

Predictive Model of Energy Consumption in Beer Production

Tiecheng Pu and Jing Bai*

College of Electrical and Information Engineering, Beihua University, Jilin, Jilin, China

Received 15 May 2013; Accepted 25 July 2013

Abstract

The predictive model of energy consumption is presented based on subtractive clustering and Adaptive-Network-Based Fuzzy Inference System (for short ANFIS) in the beer production. Using the subtractive clustering on the historical data of energy consumption, the limit of artificial experience is conquered while confirming the number of fuzzy rules. The parameters of the fuzzy inference system are acquired by the structure of adaptive network and hybrid on-line learning algorithm. The method can predict and guide the energy consumption of the factual production process. The reducing consumption scheme is provided based on the actual situation of the enterprise. Finally, using concrete examples verified the feasibility of this method comparing with the Radial Basis Functions (for short RBF) neural network predictive model.

Keywords: ANFIS, subtractive clustering, RBF, predictive model, energy consumption

1. Introduction

The beer production is a multi-working procedure, strong correlation and unequal length delay process. Generally, there are raw material crushing, saccharification, gelatinization, fermentation and filling processes, each of processes consumes water, power and steam etc, but also the consumed energy mutual influences among the production. For example, in order to saving the power in raw material crushing process, the crushing time can be shorten, then the raw material can not be crushed fully, this leads to increase energy consumption in the saccharification and gelatinization processes. In addition, every working time of process is not consistent. For example, the working time of saccharification and gelatinization processes is several hours, but the working time of fermentation process is more than ten days. How to predict the complex energy consumption has been the urgently solved practical problem.

The energy consumption and the influence factors have nonlinear uncertainty relation in the beer production process, so the prediction model cannot be built based on the physical principle. The prediction model^[1] based on the data is the solution to save energy. The complete data model cannot deal with the semantic items which exist in the qualitative influence factors. Neuro-fuzzy inference systems bring the low-level learning and computational power of neural networks into fuzzy systems and bring the high-level human-like thinking and reasoning of fuzzy systems to neural networks^[2]. The key problem in the fuzzy reasoning system is to establish the rules of model. We should face the difficulties how to establish the rules without the artificial experience and how to determine the premise parameters in

modeling the predictive model. This paper studies the prediction model of energy consumption in beer production based on the subtraction clustering and adaptive neural fuzzy inference system.

The remainder of this study is organized as follows. Section 2 presents the subtractive clustering. Section 3 describes the structure and learning algorithm of Adaptive-Network-Based Fuzzy Inference System. Section 4 describes the Radial Basis Functions neural network. Section 5 presents the simulation for concrete examples, which verifies the feasibility of this method. Section 6 draws the conclusions.

2. Subtractive Clustering

The subtractive clustering^[3] is a fast algorithm in estimating the number of cluster and cluster center during a set of data. There are normalized data (x_1, x_2, \dots, x_n) , every data can probably be the cluster center, so the density^[4] at the data point x_i is

$$D_i = \sum_{j=1}^n \exp(-a\|x_i - x_j\|^2) \quad (1)$$

We choose the data point with maximal density as the first cluster center after computing the density at every data point. Set x_{c1} as the chosen cluster center and D_{c1} as the density. Every density of data point can be amended by

$$D_i = D_i - D_{c1} \exp(-\beta\|x_i - x_{c1}\|^2) \quad (2)$$

* E-mail address: jilinbaijing@163.com

We choose the next cluster center x_{c2} after amending the every density of data point, and then amend the all density of data point. The process can be repeated until we get the enough cluster centers.

3. Adaptive-Network-Based Fuzzy Inference System

ANFIS takes the neural network technique into the fuzzy inference system [5]. It can conquer the dependence on operator in the traditional fuzzy logical inference system and the intervention by human in designing the membership function. ANFIS can automatically update the membership function based on the built fuzzy rules and receive the optimal parameters which meet the input-output relation of fuzzy inference system [6].

3.1 Structure of ANFIS

Scholars Takagi and Sugeno put forward T-S fuzzy model in 1985, for short Sugeno fuzzy system model. It is a commonly used nonlinear fuzzy reasoning model and suits for expressing the dynamic characteristics of multivariable complex system. The typical rule of inference is

$$IF\ x\ is\ A\ and\ y\ is\ B\ THEN\ z = f(x, y) \tag{3}$$

Where A and B are the premise fuzzy numbers, $z = f(x, y)$ is the conclusion accurate number and usually is an n orders polynomial about X and Y . The model is called n orders Sugeno fuzzy inference system.

Scholar Jyh-Shing Roger Jang put forward the adaptive neural network fuzzy inference system [7] in 1993. Its function is similar to the Sugeno model. The structure of two inputs and one output ANFIS with two rules is shown in figure 1.

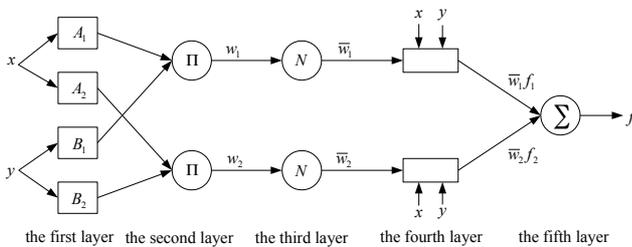


Fig. 1. Structure Diagram of Two Inputs and One Output ANFIS with Two rules

We compare these neurons in the structure with that in the conventional neural network. The same is that the neurons not only express the signal but also include operation. The difference is that there are not connection weights and thresholds among the neurons in the structure. And the parameters of block nodes need be identified. The parameters of circular nodes need not be identified. The output of j th node in i th layer is $O_{i,j}$. The nodes in the same layer have the same node function. The first layer, it is the input layer. The membership function value of fuzzy set A is $O_{1,j}$. The fuzzy set A includes A_1, A_2, B_1 and B_2 . The membership function value determines the meet degree to A for the given input X or Y , which often is represented by Bell function or Gauss function.

$$O_{1,j} = u_{A_j}(X) \quad (j=1,2) \text{ and } O_{1,j} = u_{B_{j-2}}(X) \quad (j=3,4) \tag{4}$$

$$u_{A_j}(X) = \left\{ 1 + \left(\frac{X - c_j}{a_j} \right)^{2b_j} \right\}^{-1} \text{ or } u_{A_j}(X) = \exp \left[- \left(\frac{X - c_j}{a_j} \right)^2 \right] \tag{5}$$

Where parameters a_i, b_j relate to the curve width, c_j expresses the centre of curvature of the curve. The membership function parameters in the layer are called the premise parameters. We can calculate the fuzzy membership degree values of input variables using the premise parameters.

The second layer, each of nodes in the layer is represented using the symbol Π . Multiply the input signal, the results is the layer output as follows in formula (6). The multiplication in this layer is sure any AND operator which conforms to T Standard. We can use such calculations to determine the incentive intensity of each fuzzy rule.

$$O_{2,j} = w_j = u_{A_j}(X) \times u_{B_j}(Y) \quad (j=1,2) \tag{6}$$

The third layer, each of nodes in the layer is represented using the symbol N . w_j , which is the incentive intensity of j th node to the j th rules, divides the total incentive intensity as follows in formula (7). Thus the output of the layer is called normalized incentive intensity.

$$O_{3,j} = \bar{w}_j = \frac{w_j}{w_1 + w_2} \quad (j=1, 2) \tag{7}$$

The fourth layer, each of nodes in the layer has the adaptive node function. The output is shown as follows in formula (8).

$$O_{4,j} = \bar{w}_j f_j = \bar{w}_j (p_j x + q_j y + r_j) \quad (j=1,2) \tag{8}$$

Where the normalized incentive intensity, \bar{w}_j , comes from the third layer. The linear system parameters of the Takagi-Sugeno fuzzy model are p_j, q_j and r_j . The parameters in the layer are called the consequent parameters of ANFIS.

The fifth layer, the single node in the layer is a fixed node. The total output is the sum of all input signals as follows in formula (9).

$$O_{5,j} = f = \sum_j \bar{w}_j f_j = \frac{\sum_j w_j f_j}{\sum_j w_j} = output \tag{9}$$

Therefore, for ANFIS structure, the total output is the linear combination of consequent parameters after determining the initial membership functions (or the premise parameters). The mathematically expression is

$$\begin{aligned}
 f &= \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 = \bar{w}_1 f_1 + \bar{w}_2 f_2 \\
 &= (\bar{w}_1 x) p_1 + (\bar{w}_1 y) q_1 + \bar{w}_1 r_1 + \\
 &\quad (\bar{w}_2 x) p_2 + (\bar{w}_2 y) q_2 + \bar{w}_2 r_2
 \end{aligned}
 \tag{10}$$

3.2 Learning Algorithm of ANFIS

In order to reduce the training time of the network, the recursive least squares estimation algorithm is added based on the original gradient descent method. In the forward channel, the output of each node can be forward transmitted, and the conclusion parameters can be identified using least squares estimation. At the same time in the reverse channel, the premise parameters can be modified using the gradient descent algorithm. When the premise parameters are determined the consequent parameters are optimal. The hybrid learning algorithm can reduce the search space dimension of error back propagation and training time. Firstly, the consequence parameters are on-line identified using the recursive least squares.

$$\begin{aligned}
 K(t+1) &= \frac{P(t)\varphi'(t+1)}{1 + \varphi(t+1)P(t)\varphi'(t+1)} \\
 \hat{\theta}(t+1) &= \hat{\theta}(t) + K(t+1)(z(t+1) - \varphi(t+1)\hat{\theta}(t)) \\
 P(t+1) &= P(t) - K(t+1)\varphi(t+1)P(t)
 \end{aligned}$$

Where, $\theta = (p_1, q_1, r_1, \dots, p_l, q_l, r_l)$ is the vector of consequence parameters, and $\varphi = (\bar{w}_1 x, \bar{w}_1 y, \bar{w}_1, \dots, \bar{w}_l x, \bar{w}_l y, \bar{w}_l)$ is the vector of inputs and weights. Z is the actual output data. The number of rules is l . $\hat{\theta}(0) = 0I_{3 \times 3l}$, $P(0) = \gamma I_{3 \times 3l}$, γ is a large enough positive integer. Then, the premise parameters are modified using the gradient descent algorithm [8]. Assuming there are L layers in the adaptive network, there are $\#(L)$ nodes in the L th layer. The sum of square error of the p training sample is defined

$$E^p = \sum_{i=1}^{\#(L)} (T_i^p - O_{L,i}^p)^2 \tag{11}$$

Where for the p th training sample, T_i^p is i th actual value, $O_{L,i}^p$ is the output of i th node in the L th layer. The error rate is derived by the chain rule.

$$\frac{\partial E^p}{\partial O_{L-1,i}^p} = \sum_{i=1}^{\#(L)} \frac{\partial E^p}{\partial O_{L,i}^p} \frac{\partial O_{L,i}^p}{\partial O_{L-1,i}^p} \tag{12}$$

The error rate of internal nodes can be described as the linear combination of error rate of the next layer corresponding to the nodes.

$$\frac{\partial E^p}{\partial \alpha} = \sum_{\sigma \in \Omega} \frac{\partial E^p}{\partial O^*} \frac{\partial O^*}{\partial \alpha} \tag{13}$$

Where α are the parameters of the given membership function, Ω is node set, the output of which is relevant to α . In summary, the modified formula of parameter α and learning rate η is shown as follows

$$\Delta \alpha = -\eta \frac{\partial E^p}{\partial \alpha} \tag{14}$$

4. Radial Basis Functions Neural Network

Application of artificial neural networks for prediction has been recently proposed by a number of investigators [9]. RBF neural network includes simpler structure and faster learning algorithms. It is a feed forward neural network model and has already proved its universal approximation ability with no local minima problem. The typical RBF configuration is shown in figure 2.

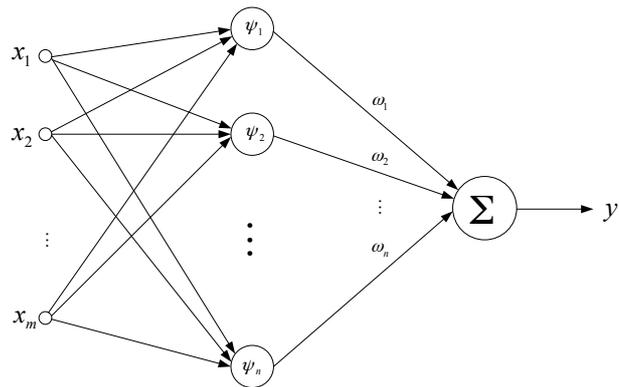


Fig. 2. Structure Diagram of RBF Neural Network

A single network output is produced linearly combining the outputs of the nonlinear activation with the weight vector ω of the output layer [10].

$$y = \sum_{i=1}^n \omega_i \psi_i \tag{15}$$

Where ω_i is the joint weighted value of the i th basis function.

The most commonly used radial base is the Gaussian function given as:

$$\psi_i(x) = \exp\left(-\frac{\|x - c_i\|^2}{\sigma_i^2}\right) \tag{16}$$

Where c_i and σ_i are center and spread of the i th RBF node.

Training the RBF neural network is done in two steps. In the first step centers are constructed by clustering the training data. The second step is basically a linear estimation of one weighting vector using ordinary least squares.

5. Example of Energy Consumption Prediction in the Beer Production

The day is always taken as a statistical unit in the actual production. The data in this study came from the production process measurement system of a beer enterprise. The measurement system real-time collects and computes the material and energy consumption in the produce. The

consumption of raw material crushing, saccharification and gelatinization processes can be computed in the same day. Fermentation process usually experiences fifteen working days and consumes the energy from the third working days. Because the working days of process influence each other, the input and output indices are chosen as follows.

Input indices: raw material consumption in the same day, hops consumption in the same day, raw material consumption before the twelfth day, yeast consumption before the twelfth day, energy consumption grade of tomorrow team.

Output index: power consumption of tomorrow production

Select 100 sets of data as training data, 50 groups of data as testing data. The deterministic data can be normalized to eliminate the dimensional effect. The qualitative data, team's energy consumption grade, can be determined using the fuzzy membership function. Such all data are transformed into the [0, 1] interval. Then the number of rules is received using the subtractive clustering. The figure 3 shows that the different distance can lead to the different number of rules. The smaller the distance, the class of clustering is more and the number of rules is also more. When the distance is 0.6 the number of rules is 49. The model of ANFIS is established using the training data. The error is shown in figure 4 after five times training. Along with the number increase, the error decrease.

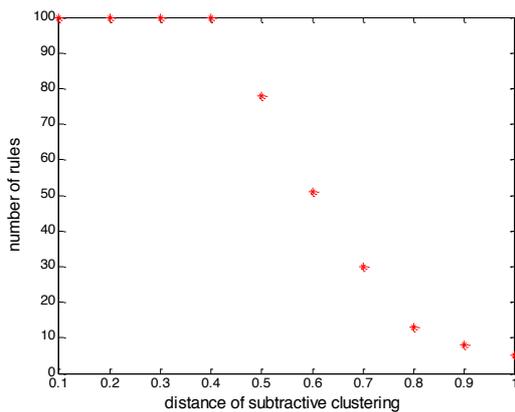


Fig. 3. The relationship between distance of subtractive clustering and number of rules

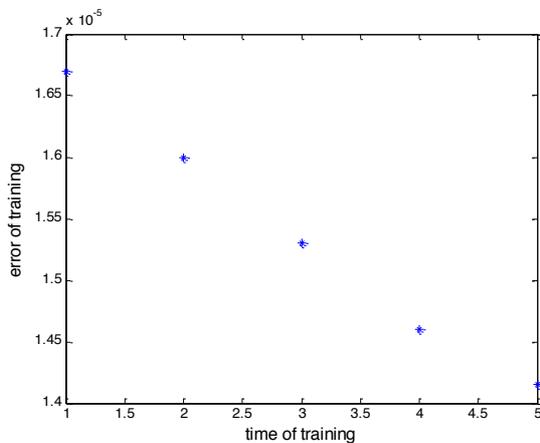
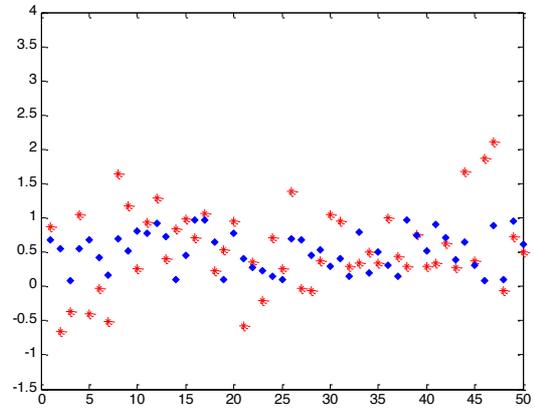
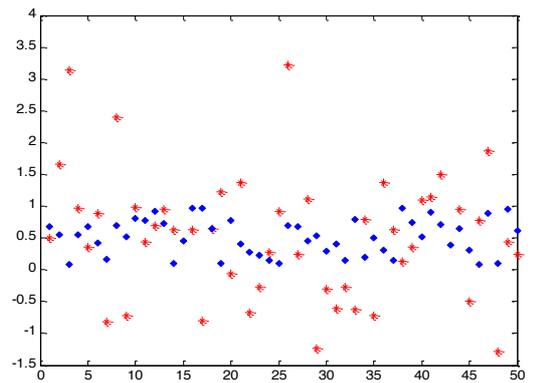


Fig. 4. The relationship between time of training and error of training in ANFIS

The model is tested using the testing data as shown in figure 5. The signal * expresses the prediction of the network, while the signal \diamond is the actual value. The figure (a) is the prediction of ANFIS model and (b) is the prediction of RBF neural network model. The mean error of ANFIS is 0.472, while the mean error of RBF is 1.868. Visibly, the accuracy of ANFIA is higher than that of RBF.



(a) ANFIS model



(b) RBF model

Fig. 5. The result between ANFIS and RBF models testing

6. Conclusion

China is a high energy consumption country in the world. It is eager to reduce energy consumption. Every industry should look for ways to save energy and improve the efficiency and avoid the waste of energy combining with their own situation. In recent years, the beer industry is developing rapidly in China. The market demand space is large, but overall energy utilization rate is not high.

The predictive model of energy consumption is built based on subtractive clustering and ANFIS in the beer production. Using the subtractive clustering on the historical data of energy consumption, the limit of artificial experience is conquered while confirming the number of fuzzy rules. The parameters of the fuzzy inference system are acquired by the structure of adaptive network and hybrid on-line learning algorithm. The concrete example shows that the precision of ANFIS model is higher than that of RBF neural network model. The method proposed in this paper can raise energy utilization rate and economic benefits for the enterprise. It also can be successfully applied to other similar problems in industry.

References

1. Melek Acar Boyacioglu, Derya Avci. "An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction of stock market return: The case of the Istanbul Stock Exchange", *Expert Systems with Applications*, 37, 2010, pp. 7908-7912.
2. Ilija Svalina, Vjekoslav Galzina, Roberto Lujić, Goran Šimunović. "An adaptive network-based fuzzy inference system (ANFIS) for the forecasting: The case of close price indices", *Expert Systems with Applications*, 40, 2013, pp. 6055-6063.
3. Cheng-Hung Chen, Yen-Yun Liao. "An efficient cluster-based tribes optimization algorithm for functional-link-based neurofuzzy inference systems", *Applied Soft Computing*, 13, 2013, pp. 2261-2271.
4. Haralambos Sarimveis, Alex Alexandridis, George Bafas. "A fast training algorithm for RBF networks based on subtractive clustering", *Neurocomputing*, 51, 2003, pp. 501-505.
5. H.M.Jiang, C.K.Kwong, W.H. Ip, T.C.Wong. "Modeling customer satisfaction for new product development using a PSO-based ANFIS approach", *Applied Soft Computing*, 12, 2012, pp. 726-734.
6. J. Fernandez de Canete, A. Garcia-Cerezo, I. Garcia-Moral, P. Del Saz, E. Ochoa. "Object-oriented approach applied to ANFIS modeling and control of a distillation column", *Expert Systems with Applications*, 40, 2013, pp.5648-5660.
7. Jyh-Shing Roger Jang. "ANFIS: Adaptive-network-based Fuzzy Inference System", *IEEE Transactions on Systems, Man, and Cybernetics*, 23, 1993, pp. 665-685.
8. Xuemei Dong, Ding-Xuan Zhou. "Learning gradients by a gradient descent algorithm", *Journal of Mathematical Analysis and Applications*, 341, 2008, pp. 1018-1027.
9. Jian-chuan Yin, Zao-jian Zou, Feng Xu. "On-line prediction of ship roll motion during maneuvering using sequential learning RBF neural networks", *Ocean Engineering*, 16, 2013, pp. 139-147.
10. Surajdeen A. Iliyas, Moustafa. Elshafei, Mohamed A. Habib, Ahmed A. Adeniran. "RBF neural network inferential sensor for process emission monitoring", *Control Engineering Practice*, 21, 2013, pp. 962-970.