A NEW INDEX FOR AIR QUALITY MONITORING USING FUZZY LOGIC

Graziela Marchi Tiago, grazielamarchi@gmail.com
Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – *Campus* São Paulo
Rua Pedro Vicente, 625 – 09011-010 – São Paulo/SP

Elaine Inacio Bueno, elainebueno@usp.br
Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – *Campus* Guarulhos
Avenida Salgado Filho, 3501 – 07115-000 – Guarulhos/SP

Paulo Roberto Barbosa, barbosapr@gmail.com
Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – *Campus* São Paulo
Rua Pedro Vicente, 625 – 09011-010 – São Paulo/SP

Abstract. *The air quality monitoring occurs due to the concern of determining the concentration levels of some pollutants which was known as indicator. The indicator choice was made because they can be found in the atmosphere and they provoke harmful effects to the environment and to the human health. The indicators are: sulphur dioxide (SO_2), carbon monoxide (CO), nitrogen dioxide (NO_2), particulate matter (MP_{10}) and ozone (O_3). There are maximum concentration limits to these indicators with the purpose to assure the health and population well-being. In Brazil, these limits had been established by the IBAMA (Environment Brazilian Institute) and approved by the CONAMA (Environment National Council) through CONAMA resolution 03/90. These indices are always monitored by using a mathematical tool whose name is Air Index Quality which converts the pollutant concentrations into scales: good, regular, inadequate, bad and very bad. The objective of this work is to develop a new system to monitor the air quality of Sao Paulo state in Brazil using Fuzzy Logic. Results show similar values for the index when compared with the classical calculation and agreement for the scales (classification) in the most of simulated cases. The differences observed are justified because of the highest sensitivity in this model which considers all the indicators in its calculations.*

Keywords*: Air quality, fuzzy logic, pollutants.*

1. INTRODUCTION

In the last few years, it has been noticed an increasing concerning with the possible harmful effects to the heath because of the pollution air exposition. These harmful effects have been observed after the first half in the last century during episodes of higher pollutant concentration which was observed in the Meuse Valley, in Belgium (Firket 1931); in Donora in the Pennsylvania (Ciocco and Thompson, 1961); and in London in the winter of 1952-1953 (Ministry of Health 1954).

These and other isolated episodes have influenced some measures to control the air pollution ambient levels in many urban centers (Gouveia, et al. 2003). The atmospheric pollution control is extremely important because some studies relate the arising of some diseases, such as: cardiovascular illnesses (Ballester, et al. 1996), respiratory illnesses (Borja, et al. 1997), and the mortality index increase, due the presence of pollutants in the atmosphere (Anderson, et al. 1996).

Due to these facts it is necessary monitoring the air quality in order to make possible to take adequate action when the pollutants level in the atmosphere exceed the maximum limits allowed, which could place in risk the population health. The Ibirapuera Park which is located in Sao Paulo city, for example, considering the stations list measured from CETESB, received the highest pollution classification in the year of 2010, where it was observed the ozone level above the recommended limit for 49 times. This pollutant is known secondary and it was formed from the reactions between nitrogen oxides and volatile organic composites, in the presence of solar light. Because of its formation process and depending on the meteorological conditions, it is more difficult to control the ozone levels. Pretty days with a hot sun and clearly sky are the worse ones to ozone.

The air quality monitoring is carried through the concentration measurement of some pollutants as: total suspended particulate matter, smoke, particulate matter, sulphur dioxide, carbon monoxide, nitrogen ozone and dioxide (CONAMA, 1990), following quality air standards which legally define the maximum limits for the concentration of these pollutants in the atmosphere.

The IBAMA (Environment Brazilian Institute) and the CONAMA (Environment National Advice) established national standards for the air quality through CONAMA Resolution 03/90 (CONAMA, 1990) (CETESB, 2011). This resolution establishes two standards to air quality: the primary and the secondary. The primary air quality standards are the maximum pollutants concentration levels, when it is exceeded, they will be able to affect the population health and constitute a short and medium time goal. The secondary air quality standards are the desired level of pollutants concentration and constitute a long time goal and it had been created with the purpose to establish the prevention politics of air quality degradation which should be applied to the preservation areas.

The Air Quality Index is used to publish the air quality, and it is a mathematical tool developed to simplify the publishing of air conditions. The parameters considered by this index structure from the CETESB are: sulphur dioxide, particulate matter, carbon monoxide, ozone, nitrogen oxide. For each pollutant measured is calculated an index and an appropriate qualification is obtained, which can be: good, regular, inadequate, bad and very bad. It is used the highest air quality index to publish to the population, it means, it is considered the worst air quality index of a determined station.

As an alternative method to the mathematical tool for the pollutants measured index calculation, it can be used the Fuzzy Logic to the Air Quality Monitoring, as it can be seen in (Khan and Sadiq, 2005) (Silvert, 2000) (Engin, Demir and Hiz, 2004). The objective of this work is the development of a new monitoring system to the air quality in Sao Paulo State using Fuzzy Logic. This system will use all the indicators concentration to classify the air quality and will be used to support in decisions taking during situations which take the population health at risk. The great differential of this developed system is the possibility to classify the air condition without using a mathematical tool to take it, and the use of all involved pollutants to get the index and to classify the air quality. This work will collaborate with pollution control programs, for example, urban toll and the car bands reduction due to the new cycle bands, because it is necessary to have some programs and not only depend on meteorology.

2. FUZZY INFERENCE SYSTEMS

Fuzzy logic was originally proposed by Lofti Zadeh in 1965 with the work “Fuzzy sets” (Zadeh, 1965) and then developed as a tool for manipulating and processing vague information in uncertain conditions. One of the main characteristics of this approach is the element partial membership which allows smooth transitions from one rule to another (Yager and Filev, 1994). In this context, the production of the membership functions, i.e., functions that define the membership degrees for each input and output of the system is called “fuzzyfication”. All fuzzy set representing the crisp (physics) variables related by membership functions are the so called “knowledge basis”.

The knowledge basis has uncertain information however significant for the system modeling. Although this uncertain is completely solved as the input and output fuzzy sets and the knowledge manipulation strategy are defined.

A fuzzy algorithm processes the membership functions for each one of the fuzzy sets and the results are aggregating through instructions or rules, producing the so called “rule basis”.

There are basically two types of fuzzy system models differentiating in the ability of representing different kinds of information, i.e, in the form of representing the rule basis. The first include the linguistic models based in collections of IF-THEN rules with vague attributes and have fuzzy reasoning. In this type of model, fuzzy quantities are associating with linguistic labels and a fuzzy model is essentially a qualitative expression of the system. The second type of model is based in the Takagi-Sugeno-Kang reasoning method (Sugeno, 1985). These models are constructed by logic rules which are combination of fuzzy and crisp models (Yager and Filev, 1994).

A set of inference rules is adopted to manipulate the knowledge basis. The most used method to represent the human knowledge is through natural language expressions as: IF (antecedent) THEN (consequent).

Since decisions are based on the testing of all of the rules in the inference system, the rules must be combined in some manner in order to make a decision. Aggregation is the process by which the fuzzy sets that represent the outputs of each rule are combined into a single fuzzy set.

One of the most used is the Mamdani implication method for inference in which the aggregated output is:

$$\mu_{B_n^k}(\alpha(i),\alpha(j)) = \max[\min[\mu_{A_{n1}^k}(\alpha(i)),\mu_{A_{n2}^k}(\alpha(j))]], \qquad \text{for } k = 1,...,r \tag{1}$$

where A_{n1}^k and A_{n2}^k represent antecedent fuzzy sets, μ represent membership functions, B_n^k represent the consequent fuzzy set for inputs $\alpha(i)$ and $\alpha(j)$.

Often fuzzy process output must be a scalar quantity and not fuzzy sets. A crisp value for the system output is obtained by the defuzzyfication of the fuzzy output set. In the literature there are some defuzzyfication methods as, for instance, centroid, bisector, middle of maximum (the average of the maximum value of the output set), largest of maximum, and smallest of maximum. Perhaps the most popular defuzzification method is the centroid calculation, which returns the center of area under the curve. In this method the crisp value is obtained by the center of area given by the gather of the output membership functions as

$$y^* = \frac{\int \mu_{B_n^k}(y)y\,dy}{\int \mu_{B_n^k}(y)\,dy} \tag{2}$$

being y^* the value obtained by the defuzzyfication and e B_n^k consequent fuzzy sets.

3. FUZZY MONITORING SYSTEM

This section describes the fuzzy model developed at Labview® software to monitoring the air quality. The experimental data was obtained from CETESB website and they refer to cities in São Paulo state – Brazil. The choose of days was made randomly. Five variables were defined as system inputs, they are: sulphur dioxide, carbon monoxide, nitrogen dioxide, particulate matter and ozone. Each one of the system inputs were constructed with 5 trapezoidal membership functions called: Good, Regular, Inadequate, Bad and Very Bad. Examples of these input fuzzy set are shown in following figures. (Fig. 1 and Fig. 2).

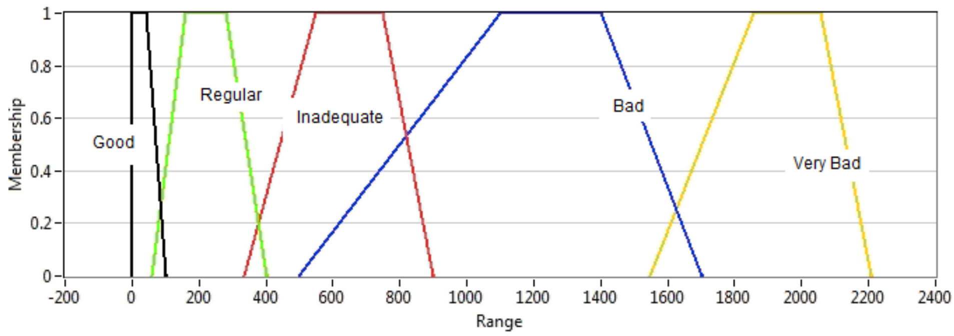

Figure 1. Membership functions for the input variable SO_2

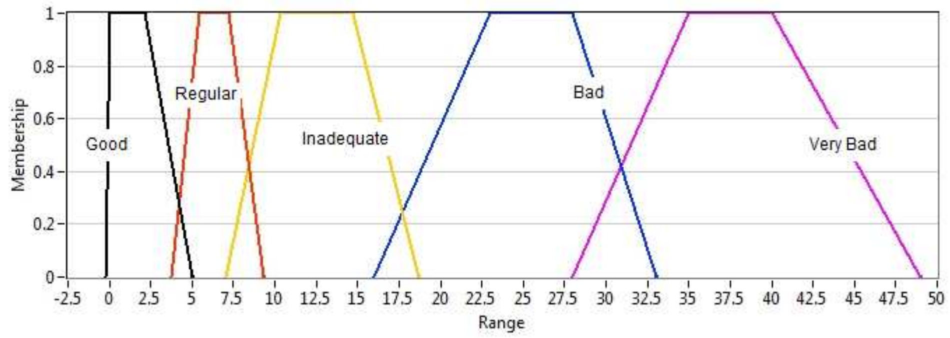

Figure 2. Membership functions for the input variable CO

The model has only one output variable, called Air Quality (Fig. 3). This variable was also composed of five trapezoidal membership functions with the same labels as the input variables (see Fig. 3).

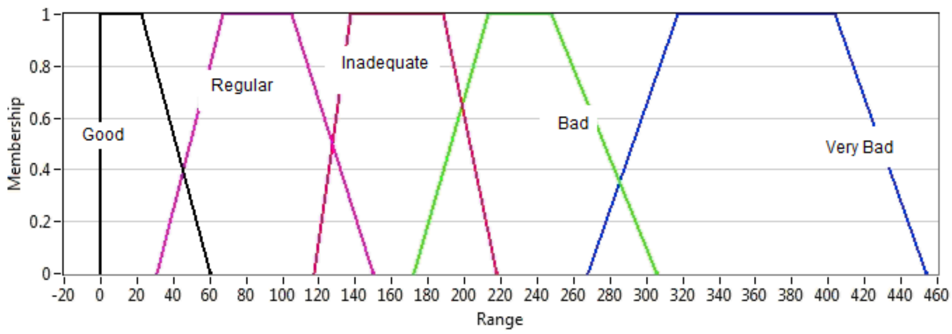

Figure 3. Membership functions for the output variable ‘air quality’

The rule base was composed by only 5 rules. All rules were design with the same weight and from simple decisions, as for example, if all input variables were considered good then the air quality must be classified as good.

The Mamdani type model developed used the centroid calculation as defuzzification method (Section 2).

The front panel of the developed software is shown in following figure (Fig. 4) and the influence of pollutants MP_{10} and O_3 can be observed in Fig. 5 through the input-output mapping surface.

Figure 4. Front panel of the developed software

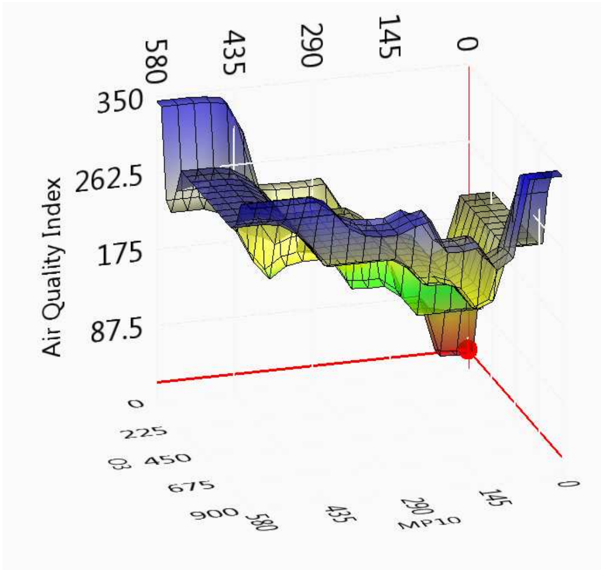

Figure 5. Input-output mapping considering O_3 and MP_{10}

4. RESULTS

Around 120 tests were performed from data obtained in CETESB website, corresponding to period from 2003 to 2011. Some of these tests are presented in the next table (Tab. 2) where they are compared with the index and classification obtained with the classical calculation.

The air quality index has been used by CETESB since 1981. In the actual calculation, called classical in this work, for each one of the measured pollutants one index is calculated and correlated with the air quality.

To purpose of divulgation, the highest index in each station is used. In other words, the air quality in a place, by the traditional method is defined by the worst case. More details can be found in Kiely (1996).

Considering the same classification used in the classical model, shown in table 1, we can verify that the Fuzzy model produced the same results in almost every realized test. The small differences that were found can be justified because the classical method calculates the index for each pollutant separately e uses the worst one. By the other side the fuzzy model considers all pollutant for calculus and consequently for classification. By this way, the fuzzy model obtains a more realistic and complete classification once it observes the contribution of all pollutants.

Table 1. Air quality classification.

Index	[0, 50]	(50, 100]	(100, 199]	(199, 299]	(299, ∞)
Classification	Good (G)	Regular (R)	Inadequate (I)	Bad (B)	Very Bad (VB)

Table 2. Results for air quality index in classical and fuzzy systems.

Locality	SO_2 $(\mu g/m^3)$	MP_{10} $(\mu g/m^3)$	NO_2 $(\mu g/m^3)$	CO (ppm)	O_3 $(\mu g/m^3)$	Classical index	Classical Classification	Fuzzy index	Fuzzy Classification
Congonhas	5	30	77	1.9	0	38	Good	22.76	Good
Parelheiros	0	38	25	2.6	73	45	Good	24.61	Good
Osasco	5	40	65	2	0	40	Good	22.12	Good
São Caetano do Sul	3	17	45	0.5	137	86	Regular	88.67	Regular
Paulínia	4	19	34	0	132	83	Regular	88.67	Regular
Parque D. Pedro II	31	138	176	5.4	60	94	Regular	92.94	Regular
Diadema	0	23	0	0	116	73	Regular	87.68	Regular
Cubatão Vila Parisi	5	91	57	0	0	71	Regular	89.37	Regular
São Caetano do Sul	20	88	150	9.4	74	107	Inadequate	123.90	Inadequate
Ibirapuera	0	56	208	1.7	188	170	Inadequate	142.17	Inadequate
Paulínia	3	24	30	0	183	158	Inadequate	172.41	Inadequate
São José dos Campos	6	16	0	0	172	130	Inadequate	165.51	Inadequate
Cubatão Centro	8	39	53	0	184	160	Inadequate	175.47	Inadequate
Parque D. Pedro II	31	161	214	7.8	3	111	Inadequate	118.82	Inadequate
Mauá	0	46	90	0	208	201	Bad	227.18	Bad
Santana	0	78	0	0	221	204	Bad	216.01	Bad
Pinheiros	0	47	102	0	262	210	Bad	206.66	Bad
Ipen USP	0	0	78	1	235	206	Bad	235.54	Bad
Ibirapuera	0	40	0	1.3	272	212	Bad	235.54	Bad
São Caetano do Sul	6	81	96	1.3	213	205	Bad	203.01	Bad
Ibirapuera	0	17	37	0.5	203	201	Bad	217.99	Bad
São José dos Campos	5	29	0	0	159	99	Regular	106.48	Inadequate
Mauá	0	48	35	0	198	195	Inadequate	205.62	Bad
Taboão da Serra	0	54	74	1.8	0	53	Regular	23.28	Good

In some situations as, for example, in the three last lines of Tab. 2, the fuzzy model classification was different of the classical model. At São José dos Campos and Mauá cases the fuzzy index was bigger than the classical due to the contribution of other variables as MP_{10}. At Taboão da Serra case the fuzzy index was lower than the classical. It was classified as Good because the pollutant MP_{10} is at the beginning of regular set and the others pollutant associated as Good contributed to place the index in a Good classification.

Once it was not found in the CETESB website measurement of Very Bad classification, some data were simulated for this specific situation. In this case we made the calculus with the mathematical expressions of the classical method and we compared with the results obtained with the fuzzy model. In all these simulations the fuzzy and the classical classifications were exactly the same.

5. CONCLUSIONS

A Fuzzy model was developed in this work using the Labview software aiming the air quality monitoring. The model uses as input the measured values obtained from CETESB. These measured values are: sulphur dioxide, particulate matter, carbon monoxide, ozone, nitrogen oxide. The main objective was to use all the cited pollutants to estimate a new index to the air quality and by using the classification adopted by IBAMA and CONAMA to compare the obtained results with the classical model (mathematical calculation).

The fuzzy system developed has some advantages when compared with the classical model, which are: it uses only 5 rules to esteem the index, easy implementation and interpretation to the user and the most important is that it considers the contribution of all the pollutants to generate the air quality index. This important contribution did not occur with classical models because they only use one pollutant (the worst value) to the index calculation.

Hence, the proposal methodology suggests a new index calculation which was able to reproduce the classical system classification in most simulated situations and has a great flexibility once it is possible to give different weights to one determined rule, for example. These elements amplify the model sensitivity significantly to air quality classification and make the proposed methodology very adaptable to any new standards.

6. REFERENCES

Anderson, H. R., Ponce, L. A., Bland, J. M. and Bower, J. S., 1996. “Air pollution and daily mortality in London: 1987-92.” British Medical Journal, pp. 665-669.

Ballester, F., Corella, D., Perez, H. S. and Hervas, A., 1996. “Air pollution and mortality in Valencia, Spain: a study using the APHEA methodology.” Journal of Epidemiology and Community Health, pp. 527-533.

Blanke, T., Hammel, U. and Schwefel, H. P., 1997. “Evolutionary computation: Comments on the history and current state.” IEEE Transactions on Evolutionary Computation, pp. 3-17.

Borja, A. H., Loomis, D. P., Bangdiwala, S. I. and Shy, C. M., 1997. “Ozone, suspended particulates, and daily mortality in Mexico City.” American Journal of Epidemiology, pp. 258-268.

CETESB, 2011. “Companhia Ambiental do Estado de São Paulo”. 3 Jan. 2011, <http://www.cetesb.sp.gov.br/>.

Ciocco, A. and Thompson, D. J., 1961. “A follow-up on Donora ten years after: methodology and findings.” American Journal of Public Health, pp. 155-164.

CONAMA - Conselho Nacional do Meio Ambiente, 1990. “RESOLUÇÃO/CONAMA/N.º 003/90.” 3 Jan. 2011, <http://www.mma.gov.br/port/conama/res/res90/res0390.html>.

Engin, G. O., Demir, I. and Hiz, H., 2004. “Assessment of urban air qualityin Istanbul using fuzzy synthetic evaluation.” Atmospheric Environment, pp. 3809-3815.

Firket, J., 1931. “The cause of the symptoms found in Meusa Valley during the fog of December 1930.” Bulletin of the Academy of Royal Medicine of Belgium, pp. 683-741.

Gouveia, N., Mendonça, G. A. S., Leon, A. P. and Correia, J. E. M., 2003. “Poluição do ar e efeitos na saúde nas populações de duas grandes metropóles brasileiras.” Epidemiologia e Serviços de Saúde, pp. 29-40.

Khan, F. I. and Sadiq, R., 2005. “Risk-based prioritization of air pollution monitoring using fuzzy synthetic evalution technique.” Environmental Monitoring and Assessment, pp. 261-283.

Kiely, G., 1996, “Environmental Engineering”, Berkshire England, McGraw-Hill.

Korbicz, .J, Koscielny, J. M., Kowalczuk, Z. and Cholewa, W., 2004. “Fault Diagnosis - Models, Artificial Intelligence, Applications”. Berlin: Springer.

Ministry of Health, 1954. “Mortality and morbidity during the London fog of December 1952.” Reports on Public Health and Medical Subjects.

Silvert, W., 2000. “Fuzzy indices of environmental conditions.” Ecological Modelling, pp. 111-119.

Sugeno, M., 1985, “Industrial applications of fuzzy control”, Elsevier Science Pub. Co.

Yager, R.R. and Filev, D.P., 1994. “Essentials of fuzzy modeling and control”. New York, John Wiley & Sons.

Zadeh, L., 1965. “Fuzzy sets. Information and control”, v. 8, pp. 338-353.

7. RESPONSABILITY NOTICE

The authors are the only responsible for the printed material included in this paper.