Computer Science and Information Technologies
Vol. 6, No. 2, July 2025, pp. 136~146
ISSN: 2722-3221, DOI: 10.11591/csit.v6i2.p136-146

Arowana cultivation water quality forecasting with multivariate fuzzy timeseries and internet of things

Alauddin Maulana Hirzan, April Firman Daru, Lenny Margaretta Huizen
Department of Information Technology, Faculty of Information Technology and Communication, Semarang University, Semarang, Indonesia

Article Info

Article history:

Received Feb 8, 2025
Revised May 5, 2025
Accepted May 23, 2025

Keywords:

Arowana
Internet of things
Multivariate fuzzy time series
Prediction
Water quality

ABSTRACT

Water quality plays a crucial role in the growth and survival of arowana fish, with imbalances in key parameters (pH, temperature, turbidity, dissolved oxygen, and conductivity) leading to increased mortality rates. While previous studies have introduced various monitoring models using Arduino IDE and intrinsic approaches, they lack predictive capabilities, leaving cultivators unable to take proactive measures. To address this gap, this study develops a predictive model integrating the internet of things (IoT) with a fuzzy time series (FTS) algorithm. Through rigorous evaluation and validation, the proposed FTS-multivariate T2 model demonstrated superior performance, achieving an exceptionally low error rate of 0.01704%, outperforming decision tree (0.13410%), FTS-multivariate T1 (0.88397%), and linear regression (20.91791%). These findings confirm that FTS-multivariate T2 not only accurately predicts water quality but also significantly reduces the mean absolute percentage error, providing a robust solution for sustainable arowana aquaculture.

Corresponding Author:

Alauddin Maulana Hirzan
Department of Information Technology, Faculty of Information Technology and Communication
Semarang University
Semarang, Indonesia
Email: maulanahirzan@usm.ac.id

1. INTRODUCTION

Arowana fish are a type of decoration fish that people usually keep to furnish their houses. People believe that keeping them may bring prosperity and wealth to the owner. Because of that simple reason, many fish farmers have started cultivating this type of fish. This fish is also known by the scientific name Scleropages formosus and dragon fish in Asian countries. However, cultivating arowana fish is not an easy task for the cultivator. To allow arowana fish to grow optimally, a specific range for each parameter: pH, temperature, dissolved oxygen, conductivity, and turbidity is required [1]–[3]. If the cultivators fail to keep these parameters balanced, this may render arowana fish growth and may cause death for the fish [4]. To mitigate this problem, many cultivators measure the water quality inside the water manually with some sensors. However, the conditions inside the water are sometimes unpredictable. The cultivators cannot predict when one or more water quality parameters are below or over the threshold. For that reason, many studies proposed a monitoring model to detect water quality with internet of things (IoT) technology [5]–[7]. This technology allows cultivators to monitor the water quality within the cultivation area automatically and remotely. So, the cultivators only need to come whenever the parameters are almost below or over the threshold.

The study from 2020 proposed a model where the model is equipped with an ultrasonic sensor and Arduino UNO. This model is capable of the condition of the aquarium and reporting the result to many devices [8]. This model was then improved in the next study of 2021. The proposed model in that year is equipped with pH, temperature, and turbidity sensors. The model successfully monitors the aquarium with minimum effort [9]. The model development did not stop there and still improving. In 2022, the next study improved the previous model by designing IoT-based water quality monitoring (SIMONAIR) [10]. The latest model was proposed in 2023 where some studies proposed a model with better accuracy. This article [7] showed that its model has a low error rate of up to 1% compared to the common sensor. There is one more article [11] with a model connected to the Thingspeak service that has accurate measurements.

Based on the previous paragraph, this study analyzed the previous models and found common weaknesses. The first problem is about the processing model used by the previous model. The proposed model from the article used Arduino UNO as the processing board. This board was not equipped with wireless communication and only worked locally [12]. Thus, the board must be connected to an additional component to allow communication to the Internet. The second problem is that there is no machine-learning algorithm to assist the model in predicting future conditions. Previous models were only limited to monitoring in real-time, but they cannot predict future water quality. Without future predictions, the cultivators cannot mitigate the future outcomes that might occur in the cultivation areas [13]. With the limitations of previous models, the cultivator may suffer severe economic loss if arowana's mortality rate increases.

The current research gap that exists within the previous studies is the missing prediction algorithm to support future prediction based on time-series data. Thus, the cultivators can mitigate what will occur in the future. For that reason, this study has the purpose of solving the problem in the previous paragraph by externally implementing a fuzzy timeseries multivariate (FTS-MV) algorithm as the prediction algorithm for the IoT. This algorithm is suitable for time series-based data and is often implemented in many situations. For example, an article [14] published in 2020 implemented a fuzzy time series (FTS) for predicting non-stationary environment data. In different articles [15], this algorithm is also implemented in solar energy prediction. The last article [16] within the same year 2020 also uses a fuzzy time-series algorithm to predict the air quality index. These articles prove that the FTS is implementable and capable to predict water quality in arowana's cultivation.

2. METHOD

In this section, this study explains how to gather and prepare the required data before designing the proposed IoT model and equipping it with a fuzzy time-series algorithm. The first step is to gather the required data from arowana's cultivation. In this case, this study uses a monitoring IoT model equipped with several sensors like PH-4502C (water acidity sensor), analog total dissolved solid (water conductivity sensor), DS18B20 (water temperature sensor), dissolved oxgen and turbidity sensors. The schematic in Figure 1 is the illustration for the data gathering as well as the prediction node.

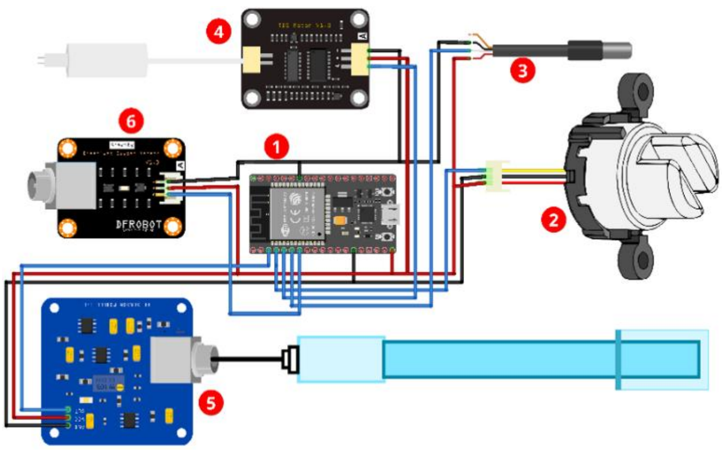


Figure 1. The schematic for monitoring and prediction model

Figure 1 is the illustration of the model's schematic. There are several components installed in the model: i) ESP32 processing board equipped with WiFi networking, ii) turbidity sensor, iii) DS18B20 temperature sensor, iv) analog TDS sensor, v) PH-4502C water acidity sensor, and vi) dissolved oxygen sensor. With that node, this study gathers the data for two days with an interval of five seconds between data in a small-sized aquarium with one arowana. Each recorded data has a timestamp in it to show when the data is recorded. After gathering the data, this study obtained 34,560 rows in a CSV format for easier access during the training phase. Table 1 contains the sample of gathered data.

Table 1 contains the sample from gathered data over two days. The first column is the timestamp of each data in UNIX epoch format [17]. Then the next columns followed by pH, temperature, turbidity, dissolved oxygen, and conductivity. After obtaining the data, this study continues the step to fuzzify the data to obtain water quality. The water quality output is in regression format [18], [19]. This study configures the fuzzy logic to produce the output between the 0 to 100 ranges. The contains of membership configurations for fuzzification process is shown in Table 2.

Table 2 contains the fuzzification table to obtain water quality. There are five features on the table as the input and one feature as the output. Each feature is divided into three different configurations that act as a threshold. The pH has three different configurations: acid, neutral and alkaline. The temperature has cold, warm and hot configurations. The turbidity, dissolved oxygen and conductivity share similar configurations: low, medium, and high. Meanwhile, the output membership (quality) has different configurations: poor, fair, and good. After the fuzzification process, the dataset will have an additional column called quality with range between 0 to 100. The next step is to train the fuzzy time-series multivariate model. This algorithm is similar to other FTS. However, this algorithm utilizes multiple features to predict instead single feature. Similar to multivariate linear regression, but for time series dataset.

This study creates two simple Python scripts that imports the PyFTS library to create two FTS multivariate models [20]. This study names the model with FTS-multivariate T1 and T2 based on the dataset’s degree of differentials. The first model (FTS-multivariate T1) was trained with the dataset's first differential degree. In contrast, the second model (FTS-multivariate T2) was trained with the dataset's second differential degree. Higher differential degrees lead to more stationary and consistent time series patterns. After that, these models are exported into binary format for server use. Figure 2 illustrates the prediction mechanism from an IoT node to the server and its database.

Table 1. Sample of gathered data

Timestamp	pH	Temperature	Turbidity	Dissolved oxygen	Conductivity
1736323200.00	5.49176	21.34176	2.74176	6.04176	340.2418
1736323205.00	7.873051	22.89593	6.069674	13.98629	328.2109
1736323210.00	8.35498	30.04688	6.934142	7.452285	371.6803
1736323215.00	5.145914	19.41082	3.020109	0.817705	243.9608
			….		
1736495985.00	8.255449	19.80852	5.491624	11.81403	416.3683
1736495990.00	9.418517	24.36847	4.489872	12.21427	298.1736
1736495995.00	5.769331	17.61775	3.455537	10.21275	278.4579

Table 2. Membership functions for quality fuzzification

Category	Indicator	Configuration 1	Configuration 2	Configuration 3
Input	pH	0–6.9 (Acid)	6.8–7.2 (Neutral)	7.1–14 (Alkaline)
	Temperature	0–25 (Cold)	24–35 (Warm)	34–100 (Hot)
	Turbidity	0–20 (Low)	15–60 (Medium)	55–1000 (High)
	Dissolved oxygen	0–4 (Low)	3–9 (Medium)	8–15 (High)
	Conductivity	0–120 (Low)	100–2020 (Medium)	2000–5000 (High)
Output	Quality	0–35 (Poor)	32–75 (Fair)	72–100 (Good)

Figure 2 explains the prediction mechanism from the proposed model, starting from an IoT node sending five parameters to the server through the ReST protocol [21], [22]. When the server receives the data, then the server does the prediction with the previously exported FTS model. The server then stores the prediction result in a database.

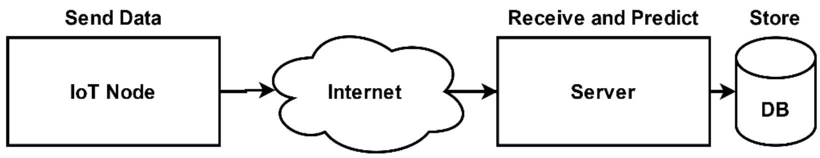


Figure 2. Prediction process from a IoT node to the server

To evaluate the proposed model, this study compares the result with another algorithm like multi-variate linear regression and decision tree. This study utilizes a statistical approach by calculating mean absolute error (MAE), mean absolute percentage error (MAPE), mean squared error (MSE), root mean squared error, R-squared (R^2), and adjusted R^2. MAE shows the average size of the errors between actual and predicted values. MAPE shows the average error as a percentage, which helps compare the results. MSE finds the average of the squared errors, giving more weight to significant errors. Root mean squared error (RMSE) gives a result in the same units as the data. R-squared (R^2) tells how much of the variation in the data. Adjusted R^2 fixes this by lowering the score. This approach is more suitable than the confusion matrix to measure the error rate [23], [24]. In (1)-(6) for each evaluation is shown as follows:

$$\mathrm{MAE} = \frac{1}{total\ data}\sum_{i=1}^{total\ data}|actual_i - prediction_i| \tag{1}$$

$$\mathrm{MAPE} = \frac{1}{total\ data}\sum_{i=1}^{total\ data}\left|\frac{actual_i - prediction_i}{actual_i}\right| x100 \tag{2}$$

$$\mathrm{MSE} = \frac{1}{total\ data}\sum_{i=1}^{total\ data}(actual_i - prediction_i)^2 \tag{3}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{total\ data}\sum_{i=1}^{total\ data}(actual_i - prediction_i)^2} \tag{4}$$

$$R^2\ Score = 1 - \frac{\sum_{i=1}^{total\ data}(actual_i - prediction_i)^2}{\sum_{i=1}^{total\ data}(actual_i - \overline{actual})^2} \tag{5}$$

$$R^2{}_{Adjusted} = 1 - \left(\frac{(1-R^2)(total\ data - 1)}{total\ data - total\ features - 1}\right) \tag{6}$$

Each variable has different meaning, where i refers to row number of datasets. The total data refers to the row number for each actual and prediction values. The actual and prediction refer to dataset stored inside the database. Meanwhile, total features refer to the number of features in a dataset. This study uses two different approaches to validate the result. The first one creates a baseline from the stored data, and the second one uses cross-validation (5 folds) results. The first approach uses a baseline for comparison with the proposed models, multi-variate linear regression [25], [26] and decision tree [27]. These algorithms are often used in many situations, including water quality predictions. Based on articles [28], [29] in 2022, both linear regression and decision tree were implemented to predict water quality in different studies. These articles are solid evidence of the application of both algorithms in water quality predictions. Meanwhile, the second approach uses the validation method to calculate the accuracy and error percentages from the proposed models trained with different lengths of training and test data.

3. RESULTS AND DISCUSSION

In this section, this study explains the results that are obtained from the evaluation and validation phases. The first explanation is about the prediction results stored in a database. Then, the second explanation is about the evaluation results where this study compares with another regression algoritms. The last explanation is about the validation with fuzzy logic as the baseline to strengthen the evaluation results. Table 3 contains the sample of the fuzzy logic water quality baseline together with predicted values from two FTS models, multivariate linear regression, and decision trees.

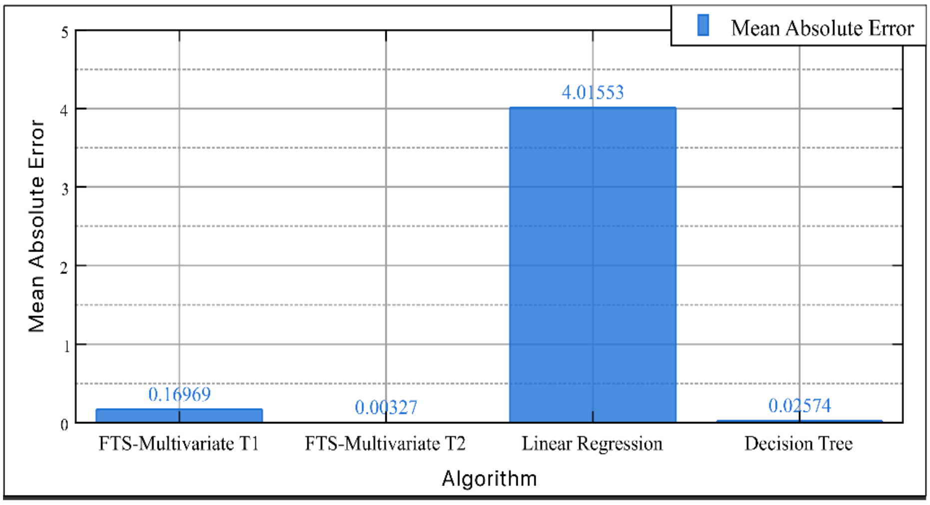
Table 3 contains the water quality predictions stored in a database. This table has several columns: The baseline column was obtained from fuzzifying features with fuzzy logic that used Table 2 as the configuration. The FTS-multivariate T1 was the first FTS multivariate model with one degree of differential. Meanwhile, the FTS-multivariate T2 was the second model with two degrees of differential. Besides that, there were two more columns: linear regression (that operated in multivariate) and decision tree. However, evaluation cannot be done alone with a table. Thus, this study evaluated the results with statistical approaches. There are six components of evaluation that this study has done to measure the performance of all models.

Figure 3 explains the evaluation results based on MAE and MAPE evaluations. Specifically, Figures 3(a) and 3(b) present the detailed results of MAE and MAPE from all algorithms. This evaluation determined the regression accuracy between the baseline and the prediction result. Thus, lower result is the target of the evaluation. According to the results, FTS-multivariate T2 has the lowest evaluation results where MAE was 0.0033 and MAPE was 0.017%. The second place was the decision tree model with MAE

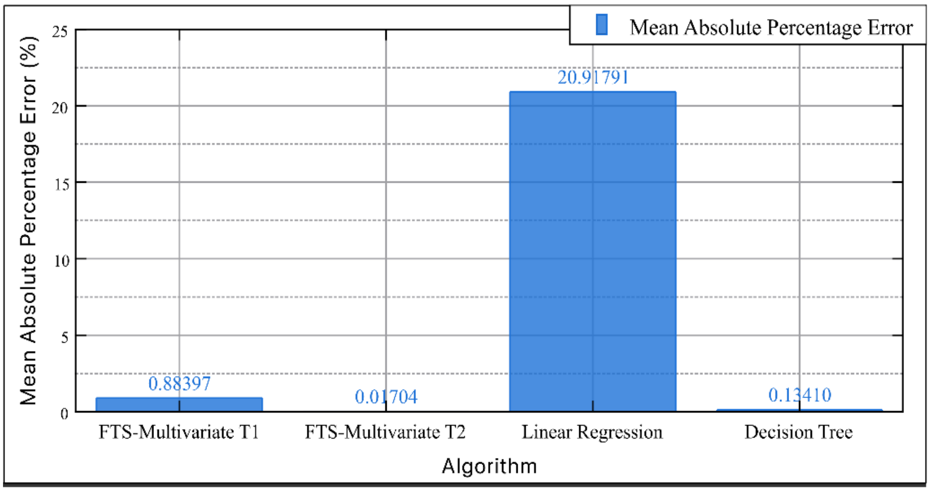
0.0257 and MAPE 0.1341%. Followed in the third place was FTS-multivariate T1 with MAE 0.1697 and MAPE 0.8839%. The last place for this evaluation was linear regression with MAE 4.0155 and MAPE 20.9179%. From these results, this study has found the highest regression accuracy model. However, the evidence was too shallow to decide which model was the best. The next evaluation was a MSE and RMSE. These evaluations were needed to evaluate the difference between the results with larger error penalties.

Table 3. Sample of water quality prediction results

Baseline	FTS-multivariate T1	FTS-multivariate T2	Linear regression	Decision tree
16.9474	16.95011	16.9474	18.52947	16.94693
16.76119	19.10911	16.76119	18.88292	16.76119
16.95728	16.95999	16.95728	18.73887	16.95743
		…		
16.76119	16.76391	16.76119	18.82109	16.76119
17.03074	17.03346	17.03074	18.90648	17.03745
16.95283	20.03482	16.95283	18.96535	16.95256



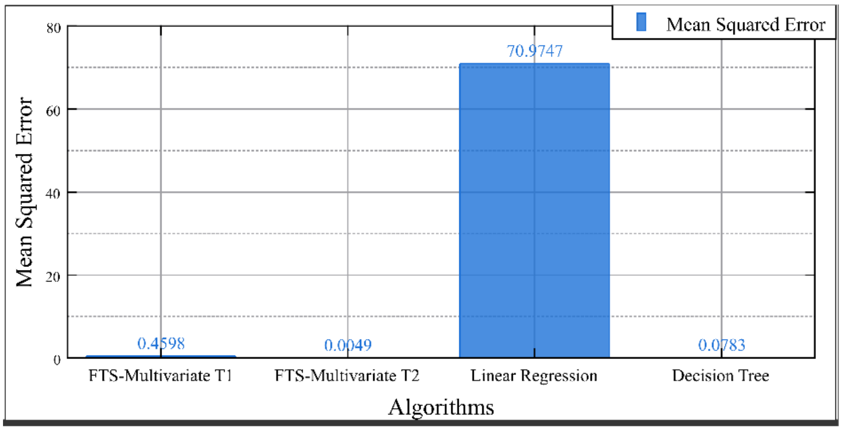
(a)



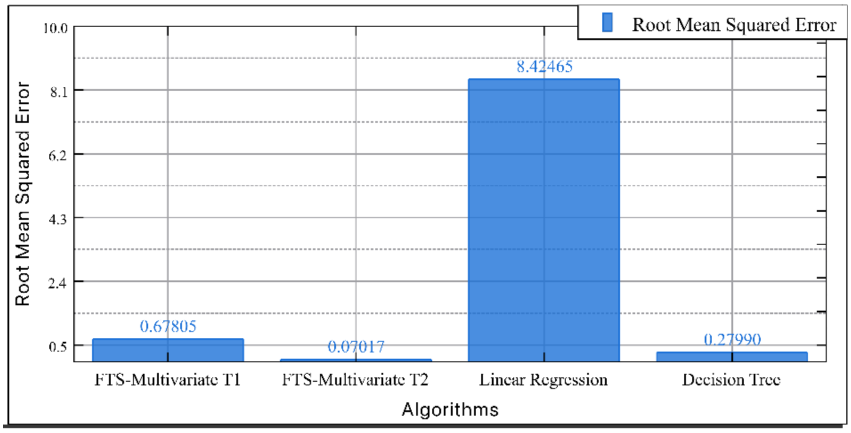
(b)

Figure 3. Evaluation results from (a) MAE and (b) MAPE

Figure 4 explains the evaluation results from MSE and RMSE aspects. Figure 4 explains the difference between the prediction and the baseline, with a larger penalty for error. Thus, the lower result also means that the penalty of error is low as well. Figure 4(a) was the result of the MSE, and Figure 4(b) was the result of the RMSE. In Figure 4(a), this study found that FTS-multivariate T2 has the lowest error penalty of 0.0049. The second place was Decision tree with an error penalty of up to 0.0783, followed by FTS-multivariate T1 with an error penalty of up to 0.4598. The last place was linear regression with an error penalty up to 70.9747. Figure 4(b) shows the simpler interpretation of Figure 4(a), where the results are similar to previous explanations.



(a)



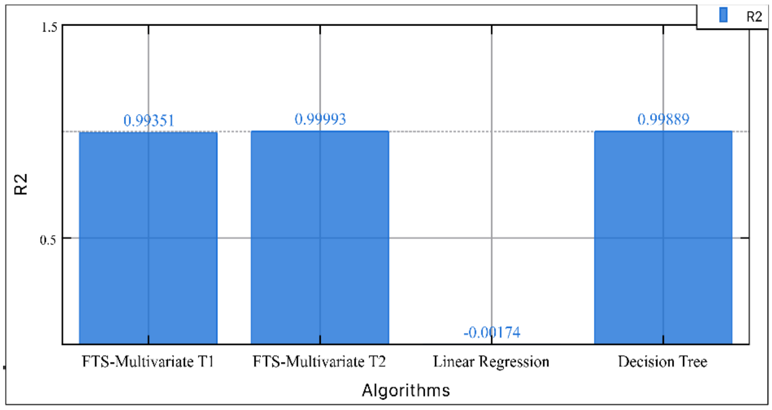
(b)

Figure 4. Evaluation results from (a) MSE and (b) RMSE

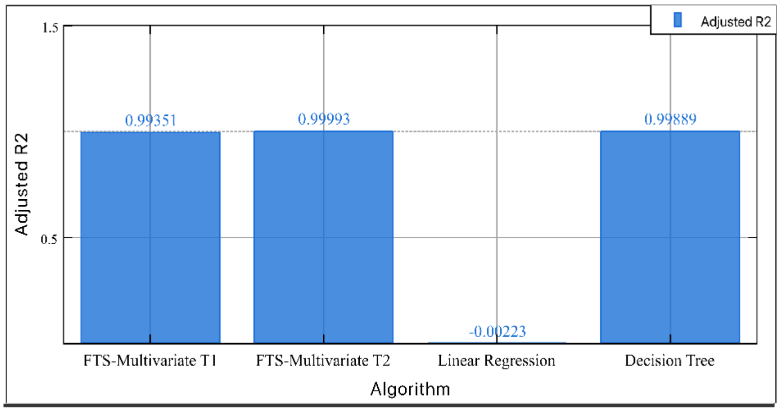
The next evaluation is R^2 and adjusted R^2. Both evaluations are used to evaluate how well the independent variable explains the variety of the dependent variables. Adjusted R^2 is more focused on the number of features to ensure a fair evaluation. Hence, a higher result is recommended. Figure 5 explains the variance results from each model for R^2 and adjusted R^2 evaluations.

According to Figure 5, this study found that all models except linear regression have high variance. As shown in Figure 5(a), all models except linear regression have near-perfect variance reaching 0.99. Precisely, 0.99993 for FTS-multivariate T2, 0.99889 for decision tree, and 0.99351 for FTS-multivariate T1. Meanwhile, the linear regression model failed to determine the variance with result -0.00174. Figure 5(b) shared a similar result with Figure 5(a) except for linear regression, where its result is still the lowest with -0.00223. To validate the evaluation results, this study compared the prediction results side-by-side with the baseline. Figure 6 shows the models' prediction results in comparison with the baseline. Since the prediction results were too many, this study took 100 sequenced rows as a sample and plotted it into a graph.

Figure 6 shows the prediction comparison results of 100 samples between the baseline and the model's predictions. According to the timestep sample shown in Figure 6, information at timestep 27 showed that all models were close to the baseline except the linear regression model. FTS-multivariate T2 has precise prediction with 29.2362, followed by FTS-multivariate T1 with 29.2389 and decision tree with 30.5335. Meanwhile, linear regression was far from prediction with result 19.0401. Thus, the result in Figure 6 validated all evaluation results and showed that FTS-multivariate T2 has accurate regression predictions. This study used the cross-validation method to evaluate the consistency of prediction accuracy. By applying 5-fold validation, the average accuracy for FTS-multivariate T2 was 99.98%, with an error percentage of 0.016%. In contrast, the accuracy for FTS-multivariate T1 was 99.13%, accompanied by an error rate of 0.867%. These results indicate that the accuracy of FTS multivariate models remains both high and stable, even when different lengths of datasets are used for training.



(a)



(b)

Figure 5. Evaluation results from (a) R^2 and (b) adjusted R^2

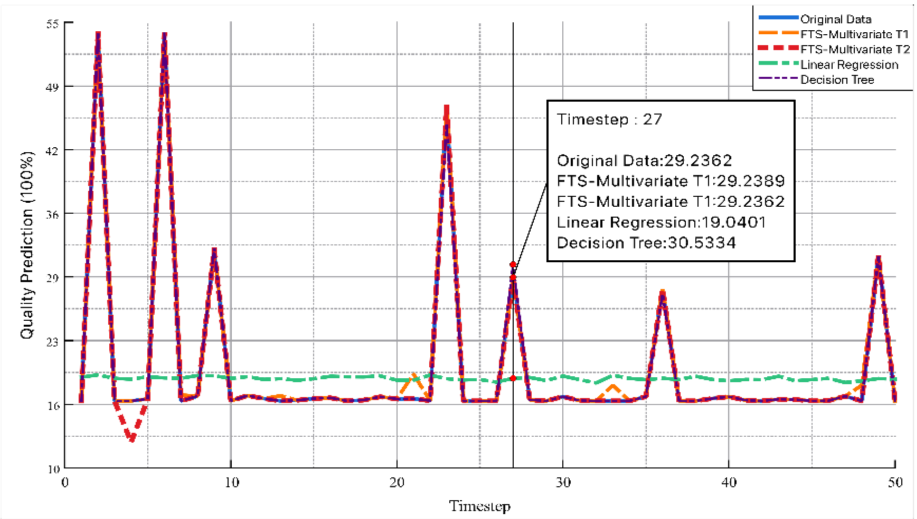


Figure 6. Prediction comparison sample results with the baseline (original data)

After explaining the evaluation and validation results, this study continues to the discussion part. In the discussion part, this study explains several points: result interpretations, comparison with past studies, implications, strength and weakness, and future study. The first discussion is about the results obtained. According to the evaluation and validation results, the proposed models showed the lowest mean absolute percentage error compared to other models. The proposed model of FTS-multivariate T2 has the lowest error of 0.0174%. This result means that the proposed model has high accuracy reaching 99%. This explanation is proved in the validation phase, where the prediction result is similar to the baseline. The runner-up based on the lowest error percentage was the decision tree model, followed by FTS-multivariate T1 and linear regression. There were other aspects like mean squared error, root mean squared error, R^2, and Adjusted R^2. The results in these aspects were similar with mean absolute percentage error. The best model was FTS-multivariate T2, followed by decision tree, FTS-multivariate T1, and linear regression.

The second discussion is about the comparison with the past models. As explained in the introduction section, the past models could not predict future situations. Thus, the cultivators cannot predict what will happen in the future. This problem has been solved with the proposed model, where this model (especially FTS multivariate T2) accurately predicts the water quality. This result has been validated with the water quality baseline. Thus, the proposed model performed better than past models with the capability to predict the water quality in the future.

The third discussion is the implication of this study in the theoretical and practical areas. The proposed model was a piece of evidence for fuzzy-based prediction. Most scholars know that fuzzy logic is mostly implemented to translate any numeric input from a device or a node into human interpretation. However, the capability of fuzzy algorithms did not stop there. A team of programmers improved the fuzzy algorithms and turned it into a regression predicting algorithm. This study has successfully proved this algorithm's accuracy by implementing it as a water quality prediction. The second implication is toward practical areas.

This study has validated its proposed models and proved how accurate the prediction was. This proposed model (FTS-multivariate T2) is implementable in arowana cultivations. It can help cultivators mitigate future water quality conditions more accurately and reduce the number of dead fish. It will lead to a better economy for the cultivator by reducing the number of dead fish. For instance, the cultivator is now able to monitor water quality proactively. This predictive approach leads to a reduced error margin and greater accuracy.

The fourth discussion is about the strengths and weaknesses of the proposed models. Based on the evaluation and validation phases, this study found that the proposed model (especially FTS-multivariate T2) has the highest accuracy compared to another algorithm. Meanwhile, the other algorithm (FTS-multivariate T1) performed poorly below the decision tree model. There was a reason why FTS-multivariate T2 performed better than FTS-multivariate T1. The key was in the degree of differentials. The degree of differential was used to remove trends inside the dataset and make it more stationary. FTS-multivariate T2 predicted more accurately than FTS-multivariate T1 is caused by leftover trends in the first degree of differential (the second degree of differential offered a cleaner and more stationary dataset). Thus, FTS-multivariate T2 can understand the seasonality of the dataset more than FTS-multivariate T1. One of this algorithm's strengths is its scalability. The FTS-multivariate model (both T1 and T2) demonstrates considerable flexibility regarding scaling. It can be implemented in a larger aquarium. As long as the necessary dataset and sensors are available, it can be scaled up for more extensive applications with only minor adjustments, such as aggregating data from multiple sensors.

However, the detection range depends entirely on the type and quality of the sensors used. However, both models suffered similar weaknesses. Both models required a time series type of dataset. An image dataset is an example of a dataset that is difficult to be used with this algorithm. Hence, this algorithm is unsuitable for that type of data. The second weakness of this proposed model is its limited application. Since it was curated with arowana's dataset and parameters, it might be unsuitable for another type of fish. Thus, retraining the model with a proper dataset is recommended. The last discussion is about the future possibility of this study. The proposed models are still growing. There are many chances to improve the current model by implementing numerous time series-based algorithms. For example, long short-term memory, autoregressive integrated moving average, or seasonal autoregressive integrated moving average.

Based on the discussions, this study can conclude that the proposed models are successfully capable to predict the water quality. The FTS-multivariate T2 model is the best model with the highest accuracy compared to other algorithms. Followed by decision tree, FTS-multivariate T1 and linear regression. In summary, this predictive model serves as a sophisticated resource for enhancing water quality management in arowana aquaculture, facilitating the adoption of more sustainable practices.

4. CONCLUSION

Water quality is an important aspect that affects arowana’s growth. Failing to balance the five parameters (pH, temperature, turbidity, dissolved oxygen, and conductivity) may render arowana's growth and increase the number of dead fish. Many past studies proposed many different models to mitigate this problem. Some proposed a monitoring model with Arduino IDE, and some used an intrinsic approach to make the monitoring results easier to read. However, there is a problem with the past models. They were not equipped with a prediction algorithm to predict what would happen in the future. Thus, the cultivators cannot mitigate when a situation occurs. To solve this problem, this study designed a prediction model based on the IoT combined with a FTS algorithm. Based on the evaluation and validation, the proposed models (especially FTS-multivariate T2) achieved a low percentage of error reaching 0.01704%. Followed by decision tree 0.13410, FTS-multivariate T1 with 0.88397 and linear regression 20.91791. These results also aligned with the baseline in the validation phase. This study concluded that the proposed model (FTS-multivariate T2) is not only capable of predicting water quality but also offers lower mean absolute percentage error compared to other algorithms.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to Lembaga Penelitian dan Pengabdian Kepada Masyarakat (LPPM) Semarang University for the invaluable support and resources provided throughout this research. We deeply appreciate the guidance, facilities, and academic environment that contributed to the successful completion of this study. Special thanks to the faculty members, research staff, and fellow colleagues for their constructive feedback and encouragement. This work would not have been possible without the university’s commitment to fostering innovation and scientific exploration.

FUNDING INFORMATION

This study is funded by Lembaga Penelitian dan Pengabdian Kepada Masyarakat Semarang University under contract number: 083/USM.H7.LPPM/L/2024

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Alauddin Maulana Hirzan		✓	✓	✓	✓		✓	✓	✓	✓			✓	✓
April Firman Daru	✓	✓			✓	✓	✓		✓			✓	✓	
Lenny Margaretta Huizen			✓	✓		✓		✓		✓	✓			✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

ETHICAL APPROVAL

The research related to animal use has been complied with all the relevant national regulations and institutional policies for the care and use of animals.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, AMH, upon reasonable request.

REFERENCES

[1] G. H. Yue, A. Chang, A. Yuzer, and A. Suwanto, "Current knowledge on the biology and aquaculture of the endangered Asian arowana," *Reviews in Fisheries Science & Aquaculture*, vol. 28, no. 2, pp. 193–210, 2020, doi: 10.1080/23308249.2019.1697641.

[2] S. R. Medipally, F. M. Yusoff, N. Sharifhuddin, and M. Shariff, "Sustainable aquaculture of Asian arowana - a review," *Journal of Environmental Biology*, vol. 37, no. 4, pp. 829–838, 2016.

[3] A. S. Panjaitan, F. Hapsari, M. Margono, and H. Al-Qoyyimah, "Variety types of natural food fish usage in the growth of super red arowana fish (Scleropages formosus)," *Eduvest-Journal of Universal Studies*, vol. 3, no. 4, pp. 842–849, Apr. 2023, doi: 10.59188/eduvest.v3i4.796.

[4] J. W. Senders and N. P. Moray, *Human error: cause, prediction, and reduction*, CRC Press, 2020.

[5] J. Trevathan and D. Nguyen, "Adaptive IoT technology for measuring salinity, dissolved oxygen, and pH in aquatic environments," *International Journal of Hyperconnectivity and the Internet of Things*, vol. 6, no. 1, pp. 1–20, Mar. 2022, doi: 10.4018/IJHIoT.294894.

[6] A. Rehman, T. Saba, M. Kashif, S. M. Fati, S. A. Bahaj, and H. Chaudhry, "A revisit of internet of things technologies for monitoring and control strategies in smart agriculture," *Agronomy*, vol. 12, no. 1, pp. 1-21, 2022, doi: 10.3390/agronomy12010127.

[7] A. F. Daru, F. W. Christanto, and V. Vydia, "Internet of things-based water pH level monitoring for arowana cultivation," *IOP Conference Series: Earth and Environmental Science*, vol. 1177, no. 1, 2023, doi: 10.1088/1755-1315/1177/1/012004.

[8] J. E. da Costa and A. W. R. Emanuel, "IOT design monitoring water tank study case: Instituto profissional de canossa (IPDC)," *International Conference on Science and Applied Science (ICSAS) 2020, AIP Conference Proceedings*, 2020, doi: 10.1063/5.0032638.

[9] M. F. Suhaimi, N. Huda, M. Tahir, S. N. Mohamad, and S. R. Aw, "IOT based automatic aquarium monitoring system for freshwater fish," *International Journal of Synergy in Engineering and Technology*, vol. 2, no. 1, pp. 125–133, 2021. [Online]. Available: https://tatiuc.edu.my/ijset/index.php/ijset/article/view/95

[10] W. Sholihah, A. Hendriana, I. Kusumanti, and I. Novianty, "Design of IoT based water monitoring system (simonair) for arwana fish cultivation," *Eduvest-Journal of Universal Studies*, vol. 2, no. 12, pp. 2872–2884, Dec. 2022, doi: 10.59188/eduvest.v2i12.708.

[11] F. L. Toruan and M. Galina, "Internet of things- based automatic feeder and monitoring of water temperature, pH, and salinity for litopenaeus vannamei shrimp," *Jurnal ELTIKOM*, vol. 7, no. 1, pp. 9–20, Jun. 2023, doi: 10.31961/eltikom.v7i1.658.

[12] A. C. Bento, "IoT: nodeMCU 12e X arduino uno, results of an experimental and comparative survey," *International Journal of Advance Research in Computer Science and Management Studies*, vol. 6, no. 1, pp. 46–56, 2018. [Online]. Available: https://sci-hub.se/https://pdfs.semanticscholar.org/84c4/71470cc3c8e27807df60aa5dde2888cc3b3f.pdf%0Awww.ijarcsms.com.

[13] B. Strauch, "John senders, human error, and system safety," *Human Factors*, vol. 65, no. 5, pp. 766–778, Aug. 2023, doi: 10.1177/00187208211001982.

[14] P. C. de L. e Silva, C. A. Severiano, M. A. Alves, R. Silva, M. W. Cohen, and F. G. Guimarães, "Forecasting in non-stationary environments with fuzzy time series," *Applied Soft Computing Journal*, vol. 97, 2020, doi: 10.1016/j.asoc.2020.106825.

[15] O. Orang, R. Silva, P. C. de L. e Silva, and F. G. Guimaraes, "Solar energy forecasting with fuzzy time series using high-order fuzzy cognitive maps," in *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Jul. 2020, pp. 1–8, doi: 10.1109/FUZZ48607.2020.9177767.

[16] Y. Alyousifi, M. Othman, R. Sokkalingam, I. Faye, and P. C. L. Silva, "Predicting daily air pollution index based on fuzzy time series markov chain model," *Symmetry*, vol. 12, no. 2, p. 293, Feb. 2020, doi: 10.3390/sym12020293.

[17] G. Cox and J. Lund, "Time.now," in *Uncertain Archives: Critical Keywords for Big Data*, MIT Press, 2021, pp. 523–532, doi: 10.7551/mitpress/12236.003.0057.

[18] G. Hesamian and M. G. Akbari, "A robust multiple regression model based on fuzzy random variables," *Journal of Computational and Applied Mathematics*, vol. 388, May 2021, doi: 10.1016/j.cam.2020.113270.

[19] J. Chachi, S. M. Taheri, and P. D'Urso, "Fuzzy regression analysis based on M-estimates," *Expert Systems with Applications*, vol. 187, Jan. 2022, doi: 10.1016/j.eswa.2021.115891.

[20] P. C. de L. e Silva *et al.*, "PYFTS/pyFTS: Stable version 1.7," *Zenodo*, 2019, doi: 10.5281/zenodo.7860231.

[21] I. Ahmad, E. Suwarni, R. I. Borman, Asmawati, F. Rossi, and Y. Jusman, "Implementation of RESTful API web services architecture in takeaway application development," in *2021 1st International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS)*, Oct. 2021, pp. 132–137, doi: 10.1109/ICE3IS54102.2021.9649679.

[22] A. Ehsan, M. A. M. E. Abuhaliqa, C. Catal, and D. Mishra, "RESTful API testing methodologies: rationale, challenges, and solution directions," *Applied Sciences*, vol. 12, no. 9, Apr. 2022, doi: 10.3390/app12094369.

[23] D. S. K. Karunasingha, "Root mean square error or mean absolute error? use their ratio as well," *Information Sciences*, vol. 585, pp. 609–629, Mar. 2022, doi: 10.1016/j.ins.2021.11.036.

[24] T. O. Hodson, "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not," *Geoscientific Model Development*, vol. 15, no. 14, pp. 5481–5487, Jul. 2022, doi: 10.5194/gmd-15-5481-2022.

[25] J. M. Hahne *et al.*, "Linear and nonlinear regression techniques for simultaneous and proportional myoelectric control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 2, pp. 269–279, Mar. 2014, doi: 10.1109/TNSRE.2014.2305520.

[26] D. M. Blei, "Linear regression, logistic regression, and generalized linear models," Columbia University, 2014.

[27] J. K. Afriyie *et al.*, "A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions," *Decision Analytics Journal*, vol. 6, Mar. 2023, doi: 10.1016/j.dajour.2023.100163.

[28] A. Mokhtar, A. Elbeltagi, Y. Gyasi-Agyei, N. Al-Ansari, and M. K. Abdel-Fattah, "Prediction of irrigation water quality indices based on machine learning and regression models," *Applied Water Science*, vol. 12, no. 4, pp. 1-15, 2022, doi: 10.1007/s13201-022-01590-x.

[29] F. Gorgan-Mohammadi, T. Rajaee, and M. Zounemat-Kermani, "Decision tree models in predicting water quality parameters of dissolved oxygen and phosphorus in lake water," *Sustainable Water Resources Management*, vol. 9, no. 1, Feb. 2023, doi: 10.1007/s40899-022-00776-0.

BIOGRAPHIES OF AUTHORS

Alauddin Maulana Hirzan holds a Master's degree in Computer Science with a specialization in Internetworking Technology, IoT, and Machine Learning from Universitas Semarang. His research focuses on the integration of IoT and artificial intelligence to develop innovative solutions in data analytics, predictive modeling, and smart systems. With a strong background in network technology and machine learning, he is dedicated to advancing intelligent automation and optimization in various technological domains. He can be contacted at email: maulanahirzan@usm.ac.id.

April Firman Daru is a Doctor of Computer Science specializing in IoT and Machine Learning from Universitas Semarang. His research focuses on the development of intelligent systems, predictive analytics, and IoT-based solutions to enhance automation and decision-making processes. With extensive expertise in machine learning algorithms and smart technology integration, he is committed to advancing innovative applications in computing and data science. He can be contacted at email: firman@usm.ac.id.

Lenny Margaretta Huizen holds a Master's degree in Computer Science with a specialization in Data Mining from Universitas Semarang. Her research focuses on data analysis, pattern recognition, and predictive modeling to extract valuable insights from complex datasets. With expertise in data-driven decision-making and machine learning applications, she is dedicated to advancing knowledge in data mining and its practical implementations across various industries. She can be contacted at email: lenny@usm.ac.id.