Model Selection in Adaptive Neuro Fuzzy Inference System (ANFIS) by using Inference of R^2 Incremental for Time Series Forecasting

Indah Puspitasari[1], Brodjol Sutijo S U[2], Suhartono[3]

[1, 2,3]Statistics Department
Sepuluh Nopember Institute of Technology Surabaya, Indonesia
[1]*indah.puspitasari11@mhs.statistika.its.ac.id*
[2]*brodjol_su@statistika.its.ac.id*
[3]*suhartono@statistika.its.ac.id*

Abstract: ***The aim of this paper is to propose a procedure for model selection in Adaptive Neuro-Fuzzy Inference System* (ANFIS) *for time series forecasting. In this paper, we focus on the model selection based on statistical inference of R^2 incremental. The selecting model is conducted by evaluating the inputs, number of membership functions and rules in architecture of ANFIS until the contribution of R^2 incremental was not significant. We use simulation data as a case study. The results show that statistical inference of R^2 incremental is an effective procedure for model selection in ANFIS for time series forecasting.***

Keywords: ANFIS, R^2 incremental, time series forecasting

1. Introduction

In the development of time series analysis, many observations are the nonlinear observation or namely the relationship between events in the past and now are nonlinear. Thus, the linear time series modeling is not sufficient and appropriate for such cases. One method based on machine learning that can be used for time series analysis is the Adaptive Neuro-Fuzzy Inference System (ANFIS). ANFIS was introduced by Jang [1] and this method can be used on linear or nonlinear function (Universal approximation) [2]. ANFIS based on fuzzy inference system trained using the learning algorithm derived from neural network system [3]. Thus, the ANFIS method has all the advantages possessed by a fuzzy inference system and neural network systems. The architecture in ANFIS consists of 5 fixed layers [4]. Because of the advantages of ANFIS, many researches in the various fields apply ANFIS for analyzing [5]-[12].

Research on procedure of input selection of ANFIS ever conducted using a combination of each two input and selecting the best input based on MSE [13]. Currently, no procedure of ANFIS selection that is combines between the input variables, number of membership function and rule of ANFIS to find the optimum ANFIS. Kaashoek and Van Dijk introduce a pruning procedure by implementing R^2 incremental to select the optimal number of units in the hidden layer and input variables of neural network [14]. The contribution of R^2 incremental is then tested by the statistical inference test. There are three steps procedure is used. First, modeling data with a simple model, then modeling data with a complex model and the final step is perform statistical test of R^2 incremental contribution [15].

Referring to these researches [14] [15], the aim of this research is to propose a procedure for obtaining optimum ANFIS model using statistical inference of R^2 incremental. R^2 incremental in ANFIS is used to determine the input variables, number of membership function and rules of ANFIS.

2. Methods

2.1 Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS is the abbreviated of Adaptive Neuro-Fuzzy Inference System. Actually, this method is like a fuzzy inference system with this different that here by using a back propagation tries to minimize the error. The performance of this method is like both neural network and fuzzy logic. In both neural network and fuzzy logic case, the input pass through the input layer (by input membership function) and the output could be seen in output layer (by output membership functions). Since, in this type of advanced fuzzy logic, neural network has been used, therefore, by using a learning algorithm the parameters have been changed until reach the optimal solution. Actually, in this type the fuzzy logic tries by using the neural network advantages to adjust its parameters. As we know, the different between real and network output in neural network is one of the common method to evaluate its performance. Therefore, ANFIS uses either back propagation or a combination of least squares estimation and back propagation for membership function parameter estimation [16].

Assume that the considered FIS has two inputs (Z_{t-1} and Z_{t-2}) and one output. For a first-order Sugeno fuzzy model, a common rule set with two fuzzy if–then rules is as follows:

If Z_{t-1} is A_1 and Z_{t-2} is B_1 then $Z_t^{(1)} = c_{1,1}Z_{t-1} + c_{1,2}Z_{t-2} + c_{1,0}$ (1)

If Z_{t-1} is A_2 and Z_{t-2} is B_2 then $Z_t^{(2)} = c_{2,1}Z_{t-1} + c_{2,2}Z_{t-2} + c_{2,0}$ (2)

Figure 1 is an example of ANFIS with two inputs (Z_{t-1} and Z_{t-2}) and two rules.

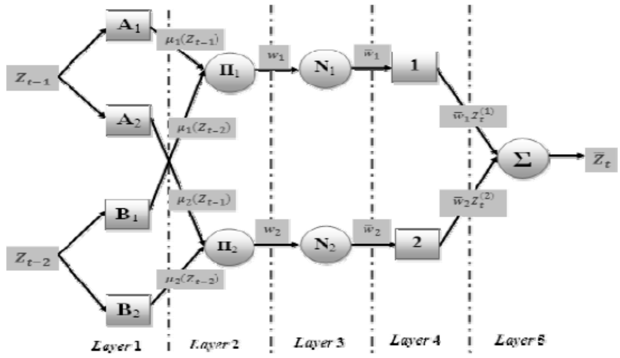


Figure 1. Framework of ANFIS

2.2 Statistically inference of R^2 incremental contribution

Kaashoek and Van Dijk [14] stated that a natural candidate for quantification of the network performance is the square of the correlation coefficient of Z and $\hat{Z}$.

$$R^2 = \frac{(\hat{Z}'Z)^2}{(Z'Z)(\hat{Z}'\hat{Z})}, \qquad (3)$$

where $\hat{Z}$ is the vector of network output points. The network performance with only one unit hidden cell deleted can be measured in a similar way. For instance, if the contribution of hidden cell h is put to zero $(\beta_h = 0)$, then the network will produce an output $\hat{Z}_{-h}$ with errors

$$e_{-h} = Z - \hat{Z}_{-h}. \qquad (4)$$

This reduced network can be measured by the square of the correlation coefficient R^2_{-h} between Z and $\hat{Z}_{-h}$,

$$R^2_{-h} = \frac{(\hat{Z}'_{-h}Z)^2}{(Z'Z)(\hat{Z}'_{-h}\hat{Z}_{-h})}. \qquad (5)$$

Now the R^2 incremental contribution of unit hidden cell h is given as

$$R^2_{(h)} = R^2 - R^2_{-h}. \qquad (6)$$

Unlike the neural network, ANFIS architecture consists of 5 fixed layers fixed. Thus, R^2 incremental is used to select the optimization on a number of rules. Additionally, R^2 incremental can be used to evaluate the number of inputs and number of membership functions.

Suhartono [15] stated a new forward procedure based on the statistical inference of R^2 incremental contribution. This approach involves three basic steps.

1. Begin with the simple model considered to be appropriate for the data, which in this context is called the reduced or restricted model. Then, fit this reduced model and obtain the error sum of squares, denoted by $SSE(R)$.
2. Next, consider the complex or full model, then fit this model and obtain the error sum of squares. The error sum of squares of this full model denoted by $SSE(F)$.
3. Finally, calculate the test statistic:

$$F^* = \frac{SSE(R) - SSE(F)}{df_R - df_F} \div \frac{SSE(F)}{df_F}, \qquad (7)$$

for large n, this test statistic is distributed approximately as $F(v1 = df_R - df_F, v2 = df_F)$ when H_0 holds, i.e. additional parameters in full model all equal to 0. The Equation (**Error! Reference source not found.**) can be written in R^2 incremental contribution as

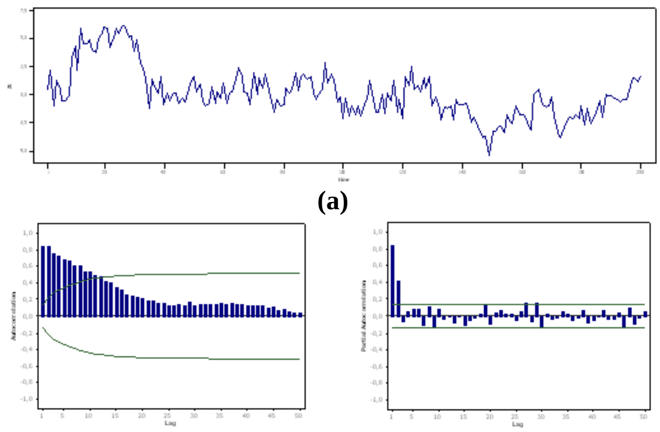
$$F^* = \frac{R^2_{(Incremental)}}{df_R - df_F} \div \frac{(1 - R^2_F)}{df_F}. \qquad (8)$$

3. Proposed Procedure for Determining Optimum ANFIS

In this research, we consider the simulated data. The simulation experiment is carried out to show how the procedure based on R^2 incremental works. Simulated data are generated as ARMA(2, 0) model, i.e.

$$Z_t = 0{,}5Z_{t-1} + 0{,}4Z_{t-2} + a_t. \qquad (9)$$

To show the importance of the size of the time series T for each strategy, we will compare different sizes, namely $T \in \{50, 100, 200\}$ and we replicate up to 50 replication. Figure 2 shows Time series, ACF and PACF plots of this simulated data sample with T = 200 observations.



(a)

(b) (c)

Figure 2. Time series Plot (a), ACF (b), PACF (c) of Simulated ARMA(2,0) Model

The optimization of ANFIS steps are stated in the following four stages:

step 1 : ANFIS in this simulation is ANFIS with Gaussian functions. ANFIS modeling is started by using only one input and the simplest number of membership functions, i.e. 2 number of membership functions with the most minimal rules, i.e. 2 rules. The 1st input is determined from the value of R^2. The best starting input is input variable with the largest R^2,

step 2 : Modeling data using two input variables then compute the inference of R^2 incremental statistically. If the contribution of R^2 incremental is significant, so we add the other input. However if it is not significant, then the optimal input of ANFIS is by using one input. The optimization of variable input quit when the contribution of R^2 incremental is not significant,

step 3 : After obtaining the optimal input, then we increase the number of membership functions, initially 2 number of membership into 3 number of membership functions with the most minimal rules. We calculate the value of R^2 incremental and also test the contribution of R^2 incremental. In this study, we only define the maximum number of membership function is 3 number of membership. Thus, if R^2 incremental is significant, then the optimal number of membership function is 3 number of membership. Additionally, if R^2 incremental is not significant, then the optimal number of membership by using 2 number of membership functions.

step 4 : The last optimization is select the optimal rules. R^2 incremental is calculated after the number of rule has been added. If it shows the significant result then we increase the number of rule again and if it not significant, then the best model is the last model before the R^2 incremental is not significant.

4. Empirical Result

First, we apply the proposed forward procedure starting with ANFIS by using 3 input variables $(Z_{t-1}, Z_{t-2}, Z_{t-3})$. The result of sample input optimization steps are reported in Table 1.

Table 1. Optimization of Input Variables

Lag input	AIC	SBC	R^2	R^2 increment	F	P-value
Z_{t-1}	56.64	48.64	0.75			
Z_{t-2}	72.01	95.82	0.73			
Z_{t-3}	131.08	173.49	0.64			
Z_{t-1}, Z_{t-2}	27.95	15.95	0.79	0.04	17.25	0.00
Z_{t-1}, Z_{t-3}	58.38	78.19	0.76	0.01	1.11	0.34
$Z_{t-1}, Z_{t-2}, Z_{t-3}$	28.05	12.05	0.80	0.01	1.89	0.15

Table 1 shows that an initial input is lag 1 because it has the largest value of R^2. Furthermore, we increase number of input variables into two and three variables. The result optimization based on inference of R^2 incremental conclude the best input variables by using lag 1and lag 2 (Z_{t-1}, Z_{t-2}) as inputs of ANFIS. Using three inputs show R^2 incremental is not significant. Then, we continue the optimization to select the optimal number of membership function.

Table 2. Optimization of Number of Membership Function

MF	AIC	SBC	R^2	R^2 increment	F	P-value
2	27.947	47.646	0.791			
3	21.426	50.975	0.804	0.013	4.112	0.007

The results of optimization steps for determining the optimal number of membership functions are reported in Table 2. It shows that using 3 number of membership function is the optimal result. Plot input variables with the optimal number of membership functions that have been obtained can be seen in Figure 2. Data are divided into 3 groups, where each group of data patterns seen clearly.

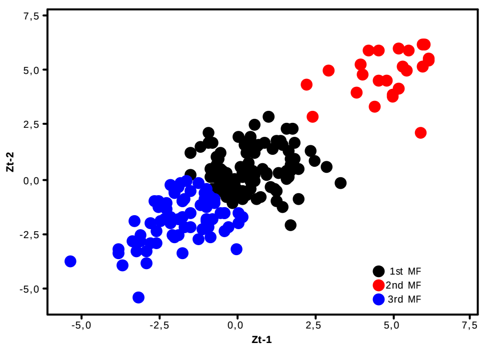


Figure 3. Plot of Lag 1 and Lag 2 with 3 numbers of Membership Functions

Last step is about rule optimization. The optimization result can be seen in Table 3. Because the optimal number of membership functions is 3 numbers of memberships, so that the most minimal rules are 3 rules. We use it as the initial rules in ANFIS and increase the number of rules into 3 rules then apply significant test of R^2 increment statistically.

Table 3. Optimization of Rules

Rule	AIC	SBC	R^2	R^2 increment	F	P-value
3	21.426	50.975	0.804			
4	20.661	60.059	0.810	0.006	0.000	0.095

The result shows that ANFIS with 3 rules is the best architecture model and further optimization runs are not needed because contribution of R^2 incremental is not significant when we add another rule. Hence, this forward procedure yields the optimal ANFIS with 2 input variables, i.e. (Z_{t-1}, Z_{t-2}), 3 number of membership function and 3 rules of ANFIS.

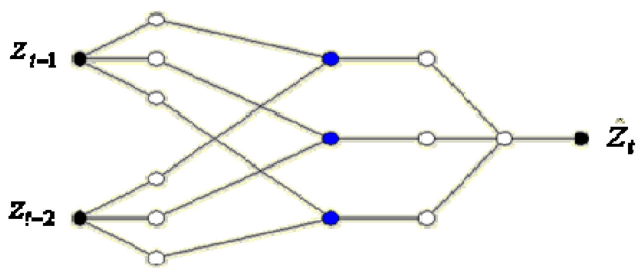


Figure 4. Optimal Architecture of ANFIS

Replications are performed for different size of data to obtain the level of significance in optimization of input variables, the number of membership functions and rules of ANFIS. The high level of significance shows that R^2 incremental tends to be significant. Conversely, if the level of significance is small, then R^2 incremental tends to be not significant.

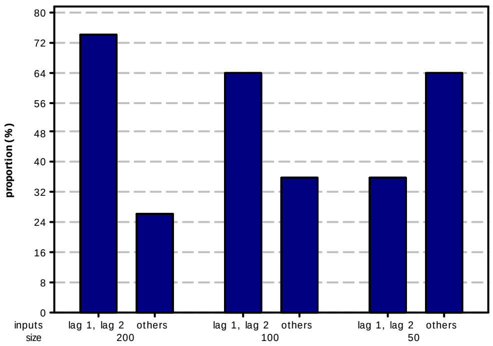


Figure 5 describes the significant proportion of R^2 incremental based on the input variables. Using 2 input variables lag 1 and lag 2 (Z_{t-1} and Z_{t-2}) produce a high proportion of significance, which is not less than 60% of R^2 incremental are significant in 200 and 100 sample size.

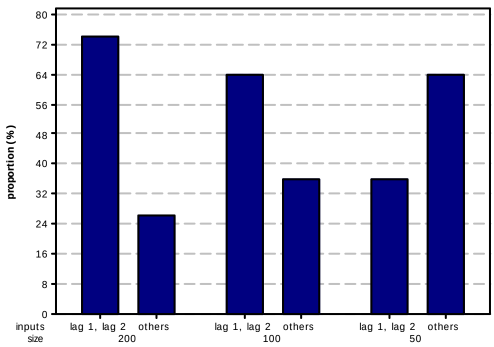


Figure 5. Significant Proportion of R^2 Increment for Input Optimization

When we have 50 data, significant proportion of R^2 incremental for using lag 1 and lag 2 (Z_{t-1}, Z_{t-2}) as input variable is 36%. The other inputs are lag 1 (Z_{t-1}) that produces 34% significant proportion of R^2 incremental and also lag 1, lag 2 and lag 3 ($Z_{t-1}, Z_{t-2}, Z_{t-3}$) with 30% proportion of significance. Based on the optimization of input variables, we can conclude that using lag 1 and lag 2 (Z_{t-1} and Z_{t-2}) as inputs of ANFIS is better than other inputs in AR(2) model.

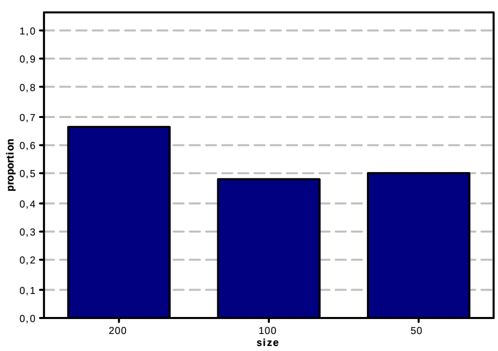


Figure 6. Significant Proportion of R^2 Increment for Number of Membership Function Optimization

For significant proportion based on the number of membership function show that ANFIS with 3 number of membership functions produce more than 60% of R^2 incremental is significant in 200 sample sizes. It means that using 3 number of membership tends to have significant of R^2 incremental.

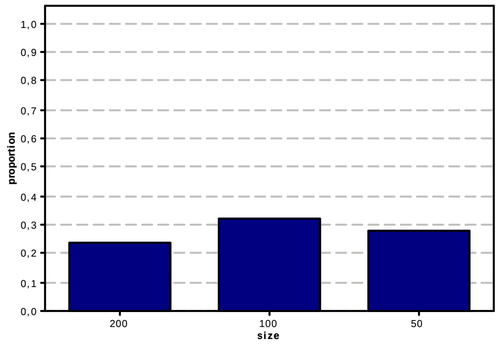


Figure 7. Significant Proportion of R^2 Increment for Rule Optimization

The rule optimization shows the addition of a new rule is not likely to have a significant effect. The significance proportion of R^2 incremental is less than 40%. Therefore, it means that optimum rule is 3 initial rules of ANFIS. Based on the analysis, it can be seen that the inference of R^2 incremental is an effective procedure for the selection of optimal ANFIS model which includes the selection of input variables, the number of membership function and rules of ANFIS.

5. Conclusion

Based on the results at the previous section, we can conclude that forward procedure by inference of R^2 incremental is an effective and efficient procedure for determining the best ANFIS model applied to time series forecasting.

References

[1] J. Jang., "Neuro-fuzzy modeling: architecture, analysis and applications," *Ph.D. Thesis, Department of Electrical Engineering and Computer Science, University of California, Berkeley, USA,* 1992.

[2] J. Jang, "Anfis: adaptive-network-based fuzzy inference system," *IEEE Transactions on System, Man, and Cybernetics,* vol. 23, no. 3, pp. 665-685, 1993.

[3] S. Kusumadewi and S. Hartati, Neuro-Fuzzy : Integrasi Sistem Fuzzy dan Jaringan Syaraf, Yogyakarta: Graha Ilmu, 2006.

[4] A. Fariza, A. Helen and A. Rayid, "Performansi Neuro Fuzzy Untuk Peramalan Data Time Series," in *Seminar Nasional Aplikasi Teknologi Informasi 2007,* Yogyakarta, 2007.

[5] A. Azadeh, S. Asadzadeh and A. Ghanbari, "An adaptive network-based fuzzy inference system for short-term natural gas demandestimation:Uncertainandcomplexenvironments," *Energy Policy,* vol. 38, pp. 1529-1536, 2010.

[6] Z. Tan, C. Quek and P. Cheng, "Stock trading with cycles: A financial application of ANFIS and reinforcement learning," *Expert Systems with Applications,* vol. 38, pp. 4741-4755, 2011.

[7] Wei. L.Y, C. T.L. and T. Ho, "A Hibryd Model Based on Adaptive-Network-based Fuzzy Inference System to Forecast Taiwan Stock Market," *Expert System with Applications,* vol. 38, pp. 13625-13631, 2011.

[8] R. Ebrahimpour, H. Nikoo, S. Masoudnia, M. R. Yousefi and M. S. Ghaemi, "Mixture of MLP-experts for trend forecasting of time series: A case study of the Tehran stock exchange," *International Journal of Forecasting,* vol. 27, pp. 804-816, 2011.

[9] F. K. Wang, K. K. Chang and C. W. Tzeng, "Using adaptive network-based fuzzy inference system to forecast automobile sales," *Expert Systems with Applications,* vol. 38, pp. 10587-10593, 2011.

[10] Y. Ho and C. Tsai, "Comparing ANFIS and SEM in linear and nonlinear forecasting of new product development performance," *Expert Systems with Applications*, vol. 38, pp. 6498-6507, 2011.
[11] M. E. Keskin and D. Taylan, "Artificial Models for Interbasin Flow Prediction in Southern Turkey," *Journal of Hydrologic Engineering*, vol. 14, pp. 752-758, 2009.
[12] A. Moghaddamnia, M. G. Gousheh, J. Piri, S. Amin and D. Han, " Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques," *Advances in Water Resources*, vol. 32, pp. 88-97, 2009.
[13] J. Jang, "Input Selection for ANFIS Learning," *Department of Computer Science, National Tsing Hua University Hsinchu, Taiwan.*
[14] J. Kaashoek and H. Van Dijk, " Neural Network Pruning Applied to Real Exchange Rate Analysis," *Journal of Forecasting*, vol. 21, pp. 559-577, 2002.
[15] Suhartono, "Feed Forward Neural Network untuk Pemodelan Runtun Waktu," *Ph.D Thesis, Department of Statistics, University of Gajah Mada, Indonesia*, 2007.
[16] P. Tahmasebi and A. Hezarkhani, "Application of Adaptive Neuro-Fuzzy Inference System for Grade Estimation; Case Study, Sarcheshmeh Porphyry Copper Deposit, Kerman, Iran," *Australian Journal of Basic and Applied Sciences*, vol.4, no.3, pp. 408-420, 2010.

Author Profile

Indah Puspitasari received S.Si degree in Statistics Department from Sepuluh Nopember Institute of Technology (ITS), Indonesia in 2012. She started her master study in the same department in 2011. Currently, she is at the last semester and completing her research paper.

Suhartono received his S.Si degree in Statistics department, ITS Indonesia and received his M.Sc in Statistical analysis and stochastic system, UMIST, UK. He finished his doctoral program in 2007 in Gajah Mada University, Indonesia. Currently, he is working as lecturer and head of postgraduate studies in Statistics Department, ITS.

Brodjol Sutijo S U received his bachelor degree in Statistics Department from Sepuluh Nopember Institute of Technology (ITS) and received his master degree in Statistics department from IPB, Indonesia. He finished his doctoral program in Gajah Mada University, Indonesia. Currently, He is a lecturer and also working as head of computational laboratory in Statistics Department, ITS.