Forecasting retail sales based on cheng fuzzy time series and particle swarm optimization clustering algorithm

R Ariyanto¹*, R H Tjahjana², T Udjiani²

¹Magister Program of Mathematics, Departement of Mathematics, Faculty of Science and Mathematics, Diponegoro University, Jl. Prof. Soedharto, SH, Tembalang Semarang 50275, Indonesia
² Departement of Mathematics, Faculty of Science and Mathematics, Diponegoro University, Jl. Prof. Soedharto, SH, Tembalang Semarang 50275, Indonesia

*Corresponding author : Rikijoe29@gmail.com

Abstract. Use of the conventional forecasting method, which is based on trend data with average sales in the last few months, results inaccurate forecasting due to a large difference in data, this is the same as fuzzy forecasting with the same interval length or static. Therefore, this paper recommends using the Cheng forecasting method combined with the Particle Swarm Optimization (PSO) algorithm. We use an artificial intelligence, i.e., PSO algorithm to set non-static length of intervals each cluster on Cheng method. The comparison of this method yields a better root mean square error (RMSE) value for each cluster on the recommended method.

1. Introduction
Company B in the Semarang city branch sends goods every day to supply the needs of the retail. Data usage based on previous daily closings. Therefore, the company is looking for ways to predict what is needed so that it is maximally fulfilled and there are no empty items. Some research related to retail sales [16,17], in other fields using the cheng method [15,18]. When using the fuzzy forecasting method there is a static interval class value that results in the accuracy of forecasting because the fuzzy time series forecasting which uses the mean value in certain interval class [1]. Many studies have modified this forecasting [2-4], looking for and optimizing the interval [5] and then developing the cluster method [4-12]. One of the optimization methods used in this paper is the artificial intelligence PSO algorithm, by forming clusters, generate the unstable or not constant length of intervals and optimizing the membership function interval so that the prediction rate is more accurate [13,14]. The results of each prediction will be analyzed by comparing each the value to RMSE (root mean square error).

2. Methods

2.1. Fuzzy Time Series
Fuzzy Time Series by [1] is a fuzzy set of real numbers about set of the universe of discourse have been determined or be a class of numbers with vague boundaries. Let $U$ is set of the universe of discourse, $U = \{U_1, U_2, ..., U_n\}$, then the fuzzy set $A$ from $U$ is defined [1] $A = \frac{f_A(U_1)}{U_1} + \frac{f_A(U_2)}{U_2} + \cdots + \frac{f_A(U_n)}{U_n}$ (1), where $f_A$ is member function of $A$, $f_A: U \rightarrow [0,1]$, with degree of membership $1 \leq i \leq n$.

Definition 2.1 [2]
Let \( Y(t) (t = \ldots, 0, 1, 2, \ldots) \) is a subset \( R \). Then \( Y(t) \) as the universe of discourse is defined as a fuzzy set \( f_i(t) \). If \( F(t) \) contains \( f_i(t) (t \in N) \), then \( F(t) \) is defined as the fuzzy time series on \( Y(t) (t = \ldots, 0, 1, 2, 3, \ldots) \), or \( F(t) \) is a function of time \( t \).

**Definition 2.2** [1]

\[
F(t) = F(t - 1) \cdot R(t - 1, t) \quad (2),
\]

is a form formula based on if \( F(t - 1) \) affects \( F(t) \), namely \( F(t - 1) \to F(t) \). Where \( R(t - 1, t) \) is a fuzzy relation between \( F(t - 1) \) and \( F(t) \). Equation (2) is called the first-order \( F(t) \) prediction model. And the fuzzy logical relation is indicated by \( F(t - 1) \to F(t) \).

### 2.2. Cheng Fuzzy Time Series

Forecasting using the Cheng method is as follows [16]:

**Step 1:** Determine the value \( D_{min}, D_{max} \) which is the minimum and maximum value from trend data, and \( D_1, D_2 \in Z^+ \), then defined \( U = [D_{min} - D_1, D_{max} + D_2] \). In determining the value of the universe of discourse \( U \) into intervals of length \( u_1, u_2, \ldots, u_n \) using Sturges formula:

\[
1 + 3.322 \log 10 (n) \quad (4)
\]

while calculating the length of the linguistic intervals using a formula:

\[
\frac{\text{Max Value} - \text{Min Value}}{\text{amount of data}}
\]

**Step 2:** Let \( A_1, A_2, \ldots, A_q \) are the fuzzy set of linguistic values of the linguistic variables. This value is obtained based on equation (6):

\[
A_1 = \frac{a_{11}}{u_{11}} + \frac{a_{12}}{u_{12}} + \cdots + \frac{a_{1r}}{u_{1r}}
\]

\[
A_2 = \frac{a_{21}}{u_{21}} + \frac{a_{22}}{u_{22}} + \cdots + \frac{a_{2r}}{u_{2r}}
\]

\[
\vdots
\]

\[
A_q = \frac{a_{q1}}{u_{q1}} + \frac{a_{q2}}{u_{q2}} + \cdots + \frac{a_{qr}}{u_{qr}}
\]

Where the values \( a_{ij} \in [0,1], 1 \leq i \leq q \) and \( 1 \leq l \leq r \). The value of \( a_{ij} \) is based on the degree of membership level \( u_j \) in the fuzzy set \( A_i \). To determine the degree of membership level, suppose the actual value is between \( A_1 \) and \( A_2 \), then the value is \( A_1 \), and so on. Then form a fuzzy logical relation \( A_1 \to A_m \), this is obtained from the previous data \( A_1 \) and the data after which \( A_m \).

**Step 3:** Make groups based on the initial membership of each data from the fuzzy logical relations

**Step 4:** Calculate the prediction value, the calculation process by following the following rules:

a) Determine the weights, the weight matrix can be normalized by applying with equation (7):

\[
w_n(t) = \begin{bmatrix} W_1, W_2, \ldots, W_q \end{bmatrix} = \left[ \frac{W_1}{\sum_{q=1}^{l} W_q}, \frac{W_2}{\sum_{q=1}^{l} W_q}, \ldots, \frac{W_q}{\sum_{q=1}^{l} W_q} \right] \quad (7)
\]

b) If the fuzzification of the data is \( m \) and has only one fuzzy logical relation formed in step 3, it is indicated as formula (8):

\[
A_m \to A_n
\]

then the forecast value is obtained from \( w_n(t) \cdot m_n \), where \( w_n(t) \) is the weight of the fuzzy logical relation group.

c) If the fuzzification of the data indicated by formula (9) has more than one fuzzy logical relation, namely \( A_1, A_2, \ldots, A_n \) then the forecast value is the total of \( (w_{n1} \cdot m_1 + w_{n2} \cdot m_2 + \cdots + w_{nm} \cdot m_n) \)

\[
A_m \to A_1
\]

\[
A_m \to A_2
\]

\[
\vdots
\]

\[
A_m \to A_n
\]

\[
(9)
\]

d) If the fuzzification has no members, the forecast is 0.

### 2.3. Particle Swarm Optimization Algorithm for Cluster

---

2
The particle swarm optimization (PSO) algorithm is basically a group of particles looking for the best position to get the best solution to the optimization problem. Changes in the method of particles based on equations (10) and (11) [14]:

\[
V_{id} = \omega \cdot V_{id} + C_1 \cdot R_1 \cdot (P_{id} - X_{id}) + C_2 \cdot R_2 \cdot (P_{best} - X_{id}) \quad (10)
\]

\[
X_{id} = X_{id} + V_{id} \quad (11)
\]

The simple PSO equation can change the interval wide, it can be written as formula:

\[
V_{id} = V_{id} + C_1 \cdot R_1 \cdot (P_{id} - X_{id}) + C_2 \cdot R_2 \cdot (P_{best} - X_{id}) \quad (12)
\]

PSO method [13] steps:

**Step 1:** Initialize the particle \((X_{id})\) and the particle velocity \((V_{id})\), the particle velocity is 0.

**Step 2:** Calculate the fitness value of each particle, then the best fitness value is \(P_{best}\) and after \(I\) iteration is obtained the best fitness value for all particles and expressed as \(G_{best}\).

**Step 3:** Calculate the velocity in each particle with equation (12), and update the position with equation (4).

**Step 4:** Repeat steps 2 and 3 until you get the optimal value.

### 2.4. Particle Swarm Optimization Forecasting Method

The class length in the Cheng forecasting method is constant or the same, therefore this paper recommends the PSO algorithm to generate the unstable or not constant length of classes, so that the value is obtained with better accuracy. The following are recommendations for the forecasting method.

**Step 1:** Determine the value \(D_{min}, D_{max}\) which is the minimum and maximum value from trend data, and \(D_1, D_2 \in Z^+\), then defined \(U = [D_{min} - D_1, D_{max} + D_2]\). In determining the value of the universe of discourse \(U\) into intervals of length \(u_1, u_2, ..., u_n\) using the PSO algorithm. The next steps are 2 to 4 following the Cheng method.

### 3. Results and Discussion

In this part, we use the recommended method of forecasting historical retail sales data. With the period 2014 to August 2020. \(n\) clusters are used, \(n = 7\). The Cheng method will be compared with the recommended method. In comparison, other \(n\) values will be used, namely \(n = 50\) and 125. The prediction procedure is as follows:

**Step 1:** The universe of discourse defines \(U\). \(D_{min} = 873\), \(D_{max} = 2955\), \(D_1 = 73\), dan \(D_2 = 45\). So \(U = [800, 3000]\).

**Step 2:** Partition each \(U\) as \(u_1, u_2, ..., u_n\). Determine the number of interval class using the Sturges equation so that in this paper we get 7 clusters. By using the PSO algorithm the value of \(U\) is obtained randomly, it is indicated in Figure 1.

**Step 3:** Let \(A_1, A_2, ..., A_7\) are the sets of linguistic values from the linguistic variable retail sales data trend. Fuzzy sets \(A_1, A_2, ..., A_7\) on the universe of discourse \(U\) are 7 linguistic values. It is indicated by equation (9):

\[
\begin{align*}
A_1 &= \frac{1}{u_1} + \frac{0.5}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} \\
A_2 &= \frac{0}{u_1} + \frac{1}{u_2} + \frac{0.5}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} \\
&\vdots \\
A_7 &= \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0.5}{u_5} + \frac{1}{u_6} + \frac{1}{u_7} 
\end{align*}
\]

The fuzzification results are indicated in Table 1. Then the (FLR) fuzzy logical relations are divided based on the current state in Table 2.

| Table 1. Fuzzy trend data |
|---------------------------|
| Month | Year | Actual | Fuzzification |
|------ |----- |------- |-------------- |
| Jan  | 2014 | 1475  | A3            |


Step 4: The forecast results are indicated in Table 3 by comparing several $n$ clusters.

Table 2. FLRG (Fuzzy Logical Relation Groups) of trend data

| G1 | A1 → A1 | A1 → A2 | A1 → A3 |
| G2 | A2 → A1 | A2 → A2 | A2 → A3 | A2 → A4 |
| G3 | A3 → A1 | A3 → A2 | A3 → A3 | A3 → A4 | A3 → A5 | A3 → A7 |
| G4 | A4 → A2 | A4 → A3 | A4 → A4 | A4 → A5 | A4 → A6 |
| G5 | A5 → A2 | A5 → A3 | A5 → A4 | A5 → A5 | A5 → A6 |
| G6 | A6 → A3 | A6 → A7 |
| G7 | A7 → A3 | A7 → A5 | A7 → A7 |

Table 3. Forecasting result

| Month | Year | Actual | Forecasting Method |
|-------|------|--------|--------------------|
|       |      |        | $n = 7$           | $n = 50$           | $n = 125$          |
|       |      |        | Cheng             | Proposed           | Cheng              | Proposed           |
| Jan   | 2014 | 1475   | 1451.79           | 1658.42            | 1954.5             | 1553.2             | 1938              | 1673.67           |
| Feb   |      | 1773   | 1793.83           | 1792.13            | 2097.5             | 2138              | 2122.5            | 2110              |
| Mar   |      | 1431   | 1793.83           | 1792.13            | 2097.5             | 2138              | 2122.5            | 2110              |
|       |      |        |                   |                   |                   |                   |                   |                   |
| Jun   |      | 2118   | 2122.50           | 2217              | 2122.5            | 2110              |
| Aug   |      | 1877   | 1793.83           | 1792.13            | 1877.5            | 1880              | 1870.5            | 1872              |

Figure 1. Forecasting result for $n = 7$ cluster

In other studies [13,14,19] get better results, and from the above results we can conclude that forecasting using the recommended method is better than the Cheng method. Then compared the average forecasting accuracy value using RMSE, indicated by formula (10):
\[ \text{MSE} = \sqrt{\frac{\sum_{i=1}^{n} (\text{forecasting value}_i - \text{actual value}_i)^2}{n}} \]  
(14)

**Table 4. Ratio of RMSE for each cluster and method**

| Cluster | Cheng method | Proposed Method |
|---------|--------------|-----------------|
| 7       | 358.798      | 356.958         |
| 50      | 261.624      | 252.926         |
| 125     | 158.913      | 133.907         |

Based on Table 4, the ratio of RMSE for each cluster is better to use the recommended method for forecasting retail sales.

**4. Conclusion**

The conclusion that can be drawn is that forecasting using the Cheng Fuzzy Time Series method by adding the Particle Swarm Optimization Algorithm produces different and optimal interval lengths. This algorithm produces forecasts close to the actual value. Forecasting with this method requires a long step and longer training time than Cheng's forecasting method, because there is a process of optimizing the interval value. For further research, it can be done to increase the number of population, add inertia (\(\omega\)) and try to optimize the value of any \(c_1\) and \(c_2\) using a ratio of 2:3, 5:5, and 3:2 \(c_1 + c_2 > 4\), to get an effect on the results of forecasting retail sales.

**References**

[1] Cheng S H, Chen SM and Jian WS 2016 Inf. Sci. 327 272
[2] Huarng K 2001 Fuzzy Sets and Sys. 123(3) 387
[3] Bai E, Wong W K, Chu W C, Xia M and Pan F Exp. Sys. with Appl. 38 2701
[4] Jilani T A, Burney S M A and Aridil C 2008 Int. J. of Comp. Intel. 4(2) 112
[5] Aladag C H, Yolcu U, Egrioglu E and Dalar A Z 2012 App. Soft Comp. 12 3291
[6] Chen SM and Tanuwijaya K 2011 Exp. Sys. with Appl. 38 10594
[7] Baskara P E, Cheninippan M and Subrahmaniam T 2020 J. of loss prev. in the process industries 66 104203
[8] Gruzauskas V, Gimzauskiene E, and Navickas V 2019 J. of cleaner production 240 118225
[9] Dincer N G and Akkus O 2018 Eco. Inf. 43 157
[10] Gupta C, Jain A, Tayal D K, and Oscar C 2018 Eng. App. of Art. Intel. 71 175
[11] Anggodo Y P and Wayan W M 2017 J. Environ. Eng. Sustain. Technol. 04(01) 1
[12] Aljarah I, Ludwig SA 2012 In Proc. of the IEEE (Fargo: North Dakota/USA) 987-1-4673-4768-6
[13] Chen S M, Zou X Y and Gunawan G C 2019 Inf. Sci. 500 127
[14] Kuo I H, Horng S J, Kao T W, Lin T L, Lee C L and Pan Y 2019 Exp. Sys. with Appl. 38 6108
[15] Eman L Nursanil A, Sukono and Fatimah 2019 In Proc. of the Int. Conf. on IEOM (Pilsen, Czech Republic)
[16] Loureiro A L D, Migueis V L, Lucas F M and Sila D 2018 Dec. Supp. Sys. 114 81
[17] Guo Z X, Wong W K and Li M 2013 Dec. Supp. Sys. 55 247
[18] Tricahya S and Rustam Z 2019 Mat. Sci. Eng. 546 052080
[19] Ningrum R W, Surarso B, Farikhin and Safarudin Y M 2018 J. Phys.: Conf. Ser. 983 012055