CGC International Journal of Contemporary Technology and Research
ISSN: 2582-0486 (online) Vol.-6, Issue-2; DOI: 10.46860/cgcijctr.2024.06.10.396

Study of different Air Quality Forecasting Models

Anuj Kumar Gupta
Dept. of Computer Science. and Engineering
Chandigarh Group of Colleges, Landran, Mohali, PB, India
Email id: cgccoe.cse.anuj@gmail.com

***Abstract* – One of the main effects of urban growth is industrialization, which adds to the strain on resources like water and the environment. In recent years, air pollution has emerged as one of the biggest environmental challenges in emerging nations. It has negative impacts on people, plants, and animals. According to a recent study, air pollution is responsible for 3.3 million premature deaths worldwide each year. Therefore, air quality forecasting is crucial since it lessens the negative effects of air pollution on human health. In this paper, the various models used to predict air quality have. Air quality forecasting is crucial since it lessens the negative effects of air pollution on human health been discussed with its limitations. The number of factors affecting the Air Quality prediction model has also been discussed. It has been found that Machine learning techniques provide a model with the help of which we can predict the air quality when several factors affecting the air quality have been considered.**

***Index Terms*-**Air Pollutants, Air Pollution, Air Quality prediction model, Pollution.

I. INTRODUCTION

Pollution is defined as the contamination of the environment by foreign substances called Pollutants. The pollutants can be introduced into the air or water and thus cause air pollution and water pollution. The sources of Air pollution include smoke produced from vehicles, industries or natural disasters like forest fires. Air Pollution leads to several consequences on human health. It also affects animals and crops. The greenhouse effect is also one of the long-term effects of Air Pollution. An accurate air quality forecast is significant as it will help lower the effect of air pollution on human health. In this paper, the various models that are used to predict air quality have been discussed with their limitations.

The rest of the paper has been structured as follows. Section 2 presents an overview of various Air Pollutants. Section 3 presents the literature survey. In section 4, the various air quality prediction models with their limitations have been discussed. A brief conclusion has been presented in the last section.

II. LITERATURE SURVEY

In the past, analysts have examined data on air pollution using a variety of data mining, machine learning, and other statistical techniques [17]. Table 1 summarizes the papers based on the technique applied for evaluating pollution data.

Table1. Summary of papers based on the technique to evaluate air quality

Title	Technique	Author	Year
Modelling PM2.5 urban pollution with specified meteorological parameters and machine learning.	Boosted Trees (BT), Linear Support Vector Machines (L-SVM), Rus Boosted Tree (RBT)	Kleine Deters, J., Zalakeviciute, R., Gonzalez, M., & Rybarczyk, Y	2017
Machine learning method to predict urban pollution.	J48, ZeroR, Naïve Bayes	Rybarczyk, Y., &Zalakeviciute, R.	2016
Recognising pollution sources and predicting urban air quality using ensemble learning methods	Decision Tree Forest (DTF), Decision Tree Boost (DTB), Single Decision Tree (SDT) and Support Vector Machines (SVM)	Singh, K. P., Gupta, S., & Rai, P	2013
Principal component analysis and support vector machines for forecasting NO and NO_2 concentrations at traffic sources	Multi-Layer Perceptron and Support Vector Regression	Juhos, I., Makra, L., &Tóth, B.	2008
Modelling SO_2 concentration at a point with statistical approaches.	Artificial Neural Networks	Nunnari et al	2020
Matching neural networks and regression model for ozone foretelling.	Regression Models	Comrie, A. C.	1997
The next group of integrated air quality models: EPA's models 3	Grid Models	Dennis et al	2016
Modelling of the air pollution in the Cracow area	Numerical Models	Juda,K.	2016
Highway modeling-I: forecast of velocity and commotion fields in the stir of vehicles.	Wind Tunnel Simulation	Eskridge, R.E. and Hunt, J.C.R	2019
Time series analysis forecasting and control.	Time Series Models	Box, G.E.P. and Jenkins, G.M.	2016
Crosswind shears effects on atmospheric diffusion.	Gaussian Models	Csanday, G.T.	2012

III. AIR POLLUTANTS

Pollutants are foreign substances that are introduced into the air by emissions from industries and vehicles. Pollutants can be solid, liquid or gases. They can be either primary pollutants emitted directly from sources or secondary pollutants formed from the mixture of primary pollutants. Carbon Monoxide emitted from vehicles and Sulphur dioxides emitted from industries are examples of primary pollutants, whereas ozone is an example of a secondary pollutant. According to the United States Environmental Protection Agency [5], the following six criteria pollutants have been identified.

Ozone (O_3)

Ozone is a secondary pollutant formed from hydrocarbon and nitrogen dioxide. It causes respiratory problems like asthma in humans and also causes damage to plants and crops.

Particulate Matter (PM)

Particulate Matter is the combination of liquid drops and solid particles in the environment. $PM_{2.5}$ refers to the particles with 2.5 µm or smaller diameter and PM_{10} is the particles with 10 µm or smaller diameter. Particulate matter causes heart problems, bronchitis and other lung problems in humans. They also lead to acid rain and Haze.

Nitrogen dioxide (NO_2)

Nitrogen dioxide is emitted from vehicles and factories. Ozone is formed from the reaction of NO_2 with hydrocarbons. NO_2 causes respiratory problems like inflammation in the airway and asthma.

.

Sulphur Dioxide (SO_2)

Sulphur Dioxide is emitted from the industries. It causes breathing problems and wheezing in asthma patients.

Carbon Monoxide (CO)

Carbon Monoxide is produced from vehicular emissions. It causes problems in people with heart problems as it lowers the capacity of blood to carry oxygen.

Lead (Pb)

Lead is emitted from the vehicles with leaded gasoline. Other sources of lead include activities such as battery recycling. It causes neurodevelopment disorders in children.

The various factors that affect the concentration of air pollutants are meteorological factors like temperature, humidity, wind speed, temporal factors like hour of the day and day of the week and the local topography like location. Among these factors, the one that has the strongest influence on the variation in the concentration of air pollutants is the meteorological factor.

IV. AIR QUALITY PREDICTION MODEL

An instrument used to anticipate the air quality, which in turn depends on a variety of elements, is an air quality prediction model. It defines the relationship between emissions that cause pollution and the air quality affected from it. The various factors that affect Air Quality Prediction are weather factors like temperature, humidity and wind speed and temporal factors like hour of the day and day of the week. Factors affecting Air Quality Prediction are depicted in Figure 1.

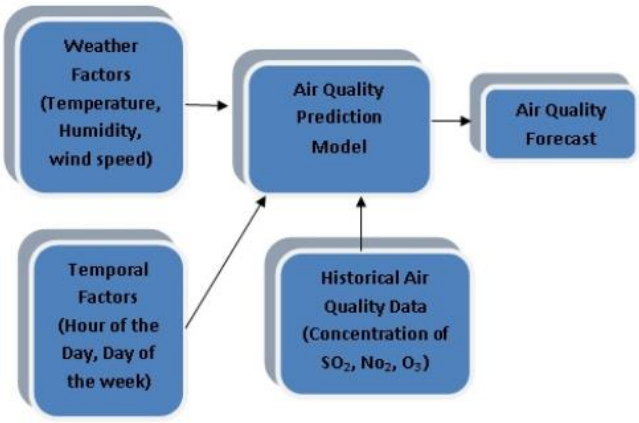

Figure 1 Factors Affecting Air Quality Prediction Model

4.1 Advantages of Air Quality Prediction

If the Air Quality Prediction is reliable, then it has several advantages:

i. **Issue of Health Alerts** – When the air quality is higher than the desired levels, health alerts can be issued to the public. This can be helpful to people who suffer from respiratory problems like asthma.

ii. **Traffic planning** – Smoke generated from forest fires or haze produced as a result of Particulate Matter can affect aviation and other traffic. Thus, a forecast for air quality is necessary for safety.

4.2 Types of Air Quality Prediction Models

In literature, several models for Air Quality Prediction have been used to forecast pollutants like Particulate Matter (PM) and other harmful gases like Carbon Monoxide (CO), Sulphur Dioxide (SO_2), and Nitrogen Dioxide (NO_2). The limitations of these models are

summarised in Table 2. The various Air Quality Prediction models are described as follows.

4.2.1 Deterministic Models

Deterministic models are the traditional models for air quality prediction [1]. They calculate the concentration of pollutants based on the solution of an equation consisting of various emission and meteorological variables that represent the physical process.

i. The Gaussian model is a widely used deterministic air quality prediction model.
ii. Numerical modelling is a deterministic model where the relationship between the variables is represented by partial differential equations.

4.2.2 Statistical Models

Statistical models find the pollutant concentration by associating a statistical relationship between the emission and meteorological variables [4].

i. The link between the independent (meteorological and emission parameters) and dependent variables is described by regression models. (Pollutant concentrations).
ii. Time series models are the statistical method applied to non-repeatable experiments.

4.2.3 Physical Models

Physical models have a high potential to predict air quality. These models use scaling methods to convert the measured concentrations to atmospheric concentrations of pollutants [2].

4.2.4 Photochemical models

Using a set of mathematical equations characterised by chemical and physical processes, photochemical models calculate the pollutant concentrations. Grid Models are the most powerful model that solves a problem by dividing the region into horizontal and vertical cells [10].

4.2.5 Machine Learning

Machine Learning has been defined as a field of study that gives computers the ability to think [11]. A model with some parameters is defined and machine learning is the process of executing a program to optimize performance based on either training data or using some past experiences. Several techniques in machine learning like Support Vector Machine, Decision trees, Naïve Bayes, Regression etc. have been used for Air Quality Prediction.

Table2. Limitations of air quality prediction models

Model Name		Limitations
Deterministic Model	Gaussian Models	It is expensive in terms of time and data storage. Additionally, it needs a lot of input data.
	Numerical Models	
Statistical Models	Regression Models	Require long historical data sets.
	Time Series Models	
Physical Models	Wind Tunnel Simulation	The cost of construction and operation is high. Real-time forecast is not possible.
Photochemical Models	Grid Models	Inaccuracy and missing values in meteorological and emission variables cause biases.

V. CONCLUSION

In this paper, the various models for Air Quality prediction with their limitations have been discussed. The different factors affecting air quality have also been discussed. If these parameters are carefully chosen, then the accuracy of the Air Quality prediction model can be improved. Even though numerous models have been suggested in the literature, most of them have substantial operating and data storage costs. Real-time forecasting is not possible in some models, and outcomes are skewed as a result of missing values. A great technique for managing air pollution, machine learning can overcome these constraints and take into account multiple characteristics in a single model.

REFERENCES

[1]. Wang, D., Wei, S., Luo, H., Yue, C., & Grunder, O. (2017). A novel hybrid model for air quality index forecasting based on a two-phase decomposition technique and modified extreme learning machine. Science of The Total Environment, 580, 719-733.

[2]. Kleine Deters, J., Zalakeviciute, R., Gonzalez, M., &Rybarczyk, Y. (2017). Modeling PM2. 5 urban pollution using machine learning and selected meteorological parameters. Journal of Electrical and Computer Engineering, 2017.

[3]. Rybarczyk, Y., &Zalakeviciute, R. (2016, October). Machine learning approach to forecasting urban pollution. In Ecuador Technical Chapters Meeting (ETCM), IEEE (pp. 1-6). IEEE.

[4]. Singh, K. P., Gupta, S., & Rai, P. (2013). Identifying pollution sources and predicting urban air quality using ensemble learning methods. Atmospheric Environment, 80, 426-437.

[5]. EPA (2005). Six common air pollutants. U. S. Environmental Protection Agency. http://www.epa.gov/air/urbanair/6 poll.html.

[6]. Nunnari, G., Dorling, S., Schlink, U., Cawley, G., Foxall, R., & Chatterton, T. (2004). Modeling SO 2 concentration at a point with statistical approaches. Environmental Modelling & Software, 19(10), 887-905.

[7]. Lu, W., Wang, W., Leung, A. Y. T., Lo, S. M., Yuen, R. K., Xu, Z., & Fan, H. (2002). Air pollutant parameter forecasting using support vector machines. In Neural

Networks, 2002. IJCNN'02. Proceedings of the 2002 International Joint Conference on (Vol. 1, pp. 630-635). IEEE.

[8]. Dennis, R. L., Byun, D. W., Novak, J. H., Galluppi, K. J., Coats, C. J., and Vouk, M. A. (1996). The next generation of integrated air quality models: EPA's models3. Atmospheric Environment, 30(12):1925–1938.

[9]. Juda, K. (1986). Modeling of the air pollution in the Cracow area. Atmospheric Environment (1967), 20(12), 2449-2458.

[10]. Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C., &Baklanov, A. (2012). Real-time air quality forecasting, part II: State of the science, current research needs, and prospects. Atmospheric Environment, 60, 656-676.

[11]. Alpaydin, E. (2014). Introduction to machine learning. MIT Press.

[12]. Eskridge, R.E. and Hunt, J.C.R., (1979). Highway Modelling-I: prediction of velocity and turbulence fields in the wake of vehicles. Journal of Applied Meteorology, 18 (4), 387- 400.

[13]. Box, G.E.P. and Jenkins, G.M., (1976). Time series analysis forecasting and control. 2nd Edition, Holden day, San Francisco.

[14]. Csanday, G.T., (1972). Crosswind shears effects on atmospheric diffusion. Atmospheric Environment, 6,221-232.

[15]. Singh, K. P., Gupta, S., & Rai, P. (2013). Identifying pollution sources and predicting urban air quality using ensemble learning methods. Atmospheric Environment, 80, 426-437.

[16]. Juhos, I., Makra, L., &Tóth, B. (2008). Forecasting of traffic origin NO and NO 2 concentrations by support vector machines and neural networks using principal component analysis. Simulation Modelling Practice and Theory, 16(9), 1488-1502.

[17]. AK Gupta, M Sharma, A Sharma, V Menon. "A Study on SARS-CoV-2 (COVID-19) and Machine Learning Based Approach to Detect COVID-19 Through X-Ray Images", International Journal of Image and Graphics, 2021,