# Construction and Analysis of a Fuzzy Sanitary Quality Index

**José Arnaldo F. Roveda**, **Antonio Cesar G. Martins**,
**Sandra R. M. M. Roveda**, **Roberto W. Lourenço**.
Depto de Engenharia Ambiental, Campus Sorocaba, UNESP,
18087-180, Sorocaba, SP
E-mail: {roveda, amartins, sandra, robertow}@sorocaba.unesp.br

**Samuel Barsanelli Costa**
Analista de Geoprocessamento
GEOJA - Mapas Digitais
E-mail: samuelbarsanelli@msn.com.

**Abstract:** *This article describes the construction of a Sanitary Quality Index that takes in to account conditions of water supply, sanitarian facilities and garbage collection of houses in a city by a fuzzy approach based on an inference system that incorporates the experience of an environmental expert using a set of rules. This methodology was applied to data from the city of Sorocaba, São Paulo State, Brazil generating index values for 547 sectors. In order to compare the obtained results, a classical statistical methodology was applied in the same data.*

**Keywords:** *Fuzzy inference systems, environmental index, sanitary quality*

## 1 Introduction

Population life condition is close related to a wide group of factors such as: quality of public services, environmental aspects and governmental programs. Within those factors, aspects that influences sanitary conditions can have great impact with important effects in human health. Lack of good quality water supply, sanitary facilities and garbage collection can lead to the development of a series of diseases such as: hepatitis, leptospirosis and polio.

The development of an index that takes in to account those three sanitary factors can indicate the degree of exposure that people are to those diseases and other risks.

Given a city divided into sectors, an index can be constructed by defining a set of classification classes for water supply, sanitary facility and garbage collection conditions and use some methodology to mix them up.

It can be easily seem that in such an approach there is a great deal of inherent uncertainty. In the construction of such classes, the environmental specialist must define them based on some characteristics. For each factor, after the number of houses in each class is counted, the related index value must be computed in a way that represents the overall distribution of conditions that are presented in that sector. For the final index, the mixing procedure must be such that the output index has a value that represents the condition of the related sector.

The use of fuzzy logic has been characterized to be conceptually easy of understanding, and based on natural language, have been successfully used to model non-linear functions, to build inference systems in top of the experience of experts, and to deal with imprecise data [1], [2].

As stressed by Silvert [3] fuzzy logics provides a convenient methodology for the development of environmental indexes, with this formalism being a significant tool to describe the environmental conditions with implicit value judgements. For instance, Ocampo-Duque [4] proposes the construction of a fuzzy water quality index, showing results that improves decision making in water management, and comments that the flexibility of fuzzy logics based on the use of natural language facilitates the correlation of subjective information.

This article is organized as following: section 2 presents the methodology for the construction of water supply, sanitary facilities and garbage destination conditions indexes, section 3 discusses the fuzzy inference system, in section 4, for a comparison purpose, a classical statistical approach for the construction of a Sanitary Quality Index is presented, section 5 shows the obtained results and finally section 6 brings the final remarks.

## 2 Classification of water supply, sanitary facilities and garbage destination conditions

To construct the Sanitary Quality Index (SQI), three factors will be taken into account: water supply, sanitary facilities and garbage destination. The condition classification of such factors will be done according to three classes denominated *good, median* and *weak*, taking into account certain characteristics of each class, described as follows.

For water supply, condition will be considered *good* if the private house has permanent water supply by a public service net, *median* if the supply is done by wells and *weak* otherwise.

Sanitary facilities condition will be considered to be *good* if the private house has a bathroom, or similar, connected to the public sewer net, *median* if the bathroom, or similar, is connected to septic tank or *weak* otherwise.

Garbage destination condition will be considered *good* if garbage from the private house is permanently and directly collected by public service, *median* if garbage is indirectly collected using collective buckets or *weak* otherwise.

In order to quantify each of those conditions in a sector, three indexes are created: Water Supply Index (WSI), sanitary Facility Index (SFI) and Garbage Destination Index (GDI). For a given sector, the number of houses in each class for each sector is counted, normalized by the total number of house in the sector and a weighted sum is calculated. The weight is defined as 0.33 for the class *good*, 0.66 for *median* and 1 for *weak*. Thus, the close to 0.33 the index, the better is the related condition.

## 3 Development of the Fuzzy Sanitary Quality Index (FSQI)

The proposed fuzzy approach incorporates the analytical knowledge of the expert when mixing WSI, SFI and GDI by the use of a set of rules and a Mandami inference system. To construct the set of rules, the expert was initially asked to evaluate each of the 27 different resulting possible mixing states as *good*, *median* and *weak*. After that, the expert was asked to group and further classify the states in each of the three classes in a way that a strength for the classification was obtained . Thus, a set of rules with weights for each class was constructed, as shown in Table 1. Those weights were used in the Mandami inference system [5].

As can be seen from Table 1, by the expert analysis the quality of sanitary facilities has a strong influence in the final classification. If, for instance, water supply is considered *weak*, sanitary facilities are *good* and garbage destination conditions is *weak*, the overall classification is *median*. Otherwise, if water supply and garbage destination are considered *good*, but sanitary facilities *weak*, the overall classification is *weak*.

For each index (WSI, SFI and GDI) the membership functions were created in the range [0.33; 1] with support interval from [0.33; 0.66] for the condition *good*, [0.33; 1] for *median* and [0.66;1] for *weak*, as shown in Fig.1. The same feature for FSQI membership functions was considered.

Since it is desired that for an input of WSI=0.33 (Good), SFI=0.33 (Good) and GDI=0.33 (Good), FSQI assumes the value 0.33 (Good), for the defuzzification process the mean of maximum method was chosen [2].

Table 1: Set of rules resulting from the expert analysis to all possible combinations of WSI, SFI and GDI. Within the same resulting class the weight closer to one indicates the strenght of the rule.

| WSI | SFI | GDI | Expert Analysis | Weight |
|---|---|---|---|---|
| Good | Good | Good | Good | 1.0 |
| Good | Good | Median | Good | 0.9 |
| Median | Good | Good | Good | 0.9 |
| Good | Good | Weak | Median | 1.0 |
| Weak | Good | Good | Median | 1.0 |
| Median | Good | Median | Median | 1.0 |
| Median | Good | Weak | Median | 1.0 |
| Weak | Good | Weak | Median | 0.9 |
| Good | Median | Good | Median | 0.9 |
| Good | Median | Median | Median | 0.9 |
| Good | Median | Weak | Median | 0.8 |
| Median | Median | Good | Median | 0.8 |
| Median | Median | Median | Median | 0.8 |
| Weak | Median | Good | Median | 0.8 |
| Median | Median | Weak | Median | 0.7 |
| Weak | Weak | Weak | Weak | 1.0 |
| Good | Weak | Weak | Weak | 0.9 |
| Median | Weak | Median | Weak | 0.9 |
| Weak | Weak | Median | Weak | 0.9 |
| Median | Weak | Weak | Weak | 0.9 |
| Weak | Weak | Good | Weak | 0.8 |
| Good | Weak | Median | Weak | 0.8 |
| Median | Weak | Good | Weak | 0.8 |
| Good | Weak | Good | Weak | 0.7 |
| Weak | Median | Weak | Weak | 0.7 |
| Weak | Good | Median | Weak | 0.6 |
| Weak | Median | Median | Weak | 0.6 |

# 4 Description of a Sanitary Quality Index (SQI) by a statistical approach

The three indexes (WSI, SFI and GDI) bring general information about the conditions in each sector. To obtain SQI with a statistical approach, the distribution of the three indexes values in the city were taken in to account.

For each index (WSI, SFI and GDI) the mean and the standard deviation were calculated from the values obtained through the city and the coefficient of variation (CV), given by the standard deviation divided by the mean, was obtained. Then, for each index, weight values were obtained as follows:

$$W_{WSI} = \frac{CV_{WSI}}{(CV_{WSI} + CV_{SFI} + CV_{GDI})} \tag{1}$$

$$W_{SFI} = \frac{CV_{WSI}}{(CV_{WSI} + CV_{SFI} + CV_{GDI})} \tag{2}$$

$$W_{GDI} = \frac{CV_{WSI}}{(CV_{WSI} + CV_{SFI} + CV_{GDI})} \tag{3}$$

where $CV_{WSI}, CV_{SFI}, CV_{GDI}$ are the coefficient of variation for water supply index, sanitary facility index and garbage destination index, respectively.

SQI is finally obtained with a normalized weighted sum of WSI, SFI and GDI by the respective weight:

$$SQI = \frac{(WSI * W_{WSI} + SFI * W_{SFI} + GDI * W_{GDI})}{(W_{WSI} + W_{SFI} + W_{GDI})} \tag{4}$$

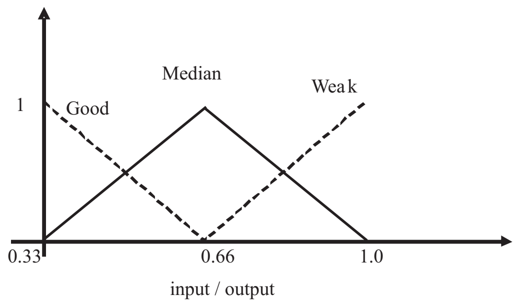


Figure 1: General features for the membership functions used for WSI, SFI, GDI and FSQI.

It can be seen from (4) that the condition with greater variance will have a stronger weight and that SQI varies from 0.33, the best SQI ,to 1, the worst condition.

It can also be noticed that the index was constructed in such a way to be as objective as possible.

It should be stretched that SQI is an index given for each sector in the city, taking in to account the overall condition of the city. This way, this is an index with the purpose of inter-sector analysis of the same city.

## 5 Results

In order to study the output from the fuzzy approach and compare with the statistical one, data of water supply, sanitary facilities and garbage destination conditions for the houses from the city of Sorocaba, in the state of São Paulo, Brasil, from year 2000 was analyzed. The city was divided into 547 sectors where the number of house in each condition (as described in section 2) was counted.

From the 547 sectors, 298 had WSI, SFI and GDI smaller than 0.35, with 192 sectors presenting WSI=0.33, SFI=0.33 and GDI=0.33.

FSQI and SQI were computed for all the 547 sectors. The output for SQI was such that 531 had values lower than 0.5 and 16 in the interval [0.50; 0.67]. In the case of FSQI, 521 sectors had an output lower than 0.5, 21 values around 0.66 and 5 values in the interval [0.89; 0.93].

A total of 32 sectors had a difference between SQI and FSQI greater than 20%. Considering absolute values, only 20 sectors had a difference equals or greater than 0.20. Table 2 shows these values, where it can be seen that SFI presents high values whereas WSI and GDI lower ones. For those cases, FQSI results in high values, what classifies the region as median or weak, whereas SQI results in lower values and a better classification for the sector.

Table 2: Values of SQI and FHQSI for city sectors were $|SQI - FSQI| \geq 0.20$.

| WSI | SFI | GDI | SQI | FHQSI | SQI-FSQI |
|---|---|---|---|---|---|
| 0,33 | 0,86 | 0,34 | 0,56 | 0,90 | 0,34 |
| 0,36 | 0,87 | 0,38 | 0,59 | 0,91 | 0,32 |
| 0,33 | 0,51 | 0,33 | 0,41 | 0,67 | 0,26 |
| 0,53 | 0,84 | 0,48 | 0,65 | 0,90 | 0,26 |
| 0,34 | 0,51 | 0,33 | 0,41 | 0,67 | 0,25 |
| 0,33 | 0,53 | 0,33 | 0,41 | 0,66 | 0,25 |
| 0,33 | 0,52 | 0,34 | 0,41 | 0,66 | 0,25 |
| 0,35 | 0,56 | 0,37 | 0,44 | 0,66 | 0,22 |
| 0,46 | 0,95 | 0,49 | 0,68 | 0,90 | 0,21 |
| 0,33 | 0,61 | 0,33 | 0,45 | 0,66 | 0,21 |
| 0,57 | 0,97 | 0,53 | 0,73 | 0,92 | 0,20 |

## 6 Conclusions

This paper discusses the construction of an index for sanitary quality condition to be used in the analyzes of city sectors by a fuzzy approach using a set of weighted rules and a Mandami inference system.

For a base of comparison a classical statistical index, based on metrics from the distribution of sector results, were used.

The proposed indexes were calculated for data from the city of Sorocaba in the state of São Paulo, Brazil, from year 2000.

The general output shows similar results for the fuzzy and statistical approaches but some important discrepancies were found. As can be seen from table 2, the fuzzy approach keeps the idea from the expert that sanitary facilities has a key importance in the overall sanitary condition, what was some how masked in the statistical approach.

It should also be pointed out that the fuzzy approach allowed a much easier interaction with the specialist that did not have a mathematical background, facilitating the construction of the inference model.

## References

[1] L.A. Zadeh. Fuzzy logic equals computing with words. *IEEE Trans Fuzzy Syst*, 4:103-111, 1996.

[2] T. J. Ross. *Fuzzy Logic with Engineering applications.* John Wiley & Sons, West Sussex, 2004.

[3] W. Silvert. Fuzzy indices of environmental conditions. *Ecological Modelling*, 130:111-119, 2000.

[4] W. Ocampo-Duque, N. Ferré-Huguet,J.L. Domingo and M. Schuhmacher. Assessing water quality in rivers with fuzzy inference systems: A case study. *Environmental International*, 32:733-742, 2006.

[5] R.V. Demicco and G.J.Klir. *Fuzzy Logic in Geology.* San Diego, California, 2004.