Journal of Education and Science (ISSN 1812-125X), Vol: 31, No: 02, 2022 (42-54)

The New Approach Optimization Markov Weighted Fuzzy Time Series Using Particle Swarm Algorithm

Nauval Satriani Siregar[1], Sugiyarto Surono[2*]

[1*,2] Department of Mathematics, Ahmad Dahlan University, Yogyakarta, Indonesia

E-mail: [1]nauval1800015003@webmail.uad.ac.id, [2*]sugiyarto@math.uad.ac.id

(Received February 16, 2022; Accepted March 29, 2022; Available online June 01, 2022)

DOI: 10.33899/edusj.2022.133052.1217, © 2022, College of Education for Pure Science, University of Mosul.
This is an open access article under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/)

ABSTRACT

Markov Weighted Fuzzy Time Series is a forecasting method that applies fuzzy logic to form linguistic variables from existing data. The formation of linguistic variables makes it possible for the forecasting process to be more accurate by considering the uncertainty aspect in decision-making. Its formation is started by grouping the data into a certain number of clusters. The next steps are fuzzification, transition matrix formation, and defuzzification for forecasting. In the process of grouping, the existing data will be grouped into several clusters so that it results in the interval length of each cluster. One of the problems of this grouping is the absence of a base standard in the clustering process so it is prone to have a different value in forecasting accuracy. The difference in the number of the class or interval length will result in different accuracy even though the clustering method that is used is the same. In this study, the author proposes the idea of using Particle Swarm Optimization to improve the interval length. The initial interval that is already obtained through the K-means clustering algorithm will be evaluated using the Particle Swarm Optimization method so that it will have a new interval that later will be used in the fuzzification process and forecasting. The accuracy of forecasting can be calculated by using Mean Absolute Percentage Error from Markov Weighted Fuzzy Time Series conventional method and Markov Weighted Fuzzy Time Series method with Particle Swarm Optimization. The result of this study gives an improvement in error value from 8.03% to 5.88%.

Keywords: Fuzzy; Optimization; Markov; Forecasting

الاسلوب الجديد لتحسين نهج ماركوف الامثل(المرجح) لسلاسل زمنية ضبابية باستخدام خوارزمية سرب الجسيمات

Nauval Satriani Siregar[1], Sugiyarto Surono[2*]

[1*,2] Department of Mathematics, Ahmad Dahlan University, Yogyakarta, Indonesia

الملخص

سلسلة ماركوف الوزنية الضبابية الزمنية هي طريقة تنبؤ تطبق المنطق الضبابي لتشكيل متغيرات لغوية من البيانات المعتمدة(الموجودة). تكوين المتغيرات اللغوية يجعل عملية التنبؤ اكثر دقة من خلال خاصية عدم اليقيين في عملية صنع القرار. والخطوة الاولى هي تشكيل المتغيرات من خلال تجميع البيانات في عدد من المجاميع. والخطوة الثانية هي التشويش وتشكيل مصفوفة الانتقال ثم الغاء التشويش للتنبؤ. يتم تجميع البيانات الموجودة التي في عدة مجموعات بحيث ينتجعنها طول الفاصل الزمني لكل مجموعة ، وتتمثل احدى مشكلات التجميع في عدم وجود معيار اساسي في عملية التجميع. لذا فهو يعطي قيم مختلفة في دقة التنبؤ. سيؤدي الاختلاف في عدد الفئات وطول الفترة الزمنية الى دقة مختلفة بالرغم من ان الطريقة هي نفسها. في هذه الدراسة يقترح الباحث فكرة استخدام خوارزمية سرب الجسيمات (particle swarm operation) لتحسين طول الفاصل الزمني. سيتم تقييم الفاصل الزمني الاول الذي تم الحصول عليه من خوارزمية التجميع (K-mean) باستخدام خوارزمية سرب الجسيمات حيث يكون لها فاصل زمني جديد يتم استخدامها لاحقا في عملية التشويش (fuzzification) والتنبؤ. يمكن حساب دقة التنبؤ باستخدام متوسط النسبة المئوية للخطا المطلق من طريقة ماركوف الوزنية الضبابية الزمنية التقليدية وطريقة تحسين سرب الجسيمات . اظهرت نتيجة الدراسة تجسنا في قيمة الخطا من %8.03 الى %5.88.

الكلمات المفتاحية: الضبابية ، الامثلية (المرجح) ، سلاسل ماركوف ، التنبؤ(التوقع)

1. Introduction

Air pollution is currently a concern for the public and the government, especially in big cities which are industrial centers. Air pollution can have a serious impact on human activities, especially on health so it will affect industrial activities. Some of the particles used as reference measurements are PM10 (particles measuring <10 microns), ozone level, Carbon Monoxide, Nitrogen Oxide, and Sulfur Dioxide. Thus, air quality forecasting becomes one of the important topics considering its relationship with human health effects [1][2].

Forecasting is a method to predict the future based on relevant data from the past. This method applies an ordered and direct work step which results in better analysis. By resulting in better analysis, it is expected to have a higher level of belief and certainty over obtained results since the deviation or error can be tested [3]. Forecasting can be divided into two types based on its type of data, qualitative method and quantitative method [4]. The qualitative method is a method that considers expert opinion over the mathematic model. This method is highly subjective, so it often appears less scientific. On the other hand, the quantitative method is a method that consists of raw data that will be processed mathematically to predict the future outcome. Several examples of the quantitative method are the Regression method [5], Econometric method [6], and Time Series method [7].

The time series method is a statistical method that uses past data from a certain period. This method assumes that what happened in the past will occur again in the future. This method is deeply related to a factor of time [8]. The data of the time period varies between days, weeks, months, or years. One of the methods of time series is the Markov Chain method that can predict the future variable value based on present variable value by using a transition matrix [9].

The developed version of the time series method is the Fuzzy Time Series method which applies fuzzy logic to data time series classification. The Fuzzy Time Series method is a method introduced by Song and Chissom in 1993 with the main concept of predicting the problem of a dataset consisting of linguistic data [10]. The upper hand of the Fuzzy Time Series method is the ability of fuzzy logic to predict linguistic data with better precision where the data could not be calculated using the conventional time series method. Aside from that, the Fuzzy Time Series method offers an easier calculation process in comparison with neural networks [7].

In the study [2], the authors combine the Markov forecasting method with Weighted Fuzzy Time Series to predict air pollution in Malaysia. The result showed a satisfying result with an error value of 1.06. The sole difference between Markov Chain Fuzzy Time Series method with Markov Weighted Fuzzy Time Series lies in the calculation of forecasting value, in which Markov Chain Fuzzy Time Series uses adjustment value while Markov Weighted Fuzzy Time Series method uses the calculation between defuzzification matrix with matrix weight [2].

The study [2] explains that one of the developments for the following research can use another optimization method like Particle Swarm Optimization, Genetic Algorithms, or Neural Networks. Based on that, the author of this paper uses the Particle Swarm Optimization method as an optimization method in determining the amount of its interval. The author assumes that by using the Particle Swarm Optimization method, we will be able to show optimal value to determine the among of partition interval that is needed so that it can repair both the value of Fuzzy Logical Relationship Group (FLRG) and the forecasting result [11][12]. PSO will work by fixing the interval formed by the clustering method. These improvements will give better MAPE results. The PSO process will take place iteratively until it reaches the stopping criteria so that the error value will decrease.

Section 2 is the method used in this research which is obtained from the literature which is presented sequentially according to the working algorithm. Section 3 is the results and discussion. In this section, a table of the dataset used along with the calculation results is shown. The results of the cluster division, the calculation of the lower and upper bounds, the determination of PSO particles, and the graph of the prediction results are also displayed. Finally, section 4 contains conclusions and future work perspectives.

2. Method

The method that is used in this research is the Markov Weighted Fuzzy Time Series method that will be optimized using the Particle Swarm Optimization method. The further details will be explained as follows :

2.1 K-Means

Partition method C-means or K-means is a method of clustering that uses the shortest distance of data point into cluster centre. The amount of k cluster can be determined according to the need of the researcher and can be determined using another method like Elbow [13]. The cluster center will overgo an updating process during iteration. The initial value of the cluster centre is decided randomly, but after going through iteration the following formula will be used [14]:

$$u_k = \frac{1}{n}\sum_{i=0}^{n_i} x_i \tag{1}$$

with u_k as cluster center, n is the amount of internal data, n_i is the amount of data in the cluster-i, and x_i is the data in cluster-i.

2.2 Fuzzy Sets

If X is a group of objects symbolized as x, the fuzzy set U in X is expressed as ordered pair set

$$U = \{(x, \mu_U(x)|x \in X\} \tag{2}$$

where $\mu_U(x)$ is the membership function x in fuzzy set U that is located between interval of [0,1].

2.3 Membership Function

Functions that can be used in representing fuzzy sets are sigmoid, gaussian, and pi. Informing membership function there are several types, this study uses triangular graph membership function because triangular function resulted in a more optimal system compared to the other types [15]:

$$\mu(x) = \begin{cases} \frac{(x-a)}{(b-a)} & a \le x \le b \\ \frac{(c-x)}{(c-b)} & b \le x \le c \\ 0 & other \end{cases} \tag{3}$$

2.4 Fuzzy Time Series

Fuzzy Time Series is a method of data forecasting that applies fuzzy principles as its basis. Fuzzy time series is often used in solving forecasting problems where the past data is linguistic values. The basic steps to design a model of fuzzy series are defining the universe of discourse U, dividing U into several intervals, fuzzification, defining the fuzzy logic relationship, determining the prediction value, and defuzzification [2][16].

Assuming U is a universe of discourse so that $U = \{u_1, u_2, \dots, u_n\}$ where $u_1 = 1,2,3, \dots, n$ is the linguistic value that might be from U. Then, fuzzy set of linguistic variable A_i from U can be defined as:

$$A_i = \left\{\frac{f_{A_i}(u_1)}{u_1} + \frac{f_{A_i}(u_2)}{u_2} + \dots + \frac{f_{A_i}(u_n)}{u_n}\right\} \quad (4)$$

Generally, the steps of Fuzzy Time Series include: (1) determining universe of discourse from past data, (2) defining fuzzy set A_i and fuzzification, (3) determining Fuzzy Logical Relationship, (4) classifying Fuzzy Logical Relationship so that Fuzzy Logical Relationship Group is formed, (5) calculating the forecasting value.

Fuzzification formation is done by calculating the Lower Bound and Upper Bound of the formed partition. Partition formation is done by sortation process ascendingly, then determining the value of cluster center using equation (5) and calculating the value of lower and Upper Bound [17]. d_j is a data set in the same cluster and r is the total number of data in the same cluster.

$$cluster\ center_m = \frac{\sum_{j=1}^{r} d_j}{r} \quad (5)$$

$$cluster\ ubound_m = \frac{cluster\ center_m + cluster\ center_{m+1}}{2} \quad (6)$$

$$cluster\ lbound_{m+1} = cluster\ ubound_m \quad (7)$$

Since the values of the prior first cluster and the terminal value after the last cluster are unknown, the calculation of Lower Bound from cluster lbound_1 for the first cluster and the calculation of cluster ubound_k for the last cluster is done by abiding the following rules:

$$cluster\ ubound_k = cluster\ center_k + (cluster\ center_k - cluster\ lbound_k) \quad (8)$$

$$luster\ lbound_1 = cluster\ center_1 - (cluster\ ubound_1 - cluster\ center_1) \quad (9)$$

2.5 Weighted Fuzzy Time Series

Weight fuzzy time series method is based on the weight that is a representation of repeating fuzzy logical relationship (FLR), where those weights will influence the forecasting result [18].

A is a mapping of F from

$$I^n = I \quad (10)$$

Where= $[0,1]$ can be defined as Ordered Weight Aggregation (OWA) operation of dimension n with F as a function, so the weight vector W can be written as:

$$W = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \quad (11)$$

The weight determination of fuzzy logic relationship (FLRs) can be calculated by assuming $A_i \rightarrow A_i, A_j, A_k, A_l$ with $i, j, k, l = 1,2,3, \dots, p$ which are FLR and A_i has n_1 relation with its own self, n_2 relationship with A_j, n_3 relationship with A_k, and n_4 relationship with A_l ($n_1, n_2, n_3, n_4 \in N$) so n_1, n_2, n_3, n_4 is the repetition of FLRs so the weight value can be written as follows:

R = repetition of FLRS = $\{n_1, n_2, n_3, n_4\}$

n = $n_1 + n_2 + n_3 + n_4$ = total of FLRs

$$w_m = \frac{repetition\ of\ A_i}{total\ of\ FLRs} \quad (12)$$

With

$$w_1 = \frac{n_1}{n}, w_2 = \frac{n_2}{n}, w_3 = \frac{n_3}{n}, w_4 = \frac{n_4}{n} \quad (13)$$

So

$$\sum_{m=1}^{4} w_m = \frac{n_1}{n} + \frac{n_2}{n} + \frac{n_3}{n} + \frac{n_4}{n} = 1 \quad (14)$$

with $m = 1,2,3, \dots, q$ and $q < n$. Generally, the above equation can be written as follows:

$$\sum_{m=1}^{q} w_m = \frac{n_1}{n} + \frac{n_2}{n} + \frac{n_3}{n} + \cdots + \frac{n_q}{n} = 1 \quad (15)$$

$$\sum_{m=1}^{q} w_m = \frac{n}{n} = 1 \quad (16)$$

2.6 Markov Rules

Markov analysis is a forecasting method introduced by Andrei A. Marcov, a Russian mathematician in 1907. Markov analysis is a method that analyzes the behavior of several variables in the present time with the purpose of predicting the behavior of said variables in the future. Markov analysis is also known as Markov Chain because of its chain-like character. Markov Chain is able to explain the future condition of the variable based on its condition in the present [19][20].

By using Fuzzy Logical Relationship (FLR), we are able to determine the probability of the future and make a Markov transition probability matrix. If A_i is transitioned into A_j with random $j = 1,2, \dots, n$ the Fuzzy Logical Relationship (FLRG) will be formed. Its transition probability can be written as follows

$$P_{ij} = \frac{M_{ij}}{Mi}, \; i,j = 1,2, \dots, n \quad (17)$$

with P_{ij} is the transition probability value from A_i to A_j, M_{ij} is the amount of transition from A_i to A_j, and Mi is the amount of data in A_i. So the transition probability matrix can be written as follows [21]:

$$R = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{bmatrix} \quad (18)$$

In particular, the calculation of the forecast value has two rules, the one-to-one case and the one-to-many case.

Case 1. If the FLRG of A_i is one-to-one where there is only one transition from A_i to A_k, where A_i and A_k are fuzzified linguistic variables, then the forecast value of $F(t)$ is m_k (the middle value of the lower and upper bounds cluster-i. Mathematically it can be written in the following equation

$$F(t+1) = m_k P_{ik} = m_k \quad (19)$$

Case 2. If the FLRG of A_i is one-to-many, where there is more than one transition from A_i to $A_1, A_2, \dots, A_n$ with $i = 1,2, \dots, n$. So if state A_i is observed in $Y(t)$ at time t, then the forecast value $F(t+1)$ can be calculated using the following equation

$$F(t+1) = m_1 p_{i1} + m_1 p_{i2} + \cdots + m_{i-1} p_{i(i-1)} + Y(t) p_{ii} + m_{i+1} p_{i(i+1)} + \cdots + m_n p_{in} \quad (20)$$

where $m_1, m_2, \dots, m_n$ is the middle value from the lower bound and the upper bound to cluster-n. m_i is a variable that will be replaced with $Y(t)$ at state A_i to get a more accurate value.

After obtaining the value of $F(t+1)$, it can be calculated the value of the forecast

$$\hat{F}(t+1) = F(t+1) \pm |diff(Y(t), m_i)| \quad (21)$$

2.7 Particle Swarm Optimization

Particle Swarm Optimization (PSO) is an optimization method that is inspired by the social behavior of a flock of birds and a group of fish in finding solutions from a nonlinear problem. The individual unit of the group is called a particle where every particle has its influence on its group (swarm). So if one particle has found the fastest route to get food, the other particles in the same group will follow its trail even though they are far from each other [22].

It can be concluded that 2 types of learning are experienced by the particle. Every particle will learn from its own experience during the movement and the experience of the other particles. The type of learning where the particle learns from its own experience is called cognitive learning, while the one where it is obtained from other particles is called social learning. The result of cognitive learning is that particles will have the memory of the best solution which is symbolized as P_{best}. While for social learning, the best result is symbolized as G_{best} [23]. The movement of particles into new point can be calculated using the following formula:

$$X_{id}^{t+1} = X_{id}^{t} + V_{id}^{t+1} \quad (22)$$

Where

$$V_{id}^{t+1} = \omega^t \times V_{id}^t + c_1 \times Rand(\,) \times \left(P_{best_{id}} - X_{id}^t\right) + c_2 \times Rand(\,) \times \left(G_{best} - X_{id}^t\right) \quad (23)$$

With

$$\omega^t = \omega_{max} - \frac{t \times (\omega_{max} - \omega_{min})}{iter\ max} \quad (24)$$

2.8 Mean Absolute Percentage Error

To evaluate the performance of a model identification of forecasting accuracy is needed so the best model is obtained. The forecasting needs to have a small error value. There are several methods to validate a model. One of those is the Mean Absolute Percentage Error (MAPE). The smaller its validation value, the more accurate the forecasting technique and it also applies otherwise. The result of forecasting can be classified as good if it has a value of MAPE less than 10% and its ability in forecasting is classified as good when its MAPE value is less than 20% [24].

$$MAPE = \left(\frac{100\%}{n} \sum_{t=1}^{n} \frac{|X_t - \widehat{X_t}|}{X_t}\right) \quad (25)$$

where n is the amount of data, X_t is the real observation data at time t, and $\widehat{X_t}$ is the forecasting value at period t [25].

2.9 Proposed Model and Algorithm

The MWFTS calculation algorithm follows the steps described in [2]. In this research, MWFTS with PSO optimization is introduced. The calculation algorithm is described as follows:

Step 1. Data preprocessing. Import data, fix missing data, fix outliers, and visualize data to be used.

Step 2. Calculation of k-means clustering

Stage 1. Determine the desired number of clusters along with their cluster centers randomly, and the maximum number of iterations.

Stage 2. Calculate the distance of each point to the center of the cluster using Euclid's formula. The point with the closest distance will be grouped in the same cluster

Stage 3. Determine the center of the new cluster using equation (1)

Stage 4. Repeat Stage 2 and Stage 3 until it reaches the maximum iteration or until there is no change in cluster members

Step 3. After forming the cluster, it is possible to calculate the cluster center, lower bound, and upper bound using equations (5), (6), and (7) respectively. The lower and upper bound values will be used as determinants of linguistic variables in the fuzzification of equation (4)

Step 4. Forming FLR and FLRG

Step 5. Formation of the Markov transition probability matrix based on the FLRG that was formed in Step 4 using equations (17) and (18)

Step 6. Calculation of weights using equation (11) - (16)

Step 7. Calculation of the forecast value $F(t+1)$ according to the existing case. Case 1 can use equation (19) and case 2 can use equation (20)

Step 8. Calculation of the $\hat{F}(t+1)$ forecast value based on step 7 using equation (21)

Step 9. Calculation of the error value based on equation (24)

Step 10. The error value (MAPE) obtained is then used as P_{best} and G_{best} to calculate PSO optimization. PSO calculation will be done iteratively by fixing the lower bound and upper bound that have been formed previously. In equation (24), the value of w is obtained from equation (23). Initial V_{id}^t values are determined randomly, where the value of X_{id}^t is obtained from the previous interval. The values of c_1 and c_2 are 2. Random values are determined randomly in the interval [0,1]. Then a new V_{id}^{t+1} is obtained to be used to calculate the new X_{id}^{t+1} in equation (22). The value of X_{id}^{t+1} is then sorted in ascending order to obtain a new lower bound and upper bound so that a new error value is obtained. This process is carried out iteratively until the desired error limit or has met the maximum number of iterations.

3. Results and Discussion

The data that is used in this study is index data of air pollution in Central Singapore between January 2021 until April 2021. The dataset is achieved from the Air Quality Historical Data Platform website (Air Quality Historical Data Platform (aqicn.org)).

Table 1. The Dataset

Date	pm25	pm10	O_3	SO_2	CO	maximum
01/01/2021	22	17	9	-	4	22
02/01/2021	19	16	3	-	-	19
03/01/2021	26	18	6	2	3	26
04/01/2021	28	19	10	2	3	28
...	...	...	...	...	...	...
27/04/2021	71	37	19	1	4	71
28/04/2021	69	32	13	1	4	69
29/04/2021	52	33	15	1	4	52
30/04/2021	64	32	23	1	4	64

Table 1 showsthe dataset. The variables from the dataset are pm10 (air particle with the size of 10 micrometers, pm25 (a particle with the size of 2.5 micrometers), SO2 (sulfide concentration), CO (carbon monoxide concentration), O3 (ozone concentration), and max (the highest concentration value from measured variable). The data processing is done using Jupyter Notebook using Python language program. The max variable is defined as the maximum value from measured variables like pm10, pm25, SO2, CO, dan O3.

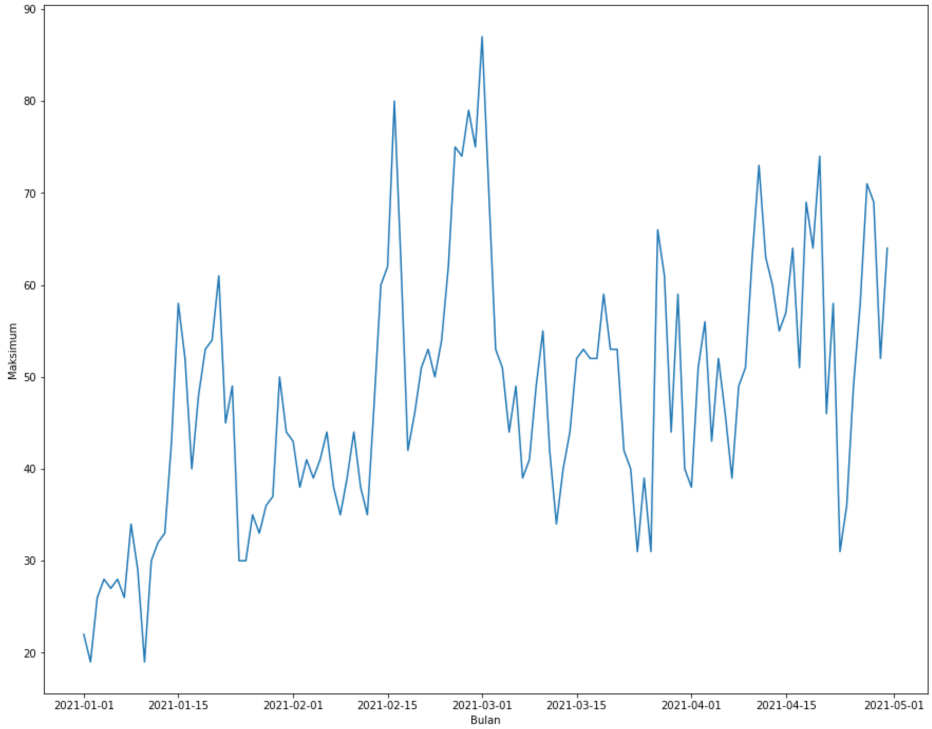

Figure 1. Plot of time series data

Figure 1 shows the graph of time series from the data that later will be used and analyzed to obtain its forecasting model.

The first step in the calculation of forecasting is by forming partition intervals. It then will be used in the fuzzification process. Interval is formed using the clustering method of k-means.

Table 2. Cluster Data 1

Date	Maximum	Cluster
2021-02-16	80	0
2021-02-25	75	0
2021-02-26	74	0
...	...	...
2021-04-20	74	0
2021-04-27	71	0
2021-04-28	69	0

Table 3. Cluster Data 2

Date	Maximum	Cluster
2021-01-14	43	1
2021-01-17	40	1
2021-01-22	45	1
...	...	...
2021-04-07	39	1
2021-04-21	46	1
2021-04-24	36	1

Table 4. Cluster Data 3

Date	Maximum	Cluster
2021-01-16	52	2
2021-01-18	48	2
2021-01-19	53	2
...	...	...
2021-04-17	51	2
2021-04-25	49	2
2021-04-29	52	2

Table 5. Cluster Data 4

Date	Maximum	Cluster
2021-01-01	22	3
2021-01-02	19	3
2021-01-03	26	3
...	...	...
2021-03-24	31	3
2021-03-26	31	3
2021-04-23	31	3

Table 6. Cluster Data 5

Date	Maximum	Cluster
2021-01-15	58	4
2021-01-21	61	4
2021-02-14	60	4
...	...	...
2021-04-22	58	4
2021-04-26	58	4
2021-04-30	64	4

The clustered data will be divided into groups based on its cluster. The grouping process is to calculate the Lower Bound and the Upper Bound of each partition interval (cluster) by calculating the cluster center. The grouping of the formed clusters are shown in Table 2 until Table 6. Table 2 shows data grouped into cluster 1. table 3 shows data grouped into cluster 2. until table 6 shows data grouped into cluster 5.

Table 7. Labelling of each Cluster

Cluster	Cluster Center	Label
Cluster 4	29.478261	Extremely good
Cluster 2	41.171429	Good
Cluster 3	51.580645	Enough
Cluster 5	61.105263	Bad
Cluster	74.666667	Extremely bad

Cluster center value is obtained using the equation (8). The result is then written in order ascendingly like how it's shown in Table 7 for later will be given a label.

Table 8. Bound of each interval based on k-means cluster

	Particle 1	
	Lower Bound	**Upper Bound**
Interval 1	23.632	35.325
Interval 2	35.325	46.376
Interval 3	46.376	56.343
Interval 4	56.343	67.886
Interval 5	67.886	81.447

Table 8 shows in detail the interval bound of particles that are formed by the clustering of k-means. The bounds of the interval will be used as a reference for fuzzification and forecasting calculation. Based on forecasting evaluation using MAPE, particle 1 has an error value of 10.9771257%. By using the PSO algorithm, the author can lessen the error value.

Table 9. Random particle

	Particle 1		Particle 2		Particle 3		Particle 4		Particle 5	
	Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**
Interval 1	23.632	35.325	23.632	29.627	23.632	26.424	23.632	25.527	23.632	27.911
Interval 2	35.325	46.376	29.627	41.772	26.424	45.148	25.527	32.925	27.911	32.930
Interval 3	46.376	56.343	41.772	66.573	45.148	53.469	32.925	51.010	32.930	60.206
Interval 4	56.343	67.886	66.573	72.920	53.469	79.934	51.010	69.707	60.206	74.347
Interval 5	67.886	81.447	72.920	81.447	79.934	81.447	69.707	81.447	74.347	81.447

The first step of the PSO method application is by using random particles. Its size is in the range of Lower Bound and Upper Bound from k-means calculation. Table 9 shows the value of the random particle that is used.

Based on those particles, fuzzification calculation and forecasting evaluation using MAPE can be done accordingly. Table 10 shows the evaluation of the MAPE calculation of each particle. Particle 1 has shown the smallest evaluation value that is ~8.03%. Figure 2 shows the plot between actual data and its forecasting value.

Table 10. MAPE Value of each Particle

	Particle 1	Particle 2	Particle 3	Particle 4	Particle 5
MAPE	10.971257	9.405541	13.434192	11.133078	8.031329

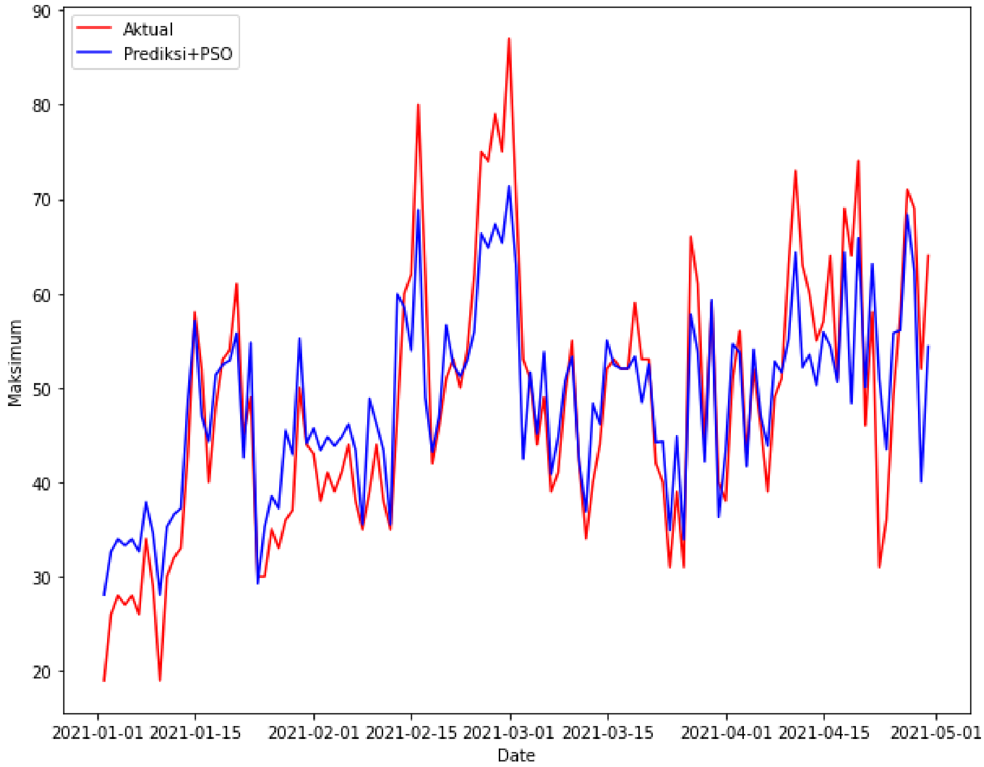

Figure 2. Graph between actual data and its prediction result

Then the calculation of velocity and the update of particle value is done by equations (21) and (20). This step is done over and over iteratively until the determined stopping criteria are fulfilled (maximum iteration or the gap of G_{best} change).

Table 11. MAPE value with PSO

	Particle 1	Particle 2	Particle 3	Particle 4	Particle 5
MAPE	5.880359	5.880359	5.904064	5.748465	5.880359

The calculation result is shown in TABLE 11 which are the final value of optimization process. It can be seen that all particle is directed into one minimum point that is ~5.88 %.

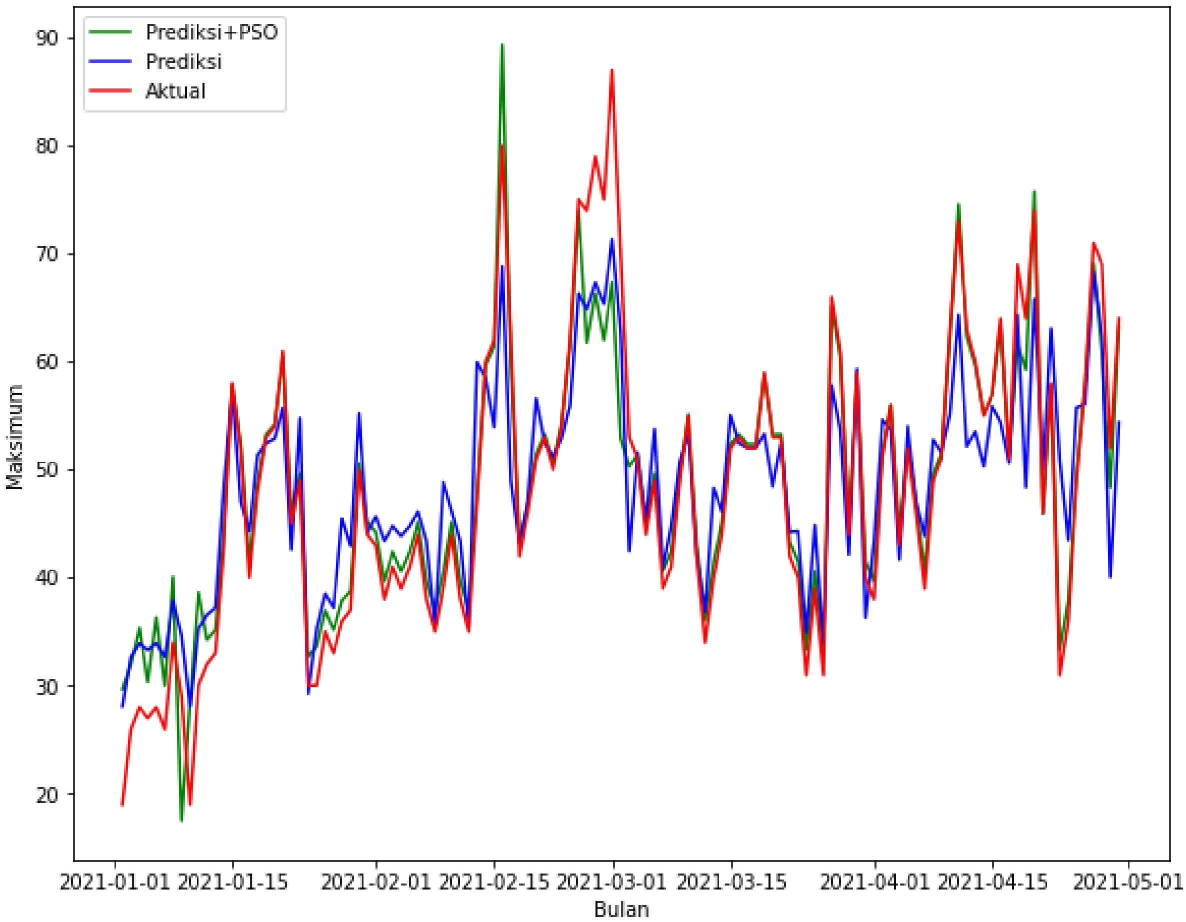

Figure 3. Graph between actual data, prediction, and prediction using PSO

Figure 3 shows the comparison of actual data, prediction value, and prediction value using the PSO algorithm. The prediction value using the PSO algorithm has a graph that is closer to the actual value.

4. Conclusion

This study proposes a Markov Weighted Fuzzy Time Series forecasting model using Particle Swarm Optimization to predict air quality or time-series data. The partition method used is k-means clustering. As far as the author knows, there is still little research related to the use of the partition method in the Markov Weighted Fuzzy Time Series model. In addition, research related to the use of Particle Swarm Optimization to optimize the Markov Weighted Fuzzy Time Series forecasting model is also still rarely found. The use of Particle Swarm Optimization will improve the interval value obtained by the k-means clustering partition. For air quality forecasting, the proposed model provides better accuracy results with Particle Swarm Optimization. In addition, the proposed model can also be used in various types of time series data. The result of forecasting using the Markov Weighted Fuzzy Time Series method and k-means clustering has an initial result of 8.03%. After going through the Particle Swarm Optimization method of optimization, the value got better with 5.88%. Based on this research, the Particel Swarm Optimization method of optimization is proven to improve the interval length and resulting in a smaller error value.

For future studies, other partitioning methods can be used to provide options for the intervals that can be formed. In addition, other optimization methods such as Neural Network, Genetic Algorithm, or Cat Mouse Based Optimizer can also be used to provide a comparison of the resulting accuracy levels. In the Particle Swarm Optimization method, several hybrid methods can also be used to further strengthen the existing accuracy values.

Acknowledgement

We thank the Head of Mathematic Laboratory for the endless support by providing the facilities needed in finishing this paper.

References

[1] Y. Yan, Y. Li, M. Sun, and Z. Wu, "Primary pollutants and air quality analysis for urban air in China: Evidence from Shanghai," *Sustain.*, vol. 11, no. 8, pp. 1–18, 2019, doi: 10.3390/su11082319.

[2] Y. Alyousifi, M. Othman, I. Faye, R. Sokkalingam, and P. C. L. Silva, "Markov Weighted Fuzzy Time-Series Model Based on an Optimum Partition Method for Forecasting Air Pollution," *Int. J. Fuzzy Syst.*, vol. 22, no. 5, pp. 1468–1486, 2020, doi: 10.1007/s40815-020-00841-w.

[3] M. Muhammad, "Sebaran dan Peramalan Mahasiswa Baru Pendidikan Matematika Universitas Muhammadiyah Purwokerto Dengan Metode Time Invariant Fuzzy Time Series," no. 345, pp. 48–58, 2020.

[4] S. M. Robial, "Perbandingan Model Statistik pada Analisis Metode Peramalan Time Series (Studi Kasus: PT. Telekomunikasi Indonesia, Tbk Kandatel Sukabumi)," *J. Ilm. SANTIKA*, vol. 8, no. 2, pp. 1–17, 2018.

[5] S. Sulistyono and W. Sulistiyowati, "Peramalan Produksi dengan Metode Regresi Linier Berganda," *PROZIMA (Productivity, Optim. Manuf. Syst. Eng.*, vol. 1, no. 2, p. 82, 2018, doi: 10.21070/prozima.v1i2.1350.

[6] E. Sumando, F. D. Hamiyani, and I. D. Purba, "Pengembangan Metode Cash Forecasting Pemerintah: Studi Kasus Saldo Kas Pemerintah 2009 – 2011," *Kaji. Ekon. Keuang.*, vol. 2, no. 1, pp. 70–93, 2018.

[7] K. Nugroho, "Model Analisis Prediksi Menggunakan Metode Fuzzy Time Series," *Infokam*, vol. 12, no. 1, pp. 46–50, 2016.

[8] L. Yuliana, "ANALISIS PERENCANAAN PENJUALAN DENGAN METODE TIME SERIES (STUDI KASUS PADA PD. SUMBER JAYA ALUMINIUM)," *J. Mitra Manaj.*, vol. 4, no. 7, pp. 780–789, 2019.

[9] F. N. Masuku, Y. A. R. Langi, and C. Mongi, "Analisis Rantai Markov Untuk Memprediksi Perpindahan Konsumen Maskapai Penerbangan Rute Manado-Jakarta Analysis of Markov Chain To Predict Consumer Movement of Airline Route Manado-Jakarta," *Ilm. Sains*, vol. 18, no. 2, pp. 1–5, 2018.

[10] B. D. Saputra, "a Fuzzy Time Series-Markov Chain Model To Forecast Fish Farming Product," *Kursor*, vol. 9, no. 4, pp. 129–138, 2019, doi: 10.28961/kursor.v9i4.167.

[11] I. H. Kuo, S. J. Horng, T. W. Kao, T. L. Lin, C. L. Lee, and Y. Pan, "An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization," *Expert Syst. Appl.*, vol. 36, no. 3 PART 2, pp. 6108–6117, 2009, doi: 10.1016/j.eswa.2008.07.043.

[12] L. Wang, Y. Liu, X. Zhao, and Y. Xu, "Particle swarm optimization for Fuzzy c-Means Clustering," *Proc. World Congr. Intell. Control Autom.*, vol. 2, pp. 6055–6058, 2006, doi: 10.1109/WCICA.2006.1714243.

[13] N. Putu, E. Merliana, and A. J. Santoso, "Analisa Penentuan Jumlah Cluster Terbaik pada Metode K-Means," pp. 978–979.

[14] C. Yuan and H. Yang, "Research on K-Value Selection Method of K-Means Clustering Algorithm," *J*, vol. 2, no. 2, pp. 226–235, 2019, doi: 10.3390/j2020016.

[15] B. Fechera, J. Kustija, and S. Elvyanti, "Optimasi Penggunaan Membership Function Logika Fuzzy Pada Kasus Identifikasi Kualitas Minyak Transformator.," *Electrans*, vol. 11, no. 2, pp. 27–35, 2012.

[16] M. T. Jatipaningrum, "Peramalan Data Produk Domestik Bruto dengan Fuzzy Time Series Markov Chain," *J. Teknol.*, vol. 9, no. 1, pp. 31–38, 2016.

[17] Z. Zhang, "Fuzzy Time Series Forecasting Based On K-Means Clustering," *Open J. Appl. Sci.*, vol. 02, no. 04, pp. 100–103, 2012, doi: 10.4236/ojapps.2012.24b024.

[18] R. Efendi, Z. Ismail, and M. M. Deris, "Improved weight fuzzy time series as used in the exchange rates forecasting of us dollar to Ringgit Malaysia," *Int. J. Comput. Intell. Appl.*, vol. 12, no. 1, pp. 1–19, 2013, doi: 10.1142/S1469026813500053.

[19] M. I. Mas'ud, A. Safitri, and A. Abassyahil, "Journal Knowledge Industrial Engineering (JKIE)," *J. Knowl. Ind. Eng.*, vol. 4, no. 1, 2017, [Online]. Available: http://jurnal.yudharta.ac.id/v2/index.php/jkie/article/view/863/727.

[20] Y. Arpan and P. C. Dewi, "Analisis Brand Switching Pengguna Transportasi Online Gojek Dengan Metode Markov Chain Di Kota Bandar Lampung," *Valid J. Ilm.*, vol. 15, no. 2, p. 152, 2018.

[21] R. C. Tsaur, "A fuzzy time series-Markov chain model with an application to forecast the exchange rate between the Taiwan and us Dollar," *Int. J. Innov. Comput. Inf. Control*, vol. 8, no. 7 B, pp. 4931–4942, 2012.

[22] B. Santosa, "Tutorial Particle Swarm Optimization," *J. Tek. Ind. ITS*, 2011.

[23] N. Van Tinh, "Enhanced Forecasting Accuracy of Fuzzy Time Series Model Based on Combined Fuzzy C-Mean Clustering with Particle Swam Optimization," *Int. J. Comput. Intell. Appl.*, vol. 19, no. 2, 2020, doi: 10.1142/S1469026820500170.

[24] M. Muhammad, S. Wahyuningsih, and M. Siringoringo, "Peramalan Nilai Tukar Petani Subsektor Peternakan Menggunakan Fuzzy Time Series Lee," *Jambura J. Math.*, vol. 3, no. 1, pp. 1–15, 2021, doi: 10.34312/jjom.v3i1.5940.

[25] R. Mubarak and E. Esyudha, "Prediksi Hasil Tangkapan Ikan Menggunakan Fuzzy Time Series Fish Catch Result Prediction Using Fuzzy Time Series," vol. 08, no. 3, pp. 303–308, 2020, doi: 10.26418/justin.v8i3.39831.