# Comparison of the Fuzzy Time Series Chen Model and the Heuristic Model in Forecasting the Number of International Tourists in West Sumatra

Rizki Akbar[1,*], Fadhilah Fitri[1], Dodi Vionanda[1], and Tessy Octavia Mukhti[1]

[1] Departement of Statistic, Faculty of Mathematics and Natural Sciences, Universitas Negeri Padang, Indonesia
*rizkiakbarstat26@gmail.com

**Abstract.** The Fuzzy Time Series Chen and Heuristic are two forecasting methods based on fuzzy logic used to predict values in time series. The FTS Chen and Heuristic models have almost identical forecasting processes, but the main difference lies in how they develop fuzzy logical relationships. The FTS Chen model uses Fuzzy Logical Relationship Groups obtained from the results of Fuzzy Logical Relationships for the forecasting process. On the other hand, the FTS Heuristic model uses Fuzzy Logical Relationships directly in the forecasting process. Fuzzy Logical Relationships are a collection of fuzzy logical relationships used to connect values in time series. By using Fuzzy Logical Relationships, the Heuristic model can predict values in time series more accurately and effectively. The forecasting is done to plan the development of tourism infrastructure, determine service needs, and optimize tourism promotion. The data shows that the number of foreign tourists visiting West Sumatra has continued to grow from 2006 to 2023. The comparison of the accuracy of the forecasting results of FTS Chen and Heuristic models for foreign tourists in West Sumatra yielded a MAPE of 0.241% for FTS model Chen and 0.194% for FTS model Heuristic. This indicates that the best forecasting model for foreign tourists is the Heuristic model due to its lower MAPE value.

Keywords: Fuzzy Time Series, Chen Model, Heuristic Model, Mean Absolute Percentage Error (MAPE).

## 1. Introduction

Forecasting aims to predict future events. According to Hyndman & Athanasopoulos (2018), there are two types of forecasting techniques: qualitative and quantitative. Qualitative techniques rely on expert opinions and do not use data in the form of numbers, while quantitative techniques use historical data in the form of numbers, known as time series data. Forecasting is carried out to plan tourism infrastructure development, determine service needs, and optimize tourism promotion. Data shows that the number of foreign tourists coming to West Sumatra has continued to grow from 2006 to 2023. By forecasting, the government and tourism stakeholders can prepare themselves to increase tourism carrying capacity, such as accommodation, transportation, and tourist attractions, to welcome growth. number of tourists in the future. According to Azmiyat and Tanjung (2017), the advantage of FTS is that it does not require large amounts of data so it does not require fulfilling assumptions.

The FTS method in forecasting is a new concept that uses fuzzy logic which can explain forecasting data in the form of linguistic variables. Linguistic variables are variables that do not use numbers, but these variables have verbal explanations. The advantage of linguistic variables is that the information obtained from the variables is easy to understand, but the weakness is that these variables are compared with variables that contain inaccurate linguistic variable numbers. Apart from that, the FTS method can be used to forecast the number of foreign tourists (tourists) in West Sumatra. West Sumatra is one of the provinces in Indonesia that has tourism that is in demand by foreign tourists after the province of Bali. This can be shown according to the Central Statistics Agency (2023) of Bali province. The number of foreign tourists reached 5,273,258 people. According to the Central Statistics Agency (2023) of West Sumatra province, the number of foreign tourists reached 56,645 people. However, the number of foreign tourists (tourists) who came to West Sumatra in the period January 2006 to December 2023 experienced quite significant fluctuations. Several factors that influence these fluctuations include the COVID-19 pandemic and regional and social

restrictions. The following are foreign tourists in West Sumatra during the period January 2006 - December 2023 according to BPS (2023) which can be seen in Figure 1.
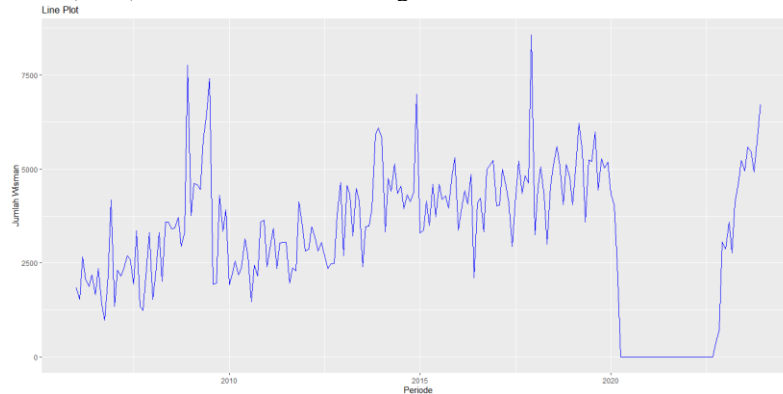


**Figure 1**. Foreign tourists in West Sumatra Period January 2006 – December 2023 (monthly)

The statistics in Figure 1 represent the number of international tourists in West Sumatra from January 2006 to December 2023. Figure 2 displays the number of persons who traveled abroad from West Sumatra in December 2017, which amounted to 8,568 individuals. This indicates a significant rise compared to the previous month and year. From April 2020 to September 2022, no foreign tourists were visiting West Sumatra due to several factors. The main reason was the COVID-19 pandemic and the limitations imposed in the region, which prevented any visits from taking place. Subsequently, during the onset of the Covid-19 epidemic, there was a notable surge in the number of international tourists visiting West Sumatra. In October 2022, there was a slight increase in the number of foreign tourists in West Sumatra, with a total of 373 people. However, in November 2022, there was a significant increase compared to October, with a total of 709 individuals. In August 2023, there was a significant and sudden surge in population, with an increase of up to 5,577 individuals. In December 2023, there was a significant surge in the number of people, with a total of 6,710 individuals, compared to other months of 2023. The occurrence is a result of the pre-existing stability that was already established during the COVID-19 epidemic, both in terms of public health and the economy. This condition is highly favorable for the tourism sector in West Sumatra since there has been a significant increase in the number of foreign tourists visiting the region. This study aims to examine the performance forecasts of two models: the Fuzzy Time Series Chen Model and the Heuristic Model. Anticipated outcomes investigation This can assist stakeholders in making informed decisions on the growth of tourism in West Sumatra.

## 2. RESEARCH METHODS

This type of research is applied research. The data used in this research is data sourced from the website of the Central Statistics Agency (BPS) of West Sumatra Province. This data is data on the number of foreign tourists in West Sumatra Province from 2006 to 2023. The data is processed and analyzed using Rstudio software. The data used is secondary data. The following are the steps in carrying out the FTS analysis of the Chen model and Singh model in Figure 2.

## 3. RESULTS AND DISCUSSION

### 3.1. Chen Model FTS Analysis

Based on data on the number of foreign tourists in West Sumatra for the period January 2006 to December 2023, a minimum value ($N_{min}$) of 0 was obtained from the period April 2020 to September 2022 and a maximum value ($N_{max}$) was 8,568 in the December 2017 period. Then for the $N_1$ and $N_2$ each is set to 0. So the universal set $U = [N_{min} - N_1; N_{max} - N_2]$ is $U = [0; 8.568]$. Next, determine the number of intervals and interval length.
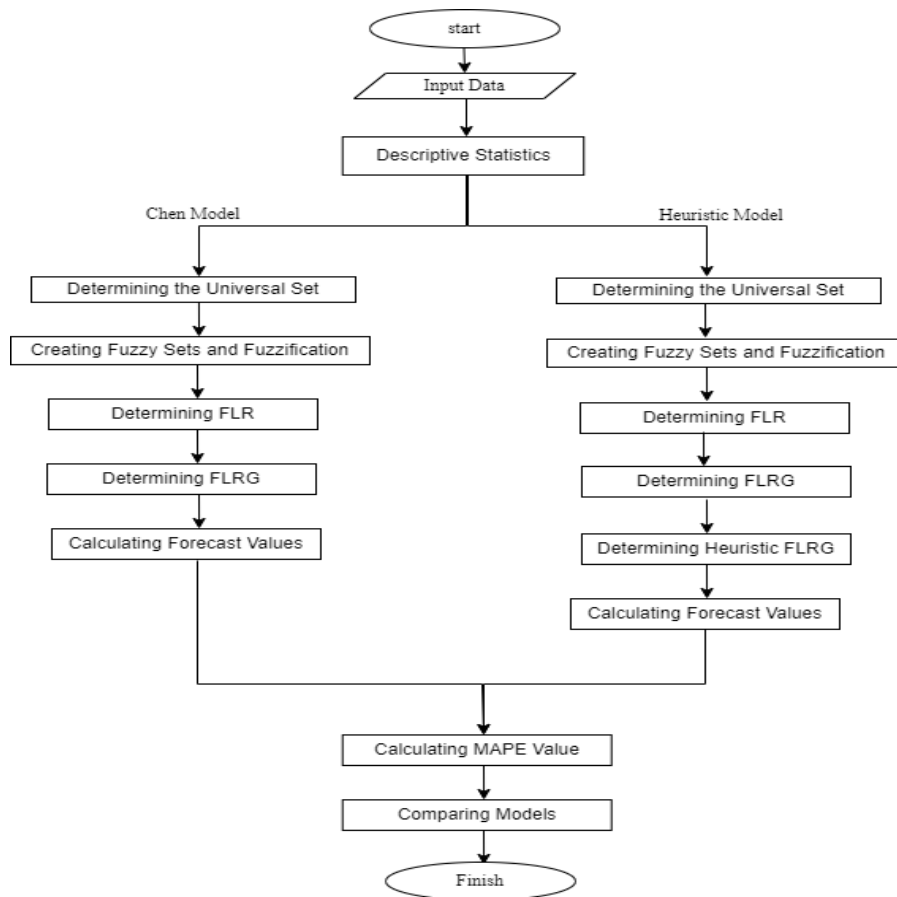
**Figure 2.** Chen Model FTS Flowchart and Heuristic Model

**Table 1**. Intervals and Fuzzification

| Intervals | Fuzzification |
|---|---|
| $U_1 = [0 ; 8,568]$ | $A_1$ |
| $U_2 = [8568 ; 17,136]$ | $A_2$ |
| $U_3 = [17.136 ; 25,704]$ | $A_3$ |
| $U_4 = [25,704 ; 34,272]$ | $A_4$ |
| $U_5 = [34.272 ; 42,840]$ | $A_5$ |
| $U_6 = [42,840 ; 51,408]$ | $A_6$ |
| $U_7 = [51,408 ; 59,976]$ | $A_7$ |
| $U_8 = [59,976 ; 68,544]$ | $A_8$ |
| $U_9 = [68,544 ; 77,112]$ | $A_9$ |

The number of intervals is used to determine how many fuzzy sets can be performed on forecasting data. The number of intervals can be calculated using Sturgess' rule as follows:

$$K = 1 + 3,3\log(n) = 1 + 3,322 \times \log(216) = 8.75 \approx 9.$$

Next, determine the length of the class using the following equation

$$l = \frac{[(D_{max} + D_2) - (D_{min} - D_1)]}{k} = \frac{(8.568 + 0) - (0 - 0)}{9} = 8.568$$

After obtaining the number of class intervals of 9 and the interval length of 8,569. So the interval and fuzzification in Table 1 are formed.

Based on Table 1 obtained results were fuzzyfication. $A_1$ is form linguistics from values on range $U_1$. Following so on until $A_9$ if data comes in in the $U_9$ interval so results in its fuzzification is $A_9$. Following results from fuzzification carried out on total data traveler abroad.

**Table 2.** Data Fuzzification

| Period | Data | Fuzzification |
|--------|------|---------------|
| Jan-06 | 1844 | A2 |
| Feb-06 | 1540 | A2 |
| Mar-06 | 2667 | A3 |
| ⋮ | ⋮ | ⋮ |
| Oct-23 | 4925 | A6 |
| Nov-23 | 5840 | A7 |
| Dec-23 | 6710 | A8 |

From Table 2 it can be seen that fuzzification entire data. Results fuzzification on Table 3 will used on the FLR stage. FLR is determined with the method connecting variable linguistics from the fuzzification process in Table 2. FLR can be written $A_i < - A_j$, Where $A_i$ is the current data observation $f(t)$ and $Aj$ is the observation before $f(t$ -1) on time series.

**Table 3**. FLR results

| Period | Data | FLR |
|--------|------|-----|
| Jan-06 | 1844 | - |
| Feb-06 | 1540 | A2<--A2 |
| Mar-06 | 2667 | A3<--A2 |
| Apr-06 | 2069 | A3<--A3 |
| ⋮ | ⋮ | ⋮ |
| Oct-23 | 4925 | A6<--A6 |
| Nov-23 | 5840 | A7<--A6 |
| Dec-23 | 6710 | A8<--A7 |

Based on FLR results in Table 3 FLRG was formed with the method grouping every FLR that has a side right the same f($t$ - 1). Results grouping or FLRG obtained based on FLR results for each data is shown in Table 4. In Table 4, the FLRG is obtained where there are nine groups with different characteristics for each group, such as group one has characteristics A1 which has a relationship with A1, while group two has characteristics A2 which has a relationship with A2, A3, A4, A5. After the FLRG is obtained, the next step is to forecast using the Chen model in Table 5.

**Table 4**. FLRG results

| Group | FLRG |
|---|---|
| 1 | A1 ->A1 |
| 2 | A2 -> A2, A3, A4, A5 |
| 3 | A3 -> A3, A4, A5, A8, |
| 4 | A4 -> A3, A4, A5, A7, A8, A9 |
| 5 | A5 -> A3, A4, A5, A6, A7 |
| 6 | A6 -> A4, A5, A6 |
| 7 | A7 -> A5, A6, A7 |
| 8 | A8 -> A5,A7 |
| 9 | A9 -> A4,A5 |

**Table 5**. Results Chen Model Forecasting

| Period | Data | Forecasting |
|---|---|---|
| Jan-06 | 1844 | - |
| Feb-06 | 1540 | 1904 |
| Mar-06 | 2667 | 1904 |
| ⋮ | ⋮ | ⋮ |
| Oct-23 | 4925 | 4284 |
| Nov-23 | 5840 | 4284 |
| Dec-23 | 6710 | 5236 |

Based on Table 5, it can be seen from the results of forecasting carried out using the *Chen model* that the first data forecast is not available because the data delays information regarding predictions of the number of foreign tourists, thereby affecting forecasts for the next period. Meanwhile, forecasts for the next three periods, namely. March 2007 to August 2007 gave the same result, namely 2380.

### 3.2. FTS Model Heuristic Analysis
Heuristic Analysis Results Universe Heuristic Results Heuristic model and Chen model have the same steps and results. There is a clear difference between the heuristic model that uses FLRG heuristic results and the Chen model that only uses FLRG results to determine its prediction results. The following are the heuristic results from the FLRG heuristic model. Where to determine the FLRG Heuristic results there are 3 rules.
**Rule 1**
If the fuzzy logical relationship group (FLRG) heuristic $A_j$ is empty ( $A_j \rightarrow$), then defuzzification $F(t)$ is obtained from the middle value of the interval that has the maximum membership value at $A_j$.
**Rule 2**
If there is only one $A_j \rightarrow A_{p1}$ fuzzy logical relationship group (FLRG) heuristic ( $A_j$), then the middle value of the interval has a maximum membership value of $A_{p1}$.
**Rule 3**
If the fuzzy logical relationship group (FLRG) heuristic $A_j$ is more than one ( $A_j \rightarrow A_{p1}$, $A_{p2}$,…, $A_{pk}$), then defuzzification $F(t)$ is obtained from the average of the middle values of each interval which has the maximum membership value in each $A_{p1}$, $A_{p2}$,…, $A_{pk}$. The FLRG Heuristic results were obtained using rule

3, namely obtaining the average of the middle values of each interval. Obtained based on group A4. The following is a table of results from group A4.

**Table 6**. FLRG Heuristic Results

| Group A4 | Forecasting |
|:---:|:---:|
| A3 | 1904 |
| A4 | 2856 |
| A5 | 3808 |
| A7 | 5712 |
| A8 | 6664 |
| A9 | 7616 |
| **Results** | **4760** |

From Table 6, the FLRG *Heuristic results* show that group A4 -> A3, A4, A5, A7, A8, A9 then use rule third And obtained results FLRG *Heuristic* is 4,760. After the FLRG *Heuristic* is obtained, the next step is to forecast using the *Heuristic model* in Table 7.

**Table 7**. Results Heuristic Model Forecasting

| Period | Data | Forecasting |
|:---:|:---:|:---:|
| Jan-06 | 1844 | - |
| Feb-06 | 1540 | 1428 |
| Mar-06 | 2667 | 1904 |
| ⋮ | ⋮ | ⋮ |
| Oct-23 | 4925 | 3808 |
| Nov-23 | 5840 | 5712 |
| Dec-23 | 6710 | 6664 |

Based on Table 7, it can be seen from the forecasting results using the Heuristic model that the first data forecasting results are not available because the data is lag data from the forecast number of foreign tourists which will influence the forecast in the next period. Meanwhile, for forecasting the next period, namely May 2020 to November 2022, the same results were obtained in 1904

### 3.3. Measuring the Level of Accuracy and Selecting the Best Model

Chen's model estimates of the number of foreign tourists in West Sumatra using Rstudio software are presented in Figure 3. Based on Figure 3, comparative information is obtained between the original data and data resulting from forecasting the number of foreign tourists in West Sumatra using the Chen model. It can be seen that the red graph is a forecast estimate and the blue graph is a graph of original data on the number of foreign tourists in West Sumatra from January 2006 to December 2023. The estimated value for the first data does not exist because it is lag data from the number of foreign tourists in West Sumatra. Then you can see the comparison between the original data and the forecasting results, the shape of the plot of the estimated results for each period is slightly different at certain points from the original data. As at this point throughout 2019, forecast estimates show an increasing graph, while the original data shows a decrease.
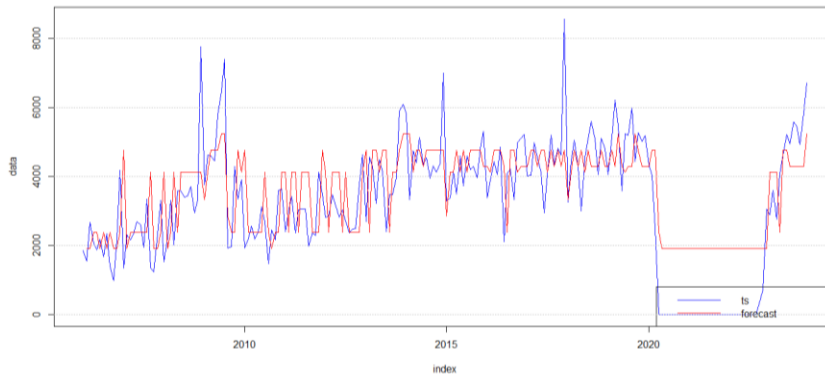
**Figure 3.** Chen Model Estimation

Next, the *Heuristic model* can be seen in the historical data plot and data from the estimated value of the number of foreign tourists in West Sumatra. The following *Heuristic model* can be seen in Figure 4.
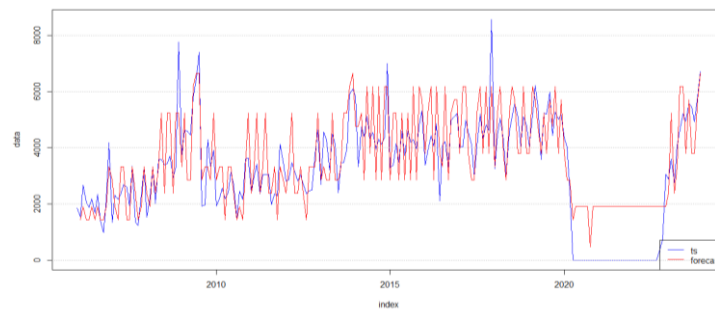


**Figure 4**. Results Heuristic Model Estimation

Based on Figure 4, comparative information is obtained between the original data and forecast data using the Singh model. The blue graph is for original data, then the red graph is for forecasting data on the number of foreign tourists in West Sumatra. Forecasting data from the first to the third point has no value because these points are lag data from the number of foreign tourists. To compare the forecasting data and the original data in Figure 4, you can see that using the Heuristic model, the forecasting results are almost close to the actual data, there are only a few points that are different, like the points in 2020, but not too far from the original data points.

To make it clearer, MAPE will be used to measure the best model. In this research, two FTS Chen and Heuristic models were compared using the assumption that if the forecasting model obtained the smallest MAPE value then that model was the best forecasting model. The MAPE values in both models can be seen in Table 8.

**Table 8**. Comparison MAPE value

| Model | MAPE |
|---|---|
| *Chen* | 0.241% |
| *Heuristics* | 0.194% |

From Table 8 it can be seen that the *Chen model* obtained a MAPE value of 0.241%, while the *Heuristic* model obtained a MAPE value of 0.194%. From the comparison of the two MAPE values, the *Heuristic model* has a smaller MAPE value than the Chen model.

## 4. Conclusion

A comparison of the two FTS Chen and Heuristic prediction methods for the number of foreign tourists shows that the FTS Heuristic model has a lower MAPE value than the Chen model. Thus, it can be concluded that the FTS Heuristic model predicts the number of foreign tourists in West Sumatra better. The forecasting results made using the FTS Heuristic model can be used as a reference and benchmark for the tourism government in making decisions to promote West Sumatra tourism in the future. Future research should explore and compare additional Fuzzy Time Series (FTS) models, such as the Cheng, Lee, Markov Chain, and Singh models. This comparison could lead to the identification of an even more effective model for the task at hand**.**

## References

[1]     Azmiyati, S., Tanjung, WN (2017), "Forecasting the Number of Fresh Fruit Bunches (FFB) for Palm Oil Using the *Chen Fuzzy Time Series Method* and the Ruey Chyn Tsur Algorithm", Vol..8, No. 1, p. 36-48.

[2]     Guo, H., Pedrycz, W. & Liu, X. (2019), *Fuzzy time series forecasting based on axiomatic fuzzy set theory . Neural Comput & Applic* 31 , 3921–3932.

[3]     Kun-Huang Huarng, Tiffany Hui-Kuang Yu, & Yu Wei Hsu. (2007). *A Multivariate Heuristic Model for Fuzzy Time-Series Forecasting. IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics), 37(4), 836–846.*

[4]     Shyi-Ming Chen, & Nai-Yi Wang. (2010). *Fuzzy Forecasting Based on Fuzzy-Trend Logical Relationship Groups . IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* , 40(5), 1343–1358.

[5]     Nisa, RK, Nur, IM, & Fauzi, F. (2021). *Chen* Model *Fuzzy Time Series* Method and *Heuristics* in Forecasting. Muhammadiyah University Semarang ..

[6]     Huang, K. 2001. *Heuristics Models of Fuzzy Time Series for Forecasting. Journal of Fuzzy Sets and Systems* .

[7]     Tsaur, R.-C., O Yang, J.-C., & Wang, H.-F. (2005 *). Fuzzy relation analysis in fuzzy time series models . Computers & Mathematics with Applications* , 49(4), 539–548

[8]     Egrioglu, E., Aladag, C. H., & Yolcu, U. (2013). *Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks. Expert Systems with Applications* , 40(3), 854–857.

[9]     Chen, S. M. (1996). *Forecasting enrollments based on fuzzy time series* . 311-319.