Optimization of Tsukamoto FIS Using Genetic Algorithm for Rainfall Prediction in Banyuasin Regency

Muhammad Rafi Akbar[a,1], Kanda Januar Miraswan[b,2,*], Desty Rodiah[b,3],
Muhammad Ali Buchari[c,4], Anna Dwi Marjusalinah[d,5]

[a]Undergraduate Student of Informatics Engineering, University of Sriwijaya, Palembang, Indonesia
[b]Lecturer of Informatics Engineering, Faculty of Computer Science, University of Sriwijaya, Palembang, Indonesia
[c]Lecturer of Computer Engineering, Faculty of Computer Science, University of Sriwijaya, Palembang, Indonesia
[d]Lecturer of Informatics Management, Faculty of Computer Science, University of Sriwijaya, Palembang, Indonesia
[1] murafba@gmail.com; [2] kandajm@ilkom.unsri.ac.id*; [3] destyrodiah@gmail.com;
[4] m.ali.buchari@unsri.ac.id; [5] annadwimarjusalinah@unsri.ac.id
* corresponding author

ARTICLE INFO

Article history
Received 31 August 2024
Revised 5 Sept 2024
Accepted 10 Sept 2024

Keywords
Rainfall Prediction
Fuzzy Inference System
Genetic Algorithm
Prediction Optimization
Fuzzy Tsukamoto
Fuzzy Optimization

ABSTRACT

Indonesia, as a tropical country with high rainfall, heavily relies on accurate rainfall predictions for various critical purposes, including water resource management and extreme weather impact mitigation. One commonly used method is the Tsukamoto Fuzzy Inference System (FIS). However, implementing the Tsukamoto FIS often leads to high error rates. This is attributed to the difficulty in determining the boundaries of fuzzy variable membership functions. To address this issue, this research proposes an innovative approach by optimizing the boundaries of fuzzy membership functions using Genetic Algorithms (GA). The study resulted in a 49.02% reduction in the error rate, decreasing from 76.82% to 27.8%. This method significantly enhances rainfall prediction accuracy and contributes to the advancement of more sophisticated prediction methods. The optimization method proposed in this study also holds potential for application across various atmospheric science contexts.

1. Introduction

Indonesia is a tropical country with high rainfall. The percentage of rainfall in Indonesia varies significantly, ranging between 8% and 37%, with an average of 22% [1]. Rainfall prediction is crucial as it plays a significant role in human life. The prediction outcomes serve as benchmarks for various future planning purposes in numerous sectors, such as irrigation management, agriculture, livestock farming, plantation, flood mitigation, aviation, urban drainage, ports, industries, etc. For instance, in the agricultural sector, there was an average reduction in potato yields in 2010 by 3,385 kg/ha due to changes in rainfall patterns [2]. Then, in the plantation sector, there was a 15% decrease in apple production in 2014 compared to 2013 [3].

Tsukamoto Fuzzy Inference System (FIS) is a commonly used approach in performing predictions. Based on the research [4], the Tsukamoto FIS can predict rainfall in the Tengger region with the smallest RMSE value of 8.64. The study conducted by Reynaldi et al. [5] stated that the Tsukamoto FIS resulted in an error rate of 8%, while the Sugeno FIS reached 38%. However, in fact, the Tsukamoto inference method in practice sometimes still yields high error values. This is due to the difficulty in determining the boundaries of the membership functions, which affect the prediction results. Azizi [6] in his research stated that the fuzzy Tsukamoto prediction still yielded a very high error rate, specifically at 76.9%. Another study conducted by Fajri et al. [7] in diagnosing dental diseases mentioned that using the Tsukamoto method without optimization only resulted in a 70% accuracy rate. Similarly, research conducted by Afrilia [8] and Paramitha [9] stated that the Tsukamoto FIS only yielded accuracy rates of 50% and 40% respectively. Therefore, it is necessary to optimize the boundaries of membership functions to obtain prediction results with a more optimal rate of accuracy.

The Genetic Algorithm (GA) is one of the population-based search algorithms widely utilized in various optimization problems [10]. It can solve complex problems and has a broad search space. Moreover, computational processes in genetic algorithms tend to be swift and yield more effective solutions [11,12]. Based on research conducted by Wahyuni and Mahmudy [13], GA can be applied to optimize the boundaries of fuzzy membership functions and reduce the prediction error rate by up to 26.45%. Rofiq et al. [14] in their research demonstrated that optimizing the Tsukamoto FIS using GA was able to increase the accuracy rate by 47%, from 47% to 94%. Therefore, based on the explanations before, this study will utilize GA to optimize the boundaryvalues in Tsukamoto's fuzzy membership functions, in hopes of obtaining rainfall prediction results with a lower error rate. Additionally, the study outcomes will be compared with a previousresearch conducted by Azizi, which optimized the Tsukamoto FIS using the Artificial Bee Colony (ABC) algorithm [6].

2. Literature Study

a. Tsukamoto Fuzzy Inference System (FIS)

Fuzzy logic is a logic that holds truth values within the range of 0 to 1 [15]. It maps input values into an output through a fuzzy inference system. There are several types of FIS, such as Mamdani, Sugeno, and Tsukamoto. In Tsukamoto FIS, each consequent of every rule needs to be represented by a fuzzy set with a monotonic membership function. Then, the inference output is explicitly given based on the fire strength [16]. The result of the defuzzification process is obtained from the weighted average formula shown in (1).

$$Z = \sum_{i=1}^{n} \frac{\alpha_i Z_i}{\alpha_i} \tag{1}$$

Zrepresents the defuzzification value, while α_i and Z_i respectively denote the fire strength resulting from the minimum implication function and the crisp value resulting from the monotonic membership function.

b. Genetic Algorithm (GA)

The GA is one of the heuristic optimization methods based on genetic principles and the natural selection process of Darwin's Theory of Evolution [17]. This method was first introduced by John Holland in the 1970s. The GA works by modeling the boundaries of membership functions values into chromosomes. These chromosomes then reproduce and are subsequently selected based on their fitness values. Chromosome with the best fitness value will be used as the boundaries for the fuzzy membership functions [13].

3. Methodology

a. Chromosome Representation

The chromosome representation is the initial stage in the GA process. This stage involves forming chromosomes based on the predetermined population size by encoding candidate solutions within the specified range of values [18]. These candidates' solutions are also referred to as genes or alleles. The chromosome representation used in this study is the real-coded or floating point. The chromosome is composed of 12 genes divided into 4 criteria. Each criterion consists of 3 genes representing the boundaries of the membership functions for each fuzzy input variable. The gene values are randomly generated and then arranged in ascending order. Fig. 1 shows the chromosome representation.

Chromosome

20	25	30	50	60	70	1009	1011	1013	1	5	7.5
Temperature			Humidity			Air Pressure			Wind Velocity		

Fig. 1. Chromosome representation

b. Fitness Function

The fitness function plays a role in measuring how well a chromosome can endure and solve a problem. This aligns with the concept of natural selection, where superior individuals have a higher probability of surviving into the next generation. Therefore, the larger the fitness value of a chromosome, the smaller its chance of being eliminated in the subsequent generation [18]. The calculation of the fitness value is expressed in (2).

$$Fitness = \frac{1}{MAPE} \tag{2}$$

Equation (3) is the formula used to determine the MAPE value. MAPE is utilized to measure the error rate of rainfall prediction resultsusing the Tsukamoto FIS, where $A_i$stands for actual data and F_i denotes the forecasted value.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{A_i - F_i}{A_i} \right| \tag{3}$$

c. Parent Selection

The parent selection method employed is the roulette wheel selection. The initial step in the roulette wheel selection involves calculating the probability and cumulative probability of each chromosome. Afterward, two random values within the range of 0 to 1 will be generatedto determine the parents based on the condition $C_{n-1}<R<C_n$ [13]. The calculation of probability and cumulative values is presented in Table 1, while the outcomes of parent selection are displayed in Table 2.

Table 1. Calculation Results for Probability and Cumulative Values of Each Chromosome

Chromosome	Fitness Value	Probability Value	Cumulative Value
1	0.0224316	0.35311	0.35311
2	0.0138389	0.217847	0.570957
3	0.0272554	0.429043	1

Table 2. The Outcomes of Roulette Wheel Selection

Generated Random Value	Selected Chromosome
0.2582804	1
0.5212309	2

d. Crossover

The crossover operator is a primary operator in GA that evaluates genes in the selected parent chromosomes to generate new chromosomes inheriting the traits of their parents [17]. The crossover method utilized in this study is the extended intermediate crossover [13]. However, before proceeding, it is necessary to determine the number of offspring formed using (4), where *cr*denotes the crossover rate input by user within the range of0 to 1.

$$Offspring = cr \times population\ size \tag{4}$$

Once the number of offspring formed is determined, the next step is to evaluate the genes in the parent chromosomes to generate offspring using (5).

$$Offspring_{i_j} = Parent_{1_j} + \alpha(Parent_{2_j} - Parent_{1_j}) \tag{5}$$

$Offspring_{i_j}$ refers to the i-th offspring with the j-th gene value, while αdenotes to a random number in the range [-0.25, 1.25]. Fig. 2 illustrates an example of intermediate crossover process.

Parent 1	20	25	30	50	60	70	1009	1011	1013	1	5	7.5
Parent 2	24	27.5	29	63	75	85	1008	1012	1015	2	4	7
α	0.59	0.48	0.63	0.8	0.57	0.473	0.251	-0.19	0.912	0.283	-0.2	0.1
Offspring 1	22.36	26.2	29.37	60.4	68.55	77.095	1008.7	1010.8	1014.8	1.28	5.2	7.45

Fig. 2. Example of intermediate crossover

e. Mutation

Mutation is the second reproduction process following crossover. The mutation process is considered a secondary operator in genetic algorithms. Similar to the natural selection process in real life, chromosomes in GA can undergo structural changes in their genes. These changes result in the formation of a mutant, which is a new chromosome genetically different from its predecessor. This operator works on a single chromosome, unlike the crossover operator, which operates on a pair of chromosomes [18]. The mutation process prevents the obtained solutions from falling into local optima because it introduces diversity into the chromosomes within a population [15].

The type of mutation employed in this study is random mutation [15].The mutation works by altering the selected gene values with small random numbers using (7), where α represents a randomly generated value within the range of [-0.1, 0.1], whilemax_y and min_y denote the value range for each variable. However, beforehand, the number of mutants formed must bedetermined using (6). The variable *mr*represents the mutation rate within the range of values from 0 to 1.

$$Mutant = mr \times population\ size \tag{6}$$

$$Mutant_{ij} = Chromosome_{ij} + \alpha(max_y - min_y) \tag{7}$$

Fig. 3 shows an example of random mutation process.

Chromosome	25	26,1	27,25	63	68,8	95,5	1010,7	1012,5	1013,4	3,9	7	8,2
α	-0,03	0,02	0,03	0,04	0,05	0,04	-0,10	0,01	0,04	0,09	0,10	0,06
Mutant	24,89	26,16	27,35	63,48	69,9	96,46	1010,3	1012,5	1013,6	4,26	7,5	8,86

Fig. 3. Example of random mutation

f. Evaluation and Elitism

The old and new chromosomes, including offspringfrom the crossover process and mutants from the mutation process, will each be evaluated using the Tsukamoto FIS to obtain their respective MAPE values. Subsequently, the fitness values for each individual will be calculated using formula (2). After evaluating the fitness values of all chromosomes, the next step is to reselect these chromosomes to replace the less superior chromosomes from the previous generation employing elitism. Elitism selection eliminates chromosomes with the lowest fitness values [14]. Assume a population consists of 3 chromosomes, with additional chromosomes comprising one offspring and one mutant, then the results of the MAPE and fitness calculations are shown in Tables 3, while Table 4 presents an example of the elitism selection results.

Table 3. The Results of The MAPE and FitnessCalculations

Chromosome	1	2	3	4	5
MAPE	44.58	72.26	36.69	49.41	59.46
Fitness	0.0224316	0.0138389	0.0272554	0.0202388	0.016818

Table 4. The Elitism Selection Results

Chromosome	3	1	4
Fitness	0.0272554	0.0224316	0.0202388

g. Stopping Condition

There are two types of stopping condition in genetic algorithms. The first involves limiting the number of generations, for instance, 50 generations [13]. The second method involves calculating consecutive chromosome replacement failures within the population. For this research, the first stopping condition will be used.

4. Result and Discussion

In this research, the monthly climate data of Banyuasin Regency from January 2018 to December 2022 will be utilized for rainfall prediction [6]. The climate data comprises attributes such as temperature, air humidity, air pressure, wind velocity, and rainfall.

There are 4 testing parameters: population size, generation size, crossover rate, and mutation rate. The testing aims to determine the optimal parameter values that yield the lowest average MAPE value. Each parameter undergoes five testing scenarios.

a. The Test Results of Tsukamoto FIS

The rainfall prediction results using the Tsukamoto FIS method resulted in a very high error value, specifically 76.82%. Table 5 presents the prediction outcomes.

Table 5. The Rainfall Prediction Results Using the Tsukamoto FIS

Period	**Actual Data**	**Forecasted Data**	**APE**
Jan-18	326.00	480.29	47.33%
Feb-18	311.00	379.72	22.1%
Mar-18	317.00	433.27	36.68%
…	…	…	…
Oct-22	477.00	269.45	43.51%
Nov-22	211.00	347.17	64.54%
Dec-22	266.00	472.36	77.58%
MAPE			76.82%

b. TheTest Results of Generation Size

The testing based on the generation size was performed five times. The generation size ranged from 10 to100, with increments of 10. Meanwhile, the parameters such as population size, crossover rate, and mutation rate were set at 20, 0.3, and 0.1, respectively. This testing resulted in the lowest average MAPE value occurring at the transition of 100 generations, which was 32.79%. These results indicate that as the generation size increases, the average MAPE value is likely to decrease. The test results are visualized in Fig. 4.

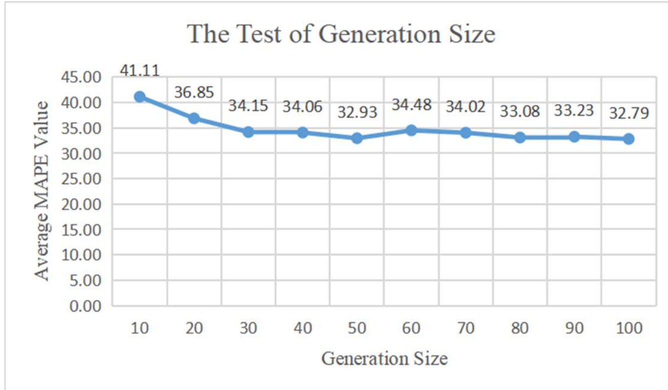

Fig. 4. The test resultsof the generation size

c. The Test Results of Population Size

The testing based on population size was also conducted five times. The population size ranged from 10 to 100, with increments of 10. The parameters such as crossover rate and mutation rate respectively were each set at 0.3 and 0.1, whereas the generation size was fixed at 100 based on the results from the previous tests. This testing resulted in the lowest average MAPE value when the population consisted of 80 chromosomes, precisely at 31.08%. These results indicate that with an increase in population size, the average MAPE value tends to decrease. However, after reaching a certain point, the decrease in the MAPE value becomes less significant. Fig. 5 illustrates the outcomes of the population size testing.

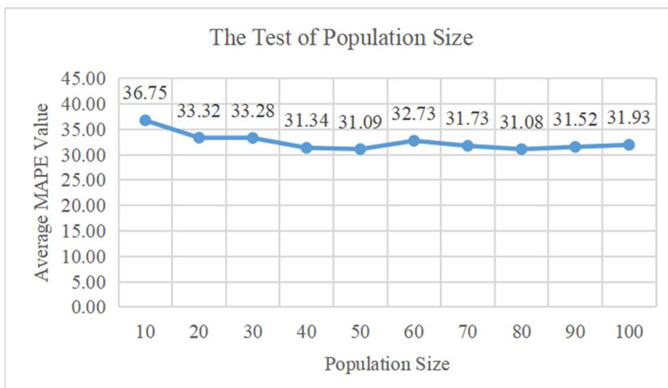

Fig. 5. The test results of the population size

d. The Test Results ofCrossover Rate

The testing based on the crossover rate parameter commenced with an input value from 0.1 to 1, incremented by 1. The mutation rate parameter was set at 0.1, while the generation size and population size were fixed at 100 and 80, respectively, derived from previous testing. This testing resulted in the lowest average MAPE value occurring at a crossover rate of 0.1, precisely at 31.07%. These findings indicate that as the crossover rate increases, the average MAPE value is also likely to increase. Fig. 6 depicts the outcomes of the testing concerning the crossover rate parameter.

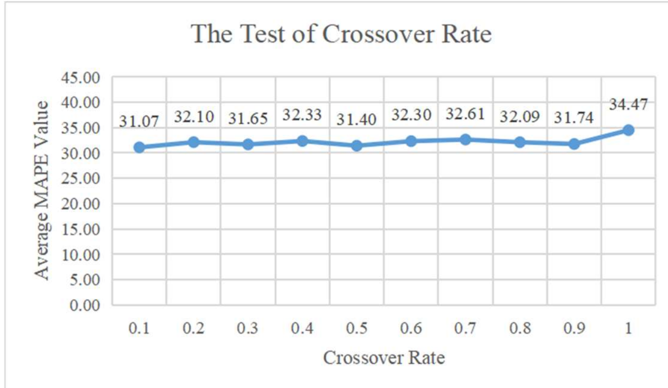

Fig. 6. The test results of the crossover rate

e. The Test Results ofMutation Rate

The final testing was based on the mutation rate, conducted five times. The input parameter for the mutation rate was the same as the crossover rate, ranging from 0.1 to 1, incremented by 1. Meanwhile, the parameters such as generation size, population seize, and crossover rate were set at 100, 80, and 0.1, respectively, derived from the previous testing. This testing resulted in the smallest average MAPE value occurring at a mutation rate of 1, precisely at 28.83%. These results indicate that as the mutation rate increases, the average MAPE value is likely to decrease. Additionally, this study found the smallest MAPE value at 27.8% in the fifth testing scenario when the mutation rate was 0.4. Fig. 7 illustrates the outcomes of the mutationrate testing.

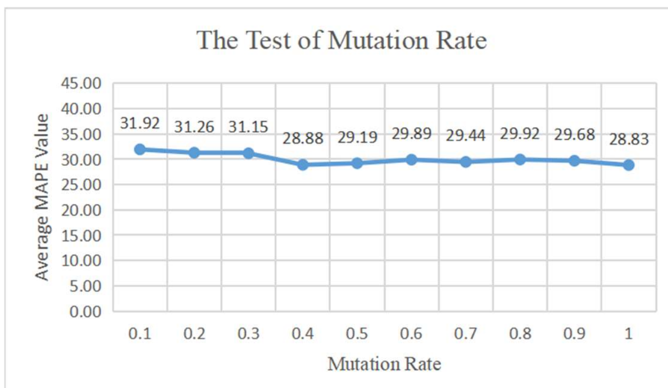

Fig. 7. The test results of the mutation rate

f. Comparative Analysis of GA and ABC

The lowest MAPE value obtained based on the tests was 27.8%. This result clearly demonstrates that GA has a significant impact on reducing the error rates in the Tsukamoto FIS. However, the optimization of the ABC algorithm showed a better result, precisely 26.02% [6]. Therefore, in predicting rainfall, the ABC algorithm slightly outperforms GA. Table 10 illustrates the MAPE values for each method.

Table 6. The Comparison of MAPE Values for Each Method

Method	MAPE
Tsukamoto FIS	76.82%
Tsukamoto FIS + GA	27.8%
Tsukamoto FIS + ABC [6]	26.02%

5. Conclusion

Based on the conducted tests, it can be concluded that the genetic algorithm can be applied to the Tsukamoto FIS to optimize the boundaries of variable membership functions for rainfall prediction and significantly impact the reduction of MAPE values. This is evidenced by the test results, which achieved much lower MAPE values after optimization using the GA compared to before optimization, reducing from 76.82% to 27.8%. These results were obtained with input parameters set at a generation size, population size, crossover rate, and mutation rate of 100, 80, 0.1, and 0.4, respectively. Furthermore, it's notable that the Tsukamoto optimization using the GA did not outperform the ABC algorithm in Azizi's research [6], which achieved the lowest MAPE value of 26.02%. However, this difference is not particularly significant.

For future research, optimizing the Tsukamoto FIS could be done using more efficient optimization algorithms such Particle Swarm Optimization (PSO) [7]. Alternatively, exploring other fuzzy inference system methods like Mamdani could be considered. Furthermore, this study could involve modifying the chromosome structure by adding output variables, thereby expanding the gene length to 16 genes.

References

[1] D. Mulyono, "Analisis karakteristik curah hujan di wilayah Kabupaten Garut Selatan," *Jurnal Konstruksi*, vol. 13, pp. 1–9, April 2016.
[2] I. Wahyuni, W.F. Mahmudy, and A. Iriany, "Rainfall prediction using hybrid adaptive neuro-fuzzy inference system (ANFIS) and genetic algorithm," *Journal of Telecommunication, Electronic, and Computer Engineering (JTEC)*, vol. 9, pp. 51–56, 2017.
[3] I. Wahyuni and F. A. Ahda, "Pemodelan fuzzy inference system Tsukamoto untuk prediksi curah hujan studi kasus Kota Batu," *Jurnal Ilmiah Teknologi Informasi Asia*, vol. 12, pp. 115–124, 2018.
[4] I. Wahyuni, W. F. Mahmudy, and A. Iriany, "Rainfall prediction in Tengger region Indonesia using Tsukamoto fuzzy inference system," *2016 1st International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, pp. 130–135, 2016.
[5] Reynaldi, W. Syafrizal, M. F. A. Hakim, "Analisis perbandingan akurasi metode fuzzy Tsukamoto dan fuzzy Sugeno dalam prediksi penentuan harga mobil bekas," *Indonesian Journal of Mathematics and Natural Science*, vol. 44, pp. 73–80, 2021.
[6] M. R. Azizi, "Optimasi fuzzy Tsukamoto dalam memprediksi curah hujan di Kabupaten Banyuasin menggunakan algoritma artificial bee colony," unpublished.
[7] D. M. N. Fajri, W. F. Mahmudy, Y. P. Anggodo, "Optimization of FIS Tsukamoto using particle swarm optimization for dental disease identification," *2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, pp. 261–268, 2017.
[8] Afrilia, "Optimasi fungsi keanggotaan fuzzy inference system Tsukamoto dengan particle swarm optimization untuk prediksi cuaca di Palembang," unpublished.
[9] P. Paramitha, "Optimasi derajat keanggotaan fuzzy Tsukamoto menggunakan genetic algorithm untuk menentukan kecukupan gizi pada pola makanan balita," unpublished.
[10] V. N. Wijayaningrum, *Konsep dan PenerapanAlgoritmaGenetika*. Banda Aceh: Syiah Kuala University Press, 2021.
[11] I. S. Bimawijaya, "Optimasi fuzzy Tsukamoto dua tahap menggunakan algoritma genetika untuk seleksi calon karyawan," unpublished.
[12] A. Josi, "Implementasi algoritma genetika pada aplikasi penjadwalan perkuliahan berbasis web dengan mengadopsi model waterfall," *Jurnal Informatika: Jurnal Pengembangan IT (JPIT)*, vol. 2, pp. 77–83, July 2017.
[13] I. Wahyuni and W. F. Mahmudy, "Rainfall prediction in Tengger Indonesia using hybrid Tsukamoto FIS and genetic algorithm," *Journal of ICT Research and Applications*, vol. 11, pp. 38–54, April 2017.
[14] M. Rofiq et al., "Integrating fuzzy logic and genetic algorithm for upwelling prediction in Maninjau Lake," *Telecommunication Computing Electronics and Control*, vol. 17, pp. 226–234, February 2019.
[15] Q. Kotimah, W. F. Mahmudy, V. N. Wijayaningrum, "Optimization of fuzzy Tsukamoto membership function using genetic algorithm to determine the river water," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 7, pp. 2838–2846, October 2017.
[16] Sujarwata, *Buku Ajar Sistem Fuzzy dan Aplikasinya*. Yogyakarta: Deepublish, 2018.
[17] Z. Zukhri, *AlgoritmaGenetika: Metode Komputasi Evolusioner untuk Menyelesaikan Masalah Optimasi*. Yogyakarta: Penerbit Andi, 2014.
[18] Y. Arkeman, K. B. Seminar, and H. Gunawan, *Algoritma Genetika: Teori dan Aplikasinya untuk Bisnis dan Industri*. Bogor: IPB Press, 2012.