An improved forecasting method of frequency density partitioning (FDP) based on fuzzy time series (FTS)

B Irawanto1,*, R W Ningrum1,2,*, B Surarso1 and Farikhin1

1 Department of Mathematics, Faculty of Science and Mathematics, Diponegoro University, Semarang, Indonesia
2 Sekolah Tinggi Ilmu Agama Islam Walisembilan, Semarang, Indonesia

*Corresponding author: b_irawanto@yahoo.co.id

Abstract. FTS is popular in many recent years. Researchers are competing to outperform existing method by making new improvement including modifications at clustering step. Here we discuss about clustering process, i.e., partitioning based metric frequency density and firefly clustering algorithm. In the simulation, we compare the forecasting results and error value of the method with previous existing methods. The modifications give better forecasting results than previous methods indicated with smaller Root Means Errors (RMSEs) and Average Forecasting Error (AFER).

1. Introduction
Forecasting science has a vital role in human life. The data often used in forecasting is time series data. FTS is one of the soft computing methods that have been used and applied in time series data analysis. The difference between FTS and conventional time series is the value used in forecasting is not a crisp value but a linguistic value.

FTS-based forecasting methods have been developed in recent years [1-12]. Abdullah [1] presented a method to forecast exchange rate using distance-based FTS. Ningrum [2] has been presented combined Chen methods and firefly algorithm to forecast IHSG. Jilani [3] has been developing an approaching method based on frequency density partitioning.

In our other study, we have modified FDP forecasting method by repairing the definition of the universe interval. The results show that modified FDP gives better forecasting results than original FDP. Then in this occasion, we modify improved FDP by adding metric approaching step in FDP forecasting algorithm, then we compare this improvement or we called it Metric approaching method with Firefly-chen method [2].

The arrangement of this paper is organized as follows: Section 2, FTS, Section 3 forecasting results of both methods, Section 4, comparison forecasting results of the proposed method with the existing methods. The conclusions are discussed in Section 5.

2. Preliminaries
In this section, we discuss about the definition of FTS, where the linguistic values are represented by FTS. A fuzzy set \( A \) in the universe interval \( N \), \( N = \{u_1, u_2, u_3, \ldots, u_n\} \), can be represented by [4]:

\[
A = f_A(u_1)/u_1 + f_A(u_2)/u_2 + \cdots + f_A(u_n)/u_n
\]

(1)

with \( f_A \) is the member’s function of the fuzzy set \( A \). \( f_A: N \to [0,1] \), \( f_A(u_i) \) is membership degree of \( u_i \).

Definition 2.1 [4]

Let \( N(t) \), a subset of \( R^1 \), be the universe interval on which fuzzy sets \( f_i(t) \) and let \( P(t) \) be a collection of \( f_1(t), f_2(t), \ldots \). Then, \( P(t) \) is called a FTS defined on \( N(t) \).
3. Comparing the Methods

Figure 1 shows the flowchart of Metric Approach forecasting method.

![Flowchart of Metric Approach Forecasting Method](image)

**Figure 1.** Flowchart of metric approach forecasting method

The advantages of Metric approaching based frequency density method is the universe interval is divide into some interval then it is evaluated by its frequency of occurrence data at its interval. Top 3 of frequencies are divided again by mean-based partitioning to be sub-intervals. The new intervals are evaluated again by counting its frequencies. The interval with zero frequency is eliminated.

While at Firefly-Chen method [2], the number of intervals has been determined at the initiation stage. However, it is undeniable that Metric approaching method has deficiency, this algorithm will be very troublesome if calculated manually, especially if the amount of historical data is very large. In the other hand, Firefly algorithm is an automatic clustering algorithm. No matter how many clusters are used, Firefly algorithm can automatically update the best interval length.

4. Results and Discussion

In the following, we apply the methods to forecast the monthly data of IHSG. In the Table 1, we compare the forecasting results of both methods, i.e., Firefly-Chen method and Metric approaching method. The Metric approaching method shows smallest value of RMSE and AFER. That is, the fuzzy metric approaching method-based frequency density partitioning can get a better forecasting result.
Table 1. The comparison of IHSG forecasting results

| Year | Month | Actual Data | FF-Chen Method | Metric Approaching Method |
|------|-------|-------------|----------------|--------------------------|
| 2015 | 1     | 5289.404    | 5275.023632    |                          |
|      | 2     | 5450.294    | 5454.845495    |                          |
|      | 3     | 5518.675    | 5512.639634    |                          |
|      | 4     | 5086.425    | 5084.81757     |                          |
|      | 5     | 5216.379    | 5222.11098     |                          |
|      | 6     | 4910.658    | 4923.691154    |                          |
|      | 7     | 4802.529    | 4812.3139      |                          |
|      | 8     | 4509.607    | 4532.916208    |                          |
|      | 9     | 4223.908    | 4312.837711    |                          |
|      | 10    | 4455.18     | 4421.622488    |                          |
|      | 11    | 4446.458    | 4421.622488    |                          |
|      | 12    | 4593.008    | 4617.332292    |                          |
| 2016 | 1     | 4615.163    | 4617.332292    |                          |
|      | 2     | 4770.956    | 4711.408892    |                          |
|      | 3     | 4845.371    | 4851.429947    |                          |
|      | 4     | 4838.583    | 4851.429947    |                          |
|      | 5     | 4796.869    | 4778.824439    |                          |
|      | 6     | 5016.647    | 5011.70863     |                          |
|      | 7     | 5215.994    | 5222.11098     |                          |
|      | 8     | 5386.082    | 5391.867322    |                          |
|      | 9     | 5364.804    | 5357.843731    |                          |
|      | 10    | 5422.542    | 5420.967346    |                          |
|      | 11    | 5148.91     | 5151.336372    |                          |
|      | 12    | 5296.711    | 5315.772625    |                          |
| 2017 | 1     | 5302.662    | 5315.772625    |                          |
|      | 2     | 5386.692    | 5391.867323    |                          |
|      | 3     | 5568.106    | 5593.784927    |                          |
|      | 4     | 5685.298    | 5666.855684    |                          |
|      | 5     | 5738.155    | 5733.390726    |                          |
|      | 6     | 5829.708    | 5806.141048    |                          |
|      | 7     | 5840.939    | 5806.141048    |                          |
|      | 8     | 5864.059    | 5861.068003    |                          |
|      | 9     | 5900.854    | 5903.402094    |                          |
|      | 10    | 6005.784    | 6087.186576    |                          |
|      | 11    | 5952.138    | 5987.871941    |                          |
|      | 12    | 6355.654    | 6250.456533    |                          |
| RMSE |       | 106.317     | 67.048         |                          |
| AFER |       | 1.2917%     | 0.0935%        |                          |

The error value show that the Metric Approaching method give the better results to forecast the IHSG with RMSE value 67.048 and AFER value 0.0935%. Here is the forecasting result’s graphic of both methods.
In Figure 2, the Metric Approaching method looks very close to the actual data. Only at some points in September 2015 and December 2017, the forecasting value is not approach the actual data. In general, we said that at case IHSG forecasting, Metric approaching method has best results than FF-Chen method. It happened because the proposed method is reducing some zero intervals, i.e., intervals which do not have actual data, or intervals with zero frequencies. If there is no interval with zero frequencies, it can optimize the number of clusters so that the forecast gives better results [3].

While FF-Chen method forecasting results is not too approaching to actual data for almost all points. At this case, we used 31 intervals at clustering process. FF-Chen method will give best results if we used more clusters, bigger than 31 intervals. It is because if we increase the number of clusters, the length of intervals will be smaller, so the forecast will get closer to the actual data [2]. While how many effective clusters should be used to obtain the best results in forecasting will be discussed in the next study.

5. Conclusion
In this paper, we have comparing Fuzzy metric approaching method and Firefly-Chen method to forecast historical data of monthly IHSG. At clustering process, both methods use same number of intervals. At Fuzzy metric approaching method, in the beginning the universe interval is partitioned to be 13 main intervals, then top 3 frequency values are divided again into sub-intervals by mean-based partitioning algorithm to be 31 intervals. Then it is evaluated again by counting each frequency. The interval with zero frequencies is discarded. Its effect is the interval has effective interval widths, while at Firefly-chen method is still possible to have intervals with zero frequencies. In Figure 2, the forecasting result of Fuzzy metric approaching method is closer to the actual value and it gives better forecasting error value of RMSE and AFER.

References
[1] Abdullah L 2013 Int. J. of Engineering and Technology (IJET) 5 452
[2] Ningrum R W, Surarso B, Farikhin and Safarudin Y M 2018 IOP Conf. Series : J. of Phys. Conf. Series 983 0120055
[3] Jilani T A, Burney S M A and Ardil C 2010 Int. J. of Computer and Inf. Engineering 4 1194
[4] Cheng S H, Chen S M and Jian W S 2016 *Inf. Sci.* **327** 272
[5] Tsaur R C 2011 *Int. J. of Innovative Computing Inf. and Control* **8** 4931
[6] Cheng C H and Yang J H 2018 *Neurocomputing* **302** 33
[7] Bisht K and Kumar S 2016 *Expert System with Applications* **64** 557
[8] Torbat S, Khaskei M and Bijari M 2018 *Economic Analysis and Policy* **58** 22
[9] Chen M Y and Chen B T 2015 *Inf. Sci.* **294** 227
[10] Martinez F, Frias M P, Perez-Godoy M D and Rivera A J 2018 *Expert Systems with Applications* **103** 38
[11] Tan L, Wang S and Wang K 2017 *Inf. Processing Letters* **127** 32
[12] Wei L Y 2016 *Applied Soft Computing* **42** 368