*i*Journals: International Journal of Software & Hardware Research in Engineering (IJSHRE)
ISSN-2347-4890
Volume 9 Issue 2 February 2021

RISK ASSESSMENT PREDICTION EXTREME RAINFALL MODEL FOR FLOOD HAZARD INFORMATION SYSTEM

Rahmat Gernowo[1]; Aris Puji Widodo[2]

[1]Department of Physics, Faculty of Science and Mathematics, Diponegoro University, Semarang, Indonesia

[2]Department of Informatic, Faculty of Science and Mathematics, Diponegoro University, Semarang, Indonesia

E-mail: gernowo@yahoo.com[1]; arispw@gmail.com[2]

DOI: 10.26821/IJSHRE.9.2.2021.9201

ABSTRACT

Hydrometeorological disasters are a major factor in natural disasters in tropical areas, so rainfall predictions are very important information for the community. There are several statistical methods that can be applied in predicting rainfall, namely the POLY-6 (Polynomial of Order Six) method and the ANFIS (Adaptive Neuro Fuzzy Inference System) method. The purpose of this study was to analyze the comparison between the ANFIS method and the POLY-6 method to predict rainfall in the next few periods. The POLY-6 method is a time series data analysis method that is often used for forecasting. Meanwhile, the ANFIS method is a forecasting method based on artificial neural networks combined with fuzzy logic, as input for time series event data which is purely linear or non-linear. The results of this study indicate that the ANFIS method is better for analyzing non-linear time series data than the POLY-6 method. The ANFIS method has better accuracy with a correlation result of 6.9811 and an RMSE value of 87.29%, while the POLY-6 method yields a correlation of 14.037 and an RMSE value of 24.92%. The POLY-6 method for the case of daily data and non-linear data is not suitable, so that the prediction results do not follow the actual data pattern and tend to have constant results or prediction results that tend to be flat.

Keywords: Poly 6, ANFIS, Time Series, linear no linear, rainfall data.

1. INTRODUCTION

Indonesia is known as a country with a tropical climate. This can be seen from the geographic location of Indonesia which is around the equator. Indonesia is surrounded by a vast expanse of sea, evaporation and collection of clouds takes place

very quickly so that the rain process occurs almost every day. This makes Indonesia a country that has high rainfall [1, 2, 3, 4].

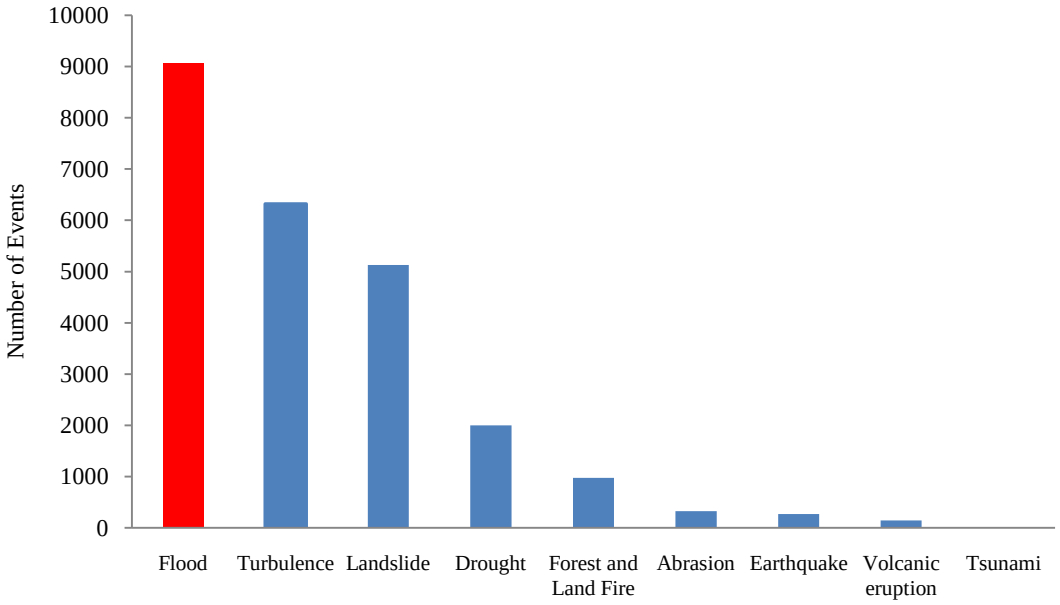

Figure 1. Number of Natural Disaster Events in Indonesia during the period 2000 to 2020 [37].

Based on BNPB data, natural disasters due to anomalies in atmospheric dynamics in Indonesia occupy the highest number, as in Figure 1. Floods followed by turbulence hazard. This is categorized as a hydrometeorology disaster, with events during the period 2000 to 2020 from BNPB data. In this research, prediction of future rainfall events was carried out based on the 6th order polynomial and ANFIS model [5, 6, 7].

Rainfall prediction is very important information for the people of Indonesia. Rainfall prediction accuracy can be used to support various community activities, such as agriculture, plantations and so on. There are several statistical methods that can be applied to forecast monthly rainfall. One of the statistical methods that have been applied by the Meteorology, Climatology and Geophysics Agency in forecasting rainfall is the POLY-6 (Polynomial of Order Six) method, which is a method based on the values of a variable that has occurred in the past and is then used to determine the historical pattern of the data, and then used to extrapolate the pattern in the future [1]. Apart from the POLY-6 method, there is the ANFIS (Adaptive Neuro Fuzzy Inference System) method. The idea of using the ANFIS method as a forecasting method is based on rarely found time series events that are purely linear or non-linear [8, 9, 10].

The ANFIS method uses a neural network to implement a fuzzy inference system. This study compares the POLY-6 method and the ANFIS method in predicting rainfall for the next several periods. The input used is in the form of daily rainfall data. Rainfall data is time series data that describes seasonal patterns. According to the research of [2], ANFIS provides better accuracy than POLY-6 with a smaller RMSE value than POLY-6 [3]. With the results of the comparison between the two methods, it can be used as input to the BMKG to switch to using the ANFIS method for forecasting in order to get more accurate results [11].

2. METHODOLOGI

2.1 POLY-6 (Polynomial of Order Six) method.

Regression is a statistical analysis method used to see the effect between two or more variables. The relationship of these variables is functional which is manifested in a mathematical model. In regression analysis, it is divided into several types, namely Simple Regression (Simple Linear and Simple Nonlinear) and Multiple or Multiple Nonlinear Regression [12, 13, 14].

The Linear Regression equation includes more than two predictors or independent variables, the regression is called Multiple Linear Regression. The Multiple Linear Regression Model has a general regression equation in the form of Equation (1) [15, 16, 17]

$$\hat{Y} = \beta_0 + \beta 1X_1 + \beta_2X_2 + \beta_3X_3 + \ldots + \beta_nX_n + \varepsilon \quad (1)$$

where Y is the dependent variable / dependent, X_1, X_2, X_3, ..., X_n are independent variables, β_0 is the intercept, and β_1, β_2, β_3, ..., β_n are the parameters of the regression coefficients, while Polynomial Regression is a Regression model Linear which is formed by adding up the effect of each predictor variable (X) which is raised to the power of increasing to the nth order. In general, the Polynomial Regression model is written in Equation (2).

$$Y = b_0 + b_1X_1 + b_2X_2 + \ldots + b_nX_n + \varepsilon \quad (2)$$

Information :

Y = critorium variable / dependent / influenced

$b_0 = b_1$ intercept,

$b_2, .., b_n$ = regression coefficients

X = variable predictor / free / affect.

ε = confounding factor that cannot be explained by the regression model

2.2 ANFIS (Adaptive Neuro Fuzzy Inference System)

ANFIS (Adaptive Neuro Fuzzy Inference System) is an adaptive network implementation in fuzy inference system. The fuzzy logic system is adaptive, meaning that the fuzzy system can be adjusted according to the conditions as desired. The ANFIS architecture is similar to a radial functional neural network with certain limitations. In other words, ANFIS is a method that can perform rule settings using a learning algorithm for a set of data. Fuzzy models can be used instead of multiple layers. In this case, the system is divided into two groups, the first is a group in the form of a neural network with fuzzy weights and a fuzzy activation function. Meanwhile, the second group is a neural network with fuzzy input at the first layer or at the second layer. However, in the second group, the weights on the neural network are not fuzzy as in the first group [18].

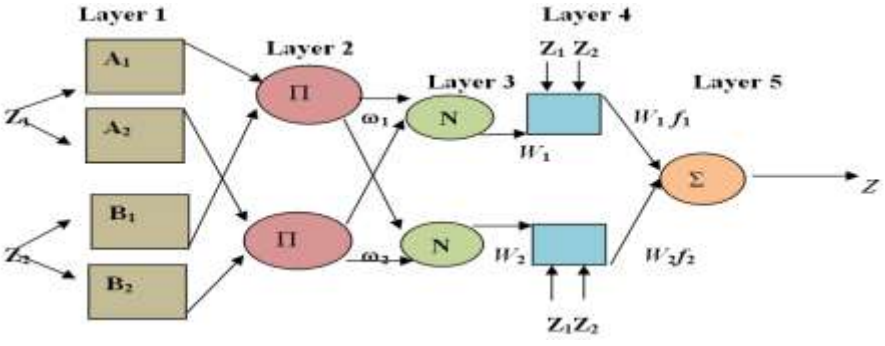

Figure 2. ANFIS network architecture

According to [19, 20, 21], as in Figure 2, the ANFIS network consists of the following layers:

Layer 1:

It is a fuzzification layer, where in this layer each neuron is adaptive to the activation parameter. The output of each neuron is a degree of membership given by the input membership function. For example, the Gaussian membership function can be written in the following equation:

$$\mu(Z) = e^{-\frac{1}{2}\left(\frac{Z-c}{\sigma}\right)^2} \quad (3)$$

Where Z is the input, in this case $Z = \{Z_{1,t}, Z_{2,t}\}$ and $\{\sigma, c\}$ are parameters. If the values of these parameters change, the shape of the curve will also change. This parameter is often called the premise parameter.

Layer 2:

In this layer, a fixed neuron (Π) is the product of all inputs which can be written in the following equation:

$$w_i = \mu_{A_i} . \mu_{B_i} \quad (4)$$

At this layer the AND operator is commonly used and the result of this calculation is called the firing strength of a rule. Where each neuron represents the ith rule.

Layer 3:

In this layer in the form of fixed neurons (N) which is the result of calculating the ratio of the firing strength (w_i) to the total firing strength in the second layer which can be written in the following equation:

$$\bar{w}_i = \frac{w_i}{w_1 + w_2}, i = 1,2 \quad (5)$$

The result of this calculation is called normalized firing strength.

Layer 4:

In this layer in the form of neurons which are adaptive neurons to an output which can be written in the following equation:

$$\bar{w}_i f_i = \bar{w}_i (p_i Z_{1,t} + q_i Z_{2,t} + r_i) \quad (6)$$

Where $\bar{w}_i$ is the normalized firing strength in the third layer and; p_i, q_i, dan r_i are consequent parameters.

Layer 5:

In this layer in the form of a single neuron (Σ) which is the result of the sum of all outputs from the fourth layer which can be written in the following equation:

$$\sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (7)$$

3. RESEARCH OF DATA

The variable used in this study is the daily rainfall data for the Jakarta from January 1, 2002 to January 31, 2020, which consists of 216 data. The data will be divided into two parts, namely training data and testing data. Training data is used for modeling. While the test data is used for comparison with the resulting forecast [22].

The equipment used in this research is the Mat'lab R2015a software which is used to create forecast models in the ANFIS, Minitab18 and eviews methods which are used for model forecasts using the POLY-6 method, and Matlab is used to calculate correlation and the RMSE value (Root

Mean Square Error) between the output results of the two methods [23].

As in (**Table 1**), shows that the major floods in Jakarta were on January 26 - February 1 2002, February 4-14 2007, January 15-24 2013, and February 9 - 12. 2015, and December 31 - January 2 2020. Precursors Floods in Jakarta are mainly caused by extreme rainfall, and a contribution to sea level rise. [2, 3].

Tabel. 1. Rainfall data areas in Jakarta [37].

Date of Occurrence	2 February 2002	2 February 2007	17 January 2013	11 February 2015	1 January 2020
Highest Rainfall (mm / day)	168	340	100	277	377
RW Flooded	353	955	599	702	390
Flooded Area (km^2)	168	455	240	281	156
Low tide> 95% inundation (days)	6	10	7	7	4

4. RESULTS AND DISCUSSIONS

4.1 POLY-6 Analysis on Rainfall Prediction.

While Figure 3 is a stationarity test of the rainfall data for the study area. Stationary data can be known visually or formally with a stationary test, so there is no need for differencing. Visual test by paying attention to the time series data plot, while the formal test can be done by using the ADF (Augmented Dickey-Fuller) test. Furthermore, in Figure 4 is the estimation of POLY-6 modeling by looking at the Polynomial Regression plots [24].

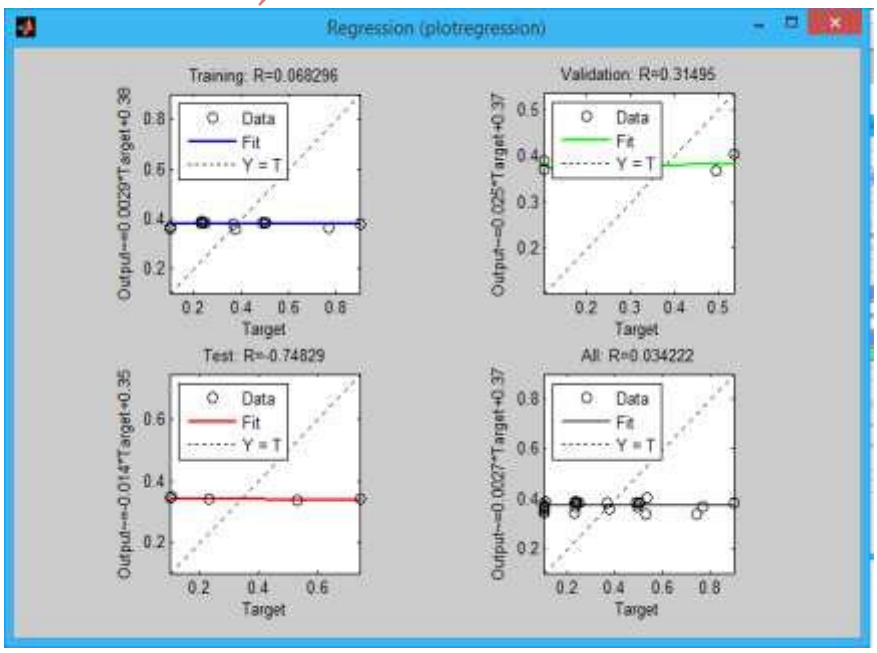

Figure 3. Polynomial Regression R= 0.068296

Based on the Polynomial Regression plots, ten provisional models were obtained and one of the best models was POLY-6 which was then followed by a significance test, residual normality test, independence test, ARCH-LM test and linearity test. The best model equation for POLY-6 can be written as follows:

$$Y_t = 0{,}471 + 0{,}9473\ X_{t-1} - 0{,}8216\ \varepsilon_{t-1} + \varepsilon_t \quad (10)$$

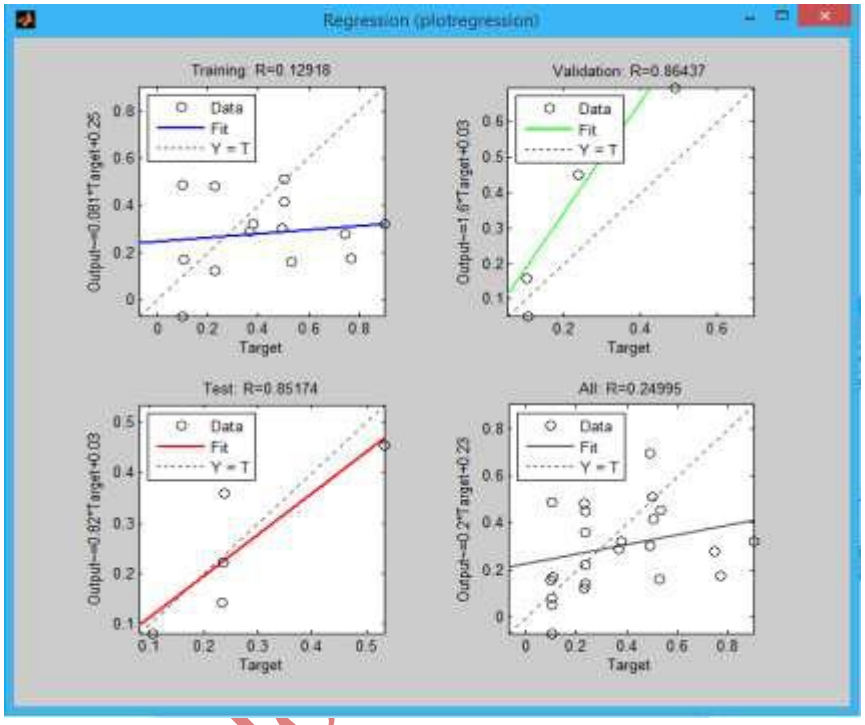

Figure 4. Polynomial Regression R= 0.12918

4.2 ANFIS Analysis on Rainfall Prediction.

ANFIS model formation consists of four inputs and one output. Inputs in the ANFIS modeling consisted of average temperature (°C)($Z_{1,t}$), average humidity (%)($Z_{2,t}$), duration of irradiation (hours) ($Z_{3,t}$) and the average wind speed (Knots))($Z_{4,t}$). Meanwhile, the output is in the form of rainfall (mm) (Z_t). An FIS (Fuzzy Inference System) is used to obtain ANFIS architecture, prior to conducting training at ANFIS the initial parameters are determined first, namely using genfis2 (Generate Fuzzy Inference System) with the type of membership function used is gaussmf (Gaussian). In genfis2, the subclust function is used to perform data clustering, train data inputs and generate fuzzy reasoning using the first order Takagi - Sugeno method. The value of radius or often referred to as range of influence (radius) is closely related to determining the input for the number of membership functions. The smaller the range of influence, the greater the number of inputs to the membership function, which means that the level of accuracy is higher, which is indicated by a smaller error. The training process uses 30 iterations (epochs). Using a range of influence value of 0.2, the ANFIS Training data see Fig. 5 [25, 26]

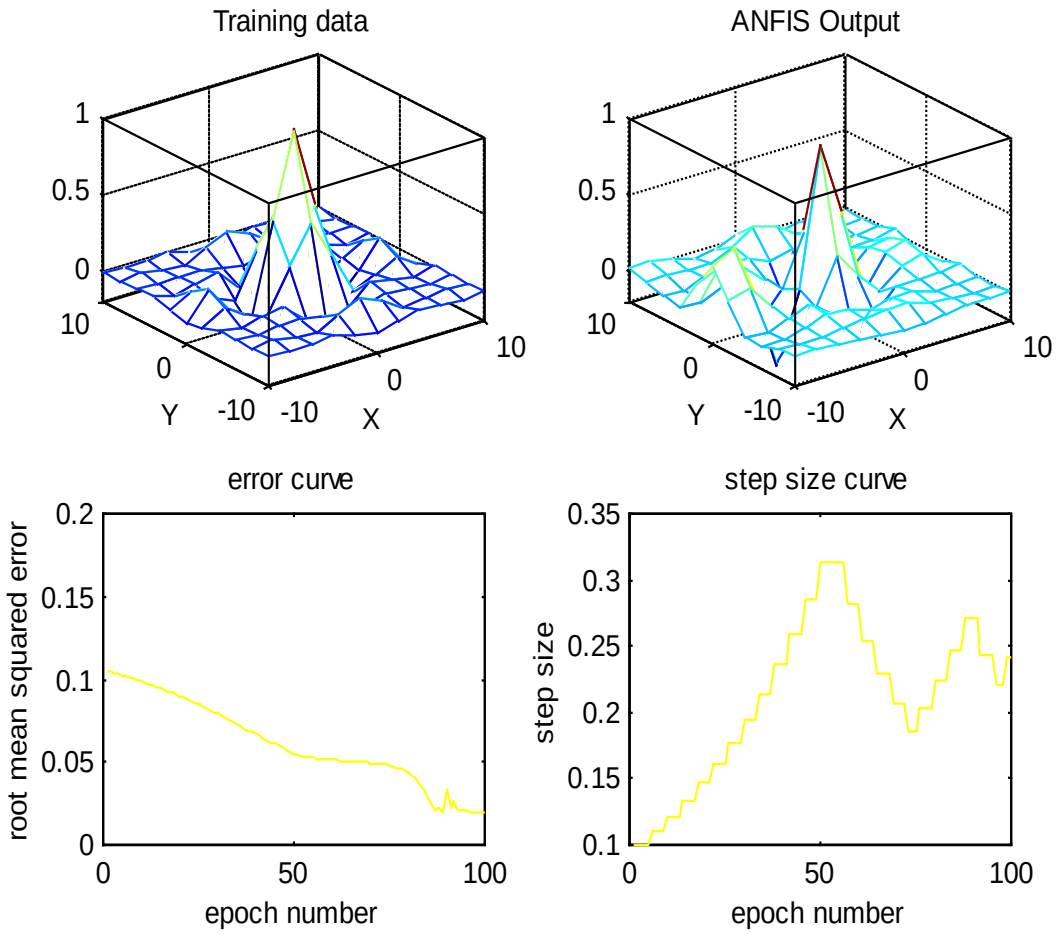

Figure 5. ANFIS Training data.

There are 40 rules with AND operator and consists of one output with 40 MF outputs formed. The lines that connect between nodes show the corresponding rules. From the training results, the total parameters used were 520 parameters consisting of 320 non-linear parameters and 200 linear parameters [27, 28, 29, 30, 31, 32].

5. ANALYSIS AND RESULTS

Based on the results obtained in Figure 7, the POLY-6 method and the ANFIS method can be seen that the ANFIS method is better than the POLY-6 method because it has very good accuracy in forecasting the period of daily rainfall than the POLY-6 method, where the method ANFIS has a greater correlation value, namely 87.29%, with a smaller RMSE value of 6.8911. Meanwhile, POLY-6 produced a correlation of 24.92% and an RMSE value of 14.037.

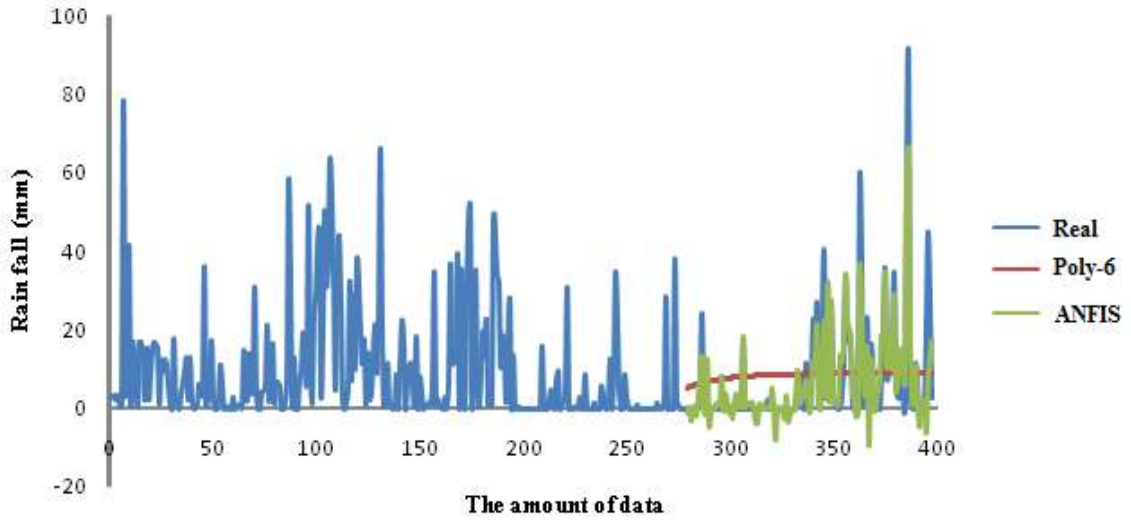

Figure 6. Comparison of the actual rainfall data with the prediction results of the POLY-6 method and the prediction results of the ANFIS method

The output of the ANFIS method almost coincides with the actual data because the ANFIS method is able to analyze non-linear time series data by changing non-linear parameters. Compared with the POLY-6 method, it cannot analyze non-linear time series data because most models of the POLY-6 method are linear [37, 38, 39, 40]. The POLY-6 method for the case of daily data is not suitable, so that the prediction results do not follow the actual data pattern and tend to have constant results or the predicted results tend to be flat. In the ANFIS training process, more data is used than for testing because the more data used for training, the better the training results will be obtained, so that the forecasting system obtained will be better. Whereas the POLY-6 method which uses too much data can result in the residual auto-correlation function not being independent [33, 34, 35, 36].

6. CONCLUSION

Based on the identification of problems and the previous discussion, it can be concluded that the ANFIS method is better for analyzing non-linear time series data than the POLY-6 method. The ANFIS method has better accuracy for rainfall forecasting, with a correlation result of 6.9811 and an RMSE value of 87.29%, while the POLY-6 method yields a correlation of 14.037 and an RMSE value of 24.92%. The POLY-6 method for the case of daily data is not suitable, so that the prediction results do not follow the actual data pattern and tend to have constant results or the predicted results tend to be flat.

7. CONFLICT OF INTEREST

There is no conflict to disclose.

8. REFERENCES

[1]. Diaz, H.F., M.P. Hoerling, and J.K. Eischeid (2001) ENSO variability, teleconnections, and climate change. *Int., J. Climatol*,21, 1845-1862.

[2]. Gernowo R., Adi K., Yulianto T., Seniyatis S., and Yatunnisa A.A. (2018) Hazard Mitigation with Cloud Model-based rainfall and Convective data. *Journal of Physics: Conference Series*, 1025 (1), 012023.

[3]. Gernowo R., Adi K., and Yulianto T. (2017) Convective Cloud model for Analyzing of Heavy rainfall of Weather Extreme at

Semarang Indonesia. *Advanced Science Letter*, 23(7), pp. 6593-6597.

[4]. Hidayat R., Juniarti M.D., and Ma'rufah U. (2018) Impact of La Niña and La Niña Modoki on Indonesia rainfall variability. *IOP Conf. Ser.: Earth Environ. Sci.* **149** 012046

[5]. Masouleh Z.P., Walker D.J. and Crowther J.M. (2019) A Long-Term Study of Sea-Breeze Characteristics: A Case Study of the Coastal City of Adelaide. *J. Appl. Meteor. Climatol.* (2019) **58** (2): 385–400. https://doi.org/10.1175/JAMC-D-17-0251.1

[6]. Prasetyo Y., Yuwono B. D., and Ramadhanis Z. (2018) Spatial Analysis of Land Subsidence and Flood Pattern Basedon DInSAR Method in Sentinel Sar Imagery and WeightingMethod in Geo-Hazard Parameters Combination in NorthJakarta Region. *IOP Conf. Series: Earth and Environmental Science*, 123 012009.

[7]. Roxana C., and Wajsowicz (2005) Forecasting extreme events in the tropical Indian ocean sector climate. *Journal dynamics of atmospheres, and ocean*, pp.1-15.

[8]. Cooley,D., D.Nychka, and P. Naveau (2007) Bayesian spatial modeling of extreme precipitation return levels. *J. Amer.Stat.Assoc.*,102,824–840,https://doi.org/10.1198/016214506000000780.

[9]. Roy I., Tedeschi R.G., and Collins M. (2019) ENSO teleconnections to the Indian summer monsoon under climate changing. *International Journal of Climatology*, Vol. 39, Issue 6, May 2019, Pages 3031-3042.

[10]. Sakurai N., Murata F., Yamanaka M.D., Mori S., Hamada J.I., Hasiguchi H., Tauhid Y.I., Sribimawati T., Suhardi B. (2005) Diurnal Cycle of Cloud System Migration over Sumatera Island. *Journal of the Meteorological* , Vol. 83, Issue 5.

[11]. Makridakis, S., Wheelwright, S. C., Hyndman, R. J., 1998. Forecasting methods and applications, New York, Wiley.

[12]. Novotný, L., 2008. Utilization of finite element method in deformation calculation of assembly platform (in Slovak), Acta Mechanica Slovaca, pp. 607-612.

[13]. Trebu□a, F., et al., 2009. The solution for life extension of casting ladle supporting structure on the continuous casting machine 2 while takingadvantage of until now available conclusions and proposals for life extension of the casting ladle supporting structure, final report, TechnicalUniversity of Košice, Košice, p. 130.

[14]. Ostertag, O., Sivák, P., 2010. Degradation processes and fatigue life prediction (in Slovak), Typopress Košice, Slovakia, ISBN 978-80-553-0486-1.

[15]. Ostertagová, E., 2005. Probability and Mathematical Statistics with examples (in Slovak), Elfa Košice, Slovakia, p. 123, ISBN 80-8086-005-X.

[16]. Ostertagová, E., 2011. Applied Statistic (in Slovak). Elfa Košice, Slovakia, p. 161, ISBN 978-80-8086-171-1.

[17]. Jang, J.S.R, C.T Sun, dan E. Mizutani, 1997, *Neuro-Fuzzy And Soft Computing: A Computational Approach To Learning And Machine Intelligence*, Prentice-Hall, Inc., London

[18]. Trebu□a, F., Maslá□ková, K., Frankovský, P., 2011. "Residual stress measurements," Modelling of Mechanical and Mechatronical Systems, Proceedings of the 4th international conference, Her□any, Slovakia, pp. 487-491.

[19]. Buragohain, M., & Mahanta, C. (2008). A novel approach for ANFIS modeling based on full factorial design. Applied Soft Computing, 8, 609–625.

[20]. BMKG (Meteorology Climatology and Geophysics Council), Climate Analysis. progress report, https://www.bmkg.go.id/

[21]. Chakraborty, S., & Dey, S. (2007). QFD-based expert system for non-traditional machining processes selection. Expert Systems with Applications, 32,1208–1217.

[22]. Ostertagová, E., 2012. Applied Statistic with MATLAB (in Slovak). Equilibria Košice, Slovakia, p. 193, ISBN 978-80-8143-006-0.

[23]. Guiqin, L., Fanhui, K., Wenle, L., Quing feng, Y., & Minglun, F. (2007). The neuralfuzzy modeling and Genetic optimization in WEDM. IEEE International Conference on Control and Automation, 1440–1443.

[24]. Jang, J. S. R. (1993). ANFIS: Adaptive network based fuzzy inference system. IEEE Transactions on Systems, Man Cybernetics, 23, 665–685.

[25]. Kaneko, T., & Onodera, T. (2004). Improvement in machining performance of diesinking EDM by using self-adjusting fuzzy control. Journal of Materials Processing Technology, 149, 204–211.

[26]. Kim, C. H., & Kruth, J. P. (2001). Influence of the electrical conductivity of dielectric on machinability of the WEDM. Transactions of the Korean Society of Mechanical Engineers (A), 27, 1800–1808.

[27]. Lei, Y., He, Z., Zi, Y. (2007). A new approach to intelligent fault diagnosis of rotating machinery. Expert Systems with Applications.

[28]. Lin, C. T., Chung, I. F., & Huang, S. Y. (2001). Improvement of machining accuracy by fuzzy logic at corner parts for wire-EDM. Fuzzy Sets and Systems, 122, 499–511.

[29]. Lin, J. L., & Lin, C. L. (2005). The use of grey fuzzy logic for the optimization of the manufacturing process. Journal of Materials Processing Technology, 160, 9–14.

[30]. Mandal, D., Pal, S. K., & Saha, P. (2007). Modeling of electrical discharge machining process using back propagation neural network and multi-objective optimization using non-dominating sorting genetic algorithm-II. Journal of Materials Processing Technology, 186, 154–162.

[31]. Ozturk, A., Arslan, A., & Hardalac, F. (2008). Comparison of neuro-fuzzy systems for classification of transcranial Doppler signals with their chaotic invariant measures. Expert Systems with Applications, 34, 1044–1055.

[32]. Salman, Ö., & Kayacan, M. C. (2008). Evolutionary programming method for modeling the EDM parameters for roughness. Journal of Materials Processing Technology, 200, 347–355.

[33]. Sengur, A. (2008a). Wavelet transform and adaptive neuro-fuzzy inference system for color texture classification. Expert Systems with Applications, 34, 2120–2128.

[34]. Sengur, A. (2008b). An expert system based on principal component analysis, artificial immune system and fuzzy k–NN for diagnosis of valvular heart diseases. Computers in Biology and Medicine, 38, 329–338.

[35]. Tarng, Y. S., Tseng, C. M., & Chung, L. K. (1997). A fuzzy pulse discriminating system for electrical discharge machining.

International Journal of Machine Tools and Manufacture, 37, 511–522.

[36]. Tzeng, Y. F., & Chen, F. C. (2007). Multi-objective optimization of high-speed electrical discharge machining process using a Taguchi fuzzy based approach. Materials and Design, 28, 1159–1168.

[37]. BNPB (National Disaster Management Authority), *Indonesian Disaster Information Data2000-2020*, (http://bnpb.cloud/dibi/grafik1a).

