COVID-19 time-varying reproduction number and its prediction by adaptive neuro-fuzzy inference system (ANFIS)

Achmad Riadi[1,a)], Yulianto Sulistyo Nugroho[2,b)], Fariz Nurwidya[3,c)]

[1]*Department of Mechanical Engineering, Faculty of Engineering, Universitas Indonesia, Kampus UI Depok, West Java, 16424, Indonesia*

[2]*Department of Mechanical Engineering, Faculty of Engineering, Universitas Indonesia, Kampus UI Depok, West Java, 16424, Indonesia*

[3]*Department of Pulmonology and Respiratory Medicine, Faculty of Medicine, Universitas Indonesia - Persahabatan Hospital, Jakarta 13230, Indonesia*

[a)]Corresponding author: achmadriadi@ui.ac.id
[b)]yul.nugroho@ui.ac.id
[c)]fariz.nurwidya @ui.ac.id

Abstract. Epidemiologists and the COVID-19 Response Acceleration Task Force in Jakarta, Indonesia continue to monitor various parameters of the outbreak in Jakarta and its surroundings. One of the parameters to monitor and predict is the effective reproduction number (R_t). We propose a machine learning approach called adaptive neuro-fuzzy inference system (ANFIS) to predict the time-varying R_t. We calculated historical data for R_t using the daily confirmed cases, daily death cases, and daily recovered cases by applying a Susceptible-Infected-Recovered-Dead (SIRD) dynamic simulation. We found that using the input data from the last seven previous days of R_t as the training data gave best prediction results based on the evaluation metric of mean absolute percentage error (MAPE). We validated the prediction results with 10 consecutive days of R_t from November 1–10, 2020. The results showed that the developed ANFIS model could predict the reproduction number for COVID-19 in Jakarta, Indonesia with an accuracy of approximately 81.27%. The transmission of COVID-19 in Jakarta, Indonesia has the potential to be controlled because most R_t values were lower than 1, at least during the period analyzed in this study.

Keywords: COVID-19, reproduction number, dynamic simulation, machine learning, neuro-fuzzy, SIRD model

1. Introduction

COVID-19 or 2019-nCoV is a disease caused by a novel coronavirus that first appeared in Wuhan, Hubei Province, China on December 31, 2019, and quickly spread to several cities in several countries. On January 30, 2020, the World Health Organization (WHO) declared this outbreak a public health emergency of international concern (WHO Emergency Committee, 2020). The outbreak has prompted all countries to implement physical distancing and quarantine policies to prevent the spread of the virus.

The capital city of Jakarta is the city with the largest number of coronavirus cases in Indonesia. Since the first confirmed case of COVID-19 in Indonesia on March 2, 2020, the Jakarta city government has implemented many policies to reduce the spread of the virus. In addition to socialization in implementing health protocols, the Jakarta city government has also implemented a large-scale social restriction (PSBB) policy. This policy was extended several times, and it was relaxed several times (transitional PSBB) when there was a decreasing trend in the number of cases. However, on September 14, 2020, an emergency brake policy (full PSBB) was implemented for approximately a month because the number of cases was increasing uncontrollably. After the number of new cases stabilized, the Jakarta city government started to lift the restrictions slowly (Aqil, 2020).

The impact of the policies implemented by the Jakarta city government in suppressing the spread of the virus should be measured using an indicator. One indicator that can be used is the coronavirus reproduction number, R_t.

This number is one important epidemic characteristic needed by public health bodies or governments to formulate possible treatment measures (Chintalapudi et al., 2020). Reproduction numbers are commonly used in assessing the successful implementation of policies for handling outbreaks. The reproduction number is also an indication of the ability of a virus to spread, as it represents the average number of new infections transmitted by infected people (Liu et al., 2020). The R_t value explains how contagious the disease is, and an accurate estimate of this value is extremely important for policymakers and epidemiologists in the COVID-19 pandemic. If the R_t values are less than 1, then the spread of the virus is slow and the epidemic has the potential to be controlled (Inglesby, 2020). R_t values higher than 1 indicate that the number of cases will increase (Caicedo-Ochoa et al., 2020).

There are several studies related to the reproduction number and its prediction, especially for the COVID-19 pandemic. One such study aimed to estimate the reproduction number of the COVID-19 virus during the first 10 days of the outbreak in seven Latin American countries (Caicedo-Ochoa et al., 2020). The study estimated R_t using the ratio of the number of new infections generated in step t, I_t, to the total number of infected individuals at time t. Another study considered the Marche region of Italy, including five provinces, to compute R_t. The study estimated virus reproduction numbers by fitting exponential infection growth rates on both a daily and a weekly basis (Chintalapudi et al., 2020). The reproduction number can also be estimated by first describing the epidemic curve (Sahafizadeh & Sartoli, 2020). One method for predicting the dynamic spread of an epidemic is the computer simulation of an epidemic mathematical model (Abdulrahman, 2020).

A machine learning approach called adaptive neuro-fuzzy inference system (ANFIS) (Jang, 1993) is widely applied in forecasting problems, including time-series predictions, and shows good, reliable results. ANFIS combines several properties of artificial neural networks (ANN) and fuzzy logic systems that provide flexibility in handling nonlinearity in time series data. There are many ANFIS applications in forecasting, and they have begun to be applied to studies related to COVID-19. One study presented the forecasting of confirmed cases of COVID-19 in China using an improved model of ANFIS called FPASSA-ANFIS which utilizes a flower pollination algorithm (FPA) and salp swarm algorithm (SSA) (Al-qaness et al., 2020). The authors evaluated the FPASSA-ANFIS model by forecasting the number of confirmed cases of COVID-19 ten days ahead, and it showed good performance. Another study proposed a hybrid machine learning method using ANFIS and a multi-layered perceptron-imperialist competitive algorithm (MLP-ICA) to predict a time series for infected individuals and the mortality rate of the COVID-19 outbreak (Pinter et al., 2020). Predicting the reproduction number of COVID-19 using ANFIS has not been the focus of many previous studies, so it needs to be further analyzed and explored.

In some studies, analytic approaches were proposed for pandemic models, including the Susceptible-Infected-Removed (SIR) model (Cooper et al., 2020; Nesteruk, 2017; Zhong et al., 2020), Susceptible-Exposed-Infected-Removed (SEIR) model (Neher et al., 2020; Roda et al., 2020), Susceptible-Infected-Recovered-Dead (SIRD) model (Anastassopoulou et al., 2020; Caccavo, 2020), and Susceptible-Exposed-Infected-Removed-Dead SEIRD (Ala'raj et al., 2021; Peng et al., 2020). The models can be used as basic models for performing a time-domain dynamic simulation of a coronavirus outbreak. In the current study, an application of an SIRD dynamic simulation with the ANFIS is proposed for simulating a coronavirus outbreak in Jakarta city and estimating the reproduction number. The main contribution of the current study is to provide an accurate prediction of the coronavirus reproduction number, which is needed by public health bodies or governments in formulating possible policies and treatments.

The rest of this study is organized as follows. The principles of SIR, SIRD, the ANFIS, the reproduction number, and the main dataset are described in Section 2. Section 3 presents the experimental setup. Section 4 presents the result and discussion. Finally, in Section 5, the conclusion of this study is presented.

2. Material and Methods

2.1. SIR and SIRD models

The SIR model is a basic model for describing disease spread. It is also a base model for several other models, such as the SIR, SEIR, SIRD, and SEIRD. The SIR model consists of three population levels, namely: Susceptible, Infected, and Removed. An individual may belong to any one of these population levels. Susceptible individuals are not immune to the disease but are not infected. If a vaccine for this disease has not been developed and applied, it can be said that the entire community has a chance of being infected by this disease. Therefore, the "Susceptible" population level can be represented by the entire population in a community. An individual in the "Susceptible" population can move into the "Infected" population through contact with an infected person. With a single transmission or movement of an individual from the "Susceptible" population to the "Infected" population, the number

of individuals in the "Susceptible" population decreases by one and the number of individuals in the "Infected" population will increase by one. The "Infected" population consists of individuals who have the disease and can spread it to susceptible people. Individuals in the "Infected" population may move into the "Removed" population by recovering from the disease. The "Removed" population includes individuals who are no longer infected and individuals who have died from the disease. Figure 1 depicts a basic SIR model, where β denotes the infection rate or transmission rate, and γ denotes the recovery or removed rate.
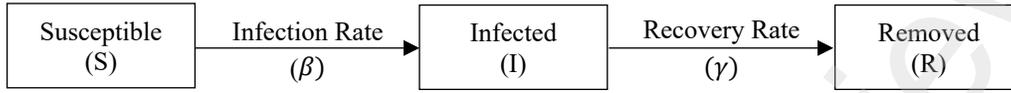

Fig. 1 Basic SIR model.

The SIR model is a dynamic system of three paired differential equations that describe the time evolution of the three populations (Ardabili et al., 2020; Cooper et al., 2020; Pinter et al., 2020). The three paired differential equations are as follows:

$$\frac{dS(t)}{dt} = -\beta S(t) I(t) \tag{1}$$

$$\frac{dI(t)}{dt} = \beta S(t) I(t) - \gamma I(t) \tag{2}$$

$$\frac{dR(t)}{dt} = \gamma I(t) \tag{3}$$

where N indicates the overall population size, $N = S + I + R$.

Another model to represent disease spread is the SIRD model. In this model, an extra population level is introduced at the end of the SIR model to differentiate between recovered and death cases in the "Removed" population. Figure 2 depicts the SIRD model, where δ denotes the death rate.
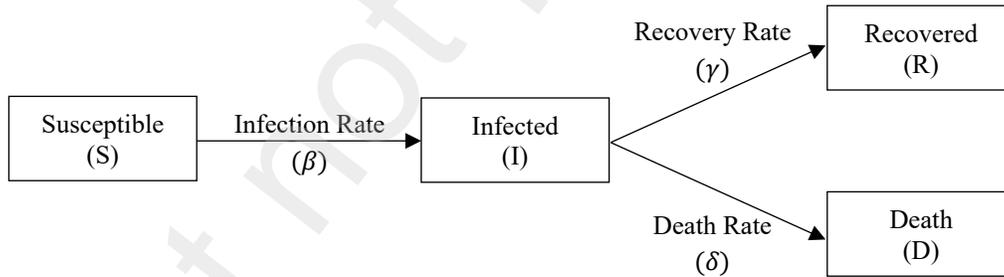

Fig. 2 SIRD model.

In the SIRD model, the overall population size, N is denoted as $N = S + I + R + D$. The differential equation for S and R remain the same as in the SIR model, while for I and D, the following differential equations are used:

$$\frac{dI(t)}{dt} = \beta S(t) I(t) - (\gamma + \delta) I(t) \tag{4}$$

$$\frac{dD(t)}{dt} = \delta I(t) \tag{5}$$

2.2. Reproduction number

There are two types of reproduction number, i.e., basic and effective types of reproduction number (Susanto et al., 2020). The basic reproduction number (R_0) is the expected number of cases directly generated by one case in a population before the pandemic spreads, while the effective reproduction number (R_t) is the number of cases generated in the current state of a population or at a specific time. In the SIRD model, R_0 is calculated using the following equation:

$$R_0 = S_0\left(\frac{\beta}{\gamma + \delta}\right) \tag{6}$$

and R_t is calculated using the following equation:

$$R_t = \left(\frac{S_t}{N}\right)R_0 \tag{7}$$

R_0 is a mathematically determined quantity. In many cases, it is important to evaluate time-dependent variations in the transmission potential of infectious diseases (Nishiura & Chowell, 2009). Therefore, one description of the time course of an epidemic can be obtained by estimating the effective reproduction number, R_t.

2.3. ANFIS

The ANFIS concepts are set out in this section. The ANFIS was developed in the early 1990s and gained widespread popularity for scientific modeling and prediction within the science community. It is an algorithm that combines two machine learning techniques: fuzzy logic and neural networks (Jang, 1993). The reason for the combination is that each technique has different advantages and disadvantages. The main advantage of ANNs is that they can improve adaptive parameters through a learning process. The disadvantage of this technique is the complexity of its structure, whereas fuzzy logic is similar in concept to human thinking. Each technique complements the other in terms of advantages and disadvantages. The structure of ANFIS is enhanced through the integration of the Takagi–Sugeno inference model, which generates a mapping between the input and output by applying IF-THEN rules (Azar, 2014; Khodaei-mehr et al., 2018). Figure 3 illustrates an example of the ANFIS structure, which has two x and y input variables. Each input variable has two membership functions (MFs), namely A_1, and A_2, and B_1 and B_2.
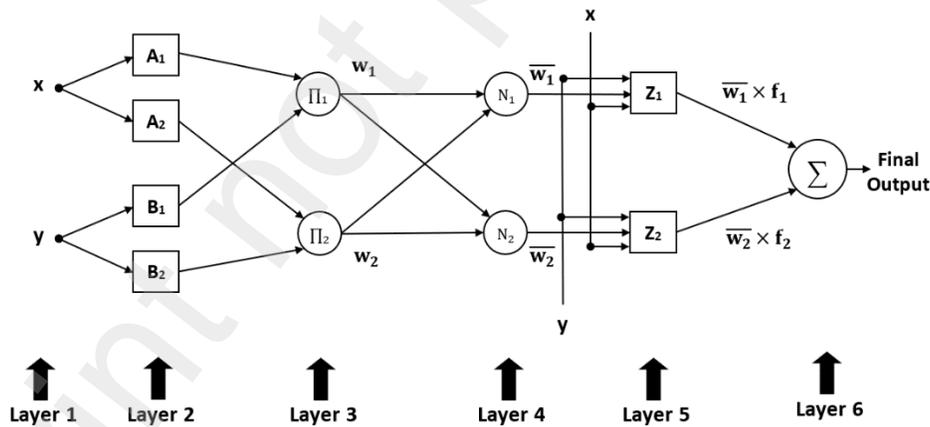

Fig. 3 ANFIS structure.

Some layers of ANFIS have the same number of nodes and the nodes in the same layer have similar functions. For convenience, the output of the ith node in layer l is denoted as $O_{l,i}$. The function of each layer is described as follows:

- Layer 1. This is the input layer. Nodes in this layer simply pass incoming signals (crisp values) to Layer 2 as $O_{1,x} = x$ and $O_{1,y} = y$.

- Layer 2. This is the fuzzification layer. The outputs of nodes in this layer are represented as:

$$O_{2,i} = \mu A_i(x), i = 1,2, \quad O_{2,i} = \mu B_{i-2}(y), i = 3,4 \tag{8}$$

μA_i (or μB_i) is a linguistic label (such as "*very low*," "*low*," "*high*," "*very high*") that is associated with this node. In other words, $O_{2,i}$ is the membership grade of a fuzzy set M ($=A_1, A_2, B_1, B_2$) and it specifies the degree to which the given input x (or y) satisfies the quantifier M. Here, the MF for M can be any appropriate parameterized membership function (triangular, trapezoidal, Gaussian, two-sided Gaussian, bell-shaped).

- Layer 3. This is the rule layer or firing strength of a rule. Each node output is the product of all incoming signals and it is calculated as follows:

$$O_{3,i} = w_i = \mu A_i(x) \times \mu B_{i-2}(y) \tag{9}$$

The output of this node is also denoted as w_i, which represents the firing strength, or the truth value.

- Layer 4. This is the normalization layer (also known as the normalized firing strength). It is calculated as follows:

$$O_{4,i} = \overline{w}_i = \frac{w_i}{\sum_{i=1}^{2} w_i} \tag{10}$$

- Layer 5. This is the defuzzification layer (also known as the adaptive node). It is calculated as follows:

$$O_{5,i} = \overline{w}_i f_i = \overline{w}_i \left(p_i x + q_i y + r_i \right) \tag{11}$$

where p_i, q_i, and r_i are the parameter sets for node i or the consequent parameters.

- Layer 6. This layer is represented by a single summation node. It contains only one node and computed as:

$$O_6 = \sum_i \overline{w}_i f_i \tag{12}$$

2.4. Main dataset

The main dataset for COVID-19 was collected from the Jakarta city government website (https://corona.jakarta.go.id/id/data-pemantauan). It contains the daily confirmed cases, daily death cases, and daily recovered cases from March 1, 2020 to November 10, 2020, as shown in Figure 4.
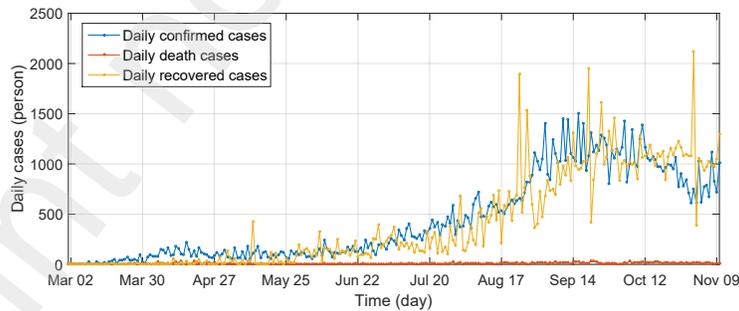

Fig. 4 COVID-19 daily data for Jakarta, Indonesia.

3. Experimental Setup

3.1. SIRD dynamic simulation

The differential equations for S, I, R, and D are presented as equations (1), (4), (3), and (5), respectively, and were used to model the time-varying behavior of the SIRD dynamic simulation to represent the spread of COVID-19. This dynamic simulation model was implemented using a spreadsheet. The data for daily confirmed cases, daily death

cases, and daily recovered cases, as shown in Figure 4, were used as input data for the dynamic simulation model. The time evolution (on a daily basis) of each population (S, I, R, and D) was obtained from the simulation. Furthermore, the infection rate (β), recovery rate (γ), and death rate (δ) could also be generated daily by using the following equations (13), (14), and (15), respectively.

$$\beta(t) = \frac{dI(t)/dt}{S(t)I(t)} \tag{13}$$

$$\gamma(t) = \frac{dR(t)/dt}{I(t)} \tag{14}$$

$$\delta(t) = \frac{dD(t)/dt}{I(t)} \tag{15}$$

The values of β, γ, and δ are depicted in Figure 5(a), 5(b), and 5(c), respectively.

(a)

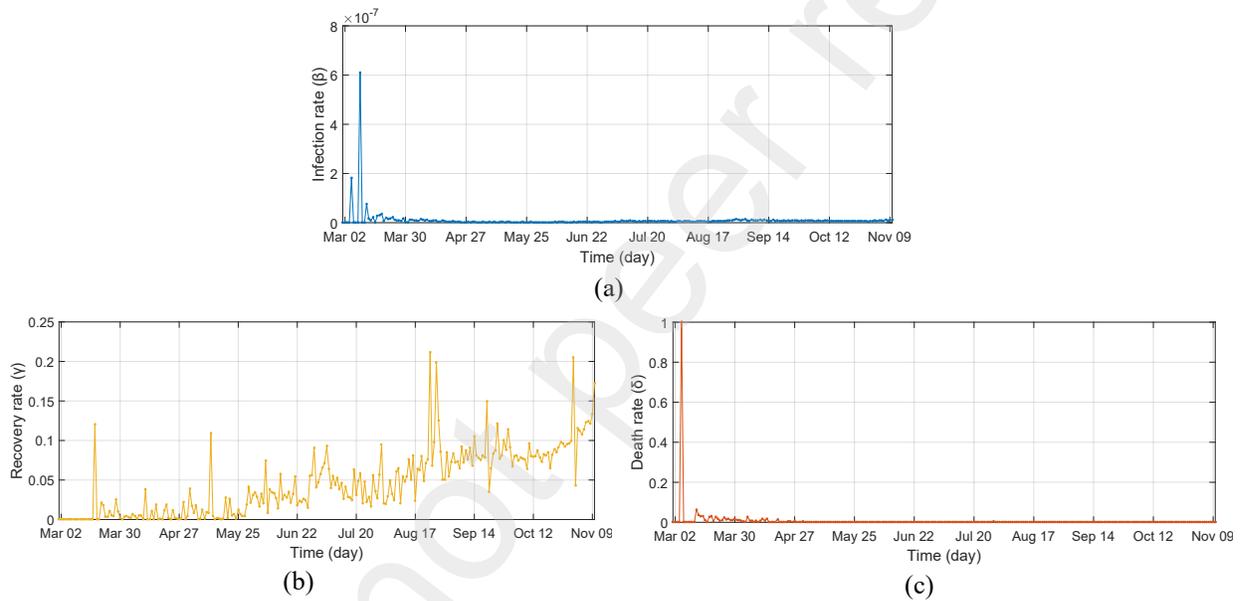(b)   (c)

Fig. 5 (a) Infection rate, (b) recovery rate, and (c) death rate.

The values of β, γ, and δ were then used for determining the values of the effective reproduction number (R_t) by applying Equation (7). There are some undefined R_t values for the observation period. This is due to division by a zero denominator in the equation for calculating R_t. The denominator is the sum of the recovery rate and death rate. If this denominator is zero, then the value of R_t is undefined. For such cases, the value of R_t is assumed to be zero. Figure 6 shows the daily movement of the effective reproduction number.
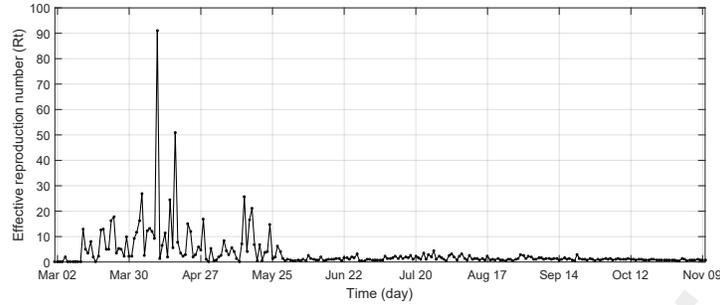
Fig. 6 Effective reproduction number (R_t).

3.2. ANFIS structure and input-output dataset preparation

The daily time series data of R_t generated from the SIRD dynamic simulation were then used as the basis for estimating the future values of the reproduction number by applying the ANFIS algorithm. ANFIS needs an input-output dataset as training data to improve its adaptive parameters through a learning process.

In the present study, the input-output dataset was prepared using three scenarios, as shown in Table 1. The first scenario categorizes the time series data for the reproduction number into seven inputs that represent the reproduction numbers for the last seven days, to be used for the prediction of the next day's reproduction number. The second scenario categorizes the time series data for the reproduction numbers into four inputs to represent the reproduction numbers for the last four consecutive odd days, to be used for the prediction of the next day's reproduction number. The third scenario categorizes the time series data for the reproduction number into four inputs the represent the reproduction numbers for the last three consecutive weeks, to be used for the prediction of the next day's reproduction number.

Table. 1 Scenario for ANFIS input-output dataset.

	Inputs	**Outputs**
Scenario 1	$R_{t-7}, R_{t-6}, R_{t-5}, R_{t-4}, R_{t-3}, R_{t-2}, R_{t-1}$	R_t
Scenario 2	$R_{t-7}, R_{t-5}, R_{t-3}, R_{t-1}$	R_t
Scenario 3	$R_{t-21}, R_{t-14}, R_{t-7}, R_{t-1}$	R_t

In the prepared input-output dataset, the output reproduction number was set within the period of April 1, 2020 to November 10, 2020. Several pieces of data collected before April 1, 2020 were used as input data, especially in the third scenario where three weeks of previous data were required. Meanwhile, the data for the first week of March 2020 were ignored, because this time was the initial period of the pandemic when the data collection and reporting system by the Jakarta government was still unstable. The data from April 1, 2020 to October 31, 2020 were used for training, and the data for November 1–10, 2020 were solely used for validation.

Before training the input-output dataset in the ANFIS, it was important to normalize the data to convert the data to a symmetrical distribution that enhances model efficiency. The following equation was used to normalize the data:

$$v' = \frac{v - min}{max - min} \tag{16}$$

Here, min and max are the minimum and the maximum values in the data, whereas v and v' are the original and the normalized values of each observation, respectively (Jain et al., 2018). Based on the above equation, the minimum and the maximum in the data are mapped to 0 and 1, respectively.

The ANFIS structure was implemented through the ANFIS toolbox in MATLAB. The input parameters were set as the independent variables in each of the above scenarios, and the output variable was the reproduction number. Some fuzzy MFs, such as triangular, trapezoidal, Gaussian, bell, and sigmoidal, were trained, and the best MF was selected. A linear output membership function was chosen due to its ability to further minimize errors. The ANFIS training was carried out using an optimal method of backpropagation with zero error tolerance.

The mean absolute percentage error (MAPE) evaluation metric was used to evaluate the accuracy of the ANFIS prediction. The following equation was used for the evaluation metric in this study.

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right| \times 100 \tag{17}$$

where n represents the number of data, A_t represents the actual value, and F_t is the forecast value.

4. Result and Discussion

4.1. Result

The training data for the COVID-19 reproduction number (from April 1, 2020 to October 31, 2020) were used to train the ANFIS algorithm to find the best set of parameters for minimal error. The ANFIS training step for each input-output dataset scenario was carried out several times to find the type and number of MFs that showed the lowest MAPE. Table 2 presents the training results of ANFIS for the COVID-19 reproduction number in Jakarta, Indonesia.

Table. 2 ANFIS training results for the COVID-19 reproduction number in Jakarta, Indonesia.

	No. of Inputs	MF type	No. of MFs	MAPE(%)
Scenario 1	7	Gaussian	2	0.31
Scenario 2	4	Bell	3	0.47
Scenario 3	4	Gaussian	2	1.05

According to Table 2, Scenario 1 provided the lowest MAPE of all the scenarios. Therefore, it can be concluded that Scenario 1 was suitable for predicting the reproduction number for COVID-19 in Jakarta, Indonesia. As for the determination coefficient, the ANFIS Scenario 1 provided a determination coefficient of 0.998, as shown in Figure 7.
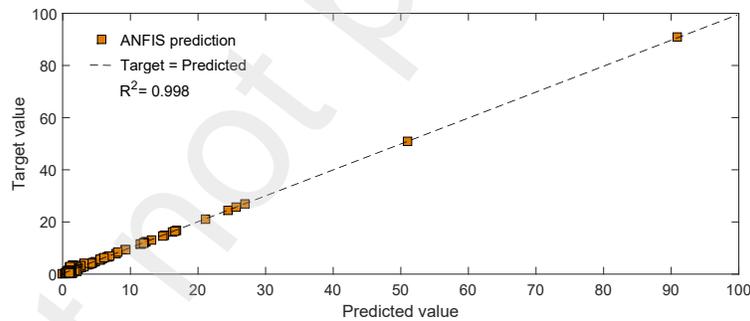

Fig. 7 Predicted vs. target value for the COVID-19 reproduction number (ANFIS Scenario 1).

After the result of the ANFIS training for each input-output dataset scenario was evaluated, the ANFIS model in the presence of Scenario 1 was selected for the prediction of the COVID-19 reproduction number in Jakarta, Indonesia. A prediction was made for the next ten days (November 1–10, 2020) and was validated using actual data. Figure 7 shows the prediction of the reproduction number resulting from the ANFIS model.
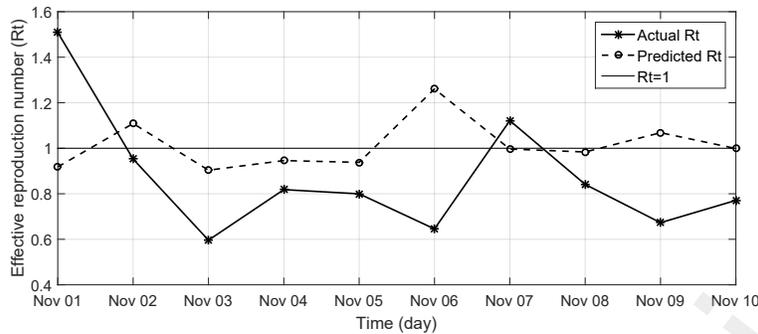
Fig. 8 Effective reproduction number (predicted vs. actual), November 1–10, 2020.

To validate the accuracy of the ANFIS prediction, the MAPE evaluation metric was calculated according to Equation (17) by incorporating the results from Figure 8. Based on the calculation, the ANFIS model prediction showed a MAPE value of 18.73%. This indicates that the developed ANFIS model could predict the reproduction number for COVID-19 in Jakarta, Indonesia with an accuracy of approximately 81.27%.

Based on the SIRD dynamic simulation, which produces reproduction number data (Figure 6), along with the prediction results using the ANFIS (Figure 8), the reproduction number for COVID-19 in Jakarta, Indonesia can potentially be controlled because most R_t values are lower than one.

4.2. Discussion

The use of the ANFIS algorithm in this study shows the potential for a machine learning approach in predicting the reproduction number for COVID-19 in Jakarta, Indonesia. More work is needed to produce more accurate predictions. Additionally, it is also necessary to ascertain whether this technique is adequate to use in the event of a significant change in the reproduction number data pattern, given the dynamic changes in the COVID-19 infection rate.

The ANFIS algorithm makes predictions using a learning process on the training data. The selection of input within the training data is one of the effective factors of the ANFIS prediction (Tarno et al., 2013). In this study, three scenarios with different input data were presented, and they resulted in different accuracies, as shown in Table 2. The selection of input data that is different from those presented in this study could produce different predictions and maybe even be more accurate.

The prediction of the reproduction number for COVID-19 benefits the government in deciding on policies for preventing the spread of the novel coronavirus. This prediction is strongly influenced by the movement of data from daily confirmed cases, daily death cases, and daily recovered cases. The existence of bias in daily data reporting would certainly affect the results of the prediction of the reproduction number for COVID-19 because predictions are made based on historical data. For Indonesia, and possibly for other emerging countries, greater data accuracy is needed to help curb the transmission of the virus (Editor Kompas, 2020). Moreover, there is still late data reporting, which causes the movement of daily data to vary from the actual data. Late data reporting can also cause data fluctuation. This phenomenon is shown in the reproduction number data for April 10 and 17, which reached 90 and 51, respectively, as shown in Figure 6. While we provide promising prediction results for the reproduction numbers for COVID-19 in Jakarta, Indonesia, we need to further consider other selections of input data as well as an uncertainty analysis of the input data.

5. Conclusion

Overall, this study tried to provide an accurate prediction of the reproduction number for COVID-19, which is useful for the governments in implementing policies to reduce the spread of the coronavirus, especially in Jakarta, Indonesia. This study demonstrated the application of a dynamic simulation to calculate historical data for the effective reproduction number, and it proposed a machine learning approach called ANFIS to predict its daily movement. The result of the ANFIS prediction was largely determined by the selection of the input data. Based on several input data scenarios that we considered, we found that the input data from the reproduction number for the previous seven days gave the best prediction results, as indicated by the smallest MAPE value. These results were then validated using

several reproduction number data for different periods. The reproduction number for COVID-19 in Jakarta, Indonesia can potentially be controlled because most of its values were lower than one, at least during the last observed period. Our study thus demonstrates promising results for the prediction of the reproduction number for COVID-19 in Jakarta, Indonesia. Increased knowledge of advanced machine learning methods for forecasting applications is needed to develop more accurate prediction results.

Conflicts of Interest: The authors declare no conflict of interest.

References

Abdulrahman, I. K. (2020). SimCOVID: Open-Source Simulation Programs for the COVID-19 Outbreak. *MedRxiv*, 2020.04.13.20063354. https://doi.org/10.1101/2020.04.13.20063354

Al-qaness, M. A. A., Ewees, A. A., Fan, H., & Abd El Aziz, M. (2020). Optimization Method for Forecasting Confirmed Cases of COVID-19 in China. *Journal of Clinical Medicine*, *9*(3), 674. https://doi.org/10.3390/jcm9030674

Ala'raj, M., Majdalawieh, M., & Nizamuddin, N. (2021). Modeling and forecasting of COVID-19 using a hybrid dynamic model based on SEIRD with ARIMA corrections. *Infectious Disease Modelling*, *6*, 98–111. https://doi.org/https://doi.org/10.1016/j.idm.2020.11.007

Anastassopoulou, C., Russo, L., Tsakris, A., & Siettos, C. (2020). Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLOS ONE*, *15*(3), e0230405. https://doi.org/10.1371/journal.pone.0230405

Aqil, A. M. I. (2020). *Jakarta returns to transitional PSBB as Anies claims 'flattened'' COVID-19 curve.'* https://www.thejakartapost.com/news/2020/10/11/jakarta-returns-to-transitional-psbb-as-anies-claims-flattened-covid-19-curve.html

Ardabili, S. F., Mosavi, A., Ghamisi, P., Ferdinand, F., Varkonyi-Koczy, A. R., Reuter, U., Rabczuk, T., & Atkinson, P. M. (2020). COVID-19 Outbreak Prediction with Machine Learning. *Algorithms*, *13*(10), 249. https://doi.org/10.3390/a13100249

Azar, A. T. (2014). Neuro-fuzzy feature selection approach based on linguistic hedges for medical diagnosis. *International Journal of Modelling, Identification and Control*, *22*(3), 195–206. https://doi.org/10.1504/IJMIC.2014.065338

Caccavo, D. (2020). Chinese and Italian COVID-19 outbreaks can be correctly described by a modified SIRD model. *MedRxiv*, 2020.03.19.20039388. https://doi.org/10.1101/2020.03.19.20039388

Caicedo-Ochoa, Y., Rebellón-Sánchez, D. E., Peñaloza-Rallón, M., Cortés-Motta, H. F., & Méndez-Fandiño, Y. R. (2020). Effective Reproductive Number estimation for initial stage of COVID-19 pandemic in Latin American Countries. *International Journal of Infectious Diseases*, *95*, 316–318. https://doi.org/https://doi.org/10.1016/j.ijid.2020.04.069

Chintalapudi, N., Battineni, G., Sagaro, G. G., & Amenta, F. (2020). COVID-19 outbreak reproduction number estimations and forecasting in Marche, Italy. *International Journal of Infectious Diseases*, *96*, 327–333. https://doi.org/https://doi.org/10.1016/j.ijid.2020.05.029

Cooper, I., Mondal, A., & Antonopoulos, C. G. (2020). A SIR model assumption for the spread of COVID-19 in different communities. *Chaos, Solitons & Fractals*, *139*, 110057. https://doi.org/https://doi.org/10.1016/j.chaos.2020.110057

Editor Kompas. (2020). *More Accurate Data is Needed*. https://kompas.id/baca/english/2020/04/27/more-accurate-data-is-needed/

Inglesby, T. V. (2020). Public Health Measures and the Reproduction Number of SARS-CoV-2. *JAMA*. https://doi.org/10.1001/jama.2020.7878

Jain, S., Shukla, S., & Wadhvani, R. (2018). Dynamic selection of normalization techniques using data complexity measures. *Expert Systems with Applications*, *106*, 252–262. https://doi.org/https://doi.org/10.1016/j.eswa.2018.04.008

Jang, J.-. R. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, *23*(3), 665–685. https://doi.org/10.1109/21.256541

Khodaei-mehr, J., Tangestanizadeh, S., Vatankhah, R., & Sharifi, M. (2018). ANFIS-Based Optimal Control of Hepatitis C virus epidemic. *IFAC-PapersOnLine*, *51*(15), 539–544. https://doi.org/10.1016/j.ifacol.2018.09.211

Liu, Y., Gayle, A. A., Wilder-Smith, A., & Rocklöv, J. (2020). The reproductive number of COVID-19 is higher

compared to SARS coronavirus. *Journal of Travel Medicine*, *27Liu*, *Y.,*(2). https://doi.org/10.1093/jtm/taaa021

Neher, R. A., Dyrdak, R., Druelle, V., Hodcroft, E. B., & Albert, J. (2020). Potential impact of seasonal forcing on a SARS-CoV-2 pandemic. *MedRxiv*, 2020.02.13.20022806. https://doi.org/10.1101/2020.02.13.20022806

Nesteruk, I. (2017). *Statistics Based Models for the Dynamics of Chernivtsi Children Disease*. 26–34. https://doi.org/10.20535/1810-0546.2017.5.108577

Nishiura, H., & Chowell, G. (2009). *The Effective Reproduction Number as a Prelude to Statistical Estimation of Time-Dependent Epidemic Trends BT - Mathematical and Statistical Estimation Approaches in Epidemiology* (G. Chowell, J. M. Hyman, L. M. A. Bettencourt, & C. Castillo-Chavez (eds.); pp. 103–121). Springer Netherlands. https://doi.org/10.1007/978-90-481-2313-1_5

Peng, L., Yang, W., Zhang, D., Zhuge, C., & Hong, L. (2020). Epidemic analysis of COVID-19 in China by dynamical modeling. *MedRxiv*, 2020.02.16.20023465. https://doi.org/10.1101/2020.02.16.20023465

Pinter, G., Felde, I., Mosavi, A., Ghamisi, P., & Gloaguen, R. (2020). COVID-19 Pandemic Prediction for Hungary; A Hybrid Machine Learning Approach. *Mathematics*, *8*(6), 890. https://doi.org/10.3390/math8060890

Roda, W. C., Varughese, M. B., Han, D., & Li, M. Y. (2020). Why is it difficult to accurately predict the COVID-19 epidemic? *Infectious Disease Modelling*, *5*, 271–281. https://doi.org/https://doi.org/10.1016/j.idm.2020.03.001

Sahafizadeh, E., & Sartoli, S. (2020). Epidemic curve and reproduction number of COVID-19 in Iran. *Journal of Travel Medicine*. https://doi.org/10.1093/jtm/taaa077

Susanto, H., Tjahjono, V. R., Hasan, A., Kasim, M. F., Nuraini, N., Putri, E. R. M., Kusdiantara, R., & Kurniawan, H. (2020). How Many Can You Infect? Simple (and Naive) Methods of Estimating the Reproduction Number. *Communication in Biomathematical Sciences; Vol 3, No 1 (2020)*. http://journals.itb.ac.id/index.php/cbms/article/view/13808

Tarno, Subanar, & Rosadi, D. (2013). *Analysis of Financial Time Series Data Using Adaptive Neuro Fuzzy Inference System (ANFIS)*.

WHO Emergency Committee. (2020). *Statement on the second meeting of the international health regulations (2005) emergency committee regarding the outbreak of novel coronavirus (2019-ncov)*. https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov)

Zhong, L., Mu, L., Li, J., Wang, J., Yin, Z., & Liu, D. (2020). Early Prediction of the 2019 Novel Coronavirus Outbreak in the Mainland China Based on Simple Mathematical Model. *IEEE Access*, *8*, 51761–51769. https://doi.org/10.1109/ACCESS.2020.2979599