Indexing the Environmental Quality Performance Based on A Fuzzy Inference Approach

Lizda Iswari
Department of Informatics, Universitas Islam Indonesia, Jalan Kaliurang Km. 14 Ngaglik Sleman Yogyakarta Indonesia
Lizda.iswari@ui.ac.id

Abstract. Environmental performance strongly deals with the quality of human life. In Indonesia, this performance is quantified through Environmental Quality Index (EQI) which consists of three indicators, i.e. river quality index, air quality index, and coverage of land cover. The current of this instrument data processing was done by averaging and weighting each index to represent the EQI at the provincial level. However, we found EQI interpretations that may contain some uncertainties and have a range of circumstances possibly less appropriate if processed under a common statistical approach. In this research, we aim to manage the indicators of EQI with a more intuitive computation technique and make some inferences related to the environmental performance in 33 provinces in Indonesia. Research was conducted in three stages of Mamdani Fuzzy Inference System (MAFIS), i.e. fuzzification, data inference, and defuzzification. Data input consists of 10 environmental parameters and the output is an index of Environmental Quality Performance (EQP). Research was applied to the environmental condition data set in 2015 and quantified the results into the scale of 0 to 100, i.e. 10 provinces at good performance with the EQP above 80 dominated by provinces in eastern part of Indonesia, 22 provinces with the EQP between 80 to 50, and one province in Java Island with the EQP below 20. This research shows that environmental quality performance can be quantified without eliminating the natures of the data set and simultaneously is able to show the environment behavior along with its spatial pattern distribution.

Keywords: Mamdani FIS, Land Cover, River Index, Water Index, Spatial

1. Introduction

Environmental degradation can be stated as a disaster that causes losses for the community, not merely in the economic aspects but a greater impact may also attach to the human lives. Efforts to reduce the rate of environmental degradation and restoration continue by both the government and community organizations.

In some countries, the environmental quality performance are monitored and assessed in many aspects, such as environmental dimension in quality of life [1], environmental awareness and behavior in terms of the social-psychological factors [2], and environmental quality health index for cities based on the noise levels and the air pollutant concentration [3].

Starting in 2009, the Indonesia Ministry of Environment has also developed a similar monitoring instrument named Environmental Quality Index (EQI) [4]. This instrument consists of three key indicators, namely river water quality, air quality, and land cover/vegetation. Currently, the
determination of EQI uses a simple weighting formula in which the water and air quality are weighted 30% respectively, and the land cover index holds 40% to represent the environmental quality performance at the provincial level [5].

Data of environmental quality index is closely related to the continuous nature of phenomenon on the earth surface. The interpretation of environmental conditions may contain some uncertainties and has a range of circumstances that possibly is partially true, i.e. data of environmental indicators that may be interpreted as “between” poor and good classification, and hence will be less appropriate if processed under a common statistical approach. As mentioned in [6], a system that has a certain degree of complexity will be impossible to provide an accurate and meaningful description of system behavior by using a quantitative method. Moreover, we analyzed that the current of EQI determination has not provided some inferences yet to inform dominant factors that lead a province to have a good or poor environmental quality performance.

One approach to overcome these problems can be done by applying the Fuzzy Inference System (FIS). It has a similar way of reasoning such as human thinking and can be effectively used to describe approximate and uncertain phenomena in the real world [7]. The FIS outcome may provide outstanding results, especially when the existing system cannot be interpreted by quantitative methods due to its complexity nor has inaccurate and uncertain qualitative information [6]. Research with regard to FIS can be found in [6] diagnosing the degradation of feed water heater and in [8] determining the quality of dates.

In this research, we aim to manage the indicators of EQI with a more intuitive computation technique that is able to handle an object with fuzzy natures and makes an inference related to the environmental quality performance. Research was conducted in three stages based on the Mamdani Fuzzy Inference System (MAFIS) approach, i.e. fuzzification, data inferencing, and defuzzification. The selection on Mamdani is based on its ability to receive human input as well as its advantages on synthesizing a set of linguistic control rules which can be obtained from experienced human operators [8].

2. Environmental Parameters
The EQI instrument used in Indonesia consists of three indicators and specified into 10 parameters [5][4].

2.1. The River Quality Parameters
The river index commonly quantified through these seven parameters as follow:

a. Total Suspended Solid (TSS): a residue of total solids retained by a filter and working to estimate the rate of sedimentation. The threshold value of good TSS is 50 mg/L.

b. Dissolved Oxygen (DO): the amount of oxygen contained in the water. Water that has a high DO showed a low level of contamination with the threshold value of 4 mg/L.

c. Chemical Oxygen Demand (COD): the amount of organic compounds in the water equivalent to the amount of oxygen to oxidize organic compounds chemically with the threshold value of 25 mg/L.

d. Biochemical Oxygen Demand (BOD): the amount of oxygen needed by microorganisms to decompose organic compounds under the aerobic conditions with the threshold value of 3 mg/L.

e. Total-Phosphate (TP): the amount of organic compounds, such as proteins, urea, and the result of the decomposition process. The higher TP leads to the decline in Dissolved Oxygen values with the threshold value of 0.2 mg/L.

f. Fecal Coli (FC): a group of bacteria found in domestic waste and consisting of Coliform, Eschericia coli, and Streptococcus faecalis bacteria. The threshold value was determined based on the data average in the last 3 years, i.e. 5000/100mL.

g. Total Coli (TC): the contamination caused by the human feces. Similar to Fecal Coli, the threshold value was also based on the data average in the last 3 years, i.e. 3000/100mL.
2.2. The Air Quality Parameters
The most common of air pollution mainly come from the combustion process which includes fossil fuels, such as Nitrogen Oxides (NO), Carbon Monoxide (CO), and Sulfur Dioxide (SO). For indicating the air quality performance, the Ministry of Environment has set NO2 and SO2 as the parameter. The threshold value of the healthy air for NO2 should not exceed 100μg/m3 and SO2 is 60μg/m3.

2.3. The Land Cover
The land cover is the condition of the earth’s surface that describes the appearance of land cover and vegetation. The state of land cover in Indonesia has been obtained as the interpretation of satellite images Landsat 8 OLI. This interpretation is divided into two major classifications, i.e. the forested areas and non-forested areas. In this research, we took the total forested areas as the indicator of the land cover with the threshold value of 40%.

3. Fuzzy Inference System
Fuzzy refers to an adjective meaning of vague or unclear [9]. The fuzzy logic is introduced as a form of basic mathematical logic development which provides a methodology to model uncertainty and inaccurate qualitative information [6]. To develop a system with fuzzy logic model, [9] described three core elements of fuzzy inference system.

3.1. Fuzzification
Fuzzification deals with a set of process to alter crisp inputs into fuzzy values using the membership function. To fuzzify the data input, we have to define the fuzzy variables along with its fuzzy sets. Fuzzy set is a description of the common condition of fuzzy variable.

We applied 10 parameters of EQI as the fuzzy input variables, i.e. Total Suspended Solids (TSS), Dissolved Oxygen (DO), Chemical Oxygen Demand (COD), Biochemical Oxygen Demand (BOD), Total Phosphate (TP), Fecal Coli (FC), Total Coli (TC) belong to the river index parameters, NO2 and SO2 belong to the air index parameters and the coverage of land cover (LC). On each of river index and land cover, we defined two types of fuzzy sets, i.e. Good and Poor. Meanwhile for the air index parameters it had Good and Medium fuzzy sets. For the fuzzy output variable, we used the terminology of Environmental Quality Performance (EQP) and defined by having 3 types of fuzzy sets, i.e. High, Moderate, and Poor.

Afterward we defined fuzzy set domain as the allowed value to be processed in a fuzzy set. We considered the threshold and standard deviation of the data set to define data domain of input variables. In these, the threshold value worked as the center value and the standard deviation worked as the range of each fuzzy set. For instance, on TSS we found that the threshold value was 50 mg/L and the standard deviation was 36.5 mg/L (rounded to 40). Hence, the data domain of TSS was in the range of 30 to 70 with 50 on the center as can be seen in Figure 1. Meanwhile, the data domain of output variable was determined manually to the interval of 0 to 100.

The last step on fuzzification was to defining a fuzzy set membership function which mapped an input to a membership degree and was represented in the real number with the interval of 0 to 1. The membership function can be illustrated as curves to describe the mapping mechanism of continuous or linear data. In this research, we found all fuzzy input variables had the continuous characteristics and hence presented as the sigmoid curves. For the fuzzy output it was defined as the linear objects and was presented using the trapezoidal curves.

In more detail, for the membership function of fuzzy input we applied the growth and depreciate of sigmoid function as illustrated in Fig. (1) with the mathematical functions that are described in Eq. (1) and (2). Meanwhile, for the fuzzy output it was defined as the trapezoidal linear increase and decrease and the triangular function with the mathematical function in Eq. (3), (4) and (5).

For example, the TSS value on a province was 47.03 mg/L. By mapping this value to its membership function as Eq. (1) and (2), we found that the membership degree of TSS on Good at 0.6374 and Poor
at 0.3625 as seen in Fig. (3), indicating that TSS of this province was closer to Good rather than Poor condition.

\[
Growth[x; a, b, c] = \begin{cases} 
0 & \text{for } x < a \\
2 \left( \frac{x-a}{c-a} \right)^2 & \text{for } a \leq x \leq b \\
1 - 2 \left( \frac{c-x}{c-a} \right)^2 & \text{for } b < x \leq c \\
0 & \text{for } x > c 
\end{cases}
\]  \hspace{1cm} (1)

\[
Depreciate[x; a, b, c] = \begin{cases} 
1 & \text{for } x < a \\
1 - 2 \left( \frac{x-a}{c-a} \right)^2 & \text{for } a \leq x < b \\
2 \left( \frac{c-x}{c-a} \right)^2 & \text{for } b \leq x \leq c \\
0 & \text{for } x > c 
\end{cases}
\]  \hspace{1cm} (2)

\[
TrapLinUp[x; a, b, c] = \begin{cases} 
0 & \text{for } x \leq b \\
\frac{x-b}{c-b} & \text{for } b < x < c \\
1 & \text{for } x \geq c 
\end{cases}
\]  \hspace{1cm} (3)

\[
TrapLinDown[x; a, b, c] = \begin{cases} 
1 & \text{for } x \leq a \\
\frac{b-x}{b-a} & \text{for } a < x < b \\
0 & \text{for } x \geq b 
\end{cases}
\]  \hspace{1cm} (4)

\[
Triangle[x; a, b, c] = \begin{cases} 
0 & \text{for } x \leq a \text{ or } x \geq c \\
\frac{x-a}{b-a} & \text{for } a < x < b \\
\frac{c-x}{c-b} & \text{for } b < x < c \\
1 & \text{for } x = c 
\end{cases}
\]  \hspace{1cm} (5)

where:
- \(a\) = bottom boundary;
- \(b\) = center point;
- \(c\) = upper boundary

\textbf{Figure 1.} The Membership Curves of Fuzzy Input

\textbf{Figure 2.} The Membership Curves of Fuzzy Output
3.2. Data Inference

Fuzzy inference strongly relies on the fuzzy rules and the membership degree of the fuzzy sets, i.e. to apply the implication function and compose all the fuzzy outputs. In this stage, we applied two levels of data reasoning. First, we applied the monotonous reasoning method for each of main environment indicators, i.e. to define the index of river quality, air quality and the coverage of land cover. Second, we combined these indexes to represent the index of Environment Quality Performance (EQP) by using the min-max inference approach.

On the first method, the monotonous reasoning, we related two fuzzy variables with simple implication rules. E.g. rule for determining a “High” level of river quality index was defined as follows:

\[ \text{IF river quality is Good THEN river index is High} \]

The value of river quality was obtained based on the average of all the components forming the index, i.e. the average of “Good” membership degree from seven parameters (TSS, DO, COD, BOD, TP, FC, TC) and mapped it to the “High” membership function to represent the river index. Similar approaches were also applied to define the “Low” river index based on the average of “Poor” values of all river parameters, and to define the “High” and the “Low” index of the air quality and the coverage of land cover.

| Rules # | Antecedents     | Consequent |
|---------|-----------------|------------|
|         | River | Air  | Land Cover | EQP          |
| 1       | High  | High | High       | High         |
| 2       | High  | High | Low        | Moderate     |
| 3       | High  | Moderate | High     | High         |
| 4       | High  | Moderate | Low      | Moderate     |
| 5       | Low   | High  | High       | Moderate     |
| 6       | Low   | High  | Low        | Moderate     |
| 7       | Low   | Moderate | High     | Moderate     |
| 8       | Low   | Moderate | Low      | Low          |

On the second reasoning method, we applied the minimum function to determine the implication of fuzzy rules and aggregated all output based on the maximum function. In this step, we had defined 8 fuzzy rules which have been derived from all combinations of environment indicator fuzzy sets as the rule’s antecedent, in which each indicator consisted of two fuzzy sets to determine the general index of Environmental Quality Performance (EQP) as the rule’s consequent as seen in Table 1. The first rules can be stated as follow:

\[ \text{IF (river index is High and air index is High and land cover is High) THEN EQP is High} \]
3.3. Defuzzification
Defuzzification is a process to obtain the crisp value as the representation of fuzzy natures. There are various methods to calculate the crisp output of fuzzy system. The most common method is the centroid function (composite moment). However, in this study we applied the Mean of Maximum (MOM) which is simple in computation and is able to give the reliable results. In MOM, the crisp value \( z^* \) is based on the average values of fuzzy sets with the maximum of membership degree as shown in Fig. (4) and Eq. (6).

\[
z^* = \frac{a+b}{2}
\]

where:
- \( a \) = bottom boundary
- \( b \) = upper boundary

4. Results and Discussion
The model for indexing the quality of environmental performance based on fuzzy inference approach in 33 provinces in Indonesia is presented here. We applied the environmental condition data set that was published by the Ministry of Environment of Indonesia in 2015.

In the beginning, we analyzed the result of fuzzification step. We compared the membership degree of 10 environmental quality parameters of 33 provinces. We used the “Good” membership degree which had the value more than 0.5 since it may represent data within the safe limits. As seen in Fig. (5), we inspected that of these 10 parameters the best circumstance was found in SO2 in which all provinces in Indonesia were within the limits of healthy air.

Meanwhile the lowest circumstance was found in the Land Cover (LC) since the number of provinces with the coverage of forested area larger than 40%, was smaller than its opposite side and had the smallest value of all parameters. Especially for the river index parameters, i.e. DO, TP, COD, TSS, FC, TC and BOD, we found the lowest circumstance in Biochemical Oxygen Demand (BOD).

We also analyzed the spatial distribution of these “Good” membership degrees. To simplify the explanation, we defined 5 groups of data based on the most well-known islands, i.e. Java, Kalimantan (abbreviated as Kali), Nusa Tenggara Papua (abbr. as Nuspa), Sulawesi (abbr. as Sula) and Sumatra (abbr. as Suma). The list of group members is shown in Table 2.

![Figure 5. Distribution of GOOD Membership Degree](image1)

![Figure 6. Map of EQP Distribution in Indonesia](image2)

In more detail, we analyzed the spatial distribution of island entities with “Good” membership degree below 0.5. This analysis was done for determining what indicators should have more consideration by these five major islands. As seen in Table 3, Java had TSS value of 50% indicating that 3 of 6 provinces had poor circumstances since their TSS were above the allowed threshold. Similarly, we inspected that
Java also had the poor circumstances on COD, BOD, TC, FC and LC. Kalimantan had the poor circumstances on BOD, Nusa Tenggara Papua (Nuspa) had the poor circumstances on BOD and TC, Sulawesi had the poor circumstances on TSS, and Sumatra had the poor circumstances on the coverage of land cover (LC).

Afterwards, we analyzed the result of data inference based on the monotonous and min-max reasoning. We found that the highest index of “High” river quality performance was 95.88 (in the scale of 0 to 100), including a province in Sumatra, and the lowest index was 10.73 which was located in a province in Java. For the index of air quality performance, we found that most provinces were having the value above 75, except one province in Java which had the air quality index below 50. For the land cover, we found the highest index was 100 which were located in 4 provinces in Nuspa and one province, respectively, in Kalimantan and Sulawesi. Meanwhile for the lowest index of land cover it was 0 and found in 8 provinces which dominantly were located in Java and Sumatra Island.

At the end, we analyzed the defuzzification results. The main output of defuzzification process was the environmental quality index of each province in Indonesia in the range of value from 0 to 100. In this final step, we found that 10 provinces had the EQP above 80 with the dominance of provinces in the group of Nusa Tenggara Papua (Nuspa) and none were located in Java Island. Meanwhile the lowest EQP was at 11.04 in a province in Java Island. The rest of 22 provinces had the EQP values between 50 and 80. To visualize the spatial distribution of EQP value among these 33 provinces in Indonesia, we illustrated them as a map as seen in Fig. (6), i.e. a darker red color indicated a lower value of EQP.

Table 2. List of Island Members

| Island | Member of Province |
|--------|--------------------|
| Java (6) | Jakarta, West Java, Central Java, Yogyakarta, East Java, Banten |
| Kali (4) | West Kalimantan, Central Kalimantan, South Kalimantan, East Kalimantan |
| Nuspa (7) | Bali, West Nusa Tenggara, East Nusa Tenggara, Maluku, North Maluka, Papua, West Papua |
| Sula (6) | North Sulawesi, Central Sulawesi, South Sulawesi, South East Sulawesi, West Sulawesi, Gorontalo |
| Suma (10) | Aceh, North Sumatra, West Sumatra, Riau, Island of Riau, Jambi, South Sumatra, Bengkulu, Lampung, Bangka Belitung |

Table 3. Percentage of Provinces with Good Membership Degree < 0.5

| Params | Java | Kali | Nuspa | Sula | Suma |
|--------|------|------|-------|------|------|
| TSS    | 50.0 | 25.0 | 14.3  | 83.3 | 20.0 |
| COD    | 66.7 | 25.0 | 14.3  | 33.4 | 30.0 |
| DO     | 33.4 | 0    | 0     | 33.4 | 10.0 |
| BOD    | 100  | 75.0 | 57.2  | 16.7 | 40.0 |
| TP     | 50.0 | 0    | 42.9  | 16.7 | 30.0 |
| TC     | 100  | 50.0 | 57.2  | 50.0 | 20.0 |
| FC     | 100  | 50.0 | 28.6  | 33.4 | 20.0 |
| SO2    | 0    | 0    | 0     | 0    | 0    |
| NO2    | 16.7 | 0    | 0     | 0    | 0    |
| LC     | 100  | 50.0 | 14.3  | 33.4 | 80.0 |

5. Conclusions
Based on the research findings, there are several conclusions that can be drawn here:

1. On the fuzzification step, there were analyses regarding environmental parameter circumstances. We found that among three types of index that form the Environmental Quality Index, the best circumstances was the air quality performance, meanwhile the coverage of the green land cover was the lowest circumstance in Indonesia.
2. Data inference has been done based on the monotonous and mix-max reasoning to quantify the performance of each indicator which was part of the environmental quality index. Meanwhile, the control rules as the center of data inference were derived from all combinations of environment indicator fuzzy sets. However, these rules only cover the three main indexes, i.e. river, air, and land cover not included the environmental parameters. Hence, we still need a better method to generate the fuzzy rules which also consider all aspects of environmental performance.

3. We also have provided an analysis of data spatial distribution, i.e. research objects were grouped into five islands, and found that Java Island had the most number of provinces with the environmental parameters below the threshold or the safe limits.

4. This research showed that the environmental quality performance can be quantified without eliminating the natures of the data set and is simultaneously able to show the environment behavior by considering the control rules as the representation of human way of thinking and reasoning along with its spatial pattern distribution.

Acknowledgment
Thank you to the Directorate of Research and Service Communities of Universitas Islam Indonesia for the funding and also for Hadi Kusuma Hidayat for providing the data sets.

References
[1] Streimikiene D 2015 Environmental indicators for the assessment of quality of life Intellect. Econ. 9 pp 67–79.
[2] Mei N S, Wai C W, and Ahamad R 2016 Environmental awareness and behaviour index for Malaysia Procedia - Soc. Behav. Sci. 222(7) pp 668–75.
[3] Silva L T 2015 Environmental quality health index for cities Habitat Int. 45(P1) pp 29–35.
[4] Environment M 2012 Report of Environmental Quality Index in Indonesia 2012 (Indeks Kualitas Lingkungan Hidup Indonesia).
[5] Environment M and Forestry M 2014 The Statistics of Environmental Quality and Forestry in Indonesia 2014 (Statistik Kementrian Lingkungan Hidup dan Kehutanan Tahun 2014).
[6] Kang Y K, Kim H, Heo G, and Song S Y 2017 Diagnosis of feedwater heater performance degradation using fuzzy inference system Expert Syst. Appl. 69 pp 239–46.
[7] Effendii H 2016 River water quality preliminary rapid assessment using pollution index Procedia Environ. Sci. 33 pp 562–7.
[8] Alavi N 2012 Using Mamdani fuzzy inference system to quality determination of Mozafati dates J. Saudi Soc. Agric. Sci. 12(2) pp 137–42.
[9] Guillaume S and Charnomordic B 2012 Fuzzy inference systems: An integrated modeling environment for collaboration between expert knowledge and data using FisPro Expert Syst. Appl. 39(10) pp 8744–55.