

## Analisis dan Peramalan Polusi Udara di Surabaya Menggunakan Metode *Fuzzy Time Series*

Wahananta Izha Mahendra<sup>1\*</sup>, Anggay Luri Pramana<sup>2</sup>

<sup>1,2</sup>Universitas Nahdlatul Ulama Sidoarjo, Indonesia

Jl. Lingkar Timur KM 5,5 Rangkah Kidul, Sidoarjo, Jawa Timur, Indonesia

\*Corresponding author e-mail : wahanantaizha@gmail.com

### ABSTRAK

Peramalan merupakan proses atau metode untuk menghasilkan estimasi atau prediksi mengenai peristiwa atau kondisi yang akan terjadi di masa depan berdasarkan informasi yang tersedia saat ini. Peramalan polusi udara ini dilakukan untuk menyikapi dampak negatif polusi udara membutuhkan pemahaman yang mendalam terhadap perubahan struktur kualitas udara di Surabaya, analisis dan peramalan polusi udara sangat penting untuk menjaga keseimbangan ekologi dan kesehatan masyarakat. Metode yang dilakukan dengan menggunakan *Fuzzy Time Series*. Uji coba dilakukan dengan dataset sebanyak 533 data dari tahun 2022 hingga pertengahan tahun 2023. Adapun hasil yang didapat dengan evaluasi *Mean Absolute Percentage Error* (MAPE) cukup baik, dengan hasil 18,71% mendapatkan interpretasi peramalan baik. Evaluasi akurasi peramalan juga menggunakan *Mean Squared Error* (MSE) dengan mendapat skor 188,32 dan nilai *Root Mean Squared Error* (RMSE) sebesar 13,72. Kesimpulan, program peramalan polusi udara ini dapat dikembangkan dan memberikan informasi terkait polusi udara di Surabaya di masa yang akan datang.

**Kata kunci** : peramalan, polusi udara, *Fuzzy Time Series*, evaluasi.

### ABSTRACT

*Forecasting is a process or method for generating estimates or predictions about events or conditions that will occur in the future based on information available today. Air pollution forecasting is done to address the negative impact of air pollution requires a deep understanding of the structural changes in air quality in Surabaya, air pollution analysis and forecasting is very important to maintain ecological balance and public health. The method is done by using Fuzzy Time Series. The trial was conducted with a dataset of 533 data from 2022 to mid-2023. The results obtained by evaluating the Mean Absolute Percentage Error (MAPE) is quite good, with the results of 18.71% get a good forecasting interpretation. Evaluation of forecasting accuracy also uses Mean Squared Error (MSE) with a score of 188.32 and Root Mean Squared Error (RMSE) of 13.72. In conclusion, this air pollution forecasting program can be developed and provide information related to air pollution in Surabaya in the future.*

**Keywords:** forecasting, air pollution, *Fuzzy Time Series*, evaluation

## I. PENDAHULUAN

Polusi udara merupakan salah satu masalah lingkungan utama di kota-kota besar, termasuk Surabaya. Karena Surabaya sebagai pusat perkotaan dan industri, peningkatan jumlah kendaraan, kegiatan industri, dan struktur konsumsi energi menyebabkan peningkatan emisi polutan ke atmosfer, yang dapat berdampak negatif terhadap kesehatan penduduk dan lingkungan.

. Dengan demikian, tujuan dari penelitian ini adalah untuk menerapkan pendekatan inovatif, yakni metode *Fuzzy Time Series*. *Fuzzy Time Series*

mengintegrasikan konsep *fuzzy logic* dengan metode analisis deret waktu, memungkinkan membuat model yang disesuaikan dengan kompleksitas dan ketidakpastian data deret waktu tentang polusi udara [1]. Memahami sifat perubahan dan faktor-faktor yang mempengaruhinya, direncanakan akan dibuat model peramalan yang akurat dan efektif yang akan memberikan informasi terkini tentang polusi udara di Surabaya.

Dalam prakteknya, peramalan digunakan untuk sejumlah tujuan yang beragam, termasuk perencanaan

strategis, manajemen rantai pasokan, pengambilan keputusan investasi, dan pengelolaan inventaris [2].

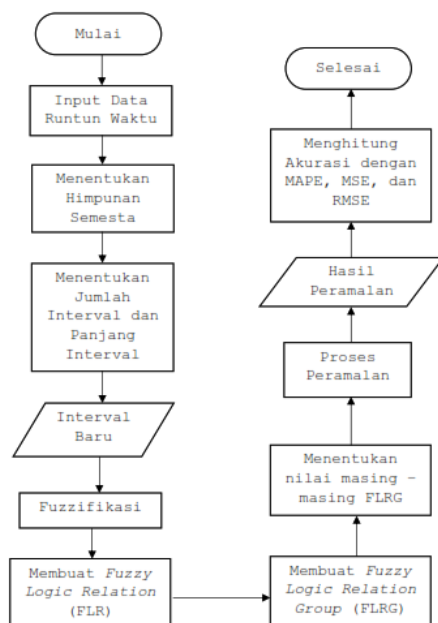
Walaupun tidak ada metode peramalan yang sempurna, upaya terus dilakukan untuk meningkatkan akurasi prediksi melalui penerapan teknologi yang lebih mutakhir, integrasi data yang lebih luas, dan pengembangan model yang lebih kompleks.

Landasan penelitian ini diperkuat oleh urgensi implementasi manajemen kualitas udara yang mengandalkan data dan prediksi yang presisi. Metode *Fuzzy Time Series*, sebagai pendekatan peramalan, unggul dalam mengatasi kompleksitas dan ketidakpastian dalam data deret waktu polusi udara. Oleh karena itu, penerapan metode ini diharapkan dapat menghasilkan pemahaman yang lebih mendalam terkait dinamika polusi udara, mendukung pengambilan keputusan yang lebih tepat oleh pihak berwenang, dan memberikan informasi berharga kepada masyarakat untuk mengurangi dampak negatifnya terhadap kesehatan [3].

Fuzzy time series adalah suatu teknik peramalan data yang merujuk pada prinsip fuzzy set sebagai dasar perhitungannya [4]. Sistem peramalan yang menggunakan metode ini beroperasi dengan menangkap pola dari data historis untuk kemudian mengaplikasikannya dalam proyeksi data di masa mendatang. Dibandingkan dengan metode lain seperti algoritma genetika dan jaringan syaraf yang kompleks, pendekatan ini memiliki proses yang lebih sederhana, menjadikannya mudah untuk diterapkan dan dikembangkan [5].

## II. METODE

Pada penelitian ini, diagram alir yang peneliti gunakan adalah sebagai berikut.



Gambar 1. Diagram Alir *Fuzzy Time Series Chen*

Dari diagram alir diatas didapat sebuah alur untuk menentukan nilai peramalan dengan metode *Fuzzy Time Series*. Langkah pertama yaitu mempersiapkan data runtun waktu [6]. Pada penelitian ini, peneliti menggunakan data runtun waktu dari open dataset Surabaya, yaitu Jumlah Indeks Standard Pencemar Udara (ISPU) yang terpublish dari 2 Stasiun Pemantau Kualitas Udara Ambien (SPKUA) Kota Surabaya dalam 18 bulan.

Tabel 1. Data Runtun Waktu Indeks Standard Pencemar Udara (ISPU) Kota Surabaya 18 Bulan

Waktu	Nilai ISPU Max
01/01/2022	67
02/01/2022	73
03/01/2022	71
...	...
28/06/2023	67
29/06/2023	64
30/06/2023	63

Langkah kedua membentuk himpunan semesta dengan lambang  $U$  dengan definisi  $U = [Z_{min} - F_1; Z_{max} + F_2]$ , dimana  $Z_{min}$  adalah nilai minimum semua data dan  $Z_{max}$  adalah nilai maksimum semua data. Nilai  $F_1$  dan  $F_2$  adalah nilai sembarang bilangan positif yang dapat ditentukan oleh peneliti.

Langkah ketiga membentuk panjang dan jumlah interval. Interval pertama menentukan dahulu jumlah interval dengan rumus :

$$Jumlah\ interval = 1 + 3,322 \log (n) \quad (1)$$

Dimana  $n$  adalah jumlah seluruh data. Setelah menemukan jumlah interval menghitung panjang interval dengan cara nilai maksimum dikurangi dengan nilai minimum lalu dibagi dengan jumlah interval, dapat ditulis dengan rumus :

$$Panjang\ interval = \frac{X_{max} - X_{min}}{jumlah\ interval} \quad (2)$$

Setelah dapat jumlah dan panjang interval, menghitung nilai tengah ( $m_1, m_2, m_3, \dots, m_n$ ) yang didapat dari nilai interval awal ditambah panjang interval dan dibagi 2.

Langkah keempat membentuk *fuzzy set* ( $A_i$ ) ditentukan sesuai dengan jumlah interval. Nilai keanggotaan himpunan *fuzzy set* ( $A_i$ ) berada diantara 0, 0.5, dan 1.

$$A_1 = \left\{ \frac{1}{u_1} + \frac{0,5}{u_2} + \frac{0}{u_3} + \dots + \frac{n}{u_n} \right\}$$

$$A_2 = \left\{ \frac{0,5}{u_1} + \frac{1}{u_2} + \frac{0,5}{u_3} + \dots + \frac{n}{u_n} \right\}$$

...

$$A_n = \left\{ \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \dots + \frac{1}{u_n} \right\} \quad (3)$$

Tahap selanjutnya adalah melakukan fuzzifikasi berdasarkan nilai interval yang diperoleh dapat ditentukan nilai linguistik sesuai dengan banyaknya interval yang terbentuk. Langkah keempat menentukan *Fuzzy Logic Relations* (FLR) dengan memperhatikan *fuzzy set*  $A_i$  dari hari ke hari. FLR dapat ditulis dengan  $A_i \rightarrow A_k$  dimana  $A_i$  adalah hasil fuzzifikasi dan  $A_k$  adalah pengamatan sesudah data sebelumnya.

Langkah kelima menentukan *Fuzzy Logic Relations Group* (FLRG) berdasarkan hasil *Fuzzy Logic Relations* (FLR) dengan cara mengelompokkan setiap FLR yang memiliki  $A_i \rightarrow A_k$  yang sama.

Langkah keenam menentukan Nilai peramalan dihitung dengan menghitung nilai tengah sesuai dengan FLRG tersebut dihitung rata – ratanya atau ditulis dengan rumus :

$$F(t) = \frac{m_1+m_2+\dots+m_n}{\text{jumlah } m} \quad (4)$$

Selanjutnya, menghitung akurasi nilai peramalan dapat dilakukan dengan *Mean Absolute Percentage Error* (MAPE) untuk menilai tingkat akurasi atau kesalahan relatif dari suatu model atau metode peramalan. MAPE memberikan gambaran tentang sejauh mana prediksi model mendekati nilai aktual dengan menghitung rata-rata persentase kesalahan dari setiap prediksi. Rumus perhitungan MAPE adalah sebagai berikut :

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{A_i - F_i}{A_i} \right| \times 100\%$$

Dimana,

$n$  = ukuran sampel

$A_i$  = nilai data aktual

$F_i$  = nilai data peramalan

Tabel 2. Hasil Interpretasi Nilai MAPE

MAPE (%)	Interpretasi
<10	Peramalan sangat akurat
10 – 20	Peramalan baik
20 – 50	Peramalan layak
>50	Peramalan kurang akurat

Juga menghitung akurasi dengan *Mean Square Error* (MSE) digunakan dalam masalah regresi untuk mengukur seberapa dekat prediksi model dengan nilai sebenarnya dari data yang diamati. Rumus MSE adalah sebagai berikut :

$$MSE = \frac{1}{n} \sum_{i=1}^n (A_i - F_i)^2$$

Dimana,

$n$  = ukuran sampel

$A_i$  = nilai data aktual

$F_i$  = nilai data peramalan

Berikut adalah langkah-langkah untuk menghitung *Mean Squared Error* (MSE):

1. Hitung selisih antara nilai sebenarnya dan nilai prediksi untuk setiap observasi.
2. Kuadratkan setiap selisih.
3. Jumlahkan semua kuadrat selisih.
4. Bagi jumlah tersebut dengan jumlah total observasi.

*Root Mean Squared Error* (RMSE) digunakan dalam statistik dan pembelajaran mesin untuk mengevaluasi seberapa akurat model dalam memprediksi nilai dengan mempertimbangkan perbedaan antara nilai yang diamati (nilai sebenarnya) dan nilai yang diprediksi oleh model. Rumus perhitungan RMSE adalah sebagai berikut.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (A_i - F_i)^2}$$

Meskipun RMSE dan MSE memiliki kesamaan, RMSE berbeda karena mengambil akar kuadrat dari MSE. Hal ini menghasilkan ukuran yang memiliki satuan yang sama dengan variabel target asli.

### III. HASIL DAN PEMBAHASAN

Data diatas diambil dari dataset Kota Surabaya Indeks Standar Polusi Udara (ISPU) yang dimana diambil nilai maksimum dari seluruh komponen.

Tabel 3. Nilai Minimum dan Maksimum serta Interval

$Z_{min}$	15
$Z_{max}$	100
<b>Panjang interval</b>	8,5
<b>Jumlah interval</b>	10

Menentukan himpunan semesta  $U = [Z_{min} - F_1; Z_{max} + F_2]$ . Nilai  $F_1$  dan  $F_2$  adalah 0. Untuk menghitung banyak kelas menggunakan rumus  $Jumlah\ interval = 1 + 3,322 \log(n)$ . lalu rentang kelas dihitung dengan nilai maksimal dikurangi nilai minimal. Lalu interval kelas dihitung dengan membagi rentang kelas dengan banyak kelas.

Tabel 4. Pembentukan Interval dan Nilai Tengah

Interval yang Terbentuk	Range	Nilai Tengah (m)
$A_1$	{15; 23,5}	19,25
$A_2$	{23,5; 32}	27,75
$A_3$	{32; 40,5}	36,25
$A_4$	{40,5; 49}	44,75
$A_5$	{49; 57,5}	53,25

$A_6$	{57,5; 66}	61,75
$A_7$	{66; 74,5}	70,25
$A_8$	{74,5; 83}	78,75
$A_9$	{83; 91,5}	87,25
$A_{10}$	{91,5; 100}	95,75

Interval terbentuk dihitung dari nilai minimal ditambah dengan interval kelas. Sampai menjadi sesuai banyak kelas dan nilai maksimal. Setelah interval sudah terbentuk. Lalu mencocokkan nilai dengan data aktual. Sehingga didapat hasil fuzzifikasi sebagai berikut.

Tabel 5. Nilai Fuzzifikasi

Waktu	Nilai ISPU Max	Fuzzifikasi
01/01/2022	67	$A_7$
02/01/2022	73	$A_7$
03/01/2022	71	$A_7$
...	...	...
28/06/2023	67	$A_7$
29/06/2023	64	$A_7$
30/06/2023	63	$A_6$

Dari hasil fuzzifikasi, diproses untuk *Fuzzy Logic Relations* (FLR). Dengan cara mencatat hasil fuzzifikasi ke 1 ditaruh ke *Fuzzy Logic Relations* (FLR) kedua. Dilakukan seterusnya sampai akhir dari data.

Tabel 6. *Fuzzy Logic Relations* (FLR)

Waktu	Nilai ISPU Max	Fuzzifikasi	FLR
01/01/2022	67	$A_7$	
02/01/2022	73	$A_7$	$A_7$
03/01/2022	71	$A_7$	$A_7$
...	...	...	...
28/06/2023	67	$A_7$	$A_8$
29/06/2023	64	$A_7$	$A_7$
30/06/2023	63	$A_6$	$A_7$

Setelah itu, menghitung nilai tengah dari setiap interval dengan menambahkan nilai interval awal ditambah nilai interval akhir dan dibagi 2. Sehingga, didapat hasil seperti di bawah ini.

Tabel 7. *Fuzzy Logic Relations Group* (FLRG)

<i>Fuzzy Logic Relation Group</i>	
$A_1$	$A_1, A_9$
$A_2$	$A_2, A_3, A_6$
$A_3$	$A_3, A_4, A_5, A_6, A_7$
$A_4$	$A_3, A_4, A_5, A_6, A_7, A_8$
$A_5$	$A_2, A_4, A_5, A_6, A_7, A_8, A_9, A_{10}$
$A_6$	$A_3, A_4, A_5, A_6, A_7, A_8, A_9$
$A_7$	$A_3, A_4, A_5, A_6, A_7, A_8, A_9, A_{10}$

$A_8$	$A_4, A_5, A_6, A_7, A_8, A_9, A_{10}$
$A_9$	$A_5, A_6, A_7, A_8, A_9, A_{10}$
$A_{10}$	$A_6, A_7, A_8, A_9, A_{10}$

Setelah menentukan FLRG dengan menghitung nilai semua relasi lalu dibagi dengan jumlah nilai relasi  $F(t) = \frac{m_1+m_2+\dots+m_n}{\text{jumlah } m}$ , sehingga mendapatkan nilai peramalan sebagai berikut.

Tabel 8. Perhitungan Nilai Peramalan

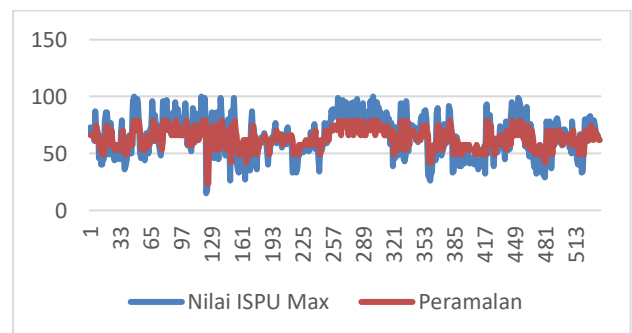
<i>Fuzzy Logic Relation Group</i>	Nilai Peramalan
$A_1$	23,5
$A_2$	41,91666667
$A_3$	49
$A_4$	53,25
$A_5$	57,5
$A_6$	61,75
$A_7$	66
$A_8$	70,25
$A_9$	74,5
$A_{10}$	78,75

Nilai peramalan diletakkan sesuai dengan FLR, dan hasil dari nilai peramalan tersebut dapat dilakukan perhitungan untuk akurasi peramalan yang kita hitung.

Tabel 9. Hasil Akhir Peramalan

Waktu	Nilai ISPU Max	FLR	Nilai Peramalan
01/01/2022	67	$A_7$	66
02/01/2022	73	$A_7$	66
03/01/2022	71	$A_7$	66
...	...	...	...
28/06/2023	67	$A_7$	66
29/06/2023	64	$A_7$	66
30/06/2023	63	$A_6$	61,75

Setelah mendapatkan nilai peramalan, dibentuk grafik yang menampilkan data aktual runtun waktu dan data nilai peramalan.



Gambar 2. Grafik Data Aktual Runtun Waktu dan Nilai Peramalan

Selanjutnya dilakukan perhitungan nilai akurasi peramalan dengan menggunakan rumus *Mean Absolute Percentage Error* (MAPE) dengan nilai 18,71%, *Mean Squared Error* (MSE) dengan nilai 188,32, dan *Root Mean Squared Error* (RMSE) dengan nilai 13,72.

Tabel 10. Hasil Akurasi Peramalan dengan MAPE

Waktu	Nilai ISPU Max	Nilai Peramalan	MAPE
01/01/2022	67	66	1.492537313
02/01/2022	73	66	7.042253521
03/01/2022	71	66	5.714285714
...	...	...	...
28/06/2023	67	66	1.492537313
29/06/2023	64	66	3.125
30/06/2023	63	61,75	1.984126984
			18,71

Tabel 11. Hasil Akurasi Peramalan dengan MSE dan RMSE

Waktu	Nilai ISPU Max	Nilai Peramalan	MSE
01/01/2022	67	66	1
02/01/2022	73	66	25
03/01/2022	71	66	16
...	...	...	...
28/06/2023	67	66	1
29/06/2023	64	66	4
30/06/2023	63	61,75	1.5625
			188,32
		RMSE	13,72

Tabel 12. Hasil Proses Peramalan *Fuzzy Time Series*

Waktu	Nilai ISPU Max	Peramalan	MAPE	MSE
01/01/2022	67			
02/01/2022	73	66	9.589041	49
03/01/2022	71	66	7.042254	25
04/01/2022	70	66	5.714286	16
05/01/2022	64	66	3.125	4
06/01/2022	61	61.75	1.229508	0.5625
07/01/2022	87	61.75	29.02299	637.5625
08/01/2022	70	74.5	6.428571	20.25
09/01/2022	70	66	5.714286	16
10/01/2022	65	66	1.538462	1
11/01/2022	46	61.75	34.23913	248.0625
12/01/2022	65	53.25	18.07692	138.0625
13/01/2022	40	61.75	54.375	473.0625
14/01/2022	40	49	22.5	81
15/01/2022	47	49	4.255319	4
16/01/2022	45	53.25	18.33333	68.0625

17/01/2022	78	53.25	31.73077	612.5625
18/01/2022	86	70.25	18.31395	248.0625
19/01/2022	86	74.5	13.37209	132.25
20/01/2022	80	74.5	6.875	30.25
21/01/2022	49	70.25	43.36735	451.5625
22/01/2022	73	57.5	21.23288	240.25
23/01/2022	77	66	14.28571	121
24/01/2022	61	70.25	15.16393	85.5625
...	...	...	...	...
21/06/2023	83	70.25	15.36145	162.5625
22/06/2023	68	74.5	9.558824	42.25
23/06/2023	64	66	3.125	4
24/06/2023	70	61.75	11.78571	68.0625
25/06/2023	79	66	16.4557	169
26/06/2023	74	70.25	5.067568	14.0625
27/06/2023	67	66	1.492537	1
28/06/2023	67	66	1.492537	1
29/06/2023	64	66	3.125	4
30/06/2023	63	61.75	1.984127	1.5625
			18.71720	188.32225
		RMSE		13.723055

#### IV. KESIMPULAN

Dengan memahami perubahan dalam struktur kualitas udara dan menganalisis dan meramalkan polusi udara, penelitian ini bertujuan untuk menangani efek negatif polusi udara. Hal ini dilakukan untuk menjaga keseimbangan ekologi dan kesehatan masyarakat. Dataset dari tahun 2022 hingga pertengahan tahun 2023 digunakan dalam penelitian ini. Hasil evaluasi dilakukan dengan menggunakan *Mean Absolute Percentage Error* (MAPE), *Mean Squared Error* (MSE), dan *Root Mean Squared Error* (RMSE). MAPE mencapai 18,71%, dan MSE mencapai 188,32, dan RMSE mencapai 13,72.

Hasilnya menunjukkan bahwa program peramalan polusi udara ini berhasil memberikan informasi yang berguna tentang polusi udara di Surabaya untuk masa depan. Ada kemungkinan bahwa program ini dapat dikembangkan lebih lanjut untuk menjadi lebih relevan dan akurat dalam menyelesaikan masalah polusi udara di daerah tersebut.

#### DAFTAR PUSTAKA

- [1] Ikhsanudin, A., Santoso, K. I., & Wahyudion, S. (2022). Metode Fuzzy Time Series Model Chen untuk Memprediksi Jumlah Kasus Aktif Covid-19 di Indonesia. *Transformasi*, 18(1).
- [2] Maulana, S. A., Batubara, S. H., & Rahman, W. K. (2023). Penerapan Metode Naïve Bayes

- dalam Peramalan Polusi Udara di Kota Jakarta. *Mutiara: Jurnal Penelitian dan Karya Ilmiah*, 1(6), 347-362.
- [3] Maharani, S., & Aryanta, W. R. (2023). Dampak Buruk Polusi Udara Bagi Kesehatan Dan Cara Meminimalkan Risikonya. *Jurnal Ecocentrism*, 3(2), 47-58.
- [4] Selasakmida, A. D., Tarno, T., & Wuryandari, T. (2021). Perbandingan Metode Double Exponential Smoothing Holt Dan Fuzzy Time Series Chen Untuk Peramalan Harga Paladium. *Jurnal Gaussian*, 10(3), 325-336.
- [5] Virgianti, V., Martha, S., & Imro'ah, N. (2021). Penerapan Fuzzy Time Series Chen Average Based Pada Peramalan Curah Hujan. *Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya*, 10(4).
- [6] Wahyuni, S. C., Arifianto, D., & Saifudin, I. (2022). Peramalan Jumlah Penduduk Miskin Di Pulau Jawa Menggunakan Metode Fuzzy Time Series Chen. *Smart Teknol*, 3(5), 133-139.