EQA - International Journal of Environmental Quality
ISSN 2281-4485 - Vol. 67 (2025): 45 - 50
Journal homepage: https://eqa.unibo.it/

published by

Prediction of $PM_{2.5}$ via precursor method using meteorological parameters

Naeem Sadiq[1*], Zaheer Uddin[2]

[1]Institute of Space Science and Technology, University of Karachi, Karachi, Pakistan
[2]Department of Physics, University of Karachi, Karachi, Pakistan

*Corresponding author Email: nsadiq@uok.edu.pk

Article info
Received 20/12/2024; received in revised form 30/12/2024; accepted 27/1/2025
DOI: 10.6092/issn.2281-4485/21009

Abstract
Air pollution is one of the major environmental concerns faced by many countries including Pakistan. Being the major component of the pollution, particulate matters 2.5μm diameter ($PM_{2.5}$), are known to highly raise health risks to people in the country. The present study investigates the modelling and prediction of particulate matter 2.5μm using its precursor values and meteorological parameters; temperature, humidity levels, and wind speeds in Lahore and Karachi. The air quality of Lahore repeatedly plummets to hazardous levels in the winter season which is a severe threat to the public health and environment. An Artificial Neural Network (ANN) architecture was designed to predict $PM_{2.5}$ by employing meteorological parameters (temperature, relative humidity & wind speed) and precursor values of $PM_{2.5}$. The model consists of an input layer with four input variables, a hidden layer with 10 neurons, and an output layer consisting of $PM_{2.5}$. The model was used for both the cities of Lahore and Karachi. The Root Mean Square Error (RMSE) value for Karachi was less than 18 and for Lahore, it was 39. The prediction of $PM_{2.5}$ via ANN was good for Lahore and Karachi. However, the results of the modeling are better for Karachi. The accuracy of results were further verified by Mean Absolute Percentage Error (MAPE), Mean Absolute Bias Error (MABE), and Chi square statistics (Chi).

Keywords: *$PM_{2.5}$, Meteorological Parameters, Artificial Neural Network, Air Quality Index, Lahore, Karachi*

Introduction

One of the serious issues regarding the prevailing environment in contemporary times is air pollution. Among the various factors influencing air pollution, $PM_{2.5}$ stands out contributing to significant health concerns in humans. Pope et al. (2009), Brook et al. (2010). The main sources that contribute to the $PM_{2.5}$ levels include the combustion of fossil fuel, industrial activities, emissions from vehicles, and natural resources like dust and wildfires (Liu et al., 2020). $PM_{2.5}$ causes various health issues, including respiratory problems, cardiovascular diseases, and in some cases premature death (Linden, (2019). The World Health Organization has also established guidelines and categorization limits for the levels of $PM_{2.5}$ (WHO, 2006). There has been recent interest in understanding the associations between $PM_{2.5}$ and meteorological parameters because they are used to develop an effective air quality management strategy (Laden et al., 2006; Zhao et al., 2017). Three different studies were carried on to understand the dependence of meteorological parameters on $PM_{2.5}$ (Jacob & Winner, 2009; Tai et al., 2010; Yang et al., 2017). The outcomes of these studies were not similar, for example, a positive correlation was found between relative humidity and $PM_{2.5}$ where whereas a negative correlation between temperature and $PM_{2.5}$ in the region of the United States (Laden et al., 2006). For China, the outcomes showed that temperature is posi-

DOI: 10.6092/issn.2281-4485/21009

tively correlated with $PM_{2.5}$ (Zhang et al., 2015). A study carried out in India showed a positive correlation for both, temperature and relative humidity with $PM_{2.5}$ (Zhang et al., 2015). $PM_{2.5}$ concentrations have also been influenced by another meteorological parameter, i.e. wind. A study in England found that wind speed had a negative correlation with $PM_{2.5}$ levels, which may be because wind speed dispersed the pollutants and hence reduced the concentrations (Harrison et al., 2012). Similarly, a study in India revealed that the direction of the wind from industrial areas plays a vital role in contributing to higher concentrations of $PM_{2.5}$ (Kumar et al., 2011). An extensive usage of Artificial Neural Networks (ANNs) has been observed, during the last few years, to simulate and predict the levels of $PM_{2.5}$; being able to represent complex nonlinear relationships among the variables (Gardner & Dorling, 1998; Kukkonen et al., 2003). For example, a study in the United States used an ANN model to predict $PM_{2.5}$ concentrations in Los Angeles, with satisfactory results (Gardner & Dorling, 1998). A study in China used an ANN model to predict $PM_{2.5}$ concentrations in Beijing, with sufficiently good results (Wang and Sun, 2019). Also, Masood and Ahmad, in 2023 used an ANN model to predict $PM_{2.5}$ concentrations in Delhi, India, with satisfactory results. A few studies were also conducted in the context of Pakistan. A study in 2019 by basheer et al., for Lahore found that $PM_{2.5}$ concentrations were high during the winter months and low during the summer months. Another study in Karachi revealed that $PM_{2.5}$ concentrations were high in industrial areas and low in residential areas (Shahid et al., 2016). In Pakistan, there is a need for more closer study of the relationships between $PM_{2.5}$ and meteorological parameters, in addition to the application of ANN models for predicting $PM_{2.5}$ concentrations. Hence, this study aims to address this gap by investigating the relationships between $PM_{2.5}$ and temperature, relative humidity, and wind speed in Lahore and Karachi, and by developing an ANN model to predict $PM_{2.5}$ concentrations in these cities.

Methodology

To study and compare the air quality of Lahore and Karachi machine learning methodology based on Artificial Neural Network has been used. ANN works in the same way as the human brain perception technique works. Like a message conveyed to various parts of the brain via neurons, ANN uses neurons to perceive the mathematical dependence of the input variable. The simplest ANN architecture consists of three layers, one input layer, where input variables are fed into the network, the second layer is the hidden layer, which consists of various numbers of neurons that are connected through input variables and random weights for each variable. The weights for neutrons are adjusted so that neurons understand complex relationships between input and output variables. In this study, we employed meteorological parameters daily average temperature, daily average humidity, and wind speed as input parameters to predict the $PM_{2.5}$ for two cities (Karachi and Lahore) in Pakistan. An additional parameter precursor value of $PM_{2.5}$ was used in the set of input variables. A single hidden layer with ten neurons was employed that connects the output parameter. The architecture of the neural network is shown in the Figure 1.

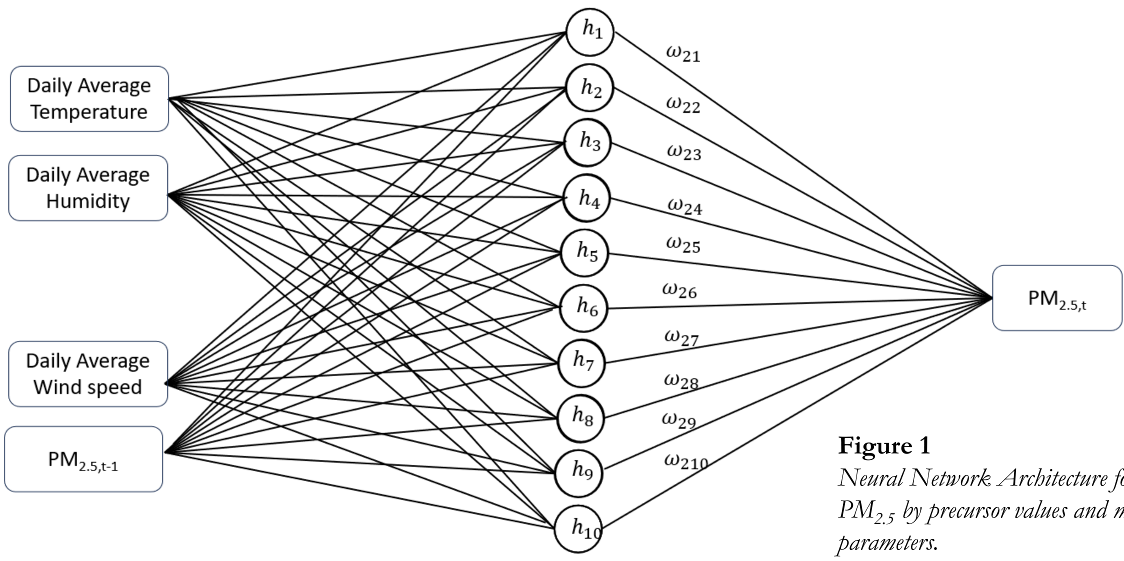

Figure 1
Neural Network Architecture for Prediction of $PM_{2.5}$ by precursor values and meteorological parameters.

DOI: 10.6092/issn.2281-4485/21009

Data

The data on particulate matter were obtained from the Pakistan Environmental Protection Agency (EPA) from January 2021 to December 2023 for Lahore and Karachi cities. While, Meteorological data, like temperature and relative humidity, were requested and collected from the Pakistan Meteorological Department (PMD) for the same time frame as $PM_{2.5}$ data. The datasets were prepared by cleaning and preprocessing the data and then addressing data gaps.

Error Analysis

To compare the errors in the predicted values of $PM_{2.5}$ by the ANN model, the following errors were calculated.

$$RMSE = \sqrt{\frac{\left((PM_{2.5})_{Recorded} - (PM_{2.5})_{Predicted}\right)^2}{n}} \quad [1]$$

$$MABE = \frac{\left|(PM_{2.5})_{Recorded} - (PM_{2.5})_{Predicted}\right|}{n} \quad [2]$$

$$MAPE = \frac{(PM_{2.5})_{Recorded} - (PM_{2.5})_{Predicted}}{(PM_{2.5})_{Recorded}} \text{x}100 \quad [3]$$

$$Chi = \frac{\left((PM_{2.5})_{Recorded} - (PM_{2.5})_{Predicted}\right)^2}{(PM_{2.5})_{Recorded}} \quad [4]$$

Results and Discussion

This study aims to predict $PM_{2.5}$ values using its precursor values together with three meteorological parameters. An ANN model was developed for this purpose with four input variables and one hidden layer. The backpropagation algorithm for ANN was employed to predict the $PM_{2.5}$ values first for Lahore and then for Karachi. Figure 2 provides a graphical representation of the machine learning regression through ANN modeling for Lahore. It is evident from Figure 2 that R^2 values for the training, validation, test, and combined datasets referred to as "All" are reasonably high. It indicates a strong correlation between the predicted and observed $PM_{2.5}$ levels. The values of R^2 are given in Table 1.

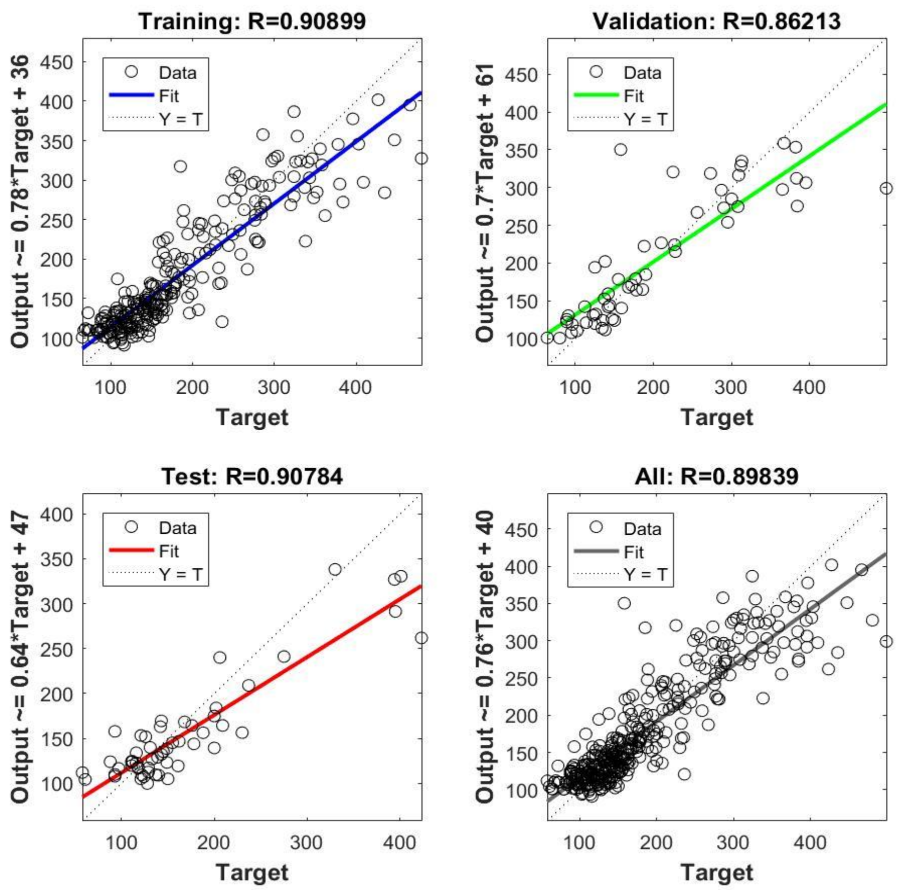

Figure 2
Predicted $PM_{2.5}$ levels in Lahore using wind speed, humidity, and temperature as input features

DOI: 10.6092/issn.2281-4485/21009

Table 1 presents the R^2 and RMSE values for ANN training, validation, testing, and overall data for Lahore and Karachi. The RMSE values are reasonably good, however, the model performance for Karachi was better than Lahore.

Table 1: *R^2 for training testing and validatioin, & error values for the regression line of the ANN Model for Lahore and Karachi*

	Training	Validation	Testing	All	RMSE	MABE	MAPE	Chi
Lahore	0.909	0.862	0.908	0.898	39.203	27.900	0.165	7.690
Karachi	0.925	0.879	0.882	0.911	18.698	12.770	0.120	2.866

The output values of ANN are compared with the recorded values of $PM_{2.5}$. Figure 3 gives the comparison between predicted (orange line) and recorded (blue line) $PM_{2.5}$ levels. It is evident from the figure that the predicted values are close to the recorded values of $PM_{2.5}$, showing a perfect overlap in most of the regions of the figure. The result indicates that the model predicts $PM_{2.5}$ for Lahore effectively using the four input variables mentioned above. Figure 4 depicts the Training, Validation, testing, and overall relationship in the form of a regression line via ANN for the prediction of $PM_{2.5}$ in Karachi. As mentioned in Table 1 and also shown in Figure 4, the R^2 values for the training, validation, testing, and all data are reasonably good. It implies that the ANN model employed in this study predicts $PM_{2.5}$ values for Karachi relatively

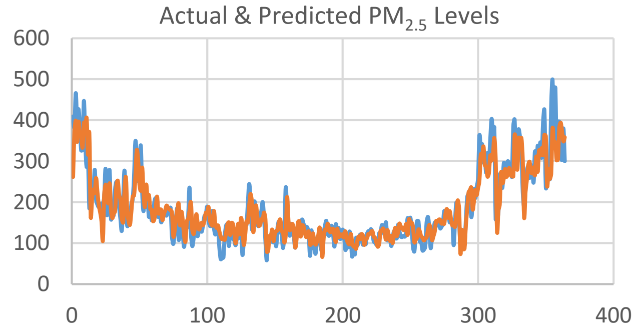

Figure 3. *Comparison of predicted and actual $PM_{2.5}$ levels in Lahore*

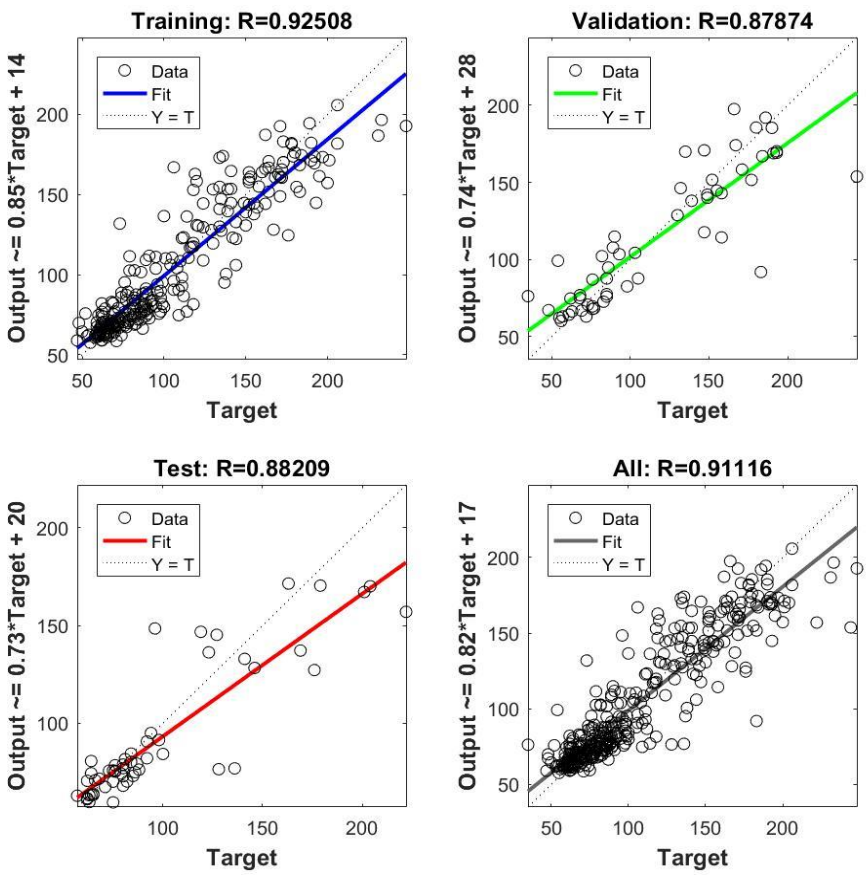

Figure 4
Predicted $PM_{2.5}$ levels in Karachi using wind speed, humidity, and temperature as input features

DOI: 10.6092/issn.2281-4485/21009

better than Lahore. It can also be supported by RMSE values which is almost half for Karachi (see Table 1). Figure 5 compares the predicted values of $PM_{2.5}$ with the corresponding recorded values. There are two sketches shown in orange (predicted) and blue (recorded) colors. As compared to Lahore, the predicted values of $PM_{2.5}$ for Karachi are relatively better and more close to the recorded values. This shows that the ANN model works well for both Lahore and Karachi. The error analysis in the model was carried out by calculation of RMSE, MABE, MAPE, and Chi-square values for both cities. The values of all the errors are reasonable so the prediction of $PM_{2.5}$ by ANN using precursor values exhibits a suitable model. Further comparison of errors within the cities displays that all the errors for Karachi are sufficiently less than those of Lahore. ANN model is the best for the prediction of $PM_{2.5}$ for Karachi using its precursor values together with the meteorological data. It can be concluded that the prediction of $PM_{2.5}$ by the ANN model with its precursor values demonstrates excellent performance. The model predicts future values of $PM_{2.5}$ using the previous day's value. An excellent correlation is seen between recorded and predicted values.

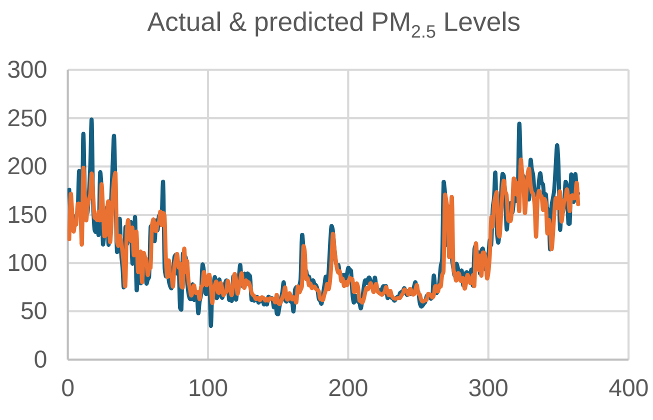

Figure 5. *Comparison of predicted and actual $PM_{2.5}$ levels in Karachi*

Conclusions

The relationships between particulate matter ($PM_{2.5}$) and influential meteorological parameters like temperature, relative humidity, and wind speed in Lahore and Karachi have been investigated. The outcomes of the study show that considered meteorological parameters influence $PM_{2.5}$ concentration. To study this relationship and to predict $PM_{2.5}$ ANN model was developed for Lahore and Karachi data. ANN model with a single hidden layer and 10 neurons demonstrated excellent fitting of $PM_{2.5}$ data for Lahore and Karachi. However, the results of the prediction of $PM_{2.5}$ were better for Karachi, where the RMSE values were about half of that of Lahore. The use of an additional parameter in the input as precursor values of $PM_{2.5}$ enables us to predict future values using present values of $PM_{2.5}$. Such findings will help to focus interventions at specific conditions for controlling and mitigating $PM_{2.5}$ pollution. Some recommendations are made based on the results of the study:

1. An air quality monitoring network should be installed in the big populated and polluted cities like Lahore and Karachi to obtain the more close results and predictions through integrated data.
2. A forecasting system should be developed to predict the concentration of particulate matter over Lahore and Karachi based on weather-related parameters.
3. Implement emission reduction strategy to reduce $PM_{2.5}$ concentrations, such as clean and green energy resources and implementing emission controls on industrial sources
4. Further research should be conducted to investigate the relationships between $PM_{2.5}$ and other air pollutants, such as NO_2 and O_3, in Lahore and Karachi.

Ethics Statement

It is hereby declared that this research was conducted in accordance with ethical standards. No primary data collection involving human subjects.

Funding

There is no funding for the publication. However, one of the authors is receiving a stipend from the Sindh Higher Education Commission.

References

BASHEER S., RASHID H., NASIR A., NAWAZ R.A. (2019) Spatial and temporal variability analysis of PM 2. 5 concentration in Lahore city. Environmental Contaminants Reviews (ECR), 2(1) : 6-10. https://doi.org/10.1016/j.atmosres.2019.104715

BROOK R.D., RAJAGOPALAN S., POPE III C.A., BROOK J.R., BHATNAGAR A., DIEZ-ROUX A.V., HOLGUIN F., HONG Y., LUEPKER R.V., MITTLEMAN M.A., PETERS A. (2010) Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. Circulation, 121(21) :.2331-2378. https://doi.org/10.1161/CIR.0b013e3181dbece1

DOI: 10.6092/issn.2281-4485/21009

GARDNER M.W., DORLING S.R. (1998) Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmospheric Environment, 32(14-15): 2627-2636. https://doi.org/10.1016/S1352-2310(97)00447-0

HARRISON R.M., LAXEN D.P.H., MOORCROFT S. (2012) Interpretation of particulate elemental and organic carbon concentrations at rural, urban and kerbside sites. Atmospheric Environment, 55: 565-575. https://doi.org/10.1016/j.atmosenv.2005.08.017

JACOB D.J., WINNER D.A. (2009) Effect of climate change on air quality. Atmospheric Environment, 43(1):51-63. https://doi.org/10.1016/j.atmosenv.2008.09.051

KUKKONEN J., PARTANEN L., KARPPINEN A., RUUSKANEN J., JUNNINEN H., KOLEHMAINEN M., NISKA H., DORLING S., CHATTERTON T., FOXALL R., CAWLEY G. (2003) Extensive evaluation of neural network models for the prediction of NO_2 and PM_{10} concentrations, compared with a deterministic modelling system and measurements in central Helsinki. Atmospheric Environment, 37(32):4539-4550. https://doi.org/10.1016/S1352-2310(03)00583-1

KUMAR A., GOYAL P. (2011) Forecasting of daily air quality index in Delhi. Science of the Total Environment, 409(24) : 5517-5523. https://doi.org/10.1016/j.scitotenv.2011.08.069

LADEN F., SCHWARTZ J., SPEIZER F.E., DOCKERY D.W. (2006) Reduction in fine particulate air pollution and mortality: Extended analysis of the Harvard Six Cities Study. American Journal of Respiratory and Critical Care Medicine, 173(6):667-672. https://doi.org/10.1164/rccm.200503-443OC

LINDEN B. (2019) Air pollution: Outdoor air quality and health. British Journal of Cardiac Nursing, 14(6):1-4. https://doi.org/10.12968/bjca.2019.0044

LIU H., CAO C., HUANG J., CHEN Z., CHEN G., LAI Y. (2020) Progress on particulate matter filtration technology: basic concepts, advanced materials, and performances. Nanoscale, 12(2):437-453. https://doi.org/10.1039/C9NR08851B

MASOOD A., AHMAD K. (2023) Prediction of $PM_{2.5}$ concentrations using soft computing techniques for the megacity Delhi, India. Stochastic Environmental Research and Risk Assessment, 37(2):625-638. https://doi.org/10.1007/s00477-022-02291-2

POPE III C.A., EZZATI M., DOCKERY D.W (2009) Fine-particulate air pollution and life expectancy in the United States. New England Journal of Medicine, 360(4):376-386.https://www.nejm.org/doi/full/10.1056/NEJMsa0805646

SHAHID I., KISTLER M., MUKHTAR A., GHAURI B.M., RAMIREZ-SANTA CRUZ C., BAUER H., PUXBAUM H. (2016) Chemical characterization and mass closure of PM_{10} and $PM_{2.5}$ at an urban site in Karachi–Pakistan. Atmospheric environment, 128:114-123. https://doi.org/10.1016/j.atmosenv.2015.12.005

TAI A.P., MICKLEY L.J., JACOB D.J. (2010) Correlations between fine particulate matter ($PM_{2.5}$) and meteorological variables in the United States: Implications for the sensitivity of $PM_{2.5}$ to climate change. Atmospheric environment, 44(32) : pp.3976-3984. https://doi.org/10.1016/j.atmosenv.2010.06.060

WANG X., SUN W. (2019) Meteorological parameters and gaseous pollutant concentrations as predictors of daily continuous $PM_{2.5}$ concentrations using deep neural network in Beijing–Tianjin–Hebei, China. Atmospheric Environment, 211:128-137. https://doi.org/10.1016/j.atmosenv.2019.05.004

WORLD HEALTH ORGANIZATION (WHO) (2006) Air Quality Guidelines: Global Update 2005. particulate matter, ozone, nitrogen dioxide, and sulfur dioxide. World Health Organization. https://books.google.com.sg/books?hl=en&lr=&id=7VbxUdlJE8wC&oi=fnd&pg=PR9&dq=WORLD+HEALTH+ORGANIZATION+(WHO)+(2006)+Air+Quality+Guidelines:+Global+Update+2005.+Available+at:+https://www.who.int/airpollution/guidelines/en/+(Accessed:+2024).&ots=w445vPP8s7&sig=iuNgnRCvtDKnqROyq91cA883kBs&redir_esc=y#v=onepage&q&f=false

YANG Q., YUAN Q., LI T., SHEN H., ZHANG L. (2017) The relationships between $PM_{2.5}$ and meteorological factors in China: seasonal and regional variations. International journal of Environmental Research and Public Health, 14(12):1510. https://doi.org/10.3390/ijerph14121510

ZHAO L., LIANG H.R., CHEN F.Y., CHEN Z., GUAN W.J., LI J.H., (2017) Association between air pollution and cardiovascular mortality in China: a systematic review and meta-analysis. Oncotarget, 8(39) : 66438. https://doi.org/10.18632/oncotarget.20090