Forecasting Analysis of the Number of Foreign Tourist Arrivals with the Chan Chan Fuzzy Time Series Model method

Devia Kartika [1,*], Wifra Safitri[2], Sri Rahmawati[3]

[1,2,3] Universitas Putra Indonesia YPTK Padang

Jl. Raya Aru, Padang, Indonesia

Article Information	Abstract
Submitted 15 Mar 2023 **Accepted :** 20 May 2023 **Published** 30 May 2023	The tourism sector greatly influences the community's economic growth, one of the factors is the visit or number of foreign tourists who now have an important role in increasing the prosperity of life which has advantages in its cultural resources and natural resources. In realizing sustainable development in the tourism sector, activities are carried out to predict the value of a variable in the future, namely forecasting. This study aims to predict the arrival of foreign tourists to West Sumatra using the Fuzzy Time Series method. The historical data used is data on foreign tourist visits to West Sumatra in 2019. The implementation of the Fuzzy Time Series on historical data produces a Mean Absolute Percentage Error (MAPE) of 11.45% with the lowest error rate of 1,184%. These results indicate that the use of the Fuzzy Time Series in data forecasting of foreign tourist visits to West Sumatra has very good results.

1. Introduction

Foreign tourists have an important role in the economic growth of the tourism sector. In increasing foreign tourist arrivals, a movement in sustainable development is needed in the tourism sector. Advancing tourism development must be in line with the growth trend of foreign tourist visits so that they are right on target, effective and efficient [1]The development of tourism is an activity that ultimately has an economic influence on life around tourism sites, the economic influence can be felt by the surrounding community, especially from an economic perspective, namely increasing income. Another advantage is the construction of facilities for easy access to tourism locations, for example transportation and sales stalls so that they can open jobs for the community [2]The positive impact of the tourism sector can benefit Indonesia as a country that has a lot of uniqueness both in terms of culture and beautiful tourist spots. The Minister of Tourism and Creative Economy, Mari Elka Pangestu revealed that Indonesia's strength in the tourism sector lies in cultural and natural resources, as assets that can be promoted [3]. One of the tourist attractions that offer cultural and historical resources is West Sumatera. Analysis in predicting or predicting one of them by using time series data is by using time series analysis

Research Methods

Fuzzy time series is a method introduced by Song and Chissom (1993), which is a new concept for forecasting using Fuzzy logic, namely the problem of time series forecasting that is able to provide an explanation for vague data and presented in linguistic values. The use of the Fuzzy time series method has previously been carried out in several studies, including research conducted by [4] entitled Implementation of the Fuzzy Time Series Method for Forecasting the Number of Visitors at Fort Rotterdam. And . in research conducted by [5]entitled a comparison of the Fuzzy time

series method and the holt double exponential smoothing in forecasting the number of new students at the Bogor Agricultural Institute. These studies show that Fuzzy time series are recommended to be used because they tend to get small error values.

Forecasting steps using Chen's Fuzzy time series model [4] are as follows:

1 Determine the Universe of Discourse U (Universe of Discourse), then divide the intervals with the same distance. If there is one amount of data in an interval that is greater than the average value of the number of data in each interval, then the interval is divided into smaller intervals by dividing by two.

$$U = [Xmin – D1, Xmax + D2]$$

Where:
Xmin = Minimum data
Xmax = Maximum data

D1 and D2 are arbitrary positive numbers determined by the researcher to determine the universe set from the Historical data set.

2. Formation of intervals.
In this study the method used to form intervals is an average based method, which has an algorithm [6]. Calculating all the absolute values of the difference between $Xt+1$ and $X1$ ($t = 1, … , n − 1$) so that the average absolute difference value is obtained:

$$\sum_{1}^{n} = 1 \frac{|Xt+1 - X1|}{n-1}$$

a. Determine half of the average obtained from the first step to then be used as the length of the interval with the equation.

$$l = \frac{rata - rata}{2}$$

b. Based on the interval length obtained, then the basis of the interval length is determined according to the basis tabulation. Table 1 describes the length of the interval and its basis.

Tabel 1. Basis Interval

Interval	Basis
0,1 – 1,0	0.1
1,1 -10	1
11 -100	10
101 - 1000	100

3. Defining the Fuzzy Ai set and fuzzifying the observed historical data. The membership degree value of µAi (ui) is determined according to the following rules (Chen, 1996):

Rule 1: If historical data Xt is included in ui, then the membership degree value for ui is 1, and ui+1 is 0.5 and if it is not ui and ui+1, it is declared zero.

Rule 2: If historical data Xt is included in ui , $1 ≤ i ≤ p$ then the membership degree value for ui is 1, for ui-1 and ui+1 is 0.5 and if not ui , ui-1 and ui+1 means declared zero.

Rule 3: If historical data Xt is included in up, then the membership degree value for up is 1, for up-1 is 0.5 and if it is not up and up-1, it means zero.

For example A1, A2, …, Ak is a Fuzzy set that has a linguistic value from a linguistic variable, the definition of the Fuzzy set A1, A2, …, Ak in the universe of speech U is as follows (Chen, 1996):

A1 = 1/u1+ 0.5/u2+ 0/u3+ 0/u4+ 0/u5… p
A2 = 0.5/u1+1/u2+ 0.5/u3+ 0/u4+ 0/u5.. p
A3 = 0/u1+ 0.5/u2+ 1/u3+ 0.5/u4+ 0/u…p
Ap = 0/u1+ 0/u2+ 0/u3 … ,5/up-1+ 1/ up

Where: ui (i=1, 2, .., p) is an element of the universe set and the number marked with the symbol ―/‖ denotes the membership degree of µAi(ui)) to Ai (i=1, 2, ..,p) where the value is 0, 0.5 or 1.

Perform and create Fuzzy Logical Relationship (FLR) tables based on historical data. Fuzzy logical relationship Aj → Ak means that if the enrollment value in year I is Aj then year i+1 is Ak. Aj as the left side of the relationship is called the current state and Ak as the right side of the relationship is called the next state. And if there is a repetition of the relationship, it will still be counted once.

Classify the FLR that has been obtained from the 3rd stage into groups to form a Fuzzy Logical Relationship Group (FLRG) and combine the same relationships.

4. Defuzzification is a calculation process from the output forecasting results to then be calculated so as to get the results from the crisp numbers, then added to the actual data at the previous time so that the forecasting results are obtained. Forecasting values in the Fuzzy time series Chen model method there are several forecasting rules that must be considered, including:

Rule 1: If the result of the fuzzification in year t is Aj and there is a Fuzzy set that does not have a Fuzzy logic relationship, for example if Ai→ Ø, where the maximum value of the membership function of Ai is in the interval ui

and the middle value of ui is mi , then the forecasting result Ft+1 is mi.

Rule 2: If the fuzzification result for year t is ui and there is only one FLR in FLRG, for example if Ai → Aj where Ai and Aj are Fuzzy sets and the maximum value of the membership function of Aj is in the interval uj and the middle value of ui is mj , then the result of forecasting Ft+1 is mi .

Rule 3: If the result of fuzzification in year t is Aj and Aj has several FLRs in FLRG, for example Ai→Aj1, Aj2, ···, Ajk where Ai , Aj1, Aj2, ···, Ajk is the Fuzzy set and the maximum value of the membership function of Aj1, Aj2, ···, Ajp are in the intervals uj1, uj2, .., ujk and mj1, mj2, …, jk, then the forecasting results of Ft+1 are as follows (Chen, 1996):

$$mi = \frac{Batas\ Atas + Batas\ Bawah}{2}$$

Forecasting Method Accuracy

The time series model is said to be good if it is close to reality. This can be seen in the smaller error. The accuracy of the forecasting model can be calculated using the Mean Absolute Percentage Error (MAPE) (Shim, 2000) and the Root Mean of Square Error (RMSE) (Shim, 2000), each formula is as follows:according to Chang in the journal [4] the MAPE criteria are presented in Table 2.

Table 2. MAPE Criteria

MAPE	Information
< 10%	Forecasting ability is very good
10% - 20%	Good forecasting ability
20% - 50%	Forecasting ability is sufficient
> 50%	Poor forecasting ability

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} e_t^2}{n}}$$

Information:

e_t = $X_t - F_t$ = error in the t-period
X_t = historical data values in the t-period
F_t = forecast value in the t-period
PE_i = $\frac{X_t F_t}{X_t}(100)$
N = lots of observations

3. Research Methods
In analyzing the research data, the authors compiled the following steps:
1. Descriptive Statistical Analysis
2. Chen's Fuzzy Time series method

Analysis steps:
a. Determine the universe of discourse U (Universe of Discourse).
b. Determine the number and length of class intervals using average-based intervals.
c. Determine the defuzzification and define the Fuzzy set on U.
d. Doing fuzzification on the data on the number of visitors to the Fort Rotterdam Fort.
e. Classifying Fuzzy logic relationship (FLR)
f. Classifying FLR into groups to form a Fuzzy Logical Relationship Group (FLRG).

Carry out the defuzzification process and perform forecasting calculations on the number of visitors to Fort Rotterdam based on forecasting rules.

A. Results and Discussion
1. Description of Research Data
The data used is monthly data on the number of foreign tourist visitors in 2019 in West Sumatra. Plot the data on the number of visitors to get a graphical form as shown in Figure 1.

Figure 1. Graph of the Number of International Tourists

In the graph above it can be seen that the number of tourists in West Sumatra for one year from January to December has experienced ups and downs. From January to February, from 4052 tourists, it increased to 5155 tourists. And the highest number of tourists was in March, namely 6220 tourists. Meanwhile, the lowest number occurred in May 2019, namely 3,582 tourists.

Table 3. International Tourist Data

Time	2019 International Tourist Data	2019 International Tourist Data (%)
January	4052	6.66
Februay	5155	8.48
March	6220	10.23
April	5471	9
May	3582	5.89
June	5237	8.61
July	5197	8.55
August	5985	9.84
September	4435	7.29
October	5276	8.68
November	5021	8.26
December	5180	8.52

2. Use of the Fuzzy Time Series Method to Forecast the Number of International Tourists in West Sumatera

The steps for obtaining a forecasting model with Chen's Fuzzy Time Series method are as follows:

1. Forming the Universal Set
Determine universe of discussion U. Data on the number of foreign tourists visiting West Sumatra has the largest number is 6220 (10.23) and the smallest is 3582 (5.89). Based on the universe of discussion using minimum data and maximum data, all data has been included in all the intervals formed, but the number of classes to be formed is 10. So it is determined $Dmin = 5.89$ and $Dmax = 10.23$. and determined D1 = 0 and D2 = 5.29, so that the universe of discussion is obtained, namely:

U = Xmin-D1, Xmax + D2
U = [5.89 − 0, 10.23 + 5.89] = [5.89 ,15.52] .

2. Determination of the number of class intervals using average-based intervals.
The length of class 1 is known, so that it determines the number of intervals, namely (15.52 - 5.89)/1 = 9.63 rounded up to 10 classes.Established class

Table 4. Linguistic Interval

Linguistik Interval	
5.89 -6.89	U1
6.89 -7.89	U2
7.89 -8.89	U3
8.89 -9.89	U4
9.89 -10.89	U5
10.89 -11.89	U6
11.89 -12.89	U7
12.89 -13.89	U8
13.89 -14.89	U9
14.89 -15.89	U10

The fuzzy set is defined based on the equation:
A1 = 1/u1 + 0.5/u2 + 0/u3 +0/u4 +0/u5 +0/u6 +0/u7 +0/u8 +0/u9 +0/u10
A2 = 0.5/u1 + 1/u2 + 0.5/u3 +0/u4 +0/u5 +0/u6 +0/u7 +0/u8 +0/u9 +0/u10
A3 = 0/u1 + 0.5/u2 + 1/u3 +0.5/u4 +0/u5 +0/u6 +0/u7 +0/u8 +0/u9 +0/u10
A4 = 0/u1 + 0/u2 + 0.5/u3 +1/u4 +05/u5 +0/u6 +0/u7 +0/u8 +0/u9 +0/u10
A5 = 0/u1 + 0/u2 + 0/u3 +0.5/u4 +1/u5 +0.5/u6 +0/u7 +0/u8 +0/u9 +0/u10
A7 = 0/u1 + 0/u2 + 0/u3 +0/u4 +0/u5 +0.5/u6 +1/u7 +0.5/u8 +0/u9 +0/u10
A8 = 0/u1 + 0/u2 + 0/u3 +0/u4 +0/u5 +0/u6 +0.5/u7 +1/u8 +0.5/u9 +0/u10
A9 = 0/u1 + 0/u2 + 0/u3 +0/u4 +0/u5 +0/u6 +0/u7 +0.5/u8 +0/u9 +0.5/u10
A10 = 0/u1 + 0/u2 + 0/u3 +0/u4 +0/u5 +0/u6 +0/u7 +0/u8 +0.5/u9 +1/u10

3. Fuzzification
The fuzzification stage based on the effective intervals obtained can be determined linguistic values according to the number of intervals formed. The results of fuzzification of foreign tourist data in linguistic numbers can be seen in table 5.

Table 5. Determination of Linguistic ValueTabel

Time	Data	Linguistic Value
January	6.66	A1
February	8.48	A3
March	10.23	A5
April	9	A4
May	5.89	A1
June	8.61	A3
July	8.55	A3
August	9.84	A4
September	7.29	A2
October	8.68	A3
November	8.26	A3
December	8.52	A3

4.Determination of Fuzzy Logical Relationship (FLR)The formed FLR takes into account the fuzzy relationship Ai from month to month for 1≤i≤10.

Table 6. Fuzzy Logical Relationship (FLR)

Time	FLR
Jan - Feb	A1=> A3
Feb - Mar	A3=> A5
Mar - April	A5=> A4
April - May	A4=> A1
May – Juny	A1=> A3
Juny – July	A3=> A3
Juli – Agust	A3=> A4
Agust – Sept	A4=> A2
Sept – Oct	A2=> A3
Oct – Nov	A3=> A3
Nov – Des	A3=> A3

5. Form a Fuzzy Logical Relationship Group (FLRG)
FLRG is done by grouping fuzzy sets that have the same current state and then grouping them into one group in the next state based on Table 7.

Table 7. Fuzzy Logical Relationship Group (FLRG)

| Group	Relasi Logika Fuzzy
1	A1=> A3
2	A2=> A3
3	A3=> A3,A4,A5
4	A4=> A1,A2
5	A5=> A4
6	A6=> 0
7	A7=> 0
8	A8=> 0
9	A9=> 0
10	A10=> 0

6. Forecasting Value Defuzzification Process
There are two stages in the defuzzification process, namely first, finding the mean value of each interval based on the equation and second, calculating the forecast value based on 3 deffuzification rules. Thus, the results of defuzzification from FLRG are obtained in Table 8.

Table 8. Forecasting Determination

Time	Data	Linguistik Value	Interval Midpoint		Forecast
Jan2019	6.66	A1=> A3	8.39		
February	8.48	A3=> A3,A4,A5	8.39, 10.39	9.39	8.39
Maret	10.23	A5 =>A4	9.39		9.39
April	9	A4=>A1,A2	6.39, 7.39		9.39
May	5.89	A1 => A3	8.39		6.89
Juny	8.61	A3 => A3,A4,A5	8.39, 10.39	9.39	8.39
July	8.55	A3 => A3,A4,A5	8.39, 10.39	9.39	8.39
August	9.84	A4=>A1,A2	6.39, 7.39		8.39
September	7.29	A2=> A3	8.39		9.39
October	8.68	A3 => A3,A4,A5	8.39, 10.39	9.39	8.39
Nov	8.26	A3 => A3,A4,A5	8.39, 10.39	9.39	8.39
Des 2019	8.52	A3 => A3,A4,A5	8.39, 10.39	9.39	8.39
Jan 2020					8.39

Historical data plots and forecast data for foreign tourists in West Sumatra 2019 using Microsoft Excel.
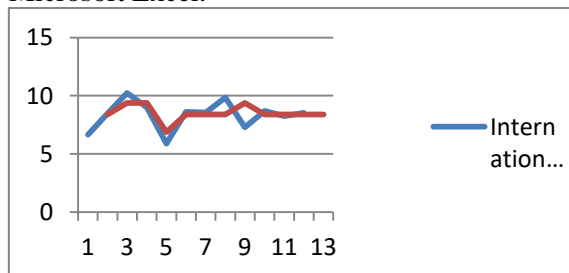
Figure 2. Graphical comparison of actual data and forecast values for foreign tourists in 2019

Based on Figure 2, the plot form of forecasting results for foreign tourists for each period has a value that is not much different from the historical data.

7. Determine the Forecasting Accuracy Level
Figure 2 presents a graphic comparison of actual data from January 2019 to December 2019 with forecast values. In general, it can be seen that the graphs almost coincide, which means that the forecast value is close to the actual value. The level of forecasting accuracy calculated using MAPE is obtained with many intervals of 10 , MAPE is 11.54% and RMSE is 1.184%.

CONCLUSION

Forecasting foreign tourist data using the Fuzzy Time Series produces a MAPE of 11.45% with the lowest error of 1,184%. These results indicate that the use of the Fuzzy Time Series in data forecasting of foreign tourist visits to West Sumatra has very good results in accordance with Table 2 regarding the MAPE criteria. This study has not used the factors that influence foreign tourist visits to West Sumatra such as the rupiah exchange rate, political and security conditions.

BIBLIOGRAPHY

[1] I. J. Thira, N. A. Mayangky, D. N. Kholifah, I. Balla, and W. Gata, "Peramalan Data Kunjungan Wisatawan Mancanegara ke Indonesia menggunakan Fuzzy Time Series," *J. Edukasi dan Penelit. Inform.*, vol. 5, no. 1, p. 18, 2019, doi: 10.26418/jp.v5i1.31074.

[2] V. M. Rizky and E. Suhartini, "Eksistensi Pariwisata Songa Adventure dan Perubahan Sosial Masyarakat Condong Kabupaten Probolinggo," *J. Entitas Sosiol.*, vol. II, pp. 1–9, 2014.

[3] M. Nawangwulan, "Pariwisata Indonesia Unggul di Budaya dan Alam," *bisnis tempo*, 2014. https://bisnis.tempo.co/read/558179/pariwisata-indonesia-unggul-di-budaya-dan-alam/full&view=ok.

[4] & M. N. Vivianti, Muhammad Kasim Aidid, "Metode Boostrap dan Jackknife dalam Mengestimasi Parameter Regresi Linear Ganda (Kasus: Data Kemiskinan Kota Makassar Tahun 2017)," *VARIANSI J. Stat. Its Appl. Teach. Res.*, vol. 1, no. 2, p. 32, 2019, doi: 10.35580/variansiunm12895.

[5] S. Nurdiati and F. Bukhari, "Perbandingan Metode Fuzzy Time Series Dan Holt Double Exponential Smoothing Pada Peramalan Jumlah Mahasiswa Baru," vol. 2, 2009.

[6] V. Virgianti and S. Martha, "Penerapan Fuzzy Time Series Chen Average Based Pada Peramalan Curah Hujan," *Bul. Ilm. Math. Stat. dan Ter.*, vol. 10, no. 4, pp. 485–494, 2021.