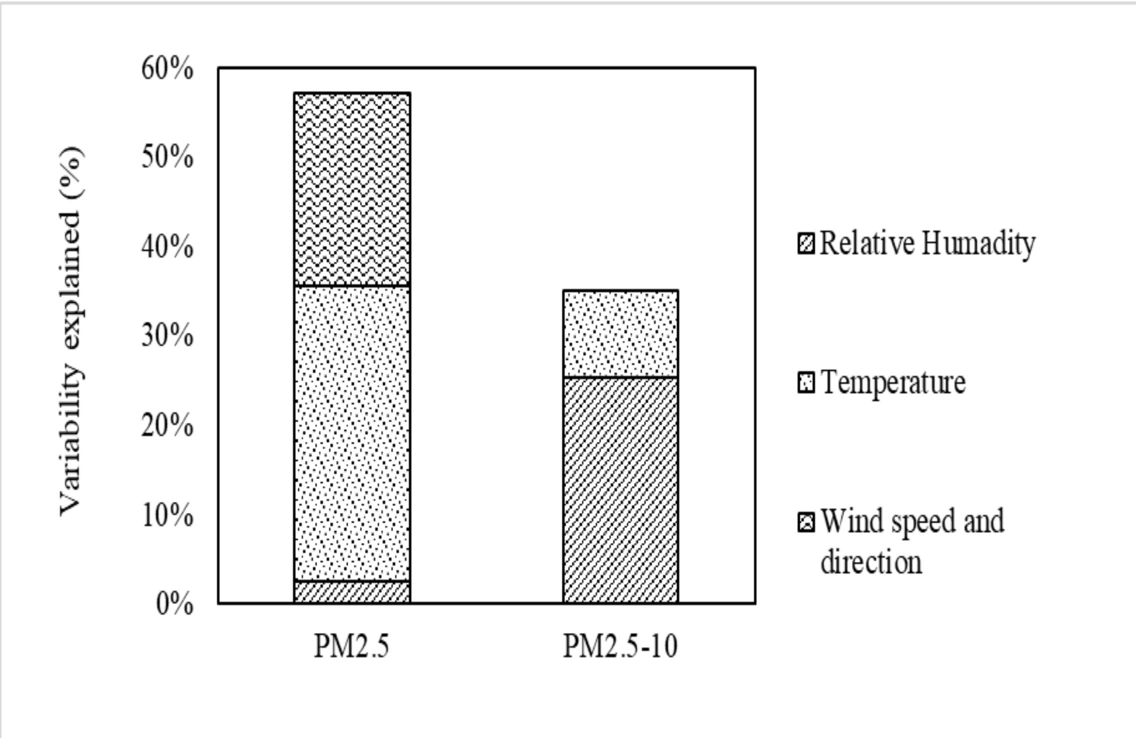

Figure 1: Relative importance of different meteorological variables in explaining the total variability

Using dominance analysis, we further quantify the proportional contribution of each variable to the variation explained directly in the model. Figure 1 represents the total variability as well as the shared variability explained by the significant meteorological parameter influencing the particulate matter concentration. Temperature can explain 58% of the total variability, whereas wind speed and wind direction altogether can explain 38% of the total variability explained by the $PM_{2.5}$ model. Although relative humidity has a small influence over $PM_{2.5}$ concentration, it can explain 72% of the total variability of coarser particles with the rest of the variability explained by temperature.

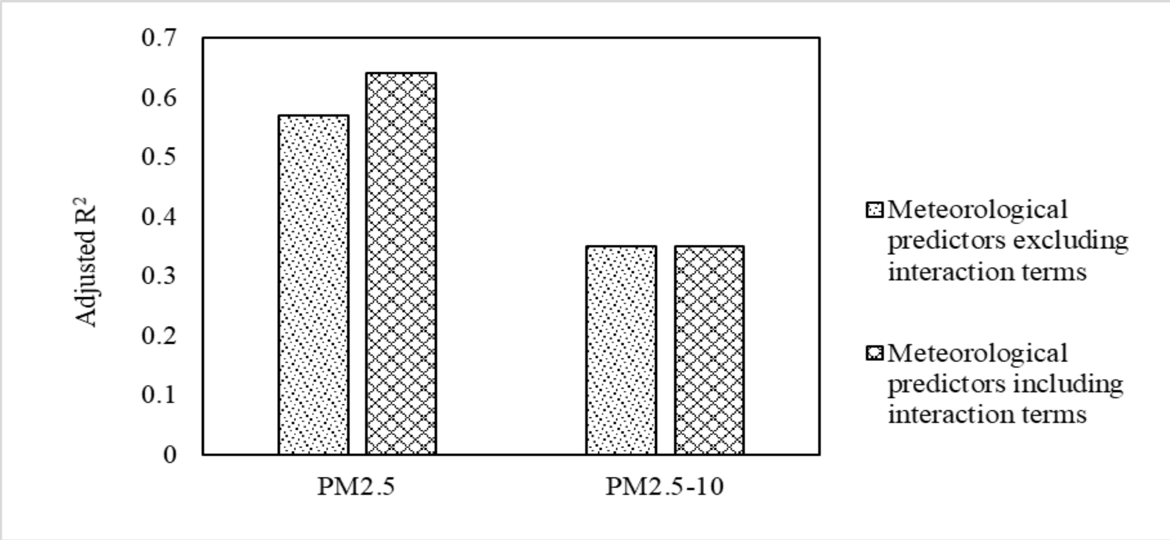

Figure 2: Influence of interactions terms among the meteorological predictors over the model performance

After incorporating all meteorological variables along with their interaction as per equation (2), the model performance for $PM_{2.5}$ increased by 12% than that if no interaction is considered among model parameters (Figure 2). Consequently, the resulting $PM_{2.5}$ model can explain 64% of the total variability of the 24-hour average $PM_{2.5}$ concentrations. In contrast, even after the addition of the interaction terms, there was no improvement in the model performance depicting negligible influence of meteorological interaction over coarse particles (Figure 2). The estimated intercept, regression coefficients and other model parameters for $PM_{2.5}$ model are summarized in Table 4. As shown in Table 4, while considering the interaction terms, all considered meteorological parameters show significant influence alone or in association with other parameters.

Table 4: Summary statistics of the stepwise regression model considering $PM_{2.5}$ as the response variable and all meteorological variables and their interaction terms as predictor variables

	Estimate	**Standard Error**	**t-Statistics**	**P-Value**
Intercept	963.43	98.89	9.74	1.1E-19
R	-228.08	96.70	-2.36	1.9E-02
T	-11.25	1.47	-7.65	2.6E-13
RH	-6.77	1.09	-6.19	2.0E-09
SR	-1.85	0.35	-5.27	2.6E-07
WS	-128.87	38.46	-3.35	9.1E-04
SinWD	107.62	35.11	3.07	2.4E-03
CosWD	12.39	3.86	3.21	1.5E-03
R:T	-9.74	2.36	-4.12	4.9E-05
R:RH	5.11	1.47	3.49	5.6E-04
R:SR	0.42	0.22	1.97	5.0E-02
T:WS	2.24	0.74	3.03	2.6E-03
RH:SR	0.02	0.00	4.81	2.3E-06
RH:WS	0.77	0.37	2.08	3.9E-02
RH:SinWD	-1.17	0.45	-2.61	9.6E-03
WS:SinWD	-18.17	6.23	-2.92	3.8E-03

4. CONCLUSIONS

Meteorology plays a significant role in explaining the variability of 24-hour average PM concentrations. In addition to the detailed understanding of the different sources of air pollution, knowledge of the meteorological drivers affecting air quality is essential to further reduce or mitigate this national problem. Similar to previous literature, the findings of the study highlight the strong negative association of temperature and relative humidity with the increase of particulate matter concentrations. Interestingly, this study also revealed the potential role of wind speed and direction over the fine particulate matter concentrations of Dhaka city, which was not explored explicitly by the previous literature. All the meteorological parameters considered in this paper altogether can explain 64% and 35% of the total variability of fine and coarse particulate matter respectively suggesting that there are other drviers of PM variability in Dhaka, and a detail source apportionment study can be done in future to identify those factors (e.g. vehicle, brick kilns, industry and construction works). Although the data used in this study is quite back-dated, the method applied in this study is powerful and gives a deeper understanding of the role of different weather parameters. Moreover, Begum et al. (2018) recently showed that the air quality of Dhaka has been stable over the past decade by analyzing the long term air quality data of Dhaka city, which justifies the use of this older dataset of this paper. In conclusion, this kind of regression model is useful for the development of particulate matter prediction models incorporating meteorological parameters. This method is also helpful for policymakers to assess the relative importance of different variables while implementing any policy in a cost-effective way.

REFERENCES

Afrin, S., Islam, M. M., Ahmed, T., & Ali, M. A. (2014). The Influence of Meteorology on Particulate Matter Concentration in the Air of Dhaka City. *2nd International Conference on Advances in Civil Engineering (ICACE-2014)*, Chittagong, Bangladesh.

Azen, R., & Budescu, D. V. (2003). The Dominance Analysis Approach for Comparing Predictors in Multiple Regression. *Psychological Methods*, *8*(2), 129–148. https://doi.org/10.1037/1082-989X.8.2.129

Bae, S., & Hong, Y. C. (2018). Health effects of particulate matter. *Journal of the Korean Medical Association*, *61*(12), 749–755. https://doi.org/10.5124/jkma.2018.61.12.749

Begum, B. A., & Hopke, P. K. (2018). Ambient air quality in Dhaka Bangladesh over two decades: Impacts of policy on air quality. *Aerosol and Air Quality Research*, *18*(7), 1910–1920. https://doi.org/10.4209/aaqr.2017.11.0465

Begum, B., Nasiruddin, M., Randal, S., Sivertsen, B., & Hopke, P. (2014). Identification and Apportionment of Sources from Air Particulate Matter at Urban Environments in Bangladesh. *British Journal of Applied Science & Technology*, *4*(27), 3930–3955. https://doi.org/10.9734/bjast/2014/11247

Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., … Forouzanfar, M. H. (2017). Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *The Lancet*, *389*(10082), 1907–1918. https://doi.org/10.1016/S0140-6736(17)30505-6

Hossain, K. M. A., & Easa, S. M. (2012). Pollutant dispersion characteristics in Dhaka city, Bangladesh. *Asia-Pacific Journal of Atmospheric Sciences*, *48*(1), 35–41. https://doi.org/10.1007/s13143-012-0004-8

Islam, M. M., & Afrin, S. (2014). Air Quality Status of Dhaka City and Effects of seasons, Weekends, Eid and Hartal Days. *2nd International Conference on Advances in Civil Engineering (ICACE-2014)*, Chittagong, Bangladesh.

Islam, M. M., & Afrin, S. (2015). The Influence of Brick Kilns on Seasonal Variation of Particulate Matter The Influence of Brick Kilns on Seasonal Variation of Particulate Matter Concentration in Dhaka City. *11th International Conference on Mechanical Engineering (ICME 2015*), Dhaka, Bangladesh.

Islam, M. M., Afrin, S., Ahmed, T., & Ali, M. A. (2015). Meteorological and seasonal influences in ambient air quality parameters of Dhaka city. *Journal of Civil Engineering*, *43*(1), 67–77.

Islam, N., Saroar, M. G., & Ahmed, T. (2018). Meteorological Influences on Atmospheric Particles in Darus Salam Area of Dhaka City. In *Proceedings of the 4thInternational Conference on Civil Engineering for Sustainable Development (ICCESD 2018),* (pp. 1–12). KUET, Khulna, Bangladesh.

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 Journal of Environmental Science and Management*, *5*(3), 265–278. https://doi.org/10.22034/gjesm.2019.03.01

Leung, D. M., Tai, A. P. K., Mickley, L. J., Moch, J. M., Van Donkelaar, A., Shen, L., & Martin, R. V. (2017). Synoptic meteorological modes of variability for fine particulate matter (PM 2.5) air quality in major metropolitan regions of China. *Atmos. Chem. Phys. Discuss. Chem. Phys. Discussion started*. https://doi.org/10.5194/acp-2017-916

Mahmood, S. (2011). Air Pollution Kills 15,000 Bangladeshis Each Year: The Role of Public Administration and Government's Integrity. *… of Public Administration and Policy Research*, *3*(May), 129–140.

Majumder, S., Sihabut, T., & Saroar, M. G. (2019). Assessment of knowledge, attitude, and practices against inhaled particulate matter among urban residents in Dhaka, Bangladesh. *Journal of Health Research*, *ahead-of-p*(ahead-of-print). https://doi.org/10.1108/jhr-12-2018-0168

Rahman, M. M., Mahamud, S., & Thurston, G. D. (2019). Recent spatial gradients and time trends in Dhaka, Bangladesh, air pollution and their human health implications. *Journal of the Air and Waste Management Association*, *69*(4), 478–501. https://doi.org/10.1080/10962247.2018.1548388

Salam, A., Hossain, T., Siddique, M. N. A., & Shafiqul Alam, A. M. (2008). Characteristics of atmospheric trace gases, particulate matter, and heavy metal pollution in Dhaka, Bangladesh. *Air Quality, Atmosphere, and Health*, *1*(2), 101–109. https://doi.org/10.1007/s11869-008-0017-8

Tai, A. P. K., Mickley, L. J., & Jacob, D. J. (2010). Correlations between fine particulate matter (PM2.5) and meteorological variables in the United States: Implications for the sensitivity of PM2.5 to climate change. *Atmospheric Environment*, *44*(32), 3976–3984. https://doi.org/10.1016/j.atmosenv.2010.06.060

Tasmin, S., Ng, C. F. S., Stickley, A., Md, N., Saroar, G., Yasumoto, S., & Watanabe, C. (2019). Effects of Short-term Exposure to Ambient Particulate Matter on the Lung Function of School Children in Dhaka, Bangladesh. *Epidemiology (Cambridge, Mass.)*, *30*, S15–S23. https://doi.org/10.1097/EDE.0000000000001012

WHO. (2016). *Monitoring Health for the SDGs, sustainable development goals, World Health Organization*. https://doi.org/10.1017/CBO9781107415324.004

Zheng, S., Pozzer, A., Cao, C. X., & Lelieveld, J. (2015). Long-term (2001-2012) concentrations of fine particulate matter (PM2.5) and the impact on human health in Beijing, China. *Atmospheric Chemistry and Physics*, *15*(10), 5715–5725. https://doi.org/10.5194/acp-15-5715-2015