Prediction of O_3 Concentration Level Using Fuzzy Non-Stationary Method

Affi Nizar Suksmawati[1], Retantyo Wardoyo[2*]

[1]Master Program of Computer Science, Gadjah Mada University, Yogyakarta, Indonesia
[2]Department of Computer Science and Electronics, Gadjah Mada University, Yogyakarta, Indonesia
[1]affinizar@mail.ugm.ac.id, [2*]corr_author: rw@ugm.ac.id

Abstract - The composition of air concentration is not constant. It constantly changes with minor changes at any time, so more than one measurement is needed to represent the air concentration level for a full day. The fuzzy non-stationary method can overcome uncertainty in an environment that is not constant or caused by minor temporal changes based on time variables. This study uses a non-stationary fuzzy method to determine the level of O_3 concentration based on the input variables of temperature, humidity, and wind speed. The tests were conducted in September, October, and November using four types of implication process interpretation, namely interpretation 1 (classical logic), interpretation 2 (classical logic), interpretation 3 (algebraic), and interpretation 3 (standard). The test results in September showed a tendency for error percentage using the MAPE amount of 19, October's amount of 25, and November's amount of 18.

Keywords: Fuzzy logic; fuzzy non-stationary; interpretation implication fuzzy; air concentration prediction; air pollution

I. INTRODUCTION

Air is a combination of gases in the Earth's atmosphere. The air has never been found to be clean without pollutants. Some gases, such as sulfur dioxide (SO_2), hydrogen sulfide (H_2S), and carbon monoxide (CO), are released into the air as a result of natural processes. These natural processes include volcanic activity, forest fires, etc. However, the contribution of pollutants to the atmosphere as a result of natural processes is relatively small and insignificant compared to the contribution of pollutants from human activities [1]. The problem caused by pollutants from human activities is that their distribution is uneven, so they are concentrated in specific locations.

The environmental service measured air concentration levels at 26 locations [2]. The air parameters measured consisted of SO_2, NO_2, and O_3. These air parameters are primarily concentrated in residential areas due to human activities. Human activities such as photocopiers, refrigerators, or air conditioners (AC) have increased O_3 pollution. Acute exposure to O_3 can irritate the nose and throat [3]. In addition, exposure to O_3 at concentrations of 1.0 to 3.0 ppm can cause headaches and loss of coordination in some sensitive people [4].

The environmental services have conducted measurements for 1 hour for the O_3 parameter. However, the measurement for 1 hour is still significantly less to represent the level of O_3 concentration in a full day (24 hours). The composition of the O_3 concentration changes over time, so it takes more than one measurement to represent measurements in one full day. Changes in the composition of the O_3 concentration are also influenced by meteorological factors such as temperature, humidity, and wind speed [5]. However, there is uncertainty in determining how much influence meteorological factors have on the level of O_3 concentration [6]. This uncertainty can be modelled using fuzzy logic [7].

Fuzzy logic can model differences in perceptions of the influence of meteorological factors on the level of concentration or air quality in the form of linguistic variables [8]. Research to determine air quality was carried out by [9][10]. Research [9] used NO_2, SO_2, PM, and CO as inputs and then represented them in the form of good, moderate, and poor linguistic variables. Research [11] used a fuzzy system to determine the carbon monoxide (CO) concentration level. In this study, a comparative analysis of type-1 and type-2 fuzzy systems was carried out. The use of fuzzy type-2 can handle many uncertainties to provide more accurate predictions. These studies can produce output in the form of concentration levels or air quality. However, the resulting output is limited to a constant or static environment.

Fuzzy logic will produce the same output for the same input, so in determining the level of O_3 concentration, it takes many measurements of input variables to have output variations. That will make it difficult for observers who must continuously carry out measurements. The non-stationary fuzzy method is proposed to be able to adapt to an environment that is not constant or non-stationary [12]. Fuzzy non-stationary can produce several variations of output for the same input [13]. The output variation produced by non-stationary fuzzy has a minor difference. That is in accordance with the nature of the composition of the air concentration, which is constantly changing with minor changes.

Research using non-stationary fuzzy is widely applied in the health field. Study [14] used fuzzy non-stationary to model expert variability in determining breast cancer treatment. Research [15] used fuzzy non-stationary to model variations of experts in diagnosing coronary heart disease. These studies can produce output variations of several defined time variables. The result of output variations can handle many uncertainties and increase the accuracy of the diagnosis. This study uses fuzzy non-stationary in the environmental field to predict the level of O_3 concentration. Meteorological factors used to consist of temperature, humidity, and wind speed. This study compares four interpretations of the implication process, consisting of interpretation 1 (classical logic), interpretation 2 (classical logic), interpretation 3 (algebraic), and interpretation 3 (standard).

II. METHOD

The research method used can be seen in Fig. 1. Based on Fig. 1, the initial stage is data preprocessing. The O_3 concentration data were obtained from the AQMS (air quality monitoring system) of the Yogyakarta city environmental service. AQMS data consists of meteorological data and air concentration level data. Data preprocessing is carried out by selecting features from meteorological data as input variables and selecting features from air concentration level data as output variables. The data cleaning process is carried out on data noise or empty values. The results of the preprocessing stage are meteorological data of temperature, humidity, and wind speed as input variables and O_3 concentration level data as output variables. In addition, meteorological and O_3 concentration level data in September, October, and November were used as test data.

AQMS data is reported in a daily period (for 24 hours) every 30 minutes. In the exploratory data analysis (EDA) stage, the O_3 data visualization process will be carried out to gain insight into the most optimal observation time. The result of this stage is an insight into the most optimal observation time starting at 11:00 to 14:00. Then proceed to the membership function design stage. At this stage, the membership function is determined for the input and output variables. Each membership function has a linguistic variable, a set of linguistic values, and a fuzzy set domain.

The next stage is the fuzzy rule base design. The fuzzy rule base was formed based on insights from the EDA process and the results of interviews with air concentration experts. Next is the non-stationary fuzzy parameter design stage. At this stage, two non-stationary fuzzy parameters are determined, which consist of variations of non-stationary form and perturbation function. Variations in non-stationary forms using location variations can be written using (1) [16].

$$\forall_{t \in T}\ \mu_A(x,t) = \mu_A\big(x + c(t)\big) \tag{1}$$

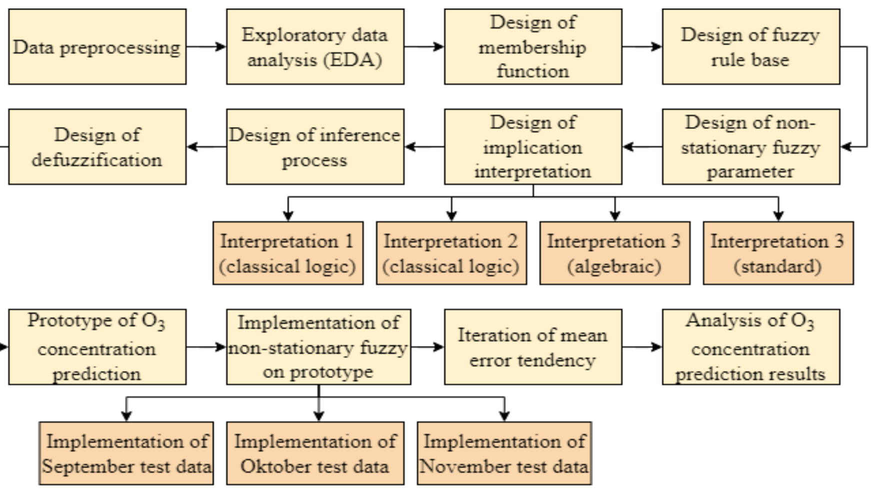

Fig. 1 Flowchart of research methodology

Based on (1), x is the parameter point of the membership function, while c(t) is the result of multiplying the constant with the perturbation function. The perturbation function used for the input variable is a uniformly distributed pseudo-random function using the LCG (linear congruential generator) method to generate random numbers. The input variable perturbation function can be written using (2) [17].

$$f(t) = \frac{\left((randomNums_{[i-1]} * a) + c\right)\%m}{m} \quad (2)$$

The variable $[i-1]$ in (2) is the sequence of pseudo-random numbers, a is the multiplier, c is the increment, while m is the modulus. The values of a, c, and m are determined by considering the randomness of the resulting numbers. There is a requirement for selecting the constant LCG method [18]. Namely the requirement for the modulus value is $(0 < m)$. In this study, the number 17 is used. The requirement for the constant α is $(0 < \alpha < m)$. In this study, the number 12 is used. The requirement for the constant c is $(0 < c < m)$. The number 6 is used in this study, and the requirement $randomNums_{[0]}$ is $\left(0 < randomNums_{[0]} < m\right)$ in this study used random integer values with an interval of [2, 10]. The perturbation function used in the output variable, namely the Sinusoidal function, can be written using (3) [19].

$$f(t) = \sin(\omega t) \quad (3)$$

Based on (3), the value ω is set at 127 while the value of t is a standard uniform random number with an interval of (0,1) [14]. The next stage is the design of the interpretation of the implication process. The implication process uses four interpretation forms, as seen in Table 1 [20].

Table I shows the interpretation of 1, 2, and 3 (algebraic) fuzzy operators used for t-norm, s-norm, and c-norm, respectively, algebraic product, algebraic sum, and standard complement. Interpretation 3 (standard) uses standard operators with t-norm (minimum) and s-norm (maximum). The next stage is the design of the fuzzy inference process. The fuzzy inference process uses GMP (generalized modus ponens), which can be written using (4).

$$\mu_{B'}(y) = \max_{x \in U} t\left(\mu_{A'}(x), \mu_{FR}(x, y)\right) \quad (4)$$

The $\mu_{A'}(x)$ in (4) is a fact while $\mu_{FR}(x, y)$ is a fuzzy rule resulting from the fuzzy implication process. The inference system used is individual based with a combination of Mamdani. The Mamdani combination uses the s-norm operator, which can be written using (5).

$$\mu_{B'}(y) = s\left(\mu_{B'1}(y), \mu_{B'2}(y), \mu_{B'3}(y), \mu_{B'4}(y), \mu_{B'5}(y)\right) \quad (5)$$

Based on (5), s is the s-norm operator, while B' is the result of the conclusion using GMP. The results of the combination process using the Mamdani combination are used for defuzzification. At the defuzzification design stage, the center average method is used, which can be written using (6) [21].

$$y^* = \frac{\sum_{i=1}^{m} y^{-i} w_i}{\sum_{i=1}^{m} w_i} \quad (6)$$

The y-i in (6) is the center of the i-fuzzy set, while w_i is the height (degree of membership) of the i-fuzzy set. After components have been defined, the next stage is the process of making a prototype prediction of O_3 concentration levels. The non-stationary fuzzy system components described are then implemented on the prototype to predict the level of O_3 concentration. The non-stationary fuzzy prototype that has been built is used for testing the prediction of O_3 concentration levels in September, October, and November. The implementation of the non-stationary fuzzy method can be seen in the flowchart in Fig. 2.

TABLE I
INTERPRETATION OF THE IMPLICATION PROCESS

Type	Interpretation of Implication	Operator	Operator Type
Interpretation 1	$\bar{P} \vee Q$	t-norm s-norm c-norm	Algebraic product Algebraic sum Standard complement
Interpretation 2	$\bar{P} \vee (P\&Q)$	t-norm s-norm c-norm	Algebraic product Algebraic sum Standard complement
Interpretation 3 (Algebraic)	$P\&Q$	t-norm s-norm	Algebraic product Algebraic sum
Interpretation 3 (Standard)	$P\&Q$	t-norm s-norm	Standard (minimum) Standard (maximum)

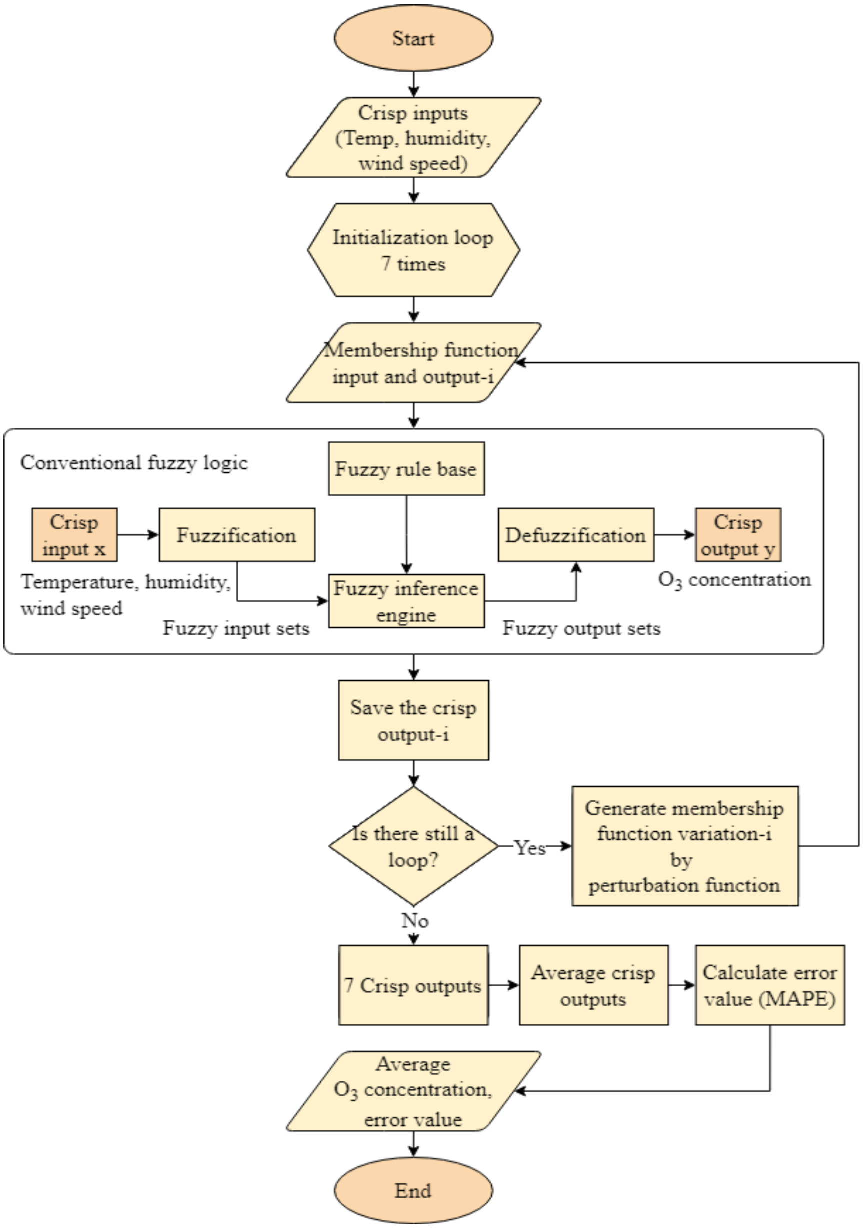

Fig. 2 Non-stationary fuzzy flowchart

Based on Fig. 2, intuitively, the non-stationary fuzzy is an iteration of the conventional fuzzy logic inference system (type-1). The non-stationary fuzzy flowchart begins by defining the input variable values for temperature, humidity, and wind speed. Then determine the number of iterations that represent changes in the time variable. The time variable used is the optimal observation time of the O_3 concentration level from 11:00 to 14:00, reporting every 30 minutes. Based on the observation time interval, there will be seven times reporting the level of O_3 concentration, so that the number of repetitions defined is seven times.

In the first iteration, the basic membership function that has been defined is used for the crisp input fuzzification process using a conventional fuzzy logic inference system. The crisp output using the basic membership function will be stored as the first prediction result. In the second iteration, the basic membership function will be shifted by the perturbation function, resulting in a variation of the membership function. The membership function variation of the second iteration is used for the crisp input fuzzification process using a conventional fuzzy logic inference system. The crisp output results using the membership function variation of the second iteration are stored as the second prediction result. The same process will be iterated for the third to seventh iterations.

After the last iteration, there will be seven crisp outputs of O_3 concentration levels. The seven crisp outputs will be calculated on an average as the average value for observing O_3 concentration levels in one day (24 hours). The average results of prediction observations with non-stationary fuzzy will be calculated for the error value against the average results of the actual observations. The calculation of the error value is done using MAPE (mean absolute percentage error), which can be written using (7) [22].

$$Mape = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A_i - F_i}{A_i}\right| \tag{7}$$

Based on (7), Ai is the average result of actual observations, while Fi is the average result of prediction observations with non-stationary fuzzy. The resulting error values will be stored for a single instance. Each instance's error value in each testing month will be calculated as the error percentage. The error percentage calculation is done by adding all the error values for each instance and then dividing by the total number of instances in that month.

The next stage is an iterative process to determine the error percentage tendency. The non-stationary fuzzy system is dynamic, so there will be a minor difference in the error percentage results in the month for each iteration. At this stage, 25 iterations are carried out to see the tendency of the error percentage over the month. Then proceed with analyzing the tendency of the error percentage results for testing in September, October, and November.

III. RESULTS AND DISCUSSION

This research uses a constant value to be multiplied by the perturbation function of the input and output variables. Referring to the (1), the determination of the constant value affects the shift value of the membership function. Based on this, the constant value is determined based on the experimental results, with the limitation of keeping the membership function shift value under a minor condition. The results of the determination of the constant values for the input and output variables can be seen in Table II.

Based on Table II, the constant value set will be multiplied by the perturbation function of the input and output variables. The results of modelling input and output linguistic variables consisting of a set of linguistic values, domains, and units can be seen in Table III.

As shown in Table III, the linguistic variable of temperature has a universe of discourse (UoD) [15, 40]. The linguistic variable humidity has a UoD [30, 95]. The linguistic variable of wind speed has UoD [0, 12], while the linguistic variable O_3 has UoD [40, 75]. The results of the formation of a fuzzy rule base based on insights from the EDA process and the results of interviews with air concentration experts can be seen in Table IV.

TABLE II
CONSTANT VALUES

Type	Variable name	Constant
Input variable	Temperature	1
	Humidity	5
	Wind speed	1
Output variable	O_3	3

TABLE III
INPUT AND OUTPUT LINGUISTIC VARIABLES

Linguistic variable	Set of linguistic value	Domain	Unit
Temperature	Low	[15, 24]	
	Medium	[22, 30]	°C
	High	[28, 40]	
Humidity	Dry	[30, 65]	%RH
	Moist	[55, 95]	
Wind speed	Calm	[0, 3]	
	Light	[2, 5]	m/s
	Gentle	[4, 12]	
O_3	Low	[40, 50]	
	Medium	[45, 65]	$\mu g/m^3$
	High	[60, 75]	

TABLE IV
FUZZY RULE BASE

Temperature	Humidity	Wind Speed	O_3
High	Dry	Calm	High
High	Dry	Light	Low
High	Dry	Gentle	High
High	Moist	Calm	Medium
High	Moist	Gentle	Low
Medium	Dry	Calm	Medium
Medium	Moist	Calm or Light or Gentle	Low

Based on Table IV, the level of O_3 concentration is strongly influenced by temperature. The correlation between temperature and humidity is inversely proportional, while the correlation between temperature and the level of O_3 concentration is directly proportional. The level of O_3 concentration is also influenced by wind speed, although the correlation is not so strong. The results of one instance test for data on September 14, 2020, with a temperature concentration of 30.4, humidity of 57, and wind speed of 1, can be seen in Table V.

The time column in Table V represents the number of observations made in one day, which is seven times. Columns "O_3 INT1", "O_3 INT2", "O_3 INT3 (Algebraic)", and "O_3 INT3 (Standard)" are the results of predicting the level of O_3 concentration using process interpretation of implications 1 (classical logic), interpretation 2 (classical logic), interpretation 3 (algebraic), and interpretation 3 (standard). Then the average calculation is carried out on the actual O_3 concentration level and the predicted results using a non-stationary fuzzy system. The average calculation aims to represent the level of O_3 concentration in one day. The results of the calculation of the average O_3 concentration level for one instance of test can be seen in Table VI.

As in Table VI, the column "Actual Concentration of O_3" is the result of the actual average level of O_3 concentration, while the column "Avg" is the result of the average prediction of the level of O_3 concentration using a non-stationary fuzzy system. The average result of the actual O_3 concentration level and the predicted average result is calculated using MAPE. The results of the error calculation for each fuzzy interpretation can be seen in the "MAPE" column. In Table VI, the result with the lowest error value is predicted using interpretation 2 (classical logic), which is 1.892. The results of modelling the membership curve of the humidity input variable and the O_3 output variable in one instance test can be seen in Fig. 3 and Fig. 4.

Fig. 3 shows the basic membership curve of the humidity input variable before and after the shift using the perturbation function, a uniformly distributed pseudo-random function. Fig. 4 shows the basic membership curve of the O_3 output variable before and after the shift using the perturbation function sinusoidal function. The test results of 30 instances in September can be seen in Table VII.

Based on Table VII, the column "Actual Average O_3" is the result of the average level of O_3 concentration in the actual data, while the columns "O_3 INT1", "O_3 INT2", "O_3 INT3 (Algebraic)", and "O_3 INT3 (Standard)" are the results of predicting the average level of O_3 concentration using a non-stationary fuzzy system. The results of the error percentage for testing 30 instances in September can be seen in Table VIII.

As presented in Table VIII, the result of the percentage error using interpretations 3 (algebraic) and (standard) is higher than the result of the percentage error using interpretations 1 and 2 (classical logic). This is because there are several data on O_3 concentration levels in September that are not in match with the defined fuzzy rule base. For example, there is a fuzzy rule that states the level of O_3 concentration is high, but in September, the data is concentrated low.

TABLE V
ONE INSTANCE TEST RESULT

Time	Actual Concentration of O_3	Non-Stationary Fuzzy Prediction Results			
		O_3 INT1	O_3 INT2	O_3 INT3 (Algebraic)	O_3 INT3 (Standard)
11:00	62	61,207	60,771	67,456	64,692
11:30	91	66,721	67,289	67,872	69,705
12:00	69	62,463	60,407	64,811	64,656
12:30	57	60,156	63,330	66,294	67,594
13:00	55	59,765	60,656	67,768	68,039
13:30	57	62,426	60,539	65,533	66,16
14:00	60	67,992	69,467	67,716	67,948

TABLE VI
ONE INSTANCE TEST AVERAGE RESULTS

Actual Concentration of O_3	O_3 INT1		O_3 INT2		O_3 INT3 (AL)		O_3 INT3 (STD)	
	Avg	MAPE	Avg	MAPE	Avg	MAPE	Avg	MAPE
64,428	62,961	2,275	63,208	1,892	66,778	3,649	66,970	3,946

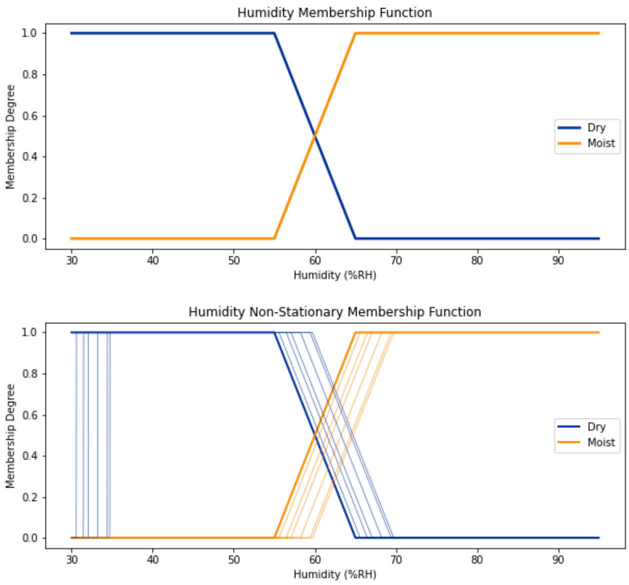

Fig. 3 Membership function of humidity variable

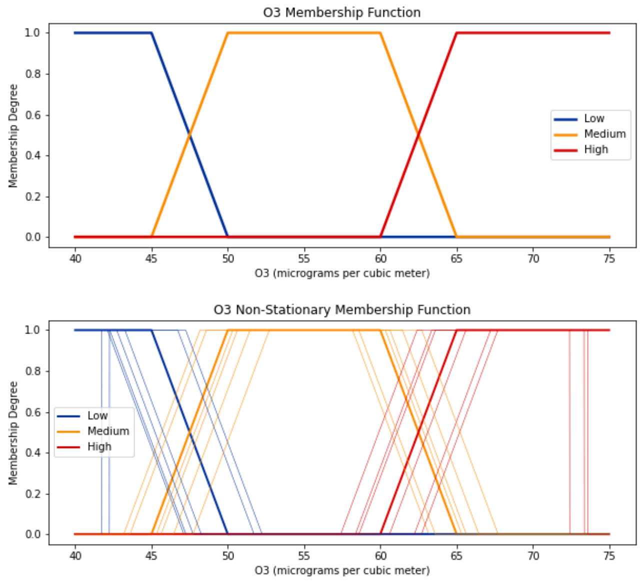

Fig. 4 Membership function of O_3 concentration level

TABLE VII
O_3 PREDICTION RESULTS FOR SEPTEMBER

Date	Actual Average O_3	Non-Stationary Fuzzy Average Result			
		O_3 INT1	O_3 INT2	O_3 INT3 (Algebraic)	O_3 INT3 (Standard)
1	62,286	60,507	59,617	65,933	63,582
2	63	57,15	50,097	45,532	47,441
3	54,143	57,176	57,141	55,587	57,492
4	53	56,316	56,113	54,736	55,831
5	57,429	56,611	55,755	54,77	55,271
6	47	55,404	55,936	47,008	50,768
7	46,857	57,498	56,981	59,19	58,353
8	68,429	65,642	66,9	67,52	67,306
9	96	56,262	57,931	56,358	55,752
10	57	58,284	57,75	58,387	60,174
11	51,857	61,764	60,334	65,679	62,006
12	114,714	59,951	59,943	66,021	63,593
13	56,286	57,06	56,619	56,981	56,919
14	64,429	63,485	60,973	66,232	64,538
15	61,429	66,859	67,888	68,219	67,529
16	62,571	60,72	60,212	65,38	65,606
17	103,286	57,216	56,429	58,597	56,431
18	98,571	68,195	67,168	68,446	68,284
19	90,429	67,991	67,983	68,366	67,9
20	77,143	68,958	68,298	68,54	68,387
21	47,857	56,374	57,044	57,795	58,249
22	55	57,215	44,247	44,309	44,409
23	29,714	55,563	54,941	56,161	55,181
24	44	51,921	50,536	45,907	47,4
25	57	46,48	47,006	44,129	44,283
26	48,857	62,969	59,953	65,932	64,868
27	40,571	57,593	57,046	58,338	57,842
28	46,571	65,258	63,976	67,548	65,399
29	48,286	55,478	55,658	44,81	53,872
30	56,714	58,387	57,858	61,063	60,888

The predictions using interpretations 3 (algebraic) and (standard) tend to conform to the defined fuzzy rule base. However, this causes many prediction errors to be found. September data tends to be low with low or medium concentrations of O_3. The results of the prediction of the average O_3 concentration level using interpretations 1 and 2 (classical logic) tend to produce lower prediction values than interpretations 3 (algebraic) and (standard) so that it can overcome some of the noise data for September.

The non-stationary fuzzy method is dynamic and gives different results each time the process runs. The prediction results of the O_3 concentration level in the first and second runs for each instance have a minor difference. Based on this, the iteration process is carried out 25 times, running on all test data instances. An iteration of 25 runs aims to get the error percentage tendency in the month. The results of the error percentage tendency for the September test can be seen in Fig. 5.

The horizontal axis in Fig. 5 is the number of iterations running 25 times, while the vertical axis is the error percentage. The results of running from 1 to 25 using interpretations 1 and 2 (classical logic) tend to get an error percentage result of 19, with the lowest error percentage using interpretation 1 (classical logic). Using interpretations of interpretation 3 (algebraic) and (standard) tends to get a percentage error of 20. The test results of 29 instances in October can be seen in Table IX.

Based on Table IX, data for October tends to be high, with the O_3 concentration being mostly concentrated in medium and high. The result of the error percentage for testing 29 instances in October can be seen in Table X. The results of the percentage error in October tend to be high. This is because, in the October data, there are several data on O_3 concentration levels that are not in accordance with the defined fuzzy rule base. For example, there is a fuzzy rule that states the level of O_3 concentration is medium, but in October data, it is concentrated at low or high. Based on data from October with a high level of O_3 concentration, prediction results using the interpretation of 3 (algebraic) and (standard) tend to produce a low level of O_3 concentration. Prediction results using interpretations 1 and 2 (classical logic) tend to be higher and can overcome some of the noise data for October. The results of the error percentage tendency for the October test can be seen in Fig. 6. In that figure, the results of running from 1 to 25 using interpretation 1 tend to get a percentage error of 26, while using interpretation 2 tend to get a percentage error of 25. The results of the percentage error using interpretation 3 (algebraic) tend to get a percentage error result of 28, while using interpretation 3 (standard) tends to get a percentage error of 27. The test results of 30 instances in November can be seen in Table XI.

Data for November as represented in Table XI tends to be high, with the O_3 concentration level being mostly concentrated in medium and high. The results of the error percentage for testing 30 instances in November can be seen in Table XII.

The results of the percentage of errors in Table XII using interpretation 2 (classical logic), interpretation 3 (algebraic) and (standard) tend to produce a higher percentage of error values than the results of the percentage of errors using interpretation 1 (classical logic). This is because the data on the level of O_3 concentration in October tends to be high. Prediction results using interpretation 2 (classical logic), interpretation 3 (algebraic) and (standard) tend to produce smaller values than prediction results using interpretation 1 (classical logic). Thus, the prediction results using interpretation 1 (classical logic) can better overcome some of the November noise data. The results of the error percentage tendency for the November test can be seen in Fig. 7.

TABLE VIII
ERROR PERCENTAGE RESULTS FOR SEPTEMBER

O_3 INT1	**O_3 INT2**	**O_3 INT3 (AL)**	**O_3 INT3 (STD)**
19,338	19,826	20,280	20,257

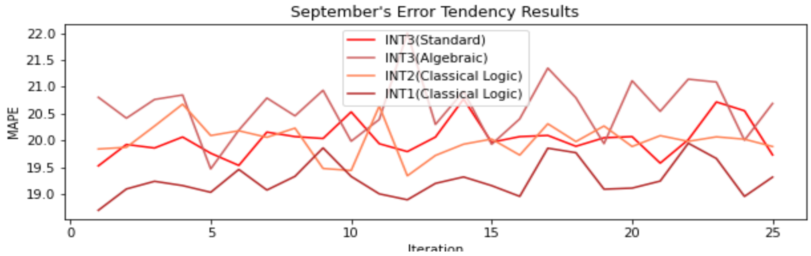

Fig. 5 Error percentage tendency results for September

TABLE IX
O_3 PREDICTION RESULTS FOR OCTOBER

Date	Actual Average O_3	Non-Stationary Fuzzy Average Result			
		O_3 INT1	O_3 INT2	O_3 INT 3 (Algebraic)	O_3 INT1
1	82,143	66,415	60,988	67,528	64,886
2	67,429	68,462	67,471	68,447	68,925
3	47,000	55,65	55,533	53,771	54,19
4	51,429	54,723	54,72	52,32	55,295
5	48,000	57,163	50,904	44,374	43,862
6	59,286	57,039	55,529	43,99	48,352
7	81,000	56,856	56,392	52,513	54,463
8	74,286	58,1	57,402	59,234	60,764
9	94,286	57,664	57,011	57,405	59,257
10	55,286	55,599	55,325	56,226	55,025
11	54,429	68,273	67,256	68,468	68,123
12	75,857	55,571	55,136	54,671	54,718
13	49,714	55,923	55,164	53,521	54,378
14	65,857	55,996	56,442	50,439	53,213
15	58,000	53,599	51,889	47,288	50,21
16	56,286	54,236	55,989	54,25	53,738
17	23,429	56,995	56,911	57,103	58,748
18	105,571	52,22	53,981	46,692	49,903
19	59,286	43,08	44,92	43,405	43,446
20	61,571	56,885	56,258	54,219	53,708
21	75,286	54,128	54,578	54,784	55,825
22	57,857	54,219	55,429	54,505	56,093
23	62,429	65,314	65,484	68,262	65,839
24	53,571	56,526	55,421	55,237	56,139
25	44,286	56,26	55,826	53,494	54,759
26	60,714	57,137	55,685	44,168	44,511
27	28,286	61,314	58,691	64,251	60,631
28	35,286	60,994	60,12	63,786	63,325
29	72,857	55,378	55,248	55,322	55,283

TABLE X
ERROR PERCENTAGE RESULTS FOR OCTOBER

O_3 INT1	O_3 INT2	O_3 INT3 (AL)	O_3 INT3 (STD)
26,554	25,647	28,816	27,988

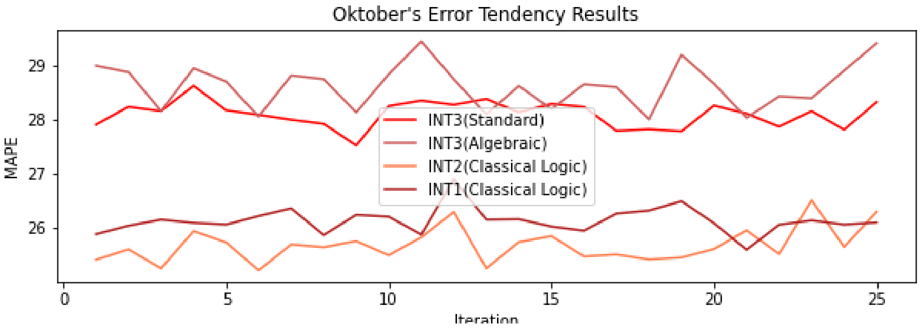

Fig. 6 Error percentage tendency results for October

TABLE XI
O_3 PREDICTION RESULTS FOR NOVEMBER

Date	Actual Average O_3	Non-Stationary Fuzzy Average Result			
		O_3 INT1	O_3 INT2	O_3 INT3 (Algebraic)	O_3 INT1
1	69,143	56,118	56,007	54,956	56,221
2	36,286	54,009	55,465	56,181	55,663
3	79,429	59,345	59,336	59,001	61,287
4	74,429	63,793	63,697	66,938	65,711
5	60,143	61,07	60,317	63,143	61,97
6	51,000	56,072	56,836	55,682	56,881
7	56,286	53,059	53,74	47,111	48,667
8	56,857	68,05	67,903	68,912	68,092
9	40,000	51,812	51,304	45,807	50,91
10	59,143	59,332	57,978	63,813	61,693
11	61,571	55,971	56,451	55,136	54,733
12	47,143	50,167	50,391	44,512	44,227
13	58,143	68,439	68,686	67,886	67,918
14	84,143	55,918	56,136	55,614	54,907
15	69,571	55,827	56,154	52,834	54,43
16	71,857	53,961	52,751	50,439	51,586
17	75,571	54,95	55,536	56,736	54,043
18	58,000	54,447	56,744	55,887	55,685
19	81,571	56,693	56,372	54,313	56,145
20	63,143	54,771	54,371	55,064	55,379
21	62,714	55,382	55,906	55,88	54,512
22	43,571	56,373	56,68	52,594	53,988
23	90,143	56,427	56,23	55,514	55,613
24	58,429	57,166	51,092	47,003	47,118
25	84,429	57,211	43,233	43,918	43,286
26	63,714	56,015	55,04	56,757	58,353
27	72,429	56,366	55,698	55,341	54,686
28	65,714	57,516	57,234	57,28	59,572
29	52,000	57,233	44,54	43,834	44,366
30	67,143	56,161	56,312	55,684	54,818

TABLE XII
ERROR PERCENTAGE RESULTS FOR NOVEMBER

O_3 INT1	O_3 INT2	O_3 INT3 (AL)	O_3 INT3 (STD)
18,288	19,448	19,941	19,985

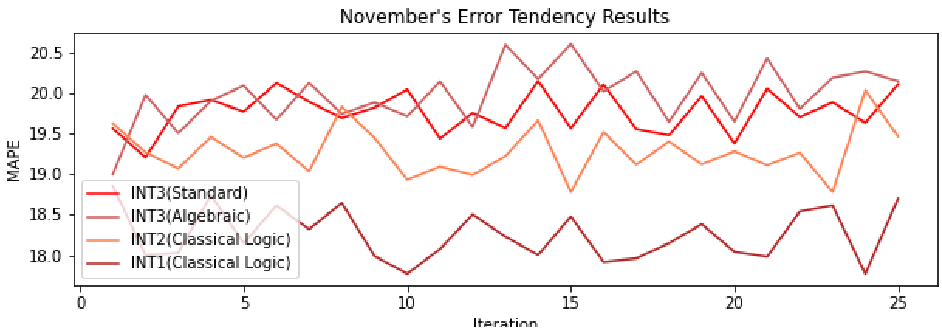

Fig. 7 Error percentage tendency results for November

Based on Fig.7, the results of running from 1 to 25 using interpretation 1 (classical logic) tend to get a percentage error of 18, while using interpretation 2 (classical logic), interpretation 3 (algebraic) and interpretation 4 (standard) tend to get an error percentage result of 19. Based on the graph of the error percentage tendency in September, October, and November, the lowest error percentage value for September uses interpretation 1 (classical logic), October uses interpretation 2 (classical logic), and November uses interpretation 1 (classical logic). These results indicate that the use of interpretations 1 and 2 can overcome some noisy data, while the use of interpretations 3 (algebraic) and (standard) is more suitable to be implemented on data with inputs and outputs that match the input and output conditions defined in the linguistic set and fuzzy rule base.

IV. CONCLUSION

The non-stationary fuzzy method can overcome the problem of uncertainty in the composition of O_3 concentration, which has minor changes over time, by forming a fuzzy set for each linguistic variable and shifting the membership function in the fuzzy set using a perturbation function. Fuzzy non-stationary is dynamic, producing seven different crisp outputs on each running system. Tests in September obtained a tendency of error percentage results using interpretations 1 and 2 (classical logic) of 19. Tests in October got a tendency of average error results using interpretation 2 (classical logic) of 25, while the November test obtained a tendency of average results with an average error using interpretation 1 (classical logic) of 18. At the data preprocessing stage for further research, it can filter out noise or outlier data so that it does not affect the prediction accuracy results too much. Further research can also be done by adding input variables that affect the prediction of O_3 concentration levels, such as solar radiation.

ACKNOWLEDGEMENT

This research is supported by the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia by the Master's Thesis Grant (Grant No. 1960/UN1/DITLIT/Dit-Lit/PT.01.03/2022). The authors would like to thank the Yogyakarta City Environmental Service (DLH), which has supported research data collection.

REFERENCES

[1] S. Fardiaz, *Polusi Air & Udara*. Yogyakarta: Penerbit PT Kanisius, 1992.

[2] DLH, *Laporan Analisa Hasil Pemantauan Kualitas Udara Kota Yogyakarta Tahun 2021*. 2021.

[3] K. Ezimand and A. A. Kakroodi, “Prediction and spatio – Temporal analysis of ozone concentration in a metropolitan area,” *Ecol. Indic.*, vol. 103, no. April, pp. 589–598, 2019, doi: 10.1016/j.ecolind.2019.04.059.

[4] M. Ando, V. Biju, and Y. Shigeri, “Development of technologies for sensing ozone in ambient air,” *Anal. Sci.*, vol. 34, no. 3, pp. 263–271, 2018, doi: 10.2116/analsci.34.263.

[5] M. Ebrahimi and F. Qaderi, “Determination of the most effective control methods of SO2 Pollution in Tehran based on adaptive neuro-fuzzy inference system,” *Chemosphere*, vol. 263, p. 128002, Jan. 2021, doi: 10.1016/j.chemosphere.2020.128002.

[6] N. Ritha and R. Wardoyo, “Implementasi Neural Fuzzy Inference System dan Algoritma Pelatihan Levenberg-Marquardt untuk Prediksi Curah Hujan,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 10, no. 2, p. 125, 2016, doi: 10.22146/ijccs.15532.

[7] S. Lestari and M. R. Romahdoni, “Decision Support System for Service Quality Using SMART and Fuzzy ServQual Methods,” *JUITA J. Inform.*, vol. 9, no. 1, p. 83, 2021, doi: 10.30595/juita.v9i1.9644.

[8] M. T. Utomo and H. Mustafidah, “Penentuan Spesifikasi Komputer Berdasarkan Kebutuhan Pemakai Dan Harga Menggunakan Basis Data Fuzzy,” *Juita*, vol. 1, pp. 28–36, 2016.

[9] R. Riyaz and P. V. Pushpa, “Air Quality Prediction in Smart Cities: A Fuzzy-logic based Approach,” *Proc. Int. Conf. Comput. Tech. Electron. Mech. Syst. CTEMS 2018*, pp. 172–178, 2018, doi: 10.1109/CTEMS.2018.8769195.

[10] A. T. Teologo, E. P. Dadios, R. Q. Neyra, and I. M. Javel, “Air Quality Index (AQI) Classification using CO and NO2 Pollutants: A Fuzzy-based Approach,” *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, vol. 2018-Octob, no. 2, pp. 194–198, 2019, doi: 10.1109/TENCON.2018.8650344.

[11] A. Dass, S. Srivastava, and G. Chaudhary, “Air pollution: A review and analysis using fuzzy techniques in Indian scenario,” *Environ. Technol. Innov.*, vol. 22, p. 101441, 2021, doi: 10.1016/j.eti.2021.101441.

[12] J. M. Garibaldi, M. Jaroszewski, and S. Musikasuwan, “Nonstationary Fuzzy Sets,” *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 4, pp. 1072–1086, Aug. 2008, doi: 10.1109/TFUZZ.2008.917308.

[13] S. Musikasuwan, “Novel Fuzzy Techniques for Modelling Human Decision Making,” University of Nottingham, 2013.

[14] J. M. Garibaldi, S.-M. Zhou, X.-Y. Wang, R. I. John, and I. O. Ellis, “Incorporation of expert variability into breast cancer treatment recommendation in designing clinical protocol guided fuzzy rule system models,” *J. Biomed. Inform.*, vol. 45, no. 3, pp. 447–459, Jun. 2012, doi:

10.1016/j.jbi.2011.12.007.

[15] V. D. V. Latumeten, “DIAGNOSA STATUS RESIKO JANTUNG KORONER MENGGUNAKAN METODE FUZZY NON STATIONARY,” *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 11, no. 1, pp. 27–38, Mar. 2017, doi: 10.30598/barekengvol11iss1pp27-38.

[16] S. Coupland and S. G. Matthews, “Using nonstationary fuzzy sets to improve the tractability of fuzzy association rules,” in *2013 IEEE Symposium on Advances in Type-2 Fuzzy Logic Systems (T2FUZZ)*, Apr. 2013, pp. 9–14, doi: 10.1109/T2FZZ.2013.6613293.

[17] O. S. Sitompul, Z. Situmorang, F. R. Naibaho, and E. B. Nababan, “Steganography with Highly Random Linear Congruential Generator for Security Enhancement,” in *2018 Third International Conference on Informatics and Computing (ICIC)*, Oct. 2018, pp. 1–6, doi: 10.1109/IAC.2018.8780445.

[18] H. E. Putra and K. Harianto, “Implementasi Linear Congruential Generator untuk Pengacakan Gambar Pada Permainan Puzzle,” *SATIN - Sains dan Teknol. Inf.*, vol. 4, no. 1, pp. 89–96, Jun. 2018, doi: 10.33372/stn.v4i1.302.

[19] S. Musikasuwan and J. Garibaldi, “Exploring gaussian and triangular primary membership functions in non-stationary fuzzy sets,” *Proc. Inf. Process. …*, 2006, [Online]. Available: http://www.math.s.chiba-u.ac.jp/~yasuda/open2all/Paris06/IPMU2006/HTML/FINALPAPERS/P388.PDF.

[20] R. Wardoyo and W. D. Yuniarti, “Analysis of Fuzzy Logic Modification for Student Assessment in e-Learning,” *IJID (International J. Informatics Dev.*, vol. 9, no. 1, p. 29, Sep. 2020, doi: 10.14421/ijid.2020.09105.

[21] N. Hasanah and R. Wardoyo, “Purwarupa Sistem Pakar dengan Mamdani Product untuk Menentukan Menu Harian Penderita DM,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 7, no. 1, p. 45, 2013, doi: 10.22146/ijccs.3051.

[22] I. Yulian, D. S. Anggraeni, and Q. Aini, “Penerapan Metode Trend Moment Dalam Forecasting Penjualan Produk CV. Rabbani Asyisa,” *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 6, no. 2, pp. 193–200, 2020.