Adaptive Neuro Fuzzy Inference System (ANFIS) approach for modeling paddy production data in Central Java

Tarno, A Rusgiyono and Sugito

1 Department of Statistics, Diponegoro University, Semarang, Indonesia
E-mail: tarno.stat@gmail.com

Abstract. The aim of this research is to develop the procedure of constructing an adaptive neuro-fuzzy inference system (ANFIS) model for time series data. The procedure of development applies statistical inference for optimizing ANFIS architecture. In this study, the procedure of Lagrange multiplier (LM) test is used for selecting input variables. Firstly, several lags which are indicated significantly different to zero are divided into 2 clusters, and these lags are selected as optimal inputs of ANFIS based on LM test. Secondly, the cluster numbers of inputs are also determined by using LM-test procedure. Based on this result, a number of rule-bases are generated. The developed model is applied for forecasting paddy production data in Central Java. This study concluded that lag-1, lag-2 and lag-5 with 2 clusters are selected as the optimal inputs. The 1-1 and 2-2 rules are selected as optimal rules. Finally, the model can work well, and generates a very satisfying result in forecasting paddy production data. Based on the root mean squares error (RMSE) and mean absolute percentage error (MAPE) values, the ANFIS performance is better than performance of Autoregressive Integrated Moving Average (ARIMA) for forecasting.

1. Introduction
Time series data are usually characterized by uncertainty, autocorrelation persistence and leptokurtic behavior. The data are usually non-stationary and non-linear [1, 2]. Autoregressive Integrated Moving Average (ARIMA) Box-Jenkins is one of the most popular models which used for time series forecasting [3]. Autoregressive Conditional Heteroscedasticity (ARCH) model proposed by [4] and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model that developed by [5] is the popular variance model. ARIMA-GARCH has been applied in a lot of research for forecasting nonlinear time series data [6]. The model still has some disadvantages when applied for forecasting non-linear time series data, because the series is not appropriate with standardized theoretical phenomenon and also the model cannot capture the uncertainty of data. Using linear models for non-linear data would cause cointegration errors in time series [7, 8].

In recent years, alternative models have been developed to analyze non-linear time series such as neural networks (NN), fuzzy system and its hybrid [9, 10]. ANFIS is a procedure that combines NN and fuzzy system. It has been implemented in many fields of time series research such as application of ANFIS based on singular spectrum analysis for forecasting chaotic time series [11]; chaotic time series prediction using improved ANFIS [12]; fuzzy time series forecasting [13]; developing a new approach for forecasting the trends of oil price [14]; forecasting of stock return[15]; forecasting of financial volatility [16]. Conclusively ANFIS is better than the other methods.

The main focus of this research is inference procedure based on LM-test for selecting the input variables, determining the number of membership function and determining the number of rules. The
procedure is extended based on LM-test for selecting NN model proposed by White in 1989 [17]. The new procedure for determining input variables, determining number of membership function and generating rule-bases of ANFIS applies statistical inference based on LM-test. Finally, the model was implemented for forecasting paddy production. Organization of this paper is as follows: Section 2 discusses the architecture of ANFIS and estimation of consequent parameters are described in Section 3; Section 4 discusses about proposed procedure for selecting model by using statistical inference based on LM-test and Section 5 discusses about application of developed model; and conclusion is discussed in Section 6.

2. ANFIS Architecture

In ANFIS, the number of hidden nodes in neural networks is as well as in a fuzzy system which consists of fuzzification (layer-1), fuzzy inference system (layer-2 and layer-3), defuzzification (layer-4) and aggregation (layer-5). The NN architecture which used in ANFIS architecture has 5 fixed-layers [9]. The ANFIS architecture for time series model with $p$ inputs $Z_{t-1}$, $Z_{t-2}$, $\cdots$, $Z_{t-p}$ and one output $Z_t$ by assuming rule-base of Sugeno first order with $m$ rules [18]

$$
\text{If } Z_{t-k} \text{ is } A_{kj} \text{ then } Z_{t}^{(j)} = \theta_{j0} + \sum_{k=1}^{p} \theta_{jk} Z_{t-k} \theta_{jk}
$$

where $Z_{t-k}$ is $A_{kj}$ as premise section, and $Z_{t}^{(j)} = \theta_{j0} + \sum_{k=1}^{p} \theta_{jk} Z_{t-k}$ as consequent section; $\theta_{jk}$ as linear parameters; $A_{kj}$ as nonlinear parameters; $j = 1,2,\ldots,m$; $k = 1,2,\ldots,p$; and if $w_1$, $w_2$, $\cdots$, $w_m$ are firing strength for $m$ values $Z_{t}^{(1)}$, $Z_{t}^{(2)}$, $\cdots$, $Z_{t}^{(m)}$ then the output $Z_t$ can be expressed as:

$$
Z_t = \overline{w}_1 Z_{t}^{(1)} + \overline{w}_2 Z_{t}^{(2)} + \cdots + \overline{w}_m Z_{t}^{(m)}
$$

(1)

The architecture of ANFIS (figure 1) consists of 5 layers which can be described as follows [9, 18].

Layer-1: Every node in the first layer is adaptive with one parametric activation function. The output is membership degree of given inputs which satisfy membership function $\mu_{A_{kj}}(Z_{t-k})$. One example of membership function is Gaussian membership function (gaussmf) which can be written as:

$$
\mu_{A_{kj}}(Z_{t-k}) = \exp\left(-\frac{1}{2} \left(\frac{Z_{t-k} - c_{jk}}{\sigma_{jk}}\right)^2\right), \quad j = 1,2,\cdots,m, \quad k = 1,2,\cdots,p;
$$

where $c_{j}$ is location parameter and $\sigma_{j}$ is the scaling parameter [9, 18]. The parameters are called as premise parameters.
Layer-2: Every node in the second layer is a fixed node, which the output of this layer is the product of incoming signal. Generally, it uses fuzzy operation AND. The output of each node represents the firing strength $w_j$ of the j-th rule [9, 18].

$$w_j = \prod_{k=1}^{p} \mu_{A_k}(Z_{t-k}), \quad j = 1,2,\ldots,m \tag{2}$$

Layer-3: Every node in the third layer is a fixed node, which computes ratio of firing strength of j-th rule relative to the sum of firing strengths of rules [9,18].

$$\overline{w}_j = \frac{w_j}{\sum_{j=1}^{w} w_j} \tag{3}$$

Layer-4: Every node in the fourth layer is an adaptive node, the output of each node [9,18]

$$\overline{w}_j Z^{(j)}_t = \overline{w}_j (\theta_{j1} Z_{t-1} + \theta_{j2} Z_{t-2} + \cdots + \theta_{jp} Z_{t-p} + \theta_{j0}) \tag{4}$$

Layer-5: Every node in the fifth layer is a fixed node which adds all of incoming signal [9,18]. The output of the whole network can be written as:

$$Z_t = \sum_{j=1}^{m} \overline{w}_j (\theta_{j1} Z_{t-1} + \theta_{j2} Z_{t-2} + \cdots + \theta_{jp} Z_{t-p} + \theta_{j0})$$

### 3. Estimating the consequent parameters of ANFIS

If the general ANFIS model is given as expression (5) then the estimation of consequent parameters $\theta_{j,k}$ ($j = 1,2,\ldots,m$, $k = 1,2,\ldots,p$) can be determined using recursive least squares (RLS) method. Estimating the parameters, a linear equation system is constructed as below.
\[ \hat{\theta}_1(\bar{w}_1 Z_{t-1}) + \hat{\theta}_2(\bar{w}_2 Z_{t-2}) + \ldots + \hat{\theta}_p(\bar{w}_p Z_{t-p}) + \hat{\theta}_0 \bar{w}_1 + \hat{\theta}_2(\bar{w}_2 Z_{t-2}) + \ldots + \hat{\theta}_p(\bar{w}_p Z_{t-p}) \]

\[ + 3 \hat{w}_1 + \ldots + \hat{\theta}_m(\bar{w}_m Z_{t-1}) + \hat{\theta}_2(\bar{w}_2 Z_{t-2}) + \ldots + \hat{\theta}_m(\bar{w}_m Z_{t-p}) + \hat{\theta}_0 \bar{w}_0 = Z_t \]

When the consequent parameters have been estimated, then the premise parameters can be updated using backpropagation method. Optimal quality is obtained by minimizing the ANFIS error function.

4. Proposed procedure for selecting model in ANFIS using Lagrange Multiplier (LM) test

Lagrange Multiplier (LM) test is used for testing hypothesis of adding variables in ANFIS modeling. Some variables should be included to the model because the new inputs, number of clusters or number of rules to be added in the model.

4.1 Procedure for adding input variables

Determining input can be identified by lag plot of data or the partial autocorrelation function (PACF) plot. Lag plot can be used for testing linearity. The lag plot or PACF plot can be used for identification of autoregressive (AR) input. Based on lag plot and PACF plot, the significant lags should be tested as input variables of ANFIS. Determining input is performed by constructing models which involve some input variables with minimum number of clusters and minimum number of rules. Firstly, construct the models with 1 input variable, 2 clusters and 2 rules and then select the model which has the largest determination coefficient. Adding input is done by using LM-test procedure to get a model with optimal inputs. If given \( p \) input variables \( Z_{t-1}, Z_{t-2}, \ldots, Z_{t-p} \) with \( m \) clusters, then the restricted model for this case can be written as:

\[ Z_t = \sum_{j=k}^{m} \theta_{jk} (\bar{w}_j Z_{t-k}) + \sum_{j=1}^{m} \theta_{j0} \bar{w}_j + \epsilon \]

where \( \epsilon \sim N(0, \sigma^2) \), and unrestricted model for adding by one input

\[ Z_t = \sum_{j=1}^{m} \hat{\theta}_{jk} (\bar{w}_j Z_{t-k}) + \sum_{j=1}^{m} \hat{\theta}_{j0} \bar{w}_j + \nu_t \]

where \( \nu_t \sim N(0, \sigma^2) \).

Formulation of null hypothesis for adding variable:

\[ H_0 : \theta_{j(p+1)} = 0; \quad j = 1, 2, \ldots, m \]

The steps for hypothesis test by using LM-test are as follows:

Step-1: Estimate the parameters of restricted model \( \hat{\theta}_{jk}, \hat{\theta}_{j0} \) for \( j = 1, 2, \ldots, m; \quad k = 0, 1, 2, \ldots, p \)

Step-2: Determining residual:

\[ \hat{\epsilon} = Z_t - \sum_{j=1}^{m} \hat{\theta}_{jk} (\bar{w}_j Z_{t-k}) - \sum_{j=1}^{m} \hat{\theta}_{j0} \bar{w}_j \]

Step-3: Regress the residual \( \hat{\epsilon} \) with: a constant, \( (\bar{w}_j Z_{t-k}), \bar{w}_j, \quad j = 1, 2, \ldots, m; \quad k = 0, 1, 2, \ldots, p \)

and determine the value of LM = \( n^* R^2_z \). The distribution of statistics LM is chi-square with degree freedom \( m \).

4.2 Procedure for adding number of clusters of inputs

Adding the number of clusters of each input can be executed using LM-test procedure when ANFIS model with optimal inputs was obtained. If given \( p \) input variables \( Z_{t-1}, Z_{t-2}, \ldots, Z_{t-p} \) with number of clusters \( m \), then the restricted model for this case can be written as:

\[ Z_t = \sum_{j=1}^{m} \sum_{k=1}^{p} \theta_{jk} (\bar{w}_j Z_{t-k}) + \sum_{j=1}^{m} \theta_{j0} \bar{w}_j + \epsilon \]

where \( \epsilon \sim N(0, \sigma^2) \), and unrestricted model for adding by one cluster.
\[ Z_t = \sum_{j=1}^{m+1} \sum_{k=1}^{p} \theta_{jk}(\bar{w}_j Z_{t-k}) + \sum_{j=1}^{m+1} \theta_j \bar{w}_j + \nu_t \]  

(10)

where \( \nu_t \sim N(0, \sigma^2_\nu) \).

The null hypothesis for testing adding variables \((\bar{w}_{m+1} Z_{t-1}), (\bar{w}_{m+1} Z_{t-2}), \ldots, (\bar{w}_{m+1} Z_{t-p})\) can be formulated as follow:

\[ H_0 : \theta_{(m+1)1} = \theta_{(m+1)2} = \cdots = \theta_{(m+1)p} = 0. \]

Procedure of hypothesis test using LM-test can be done as the following steps:

- **Step 1:** Estimate the parameters of restricted model:
  \[ \hat{\theta}_{11}, \hat{\theta}_{12}, \ldots, \hat{\theta}_{1p}, \hat{\theta}_{01}, \hat{\theta}_{21}, \hat{\theta}_{22}, \ldots, \hat{\theta}_{2p}, \hat{\theta}_{m1}, \hat{\theta}_{m2}, \ldots, \hat{\theta}_{mp}, \hat{\theta}_{m0}. \]

- **Step 2:** Determine the estimate of residual:
  \[ \hat{\epsilon}_t = Z_t - \sum_{j=1}^{m+1} \sum_{k=1}^{p} \hat{\theta}_{jk}(\bar{w}_j Z_{t-k}) - \sum_{j=1}^{m+1} \hat{\theta}_j \bar{w}_j \]

- **Step 3:** Regress the residual \( \hat{\epsilon}_t \) with a constant, \((\bar{w}_j Z_{t-k}), \bar{w}_j, j = 1, 2, \ldots, m, m+1; k = 0, 1, 2, \ldots, p \) (is called as auxiliary regression) and determine the value of \( LM = n \times R^2_{\epsilon} \). The distribution of statistics LM is chi-square with degree freedom \( p + 1 \).

### 4.3 Procedure for generating rule bases

The rule-bases of ANFIS are generated based on the optimal input variables and an optimal number of clusters. The possible combination of rules which can be generated is \( mp \), where \( m \) is number of membership functions and \( p \) is number of inputs. For instance, given \( p \) inputs \( Z_{t-1}, Z_{t-2}, \ldots, Z_{t-p} \) with \( m \) clusters, and one output \( Z_t \) by assuming first-order Sugeno fuzzy system with \( m \) rules:

- If \( Z_{t-1} \) is \( A_{1j} \) and \( Z_{t-2} \) is \( A_{2j} \) and \( Z_{t-p} \) is \( A_{pj} \) then
  \[ Z_t^{(j)} = \theta_{j1} Z_{t-1} + \theta_{j2} Z_{t-2} + \cdots + \theta_{jp} Z_{t-p} + \theta_{j0}, j = 1, 2, \ldots, m; k = 1, 2, \ldots, p \]

then the output of ANFIS is given by Eq. (11).

\[ Z_t = \sum_{j=1}^{m} \sum_{k=1}^{p} \theta_{jk}(\bar{w}_j Z_{t-k}) + \sum_{j=1}^{m} \theta_j \bar{w}_j \]  

(11)

### 5. Results and Discussion

As an implementation of ANFIS modeling for forecasting time series data, we construct ANFIS model for forecasting paddy production as a case study. In this research, the monthly paddy production data from January 2009 to December 2016 (www.bps.go.id) are used for constructing the ANFIS model. The data are transformed by logarithm function for reducing the heterogeneity of variance. Both time series and ACF plots show that the paddy production data are stationary in mean and in variance (see Figure 2). Based on the partial autocorrelation function (PACF) plot (Figure 2) lag-1, lag-2 and lag-5 are significantly different to zero.
By using LM test, the variables lag-1, lag-2 and lag-5 with 2 clusters (membership functions) can be selected as input variables in ANFIS. The results of input selection are shown in table 1.

**Table 1. Result of LM test for input selection**

| Input                          | $R^2(\ln Z_t)$ | $R^2(\epsilon)$ | LM  | p-value |
|-------------------------------|----------------|------------------|-----|---------|
| $\ln(Z_{t-1})$                | 0.414          | -                | -   | -       |
| $\ln(Z_{t-2})$                | 0.131          | -                | -   | -       |
| $\ln(Z_{t-5})$                | 0.504          | -                | -   | -       |
| $\ln(Z_{t-1}), \ln(Z_{t-2})$ | 0.526          | 0.506            | 48.12| 0.00   |
| $\ln(Z_{t-1}), \ln(Z_{t-5})$ | 0.436          | 0.423            | 40.22| 0.00   |
| $\ln(Z_{t-1}), \ln(Z_{t-2}), \ln(Z_{t-5})$ | 0.595 | 0.189 | 17.23 | 0.00 |

Based on the previous steps we can conclude that lag-1, lag-2 and lag-5 with 2 membership functions are considered as input ANFIS. The optimal inputs will be evaluated for determining the optimal number of membership functions by using LM test. The results of membership function selection are shown in table 2.

**Table 2. Result of LM test for membership functions selection**

| Input                          | $MF_5$ | $R^2(\epsilon)$ | LM  | p-value |
|-------------------------------|--------|------------------|-----|---------|
| $\ln(Z_{t-1}), \ln(Z_{t-2}), \ln(Z_{t-5})$ | 2      | 0.189            | 17.23| 0.00   |
|                               | 3      | 0.225            | 20.50| 0.00   |

According to the results in table 2 variable inputs lag-1, lag-2 and lag-5 with 3 membership functions...
can also be selected as inputs of ANFIS. By considering parsimonious principle, variable inputs lag-1, lag-2 and lag-5 with 2 clusters are determined as inputs optimal of ANFIS. If given 3 inputs \( Z_{t-1}, Z_{t-2} \) and \( Z_{t-5} \) with 2 number of clusters, and one output \( Z_t \) by assuming first-order Sugeno fuzzy system with 2 rules:

Rule-1: If \( Z_{t-1} \) is \( A_{11} \) and \( Z_{t-2} \) is \( A_{21} \) and \( Z_{t-5} \) is \( A_{51} \) then \( Z_t^{(1)} = \theta_{11}Z_{t-1} + \theta_{12}Z_{t-2} + \theta_{15}Z_{t-5} + \theta_{10} \),

Rule-2: If \( Z_{t-1} \) is \( A_{12} \) and \( Z_{t-2} \) is \( A_{22} \) and \( Z_{t-5} \) is \( A_{52} \) then \( Z_t^{(2)} = \theta_{21}Z_{t-1} + \theta_{22}Z_{t-2} + \theta_{25}Z_{t-5} + \theta_{20} \),

the optimal model can be written as follow:

\[
\ln(Z_t) = 0.043\bar{w}_{1,t}\ln(Z_{t-1}) - 0.538\bar{w}_{1,t}\ln(Z_{t-2}) - 0.08\bar{w}_{1,t}\ln(Z_{t-5}) + 21.801\bar{w}_{1,t} \\
+1.043\bar{w}_{2,t}\ln(Z_{t-1}) - 0.929\bar{w}_{2,t}\ln(Z_{t-2}) - 0.798\bar{w}_{2,t}\ln(Z_{t-5}) + 22.44\bar{w}_{2,t} \quad (12)
\]

where

\[
\bar{w}_{1,t} = \frac{w_{1,t}}{w_{1,t} + w_{2,t}} \quad \bar{w}_{2,t} = \frac{w_{2,t}}{w_{1,t} + w_{2,t}}
\]

\[
w_{2,t} = \exp \left\{ -\frac{1}{2} \left( \frac{\ln(Z_{t-1} - 13.884)}{0.481} \right)^2 + \left( \frac{\ln(Z_{t-2} - 13.794)}{0.507} \right)^2 + \left( \frac{\ln(Z_{t-5} - 13.150)}{0.604} \right)^2 \right\}
\]

\[
w_{2,t} = \exp \left\{ -\frac{1}{2} \left( \frac{\ln(Z_{t-1} - 12.974)}{0.570} \right)^2 + \left( \frac{\ln(Z_{t-2} - 13.159)}{0.548} \right)^2 + \left( \frac{\ln(Z_{t-5} - 13.742)}{0.461} \right)^2 \right\}
\]

The MAPE and RMSE values of forecasting for data using Eq.12 are 1.5% and 0.256 respectively. Whereas the MAPE and RMSE values using ARIMA([1,2,5],0,0) are 2.36% and 0.396 respectively. Therefore, the performance of ANFIS model is better than ARIMA model for forecasting paddy production data in Central Java Province. The predicted value of paddy production data based on model (12) can be seen as figure 3.

**Figure 3.** Paddy production data and its prediction.

### 6. Conclusion

Based on the empirical study can be concluded that the developed model has a good performance for forecasting paddy production data. An optimal ANFIS model for forecasting paddy production consists of 3 input variables (lag-1, lag-2 and lag-5), 2 membership functions and 2 rules. The performance of ANFIS is better than ARIMA for forecasting paddy production data in Central Java Province. The MAPE and RMSE values of forecasting for data using ANFIS are 1.5% and 0.256. Whereas the MAPE
and RMSE values using ARIMA([1,2,5],0,0) are 2.36% and 0.396 respectively. In this case performance of ANFIS is better than ARIMA.

Acknowledgments
We would like to thank Rector of Universitas Diponegoro; Directorate Research and Public Services Ministry of Research, Technology and Higher Education Republics of Indonesia for supporting this research.

References
[1] Tsay R S 2005 Analysis of Financial Time Series 2nd Edition (New York: Wiley-Interscience A John Wiley & Sons Inc. Publication).
[2] Wei W W S 2006 Time Series Analysis: Univariate and Multivariate Methods Second Edition (Boston: Pearson Education Inc.).
[3] Box G E P, Jenkins G M and Reinsel G C 1994 Time Series Analysis, Forecasting and Control, 3rd edition (New York: Englewood Cliffs: Prentice Hall).
[4] Engle R F 1982 Econometrica 50 987.
[5] Bollerslev T 1986 Journal of Econometrics 31 307
[6] Fahimifard S M, Homayounifar M, Sabouhi M and Moghaddamnia A R 2009 Journal of Applied Sciences 9 3641.
[7] Talebizadeh M and Moridnejad A 2011 Expert Systems with Applications 38 4126.
[8] Haykin 1999 Neural Networks: A Comprehensive Foundation Second Edition (New York: Prentice Hall International, Inc.).
[9] Jang J-S R, Sun C-T and Mizutani E 1997 Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligent (New York: Prentice Hall International).
[10] Daneshmand H, Tavousi T, Khosravi M and Tavakoli S 2015 Journal of the Saudi Society of Agricultural Sciences 14 33.
[11] Abdollahazade M, Miranian A, Hossein H and Iranmanesh H 2015 Information Sciences 295 107.
[12] Behmanesh M, Mohammadi M and Naeini V S 2014 International Journal of Soft Computing and Engineering (IJSCE) 4 25.
[13] Cheng S-H, Chen S-M and Jian W-S 2016 Information Sciences 327 272.
[14] Cheng C H and Wei LY 2010 Optimization and Engineering 11 303.
[15] Wei LY, Chen TL and Ho TH 2011 Systems with Applications 38 13625.
[16] Luna I and Ballini R 2012 Journal of Fuzzy and Intelligent Systems 23 27.
[17] Anders U and Korn O 1999 Neural Networks 12 309.
[18] Tarno T, Suparti S and Ispriyanti D 2018 Model Assisted Statistics and Applications 13 45.