Application of Fuzzy Pattern Recognition Optimisation Model for Air Quality Assessment

A. K. Gorai

Environmental Science & Engineering Group,
Birla Institute of Technology, Mesra, Ranchi -835215, India
amit_gorai@yahoo.co.uk

Abstract— **Air pollution monitoring program aims to monitor pollutants concentrations and its possible adverse effects at various locations over concerned area on the basis of air quality. Traditional air quality assessment is realized using air quality indices which are determined as mean values of selected air pollutants. Thus, air quality assessment depends on strictly prescribed limits without taking into account specific local conditions (like time of exposure and sensitivity of the people) and synergic relations between air pollutants. The stated limitations can be eliminated using fuzzy logic systems. Therefore, the paper presents a design of a model for air quality assessment based on fuzzy pattern recognition.**

This paper discusses the use of fuzzy pattern recognition technique in air quality risk assessment for a number of artificial dataset prepared for the present study. To demonstrate the application, common air pollutants like PM_{10}, $PM_{2.5}$, SO_2, NOx, CO, and O_3 are used as air pollutant parameters. Different air pollutants have varying in health impact and hence in air quality, the weightage of each pollutant are different. Thus, the weightage of air pollutant parameter are determined using analytical hierarchical process (AHP).

***Keywords*— Air Quality Assessment; Fuzzy Pattern Recognition; Optimisation**

I. INTRODUCTION

Concerns about the impacts of air pollution on the public extend back as far as the 14th century (Brimblecombe, 1987 [1]). After a significant increase in deaths occurred during severe air pollution episodes in Meuse Valley, Belgium in 1930, Donora, Pennsylvania in 1948 and London, England in 1952, air quality started to become an increasingly important public health issue. The event of London, England in 1952 resulted in the first modern legislation to reduce air emissions and has been followed by continuous improvements in air quality in many areas of the world to this day. Clean air is now considered to be a basic requirement for human health and well-being. However, air pollution continues to pose a significant threat to health (WHO, 2005 [2]). In the past few decades, the Air Quality Index has been developed as a tool to communicate the health risks posed by air pollution in all over world (USEPA, 1976 and 1998 [3, 4]; Malaysia, 1997 [5]; GVAQI, 1997 [6]; Ontario, 1991 [7]; ORAQI, 1970 [8]) at the urban (city) scale to communicate air quality.

In the majority of cases, the AQI is based on the ambient concentrations of common pollutants- SO_2, PM_{10}, NO_2, CO and O_3. In a few cases $PM_{2.5}$ is considered in the calculation of the index. Considerable uncertainties are involved in the process of defining air quality for designated uses. Fuzzy technique can be successfully used to model non-linear functions, and to deal with imprecise data (Mandal et. al. 2011 [9]). Thus the advantages of fuzzy logic have been applied for air quality assessment. In this paper six air pollutant parameters are considered for the fuzzy pattern recognition model. The selection of aggregation function for single index calculation is also a difficult job.

II. METHODOLOGY

The research methodology for air quality assessment using fuzzy pattern recognition involves following steps:

A. Determination of relative weightage of air pollutant parameters

Each pollutant parameter has a predetermined, fixed and relative weight that reflects its relative importance to air quality. The most significant factors have a higher weight and vice-versa. The pollutant's weights have been determined using analytical hierarchical process (AHP). The detailed methodology is explained below to determine the relative weightage of each pollutant.

The weightage of individual pollutants can be found out using Analytical Hierarchy Process (AHP). Analytical Hierarchy Process 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 Pi as compared with Pj

The comparison matrix generated by author's expertise using Saaty's scale (Satty's, 1980 [10]) is shown below in matrix A. The relative weightage can be improved by taking the experts views.

$$A = \begin{array}{c} \\ PM_{10} \\ PM_{2.5} \\ SO_2 \\ NOx \\ O_3 \\ CO \end{array} \begin{bmatrix} PM_{10} & PM_{2.5} & SO_2 & NOx & O_3 & CO \\ 1 & 1/3 & 1/2 & 1/3 & 1/3 & 1/2 \\ 3 & 1 & 5/4 & 5/4 & 1 & 5/4 \\ 2 & 4/5 & 1 & 5/4 & 2/3 & 1 \\ 3 & 4/5 & 4/5 & 1 & 1 & 5/4 \\ 3 & 1 & 3/2 & 1 & 1 & 5/4 \\ 2 & 4/5 & 1 & 4/5 & 4/5 & 1 \end{bmatrix}$$

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.

$$N = \begin{array}{c} \\ PM_{10} \\ PM_{2.5} \\ SO_2 \\ NOx \\ O_3 \\ CO \end{array} \begin{bmatrix} PM_{10} & PM_{2.5} & SO_2 & NOx & O_3 & CO \\ 0.07 & 0.07 & 0.08 & 0.06 & 0.07 & 0.08 \\ 0.21 & 0.21 & 0.21 & 0.22 & 0.21 & 0.20 \\ 0.14 & 0.17 & 0.17 & 0.22 & 0.14 & 0.16 \\ 0.21 & 0.17 & 0.13 & 0.18 & 0.21 & 0.20 \\ 0.21 & 0.21 & 0.25 & 0.18 & 0.21 & 0.20 \\ 0.14 & 0.17 & 0.17 & 0.14 & 0.17 & 0.16 \end{bmatrix}$$

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

$$W = \begin{bmatrix} w_{PM_{10}} \\ w_{PM_{2.5}} \\ w_{SO_2} \\ w_{NO_x} \\ w_{O_3} \\ w_{CO} \end{bmatrix} = \begin{bmatrix} 0.07 \\ 0.21 \\ 0.17 \\ 0.18 \\ 0.21 \\ 0.16 \end{bmatrix}$$

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

$$\sum_{i=1}^{6} W_i = 1$$

The Consistency Ratio (CR) of the matrix 'A' calculated was found to be 0.007 which is less than 0.1 as par Satty, 1980 [10] and thus the consistency of matrix A is acceptable.

B. *Fuzzy pattern recognition optimization model*

In the assessment of air quality, six governing factors are selected for the evaluation of air quality: PM_{10}, $PM_{2.5}$, SO_2, NOx, CO, and O_3. Evaluation of air quality health impact can be regarded as identification of the level to which a sample belongs according to the concentration of six air pollutant parameters values of the sample when compared with the permissible values (maximum values) listed in Table 1. So it is actually a pattern recognition problem. Here a new fuzzy pattern recognition model is proposed for assessment of air quality.

If the number of samples for assessment is n and the number of air pollutant parameters reflecting the air quality is m, the parameter matrix for the samples can be written as:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots\cdots & x_{1m} \\ x_{21} & x_{22} & \cdots\cdots & x_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots\cdots & x_{nm} \end{bmatrix}$$

Where, x_{ij} equals the concentration value of air pollutant parameter j corresponding to sample i.

For parameter j in sample i, the relative membership degree of air quality can be defined as:

$$r_{ij} = \begin{cases} 0 & \text{for } x_{ij} < x_{minj} \\ \dfrac{x_{ij} - x_{minj}}{x_{maxj} - x_{minj}} & \text{for } x_{minj} < x_{ij} < x_{maxj} \\ 1 & \text{for } x_{ij} > x_{max} \end{cases} \quad \text{----- (1)}$$

Where, x_{maxj} and x_{minj} equal the maximum and minimum value respectively of parameter j of all the samples. The minimum and maximum values for all the air pollutants parameters are reported in Table 1. The minimum value for each air pollutant parameters are taken as zero, reflecting clean air and the maximum value has been decided on the basis of permissible concentration of the pollutant in residential areas as per the regulatory body, CPCB (NAAQS, 2009 [11]) in India. The maximum values for each air pollutant are taken as four times of their corresponding permissible concentration in residential areas for the study.

By using Eq. (1), the relative membership matrix R can be derived:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots\cdots & r_{1m} \\ r_{21} & r_{22} & \cdots\cdots & r_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & \cdots\cdots & r_{nm} \end{bmatrix}$$

TABLE 1 MINIMUM AND MAXIMUM VALUES FOR SIX AIR POLLUTANTS FACTORS

Factors	Minimum Value (x_{min}) in µg/m^3	Maximum Value (x_{max}) in µg/m^3	Permissible value in µg/m^3
Particulate Matter less than 10 µm size (PM_{10})	0	400	100
Particulate Matter less than 10 µm size ($PM_{2.5}$)	0	240	60
Sulphur dioxide (SO_2)	0	320	80
Nitrogen Oxides (NOx)	0	320	80
Ozone (O_3)	0	720	180
Carbon monoxide (CO)	0	8000	2000

Assuming the relative membership degree of sample i is u_i, the normalized weighting factor of the parameter j is w_j (represented above in matrix W), the general Euclidean distance $D(r_i)$ is used to indicate the difference between sample i and the worst air quality, for which u_i= 1. This can be determined using Eq. (2). In this paper the general Euclidean distances are not calculated for the artificial datasets.

$$D(r_i) = u_i \sqrt{\sum_{j=1}^{m} [w_j (r_{ij} - 1)]^2} \quad \text{....... (2)}$$

In order to acquire the optimized solution of u_i, the objective function is established (Chen, 1998):

$$\min \left| F(u_i) = u_i^2 \sum_{j=1}^{m} \{w_j (r_{ij} - 1)\}^2 + (1 - u_i)^2 \sum_{j=1}^{m} \{w_j r_{ij}\}^2 \right|$$

To solve, $\frac{\partial F(u_i)}{\partial u_i} = 0$, then

$$u_i = \frac{1}{1 + \sum_{j=1}^{m} \{w_j (r_{ij} - 1)\}^2 \Big/ \sum_{j=1}^{m} \{w_j r_{ij}\}^2} \quad \text{...... (3)}$$

III. AIR QUALITY IMPACT ASSESSMENT USING FUZZY PATTERN RECOGNITION OPTIMIZATION MODEL

This section will demonstrate the application of fuzzy pattern optimisation model for air quality assessment. Ten artificial dataset has been prepared assuming the concentration of each air pollutant parameters for the demonstration. The concentration values for all the six air pollutant parameter in each sample are listed in Table 2. The air quality level for each sample has been assessed to offer guidance for degree of air pollution control needed.

By using Eq. (1), the relative membership degree for the air pollutant parameters of each sample is derived and listed in Table 3. For each sample, u_i is derived using Eq. (3) and compared with the threshold level of air quality (air quality value for permissible level concentration) to see the air quality of the section. This is the fuzzy pattern recognition and optimization method for assessing the air quality.

TABLE 2 ARTIFICIAL DATASET FOR THE ASSESSMENT OF AIR QUALITY IN DIFFERENT CONDITIONS

Conditions	PM_{10}	$PM_{2.5}$	NO_2	SO_2	O_3	CO
A: 4 times of the Permissible Concentration (PC) for all the pollutants	400	240	320	320	720	8000
B: Permissible Concentration (PC) for all the pollutants	100	60	80	80	180	2000
C: Pollutants Concentration (All 6P - 10% of PC)	10	6	8	8	18	200
D: Pollutants Concentration (5P - 10% of PC & 1P-100%PC)	10	6	8	8	18	2000
E: Pollutants Concentration (5P - 90% of PC & 1P-100%PC)	90	54	72	72	162	2000
F: Pollutants Concentration (5P - 99% of PC & 1P-100%PC)	99	59	79.2	79.2	178.2	2000
G: Pollutants Concentration (5P - 99% of PC & 1P-150%PC)	99	59	79.2	79.2	178.2	3000
H: Pollutants Concentration (P -10% of PC & 1P-150%PC)	10	6	8	8	18	3000
I: Pollutants Concentration (All 6P - 125% of PC)	125	75	100	100	225	2500
J: Pollutants Concentration are considered any arbitrarily value	80	60	60	30	120	1600

The air quality index obtained from the fuzzy pattern recognition model is compared with that of the index obtained from deterministic method. The methodology of deterministic method is discussed below in the next section.

TABLE 3 RELATIVE MEMBERSHIP DEGREES OF AIR QUALITY IN VARIOUS CONDITIONS

Condition	PM_{10}	$PM_{2.5}$	NO_2	SO_2	O_3	CO	Relative member-ship degree (ui)
A	1	1	1	1	1	1	1.000
B	0.250	0.250	0.250	0.250	0.250	0.250	0.100
C	0.025	0.025	0.025	0.025	0.025	0.025	0.001
D	0.025	0.025	0.025	0.025	0.025	0.250	0.008
E	0.225	0.225	0.225	0.225	0.225	0.250	0.078
F	0.248	0.248	0.248	0.248	0.248	0.250	0.096
G	0.248	0.248	0.248	0.248	0.248	0.375	0.114
H	0.025	0.025	0.025	0.025	0.025	0.375	0.018
I	0.313	0.313	0.313	0.313	0.313	0.313	0.172
J	0.200	0.250	0.188	0.094	0.167	0.200	0.048

IV. DETERMINISTIC METHOD FOR AIR QUALITY ASSESSMENT

The weighted arithmetic mean function has been used to determine the deterministic Air Quality Index (AQI). The weighted arithmetic mean function is ambiguity free function, shows small eclipsing with large number of variables and is widely used aggregation function (Bardalo et al., 2001 [12]; Kumar and Alappat 2004 [13]). The formula used to determine the aggregated air quality index is given below.

$$AQI = \sum_{i=1}^{n} W_i I_i$$

Where,

I_i is the sub-index of i_{th} pollutant

AQI is air quality index and 'n' is the number of pollutants considered.

W_i is the weightage of the i_{th} pollutant index.

The sub-index of i_{th} pollutant can be determined by

$$I_i = \frac{C_i}{C_s}$$

Where, C_i is the observed concentration of the pollutant

C_s is the concentration limit value of the pollutant as mentioned in National Ambient Air Quality Standards (NAAQS), India.

V. RESULTS AND DISCUSSION

Fig.1 shows the comparative air quality index values derived from deterministic method and fuzzy pattern recognition method. It clearly reveals that the air quality values variation of the samples from the two methods is similar. The air quality values in fuzzy pattern recognition method is reflected by the relative membership degree of the sample; membership degree of 1 represents that the sample is having worst air quality, while the membership degree of 0 is having clean air. Thus the scale of air quality in fuzzy method is 0-1. Similarly, the scale in deterministic method is 0-4; 4 represents worst air quality (maximum air pollution) and 0 represents clean air (minimum or no air pollution). The air quality values are determined using both the method (fuzzy pattern recognition method and deterministic method) for all the artificial samples listed in Table 4. The comparative air quality values for all the samples along with their rank in air quality level (worst air quality to clean air) are also shown in Table 4. The result shows that the ranking in air quality for all the samples are same in both the method. The additional advantage of fuzzy method is that it can accommodate the other subjective parameters like time of exposure and sensitivity of the people in health impact assessment.

TABLE 4 COMPARATIVE AIR QUALITY INDEX AND THE RANKING ORDER FOR VARIOUS CONDITIONS

Conditions	Deterministic AQI	FAQI	Air quality ranking order for different sample	
			Fuzzy pattern recognition method	Deterministic method
A	4.000	1.000	1	1
B	1.000	0.100	4	4
C	0.100	0.001	10	10
D	0.244	0.008	9	9
E	0.916	0.078	6	6
F	0.990	0.096	5	5
G	1.070	0.114	3	3
H	0.324	0.018	8	8
I	1.250	0.172	2	2
J	0.733	0.048	7	7

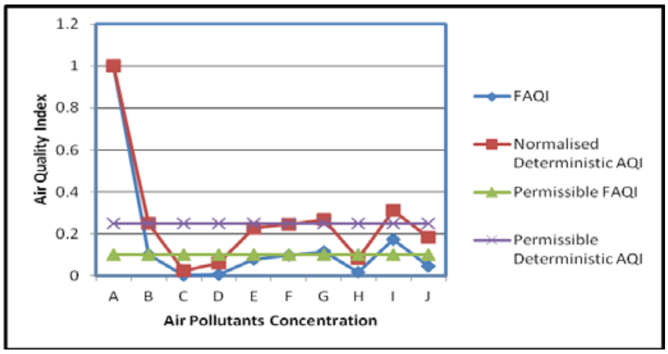

Fig. 1 COMPARATIVE VALUES OF DETERMINISTIC AQI AND FUZZY METHOD AQI

VI. CONCLUSION

The air quality assessment has been demonstrated with artificial dataset considering six major pollutant (PM_{10}, $PM_{2.5}$, SO_2, NO_2, O_3 and CO) parameters. Both the deterministic method and the fuzzy pattern recognition method are applied to a case for artificial data sample set and the results are compared and analysed. The results clearly revealed that the trends are same for both the methods. The additional advantage with the fuzzy pattern recognition model is that it can demonstrate the computing with linguistic terms within fuzzy inference system (FIS) and improves the tolerance for imprecise data. In this study, an approach is developed to rank or prioritize air pollution level in monitoring locations based on the concept of air pollution concentration levels. This will help to identify the importance of air pollution control measurement required in the concerned area. The authors believe that the fuzzy logic concepts, if used logically, could be an effective tool for air pollution control policy issues. More stringent methodologies and reliable results are then required to convince managers and policy makers to apply fuzzy model in practice.

ACKNOWLEDGEMENT

The author thanks to the Vice Chancellor of the Birla Institute of Technology, Mesra for providing the required facility to carry out the work successfully.

REFERENCES

[1] Brimblecombe, P. (1987). The Big Smoke: A History of Air Pollution in London since Medieval Times. New York: Mehuen.

[2] World Health Organization. (2005). Air Quality Guidelines - Global Update. http://www.euro.who.int/__data/assets/pdf_file/0005/78638/E90038.pdf.

[3] USEPA, (1976). USEPA, Federal Register 41 (174), Tuesday, September 7, 1976.

[4] USEPA, (1998). USEPA, Federal Register 63 (236), Wednesday, December 9, 1998.

[5] Malaysia, (1997). "A Guide to Air Pollutant Index in Malaysia", Department of Environment, Kuala Lumpur, Malaysia, 1997.

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

[7] Ontario, (1991). "A Guideline to the Ontario Air Quality Index System", Ontario Ministry of the Environment, Toronto, Ont. Canada ISBN 0-7729-8230-9 Air Resource Branch, 1991.

[8] ORAQI, (1970). "Oak Ridge Air Quality Index", In Environmental Indices Theory and Practice by Wayne R. Ott Ann Arbor Science Publisher Inc. Mich., USA, 1970.

[9] Mandal, T. & Gorai, A. K. & Pathak, G. (2011). "Development of fuzzy air quality index using soft computing approach", Environmental Monitoring and Assessment, DOI 10.1007/s10661-011-2412-0.

[10] Satty, T. L., "The Analytic Hierarchy Process", McGraw Hill International Publication, 1980.

[11] National Ambient Air Quality Standards (NAAQS) (2009). Notification of Central Pollution Control Board (CPCB), New Delhi, India, 18th Nov. 2009.

[12] Bardalo, A. A., Nilsumranchit, W., and Chalermwat, K., (2001). "Water quality and uses of the Bangpakong river, Eastern Thailand", Water Research, 35(15):3635–3642.

[13] Kumar, D., and Alappat, B. J., (2004). "Selection of the appropriate aggregation function for calculating leachate pollution index Practical Period", Hazardous Toxic Radioactive Waste Management 8(4), 253–264.

Dr. Amit Kumar Gorai was born on Jan 1, 1977 in Jharkhand, India. He graduated in Mining Engineering in the year of 2000 from Bengal Engineering College (Presently, Bengal College of Science & Engineering University), Howrah. In 2002, he obtained Master of Engineering from the same university. Soon after his M.E degree, he joined Central Mining Research Institute (CMRI), Dhanbad as a Project Fellow in 2002.

After serving few months at CMRI, he joined Centre of Mining Environment of Indian School of Mines. Dhanbad as a Junior Research Fellow in an R & D Project, sponsored by Department of Mines, Government of India in 2002. The author obtained his doctorate degree in Environmental Science & Engineering from Indian School of Mines University, Dhanbad in the year of 2007. After completion of PhD work, author joined in the Department of Mining Engineering, BIT Sindri as a Part-time Lecturer and continued till Aug. 2007 and thereafter joined in Environmental Science & Engineering Group at BIT Mesra, Ranchi and is still continuing. The author also completed MBA in Operations Management from IGNOU.