OPEN ACCESS
Indo-JC
ISSN 2460-9056
socj.telkomuniversity.ac.id/indojc

Ind. Journal on Computing
Vol. 8, Issue. 1, April 2023. pp. 23-36
doi:10.34818/indojc.2023.8.1.695

Time Series On-Board Air Quality Index

Benedictus Augusta Vianney [#1], Bayu Erfianto [*2]

 # Information Technology, Telkom University
Jl. Telekomunikasi. 1, Terusan Buahbatu – Bojongsoang, Telkom University, Sukapura, Kec. Dayeuhkolot, Kabupaten Bandung, Jawa Barat, Indonesia,

[1] benedictusveno@student.telkomuniversity.ac.id
[2] erfianto@telkomuniversity.ac.id

Abstract

With the rapid development of technology, sometimes individuals forget about their health, plus during the pandemic, the air quality in a room becomes more of a concern. Maintaining air quality to be healthy and good for humans to breathe is by keeping the amount of pollutants in the air, such as Carbon Dioxide (CO_2), Volatile Organic Compound (VOC), and Formaldehyde (HCHO), at a predetermined and agreed threshold is important. So in this research, we propose an on-board air quality index detection system for indoor and can forecast the air quality in the future. The system will use a Raspberry Pi 4 microcontroller for data acquisition and a WP6003 sensor device that will capture parameters for the air quality index. The system will run for 31 days to capture parameter data every 30 seconds. The parameter data is then analyzed using a correlation matrix to determine the parameters that affect each other. Then classified using fuzzy logic to determine the quality index based on the value of each parameter. The air quality index obtained is then forecast using the ARIMA and LSTM methods for the next 30 minutes. Then the forecasting accuracy is calculated using the RMSE and MAPE metrics. The correlation matrix results show that Carbon Dioxide (CO_2), Volatile Organic Compound (VOC), and Formaldehyde (HCHO) are related. Comparison of the forecasting results of the two methods concluded that the LSTM method outperformed ARIMA to forecast the air quality index for the next 30 minutes based on the previous 10 hours of data.

Keywords: Air Quality Index, Fuzzy, ARIMA, LSTM

Abstrak

Dengan perkembangan teknologi yang sangat cepat terkadang individu melupakan kesehatan mereka, ditambah di masa pandemi membuat kualitas udara di dalam sebuah ruangan menjadi perhatian lebih. Menjaga kualitas udara agar baik dan sehat untuk dihirup oleh manusia adalah dengan menjaga jumlah polutan pada udara, seperti Karbon Dioksida (CO_2), Volatile Organic Compound (VOC), dan Formaldehida (HCHO), pada ambang batas yang telah ditentukan dan disepakati menjadi hal yang penting dilakukan. Sehingga pada penelitian ini, kami mengusulkan sebuah sistem pendeteksi indeks kualitas udara on-board untuk dalam ruangan dan dapat meramalkan kualitas udara pada waktu yang akan datang. Sistem akan menggunakan microcontroller Raspberry Pi 4 untuk akuisisi data dan perangkat sensor WP6003 yang akan menangkap parameter untuk indeks kualitas udara. Sistem akan berjalan selama 31 hari untuk menangkap data parameter setiap 30 detik. Data parameter kemudian dianalisis menggunakan correlation matrix untuk mengetahui parameter yang saling mempengaruhi. Kemudian diklasifikasi menggunakan fuzzy logic untuk menentukan indeks kualitas berdasarkan nilai tiap parameter. Indeks kualitas udara yang diperoleh kemudian dilakukan peramalan menggunakan metode ARIMA dan LSTM untuk 30 menit yang akan datang. Kemudian akurasi peramalan dihitung menggunakan

Received on January, 2023. Accepted on March, 2023

metrik RMSE dan MAPE. Hasil correlation matrix diketahui bahwa Karbon Dioksida (CO_2), Volatile Organic Compound (VOC), dan Formaldehida (HCHO) saling berkaitan. Perbandingan hasil peramalan kedua metode disimpulkan bahwa metode LSTM mengungguli ARIMA untuk meramalkan indeks kualitas udara selama 30 menit ke depan berdasarkan data 10 jam sebelumnya.

Kata Kunci: Keywords: Air Quality Index, Fuzzy, ARIMA, LSTM

I. Introduction

WITH the rapid development of technology, people sometimes forget about important things for themselves, one of which is their health. Living in a good and healthy environment reduces the risk of falling ill. The indicator of a good and healthy environment is good air quality, which is comfortable to breathe. Furthermore, during the pandemic, the quality of the air in a room has become a greater concern for the community. Air contains several pollutants that are not good for inhalation, including Particulate Matter (PM_{10} and $PM_{2.5}$), Ozone (O_3), Carbon Dioxide (CO_2), Volatile Organic Compounds (VOCs) and Formaldehyde (HCHO) [1]. These elements are commonly used in air pollution, but there are other components of air pollution. In addition, air temperature and humidity can also be factors that affect indoor air quality. During a pandemic, it is important to maintain air quality so that it is good and healthy for people to breathe, by keeping the levels of the above-mentioned air pollutants within a predetermined and agreed threshold [2].

Several previous studies have conducted research on building fuzzy logic methods to determine the value of the air quality index and compare it with the value of the air quality index owned by the government [3]. Some studies have even used microcontrollers so that the parameter data for the air quality index is more accurate and can be directly monitored [4][5]. In addition, some studies have also conducted research to determine the indoor air quality index and provide actions based on the values obtained, such as turning on the fan or sending notifications when the air quality index is poor [6][7]. The LSTM method has good performance in forecasting the air quality index [8]. There have been several studies on forecasting the indoor air quality index with ARIMA, but its performance compared to other forecasting methods is unknown [10][12]. Forecasting of the indoor air quality index has also been done. A comparison of indoor air quality index forecasting methods to determine a suitable method is a research opportunity.

This study proposes a system that can detect the appropriate indoor air quality index and a method for predicting the air quality index. The first step is to obtain air quality index parameter data. The next step is to analyse the related parameters and through these parameters classification is performed. Finally, forecasting is performed using ARIMA and LSTM models to compare the best method. Some of the metrics used to measure the accuracy of forecasting methods are Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE).

This paper follows the following systematics: Section II discusses related research. Section III introduces the methods used in this study. Section IV is the results and discussion of our test results. Finally, Section V summarizes the key findings of this study.

II. Literature Review

Several studies have implemented IoT devices to detect parameters that affect air quality. Hanna Febryna Simorangkir [5] developed a system for monitoring air quality using IoT and displaying the sensor readings on a website. Nurul Azma Zakaria et.al [6] also conducted research to develop an IoT-based wireless device for monitoring indoor air quality using the Raspberry Pi 2 Model B and PHP and HTML-based software for the web. Jaka Prayudha et.al [11] measured air quality using the Internet of Things (IoT) and the Mamdani fuzzy logic method, and then displayed the results on the website. The difference lies in the hardware used, such as microcontrollers or sensors. Data acquisition and display methods are used as references in this research.

Several studies have used fuzzy logic to determine the air quality index based on certain parameter values. T. Mandal et.al [1] conducted research to compare air quality status values using fuzzy logic techniques with conventional techniques using only 4 parameters. Allahbakhsh Javid et.al [3] built an indoor air quality index system using fuzzy inference system on 15 parameters divided into 3 categories and the results are not much different from US-EPA. Dionova et.al [7] built an indoor air quality index system with fuzzy on 2 parts, namely IAQI for CO_2, CO, NO_2 and O_3; while TCI is an accumulation of PM, VOC, temperature and humidity data. The difference lies in the number and type of parameters used. However, the fuzzy logic method is used as a reference in this research.

Several studies have implemented quality index prediction for the future. Yan R. et.al. [8] conducted research to make forecasts of Air Quality Index (AQI) in Beijing City using 4 methods namely CNN, LTSM, CNN-LTSM and BPNN to be compared. Bedekar Gayatri et.al [10] tested for an efficient way to use the ARIMA method to forecast 4 indoor air quality index pollutants. Tingyi Liu et.al [12] analysed and predicted air quality using ARIMA and neural network methods. The difference lies in the forecasting method used for comparison. Comparison of ARIMA and LSTM in predicting air quality index is a research opportunity.

III. RESEARCH METHOD

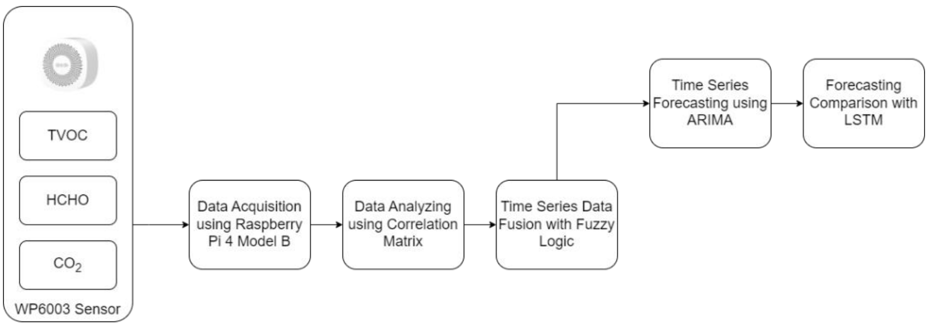

Fig. 1. System Work Diagram

The figure above shows the methodology of this research. The first step is to design the data acquisition process from the WP6003 sensor to obtain the values of air quality index parameters including air temperature, TVOC, HCHO, CO_2. The next step is to analyse the data using correlation matrix to find out the parameters that are related to each other. After that, fuzzy logic classification is used to obtain the air quality index value. Then the air quality index value is predicted using ARIMA method. Finally, the accuracy of ARIMA prediction with LSTM is compared using the test metric.

A. System On-board Design and Data Collection

Fig. 2. System Prototype Display

The figure above shows the design of our proposed on-board air quality index system. The system consists of a Raspberry Pi 4 connected to a VSON WP6003 device via Bluetooth. The sensor is powered by a microcontroller via a USB cable. Inside the WP6003 device there are already sensors to collect data on 4 parameters: temperature, TVOC, HCHO and CO_2, which affect the indoor air quality index. The Raspberry Pi retrieves the data from the sensor device, displays it in the browser, and also records and stores it in internal memory. The data is then retrieved for processing and forecasting.

B. Data Preparation and Fuzzy Logic

There are two stages to this step. The first stage is to analyse the related parameters. The analysis process uses a correlation matrix. The results of the correlation matrix are three possibilities, positively correlated, negatively correlated and have no correlation. The results of the correlation matrix are useful for determining models and rules for fuzzy logic. Parameters with good correlation, correlation values close to 1, are used in fuzzy logic rules, otherwise the parameters are ignored. The ranges used for the air quality indices are as in [13].

C. ARIMA Forecasting

Forecasting with ARIMA involves many steps. The first is data selection to determine the training data that will later be used to build the ARIMA model. The selected data is subjected to first-order differencing to make the data stationary and stable. To ensure that the data is stationary, the unit root test is performed and if the p-value is less than 5%, then the data is stationary. Plotting can also be done for the autocorrelation function (ACF) and the partial autocorrelation function (PACF). The plot helps to determine the order p and q in ARIMA (p,d,q), where the value of d is 1 because the difference has been made once. To obtain the optimal model, repeated tests of different model hypotheses are required. A good model is the model with the lowest value of Akaike's Information Criterion (AIC).

D. Testing Environment

The LSTM prediction is used to compare the accuracy of the ARIMA prediction. LSTM is used as a comparison because the LSTM model is very powerful, especially for long short-term memory retention. LSTM is an effective learning algorithm that can look at past data series and accurately predict future data series [14]. LSTM prediction must use exactly the same training and test data as the data used for ARIMA prediction. The number of iterations for the LSTM model is 50 because the data set used in this study is large. The accuracy results of both methods are compared using metrics, the RMSE equation (1) and the MAPE equation (2) [15].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=0}^{n}(A_t - F_t)^2} \tag{1}$$

$$MAPE = \frac{\sum_{t=1}^{n}\left|\left(\frac{A_t - F_t}{A_t}\right)100\right|}{n} \tag{2}$$

IV. RESULTS AND DISCUSSION

A. *WP6003 Sensor Device Test Results*

This test is carried out to determine the ability of the sensor to measure the level of each parameter affecting air quality over a given period of time, where the time specified for the sensor to read is 31 days. The test of the sensor devices is carried out by taking data from the WP6003 sensor using microcontroller, which then displays the results on the web-based application, as well as recording each sensor reading to a file.

TABLE I
AIR QUALITY INDEX DATA IN 10 MINUTES

Time	**Temperature**	**TVOC**	**HCHO**	**CO_2**
11-14-2022 15_6_3	25.8	0.22	0.034	618
11-14-2022 15_6_33	25.8	0.215	0.033	617
11-14-2022 15_7_3	25.8	0.211	0.032	616
11-14-2022 15_7_33	25.8	0.211	0.032	616
11-14-2022 15_8_3	25.5	0.204	0.031	615
11-14-2022 15_8_33	25.8	0.197	0.03	613
11-14-2022 15_9_3	25.8	0.204	0.031	615
11-14-2022 15_9_33	25.8	0.294	0.046	633
11-14-2022 15_10_3	25.5	0.338	0.053	642
11-14-2022 15_10_33	25.5	0.308	0.048	636
11-14-2022 15_11_3	25.5	0.271	0.042	629
11-14-2022 15_11_33	25.5	0.245	0.038	624
11-14-2022 15_12_3	25.5	0.266	0.041	628
11-14-2022 15_12_33	25.5	0.304	0.047	635
11-14-2022 15_13_3	25.5	0.304	0.047	635
11-14-2022 15_13_33	25.5	0.275	0.043	630
11-14-2022 15_14_3	25.5	0.275	0.043	630
11-14-2022 15_14_33	25.5	0.268	0.041	628
11-14-2022 15_15_3	25.5	0.264	0.041	627
11-14-2022 15_15_33	25.5	0.261	0.04	627

The results shown in the table above are the sensor readings for the first 10 minutes. As can be seen from the table, the sensor works by taking readings every 30 seconds, so 2 data points are obtained for each minute. This gives a total of 89,715 lines of data. The sensor can take readings for each parameter at a given time. The Time column contains the date, hour, minute and second of data collection. The Temperature column is the temperature value in Celsius at the time of data collection. The TVOC column is the value of the Total Volatile Organic Compound parameter in mg/m^3. The HCHO column is the value of the Formaldehyde parameter in mg/m^3. The CO_2 column is the value of the Carbon Dioxide parameter in ppm.

B. Fuzzy Inference System Test Results

From the sensor reading data there are 4 parameters including air temperature, TVOC, HCHO and CO_2. Before using the Fuzzy Inference System for classification and determination of AQI, it is necessary to determine in advance which parameters are used and how they are related to each other. The aim is to facilitate the creation of membership functions. Therefore, an analysis was carried out using the correlation matrix to determine the correlation between the parameters.

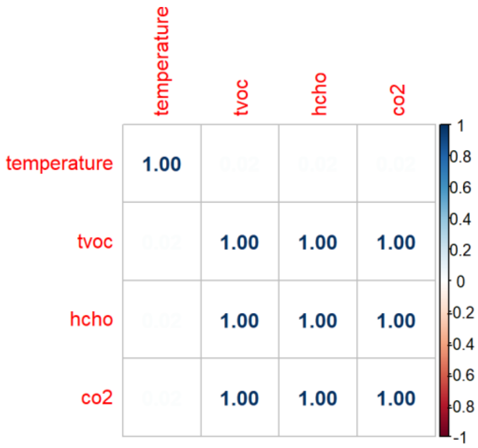

Fig. 3. Correlation matrix diagram

```
            temperature       tvoc       hcho        co2
temperature  1.00000000 0.01656598 0.01677397 0.01655535
tvoc         0.01656598 1.00000000 0.99991887 0.99994809
hcho         0.01677397 0.99991887 1.00000000 0.99986945
co2          0.01655535 0.99994809 0.99986945 1.00000000
```

Fig. 4. Correlation matrix diagram without rounding

From the data of the analysis results in the figures above, it can be concluded that there are 3 parameters that are positively correlated with each value close to 1. The three parameters are TVOC, HCHO and CO_2. As for the air temperature itself, it has almost no correlation with the other three parameters, where the correlation value is close to 0. Then the air temperature parameters can be ignored in the classification, while the TVOC, HCHO and CO_2 parameters are used in the classification process using the Fuzzy Inference System (FIS) method.

TABLE II
Membership Function of Parameters

Input	Membership Functions	Range
TVOC	Excellent quality	0 - 0,3
	Good quality	0,35 - 0,5
	Mild pollution	0,5 - 1,0

	Mediocre pollution	1,0 - 3,0
	Serious pollution	3,0 - 4,99
HCHO	Excellent quality	0 - 0,15
	Good quality	0,15 - 0,2
	Mild pollution	0,2 -0,3
	Mediocre pollution	0,3 - 1,24
	Serious pollution	1,24 - 1,99
CO_2	Excellent quality	400 - 600
	Good quality	600 - 800
	Mild pollution	800 - 1000
	Mediocre pollution	1000 - 1500
	Serious pollution	1500 - 2000
Output	**Membership Functions**	**Range**
AQI	Excellent quality	0 - 50
	Good quality	51 - 100
	Mild pollution	101 - 150
	Mediocre pollution	151 - 200
	Serious pollution	200 - 300

The table above describes the membership function along with the range of each class for each air quality index parameter. 5 categories are used for each parameter together with the results. The range is determined based on the range of values that can be obtained from the WP6003 sensor.

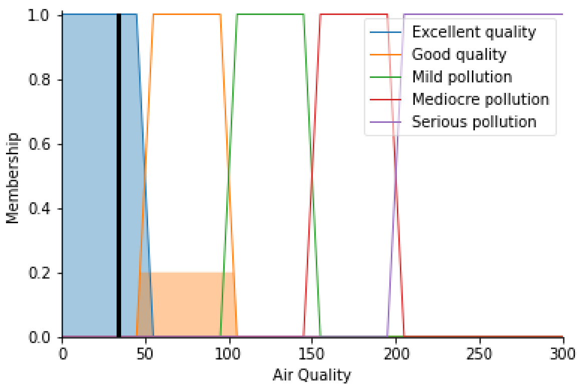

Fig. 5. Fuzzification Results Diagram

TABLE III
FUZZY PROCESS RESULTS EXAMPLE

Parameters	Value
Data TVOC	0.220
Data HCHO	0.034
Data CO2	618
Output (AQI)	33.33

The picture above is a diagram of the results of the Fuzzy Inference System process. And after the defuzzification process, the air quality index is 33.33 as shown in the table. The horizontal thick black line in the figure shows the location of the results after the defuzzification process. The line is in the "Excellent Quality" range, so based on the data in the table above, the air quality index can be classified as excellent. While the table above shows the data of each parameter that determines the fuzzy result in the previous figure.

C. Forecasting Test Results with ARIMA

In this test scenario, a forecast is made to predict the Air Quality Index for the next 30 minutes. As the system built is designed to be used in any room conditions, rooms with stable CO_2 concentrations as well as rooms with unstable CO_2 concentrations. Therefore, the 30-minute interval is set as a midpoint so that the model can be effective for many different conditions [16]. For this scenario, the predicted data is 5% of the data used to build the model (training set). Thus, the first 10 hours of data, 1200 data points in total, were used as the training set and the prediction results for the next 30 minutes, 60 data points in total, were used as the training set.

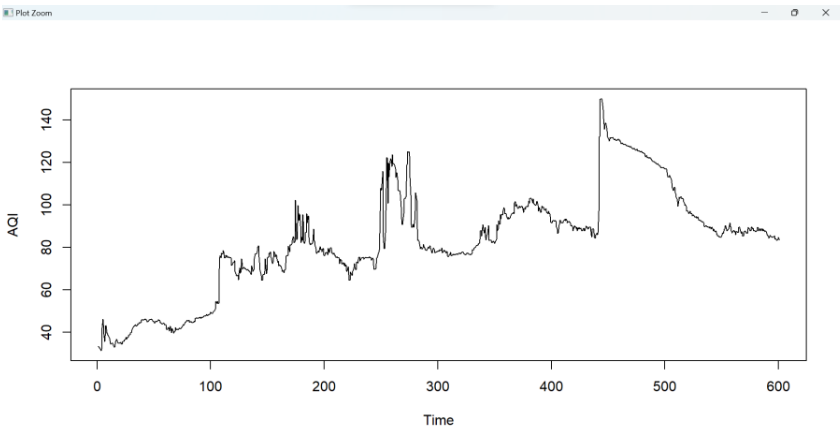

Fig. 6. 10-hour Air Quality Index Data Diagram

A graph of the results of the air quality index data for the first 10 hours is shown in the figure above. From the time series plot it can be seen that the Air Quality Index is very volatile. The graph shows that there are 2 abnormal peaks in the data, in the 400 to 600 minute range (200 - 300 in the figure) and in the 800 to 1000 minute range (400 - 500 in the figure).

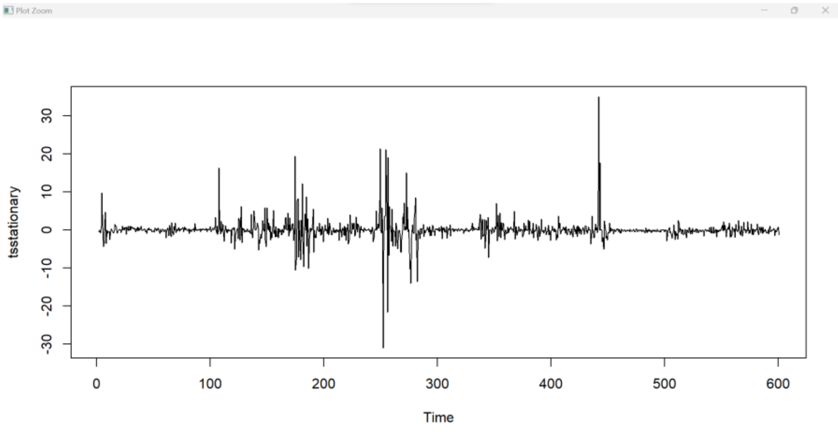

Fig. 7. Air Quality Index Data Diagram after Differentiation

The air quality index data is differentiated 1 time (first order difference) and the differentiation result diagram is shown in the figure above. From the figure after differentiation, it can be seen that the graph fluctuates at an almost constant value and shows a fairly stable condition, although there are some parts with less stable conditions. In addition, it can also be seen that there are 2 very abnormal conditions as mentioned earlier.

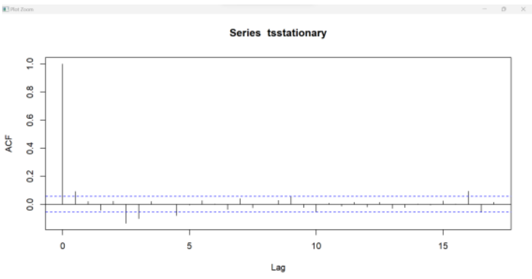

Fig. 8. Autocorrelation Function Diagram

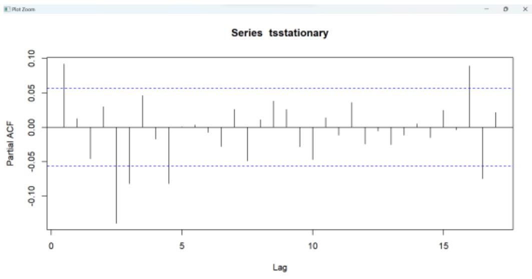

Fig. 9. Diagram Partial Autocorrelation Function

TABLE IV
Unit Root Test on Air Quality Index Data

ADF statistics	p-Value
-22.9816	< 0.01

Furthermore, the autocorrelation and partial autocorrelation of the air quality index data differentiated in the previous step are plotted. The results of plotting autocorrelation and partial autocorrelation can be seen in the figures above. And to ensure that the data is stationary and stable, the unit root test method is carried out. The result shown in the table above is that the p-value is less than 0.01 and the Null Hypothesis is rejected, so the data is stable and can be analysed further.

TABLE V
Fittings Model Results

ARIMA models	σ^2 Estimated	Log-Likelihood	AIC
ARIMA (3,1,1)	7.803	-2933.01	5876.01
ARIMA (5,1,1)	7.64	-2920.37	5854.75
ARIMA (6,1,1)	7.609	-2918.01	5852.03

As the coefficients of autocorrelation and partial autocorrelation after differentiation are still unclear, it will be difficult to determine the order of the components in ARIMA. Therefore, experiments with different combinations of the ARIMA order were carried out manually. Each trial result is then compared and analysed to determine the most appropriate order for use in forecasting. Based on Akaike's Information Criteration model, the lower the AIC value of the model, the better it is used. After repeated attempts, the ARIMA model (6,1,1) was found to be the best model. The AIC results of the experiment are shown in the AIC column of the table above.

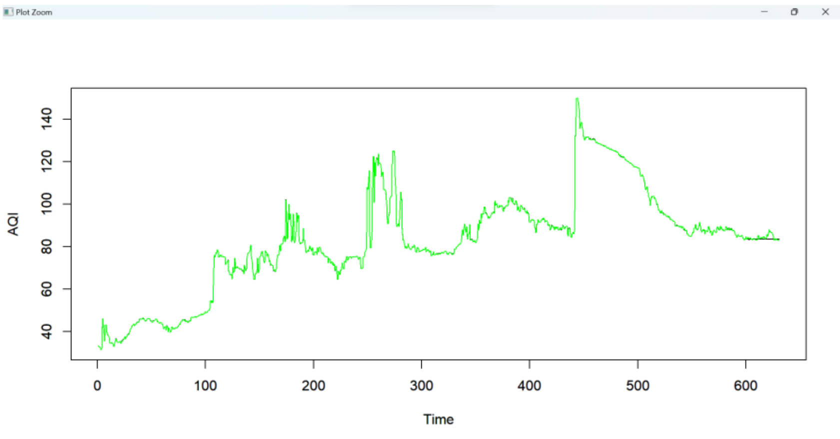

Fig. 10. ARIMA Forecasting Results Diagram

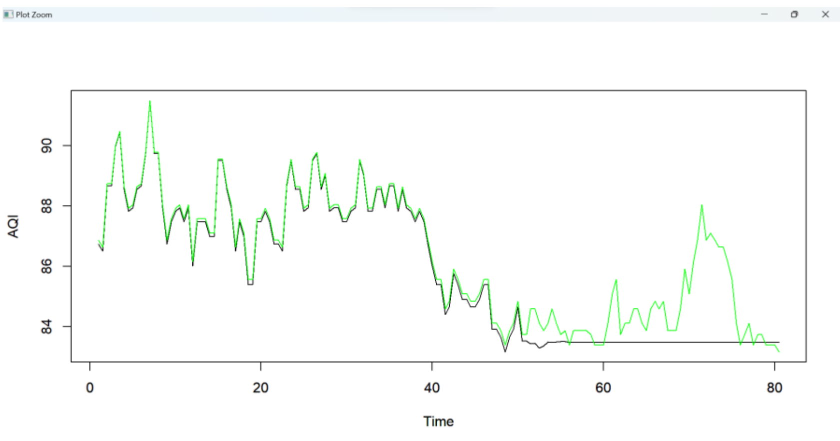

Fig. 11. ARIMA Forecasting Results Diagram with Magnification

The ARIMA model (6,1,1) is used to forecast 60 dates into the future. The prediction results are shown in the figures above. Figure 10 is an overall view of the actual and predicted data, while Figure 11 is an enlarged view of Figure 10. The green line graph is the actual data, while the black line graph is the predicted data. As can be seen, the actual data is quite volatile, but the difference in values between the actual and predicted data is not too great. To find out the effectiveness of the forecast compared to the actual values, calculations are made using the MAPE and RMSE methods. The results obtained show Root Mean Square Error and Mean Absolute Percentage Error of 2.7574 and 1.4841% respectively.

D. Forecasting Test Results with LSTM

In the LSTM test scenario, the distribution of the data set is still the same as in the ARIMA test scenario, so that the analysis of the test results becomes objective. For the air quality index data, the first 10 hours will be used as training data, where the data amount to 1200 data points. While 30 minutes after that will be a test data where the data amounts to 60 points.

First, the air quality index data were entered into the LSTM program. Then 90% of the data is defined as training data and the remaining 10% as test data (test set). A process of differentiation is also applied to the data up to 1 times (first order difference) to make the data stationary and increase the accuracy of the training data. The number of neurons in the model is set to 4. The number of iterations is set to 50 because the data set used for prediction is quite large. The model is then used to make a prediction.

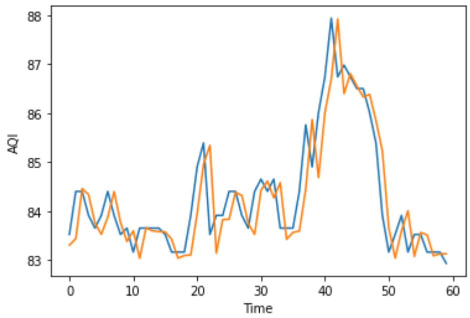

Fig. 12. LSTM Forecasting Results Diagram

A comparison chart of actual and predicted data is shown in the figure above. Where the blue line is the actual data and the orange line is the predicted result data. The data pattern of the LSTM predicted results almost follows the movement of the actual data and has a very good accuracy value. The prediction results obtained Root Mean Square Error and Mean Absolute Percentage Error values of 0.622 and 0.5505% respectively.

E. Test Analysis Results

In order to analyse the effectiveness of both methods in predicting the air quality index for the next 30 minutes, a comparison was made on the RMSE and MAPE metrics of each method. A comparison of the test metrics of each method is shown in table below.

TABLE VI
COMPARISON OF METRIC VALUES OF EACH METHOD

Type	RMSE	MAPE
ARIMA (6,1,1)	2,7574	1.4841 %
LSTM	0.622	0.5505 %

In the RMSE metric, the value of the LSTM method is much lower than that of the ARIMA method, so the LSTM method is better than the ARIMA method. The size and range of the data sets used to test the two scenarios are the same, so the relative error of the two test results can be expected to be the same. The difference in the values of the two methods is very significant because the ARIMA prediction tends to flatten out towards the end, whereas the LSTM prediction almost follows the pattern of the actual data. The LSTM method also outperforms the ARIMA method for the MAPE metric, although the LSTM value is lower than the ARIMA value. Nevertheless, the MAPE value of the ARIMA method is still quite low. MAPE values below 10% can be classified as accurate forecasts [17]. In terms of predicting the air quality index for the next 30 minutes based on 10 hours of air quality index data, the LSTM method is better than the ARIMA method.

V. Conclusion

Based on the results of the testing and analysis, it can be concluded that the Indoor Air Quality Index System has successfully read each parameter that affects the Air Quality Index for 31 days with readings taken every 30 seconds and has successfully displayed the readings on a web-based application and then saved them to a file. The parameters used in the Fuzzy Inference System (FIS) for classification are those that have a correlation with each other, and these parameters include TVOC, HCHO and CO_2. Effective forecasting methods for predicting the ordered air quality index are ARIMA with 5% sample, ARIMA with 10% sample and LSTM. For further research, a SIM module can be added so that the system can run anywhere. A cloud computing system has been added to the system so that the forecast results can be monitored by users at any time, so that they can make quick decisions if the air quality index is going to deteriorate in the future. It also uses a screen that is directly integrated with the Raspberry Pi, for example through the GPIO port, to make it easier for users to see the system's readings.

Acknowledgment

The authors would like to thank the supervisor Mr. Bayu Erfianto and Cyber Physical System Research Group of Telkom University who have supported the author to completed this research.

References

[1] Javid, Allahbakhsh et al. "Towards the Application of Fuzzy Logic for Developing a Novel Indoor Air Quality Index (FIAQI)." *Iranian journal of public health* vol. 45,2 (2016): 203-13.
[2] Ingo E. Isphording, Nico Pestel, Pandemic meets pollution: Poor air quality increases deaths by COVID-19. *Journal of Environmental Economics and Management*, Volume 108, 2021. Available: https://doi.org/10.1016/j.jeem.2021.102448.
[3] Mandal, T., Gorai, A.K. & Pathak, G. Development of fuzzy air quality index using soft computing approach. *Environ Monit Assess* 184, 6187–6196 (2012).
[4] K. Pratama and E. Setiawan, "Implementasi Monitoring Kualitas Udara Menggunakan Peramalan Exponential Smoothing dan NodeMCU Berbasis Mobile Android", *Ultima Computing : Jurnal Sistem Komputer*, vol. 9, no. 2, pp. 58-66, Apr. 2018.
[5] Hanna Febryna Simorangkir, "Rancang Bangun Pemantauan Kualitas Udara Pada Taman Wilayah Melalui Website Berbasis Arduino Menggunakan Logika Fuzzy", Vol. 1 No. 1 (2107): JATI Vol. 1 No. 1, 2017.
[6] Zakaria, Nurul Azma et al. "Wireless Internet of Things-Based Air Quality Device for Smart Pollution Monitoring." *International Journal of Advanced Computer Science and Applications* 9 (2018).

[7] Brainvendra Widi Dionova, M.N. Mohammed, S. Al-Zubaidi, Eddy Yusuf, Environment indoor air quality assessment using fuzzy inference system. *ICT Express, Volume 6, Issue 3*, 2020. Available:
https://www.sciencedirect.com/science/article/pii/S2405959520301065
[8] Rui Yan, Jiaqiang Liao, Jie Yang, Wei Sun, Mingyue Nong, Feipeng Li, Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering. *Expert Systems with Applications, Volume 169*, 2021. Available: https://www.sciencedirect.com/science/article/pii/S095741742031157X
[9] Minqiu Zhou, Amir M. Abdulghani, Muhammad A. Imran, and Qammer H. Abbasi. 2020. Internet of Things (IoT) Enabled Smart Indoor Air Quality Monitoring System. In *Proceedings of the 2020 International Conference on Computing, Networks and Internet of Things (CNIOT2020). Association for Computing Machinery*, New York, NY, USA, 89–93.
[10] Bedekar, Gayatri and Patil, R.S. and Tergundi, Parimal and Goudar, R. H., An Efficient Implementation of ARIMA Technique for Air Quality Prediction, 2021. Available: https://ssrn.com/abstract=3889537
[11] Jaka Prayudha, Ardianto Pranata, and Afdal Al Hafiz, “Implementasi Metode Fuzzy Logic Untuk Sistem Pengukuran Kualitas Udara Di Kota Medan Berbasis Internet Of Things (IoT)”, Vol 4, No 2 (2018), 2018.
[12] T. Liu and S. You, “Analysis and Forecast of Beijing’s Air Quality Index Based on ARIMA Model and Neural Network Model,” *Atmosphere*, vol. 13, no. 4, p. 512, Mar. 2022.
[13] S. M. Saad, A. Y. M. Shakaff, A. R. M. Saad, A. M. Yusof, A. M. Andrew, A. Zakaria, and A. H. Adom, "Development of indoor environmental index: Air quality index and thermal comfort index", *AIP Conference Proceedings* 1808, 2017.
[14] R. Janarthanan, P. Partheeban, K. Somasundaram, P. Navin Elamparithi. A deep learning approach for prediction of air quality index in a metropolitan city. *Sustainable Cities and Society*, Volume 67, 2021. Available: https://doi.org/10.1016/j.scs.2021.102720
[15] J. Kang and K.-I. Hwang, “A Comprehensive Real-Time Indoor Air-Quality Level Indicator,” *Sustainability*, vol. 8, no. 9, p. 881, Sep. 2016, doi: 10.3390/su8090881.
[16] Yu T-C, Lin C-C. An Intelligent Wireless Sensing and Control System to Improve Indoor Air Quality: Monitoring, Prediction, and Preaction. International Journal of Distributed Sensor Networks. 2015;11(8). doi:10.1155/2015/140978
[17] Lewis, C. D., 1982. *Industrial and Business Forecasting Methods*. London: Butterworths.