Development of a Fuzzy Pattern Recognition Model for Air Quality Assessment of Howrah City

Abhishek Upadhyay¹, Kanchan¹, Pramila Goyal², Anjaneyulu Yerramilli³, Amit Kumar Gorai¹*

⁰ Environmental Science & Engineering Group, Birla Institute of Technology, Mesra, Ranchi-835215, India
¹ Centre for Atmospheric Sciences, IIT Delhi, Delhi-110016, India
² Trent Lott Geospatial and Visualisation Research Centre, Mississippi, Jackson State University, 1230 Raymond Road, Jackson, MS39204, USA

ABSTRACT

The district of Howrah is one of the most highly industrialized districts in West Bengal, India. Howrah City continues to suffer from poor ambient air quality due to the dense siting of small scale industries without air pollution management, huge traffic congestion and high levels of human settlement. This paper presents the trends of air pollution concentration (O₃, CO, SO₂, NO₂, and PM₁₀) in Howrah City, and also demonstrates a new methodology for air quality assessment using an AHP coupled fuzzy pattern recognition model. The annual average of PM₁₀ concentration has decreased from 2009 (185.57 ± 121.16 µg/m³) to 2011 (160.01 ± 117.32 µg/m³). A similar trend was observed for the CO concentration. The eight-hour average concentration of CO in 2011 (0.939 ± 0.632 mg/m³) was found to be lower than that in 2009 (1.59 ± 0.72 mg/m³), while the reverse trend was observed for SO₂ and NO₂. The annual average concentration of SO₂ increased from 2009 (17.68 ± 20.92 µg/m³) to 2011 (43.048 ± 31.47 µg/m³). The annual average concentration of NO₂ increased from 2009 (63.87 ± 39.73 µg/m³) to 2011 (78 ± 61.51 µg/m³). There was no uniform trend observed in the annual eight-hour average concentration of ozone. An approach was developed in this study to determine fuzzy air quality based on the observed air pollution concentration. This will help to identify the air pollution control measures that are required in a certain area. The proposed method is a multi-pollutant aggregation method with varying weighting, and has the capability to consider subjective factors like sensitivity and population density. The concentrations of the five air pollutant parameters (O₃, CO, SO₂, NO₂, and PM₁₀) were used to develop the model for air quality assessment.

Keywords: Air Quality assessment; Fuzzy pattern recognition; Optimisation.

INTRODUCTION

Ambient air quality has been given continuous attention in the whole world for many years and air pollution not only affects human health, but also restricts the city’s economic development. Worldwide, many cities continuously assess air quality using monitoring networks designed to measure and record air pollution concentrations at several points deemed to represent exposure of the population to these pollutants. There has been an increasing used by national and international agencies to inform environmental policies, and quantification of the impact of air pollution on public health has gradually become a critical component in policy discussions as governments weigh options for the control of pollution. Because of the health effects associated with these conditions it is of great interest to characterize the status of air quality for several pollutants such as O₃, CO, SO₂, NO₂, particulate matter, etc. Although most of the countries in the world designed guidelines for the measurement methods for various pollutants, there are differences among the national monitoring system in terms of types and number of monitored pollutants. The purpose of such monitoring is several fold - to identify the locations of maximum air pollutant concentrations (Harrison et al., 1998; Bladauf et al., 2001, 2002), to compare measured values against guidelines or standards, to assess the success or otherwise of pollution reduction strategies and to monitor medium and long-term trends, among others. The monitoring data (pollutant concentrations) conveyed to the public through periodic reports that include concentration time series, and a comparison of the measured values for each pollutant against the applicable guideline. The practice of risk assessment is evolving away from a focus on the potential of a single pollutant in one environmental medium towards integrated assessments involving suites of pollutants in several media that may cause a variety of adverse effect on humans.
plants, animals, or even effects on ecological systems and their processes and functions (USEPA, 1997). Baldauf et al. (2002) proposed risk assessment method for establishing air quality monitoring networks based on maximum concentration impacts or maximum total population.

Current research indicates that guideline values cannot be regarded as threshold values below which a zero adverse response may be expected (Koenig and Mar, 2000; WHO, 2000a; Gent et al., 2003). Therefore, the simplistic comparison of measured values against guidelines may mislead unless appropriately quantified. Thus, a more sophisticated tool has been developed to communicate the health risk of ambient concentrations by using an air pollution (or air quality, AQI) index (API). Various air quality indices have been developed to integrate air quality variables worldwide (ORAQI, 1970; USEPA, 1976; Ontario, 1991; GVAQI, 1997; Malaysia, 1997; UK, 1998; USEPA, 1998) but most of the methods have different characteristics not only with the type of aggregation but also with the number and type of pollutants.

Bishoi et al. (2009) proposed an approach for AQI based on factor analysis. The proposed index was obtained as a function of the first three principal components of the observed concentration matrix and had no relation to the health of people. A multi-pollutant index was proposed by Cogliani in 2001. The index was based only on the hourly highest concentrations in a day of NO₂, CO and O₃. A multi-pollutant index based on the combined levels of three pollutants, PM; SO₂; NO₂, respecting WHO guidelines for air quality, was also presented by Gurjar et al. (2008). A completely different approach was presented by Cairncross et al. (2007) in 2007 for communicating the air pollution health risk. Khanna (2000) proposed an index of pollution based on the epidemiological dose-response function associated with each pollutant, together with the welfare loss due to pollution exposure. Swamee and Tyagi (1999) reported the concepts of “ambiguity” and “eclipsicity”, which was originally introduced by Ott in 1978. Lagona (2005) and Chiu et al. (2007) had proposed an approach for determining air pollution index by means of the latent factors of a Hidden Markov Model.

Fuzzy logic is also applied by many researchers for air quality assessment. It allows processing semantics of natural language in these science branches. The main characteristics of natural language semantics is its uncertainty. Uncertainty in fuzzy sets theory can be quantified (Zadeh, 1978; Lee, 1990). Sadiq and Tesfamariam (2009) proposed the concept of intuitionist fuzzy set was applied to handle both vagueness and ambiguity related uncertainties in the environmental decision-making process. Fisher (2003) illustrated that the use of fuzzy sets formalizes the underlying uncertainty and therefore leads to better decision making. Li et al. (2008) proposed an integrated fuzzy-stochastic modelling approach for quantifying uncertainties associated with both source/medium conditions and evaluation criteria and thus assessing air pollution risks. Hajek and Olej (2009) presented a design of AQIs based on tree/cascade hierarchical fuzzy inference systems.

The objective of this work is to develop a fuzzy synthetic evaluation (FSE) coupled analytical hierarchical process (AHP) model for conducting the air pollution cumulative risk assessment associated with multi-source and multi-factor under uncertainty. The developed model will be applied for air quality assessment in Howrah city, West Bengal, India. To demonstrate the application, common air pollutants like PM₁₀, SO₂, CO, O₃ and NO₂ were considered for determination of fuzzy air quality index. Since different air pollutants have varying in health impacts and hence the weights of air pollutants considered in determination of air quality index were different. The weighting of air pollutant parameters was determined using analytical hierarchical process (AHP).

**STUDY AREA**

Howrah is an industrial city situated on west bank of the Hoogli River in West Bengal, India. Air pollution monitoring station at Howrah city is shown in Fig. 1. Howrah district

![Fig. 1. Map showing district boundary of Howrah district, monitoring station and small scale industry zones.](image)
lies between 22°48′N and 22°12′N latitudes and between 88°23′E and 87°50′E longitudes. As per 2011 population census, the total population of Howrah district was 4,273,099. The population growth rate in Howrah was low (13.31%) as compared to national growth rate (21.15%) in the last decade (2001 to 2011). The initial provisional data released by census India 2011, shows that density of Howrah district for 2011 is 3,300 people per sq. km. In 2001, Howrah district density was at 2,913 people per sq. km. Howrah district administers 1,467 square kilometres of areas. The district of Howrah once occupied the prime position and was considered to be one of the highly industrialized districts in West Bengal. A number of industries like engineering, casting, steel fabrication, ship building, consumer goods industries, construction pressure die casting, forging, electric installations, manufacturing of industrial electrical goods accessories etc. had a large concentration in the district. Howrah city continues to defile ambient air quality due to dense conglomerate of small scale industries without air pollution management, huge traffic congestion and human settlement.

This paper presents air pollution trends at the monitoring station (located in Howrah city) for three consecutive years (2009 to 2011). The air pollution data (2009 to 2011) of Howrah city were used for determination of fuzzy air quality index (FAQI). The data used for assessment was taken from the West Bengal Pollution Control Board (WBPCB) website (Source: http://emis.wbpcb.gov.in/airquality/citizenreport.do). The pollution concentrations of five major air pollution parameters (SO\textsubscript{2}, NO\textsubscript{2}, PM\textsubscript{10}, O\textsubscript{3}, and CO) were selected to represent air quality assessment. The diurnal average (24 hours) concentrations of SO\textsubscript{2}, NO\textsubscript{2}, PM\textsubscript{10} and maximum daily 8 hourly average concentrations O\textsubscript{3}, and CO during 2009, 2010, and 2011 are represented in Figs. 2(a)–2(e).

Figs. 2(a)–2(e) clearly reveals that the observed concentration of SO\textsubscript{2}, NO\textsubscript{2}, PM\textsubscript{10}, and CO were high during winter months (November to February) in each year. But the concentration level of O\textsubscript{3} was found to be higher during summer. This is because O\textsubscript{3} concentration depends on the precursor sources (NO\textsubscript{x}, VOC’s) and solar radiation. Since the solar radiation was normally high during summer months and thus the formation of O\textsubscript{3} during summer was also high. The concentration levels of SO\textsubscript{2}, NO\textsubscript{2}, PM\textsubscript{10}, and CO were found to be low during summer due to unstable atmospheric conditions.

The descriptive statistics of the pollutant concentration variations is represented in Table 1. The annual average of PM\textsubscript{10} concentration in 2011 (160.01 ± 117.32 µg/m\textsuperscript{3}) was found to be lower than 2010 (175.66 ±138.52 µg/m\textsuperscript{3}). The annual average PM\textsubscript{10} concentration level in 2010 was found to be lower than 2009 (185.57 ± 121.16 µg/m\textsuperscript{3}). This clearly indicates that the annual average concentration of PM\textsubscript{10} has been decreased during 2009 to 2011. Similar trend was observed for CO concentration. The annual average of maximum daily 8 hours average CO concentration in 2011 (0.939 ± 0.632 mg/m\textsuperscript{3}) was found to be lower than 2010 (1.344 ± 0.961 mg/m\textsuperscript{3}). The annual average of 8 hours average concentration of CO in 2010 was found to be lower than 2009 (1.59 ± 0.72 mg/m\textsuperscript{3}). Fig. 2(d) clearly reveals that the maximum 8 hours average concentration of CO exceeded the permissible limit [National Ambient Air

Fig. 2. Diurnal variation of pollutants concentration in three consecutive years (a) SO\textsubscript{2} (b) NO\textsubscript{2} (c) PM\textsubscript{10} (d) CO (e) O\textsubscript{3}.
Table 1. Descriptive statistics of air pollutant concentration variation.

<table>
<thead>
<tr>
<th>Year 2009</th>
<th>SO$_2$ (µg/m$^3$)</th>
<th>NO$_2$ (µg/m$^3$)</th>
<th>PM$_{10}$ (µg/m$^3$)</th>
<th>CO (µg/m$^3$)</th>
<th>O$_3$ (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>1.49</td>
<td>6.00</td>
<td>27.8</td>
<td>0.42</td>
<td>0.003</td>
</tr>
<tr>
<td>Maximum</td>
<td>132.56</td>
<td>202.0</td>
<td>692</td>
<td>3.83</td>
<td>0.103</td>
</tr>
<tr>
<td>Mean</td>
<td>17.68</td>
<td>63.87</td>
<td>185.57</td>
<td>1.59</td>
<td>0.033</td>
</tr>
<tr>
<td>Standard Dev</td>
<td>20.92</td>
<td>39.73</td>
<td>121.16</td>
<td>0.72</td>
<td>0.020</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year 2010</th>
<th>SO$_2$ (µg/m$^3$)</th>
<th>NO$_2$ (µg/m$^3$)</th>
<th>PM$_{10}$ (µg/m$^3$)</th>
<th>CO (µg/m$^3$)</th>
<th>O$_3$ (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>4.91</td>
<td>13.00</td>
<td>30.4</td>
<td>0.29</td>
<td>0.0025</td>
</tr>
<tr>
<td>Maximum</td>
<td>224.02</td>
<td>349.96</td>
<td>224.02</td>
<td>5.04</td>
<td>0.35</td>
</tr>
<tr>
<td>Mean</td>
<td>28.954</td>
<td>75.84</td>
<td>175.66</td>
<td>1.344</td>
<td>0.0213</td>
</tr>
<tr>
<td>Standard Dev</td>
<td>29.74</td>
<td>49.25</td>
<td>138.52</td>
<td>0.961</td>
<td>0.0445</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year 2011</th>
<th>SO$_2$ (µg/m$^3$)</th>
<th>NO$_2$ (µg/m$^3$)</th>
<th>PM$_{10}$ (µg/m$^3$)</th>
<th>CO (µg/m$^3$)</th>
<th>O$_3$ (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>4.71</td>
<td>5.08</td>
<td>31.71</td>
<td>0.12</td>
<td>0.0051</td>
</tr>
<tr>
<td>Maximum</td>
<td>215.64</td>
<td>426.03</td>
<td>1281.57</td>
<td>3.96</td>
<td>0.426</td>
</tr>
<tr>
<td>Mean</td>
<td>43.048</td>
<td>78.0</td>
<td>160.012</td>
<td>0.939</td>
<td>0.078</td>
</tr>
<tr>
<td>Standard Dev</td>
<td>31.471</td>
<td>61.51</td>
<td>117.32</td>
<td>0.632</td>
<td>0.061</td>
</tr>
</tbody>
</table>
Quality Standard (NAAQS) for CO is 2 mg/m$^3$ (maximum 8 hrs. average) in India during winter months in every observed years. But the reverse trend was observed for SO$_2$ and NO$_2$. The annual average concentration of SO$_2$ has been increased from 17.68 µg/m$^3$ in 2009 to 43.048 µg/m$^3$ in 2011. The best thing is that the annual average concentration of SO$_2$ was within the permissible limit (50 µg/m$^3$) in each of the observed year. But the Fig. 2(a) indicates that the diurnal average concentration of SO$_2$ exceeded the permissible limit (80 µg/m$^3$) during winter months. The annual average concentration of NO$_2$ has been increased from 63.87 µg/m$^3$ in 2009 to 78 µg/m$^3$ in 2011. The worst thing is that the annual average concentration of NO$_2$ exceeded the permissible limits (40 µg/m$^3$). Fig. 2(b) indicates that the diurnal average concentration of NO$_2$ exceeded the permissible limit (80 µg/m$^3$) during winter months. The annual average of maximum daily 8 hours average concentration of ozone was found to decrease from 0.033 ppm in 2009 to 0.0213 ppm in 2010 whereas the same was increased to 0.078 ppm in 2011. Thus there was no uniform trend observed in ozone concentration. Fig. 2(e) clearly indicates that the 8 hours average concentration of ozone was within the permissible limit (0.0509) in every season of each observed years except one occasion in 2011. This is because ozone is secondary pollutant and the formation and destruction of ozone depends on many factors like NO$_x$ concentration, solar radiation, volatile organic compounds (VOC’s) concentration etc. Table 1 also clearly reveals that the fluctuation of SO$_2$, NO$_2$ and O$_3$ concentration during the observed years, which has increased continuously as the standard deviation value has been increased. But there was no uniform trend in variation (standard deviation) for PM$_{10}$ and CO concentration.

Meteorological factors play an important role in variations of air pollution concentration, particularly in pollutant transport irrespective of their entry into the environment. The seasonal windrose diagrams of the monitoring location are represented in Figs. 3(a)–3(i) for three consecutive years 2009 to 2011. Windrose diagrams reveal that the average wind velocity were in the range of 0–1 m/s (calm conditions) or 1–2 m/s in all the three seasons except few occasions during summer months. But the frequency of wind velocity (in the range of 1–2 m/s) was very high during summer months (March to June) in comparison to that of the winter months (November to February). This clearly shows an unstable atmospheric condition in summer months and stagnant air in winter month. Increased atmospheric stability leads to less general circulation and thus more stagnant air masses. Stagnant air masses allow more accumulation of pollutants in any given area. The frequency of calm conditions (0–1 m/s) during winter months was relatively high in comparison to that of the summer and monsoon months (July–October). The prevailing calm conditions

![Wind-rose diagrams](image-url)

**Fig. 3.** Wind-rose diagram for monitoring location. (a) Summer 2009 (b) Monsoon 2009 (c) Winter 2009 (d) Summer 2010 (e) Monsoon 2010 (f) Winter 2010 (g) Summer 2011 (h) Monsoon 2011 (i) Winter 2011.
facilitate more stability to atmosphere and consequently slow dispersion of pollutants generated and helps in build-up of pollutants in vicinity of the pollutant sources. As a result the mixing of the pollutants in summer months was high in comparison to that of the winter months. Again, the variation in pollution concentration may be due to the transport of the pollution away from the sources with the influence of wind movement. Fig. 1 clearly indicates that most of the small scale industries are situated in the north and south of the air pollution monitoring location. The seasonal wind-rose depicts that the prevailing wind direction is from the Northwest (NW) sectors during winter season in each of the three observed years whereas the predominant direction was towards SE during summer and monsoon seasons. Thus there is little influence in pollution concentration in monitoring location due to transport of pollutants away from the sources. Besides this, the monsoons result in large amount of precipitation to lower atmospheric pollution via associated wet deposition processes.

**DEVELOPMENT OF FUZZY PATTERN RECOGNITION MODEL**

In this work, fuzzy pattern recognition model was developed on the basis of the breakpoint concentration level of pollutants assigned in United States Environmental Protection Agency (USEPA, 2006) air quality index guideline for reporting daily air quality Index. The Air Quality Index...
(AQI) was divided into five levels (Good: 0–50; Moderate: 50–100; Unhealthy for Sensitive Groups: 101–150; Unhealthy: 151–200; Very unhealthy: 201–300; and Hazardous: 301–500) with a yardstick that runs from 0 to 500. Lower AQI value represents good air quality whereas the higher AQI value represents poor air quality. Thus the higher AQI value is having greater health concern. The present model is demonstrated with only five criteria pollutants, namely, PM$_{10}$, SO$_2$, CO, O$_3$ and NO$_2$ due to limitations in data availability. The USEPA AQI system is based on the maximum operator function [AQI = Max (AQI$_1$, AQI$_2$, ..., AQI$_n$)]. That is, there is no aggregation of multi-pollutant in the existing system and hence no synergistic effects represented in the AQI. AQI values and their corresponding pollution concentration range with respect to above mentioned five pollutants are represented in Table 2.

Based on Table 2, the concentration intervals were set as for classifying representative values ($e_i$) in fuzzy pattern recognition model. Since the model was developed for the air quality assessment of a residential/commercial zone and thus the benchmark ($p_i$) values were considered according to the permissible value for residential/commercial zone (CPCB, New Delhi, 2009). Six risk classes (‘Clean air’, ‘Good’, ‘Unhealthy for sensitive people’, ‘Unhealthy’, ‘Very unhealthy’ and ‘Hazardous’) were defined to represent air quality index in the model. The classifying representative values ($e_i$) in fuzzy pattern recognition model are represented in Table 2.

Here a new fuzzy pattern recognition model is proposed for air quality assessment. Based on the above classifying representative values new fuzzy air quality index (FAQI) classes were developed for the evaluation of air quality status in the monitoring location. The standards range of fuzzy air quality index (FAQI) was determined on the basis of benchmark values ($p_i$) defined in Table 3. Evaluation of air quality impact can be regarded as identification of the level to which a monitoring location belongs. So it is actually a pattern recognition problem. Each of the pollutants considered in the model has varying health impacts and hence their relative weights are also different. The weights of the pollutant parameters were determined using analytical hierarchical process (AHP) as explained in the next section. Based on Table 3, the standard value matrix concentrations, $Y_{ih}$ is formed as:

\[
Y_{ih} = \begin{bmatrix}
SO_2 \\
NO_2 \\
PM_{10} \\
CO \\
O_3
\end{bmatrix}
\]

\[
= \begin{bmatrix}
0 & 0 & 0 & 0 & 0.34 \\
0 & 0 & 0 & 0 & 0.64 \\
54 & 154 & 254 & 354 & 424 \\
4.4 & 9.4 & 12.4 & 15.4 & 30.4 \\
0.059 & 0.075 & 0.095 & 0.115 & 0.404 & 0.604
\end{bmatrix}
\]

** Areas are required to report the AQI based on 8-hour ozone values. However, there are areas where an AQI based on 1-hour ozone values would be more protective. In these cases the index for both the 8-hour and the 1-hour ozone values may be calculated and the maximum AQI reported.

* NO$_2$ has no short-term NAAQS and can generate an AQI only above a value of 200.

** 8-hour O$_3$ values do not define higher AQI values ($\geq$ 301). AQI values of 301 or higher are calculated with 1-hour O$_3$ concentrations.

### Table 2. AQI and corresponding concentrations interval of pollution factors.

<table>
<thead>
<tr>
<th>Pollutants</th>
<th>0–50</th>
<th>51–100</th>
<th>101–150</th>
<th>151–200</th>
<th>201–300</th>
<th>301–500</th>
</tr>
</thead>
<tbody>
<tr>
<td>SO$_2$ (ppm)</td>
<td>0.000–0.034</td>
<td>0.035–0.144</td>
<td>0.145–0.224</td>
<td>0.225–0.304</td>
<td>0.305–0.604</td>
<td>0.605–1.004</td>
</tr>
<tr>
<td>NO$_2$ (ppm)*</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.65–1.24</td>
<td>1.25–2.04</td>
<td></td>
</tr>
<tr>
<td>PM$_{10}$ (µg/m$^3$)</td>
<td>0–54</td>
<td>55–154</td>
<td>155–254</td>
<td>255–354</td>
<td>355–424</td>
<td>425–604</td>
</tr>
<tr>
<td>CO (ppm)</td>
<td>0–4.4</td>
<td>4.5–9.4</td>
<td>9.5–12.4</td>
<td>12.5–15.4</td>
<td>15.5–30.4</td>
<td>30.5–50.4</td>
</tr>
<tr>
<td>O$_3$ (ppm)</td>
<td>0.000–0.059</td>
<td>0.060–0.075</td>
<td>0.076–0.095</td>
<td>0.096–0.115</td>
<td>0.116–0.404</td>
<td>0.405–0.604**</td>
</tr>
</tbody>
</table>

** Areas are required to report the AQI based on 8-hour ozone values. However, there are areas where an AQI based on 1-hour ozone values would be more protective. In these cases the index for both the 8-hour and the 1-hour ozone values may be calculated and the maximum AQI reported.

* NO$_2$ has no short-term NAAQS and can generate an AQI only above a value of 200.

** 8-hour O$_3$ values do not define higher AQI values ($\geq$ 301). AQI values of 301 or higher are calculated with 1-hour O$_3$ concentrations.

### Table 3. Classifying representative values and benchmarks.

<table>
<thead>
<tr>
<th>Pollution Factor</th>
<th>Classifying Representative Values ($e_i$)</th>
<th>NAAQS (CPCB, 2009, New Delhi, India)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_1$</td>
<td>$e_2$</td>
<td>$e_3$</td>
</tr>
<tr>
<td>SO$_2$ (ppm)</td>
<td>0</td>
<td>0.034</td>
</tr>
<tr>
<td>NO$_2$ (ppm)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PM$_{10}$ (µg/m$^3$)</td>
<td>0</td>
<td>0.64</td>
</tr>
<tr>
<td>CO (ppm)</td>
<td>0</td>
<td>0.059</td>
</tr>
<tr>
<td>O$_3$(ppm)</td>
<td>0</td>
<td>0.059</td>
</tr>
</tbody>
</table>

* 24 hours average concentration.

** 8 hours average concentration.
Here \( Y_{i,h} \) is standard value of level \( h \) with regard to factor \( i \); \( i = 1, 2, 3, 4, 5 \) denotes five air pollutants \( \text{SO}_2, \text{NO}_2, \text{PM}_{10}, \text{CO} \) and \( \text{O}_3 \) respectively and \( h = 1, 2, \ldots, 7 \) denotes number of standard levels of concentration of these air pollutants to determine criteria for air quality assessment.

Membership degree of \( Y_{i,h} \) with regard to zero pollutant concentration is \( S_{i,h} \) can be computed using Eq. (1) as:

\[
S_{i,h} = \frac{Y_{i,h} - Y_{i,1}}{Y_{i,7} - Y_{i,1}}
\]

In the above equation, \( Y_{i,7} \) and \( Y_{i,1} \) represents the concentration values of the last level (seventh level) and first level as given in the \( S_{i,h} \) matrix. Calculation of \( S_{i,h} \) value using standard value matrix, \( Y_{i,h} \), gave membership degree matrix, \( S_{i,h} \) of standard values. This will reflect continuous transition of harmful ability of pollutant which varies from 0 to 1. By using above equation, the relative membership degree matrix \( S \) can be derived as:

\[
S_{i,h} = \begin{bmatrix}
0 & 0.03 & 0.14 & 0.22 & 0.3 & 0.6 & 1 \\
0 & 0 & 0 & 0.31 & 0.6 & 1 \\
0 & 0.09 & 0.25 & 0.42 & 0.59 & 0.7 & 1 \\
0 & 0.09 & 0.19 & 0.24 & 0.31 & 0.6 & 1 \\
0 & 0.1 & 0.12 & 0.16 & 0.19 & 0.67 & 1
\end{bmatrix}
\]

Similarly, air pollution monitoring data (24/8 hours average air pollution concentration) of pollutants can be represented in matrix, \( X \). If the number of samples (24/8 hours average air pollution concentration) for air quality assessment are \( n \) and the number of air pollutant parameters reflecting the air quality are \( m \) (five in this case), the concentration matrix (\( X \)) for the samples can be written as:

\[
X = (x_{ij})_{n \times m} = \begin{bmatrix}
x_{11} & x_{12} & \ldots & x_{1m} \\
x_{21} & x_{22} & \ldots & x_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
x_{n1} & x_{n2} & \ldots & x_{nm}
\end{bmatrix}
\]

In the above matrix, \( x_{ij} \) is air pollution concentration value of \( j \)th pollutant parameter of \( i \)th sample.

The relative membership degree \( (r_{ij}) \) of air pollution concentration value of \( j \)th pollutant parameter of \( i \)th sample can be determined by Eq. (2) as given below:

\[
r_{ij} = \frac{(x_{ij} - y_{i,1})}{(y_{i,7} - y_{i,1})}
\]

The membership degree \( (r_{ij}) \) for a sample having pollution concentration \( (x_{ij}) \) less than that of the concentration level 1 (i.e., \( x_{ij} < y_{i,1} \)) is zero.

Similarly, the membership degree \( (r_{ij}) \) for a sample having concentration \( (x_{ij}) \) higher than that of the concentration level 7 (i.e., \( x_{ij} > y_{i,7} \)) is one.

By using above equation, the relative membership degree matrix \( R \) can be derived as:

\[
R = (r_{ij})_{n \times m} = \begin{bmatrix}
r_{11} & r_{12} & \ldots & r_{1m} \\
r_{21} & r_{22} & \ldots & r_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
r_{n1} & r_{n2} & \ldots & r_{nm}
\end{bmatrix}
\]

where \( r_{ij} \) is the relative membership degree of \( j \)th pollutant parameter (\( j = 1, 2, \ldots, m \)) of \( i \)th sample (\( i = 1, 2, \ldots, n \)). Here \( m \) is equal to five as the model considered only five air pollutant parameters. In matrix \( R \), if \( r_{ij} = 1 \), the alternative \( i \) is the worst and if \( r_{ij} = 0 \), the alternative \( i \) is the best according to the standard matrix \( S \) defined above. Supposing that there is a best alternative in which all the \( m \) pollutant’s membership degrees are equal to 0, denoted by \( G = (g_1, g_2, \ldots, g_m) = (0, 0, \ldots, 0) \), the worst alternative is expressed as \( B = (b_1, b_2, \ldots, b_m) = (1, 1, \ldots, 1) \). In this case, the decision-making problem becomes a fuzzy pattern recognition problem, i.e., evaluating to class of the membership degree each alternative in matrix \( R \). In matrix \( R \), alternative \( i \) can be expressed as represented in Eq. (3)

\[
r_i = (r_{i1}, r_{i2}, \ldots, r_{im})
\]

The distance of sample \( i \) to level \( h \) can be described by Eq. (4).

\[
d_{i,h} = \sqrt{\sum_{j=1}^{m} [w_j (r_{ij} - s_{i,h})]^p}
\]

In Eq. (4), \( p \) is a distance parameter. For \( p = 1 \), the distance is called Hamming distance and \( p = 2 \), the distance is called Euclidean distance, which are commonly used for degree of differences in impact. \( W_j \) is the relative weights of \( j \)th air pollutant parameter. This is determined using analytical hierarchical process (AHP). This is explained in detail in the next section. In the view of fuzzy sets, the membership degree \( u_{ih} \) may be regarded as weight for distance \( d_{i,h} \). The synthetic weight distance will better describe the differences in air pollution impacts of \( i \)th sample (existing condition) from \( h \)th level (standard level defined). The weighted distance \( (D_{i,h}) \) to level \( h \) of sample \( i \), determined by Eq. (5) as:

\[
D_{i,h} = u_{ih} d_{i,h} = u_{ih} \sqrt{\sum_{j=1}^{m} [w_j (r_{ij} - s_{i,h})]^p}
\]

where \( u_{ih} \) is membership degree of \( i \)th sample belonging to \( h \)th level; \( i = 1, 2, 3, \ldots, n \), and \( h = 1, 2, \ldots, 7 \). According to fuzzy set theory, \( u_{ih} \) is subjected to the following constraint as represented in Eq. (6):

\[
\sum_{h=1}^{7} u_{ih} = 1
\]

In order to determine the value of \( u_{ih} \), the following objective function is established:
\[
F(u_{i,h}) = \sum_{h=1}^{7} D_{i,h}^2
\]  
(7)

Using the Eqs. (6) and (7), the following Lagrange function can be derived:

\[
L(u_{i,h}, \lambda_i) = \sum_{h=1}^{7} r_{i,h}^2 d_{i,h}^2 - \lambda_i \left( \sum_{h=1}^{7} u_{i,h} - 1 \right)
\]  
(8)

where \( \lambda_i \) is Langrange's multiplier. Setting the partial derivatives of the Langrange function with respect to each and equal to zero, respectively, i.e.,

\[
\frac{\partial L(u_{i,h}, \lambda_i)}{\partial u_{i,h}} = 0
\]  
(9)

\[
\frac{\partial L(u_{i,h}, \lambda_i)}{\partial \lambda_i} = 0
\]  
(10)

It follows that the formula for calculating the membership degree of sample \( i \) belonging to level \( h \) is:

\[
u_{i,h} = \frac{1}{d_{i,h}^2 \sum_{i=1}^{7} d_{i,h}^2} \left[ d_{i,h} \neq 0 \right]
\]  
(11)

when \( d_{i,h} = 0 \), i.e., \( r_{i,h} = s_{i,h} \), it means that sample \( i \) completely belongs to level \( h \), so

\[
u_{i,h} = 1
\]  
(12)

The membership degree matrix of each sample \( i \) belonging to each level can be represented by matrix 'U' as follows:

\[U = (u_{i,h})_{n \times 7}
\]  
(13)

Eq. (13) is actually a 7 level fuzzy pattern recognition model, where \( h = 1, 2, ..., 7 \). This will give the membership degree matrix of each sample belonging to each level is as follow.

\[
U = \begin{bmatrix}
u_{1,1} & \nu_{1,2} & \cdots & \nu_{1,7} \\
u_{2,1} & \nu_{2,2} & \cdots & \nu_{2,7} \\
\vdots & \vdots & \ddots & \vdots \\
\nu_{n,1} & \nu_{n,2} & \cdots & \nu_{n,7}
\end{bmatrix}
\]

Thus the membership degree matrix \( (U_s) \) for standard value matrix \( S \) is as follow:

\[
U_s = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

Above membership degree matrix \( (U_s) \) clearly reveals that the pollutants concentration in each level of the standard value matrix \( S \) exactly belongs to the level, that is, when \( i = h \), \( r_{i,h} = s_{i,h} \) and hence \( u_{i,h} = 1 \).

However, the evaluation level of a sample cannot be determined directly in the matrix, so the rank feature value is defined to solve the problem. Express the distribution of the membership degree of sample \( i \) on the axis of level variable \( h \). Rank feature value for the sample \( i \) can be computed using Eq. (15) as follows:

\[
H_i = \sum_{h=1}^{7} u_{i,h} \times h
\]  
(15)

The rank feature of row vector \( i \), \( H_i \), can be considered as fuzzy pattern air quality index for sample \( i \). This rank value can be determined for standard concentration which will give evaluation level that will give a basis comparison with sample. This can determine air quality of a particular region.

The rank matrix for the above standard membership degree matrix \( (U_s) \) is represented by matrix \( H_s \) as follows:

\[H_s = \begin{bmatrix}
1 \\
2 \\
3 \\
4 \\
5 \\
6 \\
7
\end{bmatrix}
\]

Thus, the scale for fuzzy air quality index (FAQI) for the evaluation of sample is represented in Table 4.

DETERMINATION OF RELATIVE WEIGHTS OF AIR POLLUTANT PARAMETERS

Analytical Hierarchical Process (AHP) is an approach to decision making that involves structuring multiple choice criteria onto a hierarchy, assessing the relative importance of these criteria. The foundation of the Analytic Hierarchy Process (AHP) is a set of axioms that carefully delimits the scope of the problem environment (Satty, 1986). It is based...
on the well-defined mathematical structure of consistent matrices and their associated right eigenvector’s ability to generate true or approximate weights (Satty, 1980, 1994). The AHP methodology compares criteria, or alternatives with respect to a criterion, in a natural, pair wise mode. To do so, the AHP uses a fundamental scale of absolute numbers that have been proven in practice and validated by physical and decision problem experiments. The fundamental scale has been shown to be a scale that captures individual preferences with respect to quantitative and qualitative attributes just as well or better than other scales (Satty, 1980, 1994). It converts individual preferences into ratio scale weights that can be combined into a linear additive weight \( w(a) \) for each alternative \( a \). The resultant \( w(a) \) can be used to compare and rank the alternatives, and hence assist the decision maker in making a choice. Given that the three basic steps are reasonable descriptors of how an individual comes naturally to resolve a multi-criteria decision problem, then the AHP can be considered to be both a descriptive and prescriptive model of decision making. The AHP is perhaps the most widely used decision making approach in the world today. Its validity is based on the many hundreds (now thousands) of actual applications in which the AHP results were accepted and used by the cognizant decision makers (DMs) (Satty, 1994).

Each pollutant parameter is of different importance in relation to health impact and hence different weights are attributed to different pollutant parameters. The most significant factor has a higher weight and vice-versa. The pollutant’s weights are determined using analytical hierarchical process (AHP). The detailed methodology of this is explained below.

The weights of individual pollutants can be found out using Analytical Hierarchy Process (AHP). AHP is a systematic method for comparing a list of objectives or alternatives. This method form a pair-wise comparison matrix ‘\( A \)’ as shown below, where the number in the \( i^{th} \) row and \( j^{th} \) column gives the relative importance of individual air pollutant \( P_i \) as compared with \( P_j \).

The comparison matrix generated by author’s expertise using Satty’s scale (Satty’s, 1980) is shown below in matrix \( A \). The relative weights can be improved by taking the experts views.

<table>
<thead>
<tr>
<th>[ R ]</th>
<th>[ \text{PM}_{10} ]</th>
<th>[ \text{SO}_2 ]</th>
<th>[ \text{NO}_2 ]</th>
<th>[ \text{O}_3 ]</th>
<th>[ \text{CO} ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \text{PM}_{10} ]</td>
<td>1</td>
<td>1/2</td>
<td>1/3</td>
<td>1/3</td>
<td>1/2</td>
</tr>
<tr>
<td>[ \text{SO}_2 ]</td>
<td>2</td>
<td>1</td>
<td>11/4</td>
<td>2/3</td>
<td>1</td>
</tr>
<tr>
<td>[ \text{NO}_2 ]</td>
<td>3</td>
<td>4/5</td>
<td>1</td>
<td>1</td>
<td>11/4</td>
</tr>
<tr>
<td>[ \text{O}_3 ]</td>
<td>3</td>
<td>11/2</td>
<td>1</td>
<td>1</td>
<td>11/4</td>
</tr>
<tr>
<td>[ \text{CO} ]</td>
<td>2</td>
<td>1</td>
<td>4/5</td>
<td>4/5</td>
<td>1</td>
</tr>
</tbody>
</table>

The sum of each column and then divide each column by the corresponding sum are computed to obtain the normalize weights, the normalized matrix \( N \), thus obtained is represented in matrix \( N \) as given below.

\[
N = \begin{bmatrix}
0.09 & 0.1 & 0.08 & 0.09 & 0.1 \\
0.18 & 0.2 & 0.28 & 0.17 & 0.2 \\
0.27 & 0.16 & 0.23 & 0.26 & 0.25 \\
0.27 & 0.3 & 0.23 & 0.26 & 0.25 \\
0.18 & 0.21 & 0.18 & 0.21 & 0.2 \\
\end{bmatrix}
\]

The relative weight vector \( W \) for the pollutants is given by the average of the row elements in matrix \( N \) as

\[
W = \begin{bmatrix}
W_{\text{PM}_{10}} \\
W_{\text{SO}_2} \\
W_{\text{NO}_2} \\
W_{\text{O}_3} \\
W_{\text{CO}} \\
\end{bmatrix} = \begin{bmatrix}
0.09 \\
0.21 \\
0.24 \\
0.26 \\
0.20 \\
\end{bmatrix}
\]

Thus, the sum of the weights of the pollutants obtained as

\[
\sum_{i=1}^{5} W_i = 1
\]

The Consistency Ratio (CR) of the matrix ‘\( A \)’ calculated was found to be 0.007 which is less than 0.1 as per Satty, 1980 and thus the consistency of matrix \( A \) is acceptable

**APPLICATION OF THIS MODEL FOR AIR QUALITY ASSESSMENT IN HOWRAH CITY**

Air pollution concentrations of a continuous monitoring station in Howrah city were assessed using the developed model. Air pollution concentration of the five pollutant parameters for three consecutive years (2009, 2010, and 2011) were represented in the factor value matrix \( X_1, X_2, \) and \( X_3 \), respectively as:

\[
X_1 = \begin{bmatrix}
x_{1,1} & x_{1,2} & x_{1,3} & x_{1,4} & x_{1,5} \\
x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} & x_{2,4} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
x_{365,1} & x_{365,2} & x_{365,3} & x_{365,4} & x_{365,4} \\
\end{bmatrix},
\]

\[
X_2 = \begin{bmatrix}
0.011 & 0.065 & 322.6 & 2.06 & 0.006 \\
0.007 & 0.024 & 294 & 1.519 & 0.007 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
0.0127 & 0.059 & 344 & 1.100 & 0.006 \\
\end{bmatrix}
\]

\[
X_3 = \begin{bmatrix}
\text{Table 4. Standard Class of Fuzzy Air Quality Index.}
\begin{array}{|c|c|c|c|c|c|}
\hline
\text{Scale} & 1-2 & 2-3 & 3-4 & 4-5 & 5-6 & 6-7 \\
\hline
\text{Class} & \text{Clean air} & \text{Good} & \text{Unhealthy for sensitive people} & \text{Unhealthy} & \text{Very Unhealthy} & \text{Hazardous} \\
\hline
\end{array}
\end{bmatrix}
\]
By using Eq. (2), the above matrices were converted into the following membership degree matrices $R_1$, $R_2$, and $R_3$ respectively for 2009, 2010 and 2011 as:

$$R_1 = \begin{bmatrix}
0.011 & 0.032 & 0.534 & 0.041 & 0.010 \\
0.007 & 0.012 & 0.487 & 0.030 & 0.012 \\
: & : & : & : & : \\
0.013 & 0.029 & 0.569 & 0.022 & 0.011
\end{bmatrix}$$

$$R_2 = \begin{bmatrix}
0.009 & 0.032 & 0.473 & 0.026 & 0.010 \\
0.006 & 0.026 & 0.407 & 0.031 & 0.010 \\
: & : & : & : & : \\
0.032 & 0.090 & 0.659 & 0.073 & 0.024
\end{bmatrix}$$

$$R_3 = \begin{bmatrix}
0.045 & 0.035 & 0.411 & 0.036 & 0.027 \\
0.017 & 0.025 & 0.255 & 0.023 & 0.032 \\
: & : & : & : & : \\
0.021 & 0.034 & 0.314 & 0.006 & 0.035
\end{bmatrix}$$

By using Eq. (10) and $p = 2$, the membership degree matrices $U_1$, $U_2$ and $U_3$ of the samples for consecutive three years respectively are obtained as follows:

$$U_1 = \begin{bmatrix}
0.283 & 0.303 & 0.218 & 0.131 & 0.052 & 0.009 & 0.003 \\
0.313 & 0.317 & 0.201 & 0.115 & 0.043 & 0.008 & 0.003 \\
: & : & : & : & : & : & : \\
0.284 & 0.293 & 0.217 & 0.136 & 0.056 & 0.001 & 0.004
\end{bmatrix}$$

$$U_2 = \begin{bmatrix}
0.323 & 0.315 & 0.196 & 0.111 & 0.045 & 0.008 & 0.003 \\
0.355 & 0.331 & 0.176 & 0.094 & 0.036 & 0.006 & 0.002 \\
: & : & : & : & : & : & : \\
0.207 & 0.255 & 0.253 & 0.180 & 0.087 & 0.013 & 0.005
\end{bmatrix}$$

$$U_3 = \begin{bmatrix}
0.288 & 0.347 & 0.211 & 0.107 & 0.038 & 0.006 & 0.002 \\
0.423 & 0.381 & 0.117 & 0.054 & 0.019 & 0.004 & 0.001 \\
: & : & : & : & : & : & : \\
0.385 & 0.367 & 0.144 & 0.071 & 0.027 & 0.005 & 0.002
\end{bmatrix}$$

By using Eq. (14), the rank matrices $H_1$, $H_2$, and $H_3$ for the three years are determined from the above membership degree matrices $U_1$, $U_2$ and $U_3$ respectively as:

$$H_1 = \begin{bmatrix}
2.41 \\
2.30 \\
: & : & : & : & : & : & : \\
Day 1 & Day 1 & Day 2 & Day 2 & Day 365 & Day 365 & Day 365
\end{bmatrix}$$

$$H_2 = \begin{bmatrix}
2.28 \\
2.16 \\
: & : & : & : & : & : & : \\
Day 1 & Day 2 & Day 365 & Day 365 & Day 365 & Day 365
\end{bmatrix}$$

$$H_3 = \begin{bmatrix}
2.01 \\
1.88 \\
: & : & : & : & : & : & : \\
Day 1 & Day 2 & Day 365 & Day 365 & Day 365 & Day 365 & Day 365
\end{bmatrix}$$

The detailed rank values of the above matrices are represented in Figs. 3(a)–3(c) respectively for $H_1$, $H_2$, and $H_3$.

**RESULTS AND DISCUSSION**

The fuzzy air quality index values in fuzzy pattern recognition method reflected by the relative membership degree of the pollution concentration data. The membership degree of each air pollution parameters ranged from 0 to 1, that is, zero air pollution concentration to severe air pollution condition. Since our objective is to determine the single index value by aggregating the multi-pollutant parameters. The developed fuzzy pattern recognition model has been used for aggregating the five pollutant parameters (PM$_{10}$, CO, SO$_2$, NO$_2$ and O$_3$).

The model results can be used for index calculation for standard concentration which will give evaluation level that will give a basis for comparison with the observed concentration. The standard class interval for the fuzzy air quality index (FAQI) ranged from 1 to 7 as represented in Table 4. Higher index value represents high pollution level
and higher impacts on health and vice-versa. The range of the fuzzy air quality index has been classified into 6 classes (1–2, 2–3, 3–4, 4–5, 5–6, and 6–7) with equal interval. The six classes are designated as Clean air, Unhealthy for sensitive people, Unhealthy, Very unhealthy, and Hazardous respectively as shown in Table 4. The fuzzy air quality index also determined for maximum limit concentration (NAAQS) of the five pollutant parameters for assessing the health impacts. The standard values of the five pollutant parameters are represented in the Figs. 2(a)–2(e). The fuzzy air quality index value for the NAAQS was found to be 2. The same model now used for determination of fuzzy air quality index for observed daily average concentrations as shown in the graph 2(a)–2(e). The daily average observed concentration for three years (2009, 2010, and 2011) are also represented in matrices X1, X2, and X3 respectively. The corresponding air quality indices are represented in matrices H1, H2, and H3 respectively. The index values for the three years are represented graphically in Figs. 4(a)–4(e). Figs. 4(a)–4(e) clearly reveals that the air quality index values lies mostly in two classes (Clean air and unhealthy for sensitive people) except few occasions in each of the observed years. All three graphs, also clearly indicates that the air quality level was poor during winter months (November to February). This is due to the fact that the pollution concentration level was high during winter months as already explained in the first section. Since the permissible values of fuzzy air quality index (as calculated for NAAQS) is 2, more than this can be harmful to people, if exposed. Figs. 4(a)–(c) clearly indicates that the observed fuzzy air quality indices are exceeded during winter months whereas the same are within the permissible limit during summer and monsoon months. Fig. 4(d) represents the comparative values of fuzzy air quality index for three observed years and the figure clearly says that the trend of air quality index is more or less same. Though, the concentration of few pollutants has been increased during the observed years, it does not have much influence in the overall air quality index. This is because the variations of the pollution concentration were not following the uniform trend.

CONCLUSIONS

The air quality assessment of Howrah city has been demonstrated with the air pollution data (PM10, CO, SO2, NO2 and O3) for three consecutive years (2009, 2010 and 2011). The decreasing trend was observed in annual average concentration of PM10 and CO during 2009 to 2011. But the reverse trend was observed for SO2 and NO2. There was no uniform trend for ozone concentration. As indicated above, five pollutants were used as a summary indicator of air quality, given its importance as a health determinant. For assessing the health impact of multi-pollutants parameters using a single index, fuzzy logic approach used for air quality assessment. In this study, an approach is developed to determine fuzzy air quality in monitoring location based on the air pollution concentration levels.
This study will help individuals as well as decision makers and air quality manager to make decisions or estimate the health impact in a particular city condition. Depending on the risk level (fuzzy air quality index), air quality manager can take preventive actions for reducing the level of index. The permissible value of fuzzy air quality index (FAQI) for residential areas is 2. It was observed that FAQI values are above 2 (permissible limit for residential areas) in most of the day during winter seasons (mid of October to February) in each of the observed year. FAQI values were found to be within the permissible limits in most of the day of remaining months (March to mid of October) in each of the observed year. This indicates higher degree of air pollution related health problem (like Asthma) occurs during winter months.

ACKNOWLEDGEMENT

The authors are grateful to West Bengal State Pollution Control Board, India for providing the air quality data used in this study.

REFERENCES


GVAQI (1997). Greater Vancouver Regional District Air
Quality and Source Control Department, Burnaby, BC, Canada.


National Ambient Air Quality Standards (2009). Central Pollution Control Board Notification, New Delhi, 18th November 2009.


West Bengal Pollution Control Board (http://emis.wbpcb.gov.in/airquality/citizenreport.do).


Received for review, April 11, 2013

Accepted, January 6, 2014