Intelligent Automation And Soft Computing, 2019
Copyright © 2019, TSI® Press
Vol. 25, no. 2, 319–327
https://doi.org/10.31209/2019.100000093

Development of a Data-Driven ANFIS Model by Using PSO-LSE Method for Nonlinear System Identification

Ching-Yi Chen and Yi-Jen Lin

Department of Information and Telecommunications Engineering, Ming Chuan University, Taoyuan, Taiwan, ROC

ABSTRACT

In this study, a systematic data-driven adaptive neuro-fuzzy inference system (ANFIS) modelling methodology is proposed. The new methodology employs an unsupervised competitive learning scheme to build an initial ANFIS structure from input-output data, and a high-performance PSO-LSE method is developed to improve the structure and to identify the consequent parameters of ANFIS model. This proposed modelling approach is evaluated using several nonlinear systems and is shown to outperform other modelling approaches. The experimental results demonstrate that our proposed approach is able to find the most suitable architecture with better results compared with other methods from the literature.

KEY WORDS: ANFIS, competitive learning scheme, PSO-LSE method.

1 INTRODUCTION

NONLINEAR system identification was a method to use input-output dataset to estimate system's nonlinear mathematical model (Sha & Bajic, 2013; Alci & Beyhan, 2016). Among many system identification methods, artificial neural network (ANN) was frequently used to simulate the dynamic relationship between output and input of nonlinear system, and the effect of its application in function approximation was pretty good (Hayashi & Buckley, 1994; Li & Chen, 2000), it not only had good fault-tolerant ability, but also had good generalization ability, however, its learned result was a black-box, therefore, it did not have ability of explanation to the user, and ANN's applications in many fields were limited in certain degree. Relative to ANN, fuzzy inference system (FIS) was another alternative that can be applied in system identification (Takagi & Sugeno, 1985; Eksin & Erol, 2000); through fuzzy inference, human's knowledge and experiences could be converted into fuzzy rules that are easy to understand. However, although the behavior of the use of fuzzy system to simulate nonlinear system had feature of linguistic information, yet it was lack of accurate quantitative analysis and learning ability of numerical value calibration.

Traditional fuzzy technique had to rely on expert's experience provided by expert to set up fuzzy system, however, ANN's ability to conduct learning and extract knowledge through training data had created complementary characteristic to fuzzy technique, and neuro-fuzzy systems was developed under such design concept. The integrated systems can combine the parallel computation and learning abilities of neural networks with the human-like knowledge representation and explanation abilities of fuzzy systems (Kaur, Sangal, & Kumar, 2017). Jang, in 1993, had associated two algorithms of fuzzy theory and ANN to propose an innovative architecture of adaptive network-based fuzzy inference system (ANFIS) (Jang, 1993). Generally speaking, ANFIS used back propagation (BP) algorithm and least square estimation (LSE) method to adjust the parameters of membership function so that it can fully exploit model's processing capability on system uncertainty and imprecision, and lots of literature had shown that ANFIS had good result on the identification of nonlinear system (Jang, 1993; Babuska, Verbruggen, 2003; Kaur, Sangal, & Kumar, 2017; Marzi, Darwish, & Helfawi, 2017).

The architecture design of ANFIS model can generally be divided into two stages of structure identification and parameters identification. In structure identification stage, theoretically, when more rules were used, it will be more helpful to construct a more complicated system, however, it will cause the increase of calculation amount at the same time.

CONTACT Ching-Yi Chen, chingyi@mail.mcu.edu.tw

Therefore, when we were constructing fuzzy system, the first thing needed to do was to summarize input spaces with similar outputs and to use appropriate fuzzy set to describe each input space with similar characteristic, therefore, it will be easy to use several fuzzy rules to construct a complicated nonlinear system. For ANFIS design method associating clustering technique, each cluster center obtained through cluster analysis can generate a fuzzy region, and can be mapped to a fuzzy rule in ANFIS architecture. Competitive learning algorithm was a self-organization learning method in its nature, and it can find from unlabeled samples for some similar features, rules or relationship, then these samples with similar features were gathered into the same class. In other words, competitive learning algorithm can find automatically from training data the inherent class rule, and the similarity of the processed data or its distribution state in high dimensional space can also be displayed; therefore, competitive learning was also very suitable to be applied in structure identification of ANFIS.

The parameter identification of traditional ANFIS has adopted ANN learning algorithm to obtain network parameters. Presently, all ANN designs almost adopted design procedure of design-evaluate-test cycle; in the structure design stage, it was needed to set up first parameters such as network structure, connection topology, transfer function or learning rate; in the evaluation stage, the learning example was conducted with simulation and evaluation; in the final test stage, unlearned data were used for the test. If the obtained results were not perfect, then the original design structure needed to be changed, that is, to start a new design cycle. Such design procedure meant that ANN designer was in an all possible network configuration space to use random method to search an optimized network configuration, and this also explained that ANN design can be seen as an optimization problem; therefore, many researches further proposed design method associating evolutionary computation to enhance the effectiveness of ANFIS (Juang, 2002; Ho et al., 2011; Chen, 2013; Fathzadeh et al., 2017; Marzi, Darwish, & Helfawi, 2017).

In this study, evolutionary ANFIS modeling approach was proposed based on competitive learning. First, competitive learning rule was used to conduct input space partitioning of FIS so as to explore effectively the clustering distribution state of training data, meanwhile, the obtained result was used in fulfilling coarse-level structure identification of ANFIS. After finishing structure identification of ANFIS, through hybrid learning scheme associating particle swarm optimization (PSO) and LSE method, premise parameters were finely tuned, and consequent parameters were learned, finally, parameter identification work of ANFIS model was finished.

2 ANFIS SYSTEM

THE proposed ANFIS model that has multiple inputs and single output is shown in Figure 1. It represents a Takagi-Sugeno-Kang (TSK)-type fuzzy system in a special five-layer feedforward network architecture.

Layer 1: This layer is called as the fuzzification layer, in which every node is an adaptive node with node function as

$$\mu_{A_{mk}}(x_{ik}) = \exp\left(-\frac{(x_{ik} - c_{mk})^2}{2\delta_{mk}^2}\right), \qquad (1)$$
$$m = 1,2,\ldots,r;\ k = 1,2,\ldots,n.$$

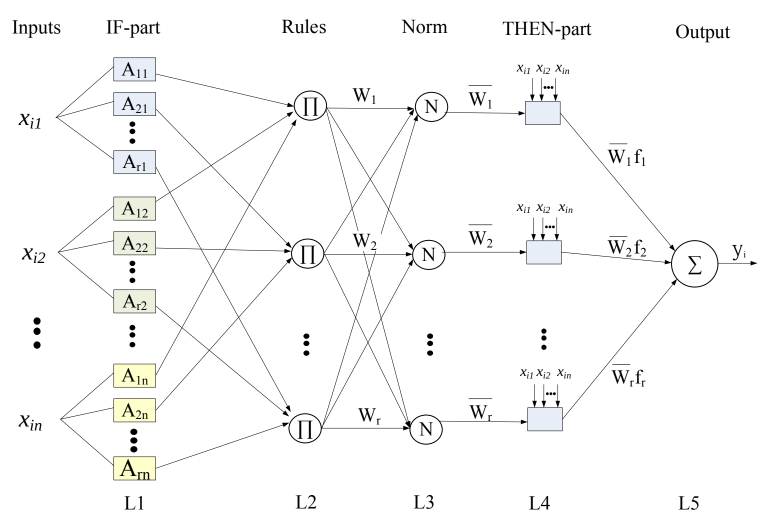

Figure 1. Structure of ANFIS.

Layer 2: Layer 2 is the production layer, in which every node is a fixed node with node function to compute the firing strength of each rule:

$$W_m = \prod_{k=1}^{n} \mu_{A_{mk}}(x_{ik}),\ m = 1,2,...,r. \qquad (2)$$

Layer 3: Each node in this layer calculates the normalized matching degree for each rule.

$$\overline{W_m}(x_i) = \frac{W_m(x_i)}{\sum_{m=1}^{r} W_j(x_i)} = \frac{\prod_{k=1}^{n} \mu_{A_{jk}}(x_{ik})}{\sum_{m=1}^{r}\prod_{k=1}^{n} \mu_{A_{mk}}(x_{ik})} = \frac{\prod_{k=1}^{n} \exp(-\frac{(x_{ik}-c_{mk})^2}{2\delta_{mk}^2})}{\sum_{m=1}^{r}\prod_{k=1}^{n} \exp(-\frac{(x_{ik}-c_{mk})^2}{2\delta_{mk}^2})} \qquad (3)$$

Layer 4: This layer known as the defuzzification layer. It calculates the conclusion inferred by each fuzzy rule.

$$f_m = b_{m0} + b_{m1}x_{i1} + ... + b_{mn}x_{in},\ m = 1,2,..,r \qquad (4)$$

Layer 5: This layer is known as the output layer. It has only one node and it calculates the overall output.

$$y_i = \sum_{m=1}^{r} \overline{W_m}(x_i) \cdot f_m \qquad (5)$$

The above ANFIS structure can be skilled to develop a TSK fuzzy inference system and determine membership functions for input and output variables of the system. For N input-output data of system to be identified x_i^{in}, i = 1,2,..,N, where $x_i^{in} = (x_{i1}, x_{i2}, ..., x_{in})$ is n dimensional data point, and y_i is its corresponding output, then the TSK fuzzy model consists of If-Then rules that has the following form:

Rule m: IF x_i^{in} is W_m then y is

$$y_i = b_{m0} + \sum_{k=1}^{n} b_{mk}x_k,\ m=1,2,\ldots,r. \qquad (6)$$

where r is the total number of fuzzy rules, b_{m0} and b_{mk} are the offsets and linear weights respectively, and W_m is defined by

$$\begin{aligned} W_m(x_i^{in}) &= \prod_{k=1}^{n} \exp(-\frac{(x_{ik}-c_{mk})^2}{2\delta_{mk}^2}) \\ &= \exp(-(\frac{(x_{i1}-c_{m1})^2}{\beta_{m1}^2} + ... + \frac{(x_{in}-c_{mn})^2}{\beta_{mn}^2})) \end{aligned} \qquad (7)$$

where $\beta_{mk} = \sqrt{2}\delta_{mk}$, W_m is a hyper-ellipsoid membership function belonging to the interval [0,1], c_{mk} and β_{mk} represent the center and the length of the kth principal axis for the related mth hyper-ellipsoid membership function, respectively. It is noted that c_{mk} and β_{mk} are both required parameters to be selected in the premised part of the fuzzy rules.

3 HYBRID TRAINING PROCESS OF ANFIS

IN this study, ANFIS model design was finished in two stages. In the first stage, self-organizing feature map (SOFM) neural network was used to partition the input space into appropriate fuzzy regions, meanwhile, these fuzzy regions were used to set up coarse ANFIS architecture that can roughly meet input and output data behavior; in the second stage, precise approximation to the behavior of input and output data was used as the goal, then learning algorithm was used to make a series of training on the premise parameters and consequent parameters of ANFIS, therefore, the designed ANFIS can approximate as much as possible system to be identified. It is as shown in Figure 2.

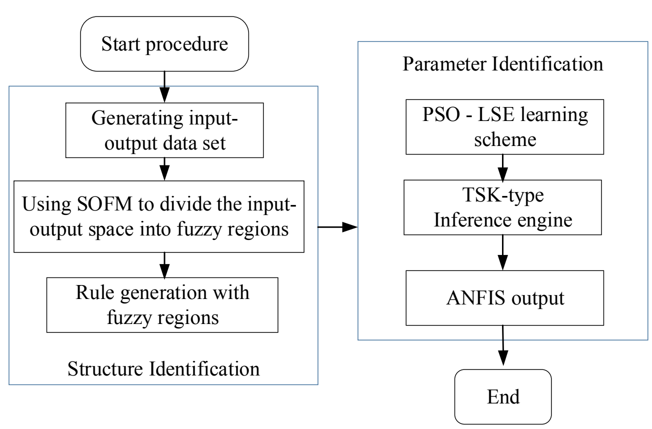

Figure 2. The training process of the proposed ANFIS model.

3.1 Structural Identification

3.1.1 Self-organizing feature map

SOFM is a neural network based on "competitive learning", that is, the neurons of output layer will compete with each other so as to get activated opportunity (Kohonen, 1989; Chen et al., 2014; Cao & Zhu, 2015). Generally, in competitive learning neural network, the winner will be selected from competitive phase, and the weight vector of the winner will be adjusted in the reward phase, which is as shown in equation (8)-(9).

$$W_j(t+1) = W_j(t) + \eta(x_i - W_j), \text{ If } j = j^* \quad (8)$$

where $j^* = \arg\min_j \|x_i - W_j\|, \; j = 1,2,...r \quad (9)$

However, the difference between SOFM and general competitive learning neural network is that in the competitive process, co-learning is adopted between the winning neurons and neurons which are the neighborhoods. In other words, in SOFM, after the competition, not only the winning neurons have the chance to learn, the neurons which are the neighborhoods can also have chance to learn. Figure 3 shows the structure of a SOFM network.

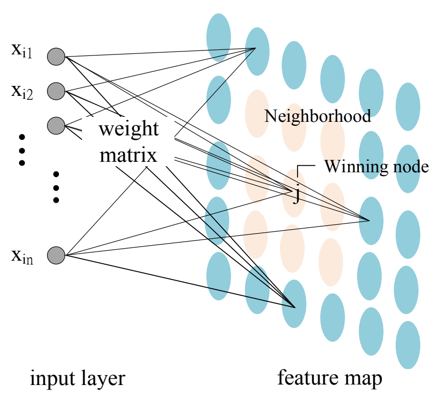

Figure 3. The architecture of a SOFM network.

In order to meet better the biological view point, the SOFM usually uses Gaussian function to decide the strength that SOFM neighborhood function is activated (Juang, 2002):

$$h_{j,j^*} = \exp\left(-\frac{d_{j,j^*}^2}{2\sigma^2}\right) \quad (10)$$

where d_{j,j^*} represents the side linking distance between jth neuron and winning neuron j^*. In addition, the effective width $\sigma(t)$ and learning parameter $\eta(t)$ of the neighborhood function is set up respectively as:

$$\eta(t) = \eta_0 \exp\left(-\frac{t}{\tau_1}\right), \; \sigma(t) = \sigma_0 \exp\left(-\frac{t}{\tau_2}\right) \quad (11)$$

where the constant σ_0 and learning parameter η_0 is set up by the initial value, and τ_1 and τ_2 are constants. Therefore, we can induce the improved SOFM algorithm as in the Table 1 (Chen et al., 2014):

Table 1. The improved SOFM algorithm.

1. *Select output layer network topology*
2. *Initialize all the connection weights to small random values*
3. *Repeat until convergence*
(1) *Select the next input vector x_i from the data set*
a. *Find the unit W_{j*} that best matches the input vector x_i*
$\|x_i - W_{j^*}\| = \min_j \|x_i - W_j\|, \; j = 1,2,..,r$
b. *Update the weights of the winner W_{j*} and all its neighbors W_D*
$W_D = W_D + \eta(t) \cdot h_{j^*D}(t) \cdot (x_i - W_D)$
(2) *Decrease the learning rate $\eta(t)$ and the neighborhood size $\sigma(t)$*

3.1.2 Using SOFM to realize the structural identification of ANFIS

In ANFIS modeling procedure, each output neuron of SOFM will be mapped to one fuzzy rule of ANFIS, and its weight vector will be used as central value c_{mk} of premise membership function of fuzzy rule. Variable δ_{mk} of eq.(3) was width of Gaussian function, in order to get appropriate δ_{mk}, we let

$$W_m(x_m^*) = \alpha \quad (12)$$

where $x_m^* = (x_{m1}^*, \ldots, x_{mn}^*)$ was input data in mth cluster that had the longest distance to cluster center c_m of mth cluster. Therefore, according to eq.(3) and eq.(12), δ_{mk} of membership function width of fuzzy set can be obtained as follows (Chen, 2000):

$$\delta_{mk} = \sqrt{-\frac{\sum_{k=1}^{n}(x_{mk}^* - c_{mk})^2}{2\ln(\alpha)}} \quad (13)$$

3.2 Parameter Identification

3.2.1 PSO basics

PSO is a population based stochastic optimization technique developed by Dr. Eberhart and Dr. Kennedy in 1995, inspired by social behavior of bird flocking or fish schooling. In PSO model, the flying of a particle is described by its velocity and position, the formula is described as follows (Kennedy & Eberhart, 1995):

$$V_{id} = \omega \cdot V_{id} + c_1 \times rand() \times (P_{id} - X_{id}) + c_2 \times rand() \times (P_{gd} - X_{id}) \quad (14)$$

$$X_{id} = X_{id} + V_{id} \quad (15)$$

where P_i and P_g represent the personal best position and the global best position, d is the dimensional number, i denotes the ith particle in the population, V is the velocity vector, X is the position vector, ω is the inertia factor, c_1 and c_2 are the cognitive and social learning rates, respectively.

3.2.2 PSO-LSE learning scheme for parameter identification

According to eq.(6) and eq.(7), it was clear that premise parameters that needed to be decided by ANFIS were decided by c_{mk} and δ_{mk}, and these parameters will be used to construct membership function W_m. In this study, PSO was used to make fine tuning on premise parameters obtained from SOFM, then accompanied with LSE method, consequent parameters $\{(b_{m0}, b_{mk})$, m=1,..,r; k=1,..,$n$$\}$ of ANFIS were obtained, therefore, the behavior of designed ANFIS model can approximate system to be identified. In PSO-LSE learning scheme, each particle of PSO represented a premise parameter set $\{(b_{m0}, b_{mk})$, m=1,..,r; k=1,..,$n$$\}$. Wherein c_{mk} obtained from SOFM and δ_{mk} obtained from eq.(13) will be defaulted as initial solution of P_g of PSO. For most ideal ANFIS model, its output value should be perfectly the same as desired output value, that is

$$\mathrm{SSE} = \sum_{i=1}^{N} (y_i - y_i^d)^2 = 0 \quad (16)$$

where y_i^d was the desired output value of input data x_i^{in}, and y_i was output value of ANFIS of the same input. Meanwhile, the output of ANFIS can be represented as

$$y_i = \frac{\sum_{m=1}^{r} W_m(x_i^{in}) \cdot f_m}{\sum_{m=1}^{r} W_m(x_i^{in})} = \sum_{m=1}^{r} g_{mi} \cdot f_m = \sum_{m=1}^{r} g_{mi} \cdot [b_{m0} + \sum_{j=1}^{n} b_{mj} x_{ij}] \quad (17)$$

where g_{mi} was normalized membership grade of mth rule on input data x_i^{in}, and its definition was

$$g_{mi} = \frac{W_m(x_i^{in})}{\sum_{j=1}^{r} W_j(x_i^{in})}, m = 1, 2, .., r \quad (18)$$

The most ideal situation of ANFIS was that its output and desired output was perfectly the same, that is

$$y_i^d = \sum_{m=1}^{r} g_{mi} \cdot [b_{m0} + \sum_{j=1}^{n} b_{mj} x_{ij}] = \sum_{m=1}^{r} [g_{mi} \cdot b_{m0} + g_{mi} \cdot \sum_{j=1}^{n} b_{mj} x_{ij}],\ i = 1,2,..., N \quad (19)$$

Therefore, eq.(19) can represented in the following matrix equation and is described by

$$\mathbf{Y} = \mathbf{WB} \quad (20)$$

where $\mathbf{Y} = [\mathrm{y}_1, \mathrm{y}_2, ..., \mathrm{y}_\mathrm{N}]^\mathrm{T}$, (21)

$$\mathbf{B} = [b_{10}, b_{11}, ..., b_{1n}, ..., b_{r0}, b_{r1}, ..., b_{rn}]^\mathrm{T} \quad (22)$$

$$\mathbf{W} = \begin{bmatrix} g_{11} & g_{11}x_{11} & \cdots & g_{11}x_{1n} & \cdots & g_{r1} & g_{r1}x_{11} & \cdots & g_{r1}x_{1n} \\ g_{12} & g_{12}x_{21} & \cdots & g_{12}x_{2n} & \cdots & g_{r2} & g_{r2}x_{21} & \cdots & g_{r2}x_{2n} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ g_{1N} & g_{1N}x_{N1} & \cdots & g_{1N}x_{Nn} & \cdots & g_{rN} & g_{rN}x_{N1} & \cdots & g_{rN}x_{Nn} \end{bmatrix} \quad (23)$$

where g_m is the normalized firing strength and N is the number of input-output data set.

Then, eq.(16) can be rewritten as

$$\mathrm{SSE} = \|\mathbf{Y} - \mathbf{WB}\|^2 = [\mathbf{Y} - \mathbf{WB}]^\mathrm{T}[\mathbf{Y} - \mathbf{WB}] \quad (24)$$

If the parameters of premise part are predetermined, the only unknown component in SSE is the consequent parameter vector **B** whose elements are the parameters in the linear regression equations of the ANFIS model. We can use the LSE method to solve the parameter vector, whose solution can be obtained by setting the derivative of this eq.(24) with respect to **B** to zero (Yoo, Bang, & Lee, 2004; Angelov, & Filev, 2004):

$$\frac{\partial}{\partial \mathbf{B}}(\text{SSE}) = -2\mathbf{W}^T[\mathbf{Y} - \mathbf{WB}] = 0 \tag{25}$$

Then the consequent parameter vector **B** minimizing SSE could be obtained by the pseudo-inversion:

$$\mathbf{B} = [\mathbf{W}^T\mathbf{W}]^{-1}\mathbf{W}^T\mathbf{Y} \tag{26}$$

After the LSE learning stage is completed and the consequent parameters obtained, the mean square error (*MSE*) between the ANFIS model and the desired outputs is determined to evaluate the efficiency of the ANFIS model in the parameters identification stage. The *MSE* is calculated as follows:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i^d - \frac{\sum_{m=1}^{r} W_m(x_i^{in}) \cdot f_m}{\sum_{m=1}^{r} W_m(x_i^{in})} \right)^2 \tag{27}$$

where N is the total numbers of input-output pairs, y_i^d denotes the desired output of the ith input data x_i^{in}.

In PSO-LSE learning scheme, the presented fitness function is exp(-*MSE*/5); thus, the PSO-LSE is determined to approach the maximal fitness value. The proposed PSO-LSE learning scheme is applied based on the fitness function's direction to select the optimal parameter set from the ANFIS models to minimize the *MSE*.

4 EXPERIMENTAL RESULTS AND DISCUSSION

IN the following subsections, we applied the proposed method to three kinds of problems: Box–Jenkins model identification (Oh & Pedrycz, 2000; Park, Pedrycz, & Oh, 2001; Rezaee & Zarandi, 2010), Chaotic Mackey–Glass time series prediction problem (Oh, Pedrycz, & Park, 2007; Choi, Oh, & Pedrycz, 2008), and daily Taiwan stock indexes (TAIEX) dataset. We use the standard performance index of the *MSE* and root mean square error (*RMSE*) as expressed by (27) and (28):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i^d - \frac{\sum_{m=1}^{r} W_m(x_i^{in}) \cdot f_m}{\sum_{m=1}^{r} W_m(x_i^{in})} \right)^2} \tag{28}$$

4.1 Box–Jenkins gas furnace data

In this subsection, Box-Jenkins gas furnace data with 296 I/O measurements data is used (Oh & Pedrycz, 2000; Park, Pedrycz, & Oh, 2001; Rezaee & Zarandi, 2010). The input measurement $u(t)$ is gas flow rate into the furnace and the output measurement $y(t)$ is CO_2 concentration in outlet gas. In order to directly compare our method with other approaches in Oh et al., (2000), Park et al., (2001), and Rezaee et al., (2010), $u(t\text{-}3)$ and $y(t\text{-}1)$ are selected as input variables to the proposed ANFIS model. The first portion of the dataset (148 pairs) is used for training, and the remaining part of the time series serves as a testing dataset. Figure 4 shows the comparison of the actual output and the output produced by our model with 6 rules. Table 2 compare out results with other models on this dataset.

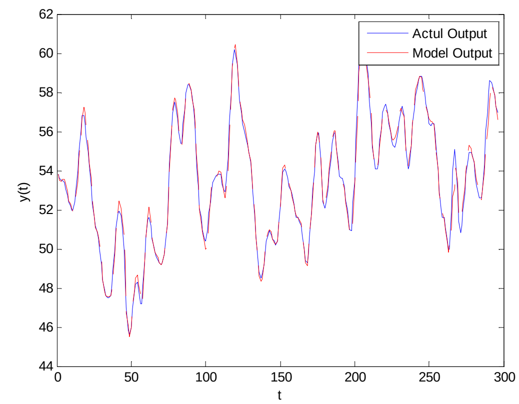

Figure 4. Desired and obtained output for the Box–Jenkins gas furnace data.

Table 2. Comparison results for the Box–Jenkins gas furnace data.

Model	No. of rules	*MSE*
Oh & Pedrycz's model (2000)	6	0.364
Park, Pedrycz, & Oh's model (2001)	6	0.333
Rezaee & Zarandi's model (2010)	6	0.2741
Our model	6	0.1150

4.2 Chaotic Mackey–Glass time series

Chaotic Mackey-Glass time series is a differential delay equation defined as follows (Oh, Pedrycz, & Park, 2007; Choi, Oh, & Pedrycz, 2008):

$$\dot{x}(t) = \frac{0.2 \cdot x(t-\tau)}{1 + x^{10}(t-\tau)} - 0.1 \cdot x(t) \tag{29}$$

This is a non-periodic and non-convergent time series that is very sensitive to initial conditions. In this experiment, we extracted 1000 input-output data pairs from the Mackey-Glass time series with t in [118-1117]. The first 500 data pairs were used in the training phase, whereas the last 500 in the testing. In order to compare out method with other approaches, we perform two experimental cased. In case 1, the proposed method has been used to design a ANFIS model with four inputs: $x(t\text{-}30)$, $x(t\text{-}18)$, $x(t\text{-}12)$, $x(t)$.

In case 2, $x(t\text{-}30)$, $x(t\text{-}12)$, $x(t)$ are selected as input variables. Figure 5 shows the comparison of the actual output and the output produced by the model with 9 rules for case 2. The root mean square error (*RMSE*) comparison with other modeling methods is shown in Table 3.

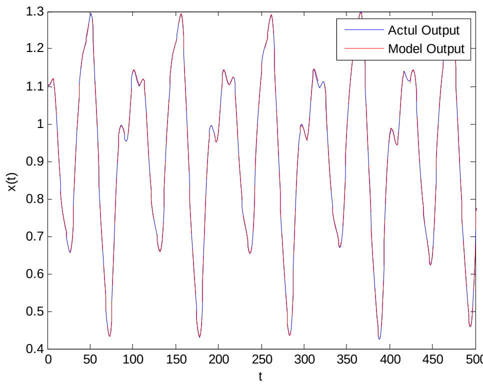

Figure 5. Desired and obtained output for the Chaotic Mackey-Glass time series.

Table 3. Comparison results for the Chaotic Mackey-Glass time series.

Model	No. of rules	*RMSE*	Input variables
Oh et al.'s model (2007)	20	0.00026	$x(t\text{-}30)$, $x(t\text{-}18)$, $x(t\text{-}12)$, $x(t)$
Our model	20	1.3828e-006	$x(t\text{-}30)$, $x(t\text{-}18)$, $x(t\text{-}12)$, $x(t)$
Choi et al.'s model (2008)	9	0.00311	$x(t\text{-}30)$, $x(t\text{-}12)$, $x(t)$
Our model	9	0.0025	$x(t\text{-}30)$, $x(t\text{-}12)$, $x(t)$

4.3 Forecasting for TAIEX

In this experiment, a five-year period of the daily Taiwan stock indexes (TAIEX) dataset from 2000/1/4 to 2004/12/31 is selected to test the forecasting performance of the proposed ANFIS model. In this simulation of the daily curve, the active period of training data ranges from 2000/1/4 to 2002/12/31, and the testing phase ranges from 2003/1/4 to 2004/12/31. Figure 6(a) and Figure 6(b) show the simulation results in training and testing periods for the TAIEX stock index based on the proposed ANFIS model, respectively. Evaluated performance comparison with the other modeling methods is illustrated in Table 4. Comparison results with the other modeling methods indicate that the proposed model can obtain the smallest *RMSE* value to approximate to the stock index curve of TAIEX in the testing phase.

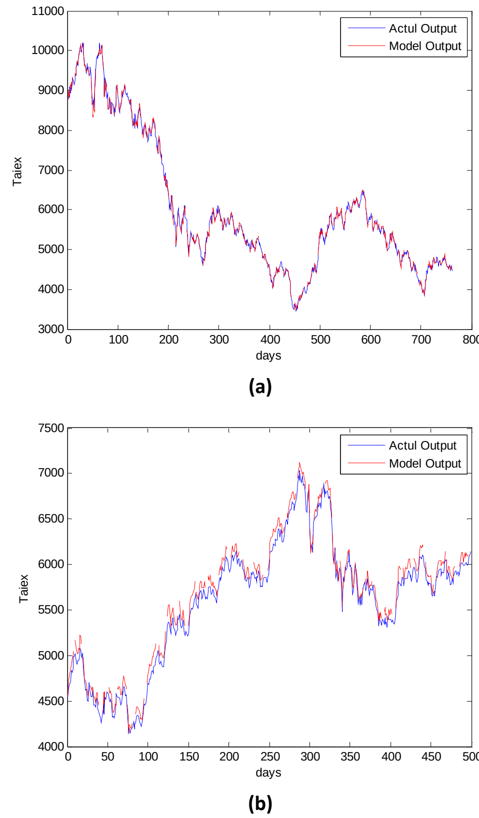

Figure 6. (a) Simulation results of the training phase for TAIEX dataset. (b) Forecasting results of the testing phase for TAIEX dataset.

Table 4. Comparison results for the TAIEX dataset.

Model	No. of rules	*RMSE*
Chen & Hwang's model (2000)	12	119
Huarng & Yu's model (2005)	12	187
Chu et al.'s model (2009)	12	84
Our model	12	66.82

5 CONCLUSION

IN this paper, an ANFIS modeling method using SOFM to conduct structure identification and using hybrid PSO-LSE learning scheme to conduct parameter identification was introduced. SOFM was one type of competitive learning algorithm, and it can, through feature mapping method, convert high dimensional vector, through nonlinear projection method, into low dimensional output vector space, meanwhile, clustering rule, distribution state and similarity could be learned from the training data; when SOFM was applied in structure identification of ANFIS model, input space can be partitioned into many nonhomogeneous fuzzy regions effectively to generate meaningful fuzzy rule to create quickly ANFIS structure that can coarsely meet the behavior of the system to be identified. In the learning stage of parameter identification, PSO algorithm will be used to make fine tuning the premise parameters of ANFIS model, meanwhile, it can associate LSE method to learn the coefficient values of consequent linear equation, therefore, the constructed model can meet the behavior of system to be identified more precisely. From the real test result, it can be seen that the modeling method proposed in this paper can, under the condition that only input-output data was known, make effective automatic extraction of FIS parameters to complete the system modeling work.

6 REFERENCES

M. Alci and S. Beyhan, (2016). Fuzzy functions with function expansion model for nonlinear system identification. *Intelligent Automation & Soft Computing*, 23(1), 87-94.

P.P. Angelov and D.P. Filev, (2004). An Approach to Online Identification of Takagi-Sugeno Fuzzy Models. *IEEE Trans. on System, Man, and Cybernetics- Part B*, 34(1), 484-498.

R. Babuska and H. Verbruggen, (2003). Neuro-fuzzy methods for nonlinear system identification. *Annual Reviews in Control,* 27, 73–85.

X. Cao and D. Zhu, (2015). Multi-AUV task assignment and path planning with ocean current based on biological inspired self-organizing map and velocity synthesis algorithm. *Intelligent Automation & Soft Computing*, 23(1), 31-39.

C.Y. Chen, J.S. Chiang, K.Y. Chen, T.K. Liu, and C.C. Wong, (2014). An Approach for Fuzzy Modeling Based on Self-Organizing Feature Maps Neural Network. *Applied Mathematics & Information Sciences.* 8(3), 1207-1215.

C.C. Chen, (2000). *Design of Fuzzy Systems based on Partitioning Input Spaces*. Ph. D. Thesis, University of Tamkang.

M.Y. Chen, (2013). A hybrid ANFIS model for business failure prediction utilizing particle swarm optimization and subtractive clustering. *Inform Sci*, (220), 180-195.

S.M. Chen and J.R. Hwang, (2000). Temperature Prediction Using Fuzzy Time Series. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics,* 30(2), 263-275.

J.N. Choi, S.K. Oh, and W. Pedrycz, (2008). Identification of fuzzy models using a successive tuning method with a variant identification ratio. *Fuzzy Set*, 159, 2873-2889.

H.H. Chu, T.L. Chen, C.H. Cheng, and C.C. Huang, (2009). Fuzzy dual-factor time-series for stock index forecasting. *Expert Systems with Applications*, 36(1), 165-171.

I. Eksin and O.K. Erol, (2000). A Fuzzy Identification Method for Nonlinear Systems. *Turk J Elec Engin*. 8(2), 125-135.

A. Fathzadeh, A. Jaydari, and R. Taghizadeh-Mehrjardi, (2016). Comparison of different methods for reconstruction of instantaneous peak flow data. *Intelligent Automation & Soft Computing*, 23(1), 41-49.

Y. Hayashi and J.J. Buckley, (1994). Approximation between fuzzy expert systems and neural networks. *Int J Approx Reason*, 10(1), 63-73.

W.H. Ho, J.X. Chen, I. Lee, and H.C. Su (2011). An ANFIS-based model for predicting adequacy of vancomycin regimen using improved genetic algorithm. *Expert Syst Appl*, 38, 13050-13056.

K. Huarng and H.K. Yu, (2005). A Type 2 Fuzzy Time Series Model for Stock Index Forecasting. *Physica A: Statistical Mechanics and its Applications*, 353, 445-462.

R. Jang, (1993). ANFIS: Adaptive network-based fuzzy inference system. *IEEE Trans. Syst., Man, Cybern.*, 23(3), 665-685.

C.F. Juang, (2002). A TSK-type recurrent fuzzy network for dynamic systems processing by neural network and genetic algorithms. *IEEE Trans. Fuzzy Syst.*, 10(2), 155–170.

R. Kaur, A. L. Sangal, and K. Kumar, (2017). Modeling and simulation of adaptive Neuro-fuzzy based intelligent system for predictive stabilization in structured overlay networks. *Eng. Sci. Technol. Int. J.*,, 20(1), 310-320.

J. Kennedy and R. Eberhart, (1995). Particle Swarm Optimization. *Proceedings of IEEE International Conference on Neural Networks*, Perth, Australia, 1942-1948.

T. Kohonen, (1989). *Self-Organization and Associative Memory*. Springer, Berlin.

X.H. Li and C.L.P. Chen, (2000). The equivalence between fuzzy logic systems and feed forward neural networks. *IEEE Trans Neural Netw*, 11(2), 356-365.

H. Marzi, A.H. Darwish, and H. Helfawi, (2017). Training ANFIS Using the Enhanced Bees Algorithm and Least Squares Estimation. *Intelligent Automation & Soft Computing*, 23(2), 227-234.

S.K. Oh and W. Pedrycz, (2000). Identification of fuzzy systems by means of an auto-tuning algorithm and its application to nonlinear systems. *Fuzzy Sets and Systems,* 115(2), 205–230.

S.K. Oh, W. Pedrycz, and K.J. Park, (2007). Structural developments of fuzzy systems with the aim of information granulation. *Simul. Model. Pract. Theor,* 15, 1292–1309.

B.J. Park, W. Pedrycz, and S.K. Oh, (2001). Identification of fuzzy models with the aid of evolutionary data granulation. *IEE Proc-Control Theory Appl*, 148(5), 406–418.

B. Rezaee and M.F. Zarandi, (2010). Data-driven fuzzy modeling for Takagi–Sugeno–Kang fuzzy system," *Inform Sci*, 180(2), 241-255.

D. Sha and V.B. Bajic, (2013). An Optimized Recursive Learning Algorithm for Three-Layer Feedforward Neural Networks For Mimo Nonlinear System Identifications. *Intelligent Automation & Soft Computing*, 17(2), 133-147.

T. Takagi and M. Sugeno, (1985). Fuzzy identification of systems and its application to modeling and control. *IEEE Trans. Syst. Man Cybern.*, 15(1), 116-132.

C.K. Yoo, Y.H. Bang, I.B. Lee, P.A. Vanrolleghem, and R. Rosen, (2004). Application of fuzzy partial least squares (FPLS) modeling nonlinear biological processes. *Korean Journal of Chemical Engineering*, 21(6), 1087-1097.

NOTES ON CONTRIBUTORS

Ching-Yi Chen received his Ph.D. degree in Electrical Engineering from Tamkang University, New Taipei City, Taiwan, R.O.C., in 2006. He joined the Department of Information and Telecommunications Engineering, Ming Chuan University in 2007 and is now an Associate Professor. His main research interests include swarm intelligence, pattern recognition, and embedded systems.

Yi-Jen Lin received the B.S. and M.S. degrees in Information and Telecommunications Engineering from Ming Chuan University, Taiwan, R.O.C., in 2010 and 2013, respectively. His research interests include data clustering and fuzzy modeling.