Sriwijaya Journal of Informatic and Applications
Vol. 2, No. 2, July 2022, pp. 45-52
e-ISSN 2807-2391

Fuzzy Time Series Optimization using Particle Swarm Optimization for Forecasting the Number of Fresh Fruit Bunches (FBB) of Palm Oil

Aisyah Filza Aliyah [a,1], Alvi Syahrini Utami [a,2*], Nabila Rizky Oktadini [a,3]

[a] Artificial Intelligence Lab, Faculty of Computer Science, Sriwijaya University, Palembang, Indonesia
[1]aisyahfilza1910@gmail.com, [2]alvisyahrini@ilkom.unsri.ac.id, nabila@ilkom.unsri.ac.id

ARTICLE INFO

Article history
Received 10 May 2021
Revised 30 May 2021
Accepted 10 June 2021

Keywords
Fuzzy time series
Forecasting
Chen-Hsu
Particle swarm optimization

ABSTRACT

Palm oil is a reliable vegetable oil producer because the oil produced has advantages than oils from other plants. The amount of Fresh Fruit Bunches (FFB) raw material from Palm oil has a significant impact on the palm oil production process. Therefore, we need a method to forecasting the amount of palm oil (FFB). One of the suitable forecasting methods is fuzzy time series (FTS). However, FTS still has shortcomings such as inaccurate determination of the interval length. For this reason, we need to optimize FTS interval to get optimal forecasting. This research implements Particle Swarm Optimization as the optimization method, FTS Chen-Hsu as the forecast method, and Mean Absolute Percentage Error (MAPE) as the measurement of error. The optimization result using PSO produce an error value of 2.0262% smaller than FTS 3.7108%.

1. Introduction

The amount of raw material for Palm Oil (FFB) has a significant impact on the production process of Crude Palm Oil because if the raw material inventory is too much, it will cause additional storage costs and a decrease in product quality, whereas if there is lack of raw material supplies, it will increase procurement costs and cause inefficient activities [4]. It is necessary to forecast the raw material of FFB for Palm Oil to anticipate too much or too little inventory.

One of the suitable forecasting methods for forecasting the amount of palm oil (FFB) is fuzzy time series. FTS is computationally light compared to other algorithms such as Neural Networks so it is easy to develop and can solve historical data forecasting using linguistic values [13]. FTS is a time series forecasting method that has a small error result and can forecast in both short-term and long-term times [10]. The FTS method used is FTS Chen-Hsu which has a smaller error result compared to the previous FTS method [5,6].

In the FTS forecasting method, the length of the interval has been determined at the beginning of the calculation process, while the determination of the length of the interval is very influential in the formation of a fuzzy relationship that will give different results of forecasting calculations. Therefore, the formation of fuzzy relationship and the determination of the length of the interval must be precise [14]. For this reason, one of the effective ways to determine the length of the interval in the FTS is to optimize the length of the FTS interval [2].

Based on previous research, forecasting using FTS optimized with PSO resulted in a smaller MAPE error of 0.73% in the prediction of enrollment at the University of Alabama compared to using the FTS method without being optimized with PSO [7]. In addition, PSO has the advantage that it can quickly achieve convergence, is easy to implement, has few operating functions, and few parameters are specified [1].

In solving the interval problem on FTS Chen-Hsu, the particle population in the PSO is generated in the first iteration, then the particles improve their position towards the best position in each iteration until the iteration is complete [1]. This best position will later be the solution to the interval problem on FTS Chen-Hsu. From this description, this study will optimize FTS Chen-Hsu using Particle Swarm Optimization in forecasting the amount of Palm Oil (FFB).

2. Literature Study / Hypotheses Development

1. Dataset

The data used is secondary data in a period of one month with a span of 7 years (2014-2020) totaling 80 data on net weight (net) of Oil Palm (FFB) in tons.

Table 1. Dataset

DATASET			
Date	No of TBS (ton)	Date	No of TBS (ton)
1/25/2014	1225506.75	1/25/2015	2604585.75
2/25/2014	1135845.75	2/25/2015	2782630.50
3/25/2014	2648480.25	3/25/2015	3054197.25
4/25/2014	3152721.00	4/25/2015	3266727.75
5/25/2014	4045411.50	5/25/2015	3534297.00
6/25/2014	3145681.50	6/25/2015	4221233.25
7/25/2014	2745824.25	7/25/2015	3486121.75
8/25/2014	4210118.25	8/25/2015	4808368.50
9/25/2014	3963121.50	9/25/2015	5212798.50
10/25/2014	4475308.50	10/25/2015	6291275.25
11/25/2014	4075236.75	11/25/2015	5453730.75
12/25/2014	2810174.25	12/25/2015	4269515.25

2. Fuzzy Time Series (FTS) Chen-Hsu

Fuzzy time series (FTS) is a data forecasting method that uses fuzzy principles as a basis. Forecasting systems with fuzzy time series capture patterns from previous data and then use them to project future data [8]. The Classic Time Series method cannot represent data in the form of linguistic value but FTS has a solution for that because FTS can represent historical data into linguistic value [5].

Fuzzy Time Series Chen-Hsu has a solution so that the linguistic value can be more varied by partitioning the length of the interval that has been formed previously based on the distribution of the most data and using rules to estimate the forecasting trend [5,6].

3. Particle Swarm Optimization

Particle Swarm Optimization is a population-based heuristic method developed by Kennedy and Eberhart in 1995, which was inspired by the collective movement of organisms such as flocks of birds and fish schools to simulate their behavior in searching for food sources. In the PSO implementation, the particle population is generated in the first iteration, then the particles improve their position towards the best position in each iteration, so that the optimum position is obtained [1].

The PSO algorithm has the advantage that when compared to other optimization methods PSO is faster to achieve convergence, besides that PSO is also easy to implement, has few operating functions, and few parameters to be determined [1]. However, the PSO algorithm that is too fast in achieving convergence can result in premature convergence. Premature convergence occurs when the PSO algorithm approaches the final solution, so the algorithm is unable to find a new, wider solution space [12].

3. Methodology

This section describes the research approach used in the study, underlines the data collection process, respondents list (for empirical study), and analysis process.

1. Fuzzy Time Series (FTS) Chen-Hsu

According to [5,6], the steps of FTS Chen-Hsu method as follows.

a. Universe of Discourse

At this stage, the Universe of Discourse is defined.

$$U = [D_{min}, D_{max}] \tag{1}$$

Universe of Discourse (U) is the universe of speakers where is the minimum value and is the maximum value from historical data.

b. Interval

At this stage, the value of U will be partitioned into several intervals of the same length. The next step is to calculate the frequency of occurrence of data from each divided interval and sort the intervals from the highest to the lowest frequency. The interval that has the highest frequency of occurrence of data is divided into four sub-intervals of the same length, the second highest is divided into three sub-intervals, and the third highest is divided into two sub-intervals. For the next interval, there is no need to change and if there is no data distributed in an interval then remove the interval.

c. Fuzzification

At this stage a fuzzy set is defined (A_1, A_2, A_3,…, A_i) to the universal set U which turns into linguistic data, that is, if a data is included in the u_i interval, then the data has a linguistic value A_i.

d. Fuzzy Logic Relationship

At this stage, the relationship between each data is defined according to the linguistic value. Assume at the (n-1) time, the fuzzification of the data is Ai and at the nth time, the fuzzification of the data is Aj, then the fuzzy logic relationship¬ is "Ai → Aj".

e. Defuzzification

At this stage, defuzzification is carried out to calculate the fuzzy value into a crisp value. In the defuzzification process, each interval is divided by the same length, where 0.25 points and 0.75 points from each interval are used as upward and downwards for forecasting, while 0.5 points as the middle value. There are conditions and rules used in determining the forecasting trend to be up, middle, or down. The following are the conditions and rules for the defuzzification process.

CONDITION:

1. **IF** (j>i) **and** ((Data $_{n-1}$ – Data $_{n-2}$) – (Data $_{n-2}$ – Data $_{n-3}$) >= 0))
   THEN RULE 2
2. **IF** (j>i) **and** ((Data $_{n-1}$ – Data $_{n-2}$) – (Data $_{n-2}$ – Data $_{n-3}$) < 0))
   THEN RULE 3
3. **IF** (j<i) **and** ((Data $_{n-1}$ – Data $_{n-2}$) – (Data $_{n-2}$ – Data $_{n-3}$) >= 0))
   THEN RULE 2
4. **IF** (j<i) **and** ((Data $_{n-1}$ – Data $_{n-2}$) – (Data $_{n-2}$ – Data $_{n-3}$) < 0))
   THEN RULE 3
5. **IF** (j==i) **and** ((Data $_{n-1}$ – Data $_{n-2}$) – (Data $_{n-2}$ – Data $_{n-3}$) >= 0))
   THEN RULE 2
5. **IF** (j==i) **and** ((Data $_{n-1}$ – Data $_{n-2}$) – (Data $_{n-2}$ – Data $_{n-3}$) < 0))
   THEN RULE 3

Fig. 1. Condition

RULE 1:

1. **IF** $|(\text{Data}_{n-1} - \text{Data}_{n-2})| / 2 > (0{,}5 * \text{distance (Interval } A_j))$
   THEN 0,75 point
2. **IF** $|(\text{Data}_{n-1} - \text{Data}_{n-2})| / 2 == (0{,}5 * \text{distance (Interval } A_j))$
   THEN 0,5 point
3. **IF** $|(\text{Data}_{n-1} - \text{Data}_{n-2})| / 2 < (0{,}5 * \text{distance (Interval } A_j))$
   THEN 0,25 point

Fig. 2. Rule 1

RULE 2:

1. **IF** ($|(\text{Data}_{n-1} - \text{Data}_{n-2}) - (\text{Data}_{n-2} - \text{Data}_{n-3})| * 2 + \text{Data}_{n-1}$) in Interval A_j **OR** ($\text{Data}_{n-1} - |(\text{Data}_{n-1} - \text{Data}_{n-2}) - (\text{Data}_{n-2} - \text{Data}_{n-3}) * 2$) in Interval A_j
   THEN 0,75 point
2. **IF** ($|(\text{Data}_{n-1} - \text{Data}_{n-2}) - (\text{Data}_{n-2} - \text{Data}_{n-3})| / 2 + \text{Data}_{n-1}$) in Interval A_j **OR** ($\text{Data}_{n-1} - |(\text{Data}_{n-1} - \text{Data}_{n-2}) - (\text{Data}_{n-2} - \text{Data}_{n-3}) / 2$) in Interval A_j
   THEN 0,25 point
3. **ELSE** (0,5 point)

Fig. 3. Rule 2

RULE 3:

1. **IF** ($|(\text{Data}_{n-1} - \text{Data}_{n-2}) - (\text{Data}_{n-2} - \text{Data}_{n-3})| / 2 + \text{Data}_{n-1}$) in Interval A_j **OR** ($\text{Data}_{n-1} - |(\text{Data}_{n-1} - \text{Data}_{n-2}) - (\text{Data}_{n-2} - \text{Data}_{n-3}) / 2$) in Interval A_j
   THEN 0,25 point
2. **IF** ($|(\text{Data}_{n-1} - \text{Data}_{n-2}) - (\text{Data}_{n-2} - \text{Data}_{n-3})| * 2 + \text{Data}_{n-1}$) in Interval A_j **OR** ($\text{Data}_{n-1} - |(\text{Data}_{n-1} - \text{Data}_{n-2}) - (\text{Data}_{n-2} - \text{Data}_{n-3}) * 2$) in Interval A_j
   THEN 0,75 point
3. **ELSE** (0,5 point)

Fig. 4. Rule 3

2. Particle Swarm Optimization

a. Population Initialization

At this stage, determine the parameter values needed for PSO, namely the number of iterations, the number of particles, the values of c_1 and c_2, initialization of the initial velocity of the particles, and initialization of the initial position of the particles. The initial position is taken randomly with the following equation [3].

$$x_i = random() \quad (2)$$

Where x_i = position of i particle.

While the initial speed is a random value between the maximum speed (Vmax) and the minimum speed (Vmin) from the following formula [9].

$$\text{Vmax} = 0{,}1 * (\text{data max} - \text{data min}) \quad (3)$$
$$\text{Vmin} = -[0{,}1 * (\text{data max} - \text{data min})] \quad (4)$$

b. Update Velocity & Update Particle Position

The Pbest value is obtained by comparing the Pbest value in the previous iteration with the results from the calculation of the new position update (new Pbest). The better fitness value will be the Pbest value for the particle. To get the Gbest value, it can be done by looking at the Pbest comparison of each particle that has the best fitness value. The fitness value is a process to get the MAPE error value in FTS Chen-Hsu calculation.

$$v\ new_{i,j} = \omega . v_{i,j} + c1.r1\left(Pbest_{i,j} - x_{i,j}\right) + c2.r2\left(Gbest_{j} - x_{i,j}\right) \quad (5)$$

According to [11], the value of ω is obtained from the following formula.

$$\omega = \omega_{max} - \omega_{min}) * G/maxG \quad (6)$$

Information:

$v\ new_{i,j}$: particle velocity i dimension j
ω : inertial weight
ω_{max} : maximum inertial weight
ω_{min} : minimum inertial weight
G : current iteratiom
maxG : maximum iteration
c : coefficient
r : random value ∈ [0,1]
$Pbest_{i,j}$: the best position of the particle i dimension j
$Gbest_{j}$: global optimal j
$x_{i,j}$: particle position i dimension j

c. Update Personal Best & Global Best

The Pbest value is obtained by comparing the Pbest value in the previous iteration with the results from the calculation of the new position update (new Pbest). The better fitness value will be the Pbest value for the particle. To get the Gbest value, it can be done by looking at the Pbest comparison of each particle that has the best fitness value. The fitness value is a process to get the MAPE error value in FTS Chen-Hsu calculation.

d. Repeating The Process Until Conditions Are Fulfilled

Before the conditions are met or reach the maximum number of iterations the system will repeat the process from stage b until the conditions are met (getting the Gbest value with the best fitness).

3. Mean Absolute Percentage Error (MAPE)

At this stage, the error calculation of the forecasting results is carried out using the Mean Absolute Percentage Error (MAPE) value [5]. The formula used is as follows.

$$MAPE = \frac{\sum_{i=1}^{n} \frac{|Yt - Yf|}{Yt}}{n}\ x\ 100 \quad (7)$$

Where n is the amount of data, Yt is the original data value (actual value) and Yf is the forecast value data.

4. Result and Discussion

1. Calculation Result of FTS Chen-Hsu Error Value

Table 2. Error Calculation FTS Chen-Hsu

No.	Date	TBS of Oil Palm Num	*Forecasting Result*
1.	Jan 2014	1225506.75	-
2.	Feb 2014	1135845.75	1166442.18
3.	Mar 2014	2648480.25	2588270.87
…	….	…	…
n	Augst 2020	1066546.75	1166442.18
MAPE			3.7108%

2. Test Result of Number of Iterations

Testing the number of iterations was carried out each 5 times. The result of testing the number of iterations shows that the higher the number of iterations, the smaller the MAPE value does not necessarily mean. This can be seen in fig. 5 that the 8th test with 400 iterations resulted in the smallest MAPE value.

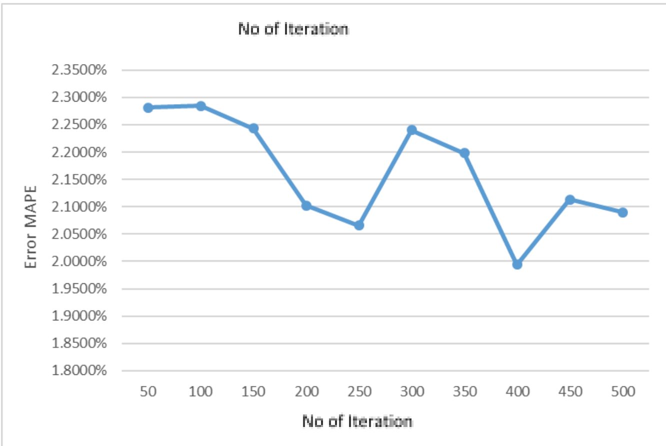

Fig. 5. Graph of MAPE Value on Testing Number of Iterations

3. Test Result of Number of Particle

The number of particles was tested 5 times each. The results of the particle count test show that the greater the number of particles does not necessarily have a smaller MAPE value. This can be seen in fig. 6 that the 5th test with 50 particles resulted in the smallest MAPE value.

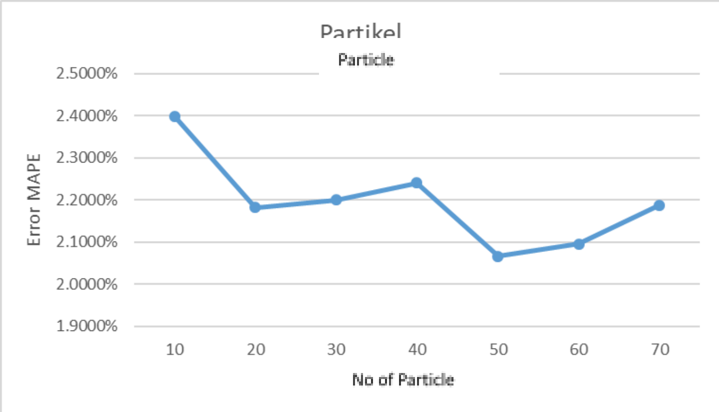

Fig. 6. Graph of MAPE value on Testing Number of Particle

4. Test Result of c_1 and c_2

Testing the values of c1 and c2 was carried out 5 times each. The results of testing the values of c1 and c2 show that the smaller the values of c1 and c2, the smaller the MAPE value. This can be seen in fig. 7 that in the 1st test with a value of 1 for c1 and c2 resulted in the smallest MAPE value

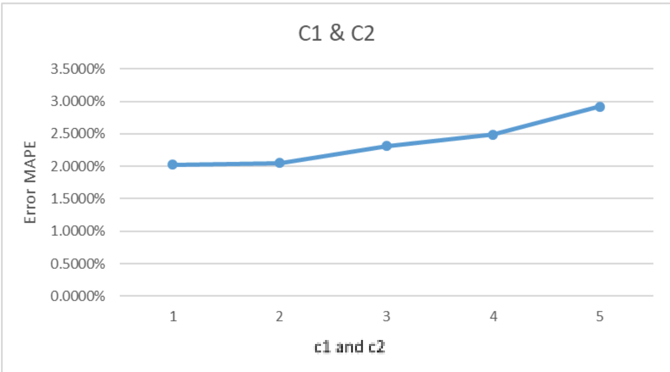

Fig. 7. Graph of MAPE Value on Testing the values of c1 and c2

5. Conclusion

Based on the tests that have been carried out, the best parameters are 400 iterations, 50 particles, and a value of 1 for c1 and c2. The MAPE FTS Chen-Hsu value produces an error of 3.7108% while the MAPE FTS Chen-Hsu value with PSO produces a smaller error of 2.0262%.

The results of the comparison test show that FTS Chen-Hsu with PSO produces a smaller MAPE value than FTS Chen-Hsu, so the prediction results using FTS Chen-Hsu with PSO are better than FTS Chen-Hsu without using optimization.

References

[1] Abduh, M., Regasari, R., Putri, M. & Muflikhah, L. 2017. Optimasi Pembagian Tugas Dosen Pengampu Mata Kuliah Dengan Metode Particle Swarm Optimization. 1(10): 989–999.

[2] Adi Prasojo, C. & Darma Setiawan, B. 2018. Optimasi Fuzzy Time Series Menggunakan Algoritma Particle Swarm Optimization Untuk Peramalan Jumlah Penduduk Di Kabupaten Probolinggo. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya, 2(8): 2548–964.

[3] Alghifari, D.R., Rahayudi, B. & Dewi, C. 2019. Optimasi Fuzzy Time Series Menggunakan Algoritme Particle Swarm Optimization Untuk Peramalan Produk Domestik Bruto (PDB) Indonesia. … Teknologi Informasi dan Ilmu Komputer, 3(4).

[4] Amalia, R.R. & Hairiyah, N. 2018. Peramalan Kebutuhan Bahan Baku Tandan Buah Segar (TBS) Menggunakan Metode Exponential Smoothing dan Linier Regresion di PT. Pola Kahuripan Intisawit. Jurnal Teknologi Agro-Industri, 5(2): 101.

[5] Anggraeni, W. & Suyahya, I. 2016. Prediksi Kurs Rupiah Terhadap Dolar Amerika Menggunakan Metode Fuzzy Time Series Chen dan HSU. STRING (Satuan Tulisan Riset dan Inovasi Teknologi), 1(1): 19–28.

[6] Chen, S.-M. & Hsu, C.-C. 2004. A new method to forecast enrollments using fuzzy time series. International Journal of Applied Science and Engineering, 3: 234–244.

[7] Chen, S.M., Zou, X.Y. & Gunawan, G.C. 2019. Fuzzy time series forecasting based on proportions of intervals and particle swarm optimization techniques. Information Sciences, 500: 127–139.

[8] Elisawati & Masrizal 2017. Penerapan Fuzzy Time Series Model Chen Untuk Memprediksi Jumlah Penduduk. Journal of Chemical Information and Modeling, 1(1): 259–267.

[9] Farsadi, M., Hosseinnejad, H. & Dizaji, T.S. 2016. Solving unit commitment and economic dispatch simultaneously considering generator constraints by using nested PSO. ELECO 2015 - 9th International Conference on Electrical and Electronics Engineering, 493–499.

[10] Ganguly, P., Kalam, A. & Zayegh, A. 2017. Short term load forecasting. Melecon, (October): 1470–1473.

[11] Jie, Z., Chaozan, F. & Bo, L. 2016. An Improved Particle Swarm Optimization Based on Repulsion Factor. (January 2012).

[12] Kurniawan, M. & Suciati, N. 2017. Modifikasi Kombinasi Particle Swarm Optimization dan Genetic Algorithm untuk Permasalahan Fungsi Non-Linier. INTEGER: Journal of Information Technology, (1998): 31–40.

[13] Nugroho, K. 2016. Model Analisis Prediksi Menggunakan Metode Fuzzy Time Series. Infokam, 12(1): 46–50.

[14] Xihao, S., L.Y. 2008. Average-based fuzzy time series models for forecasting Shanghai compound. World Journal of Modelling and Simulation, 4(2): 104–111.