

A Comparative Analysis of Interval Determination Using Sturge's Rule and the Average-Based Method in Chen's Fuzzy Time Series Model for Forecasting Indonesia's CO_2 Emissions

Vivianti¹, Aswi Aswi^{2*}

¹ Postgraduate Program in Statistics, Universitas Negeri Makassar, Indonesia

² Statistics Department, Universitas Negeri Makassar, Indonesia

*Corresponding author: aswi@unm.ac.id

*Submission date: 25 June 2025, Revision: 06 February 2026, Accepted: 01 August 2026

ABSTRACT

Forecasting carbon dioxide (CO_2) emissions plays a crucial role in supporting mitigation efforts against global climate change. One widely used method for forecasting nonlinear time series data is the Fuzzy Time Series (FTS) Chen model, whose accuracy is strongly influenced by the interval determination method applied during fuzzification. This study compares the forecasting accuracy of the FTS Chen model using two interval determination approaches, Sturges' rule and the average-based method of Xihao and Yimin (2008), which sets the interval length to one half the mean of the absolute first differences between consecutive observations. The data used are Indonesia's annual CO_2 emissions from 1960 to 2022 (63 observations), obtained from Our World in Data. The average based method produced substantially more, and narrower, intervals (88) than Sturges' rule (7), which in turn allowed the fuzzy sets to represent the data pattern with less quantization error. This is reflected in the forecasting accuracy, with the average based method achieving a MAPE of 5.602% and an RMSE of 10.056, compared with a MAPE of 86.986% and an RMSE of 59.792 for Sturges' rule. Using the same vintage 2023 actual observation as a holdout check, the average-based forecast of 724.86 $MtCO_2$ remained substantially closer to the subsequently observed value than the Sturges forecast of 577.28 $MtCO_2$. These findings indicate that the choice of interval determination method, and the resulting number of intervals, plays a critical role in the forecasting accuracy of the FTS Chen model.

KEYWORDS

fuzzy Time Series Chen, interval determination, Sturges method, mean-based method, greenhouse gas emissions, CO_2 emissions, time series forecasting.

1. INTRODUCTION

Global climate change is one of the major challenges faced by humanity in the twenty-first century, affecting human health, agriculture, water resources, biodiversity, and coastal areas [1]. It is directly driven by greenhouse gas (GHG) emissions, primarily carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O), which differ in their heat-trapping effectiveness depending on their atmospheric concentration and turnover rate [2]. Because of this influence on human life and the natural environment, this study focuses on forecasting CO_2 emissions to support the formulation of mitigation policies. The forecasting approach used is the Fuzzy Time Series (FTS) Chen model, an effective method for handling nonlinear time series data [3]. Among several FTS models, the Chen model is one of the most fundamental, developed using simple arithmetic operations and

first applied to forecast new-student enrollments at the University of Alabama [4].

A key determinant of FTS forecasting accuracy is the method used to partition the universe of discourse into intervals during fuzzification: the number and width of the intervals determine how finely the fuzzy sets can represent the underlying data, and therefore how much information is lost when a numerical observation is mapped onto a single linguistic value. This study compares two interval determination methods, Sturges' rule and the average-based method, to examine how this methodological choice affects the accuracy of CO_2 emission forecasts for Indonesia.

2. LITERATURE REVIEW

Several studies have demonstrated that the average based interval method outperforms Sturges rule in forecasting using the FTS approach. Zeidi et al. [5] obtained a MAPE of 1.37% for the average based interval compared to 1.81% for Sturges on GGRM stock data. Similarly, research by Lathifah and Yanti [6] on ISPA cases showed that the average based interval, with a MAPE of 12%, was more accurate than Sturges. Furthermore, Kartika [7] compared three approaches distribution based, Sturges, and average based interval and found that the High Order Fuzzy Time Series Markov Chain with average-based interval determination yielded the best performance in forecasting PT BCA Tbk. stock prices.

Although these studies compared Sturges' and average-based interval determination within the FTS framework, none applied them to climate data, particularly CO_2 emissions. This study addresses that gap by examining the effectiveness of Sturges' rule and the average-based method in forecasting Indonesia's CO_2 emissions using the FTS Chen model. The reason interval determination matters can be stated more precisely in terms of the fuzzification mechanism itself. Sturges' rule, $k = 1 + 3.322\log(n)$, was originally devised to choose the number of classes for a frequency histogram based only on sample size n , it does not use any information about how much the series actually changes from one period to the next. Applied to a 63-year series, it necessarily returns a small, fixed number of wide intervals, regardless of the underlying volatility of the data. Because Chen's model assigns each observation to whichever interval contains it, wide intervals cause many distinct numerical values to collapse onto the same fuzzy set A_i , which (a) increases the quantization error carried into the fuzzy logical relationships (FLR), and (b) increases the chance that consecutive observations map to the same current state, producing repeated or many to one FLRs that are then merged into large, coarse fuzzy logical relationship groups (FLRG). The average-based method, by contrast, sets the interval length from the mean absolute period to period change in the series itself, so the resulting intervals are calibrated to the scale on which the data actually moves. This tends to produce narrower, more numerous intervals, which (a) reduces quantization error at the fuzzification step, and (b) yields FLRG structures in which each current state is associated with a small number of well separated next states, so the weighted average defuzzification rule draws on midpoints that are close to the true value rather than averaged across a wide range.

3. METHODOLOGY

3.1 Data

This study employs a quantitative approach focusing on numerical data analysis. It utilizes secondary data on greenhouse gas emissions published by Our World in Data. The data are in the form of time series, with the primary variable being global greenhouse gas emissions. The data set spans a long historical period, from January 1850 to 2022, thereby providing a comprehensive depiction of the dynamics of greenhouse gas emissions from the preindustrial era to the modern age.

3.2 Methods

The method applied in this study is the FTS Chen model, which involves the following stages:

1. Defining the Universe of Discourse (U).

The universe of discourse is defined as the range of GHG emission data to be analyzed, serving as the basis for interval formation and fuzzy set construction.

2. Determining the number and length of intervals.

Intervals can be determined using two approaches:

- (a) Sturge rule, which applies the formula $k = 1 + 3.322 \log(n)$ to determine the number of classes based on sample size.
 - (b) The average-based interval method of Xihao and Yimin [8] (1) compute the absolute first differences $|X_{t+1} - X_t|$ for every consecutive pair of observations and take their mean; (2) set the trial interval length to one half of that mean; and (3) round this trial length to a base value using the base- mapping table of Xihao and Yimin [8] (**Table 1**), after which the number of intervals follows from dividing the data range by the rounded interval length.
3. Defining fuzzy sets on U.
Each interval is defined as a fuzzy set, complete with its membership function, to represent the data in fuzzy linguistic form.
 4. Fuzzification of data.
The numerical GHG emission data are transformed into fuzzy form according to the defined fuzzy sets.
 5. Establishing Fuzzy Logical Relationships (FLR).
The fuzzification data are used to establish logical relationships among fuzzy sets, reflecting the dynamics of data changes.
 6. Grouping FLRs into Fuzzy Logical Relationship Groups (FLRG).
These groups are formed to organize the fuzzy logical relationships, making them more practical for forecasting.
 7. Defuzzification and forecasting.
The FLRG results are transformed back into numerical values through the defuzzification process, which are then used to generate forecasts of GHG emissions.

4. RESULT AND DISCUSSION

4.1 Time Series Data $MtCO_2$

The data used in this study consist of carbon emission records from 1960 to 2022, measured in $MtCO_2$ [9]. The dataset comprises 63 observations. The pattern of $MtCO_2$ emissions can be observed in Figure 1.

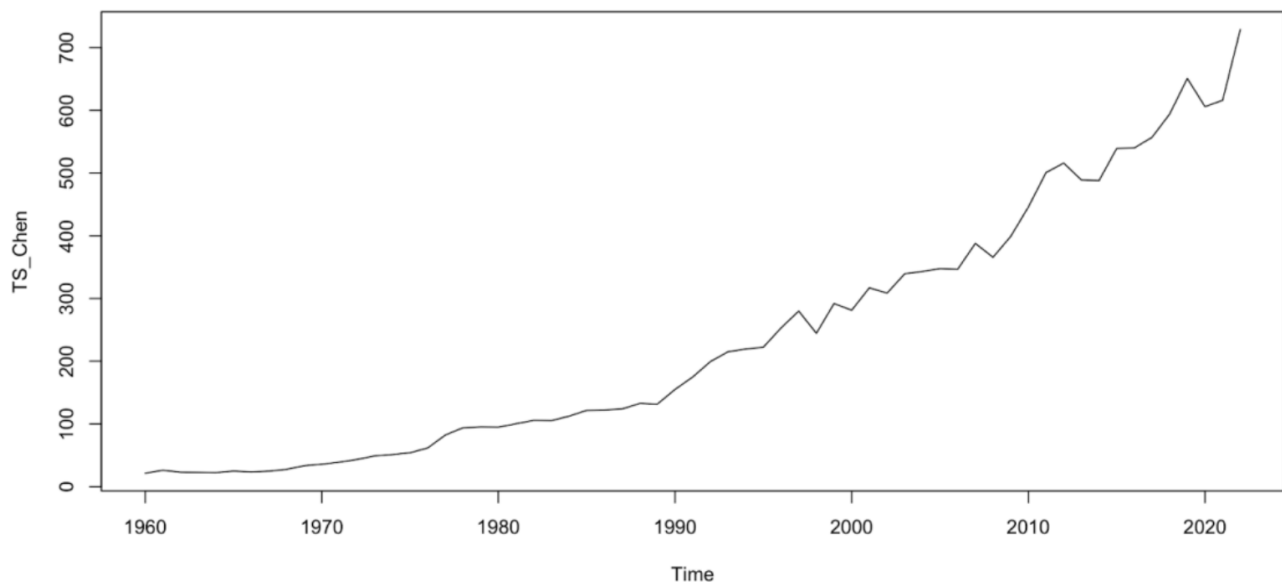


Figure 1. Data Pattern

From the graph, it can be observed that carbon emissions have experienced a consistent increase throughout the observation period. This growth appears relatively stable since the 1970s and becomes more pronounced after 2000, indicating a sharper

acceleration of emissions. The visible pattern demonstrates an upward trend, reflecting the rise in economic activity and fossil fuel based energy consumption over time. The annual fluctuations may be attributed to external factors such as global economic conditions, energy policies, and variations in both domestic and international energy consumption.

4.2 Fuzzy Time Series Chen Model Analysis

4.2.1 Determination of the Universe of Discourse (U)

The formula for determining U is given in Equation (1) [4].

$$U = [X_{min} - D_1, X_{max} + D_2] \quad (1)$$

where:

X_{min} = Minimum data value

X_{max} = Maximum data value

D_1 and D_2 are arbitrary positive constants determined by the study to define the universe of discourse based on historical data. Based on the actual carbon data, the maximum value obtained was 728.8833 and the minimum value was 21.3852. The values of $D_1 = 0$ and $D_2 = 0$ were selected, as there were no specific considerations for expanding the universe of discourse. Based on the data, $X_{max} = 728.8833$ and $X_{min} = 21.3852$. $D_1 = D_2 = 0$ were chosen for this study. This choice is appropriate here because the purpose of U in this study is to partition the historically observed range for in-sample fuzzification and one-step-ahead forecasting, and the historical data already span the full range needed for that purpose; no external estimate of how far emissions might extend beyond 2022 was available to justify a specific non-zero buffer.

4.2.2 Determination of the Number of Intervals

This study aims to compare two interval determination methods, namely Sturges Rule and the Average Based Interval method.

1. Sturge's rule

Sturges' Rule is expressed in Equation (2).

$$k = 1 + 3.322 \log(n) \quad (2)$$

where:

k = Number of classes

n = Total number of observations

By applying sturges rule, it was found that the number of intervals formed is 7 classes.

2. Average-Based Interval Method

The steps are as follows [8]:

- (a) Calculate all absolute differences between X_{t+1} dan X_t ($t = 1, \dots, n-1$) then compute the mean of these absolute differences as in Equation (3) [8].

$$mean = \frac{\sum_{i=1}^{n-1} |X_{t+1} - X_t|}{n-1} \quad (3)$$

where:

$Mean$ = Average value

n = Number of observations

X_t = data at time t

- (b) Determine half of the mean value obtained in step (1), which will be used as the interval length, using the in Equation (4) [8].

$$\ell = \frac{Mean}{n} \quad (4)$$

where ℓ = interval length

- (c) Based on the interval length obtained from step (2), the base length of the interval is determined according to the tabulation shown in **Table 1** [8]. The interval length is then rounded according to the interval basis table. Based

Table 1. Interval Base Table

Range	Base
0.1–1.0	0.1
1.1–10	1
11–100	10
101–1000	100

on the data used, the first step was to calculate the absolute differences between $(X_i + 1) - X_i$. The mean of these differences was then obtained, amounting to 16.48776. Dividing this average by two resulted in an interval length that produced 8 intervals. To determine the total number of intervals, the data range was divided by the interval length, yielding 88 classes. Since two different numbers of classes will be tested, the subsequent process will be conducted separately for each case.

4.3 Forecasting Using the Number of Classes Determined by the Average-Based Interval

1. Performing fuzzification on data

The fuzzification process is conducted by defining fuzzy sets A_i and transforming the observed historical data into corresponding fuzzy values. Let $A_1, A_2, \dots, A_k$ denote fuzzy sets representing the linguistic values of a given linguistic variable. The definition of the fuzzy sets $A_1, A_2, \dots, A_k$ within the universe of discourse U is formulated as in Equation (5) [4].

$$\begin{aligned} A_1 &= 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + \dots + 0/u_p \\ A_2 &= 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + 0/u_5 + \dots + 0/u_p \\ A_3 &= 0/u_1 + 0.5/u_2 + 1/u_3 + 0.5/u_4 + 0/u_5 + \dots + 0/u_p \\ &\vdots \\ A_p &= 0/u_1 + 0/u_2 + 0/u_3 + \dots + 0.5/u_{p-1} + 1/u_p \end{aligned} \quad (5)$$

where: u_i ($i = 1, 2, \dots, p$) is an element of the universal set and the number symbolised by “/” represents the degree of membership $\mu_{A_i}(u_i)$ to A_i ($i = 1, 2, \dots, p$) where the value is 0, 0.5 or 1. Define the fuzzy set A_i and perform fuzzification on the observed historical data using the number of classes based on the determination of the average based interval. The intervals formed are in **Table 2**.

Using the rules for determining fuzzy membership degrees, the fuzzification process is carried out in **Table 3**. After fuzzification, Fuzzy Logical Relationships (FLR) can be formed. FLR is the relationship between previous data and subsequent data in the form of fuzzy sets. The relationship is determined based on the fuzzification value of historical data. If F_{t-1} is fuzzyfied as A_i and F_t as A_j , then A_i is related to A_j . For example, in 1960, the fuzzification is A_1 and in 1961, the fuzzification is A_1 , so the FLR from 1960 to 1961 is $A_1 \rightarrow A_1$. A total of 62 FLR were formed because there were 63 pieces of historical data, each of which would be related. For 1960, there was no previous data, so there were no relationships. Therefore, the FLR was formed directly in 1961 with F_{t-1} being the fuzzification in 1960 and F_t being 1961. The FLR formed are presented in the following **Table 4**.

Table 2. Interval Table and Midpoint Values

Set	Dow	Up	Mid
A_1	21.3852	29.42495	25.40508
A_2	29.42495	37.4647	33.44483
A_3	37.4647	45.50445	41.48458
$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_{84}	688.68454	696.7243	692.70442
A_{85}	696.7243	704.76405	700.74417
A_{86}	704.76405	712.8038	708.78392
A_{87}	712.8038	720.84355	716.82367
A_{88}	720.84355	728.8833	724.86342

Table 3. fuzzification results for historical data

Relative	Ts	Fuzzyfication
1960	21,3852	A_1
1961	26,0016	A_1
1962	22,9749	A_1
$\vdots$	$\vdots$	$\vdots$
2020	605,9845	A_{73}
2021	615,9234	A_{74}
2022	728,8833	A_{88}

Table 4. Fuzzy Logic Relationship (FLR)

Current State		Next State
NA	$\rightarrow$	A_1
A_1	$\rightarrow$	A_1
A_1	$\rightarrow$	A_1
...	...	...
A_{79}	$\rightarrow$	A_{73}
A_{73}	$\rightarrow$	A_{74}
A_{74}	$\rightarrow$	A_{88}

2. Classifying FLR into groups to form Fuzzy Logical Relationship Groups (FLRG)

FLRG is performed by grouping Fuzzy Sets that have the same Current state. After that, they are grouped into one group based on FLR in **Table 5**. For example, for relation A_1 , namely $A_1 \rightarrow A_1$, $A_1 \rightarrow A_2$, the FLRG formed is $A_1 \rightarrow A_1, A_2$. The FLRG obtained based on the FLR formed is presented in **Table 5**. In determining FLRG, even if there is a repetition of relations, it is still calculated once.

3. Performing the defuzzification process and making forecasting calculations based on forecasting rules.

There are two stages in the defuzzification process. The first is to find the middle value of each interval using in Equation (6) [4].

$$m_1 = \frac{U_{lower} + U_{upper}}{2} \quad (6)$$

Second, calculate the forecast based on three defuzzification rules.

Rule 1: If the fuzzification result in year t is A_j and there is a fuzzy set that does not have a fuzzy logic relation, for

Table 5. Fuzzy Logical Relationship Group (FLRG)

Group 1	Fuzzy Logic Relation
1	$A_1 \rightarrow A_1, A_2$
2	$A_2 \rightarrow A_2, A_3$
3	$A_3 \rightarrow A_3, A_4$
$\vdots$	$\vdots$
87	$A_8 \rightarrow \emptyset$
88	$A_{88} \rightarrow \emptyset$

example if $A_i \rightarrow \emptyset$, where the maximum membership function value of A_i is in the interval u_i and the midpoint of u_i is m_i , then the forecast result F_{t+1} is m_i .

Rule 2: If the fuzzification result in year t is u_i and there is only one FLR in FLRG, for example, if $A_i \rightarrow A_j$, where A_i and A_j are fuzzy sets and the maximum membership function value of A_j is in the interval u_j and the middle value of u_j is m_j , then the prediction result F_{t+1} is m_j .

Rule 3: If the fuzzification result in year t is A_j and A_j has several FLR in the FLRG, for example $A_i \rightarrow A_{j1}, A_{j2}, \dots, A_{jk}$ where $A_i, A_{j1}, A_{j2}, \dots, A_{jk}$ are fuzzy sets and the maximum membership function value of $A_{j1}, A_{j2}, \dots, A_{jp}$ are in the intervals $u_{j1}, u_{j2}, \dots, u_{jk}$ and $m_{j1}, m_{j2}, \dots, m_{jk}$, then the F_{t+1} prediction result in Equation (7) [4].

$$F_{t+1} = \frac{(m_{j1} + m_{j2} + \dots + m_{jk})}{k} \quad (7)$$

where k is the number of distinct next-state fuzzy sets recorded for A_i in the FLRG (i.e. the count of items on the right-hand side of A_i 's FLRG entry). The mean values of each interval are presented in **Table 6**. the m_j terms are the midpoints of those k next-state intervals, each computed from Equation (6).

Table 6. Midpoint values for each interval

Set	Dow	Up	Mid
A_1	21.3852	29.42495	25.40508
A_2	29.42495	37.4647	33.44483
A_3	37.4647	45.50445	41.48458
A_4	45.50445	53.5442	49.52433
$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_{84}	688.68454	696.7243	692.70442
A_{85}	696.7243	704.76405	700.74417
A_{86}	704.76405	712.8038	708.78392
A_{87}	712.8038	720.84355	716.82367
A_{88}	720.84355	728.8833	724.86342

After determining the midpoint value of each interval, the next step is to calculate the forecast based on three defuzzification rules, resulting in the forecast based on FLRG presented in **Table 7**.

As an illustration, the forecast for the year 1961 is derived from the FLR $A_1 \rightarrow A_1$, which belongs to Group 1 in the FLRG, namely $A_1 \rightarrow A_1, A_2$. Based on the fuzzification rule, A_1 contains two FLRG relationships. The corresponding next states have midpoint values of 25.40508 and 33.44483. These two midpoint values are summed and then averaged, yielding a forecast value of 29.42495. Applying the same procedure, the forecast value for the year 2023 is obtained as 724.86342. The forecasting results are subsequently presented in graphical form in **Figure 2**.

Based on the forecast results, it is known that the amount of $MtCO_2$ will continue to increase in the coming years. This may be due to increased industrial activity, consumption of fossil fuels (such as coal, oil and natural gas), and growth in

Table 7. Forecasting Results

Point	Actual Data	Relative	Forecast
1960	21.3852	A_1-x-NA	NA
1961	26.0016	$A_1 \leftarrow A_1$	29.42495
1962	22.9749	$A_1 \leftarrow A_1$	29.42495
1963	22.7631	$A_1 \leftarrow A_1$	29.42495
...	...	...	...
2018	594.1014	$A_{72} \leftarrow A_{67}$	596.22741
2019	650.9059	$A_{79} \leftarrow A_{72}$	652.50566
2020	605.9845	$A_{73} \leftarrow A_{79}$	604.26716
2021	615.9234	$A_{74} \leftarrow A_{73}$	612.30691
2022	728.8833	$A_{88} \leftarrow -A_{74}$	724.86342
2023			724.86342

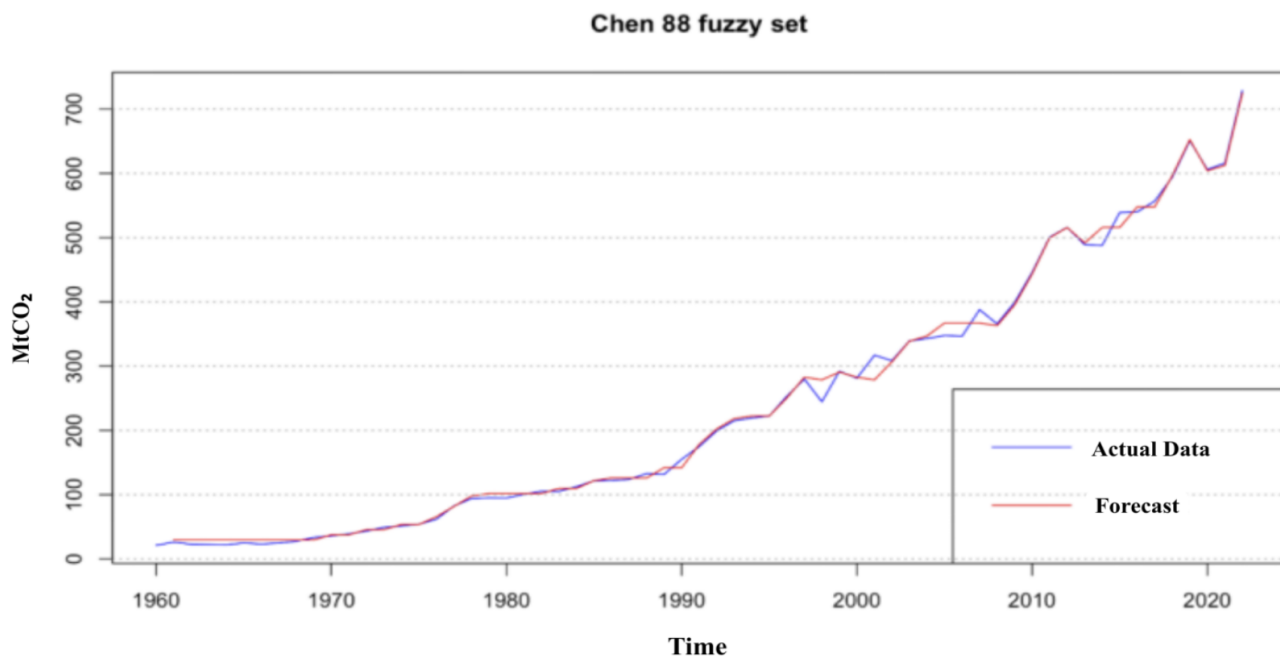


Figure 2. Comparison of Forecasting Results and Historical Data

the transport sector and urbanisation, which have not yet fully switched to renewable energy. This increase is in line with the Environmental Kuznets Curve (EKC) theory, which states that environmental emissions tend to increase in the early stages of economic growth before declining when countries begin to implement environmentally friendly technologies and sustainability policies [10].

4.4 Forecasting Using the Number of Classes from the Sturges Method

In this forecasting process, the process is the same as using the number of classes based on average based interval, the only difference being in the number of classes formed.

1. Performing fuzzification on the data

In the previous rule, the number of classes formed was 88, whereas in this Sturges analysis, there were 7 classes. The following 7 classes formed are presented in **Table 8**. In the previous rule, the number of classes formed was 88, whereas in this Sturges analysis, there were 7 classes. The following 7 classes formed are presented in **Table 8**.

Table 8. Table of Intervals and Midpoint Values

Set	Dow	Up	Mid
A_1	21.3852	122.4564	71.92078
A_2	122.4564	223.5275	172.99194
A_3	223.5275	324.5987	274.06309
A_4	324.5987	425.6698	375.13425
A_5	425.6698	526.7410	476.20541
A_6	526.7410	627.8121	577.27656
A_7	627.8121	728.8833	678.34772

Using the rules for determining fuzzy membership degrees, the fuzzification process is carried out in **Table 9**.

Table 9. fuzzification Results for Historical Data

Relative	Actual Data	Fuzzyfication
1960	21.3852	A_1
1961	26.0016	A_1
1962	22.9749	A_1
...	...	...
2020	605.9845	A_6
2021	615.9234	A_6
2022	728.8833	A_7

A total of 62 FLRs were formed because there were 63 historical data points, each of which would be related. For 1960, there was no previous data, so there were no related data points. Therefore, the FLR was formed directly in 1961 with F_{t-1} being the fuzzification in 1960 and F_t being 1961. The FLR formed is presented in **Table 10**.

Table 10. Fuzzy Logic Relationship (FLR) Table

Current State		Next State
NA	→	A_1
A_1	→	A_1
A_1	→	A_1
...	...	...
A_7	→	A_6
A_6	→	A_6
A_6	→	A_7

2. Classifying FLR into groups to form Fuzzy Logical Relationship Groups (FLRG)

The resulting FLRG is presented in **Table 11**.

Table 11. Fuzzy Logical Relationship Group (FLRG) Table

Group 1	Fuzzy Logic Relation
1	$A_1 \rightarrow A_1, A_2$
2	$A_2 \rightarrow A_2, A_3$
3	$A_3 \rightarrow A_3, A_4$
4	$A_4 \rightarrow A_4, A_5$
5	$A_5 \rightarrow A_5, A_6$
6	$A_6 \rightarrow A_6, A_7$
7	$A_7 \rightarrow A_6$

3. Performing the defuzzification process and making forecasting calculations based on forecasting rules.

The midpoint of each interval is presented in **Table 12**. Calculating the forecast based on three defuzzification rules, the

Table 12. The midpoint of each interval

Set	Dow	Up	Mid
A_1	21.3852	122.4564	71.92078
A_2	122.4564	223.5275	172.99194
A_3	223.5275	324.5987	274.06309
A_4	324.5987	425.6698	375.13425
A_5	425.6698	526.741	476.20541
A_6	526.741	627.8121	577.27656
A_7	627.8121	728.8833	678.34772

forecast results based on FLRG are presented in **Table 13**. A description of the forecast using graphs is presented in

Table 13. Forecasting Results

Point	Ts	Relative	Forecast
1960	21.3852	A_1 -x-NA	NA
1961	26.0016	$A_1 \leftarrow A_1$	122.4564
1962	22.9749	$A_1 \leftarrow A_1$	122.4564
1963	22.7631	$A_1 \leftarrow A_1$	122.4564
...	...	...	...
2018	594.1014	$A_6 \leftarrow A_6$	627.8121
2019	650.9059	$A_7 \leftarrow A_6$	627.8121
2020	605.9845	$A_6 \leftarrow A_7$	577.2766
2021	615.9234	$A_6 \leftarrow A_6$	627.8121
2022	728.8833	$A_7 \leftarrow A_6$	627.8121
2023			577.2766

Figure 3. Based on the results of data forecasts, it is known that the amount of $MtCO_2$ will continue to increase in the coming years, as illustrated in forecasts using average-based intervals. Both forecasts, using average-based intervals and Sturges, show a trend of continuous increase in $MtCO_2$ emissions. This finding is in line with the research by Cahyono

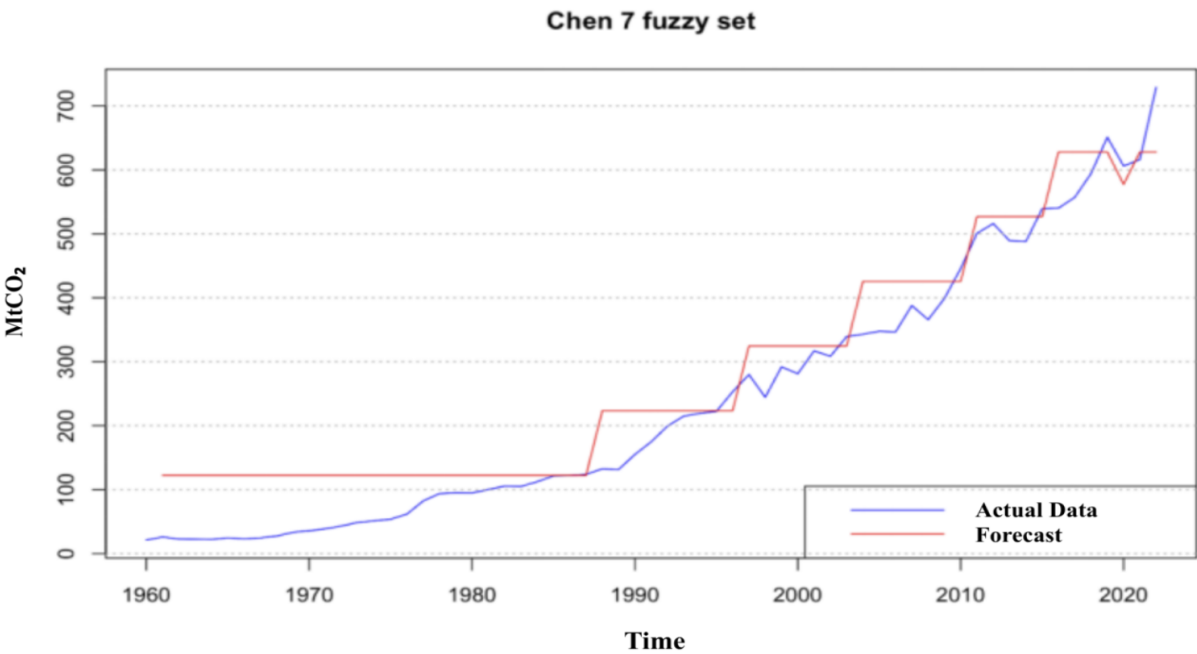


Figure 3. Comparison of Forecasting Results and Historical Data

and Sulistyowati, which explains that CO_2 emission projections in Indonesia will continue to increase until 2050, driven by increased consumption of fossil fuels, particularly coal in the power generation sector [11] A similar finding was also expressed by Aripriyanto, Khairani, and Hartono, who stated that without significant mitigation policies, Indonesia's CO_2 emissions are predicted to exceed one billion tonnes in the coming years due to economic growth and increasing energy needs [12].

4.5 Comparing the Forecasting Results of the Two Interval Determination Processes

After obtaining the forecasting results from the Fuzzy Time Series method with average-based interval determination and the Sturgess method for the amount of $MtCO_2$ from 1960 to 2022, a comparison of the accuracy of the forecasting results was carried out using MAPE and RMSE. The MAPE and RMSE of each method are presented in **Table 14**.

Table 14. Comparison Table of 2 Methods

Method	MAPE	RMSE	Conclusion
Average-based interval (X)	5,602	10,056	$X < Y$
Sturges (Y)	86,986	59,792	

Based on **Table 14**, it can be seen that to predict the amount of $MtCO_2$ emissions each year, the Fuzzy Time Series method with average-based interval determination produces lower MAPE and RMSE values compared to the Sturges method, so it can be concluded that it is better in the context of this study. This finding is in line with the statement by Elfajar, Setiawan, and Dewi that the determination of interval length greatly influences the formation of fuzzy relationships, which will certainly have an impact on the results of the forecast calculation [13]. This difference becomes more apparent when more intervals are formed. Several previous studies [5], [6], [7], [12], [13], [14], [15], [16] also show that the process of forming optimal intervals whether through an average-based approach or data distribution can improve the accuracy of forecasting results because the intervals formed can represent data fluctuation patterns in a more balanced and adaptive manner. However, this study has not been able to determine with certainty the maximum ideal number of intervals for optimal forecasting results. Visually, the

difference in error levels is also clearly seen through the significant MAPE and RMSE values. The following graph illustrates the comparison of the forecasting results presented in **Figure 4**.

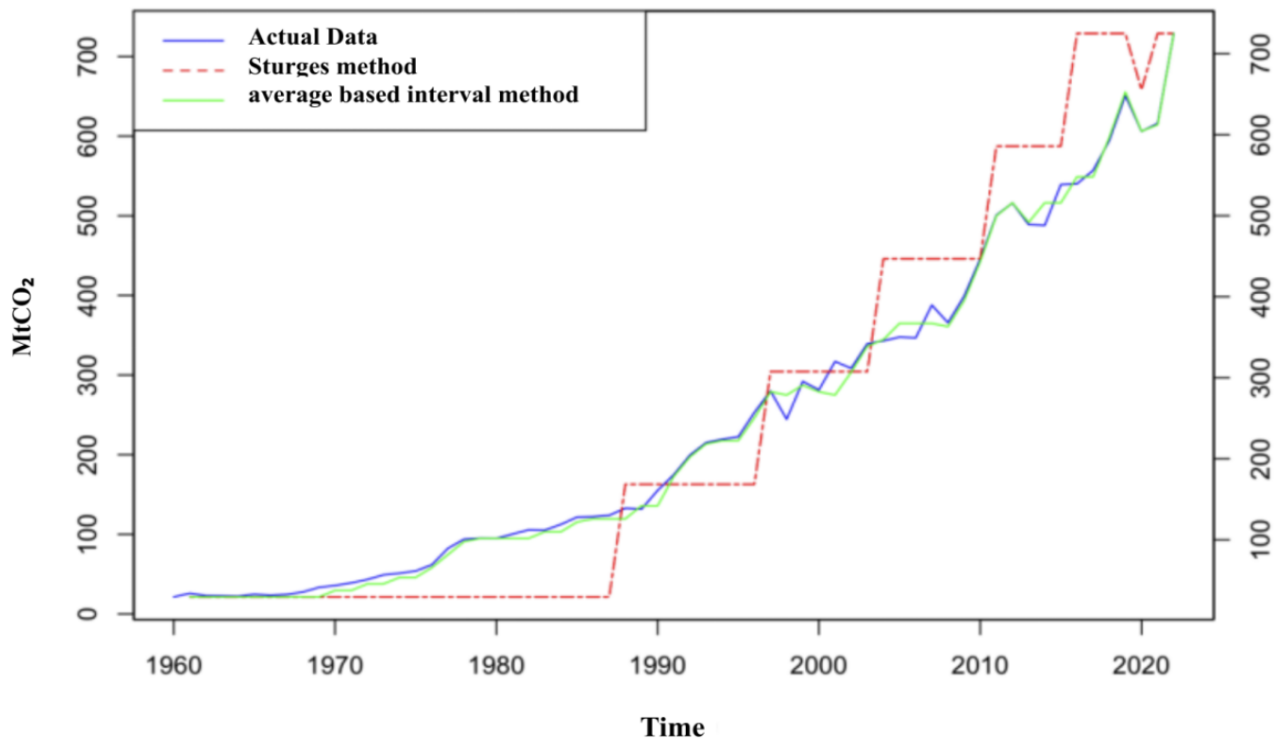


Figure 4. Comparison of Forecasting Results and Historical Data

5. CONCLUSION

Based on the discussion above, it can be concluded that the selection of the interval determination method in Chen's Fuzzy Time Series model greatly affects the accuracy of the forecasting results. The average based interval method produces more detailed intervals (88 classes) compared to the Sturges method (7 classes), thus being able to capture data patterns better. This is reflected in the forecasting accuracy values, where the average based interval method shows much higher performance with a MAPE of 5.602 and an RMSE of 10.056, while the Sturges method only produces a MAPE of 86.986 and an RMSE of 59.792. In addition, the results of Indonesia's carbon emissions prediction for 2023 using the average based interval method of 724.86 $MtCO_2$ proved to be closer to the historical data trend compared to the Sturges method, which produced 577.28 $MtCO_2$. The average based interval method is recommended in environmental time series forecasting because it provides more accurate and representative results.

This study has several limitations. It examines only the Chen FTS model, without comparison to higher-order or Markov-chain FTS variants; it compares only two interval determination methods among the many proposed in the literature (e.g., distribution-based, ratio-based, and optimisation-based approaches); it uses a single, annual, univariate CO_2 series for one country, so the results may not generalise to other emission series, other gases, other frequencies, or other countries; the universe of discourse was bounded at the historical maximum ($D_1 = D_2 = 0$), which constrains the model's ability to forecast values beyond the observed range; and the interval-count confound noted in Section 4.5 means the comparison should be extended, in future work, to a matched number of classes across methods before the average-based method's superiority can be attributed specifically to its interval-placement logic rather than to its interval count. With these caveats, the average-based interval method is recommended over Sturges' rule for FTS-based environmental time series forecasting of this kind.

REFERENCES

- [1] S. Keman, "Perubahan iklim global, kesehatan manusia dan pembangunan berkelanjutan," *Jurnal Kesehatan Lingkungan Unair*, vol. 3, no. 2, p. 3934, 2007.
- [2] B. Ma, M. S. Karimi, K. S. Mohammed, I. Shahzadi, and J. Dai, "Nexus between climate change, agricultural output, fertilizer use, agriculture soil emissions: Novel implications in the context of environmental management," *Journal of Cleaner Production*, vol. 450, p. 141801, 2024.
- [3] I. N. Neyun, W. Sulandari, and I. Slamet, "The autoregressive integrated moving average and fuzzy time series cheng hybrid for predicting stock price," *Jurnal Bumigora Information Technology (BITe)*, vol. 5, no. 2, pp. 139–150, 2023.
- [4] S. M. Chen, "Forecasting enrollments based on fuzzy time series," *Fuzzy Sets and Systems*, vol. 81, no. 3, pp. 311–319, 1996.
- [5] A. Zeidi, D. Kusnandar, and N. N. Debataraja, "Perbandingan average based dan sturges pada fuzzy time series chen untuk peramalan harga saham," *Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya*, vol. 12, no. 1, 2023.
- [6] T. S. Yanti, "Perbandingan interval pada fuzzy time series cheng untuk peramalan jumlah kasus ispa," in *Bandung Conference Series: Statistics*, vol. 4, no. 2, 2024, pp. 340–350.
- [7] Y. Kartika, "Penerapan metode high-order fuzzy time series markov chain dengan penentuan interval berbasis distribusi, sturges dan rata-rata pada peramalan harga saham (studi kasus harga penutupan saham harian pt. bank central asia tbk.)," Disertasi, Universitas Pendidikan Indonesia, 2022.
- [8] S. Xihao and L. Yimin, "Average-based fuzzy time series models for forecasting shanghai compound index," *World Journal of Modelling and Simulation*, vol. 4, pp. 104–111, 2008.
- [9] H. Ritchie, P. Rosado, and M. Roser, "Co₂ and greenhouse gas emissions," <https://ourworldindata.org/co2-and-greenhouse-gas-emissions>, 2023, our World in Data. Diakses: 21 September 2025.
- [10] S. Dinda, "Environmental kuznets curve hypothesis: A survey," *Ecological Economics*, vol. 49, no. 4, pp. 431–455, 2004.
- [11] W. E. Cahyono and A. R. Sulistyowati, "Projection of co₂ emissions in indonesia," *Energy Reports*, vol. 8, pp. 11 860–11 875, 2022.
- [12] S. Aripiyanto, D. Khairani, and A. Hartono, "Co₂ emission forecasting in indonesia until 2030: Evaluation of ets smoothing prediction models and their implications for global climate change mitigation," *Journal of Applied Data Sciences*, vol. 5, no. 1, pp. 46–55, 2024.
- [13] A. B. Elfajar, B. D. Setiawan, and C. Dewi, "Peramalan jumlah kunjungan wisatawan kota batu menggunakan metode time invariant fuzzy time series," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 1, no. 2, pp. 85–94, 2017.
- [14] O. Akay, "A computational method based on interval length for fuzzy time series forecasting," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 9, no. 1, pp. 21–26, 2021.
- [15] N. V. Tinh, N. C. Dieu, N. T. Duy, and T. T. Thanh, "Improved fuzzy time series forecasting model based on optimal lengths of intervals using hedge algebras and particle swarm optimization," *Advances in Science, Technology and Engineering Systems Journal*, vol. 6, no. 1, pp. 1286–1297, 2021.
- [16] H. Chen, X. Gao, and Q. Wu, "An enhanced fuzzy time series forecasting model integrating fuzzy c-means clustering, the principle of justifiable granularity, and particle swarm optimization," *Symmetry*, vol. 17, no. 5, p. 753, 2025.