International Journal of Soft Computing 3 (6): 467-474, 2008
ISSN: 1816-9503

Requirement Analysis a Fuzzy Logic Process

[1]Lala Mathai, [2]V. Ramchandran and [1]S. Sivanesana
[1]Department of Chemical Engineering, A.C.Tech, Anna University, Chennai 600 025, India
[2]Department of Computer Science, Anna University, Chennai 600 025, India

Abstract: Traditional methods of requirements engineering has not been able to really capture the elasticity and all the information contained in requirements as expressed by clients who are only domain experienced or even experts. The system analyst which is the first interface of requirements capture, traditionally a person recognizes the requirements from his perspective of possible computer aided solutions. From their digital background the decisions normally are to ignore or workaround requirements that seem imprecise or conflicting. A lot of information contained in the informal human communication is lost. In this study, we describes a novel process using Fuzzy Approximation Theorem. First, we discuss the current methodology and why it has not been able to represent the information contained in the inexact, imprecise and conflicting human language of the client or domain expert. We then examine the problems faced in requirement engineering in Pollution forecasting and present a process using Fuzzy logic.

Key words: Pollution mapping, ozone concentration, digital logic, fuzzy approximation theorem, fuzzy associative memory

INTRODUCTION

Understanding others requirements has always been a challenge, the very nature of the human thinking process is such. The goal of requirements engineering is to produce a set of system requirements which, as far as possible, is complete, consistent, relevant and reflects what the customer actually wants (Sommerville and Sawyer, 1997). The "voice of the customers," identified and translated into design targets, is usually vague and conflicting, especially from multiple perspectives (John *et al*., 1994). The inherent properties of modern software systems are complexity, conformity, changeability and invisibility (Frederick, 1987). Most IT systems fail to meet expectations. They don't meet the business goals and don't support users efficiently (Lauesen, 2002). The programs that developers think they've agreed to build often will not be the same products their customers are expecting (Wiegers, 2004).

Neurology has now came to accept that our basic brain function can be similar to the neural networks Neural Networks are a different paradigm for computing. Neural networks are a form of multiprocessor computer system, with simple processing elements. A neural network is a powerful data modeling tool that is able to capture and represent complex input/output relationships. The motivation for the development of neural network technology stemmed from the desire to develop an artificial system that could perform "intelligent" tasks similar to those performed by the human brain. Neural networks resemble the human brain in the following two ways: A neural network acquires knowledge through learning. A neural network's knowledge is stored within inter-neuron connection strengths known as synaptic weights.

A major problem in requirements is obtaining requirements that address the concerns of multiple stakeholders (Boehm *et al*., 1997). Requirements Engineering is the weak link in most projects (Davis, 1990). A good requirements specification should consider cost, quality feasibility, speed of completion, risk profiles, data availability, robustness, reliability, simplicity, manpower availability, maintainability among others. Digital Logic works on the basis of known, discovered and calculated formulae. So, all new thinking has to be clouded by previous knowledge. Using Fuzzy Logic, which works on real world data acquired by sensors, actual interactions and rules formulated by involuntary behavior of the various experts in the project. In this study, we describes a novel process using Fuzzy Approximation Theorem. We discuss the current methodology and why it has not been able to represent

Corresponding Author: S. Sivanesana, Department of Chemical Engineering, A.C.Tech, Anna University, Chennai 600 025, India

the information contained in the inexact, imprecise and conflicting human language of the client or domain expert. We then examine the problems faced in requirement engineering in Pollution forecasting and present a process using Fuzzy logic.

FUZZY SETS

Words stand for sets we think in sets Thought is set play. The same word can stand for different members of the set. This does not stop with nouns. Classifying with adjectives just increases vagueness. Meaning representation may be regarded as a process which, identifies the variables which are constrained and characterizes the constraints to which they are subjected (Zadeh, 1986). Our language itself makes our communication fuzzy. It is rare that another thinks of the same member of the set when we speak. We structure words down to symbols to get more precision, but this causes even more confusion. Our brains group things into loose sets and then play with the groups. These loose or fuzzy sets are expressive (Zimmermann, 1991). The expressive power of the fuzzy set fits better with our words. As the complexity of a system increases our ability to make precise statements about its behavior diminishes. When faced with a complex decision the natural human reaction is to cluster the decision elements according to, their common characteristics (Saaty, 1980). Language is sloppy and so a system to structure it is needed linguistic inventions have to be followed. They have envolved to closely represent the real world and have become communication standards. Words stand for sets of things. Sentences have many words. Sentences relate these words. In this way we reason-common sense But what we sense is really not common to all of us. The same signals of sight, sound, feel, hit us, but to each is a unique perception. A group of sentences give a system -a fuzzy system. A system is anything that maps inputs to outputs. If knowledge can be taken from this fuzzy system to model the real world. Then we can have computer systems that adjust to humans rather than reverse. Fuzzy systems allow us to model systems in words. we derive approximation. However, binary logic runs our computers, it is simple, works and has served us for long. Present computers process information on binary logic which is very different to the human thinking process. Hence, evaluation based on common sense and flexible judgment is very difficult to achieve in the computer process. Uncertainty contained in the meaning of each word makes if difficult to have an intimate relationship between man and computer. In this study, is an attempt to use the knowledge contained in human representations using Fuzzy logic.

AIR QUALITY-SURFACE OZONE

This approach is used in an air quality monitoring study at Chennai city done in AC Tech, Anna University. by Pulikesi (2005) (Table 1). This is a study of concentrations of various pollutants mainly surface ozone (O_3) and also others such as oxides of nitrogen (NOx), respirable suspended particulate matter (RSPM) and total suspended particulate matter (TSPM), under various climatic conditions viz. Temperature, time of day, relative humidity, wind speed, wind direction etc at different relevant geographical locations near the sea coast, low vegetation, far inland etc with different usage patterns viz. industrial, residential, dense human population, high vehicle traffic etc.

Any representative pollution mapping and useful projection from this study to judge total air quality will need a complicated mathematical formulae involving many variables as above. This can of course be handled by a computer program processed in a lab PC. But we are all aware that such forecasting attempts involving climate factors have met with limited success since there are too many other variables which are relevant albeit in a much smaller way but still sometimes very effective.

Further to establish the relationship to this pollution mapping to Public Health needs, they have to be related to the National Ambient Air Quality Standards (NAAQS) which specifies the quality necessary with an adequate margin of safety for the

Table 1: Details of Air quality monitoring stations in Chennai, India

Site	Code	Latitude (N)	Longitude (E)	Site classification	Remarks
Kodungaiur	S_1	13.136851	80.2538	Industrial area	High Industrial emissions/ effluents
Koyambedu	S_2	13.07224	80.20174	Commercial area	Service Industries/ Interior
Mandaveli	S_3	13.02596	80.26618	Commercial Area	Service Industries/ Coastal
Taramani	S_4	12.98112	80.23949	Residential area	Dense population
Vallalar nagar	S_5	13.10571	80.28016	Traffic island	High motor vehicle emission

protection of public health, sometimes for sensitive classes of the population, vegetation and property.

In this study, endeavors to exhibit how by using Fuzzy Logic and it can be implemented in a simple process. This also proved that why this will be a more reliable process, since even after a large number of variables are accommodated in a mathematical formulae, we can never have understood the interactions between these and also other unknown variables.

FAT THEOREM-FUZZY APPROXIMATION THEOREM WITH FUZZY ASSOCIATIVE MEMORY

Pulikesi (2005) investigated during the summer, concentrations of air pollutions different conditions of Relative Humidity (RH), Wind Speed (WS) and Wind Direction (WD) were collected over successive periods of about 24 h at 5 sites. which were different in geographic characteristics and usage patterns. The readings as tables and graphical representation are given in Annexure at the end. The National Ambient Air Quality standards (NAAQS) has specified quality necessary with an adequate margin of safety for the protection of the public health, vegetation and property. The NAAQS Standards for the air pollutants i.e. nitrogen oxides (NO_x), Respirrable Suspended Particulate Matter (RSPM) and Total Suspended Particular Matter (TSPM) are summarized in Table 2 and 3.

In Fig. 1 and 2 (Annexure), hourly O_3 concentrations are shown as a function of relative humidity. The other meteorological data i.e. wind speed and wind direction was not available on that day. In the early morning hours, O_3 concentration showed a sharp increase in ambient air and reached its first peak in 11.00-

Table 2: National Ambient Air Quality Standards (NAAQS) Air pollutants

		Concentration in Ambient Air (ppb)		
Pollutants	Averaging time	Industrial area	Residential, rural and other areas	Sensitive area
Oxides of Nitrogen NO_2	Annual Mean	40.8	30.6	7.65
	24 h	61.2	40.8	15.3
Respirrable Suspended Particulate Mater (RSPM)	Annual Mean	183.6	71.4	35.7
	24 h	255	102	51
Total Suspended Particulate Matter (TSPM)	Annual Mean	61.2	30.6	25.5
	24 h	76.5	51	38.25

Table 3: Standards for the air pollutants i.e., nitrogen oxides (NO_x),s Respirrable Suspended Particulate Matter (RSPM) and Total Suspended Particular Matter (TSPM)

	Pollutant (ppb)			
Time	O_3	NO_x	RSPM	TSPM
06-14	37	18.3	35.19	139.74
14-22	22	9.02	28.56	52.02
22-06	18	3.21	18.87	33.15
24 h. Avg.	25	10.2	27.54	74.97
Koyambedu (S_2)				
06-14	28	12.03	20.4	123.42
14-22	21	9.84	31.62	117.81
22-06	8	9.33	11.73	28.05
24 Hrs. Avg.	18	10.4	21.42	89.76
Mandaveli (S_3)				
06-14	38	12.64	20.4	107.61
14-22	35	10.25	17.34	59.93
22-06	22	9.58	24.48	61.2
24 h. Avg.	31	10.81	20.91	76.5
Taramani (S_4)				
06-14	36	4.69	29.58	110.67
14-22	40	4.43	28.56	90.27
22-06	18	2.6	24.99	55.59
24 h. Avg.	30	3.92	27.54	85.17
Vallalar nagar (S_5)				
06-14	10	27	87.72	212.16
14-22	16	15.81	24.99	117.81
22-06	4	8.16	33.15	76.5
24 h. Avg.	10	16.93	48.45	135.66

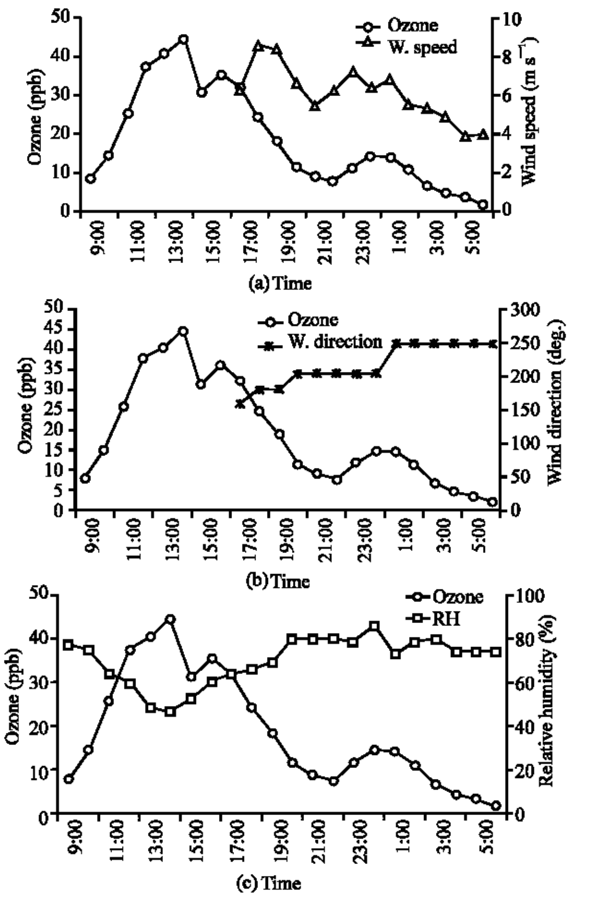

Fig. 1: Hourly variations of ozone and meteorological parameters obtained from 25-26 May 2005 at Koyambedu. (a) Ozone and wind speed, (b) ozone and wind direction and (c) ozone and relative humidity dity

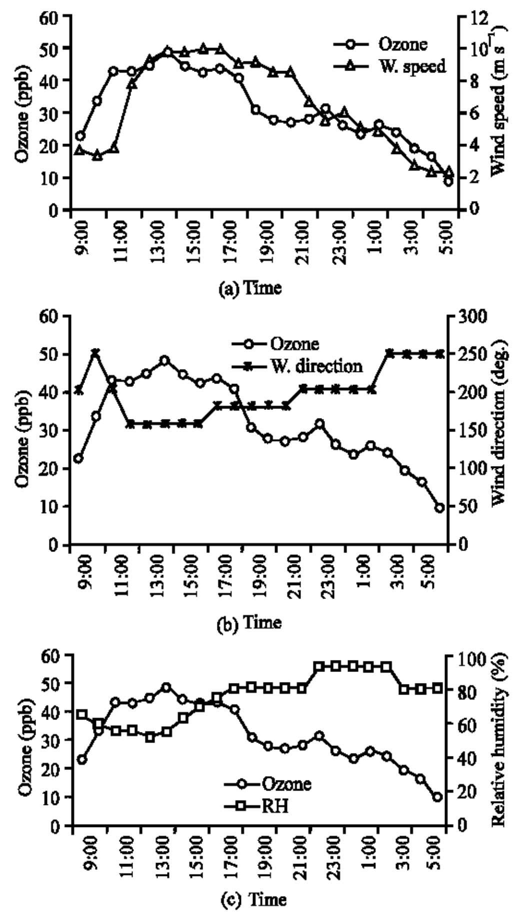

Fig. 2: Hourly variations of ozone and meteorological parameters obtained from 31-1 May June 2005 at Mandaveli. (a) Ozone and wind speed, (b) ozone and wind direction and (c) ozone and relative humidity

14.00 as also recorded by National Research Council (NRC). The first peak occurred at temperature of 42°C and relative humidity of 47% (Fig. 2). Subsequently decrease of O_3 concentration may be due to the cloudy weather conditions. As mentioned by NRC a second peak was also observed in the 16.00-20.00. After that, a gradual decrease of O_3 concentration continued until the early morning of the next day.

In Fig. 3 (Annexure), hourly variations of O_3 concentrations are shown as a function of meteorological parameters obtained i.e. Ozone and wind speed, Ozone and wind direction, Ozone and relative humidity during the period 25-26 May 2005.

In Fig. 4 (Annexure), hourly variation of ozone concentrations are shown as a function of meteorological parameters obtained i.e., Ozone and wind speed, ozone and wind direction, Ozone and relative humidity.

In a digital system we can relate the changes in concentrations of each pollutant to each condition separately or we can derive mathematical formulae to try and relate the pollutant concentrations to changes in all these conditions at each time. These formulae will be very complicated and for easier handling we may have to approximate to lower degrees. When we work in the reactions between each of these pollutants at different concentrations the formulae become even more complex. In the fuzzy system, we first establish rules or trends in the variation of each of these pollutants either from experiences of experts in the field or by an analysis of our readings as also other similar work.

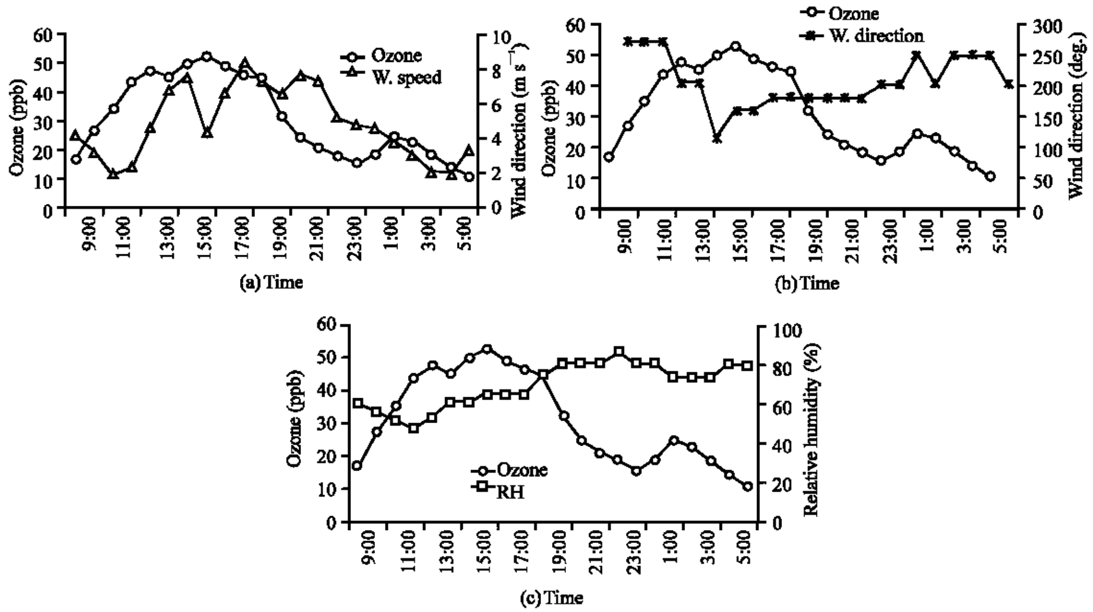

Fig. 3: Hourly variations of ozone and meteorological parameters obtained from 03-04 June 2005 at Taramani. (a) Ozone and wind speed, (b) ozone and wind direction and (c) ozone and relative humidity

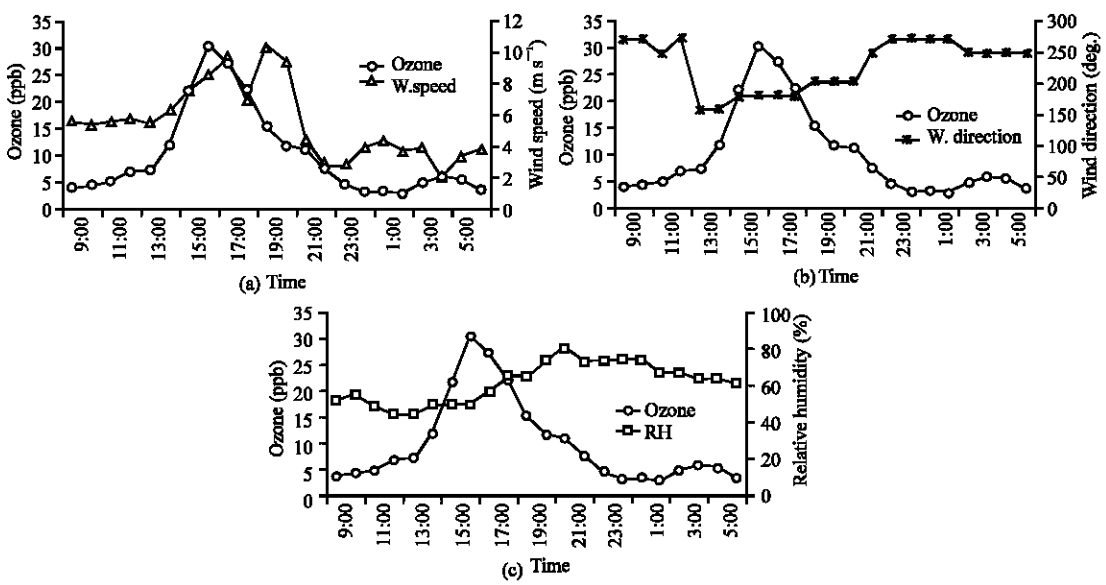

Fig. 4: Hourly variations of ozone and meteorological parameters obtained from 04-05 July 2005 at Vallalar Nagar. (a) Ozone and wind speed, (b) ozone and wind direction and (c) ozone and relative humidity

POLLUTION MAPPING AND FORECASTING MODEL

This would be a program to give results based on which decisions can be made for Locating population centers, industrial areas either for polluting or non-polluting processes, outdoor entertainment sports sites etc. The result can also be on a weighted average of the acceptable limits allowed by the National Ambient Air Quality standards (NAAQS) of all the pollutants studied. The inputs into the program would be the various readings taken as done in the study. More the readings taken over longer periods would give more reliable results. The data for each pollutant under different conditions are merged by a process known as Fuzzy Associative Memory which is explained below in detail.

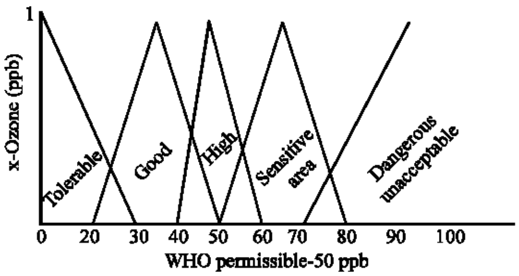

Fig. 5: The foll statements WHO standard for a high O_3 concentration is 50 ppb

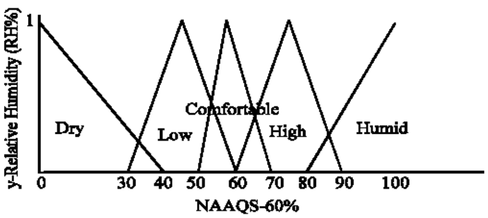

Fig. 6: The relative humidity regions of comfort, humid, dry etc with the NAAQS being 60% for maximum comfort

The resultant data for each pollutant is again fed into another module to compare with the allowable standards of NAAQS where we use another FAM to arrive at an indication of Total Air Quality. In this analysis example at first we shall take only one pollutant ie Ozone (O_3) and track its variations against only one of the climatic variables i.e., relative Humidity (RH). We now pick the variables say x-O_3 and y-RH.

We use a geometric representation rather than mathematical symbols as this more easily striking to a larger no of people in the climate forecasting and pollution study fields.

Figure 5 below is a representation of the foll statements WHO standard for a high O_3 concentration is 50 ppb. ie from 40-60 can be considered in the high region with 50 ppb being definitely only in the high region and 40 in the high as well as in the good region. Similarly, the other areas signify their membership in a scale in [0,1].

This representation captures the real fact that each reading can be " somewhat " in other regions also Similarly, Fig. 6 depicts the Relative Humidity regions of comfort, humid, dry etc with the NAAQS being 60% for maximum comfort.

Figure 7 depicts the region where the Rule " if RH is high then O_3 is good " is satisfied. After picking the variables x and y, we now have to pick the fuzzy sets for the variables x and y ie here for O_3 and RH. The fuzzy sets we can choose for O_3 are tolerable, good, high sensitive area and unacceptable/dangerous.

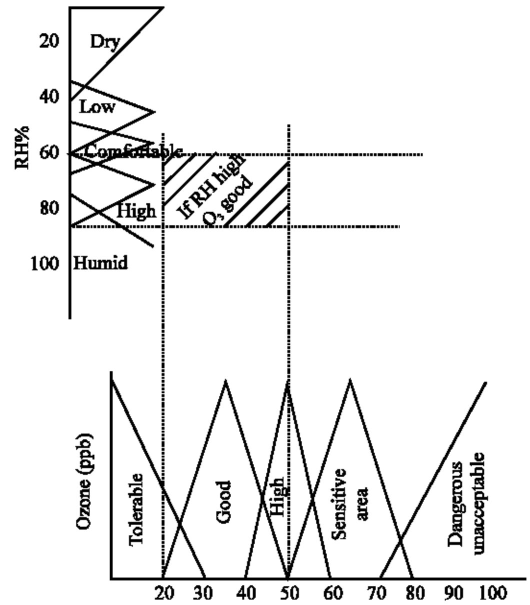

Fig. 7: The region where the rule and y ie here for O_3 and RH

The fuzzy sets we can choose for RH are dry, low, comfortable, high and humid. After choosing the appropriate fuzzy sets we have to find the rules. These are found from trends in readings expert opinions and also adaptive learning systems where available.

When we study the readings taken as shown in the Annexure and discussions, we can assume many trends. Some are given for example:

- When RH is high O_3 is low- mostly inversely proportional provided all other factors are same.
- When temperature is high NOx emissions from soil are high.
- Wind direction from the sea brings in O_3 and hence concentration increases.

These trends are also obvious from a casual study of the graphical representation of the comparative readings (Annexure). From such trends and discussions with domain experts we formulate rules. As an example we can work with the follow five rules. If RH is dry then O_3 is dangerously high and unacceptable.

- If RH is low then O_3 is higher than max and unacceptable to sensitive sections like aged and sick.
- If RH is comfortable then O_3 is high around WHO max.
- If RH is high then O_3 is in the good region.
- If RH is humid then O_3 is low and not healthy, but tolerable.

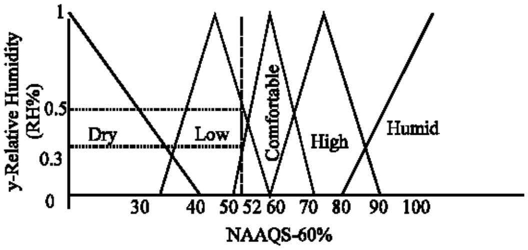

Fig. 8: Comparision of comfortable region and low region

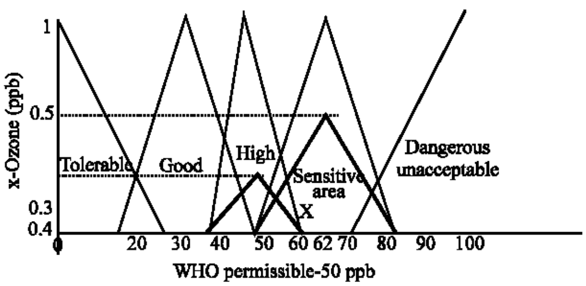

Fig. 9: The effective weight in the corresponding region of O_3

A graphical representation of the above rules can be seen in Fig. 7. Rule no 4, if RH is high then O_3 is in the good region applies to all the points lying within the bounded shaded area. Such rules can be formulated independently for pairs of all other variables which have a bearing on O_3 concentration such as Wind direction to O_3, Nox to O_3, temperature to O_3, seasons-winter, summer etc to O_3 etc.

We see from Fig. 7 that a reading of RH say 52 can lie in both the low and comfortable regions each to some extant. Hence, for a RH reading of 52 both Rule 2 and Rule 3 will fire simultaneously each to a certain extant. We can accommodate this in fuzzy logic but not in digital logic where it is EITHER-OR. To get a representative possible value of O_3 we use the FAM. This is explained below and represented geometrically in Fig. 8 and 9.

From Fig. 8 we notice that when RH is 52 it is 30% in the comfortable region and 50% in the low region; ie Rule 2 kicks in for 50% effect and Rule 3 for 30% effect. This is shown in Fig. 9 where the effective weight in the corresponding regions of O_3 are shown., ie there is a 30% effect the high O_3 region and a 50% effect in the sensitive area. This can also be called the Degree of Belief in a conclusion. We can use weighted Inference Equations (basic Neural Equations) to obtain the defuzzified result.

However, our geometric representation is more intuitive and hence used here. In the geometric representation this is achieved by shrinking down the corresponding triangles by the respective values.

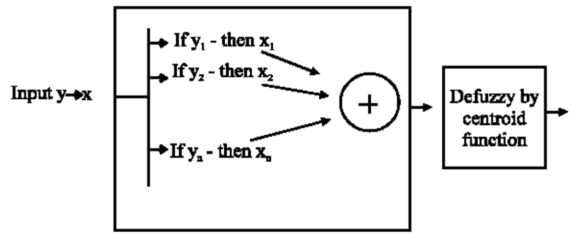

Fig. 10: Fuzzy associative memory

The value of O_3 is the resultant fuzzy set under the dark line triangles. i.e., O_3 value can fall anywhere in the area. However to get the most probable representative value we need to defuzzify this result.. We use the foll: method.

It has been found that taking the centroid of the area under the dark line gives the best defuzzified value. The centroid at point "x" is at the value 62 ppb. This is the defuzzification method found to be the most relevant and accurate in practice.

Similarly, we can take a fuzzy weighted average for each of the climatic conditions which affect the ozone concentration.

All these weighted averages can again be fed into a FAM module as shown in Fig. 10, to get the final defuzzified figure which will be the most likely O_3 concentration for the values of variables fed in.

CONCLUSION

All Data are taken from actual real life studies in the particular area concerned for a long period of time, even the other factors in nature which affect the pollutant; known and unknown parameters are also effected in the readings. The unknown parameters will cause a skew in the result, but since each of these natural phenomenon are also interrelated, much of the effect of such unknown parameters would be reflected in readings of the known parameters. The skew will anyhow be much less than any mathematical model which has to ignore unknown parameters completely. Using Fuzzy Cognitive Mapping, the interrelationships between parameters can also be more directly effected.

In the normal digital methods, we have to devise a mathematical model and get to a formula relating the input to output. This further to being complicated, we need a very high order of mathematical resource as also very large computing resources. (most of the supercomputers manufactured are working on such climate forecasting programs). This mapping and forecast can be further refined by an adaptive neural net after many seasons of data to increase the reliability of the forecast.

REFERENCES

Boehm, B., *et al.*, 1997. Software Requirements Negotiation and Renegotiation Aids. Proc. 17th Int. Conf. Software Eng., pp: 243-253.

Davis, A., 1990. Software Requirements: Analysis and Specification, Prentice-Hall, Englewood Cliffs, NJ.

Frederick, P. Brooks, 1987. No silver bullet: Essence and accidents of software engineering. Computer, 15 (1): 10-18.

John Yen, Xiaoqing Frank Liu and Swee Hor Teh, 1994. A fuzzy logic-based methodology for the acquisition and analysis of imprecise requirements. Int. J. Concurrent Eng. Res. Appl. (CERA), 2 (4): 265-277.

Lauesen, S., 2002. Software Requirements-Styles and Techniques, Addison Wesley.

Saaty, 1980. ThenbAnalytic Hierarchy Process. McGraw-Hill.

Pulikesi, M., 2005. Air quality monitoring in Chennai. In the summer of 2005. J. Hazardous Materials.

Sommerville, I. and P. Sawyer, 1997. Requirements Engineering-A Good Practice Guide. John Wiley and Sons Ltd.

Wiegers, K., 2004. Software Requirements. International Symposium on Software Engineering.

Zadeh, L.A., 1986. Test-score semantics as a basis for a computational approach to the representation of meaning. Literacy Linguistic Comput., 1 (5): 24-35.

Zimmermann, H.J., 1991. Fuzzy Set Theory and Its Applications. Kluwer Academic, Boston, MA.