FUZZY INFERENCE SYSTEMS FOR ESTIMATION OF AIR QUALITY INDEX

Daniel Dunea, Alexandru Alin Pohoată, Emil Lungu
Valahia University of Târgoviște, Romania
ddunea@yahoo.com, alinpohoata@yahoo.com, emil_lungu@hotmail.ro

Abstract
The paper provides a reliable method in assessing the Air Quality Index (AQI) by employing fuzzy logic. Long-time series of sulfur dioxide (SO2), nitrogen dioxide (NO2) and particulate matter (PM10) obtained through continuous monitoring at Romanian Târgoviște DB-1 station were inputted in a Mamdani fuzzy inference system (FIS) obtaining the AQI output in a more comprehensive matter. The rules were structured by level of air pollutants concentration. All defined rules were evaluated in parallel in a random order. The triangular and trapezoidal membership functions fitted to the intended purpose of computational efficiency. The AQI synthetic indicator allows the user a quick interpretation of the status of the surrounding air in residential areas.

Keywords: fuzzy inference, air pollution, air quality index.
2010 MSC: 00A69.
Presented at CAIM 2011.

1. INTRODUCTION

This work presents the use of fuzzy techniques for Air Quality Index (AQI) assessments. In Romania, AQI is established on a scale from 1 to 6 (1-excellent, 6-very bad) using data acquired by automated stations of the National Network of Air Quality Monitoring [1].

RNMCA comprises more than 140 automated stations for continuous monitoring of air quality managed by 41 local centers, which collect and transmit data from stations and AQI information to the public panel. The information is also transmitted after primary validation for certification at the National Reference Laboratory (LNR) in Bucharest.

The selection of AQI is made taking into account the worst value from a set of specific indices. Specific indices are established using 6 intervals of pollutant concentrations such as sulfur dioxide (SO2), nitrogen dioxide (NO2), Ozone (O3), Carbon monoxide (CO) and particulate matter (PM10). At least three specific indices should generally be available to calculate the general AQI. Fuzzy inference methods assume that all the rules are activated at every cycle and contribute collectively to the solution. It is a parallel one-shot inference, but the inference process can continue as the new inferred results can be fed again as inputs [5]. Methods based on the fuzzy sets theory should be applied in the context of environmental numbers. The boundaries
between an acceptable and an unacceptable concentration is not to be considered as sharp, but as fuzzy, with implications for subsequent action plans. The use of fuzzy numbers is proposed as a suitable technique for handling environmental criteria and tackling decisions made under uncertainty [8]. Fuzzy logic was successfully applied for various air quality assessments and predictions [2], [3], [4], [6], [7]. The literature shows that fuzzy inference systems are frequently considered for short-term forecasting applied to emissions control or to AQI computations. These methods have some advantages over deterministic approaches [9]. In this paper, large time series of SO2, NO2 and PM10 were used as inputs in a Mamdani fuzzy inference system (FIS) to model the AQI output in a more intuitive manner.

2. TECHNICAL DESCRIPTION

Specific indices are established using 6 intervals of pollutant concentrations such as:

- Sulfur dioxide (SO2),
- Nitrogen dioxide (NO2),
- Ozone (O3),
- Carbon monoxide (CO),
- and particulate matter (PM10).

The most important air pollutants variables were selected for analysis (SO2, NO2 and PM10) because they are primary pollutants emitted into the atmosphere directly from various sources retaining the same chemical form and significantly affecting the human health and components of the ecosystem. Figure 1 presents six intervals of concentration corresponding to each index.

<table>
<thead>
<tr>
<th>SO2 concentration (μg/m³)</th>
<th>Specific Index</th>
<th>SO2 concentration (μg/m³)</th>
<th>Specific Index</th>
<th>PM10 concentration (μg/m³)</th>
<th>Specific Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-49,(9)</td>
<td>1</td>
<td>0-49,(9)</td>
<td>1</td>
<td>0-9,(9)</td>
<td>1</td>
</tr>
<tr>
<td>50-74,(9)</td>
<td>2</td>
<td>50-99,(9)</td>
<td>2</td>
<td>10-19,(9)</td>
<td>2</td>
</tr>
<tr>
<td>75-124,(9)</td>
<td>3</td>
<td>100-135,(9)</td>
<td>3</td>
<td>20-29,9)</td>
<td>3</td>
</tr>
<tr>
<td>125-149,(9)</td>
<td>4</td>
<td>140-199,(9)</td>
<td>4</td>
<td>30-49,9)</td>
<td>4</td>
</tr>
<tr>
<td>350-499,(9)</td>
<td>5</td>
<td>210-319,(9)</td>
<td>5</td>
<td>50-99,9)</td>
<td>5</td>
</tr>
<tr>
<td>&gt;500</td>
<td>6</td>
<td>&gt;400</td>
<td>6</td>
<td>&gt;100</td>
<td>6</td>
</tr>
</tbody>
</table>

Fig. 1. Intervals of concentration for each variable and their corresponding specific index - AQI (1-excellent, 6-very bad)

Fuzzy inference takes inputs, applies fuzzy rules, and produces explicit outputs. The FISs of Mamdani type are suitable for air quality evaluation as both the inputs and the outputs of the FIS are represented by the values of linguistic variables [4].
The primary mechanism relying on If-Then (facts - state/action) rules has defined an input space into output space. Fuzzy inference permitted the reading of the input vector values (3 inputs) and based on the set of rules, allocated the values of vector output (a single output, which is the AQI). Numerical input values can be easily translated into descriptive words (e.g. excellent, very good, good, favorable, bad and very bad).

Membership function Mf shows the extent to which a value from a domain is included in a fuzzy concept. There is no general method for designing form, number and parameters of input and output membership functions. The design of input and output membership functions is based on the limits set by the government (see Figure 1) and on the recommendations of environmental experts. The triangular and trapezoidal membership functions fitted to the intended purpose. Consequently, selected membership functions, as specific curves, defined how each entry point in space belonged to a degree of membership in the range 0 and 1.

Figure 2 shows five triangular and one-side trapezoidal Mfs for each of the three inputs.

![Membership functions of the input variables (nitrogen dioxide - NO2; sulfur dioxide - SO2; particulate matter - PM10)](image)

Figure 3 depicts the triangular functions responsible for the AQI modeling by level of pollution.

The structure of production rules was developed using Fuzzy Logic Toolbox module of MATLAB. The rules were structured by level of air pollutants concentration (more than 80 rules). Figures 4 and 5 highlight the windows of rule editor and rule viewer in Matlab showing some of the developed rules and the relationship between inputs and AQI output.
2.1. LARGEST OF MAXIMUM (LOM) DEFUZZIFICATION

There are numerous defuzzification methods. Each defuzzification method outputs different results. There is no exact rule on selection of defuzzification for certain applications.
Fuzzy inference systems for estimation of Air Quality Index

LOM was selected due to the nature of the problem: the air quality is established by the maximum of the pollution indices that makes the usual Center Of Gravity (COG) method unsuitable.

In fuzzification, an operator transforms crisp data into fuzzy sets, so that data can be processed by the rule-base. The fuzzification process can be described as:

$$
\mathcal{A} = \text{fuzzifier}(x_0)
$$

where \( x_0 = [x_1, x_2, ..., x_n]^T \) is an input vector and \( \mathcal{A} = [\mathcal{A}_1, \mathcal{A}_2, ... \mathcal{A}_n]^T \) are fuzzy sets, and fuzzifier represents a fuzzification operator. The Mamdani implication of max-min fuzzy inference is given by:

$$
\mu_{\mathcal{B}}(z) = \max[\min[\mu_{\mathcal{A}_1}(input(x_1)), \mu_{\mathcal{A}_2}(input(x_2))]], \quad k = 1, ..., r
$$

where \( \mu_{\mathcal{B}}(z) \) is the height of the aggregated fuzzy set for the \( r \) rules. The aggregated fuzzy set is defuzzified to yield crisp output as represented by:

$$
z^* = \text{defuzzifier}(\mathcal{Z}_k)
$$

where \( z^* \) is the crisp output and \( \mathcal{Z}_k \) is the fuzzy set resulted from aggregation, and “defuzzifier” represents the defuzzification operator.

The LOM defuzzification is done in two steps. First the largest height in the union is determined:

$$
hgt(\mathcal{Z}_k) = \sup_{z \in \mathcal{Z}} \mu_{\mathcal{B}}(z)
$$

Where supremum (sup) is the least upper bound. Then, the largest of maximum is calculated:

$$
z^* = \sup_{z \in \mathcal{Z}} \left\{ z \in \mathcal{Z} \mid \mu_{\mathcal{B}}(z) = hgt(\mathcal{Z}_k) \right\}
$$

3. EXPERIMENTAL RESULTS

The set of fuzzy rules algorithm developed in Matlab has performed the following steps:

- scalar representing the system input - 3 variables were transformed into membership functions through the fuzzifying functions;
- the information was transferred to the inference engine;
- values of membership functions were transformed into output by defuzzification of the scalar value, representing the output indicator which evaluates the air quality status (0 - excellent quality; 1 - bad quality).

In this paper, we have selected for presentation only a segment of 1871 values for each input from the existing database in order to show a clear image of the FIS
capabilities. Figure 6 offers information regarding the fluctuations of AQI during 78 days randomly selected from the database. Numerous exceedings over the 0.5 level have occurred corresponding to low quality of the surrounding air near the station. FIS output was statistically analyzed determining central tendency, dispersion and distribution indicators.

Figure 7 shows the AQI frequency. Of particular interest from the environmental point of view are the values occurring in the upper intervals. The presented FIS might be able to “trigger” alarms helping residents through early warnings when bad air quality conditions might occur. Analyzed data showed the occurrence of worst conditions (8 times) and also worrisome situations when AQI was in bad conditions (55 times). The developed FIS can be a useful tool for air quality control taking into account that the analyzed pollutant time series were randomly selected, but situations of concern were detected.

The standardized skewness and kurtosis values were within the range expected for data from a normal distribution (-2, 2).
4. CONCLUSION

The concepts underlying fuzzy technology are successfully used in air quality modeling, allowing an alternative approach in solving specific environmental problems when the objectives or constraints are not precisely defined, and necessary information is missing, is sporadic or discontinuous. The AQI synthetic indicator allows the user a quick interpretation of the status of the surrounding air (imissions). By integrating three pollutant time series into single indicator, the FIS provided a better and comprehensive image of the air pollution status within the area of interest compared to classical pollution report using individual pollutant concentrations. Some issues might occur if using more than three variables due to the number increment of adjacent rules. Future work will focus on the following directions: Due to the large number of FIS rules a simplification is required and thus we will consider the utilization of parallel fuzzy systems that will reduce both the rules number as well as the processing time. We will test the adequacy of a time delay neural network implementation selecting the structure according to the best forecasting properties using FIS-generated AQI as input.

References


