A Study of Various Air Quality Prediction Models

Jasleen Kaur Sethi  
Research Scholar  
USICT, GGSIPU  
Dwarka  
New Delhi, India

Mamta Mittal  
Assistant Professor  
G. B. Pant Govt. Engg College  
Okhla  
New Delhi, India

ABSTRACT

Industrialization is one of the major consequences of urban development which poses an additional burden on resources such as water and environment. In the recent years, one of the greatest environmental threats in developing countries is Air Pollution. It causes adverse effects to humans, plants and animals. A recent study has estimated that 3.3 million annual premature deaths in the world are caused due to air pollution. Thus, Air Quality Prediction is important as it helps lower the effect of air pollution on human health. In this paper, the various models used to predict air quality have been discussed with their limitations. The number of factors affecting Air Quality prediction model have 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 numbers of factors affecting the air quality have been considered.

Keywords

Air Pollutants, Air Pollution, Air Quality prediction model, Pollution

1. INTRODUCTION

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

The remaining paper has been organized 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.

2. LITERATURE SURVEY

In the past, Analysts have used number of techniques of data mining and Machine Learning and other statistical techniques to analyze air pollution data. Table 1 summarizes the papers based on the technique applied for evaluating pollution data.

<table>
<thead>
<tr>
<th>Title</th>
<th>Technique</th>
<th>Author</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modeling PM 2.5 urban pollution using machine learning and selected meteorological parameters.</td>
<td>Boosted Trees (BT), Linear Support Vector Machines (L-SVM), Rus Boosted Tree (RBT)</td>
<td>Kleine Deters, J., Zalakeviciute, R., Gonzalez, M., &amp; Rybarczyk, Y</td>
<td>2017</td>
</tr>
<tr>
<td>Identifying pollution sources and predicting urban air quality using ensemble learning methods</td>
<td>Decision Tree Forest (DTF), Decision Tree Boost (DTB), Single Decision Tree (SDT) and Support Vector Machines (SVM)</td>
<td>Singh, K. P., Gupta, S., &amp; Rai, P</td>
<td>2013</td>
</tr>
<tr>
<td>Forecasting of traffic origin NO and NO2 concentrations by support vector machines and neural networks using principal component analysis</td>
<td>Multi-Layer Perceptron and Support Vector Regression</td>
<td>Juhos, I., Makra, L., &amp; Tóth, B.</td>
<td>2008</td>
</tr>
<tr>
<td>Modelling SO2 concentration at a point with statistical approaches.</td>
<td>Artificial Neural Networks</td>
<td>Nunnari et al</td>
<td>2004</td>
</tr>
</tbody>
</table>
The next generation of integrated air quality models: EPA’s models 3

| Modelling of the air pollution in the cracow area | Grid Models | Dennis et al | 1996 |
| Modelling of the air pollution in the cracow area | Numerical Models | Juda.K. | 1986 |
| Time series analysis forecasting and control. | Time Series Models | Box, G.E.P. and Jenkins, G.M. | 1976 |
| Crosswind shear effects on atmospheric diffusion. | Gaussian Models | Csanday, G.T. | 1972 |

3. AIR POLLUTANTS

Pollutants are foreign substances that are introduced into 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 dioxide emitted from industries are examples of primary pollutants, whereas ozone is an example of secondary pollutant. According to United States Environmental Protection Agency [5], the following six criteria pollutants have been identified.

1. Ozone (O₃)
   Ozone is a secondary pollutants formed from hydrocarbon and nitrogen dioxide. It causes respiratory problems like asthma in humans and also causes damage to plants and crops.

2. 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} are 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.

3. Nitrogen dioxide (NO₂)
   Nitrogen dioxide is emitted from vehicles and factories. Ozone is formed from the reaction of NO₂ with hydrocarbons. NO₂ causes respiratory problems like inflammation in airway and asthma.

4. Sulphur Dioxide (SO₂)
   Sulphur Dioxide is emitted from the industries. It causes breathing problems and wheezing in asthma patients.

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

6. Lead
   Lead is emitted from the vehicles with leaded gasoline. Other sources of lead include activities such as battery recycling. It causes neuro development 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 which has the strongest influence on the variation in the concentration of air pollutants are the meteorological factors.

4. AIR QUALITY PREDICTION MODEL

An Air Quality prediction model is a tool used to forecast the air quality which in turn depends on number of factors. 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.

![Fig 1: Factors affecting Air Quality Prediction Model](image-url)
4.1 Advantages of Air Quality Prediction

If the Air Quality Prediction is reliable, then it have number of advantages like:

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

2. Traffic planning – Smoke generated from forest fires or haze produced as 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, a number of 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 (SO₂), Nitrogen Dioxide (NO₂). The limitations of these models are summarized 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 equation consisting of various emission and meteorological variables that represent the physical process.

- Gaussian model is widely used deterministic air quality prediction model.
- 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].

- Regression models describes the relationship between independent (meteorological and emission parameters) variables and dependent variable (pollutant concentrations).
- Time series models are the statistical method applied to non repeatable experiments.

4.2.3 Physical Models

Physical models have 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

Photochemical models calculate the pollutant concentrations by using a set of mathematical equations characterized by chemical and physical processes. Grid Models are most powerful model which 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 the 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. A number of techniques in machine learning like Support Vector Machine, Decision trees, Naïve Bayes, Regression etc. have been used for Air Quality Prediction.

<table>
<thead>
<tr>
<th>Model Name</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deterministic Model</td>
<td>Gaussian Models</td>
</tr>
<tr>
<td></td>
<td>Numerical Models</td>
</tr>
<tr>
<td>Statistical Models</td>
<td>Regression Models</td>
</tr>
<tr>
<td>Physical Models</td>
<td>Wind Tunnel Simulation</td>
</tr>
<tr>
<td>Photochemical Models</td>
<td>Grid Models</td>
</tr>
</tbody>
</table>

5. 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 Air Quality prediction model can be improved. Although many models have been proposed in literature, most of the models have high operational and data storage cost. In some models, real time forecast is not possible and biased results are produced due to missing values. Machine learning approach can address these limitations and consider several parameters in a single model and provides an excellent tool for solving air pollution management problems.

6. REFERENCES


