International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 /doi:10.14445/22315381/IJETT-V68I12P220

Volume 68 Issue 12, 118-122, December 2020

The Developing of Fuzzy System for Multiple Time Series Forecasting with Generated Rule Bases and Optimized Consequence Part

Samingun Handoyo[#1], Ying-Ping Chen[*2]

[#]*Department of Statistics, Faculty of Mathematics and Natural Science, Brawijaya University, INDONESIA,*
[#]*EECS International Graduate Program, National Chiao Tung University, TAIWAN*
[*]*Department of Computer Science, National Chiao Tung University, TAIWAN*

[1]samistat@ub.ac.id, [2]ypchen@cs.nctu.edu.tw

Abstract *— The paper aims to build and implement a predictive model called a fuzzy system. The fuzzy rule bases component is generated by using the input-output data pairs. Its consequence part is optimized by using ordinary least squares. The initial structure model is needed to create the input-output data pairs based on the multiple time series. The rule bases are generated by using table lookup schema in which each input-output pairs has a contribution as a candidate rule. The obtaining rule base is modified to be an efficient one by optimizing its consequence part. As a case study is used, the 2-time series assumed which they have a causality effect. The data are the soybean price of both domestic and abroad. The developed fuzzy system is used in the forecasting of the domestic soybean price. The fuzzy system's performance is very satisfying, assessed according to the R-squared and mean squared error of criteria.*

Keywords — *fuzzy system, optimized rule bases, predictive model, times series forecasting.*

I. INTRODUCTION

Today, machine learning is a popular decision problem in any area ranging from engineering to social science. In general, machine learning has two types that are supervised learning and unsupervised learning. The predictive model has resulted from supervised learning. A predictive model is called a classification model if the output variable is categorical. On the other side, if the output variable is a numerical value, the predictive model is called the regression model [1-3]. A predictive model can support decision making to achieve both transparent and optimal decision [4].

There are many types of predictive models ranging from a simple model to complicated ones. Because decision-making aims to make the right decision, the model accuracy is a central issue. The phenomena motivate many researchers to develop the highest accuracy model, although it's a complicated one as a tradeoff. Some complicated models such as that are a hybrid between the wavelet decomposition and radial bases function, Kusdarwati and Handoyo [5] do neural network, the model performance comparing between logistic regression and support vector machine done by Widodo and Handoyo [6], and also Nugroho et al. [7] have compared the accuracy performance between logistic regression and linear vector quantization. The developed models are similar to a causality system that are predictor variables affecting the response variable.

In nature, based on some reasons, there are many systems considered a black box. The black box system output observed interval or period simultaneously will produce time-series data [8]. The time series modeling has grown vastly caused by forecasting from society, either an individual or corporate. Recently, modeling time series uses the heuristic approach in machine learning methods such as neural network and fuzzy logic have done by many researchers. The implementation of a fuzzy system with generated fuzzy rule bases by using empirical data has done by Handoyo and Marji [9], Handoyo et al. [10], Efendy et al. [11], Handoyo and Kusdarwati [12] and Kacimi, et al. [13] where their research focused on a single time-series data. In real life, the time series data can depend on the other one, such as the soybean domestics prices in Indonesia depend on the soybean aboard prices at the same as interval time of observed values.

The considering of other time series on the model building is a multivariate time series domain that is motivated to find a more realistic forecasting model [14-15]. Based on the description above, the research goals are Generating and implementing fuzzy rule bases. The consequence part is optimized on the multiple time series data in building a predictive model for regression purposes.

II. THE PROPOSED METHOD

The various kinds of time series data served by many sources have stimulated many researchers to build a model that involves some time series data. The autoregressive distributed lag (ARDL) model is built considering multiple time series that have causality of one series to another series [16]. The structure of an ARDL model's input and output is similar to the multiple regression model, but the predictor variables are time series of some lags sequence. Equation (1) presents the input and output pairs of an ARDL model.

$$A = \begin{bmatrix} x_{t-1} & x_{t-2} & x_{t-3} & y_{t-1} & \cdots & y_{t-3} & y_{t-4} \\ x_{t-2} & x_{t-3} & x_{t-2} & y_{t-2} & & y_{t-4} & y_{t-5} \\ & \vdots & & & \ddots & & \vdots \\ x_{t-p} & x_{t-(p-1)} & x_{t-(p-2)} & y_{t-p} & \cdots & y_{t-(p-2)} & y_{t-(p-3)} \end{bmatrix} \qquad \ldots\ldots\ldots\ldots..(1)$$

In equation (1), the matrix A is composed of 2 time series of Xt and Yt where both time series have a causality relationship, and the series Xt influences the series Y_t. The most important component of the fuzzy system is the fuzzy rule bases, which will be generated by the empirical data that have the format, such as in the equation (1).

Handoyo and Marji [10] have developed the fuzzy system with the case of the single time series. In the fuzzy system, the structure of the rule bases is given such as on the equation (2) as follows

$$R_1: IF\left(x_{1,1} = A_{1,1}\ and \ldots and\ x_{p,1} = A_{p,1} \right)$$
$$THEN\ (y_1 = \theta_{01})$$
$$R_2: IF\left(x_{1,2} = A_{1,2}\ and \ldots and\ x_{p,2} = A_{p,2} \right)$$
$$THEN\ (y_2 = \theta_{02})$$
$$R_q: IF\left(x_{1,n} = A_{1,n}\ and \ldots and\ x_{p,n} = A_{p,n} \right)$$
$$THEN\ (y_n = \theta_{0q}) \qquad \ldots\ldots\ldots\ldots(2)$$

The rule bases consist of q rules where the parts following "THEN" are called the consequence part that will be optimized by ordinary least square. [18].

Consider the equation(3) as the following

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix} = \begin{bmatrix} \bar{\alpha}_1^1 & \bar{\alpha}_1^2 & \bar{\alpha}_1^3 & \cdots & \bar{\alpha}_1^q \\ \bar{\alpha}_2^1 & \bar{\alpha}_2^2 & \bar{\alpha}_2^3 & & \bar{\alpha}_2^q \\ & \vdots & & \ddots & \vdots \\ \bar{\alpha}_n^1 & \bar{\alpha}_n^2 & \bar{\alpha}_n^3 & \cdots & \bar{\alpha}_n^q \end{bmatrix} * \begin{bmatrix} \theta_0^1 \\ \theta_0^2 \\ \vdots \\ \theta_0^q \end{bmatrix} \qquad \ldots\ldots\ldots\ldots(3)$$

Equation (3) represents Takagi Sugeno's fuzzy system that is expressed in the system linear equation. The system has q fuzzy rules in the fuzzy rule bases where the optimal consequence part is obtained by ordinary least square (OLS). The consequence part is denoted by θ_0 (θ_0^q is the optimal consequence part of the rule number q). The fire strength of rule number q with the first data input is stated by $\bar{\alpha}_1^q$ (the alpha-cut of the first input of rule number q), whereas $\hat{y}$ denoted the predicted value of the example number n.

In this work, the optimal consequence part (θ_0) is calculated bypassing the input of the training set to the fuzzy rule bases. The computation yields the alpha-cut matrix. By considering the target data as the response variable, such as on the equation (3), the linear system equation describes a fuzzy system that has unknown parameters of θ_0. Unfortunately, the alpha-cut matrix is a singular one causing the close form solution is not directly available. The singular value decomposition (SVD) is applied for obtaining the optimal θ_0 . Finally, the predicted values of the testing set are straightforwardly obtained, bypassing the input of the testing set to the system

III. THE DATA AND FUZZY SYSTEM DEVELOPMENT PROCESS

The section will be described the soybean prices graphically on both domestic and aboard as a case study on Indonesia's East Java province. It is also presented the fuzzy system development steps in sequential order briefly.

A. *Presenting data graphically*

The research uses the domestic and aboard of the daily soybean prices data. The case study Data were picked up from the "Perum BULOG" in the East Java regional division of Indonesia (from 2014 March to 2017 December). The time series plot of both data is presented in figure 1 as follows

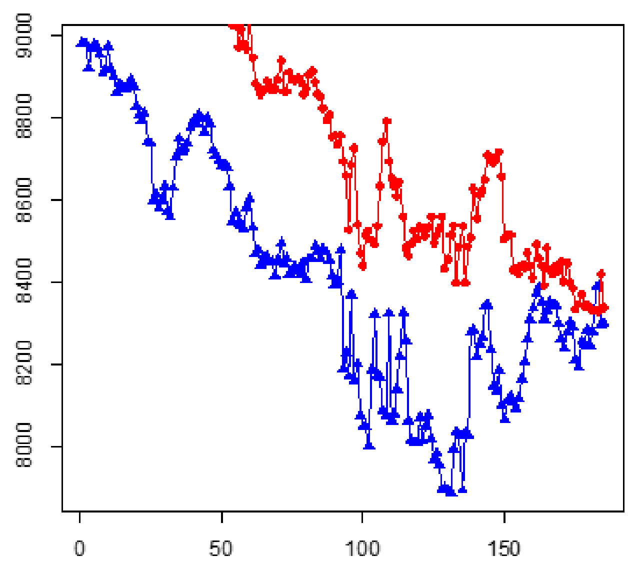

Fig. 1. Domestic and aboard of the soybean prices

In figure 1 above, both curves described the domestic soybean prices at the above part, and the lower part is the soybean aboard prices. Both soybean prices have a similar pattern, but there is a gap between them and some decreasing and increasing fluctuation gradually. The large gap fluctuating across time shows that both prices have significantly different prices.

B. The fuzzy system development steps

To assess the fuzzy system performance, the first analysis step must be determined the subset data as the training part and the other subset as the testing part. The developed fuzzy system only relates to the training part. The development steps are as follows

a. determine a membership function and number of fuzzy set covering the input domain
b. determine a structure model that describes the dependency between both time series
c. set the parameters of membership function
d. determine its membership degree of each observed value of the training part
e. arrange both time series based on the structure model as input-output data pairs
f. substitute the observed value with its associated fuzzy set to form a candidate rule
g. compute the degree of each candidate rule
h. use the table lookup schema to produce the fuzzy rule bases
i. optimize the consequence part of each rule using ordinary least square
j. run the fuzzy system on the training data
k. run the fuzzy system on the testing data
l. calculate the accuracy measure both of MAPE and R-squared.

IV. RESULTS AND DISCUSSION

This section will present the generating of fuzzy rule rules using the table lookup schema's training data. The estimation of the constant values as the consequent parts of each fuzzy rule is conducted by the OLS method. The implementation of the optimal fuzzy rule bases on both the training and testing data will be followed by the calculating system performance on both the MAPE and R-squared criteria.

A. Generating fuzzy rule candidates uses input-output data pairs

In the developing fuzzy system of the research, two parameters will be tuned. The first parameter that will be set equal to 3 is the relationship of input-output pairs as the structural model describes the dependency of output to inputs. The second parameter that will be set equal to 7 is the number of fuzzy sets used in the fuzzification phase, followed by associating the membership function type (the work will be used the Gauss MF or Gaussian membership function). Table 1 presents the Gaussian parameters (spread and mean) that must cover the training data set's domain value.

TABLE 1. The Gaussian parameters of both time series

	Domestic Price		**Aboard Price**	
Fuzzy Set	**Spread**	**mean**	**spread**	**Mean**
Small_3(S3)	91.92	8398	94.67	7885
Small_2(S2)	91.92	8581	94.67	8074
Small_1(S1)	91.92	8765	94.67	8263
Normal(N)	91.92	8949	94.67	8453
Big_1(B1)	91.92	9133	94.67	8642
Big_2(B2)	91.92	9317	94.67	8831
B9g_4(B3)	91.92	9501	94.67	9021

For simplicity purposes, the spread parameter is set to the same constant, and the mean parameter increases with factor 2 time to spread between two nearby fuzzy sets. Each observed value of training data is transformed with all of Gauss MF in Table 1. Let's consider an observed value (obs_1) with the highest degree of the B2, which means that the fuzzy set of the B2 will represent obs_1. The fuzzification resulted from the first 5 example points is presented in Table 2, where the fuzzy set notations of S3, S2, S1, N, B1, B2, B3 are replaced by the values of 1,2,3,4,5,6 and 7 respectively for the simple term.

TABLE 2. The first 5 example points after the fuzzification process

	Domestic		**Aboard**	
Obs_no	**Highest degree**	**Fuzzy set**	**Highest degree**	**Fuzzy set**
1	0.8823	7	1	7
2	1	7	1	7
3	0.6782	7	0.9186	7
4	0.7728	7	0.9226	7
5	0.9742	7	0.647	6

For example, on row number 5, it can be stated that the ov_5 (observation value number 5) is a member of fuzzy set B3 with its membership degree of 0.9742 for the domestic soybean prices, and the aboard soybean price is a member of fussy set B2 with its membership degree of 0.647. It also is noted that ov_5 means the observation is done at a time sequence of 5. Based on the determined structure model, which is the input function to produce an output, each input-output pair will be a rule candidate. For example, a structure model involves the X's lag of 3 and the Y's lag of 3. The training data is arranged as a table (matrix) data structure. Each matrix's row is an input and output pairs. The structured input and output will create the corresponding matrix of candidate rule bases, presented in table 3.

TABLE 3. the last 7 rows of the rule candidates of the fuzzy rule bases

CR_no	X3	X2	X1	Y3	Y2	Y1	Yt	degree
125	2	2	2	1	2	1	1	0.134
126	2	2	2	1	1	2	1	0.289
127	1	2	2	1	1	1	1	0.478
128	1	1	2	1	1	1	1	0.568
129	2	1	1	1	1	1	2	0.371
130	2	2	1	2	1	1	2	0.375
131	1	2	2	2	2	1	2	0.260

The total rule candidates depend on the number of the observed value and the number of lag involved in the model structure. Table 3 consists of 131 rule candidates generated from 134 observed values of training data with the 3 lags considered to affect the current value of Yt (the domestic soybean price at the time of t).

B. Optimizing consequence part of each rule

By using the table lookup schema for obtaining the fuzzy rule bases, it is selected as many as 66 rules as the member of the fuzzy rule bases presented in table 4 as follows:

TABLE 4. The first 7 fuzzy rules of the fuzzy rule bases

CR_no	X3	X2	X1	Y3	Y2	Y1	Yt	degree
1	6	6	6	6	6	6	6	0.820
2	3	3	3	4	4	4	4	0.695
3	7	7	7	7	7	7	7	0.631
4	1	1	2	1	1	1	1	0.568
5	4	4	4	4	4	4	4	0.516
6	1	2	2	1	1	1	1	0.478
7	6	6	7	6	6	6	6	0.453

Consider the first row of Table 4, and the fuzzy rule can be read as " IF (all of the inputs systems including X3, X2, X1, Y3, Y2, and Y1 have linguistic value notated by the digit of 6 that means the fuzzy set of B2 withstand of Big2) THEN the output (Yt) is the linguistic value of 6 ". It is necessary to note that the generated fuzzy rule bases are fuzzy rule bases of Mamdani's fuzzy system because the consequence part's values are a linguistic value. The next step is to assign the Gaussian parameter in Table 1 to the corresponding fuzzy set in the rule bases in Table 4. The process will transform each cell's elements in the rule bases to be the pairs of fuzzy sets and Gaussian parameters. The fuzzification of the rule bases is done by assigning each input data to the whole of the rule bases' fuzzy rules. For example, the first row of data input is conducted the fuzzification process on the first rule that will yield the membership degree on its corresponding fuzzy set. The number of membership degrees is 6 values because there are 6 predictor variables in the system input. Finally, each input data (a row of the matrix input-output pairs that the output variable is excluded) through the fuzzification process will be transformed into a matrix whose elements are membership degree with the dimension of the matrix of 61x6. Using logic operator (AND) on each rule, there is a value on each rule antecedent part called the fire strength or alpha-cut. Furthermore, the matrix becomes to be a vector with a length of 61.

When all of the training set's input data (131 rows) have been completed in the fuzzification processing on the rule bases, the result produced is a matrix with the dimension of 131x61. Each row of the last matrix has a corresponding fuzzification output to each row of the data input matrix on the fuzzy rule bases. The last matrix is mapped for the resulting output target y. Because each column of the last matrix is related to a rule's consequence in the fuzzy rule bases, its consequence value can be obtained by the OLS method. The training process's final output is a constant value due to each rule of the fuzzy rule bases.

C. Performance of fuzzy system on both of training and testing set

The fuzzy system's training process has completed when the optimized fuzzy rule bases component has been obtained. The output system will be produced when input data is entered into the fuzzy system. The input data will conduct the fuzzification process on each rule in the fuzzy rule bases. The antecedent part of each rule will produce a firing strength. The output of the fuzzy system is yielded by the dot product between the fire strength vector and the optimized consequence vector. The output system of the training set is presented in figure 2.

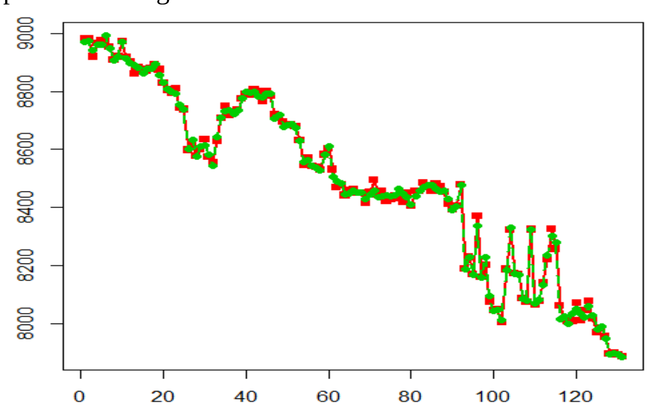

Fig. 2. Plot the actual and predicted values of the training data set

Figure 2 is the plot that describes how each sample point observed can be predicted almost perfectly by the output of the system (predicted values). Both curves are very close to each other, and it is difficult to capture the gap between the observed values and predicted values.

The accuracy measures as a system performance measure are the MAPE of 0.1155% and the R-squared of 99.8%. Both accuracy measures show that the system has very high performance on the training data.

On the other hand, Figure 3 describes how the observed values of testing data are predicted by the output system (predicted values). The gap between observed and predicted values is still easy to identify. Nevertheless, the pattern of both curves is very similar. The results give an insight that the system can forecast future values well.

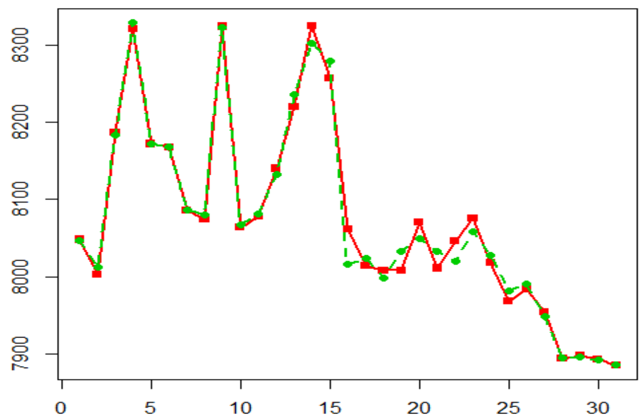

Fig. 3. Plot the actual and predictive values on the testing data

The system performance on the testing data is good enough, with the MAPE of 0.49% and the R^2 of 73%. Both accuracy measures show a significant difference with the system performance on the training data. The testing data's MAPE value is almost 4 times larger than the MAPE value of training data. The R^2 of testing data is about 27% lower than the R^2 of the training data. The results indicate that there is a possibility of overfitting on the fuzzy system training.

V. CONCLUSION

This study has implemented a fuzzy system that used Gaussian MF to consider the inputs of the 3 lags in both time series data and determine the fuzzy set number of 7 linguistic values. The table lookup schema's fuzzy rule bases consist of the 66 fuzzy rules that each of its consequence parts optimized using the ordinary least square method. Furthermore, the fuzzy system can be viewed as a linear model, so the system's output is produced by the dot product between the optimal consequence part and the fire strength of each rule. In fitting the training data, the system has performance measures of the MAPE = 0.1155% and R^2 = 99.8%. Nevertheless, the system performances on the testing data are 0.49% and 73% of the MAPE and $R^{2,}$ respectively

REFERENCES

[1] Z. Wei, et al. Large sample size, wide variant spectrum, and advanced machine-learning technique boost risk prediction for inflammatory bowel disease The American Journal of Human Genetics, 92(6) (2013) 1008-1012

[2] S. Regina, L. Suganthi, U K Sneha, S. Shwetha, Diabetic Retinopathy Classification Using Machine Learning Techniques International Journal of Engineering Trends and Technology 68(1) (2020) 51-56.

[3] J. Sadhasivam, A. Rathore, I. Bose, S. Bhattacharjee, S. Jayavel, A Survey of Machine Learning Algorithms International Journal of Engineering Trends and Technology 68(4)(2020) 64-71.

[4] I.N. Purwanto, A. Widodo, and S. Handoyo. System For Selection Starting Lineup Of A Football Players yy Using Analytical Hierarchy Process Journal of Theoretical & Applied Information Technology, , 97(1) (2018) 19-31.

[5] H. Kusdarwati and S. Handoyo. System for Prediction of Non-Stationary Time Series based on the Wavelet Radial Bases Function Neural Network Model Int J Elec & Comp Eng (IJECE), 8(4) (2018) 2327-2337.

[6] A. Widodo and S. Handoyo. The Classification Performance Using Logistic Regression and Support Vector Machine Journal of Theoretical & Applied Information Technology, 95(19) (2018) 5184-5193.

[7] W. H. Nugroho, S. Handoyo, and Y.J. Akri, An Influence of Measurement Scale of Predictor Variable on Logistic Regression Modeling and Learning Vector Quantization Modeling for Object Classification International Journal of Electrical and Computer Engineering. (IJECE), 8(1) (2018) 333-343

[8] J. Qiu, S.R. Jammalamadaka, and N. Ning, Multivariate time series analysis from a Bayesian machine learning perspective. Annals of Mathematics and Artificial Intelligence, 88(10) (2020) 1061-1082.

[9] S. Handoyo, and Marji, The Fuzzy Inference System with Least Square Optimization for Time Series Forecasting. Indonesian Journal of Electrical Engineering and Computer Science (IJEECS), 7(3) (2018) 1015-1026.

[10] S. Handoyo, Marji, I.N. Purwanto, and F. Jie, The Fuzzy Inference System with Rule Bases Generated using the Fuzzy C-Means to Predict Regional Minimum Wage in Indonesia International J. of Opers. And Quant. Management (IJOQM) , 24(4) (2018) 277-292.

[11] A. Efendi S. Handoyo A.P.S. Prasojo, and Marji, The Implementation Of The Optimal Rule Bases Generated By Hybrid Fuzzy C-Mean And Particle Swarm Optimization Journal of Theoretical & Applied Information Technology, 97(16) (2019) 4453-4453.

[12] S. Handoyo, and. H. Kusdarwati, Implementation of Fuzzy Inference System for Classification of Dengue Fever on the Malang villages in IOP Conference Series: Materials Science and Engineering, 546(5) (2019) 052026 IOP Publishing.

[13] M.A. Kacimi, O. Guenounou, L. Brikh, F. Yahiaoui, and N. Hadid, New mixed-coding PSO algorithm for a self-adaptive and automatic learning of Mamdani fuzzy rules. Engineering Applications of Artificial Intelligence, 2020, 89: 103417.

[14] H. Kusdarwati, and S. Handoyo, Modeling Threshold Liner in Transfer Function to Overcome Non Normality of the Errors In IOP Conference Series: Materials Science and Engineering, 546(5) (2019) 052039. IOP Publishing

[15] J. Zhao, S. Huang., Q. Huang, G. Leng, H. Wang., and P. Li, Watershed water-energy balance dynamics and their association with diverse influencing factors at multiple time scales. Science of The Total Environment, 711 (2020) 135189.

[16] S. Handoyo, The Multiple Time Series Modeling with Autoregressive Distributed Lag. Int. Conference on App. Analysis and Math. Modeling, ICAAMM 2019, Istambul-Turkey, (2020) 120-124.

[17] Lakshmana Phaneendra Maguluri, R. Ragupathy Comparative Analysis of Companies Stock Price Prediction Using Time Series Algorithm, International Journal of Engineering Trends and Technology 68(11) (2020) 9-15.

[18] S. Handoyo, A.P.S. Prasojo, Applied Fuzzy System with R Software (Sistem Fuzzy terapan dengan Software R). First Edition. Malang. UBPress. (2017) 81-135.