Research on Ethanol Catalysis Based on Neural Network Genetic Algorithm and TOPSIS Model

Bo Qu, Xiaoying Huang, Ya Gao, Xintian Chen, Luxin Huang, Yandong Jiang, Meijun Zhu*

School of Aeronautical Engineering, Taizhou University, Zhejiang, Taizhou, 318000, China
**Corresponding author*

Abstract: *This paper studies the relationship between catalyst combination, temperature and various products, based on the chemical experimental data. Firstly, the evaluation matrix is constructed, and the TOPSIS model is constructed by nesting the three evaluation models of G1 method (order relationship analysis method), entropy weight method and critical method respectively, so that the evaluation object is weighted by subjective and objective combination. Through the ranking of 114 weight values, it can be concluded that the best scheme is 200mg1wt%Co/SiO2-200mgHAP-ethanol concentration of 0.9ml/min; T=400℃. Finally, the functional relationship between catalyst combination, temperature and C4 olefin yield is obtained by using neural network as fitting model. Genetic algorithm is selected for solution, and the best experimental conditions are: 200mg0.5wt%Co/SiO2-200mgHAP-ethanol concentration of 0.6ml/min; T=450℃.*

Keywords: *Entropy weight method, critical method, TOPSIS model, neural network, genetic algorithm*

1. Introduction

C_4 olefins are widely used in the production of chemical products and pharmaceutical intermediates. Ethanol is the raw material for the production of C_4 olefins. The combination of catalyst (Co loading, Co/SiO_2 and HAP loading ratio, ethanol concentration) and temperature will affect the selectivity and yield of C_4 olefins in the production of C_4 olefins. In order to explore the selectivity of products after ethanol catalytic coupling preparation, a chemical laboratory conducted a series of experiments at different temperatures for different catalyst combinations.

2. TOPSIS model based on subjective and objective combination weighting

2.1 Construction of evaluation matrix

The ethanol conversion N and C4 olefin selectivity n = 2 variables were used as indexes, and a total of 114 experiments (i.e. schemes) with different catalyst combinations and temperatures were used as 114 objects to be evaluated; Two indicators are used to rank 114 objects to be evaluated to construct evaluation matrix B, namely:

$$B = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \\ \cdots & \cdots \\ \omega_{1141} & \omega_{1142} \end{bmatrix}, \tag{1}$$

Among them, the column represents 114 combinations of catalyst and temperature, the row represents two indicators, and the i-th row and the j-th column represent the ethanol conversion n or C4 olefin selectivity under certain experimental conditions.

Substitute the experimental data into matrix B. see the following matrix for some data:

$$B = \begin{bmatrix} 2.06 & 34.05 \\ 5.85 & 37.43 \\ \cdots & \cdots \\ 69.40 & 38.17 \end{bmatrix}, \tag{2}$$

ISSN 2616-5767 Vol.4, Issue 6: 39-45, DOI: 10.25236/AJETS.2021.040607

2.2 Construction of evaluation matrix

The more the chemical reaction products, the higher the purity, the better. The research experiment in this paper is the preparation of C4 olefins by ethanol coupling. It can be seen that the standard to evaluate the experiment is to look at the two values of ethanol conversion N and C4 olefins. The larger the value, the better. It can be seen that the two indicators are very large indicators.

2.3 Dimensionless processing of data

Different index values of the evaluated object follow different dimensions, so the data cannot be calculated according to the same standard. The data should be standardized (dimensionless). After consulting the literature, it is decided to refer to the efficacy coefficient method to process the data. The processing formula is as follows:

$$\pi_{ij} = \frac{\omega_{ij} - min(\omega_{ij})}{max(\omega_{ij}) - min(\omega_{ij})} \times 0.5 + 0.5 \ , \tag{3}$$

where index a is dimensionless and recorded as B

Some data after dimensionless are as follows:

$$B = \begin{bmatrix} 0.61 & 0.85 \\ 0.63 & 0.87 \\ \cdots & \ldots \\ 0.91 & 0.88 \end{bmatrix} , \tag{4}$$

2.4 Dimensionless processing of data

2.4.1 Order relation analysis (G1 method)

Order relation analysis method, also known as G1 method, is a new method for calculating index weight after improving the eigenvalue method. This method does not need to construct judgment matrix and consistency test[3]. In addition, it can also keep the order of the index, which is more simple, fast, intuitive and effective in practical application[4].

1) Sort the indicators and evaluate the importance of the bright indicators with reference to experts in multiple related fields.

2) Assign the importance of the two indicators. The weight of the k-th indicator is recorded as β_k, and the importance of the two indicators can be expressed by O_k:

$$\varsigma_k = \beta_{k-1}/\beta_k (k = n, n-1, \cdots, 2). \tag{5}$$

3) Calculate weight.

According to formula (5), the weights of the two indicators are 0.5 and 0.5 respectively:

$$\beta_n = \frac{1}{1+\sum_{k=2}^{n}(\prod_{l=k}^{n} o_l)} \ , \tag{6}$$

Where o_l indicates the importance of comparison between the two indicators.

2.4.2 Entropy weight method and critical objective weighting

Critical method is an objective weighting method that uses the contrast strength and conflict between indicators to measure the weight of indicators. The contrast strength is measured by the standard deviation of indicators, and the conflict is measured by the correlation between indicators. If there is a strong negative correlation between indicators, the conflict between them is high [5].

Weight calculation formula:

$$\beta_j^1 = \frac{\sigma_i \sum_{i=1}^{m}(1-x_{ij})}{\sum_{j=1}^{n} \sigma_j \sum_{i=1}^{m}(1-x_{ij})} \ . \tag{7}$$

ISSN 2616-5767 Vol.4, Issue 6: 39-45, DOI: 10.25236/AJETS.2021.040607

Bring the value into formula (7) to obtain $\beta_1^1 = 0.4789, \beta_2^1 = 0.5211$.

Entropy was originally a thermodynamic concept. It was first introduced into information theory by Shan non to express the dispersion degree of an index. For example, the greater the dispersion degree of an index, the smaller the information entropy of the index [6].

The entropy calculation formula is:

$$\mu_j = -\frac{1}{\ln(m)}\sum_i^m \left(Q_{ij} \times \ln Q_{ij}\right) \quad . \tag{8}$$

Bring the value into formula (8) to obtain $\mu_1 = 0.8073, \mu_2 = 0.8664$.

The weight calculation formula of entropy weight method is:

$$\beta_j^2 = \frac{(1-\mu_j)}{\sum_{j=1}^n (1-\mu_j)} \quad . \tag{9}$$

Bring the value into formula (9) to obtain $\beta_j^2 = 0.5906, \mu_2 = 0.4094$.

Critical method takes into account the correlation and contrast between indicators, but does not consider the dispersion between indicators. Therefore, the combination of critical method and entropy weight method can fully consider the correlation, contrast and dispersion of indicators. The weight calculation formula after the combination of critical method and entropy weight method is shown in formula (10).

$$\varpi_j = \frac{(\sigma_j+\mu_j)\sum_{i=1}^n (1-x_{ij})}{\sum_{j=1}^m (\sigma_j+\mu_j)\sum_{i=1}^n (1-x_{ij})} \quad . \tag{10}$$

Bring the value into formula (9) to obtain $\beta_j^2 = 0.5906, \mu_2 = 0.4094$.

2.4.3 Subjective and objective combination weight

In order to avoid the limitation of a single kind of weighting, the combination of subjective and objective weighting can not only reflect the subjective intention of experts, but also reflect the objective characteristics of indicators. In this study, G1 method (subjective weighting method) is combined with the combined critical entropy weight method (objective weighting method), and the optimal combination weight is calculated by using the principle of minimum relative information entropy and Lagrange multiplication optimization method:

$$\overline{\omega}_J = \left[\frac{\sqrt{(\beta_1\sigma_1)}}{\sum_{j=1}^n \sqrt{(\beta_j\varpi_j)}}, \frac{\sqrt{(\beta_2\varpi_2)}}{\sum_{j=1}^n \sqrt{(\beta_j\sigma_j)}}, \cdots, \frac{\sqrt{(\beta_n\varpi_n)}}{\sum_{j=1}^n \sqrt{(\beta_j\pi_j)}}\right] \quad . \tag{11}$$

2.5 Construct weighted evaluation matrix

Construct a weighted evaluation matrix such as $D=d_{ij}$, including

$$d_{ij} = y_{ij}\overline{\omega}_J \quad . \tag{12}$$

2.6 Determine sample

Construct and determine the ideal sample and negative ideal sample. The indicators have been consistent. The ideal sample can be composed of the maximum value of each indicator in all samples, represented by D^+; The negative ideal sample can be used as the index in the sample composition of minimum value, expressed by D^-.

2.7 Calculate the distance between the sample and the ideal and negative ideal samples respectively

Compare each sample object with D^+ and D^-.

ISSN 2616-5767 Vol.4, Issue 6: 39-45, DOI: 10.25236/AJETS.2021.040607

$$L_i^+ = \sqrt{\sum_{j=1}^{n} \left(d_{ij} - D_j^+\right)^2}, L_i^- = \sqrt{\sum_{j=1}^{n} \left(d_{ij} - D_j^-\right)^2} \ . \tag{13}$$

Considering the equilibrium and approximation of the evaluation indexes of the two variables, this paper considers that the maximum evaluation value can be obtained only if each index is balanced and far enough from the negative ideal sample. Therefore, the relative proximity (H_i) from the sample to the ideal sample is used as the evaluation basis.

$$H_i = \frac{L_i^-}{L_i^+ + L_i^-}, (i = 1,2,\cdots,114) \ . \tag{14}$$

Finally, the relative proximity is used as the basis for the final evaluation weight to rank the evaluated objects.

3. Nested model of neural network fitting and genetic algorithm

It is required to find the combination of catalyst and temperature to make the C4 olefin yield as high as possible. The difference is that the second small question increases the constraint condition of T < 350℃. The existing experimental data are only discrete part of the data. If you want to find the optimal solution, you must find the relationship between these discrete data by fitting, and then find the optimal value through the iterative operation in the algorithm. Genetic algorithm has inherent implicit parallelism and better global optimization ability; It adopts the optimization method of probability theory, which can automatically obtain and guide the optimized search space, adjust the search direction intelligently, and do not need to determine the rules. It is impossible to calculate the global optimal solution in an acceptable time with an accurate algorithm. Genetic algorithm has good convergence. It is a global optimization algorithm with less calculation time under the same calculation accuracy. The biggest feature of genetic algorithm is that it makes use of the value information of objective function without gradient and other auxiliary information. It is only used for large-scale, highly nonlinear and non analytical objective function optimization problems. It has strong search ability. It is proved theoretically that it can completely converge to the global optimal solution; Using random operators instead of strict deterministic operations can directly approach the solution objectives of nonlinear, multi constraint and multi-objective optimization problems. The characteristics of genetic algorithm make genetic algorithm more suitable for the optimization of fitting function of multilayer neural network than other optimization algorithms. When solving this kind of optimization problem, genetic algorithm can solve many model defects, so it is most appropriate to combine genetic algorithm with neural network optimization to obtain the optimal solution.

3.1 Neural network fitting

The BP neural network model is established, the input layer data is set as X, the parameter connection weight and bias items from the input layer to the hidden layer are w and b_1 respectively, the parameter connection weight and bias items from the hidden layer to the output layer are V and b_2 respectively, and the activation functions are f_1 and f_2 (soft max).

The built model (i.e. activation function) is as follows:

$$\hat{y} = f_2(\text{ net }_2) = f_2(\tau^T h + b2) = f_2(\tau^T(W^T x + b_1) + b_2) \ . \tag{15}$$

Since the output of the neural network is a one-dimensional array, the following loss function is constructed:

$$E(\theta) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \ . \tag{16}$$

Repeat the iterative steps until the loss function is less than the preset threshold or the number of iterations is used up. The output parameters at this time are the best parameters at present. At this time, the fitting effect of neural network is better.

After many iterations, more than 50000 data and more predicted values are obtained, which provides more convenience for finding the global optimal solution. It is found that the error value is very small. Within the allowable range, the value predicted by neural network is reliable.

By constructing the neural network fitting model, the functional mapping relationship y=f (x1, x2, x3, x4, x5) between five factors such as catalyst combination and temperature and C4 olefin yield is obtained, where y represents C4 olefin yield, and the five factors are x1-x5 respectively. In order to find the optimal values of the five independent variables with the highest C4 olefin yield under the same experimental conditions, it is necessary to find the maximum value of the function mapping relationship, but at this time, the function relationship is more complex, and the conventional optimization algorithm can not be solved iteratively, so genetic algorithm is used for optimization.

3.2 Genetic algorithm optimization

1) Initialization of genetic algorithm parameters: the number of initialization iterations is T= 2000, the population size of each variable is n=40, the chromosome crossover probability is P_r=0.3, and the mutation probability is P_m=0.1

2) Divide the solution space of the optimization function: binary code each catalyst variable and temperature independent variable in the solution space. The variable range is the minimum range and maximum range of each variable in Annex 1. The solution space of each variable is taken at a certain interval to form a set of corresponding solution spaces.

3) Calculate the fitness through the selection function: take the fitting function of the neural network as the selection (fitness) function, bring the solution space value of each variable into the neural network model for prediction, and obtain the fitness value.

4) The selection probability of each individual in the population was calculated:

$$P(x_i) = \frac{f(x_i)}{\sum_{j=1}^{N} f(x_i)} \quad . \tag{17}$$

5) The roulette method is selected as the evolution direction of the population, the population size is taken as the number of roulette rotation, and the random simulation value is taken. The selected population is taken as the selected population, and the variable value and fitness size with the highest fitness after selection are saved.

6) Carry out crossover and mutation operations on the selected population to randomly generate a random number. If the random number is less than the crossover probability, the code exchange on the corresponding bit of the population in a single variable will be carried out. On the contrary, no crossover will be carried out; According to the calculation formula of the selected variation coefficient: $L = P_m \cdot N \cdot S$, where the variation coefficient P_m, population size N and S are the coding digits. If the selected variation coefficient is greater than 1, the variation operation will be carried out, the random value operation will be carried out on the coding bits of the variable, and the number of iterations will be added automatically

7) Repeat the operations from step 3 to step 6 for the population after selection and mutation. If the number of iterations reaches 2000, terminate the iteration.

8) Output the optimal Co loading, HAP catalyst quality, Co / SiO2 and HAP loading ratio, parameter values of ethanol concentration and temperature, and corresponding fitness values, that is, the maximum C4 olefin yield.

Through the calculation of genetic algorithm steps 1) - 8), the experimental conditions (optimal parameters) that make the C4 olefin yield the highest when the temperature is not limited can be obtained, that is:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 200.0 \\ 0.5000 \\ 200.000 \\ 0.6000 \\ 450.000 \end{pmatrix} . \tag{18}$$

At this time, the C4 olefin yield is $y_{max} = 85.0710$.

The experimental environment is:

200mg0.5wtCo / SiO2 - 200mgHAP - Ethanol concentration 0.6ml / min

3.3 Genetic algorithm optimization

The above is the experimental condition for the highest C4 olefin yield when the temperature is not limited. When the temperature is less than 350 ℃, that is, the constraint condition is added in the iteration. After the operation of the same process, the best parameters can be obtained.

When the temperature is less than 350 ℃, the experimental conditions (optimal parameters) that maximize the yield of C4 olefins are:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 200.0 \\ 5.0000 \\ 10.0000 \\ 0.3000 \\ 350.000 \end{pmatrix}. \tag{19}$$

At this time, the C4 olefin yield is $y_{max} = 52.4333$.

The experimental environment is:

200mg0.5wtCo / SiO2 - 10mgHAP - Ethanol concentration 0.3ml / min

4. Model Evaluation

4.1 Advantages of the model

1) TOPSIS Model Based on G1 method, entropy weight method and critical method.

This is an evaluation method combining subjective and objective weighting. By selecting one subjective weighting method and two objective weighting methods and referring to the weight value given by experts, almost most of them are scientific evaluations based on the objective attributes of the data itself, and the evaluation results will be very reliable.

2) Eural network.

The reference data in this paper is limited. The prediction and fitting ability of neural network can solve the problem of less reference data, which is helpful to optimize using genetic algorithm.

3) Genetic algorithm

Solving the optimal problem is most afraid of falling into local free solution, but genetic algorithm can avoid falling into local free solution through mutation mechanism, and its search ability is strong; Compared with other optimization models, genetic algorithm has strong expansibility and is easy to be combined with other algorithms. In a large number of data trained by neural network, the optimal solution can be found iteratively.

4.2 Disadvantages of the model

The data of expert evaluation index weight is difficult to obtain, and the neural network training time is too long, so it will be very troublesome to process the data.

Acknowledgments

The authors gratefully acknowledge the financial support by National Undergraduate Training Program for Innovation and Entrepreneurship.

References

[1] Xu Xiaohua, Hu Xiaofei, Zhao Yanchun. Discussion on nonlinear fitting based on polynomial [J]. Journal of Chifeng University (NATURAL SCIENCE EDITION), 2014, 30 (02): 37-38.
[2] Chu Shuzhen, Zhu Yanmei, Chen Yanyan. Comprehensive competitiveness evaluation of chain pharmacies based on entropy weight method [J]. China pharmacy, 2013, 24 (17): 1630-1632
[3] Liu Ye, Zhang Hongyuan, Li Wanqing, Zeng Lili. Evaluation of sustainable growth ability of high-

ISSN 2616-5767 Vol.4, Issue 6: 39-45, DOI: 10.25236/AJETS.2021.040607

tech enterprises in entrepreneurial period——From the perspective of G1 (order relation analysis) [J]. Accounting communication, 2016 (27): 110-113.
[4] Bian Zhiyuan. Research on comprehensive evaluation system of machining accuracy of five axis NC machine tools based on "s" part [D]. Chengdu: University of Electronic Science and technology, 2015
[5] Song Dongmei, Liu Chunxiao, Shen Chen, et al. Multi objective and multi-attribute decision making method based on subjective and objective weighting method [J]. Journal of Shandong University (Engineering Edition), 2015, 45 (4): 1-9.
[6] Li Man, Li Shifeng, Ouyang Yinghong. Evaluation and analysis of modern agricultural development level in Zhuolu County Based on entropy weight method [J]. Journal of China Agricultural University, 2014, 19 (5): 236-243.