Application of nonlinear partial least square in catalyst modeling

Huang Kai[1,3] Luo Zhenghong[2] Chen Fengqiu[3] Lü Dewei[3]

([1] Department of Chemistry and Chemical Engineering, Southeast University, Nanjing 210096, China)

([2] Department of Chemical Engineering, Xiamen University, Xiamen 361005, China)

([3] Department of Chemical Engineering, Zhejiang University, Hangzhou 310027, China)

Abstract: In this paper neural network partial least square (NNPLS) was used to establish a robust reaction model for a multi-component catalyst of methane oxidative coupling. The details, including the learning algorithm, the number of hidden units of the inner network, activation function, initialization of the network weights and the principal components, are discussed. The results show that the structural organizations of inner neural network are 1-10-5-1, 1-8-4-1, 1-8-5-1, 1-7-4-1, 1-8-4-1, 1-8-6-1, respectively. The Levenberg-Marquardt method was used in the learning algorithm, and the central sigmoidal function is the activation function. Calculation results show that four principal components are convenient in the use of the multi-component catalyst modeling of methane oxidative coupling. Therefore a robust reaction model expressed by NNPLS succeeds in correlating the relations between elements in catalyst and catalytic reaction results. Compared with the direct network modeling, NNPLS model can be adjusted by experimental data conveniently and the calculation of the model is simpler and faster than that of the direct network model.

Key words: partial least square; catalyst; oxidative coupling of methane; neural network; modeling

Statistical data analysis is widely applied in establishing a model from experimental or historical data. A typical method is the partial least square (PLS) method, which was proposed by Wold[1] to generalize a robust model. Although the PLS regression method provides a good remedy for the problems of correlated inputs and limited observations, its major restriction is that only linear information can be extracted from data. On the other hand, multi-layer neural networks[2,3] have been proven as universal approximators of any continuous function with arbitrarily desired accuracy[4]. But it should be pointed out that the number of weights in a multi-layer network could be larger than the number of observations in the case of limited data. Therefore, some of the weights cannot be uniquely determined from the observed data, and the direct network model will lead to overfitting. Based on the above consideration, the PLS regression and neural network are integrated to formulate an approach (NNPLS)[5,6], that can handle nonlinearity, correlated inputs and limited observations.

In this paper, NNPLS is applied in establishing a robust model for a multi-component catalyst of methane oxidative coupling. The NNPLS model is further research of the direct network model for the computer aided-design catalyst (CADC) method.

Received 2003-11-24.

Biography: Huang Kai (1973—), male, doctor, lecturer, huangk@seu.edu.cn.

1 Catalyst for Methane Oxidative Coupling

In the previous research[7], a multi-component catalyst for the oxidative coupling of methane (OCM), which contained Na, S, W, P, Zr and Mn elements and supported on SiO_2, was introduced, and prepared by the gelatification-impregnation method. A better catalyst, which C_2 hydrocarbon reached 20.77%, was designed by using direct neural network modeling. In further research[8], a catalyst, which C_2 hydrocarbon reached 27.78%, was designed. 100 catalysts were studied in those researches, and the experimental data, including the composition of catalysts and experimental results, are used in this paper.

2 NNPLS Method

2.1 NNPLS algorithm

Experimental data is divided into two matrices, in which $\boldsymbol{y}_i$ ($i = 1, 2, \cdots, p$) are p different quality indices, and $\boldsymbol{x}_j$ ($j = 1, 2, \cdots, m$) are m causal variables, and there are n samples of data being observed. Two data matrices can be expressed as

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \in \mathbf{R}^{n \times p}$$

$$\boldsymbol{Y}=\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1m} \\ y_{21} & y_{22} & \cdots & y_{2m} \\ \vdots & \vdots & & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nm} \end{bmatrix} \in \mathbf{R}^{n\times m}$$

in which each row is composed of one observation, and two columns in $\boldsymbol{Y}$ are the conversion of methane and selectivity of C_2, respectively. The data analysis problem is to relate the matrix $\boldsymbol{Y}$ as some function of the matrix $\boldsymbol{X}$ so as to predict $\boldsymbol{Y}$ using the data of $\boldsymbol{X}$.

The PLS decomposes matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ into bilinear products plus residual matrices:

$$\boldsymbol{X}=\boldsymbol{t}_1\boldsymbol{p}_1^{\mathrm{T}}+\boldsymbol{E}_1 \tag{1}$$

$$\boldsymbol{Y}=\boldsymbol{u}_1\boldsymbol{q}_1^{\mathrm{T}}+\boldsymbol{F}_1 \tag{2}$$

where $\boldsymbol{t}_1$, $\boldsymbol{u}_1 \in \mathbf{R}^n$ are score vectors of the first principal factor; and $\boldsymbol{p}_1 \in \mathbf{R}^m$, $\boldsymbol{q}_1 \in \mathbf{R}^p$ are loading vectors corresponding to this factor. All the four vectors are determined so that the residual matrices $\boldsymbol{E}_1$ and $\boldsymbol{F}_1$ are minimized. Formulas (1) and (2) formulate a PLS outer model.

Since the relations among different components in the catalyst are very complex and could be high nonlinear, the score vectors cannot be related by a linear function or an analytic function. Based on the nonlinear property of a neural network, the PLS inner model can be expressed by a neural network as

$$\boldsymbol{u}_1=\boldsymbol{N}(\boldsymbol{t}_1)+\boldsymbol{r}_1 \tag{3}$$

where $\boldsymbol{N}(\cdot)$ expresses the nonlinear relation represented by a neural network, which is determined by minimizing the residual $\boldsymbol{r}_1$.

After going through the above calculation, the residual matrices are calculated as

$$\boldsymbol{E}_1=\boldsymbol{X}-t_1p_1^{\mathrm{T}} \tag{4}$$

$$\boldsymbol{F}_1=\boldsymbol{Y}-N(t_1)q_1^{\mathrm{T}} \tag{5}$$

Then the second factor is calculated based on the residuals $\boldsymbol{E}_1$ and $\boldsymbol{F}_1$ by going through the same procedure as the first factor. The same procedure is repeated until the last factor a is calculated, which leaves almost no information in the residual matrices $\boldsymbol{E}_h$ and $\boldsymbol{F}_h$.

Based on the above discussion, the robust reaction model of a multi-component catalyst activity can be rewritten by iteration as

$$\boldsymbol{Y}=\sum_{h=1}^{a}\hat{\boldsymbol{u}}_h\boldsymbol{q}_h^{\mathrm{T}}+\boldsymbol{F}_a \tag{6}$$

where $\hat{\boldsymbol{u}}_h$ is expressed by formula (3).

The structure of the NNPLS algorithm is shown in Fig. 1[5], where the data is transformed to latent scores, then neural network is used to learn the scores.

For the 6-component catalyst, $m = 6$; methane conversion and C_2 hydrocarbon selectivity are reaction performance studied, $p = 2$; and 100 catalysts and experimental results are obtained in the previous research, $n = 100$.

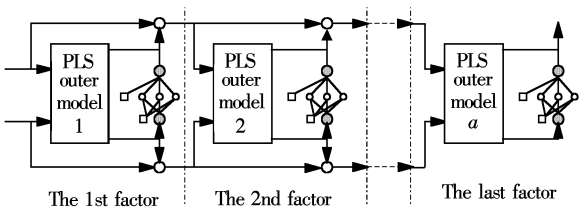

Fig.1 A schematic illustration of the NNPLS method

2.2 Standardization of initial data

In order to be handled conveniently, the initial experimental data should be standardized. In this paper, $\boldsymbol{X}$ and $\boldsymbol{Y}$ are scaled to zero-mean.

2.3 Choosing an activation function

In this particular application, the network is used to simulate the relation between $\boldsymbol{u}_h$ and $\boldsymbol{t}_h$ as shown in formula (3). It is known that both $\boldsymbol{u}_h$ and $\boldsymbol{t}_h$ have the following property[9]:

$$\sum_{i=1}^{n}u_{hi}=0 \tag{7}$$

$$\sum_{i=1}^{n}t_{hi}=0 \tag{8}$$

where u_{hi} and t_{hi} are the i-th elements of $\boldsymbol{u}_h$ and $\boldsymbol{t}_h$, respectively. Therefore, the following centered sigmoidal function is chosen to model the inner relation because zero input to the sigmoid gives zero output.

$$f(x)=\frac{1-\mathrm{e}^{-x}}{1+\mathrm{e}^{-x}} \tag{9}$$

2.4 Choosing a learning algorithm

The neural network used in the NNPLS method would be based on a back-propagation algorithm. However, the learning speed of the traditional gradient-descend method is very slow, compared with other improved learning algorithms, and that of the Levenberg-Marquardt algorithm is faster. On the other hand, the network used for the inner model in NNPLS is single input and single output (SISO), so the complexity of calculation is not great. It is convenient to use the Levenberg-Marquardt method as a learning algorithm for the NNPLS method.

2.5 The number of hidden layers and hidden units

The numbers of hidden layers and hidden units are determined by the complexity of input-output relationship. For the SISO network used in NNPLS, 1 or 2 hidden layer(s) is convenient. But, it should be stated that adding more hidden units cannot obtain

better generalization. Therefore, it is important to properly choose the number of hidden units, which is determined by a simplified cross-validation method scheme[5]. The experimental data would be divided into a training set and a testing set, which can be expressed as

$$\boldsymbol{u}_h^{\text{test}} = \boldsymbol{E}_{h-1}^{\text{test}} \boldsymbol{w}_h \tag{10}$$

$$\boldsymbol{t}_h^{\text{test}} = \boldsymbol{F}_{h-1}^{\text{test}} \boldsymbol{q}_h \tag{11}$$

where

$$\boldsymbol{E}_h^{\text{test}} = \boldsymbol{E}_{h-1}^{\text{test}} - \boldsymbol{t}_h^{\text{test}} \boldsymbol{p}_h^{\text{T}},\ \boldsymbol{E}_0^{\text{test}} = \boldsymbol{X}^{\text{test}} \tag{12}$$

$$\boldsymbol{F}_h^{\text{test}} = \boldsymbol{F}_{h-1}^{\text{test}} - \boldsymbol{N}(t_h^{\text{test}}) \boldsymbol{q}_h^{\text{T}},\ \boldsymbol{F}_0^{\text{test}} = \boldsymbol{Y}^{\text{test}} \tag{13}$$

where $\boldsymbol{w}_h$, $\boldsymbol{p}_h$, $\boldsymbol{q}_h$ have been calculated in the NNPLS algorithm. The number of hidden units is determined by the testing results. If the error between the actual data and predicted data is large enough, more hidden units should be added.

2.6 Choosing initial weights of the inner network

It is important to choose properly the initial weights, which have an effect on convergence and learning speed of the inner network. A usual approach is to choose the initial weights randomly, but it is not reliable for training the network. In this paper, a simple method is introduced to initialize the weights. $\boldsymbol{u}_h$ and $\boldsymbol{t}_h$ are a correlated linear function:

$$\boldsymbol{u}_h = \boldsymbol{b}_h \boldsymbol{t}_h + \boldsymbol{r}_h \tag{14}$$

where $\boldsymbol{b}_h$ is a coefficient which is determined by minimizing the residual $\boldsymbol{r}_h$, and calculated by

$$\boldsymbol{b}_h = \frac{\boldsymbol{t}_h^{\text{T}} \boldsymbol{u}_h}{\boldsymbol{t}_h^{\text{T}} \boldsymbol{t}_h} \tag{15}$$

Then the network which has the same structural organization is used to simulate formula (14), and some network weights can be obtained. The weights would be used as the initial weights for the inner model of NNPLS.

2.7 Estimation of the number of principal components

It is not true that the more principal components there are, the more accurate the NNPLS model is. Too many principal components would result in an over-parameterized model, while too few principal components under-parameterize the problem. Therefore, a crossvalidation method is used to estimate the number of principal components. The experimental data is divided into g groups; the data in $g-1$ groups is used to found the model, which would be tested by the additional data every time. Based on the above method, all data would be predicted, and the root mean squares errors of the quality matrix are calculated and used as the proof for estimation of the number of principal components.

3 Catalyst Modeling

Based on the above consideration, the NNPLS method is applied in catalyst modeling for methane oxidative coupling. Principal inner models for catalyst modeling are shown in Fig.2.

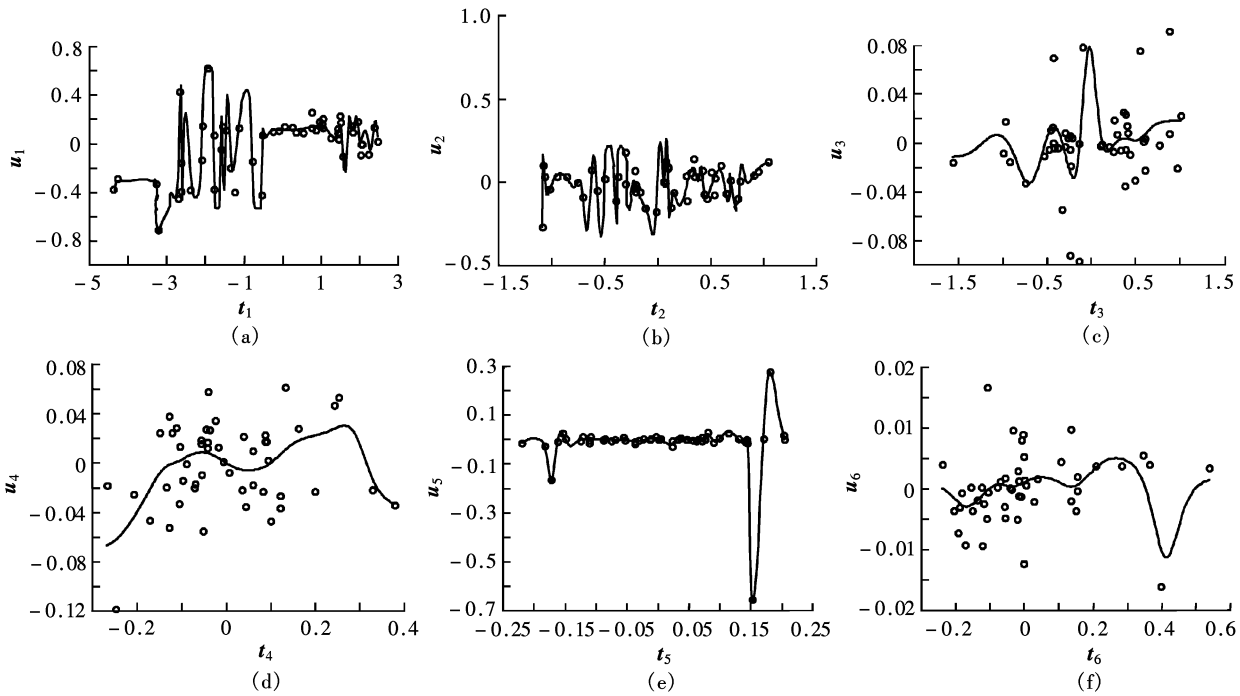

Fig.2 Principal inner models in the NNPLS catalytic robust model of OCM

The structural organizations of six principal components are 1-10-5-1 (one input unit, 10 first hidden units, 5 second hidden units, one output unit), 1-8-4-1, 1-8-5-1, 1-7-4-1, 1-8-4-1 and 1-8-6-1, respectively.

From Fig.2, by increasing the principal components, the accuracy of fitting would be improved, which also is shown in Fig.3(a). But too many principal components would lead to overfitting. It can be found in Fig.3(b), the prediction errors do not decrease with the increase of the principal components, and the best number of the principal components is 4.

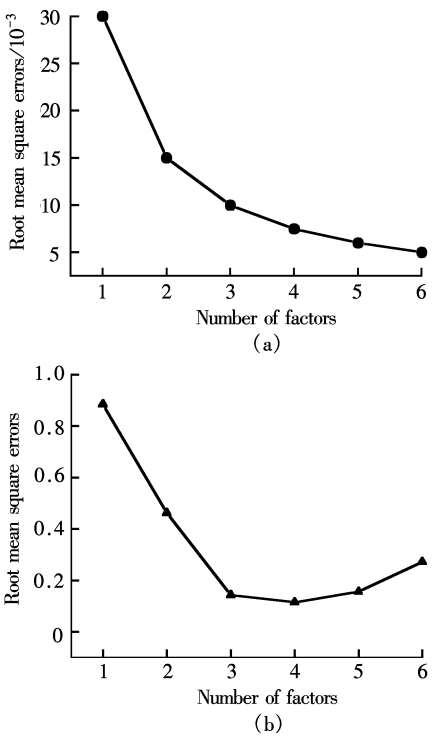

Fig.3 Cross-validation results in NNPLS method. (a) Training errors; (b) Prediction errors

Based on the above discussions, the relation between catalytic performance and catalyst components of catalyst for methane oxidative coupling can be expressed as

Outer model:

$$[\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4, \boldsymbol{x}_5, \boldsymbol{x}_6] = \boldsymbol{T}\boldsymbol{P}^{\mathrm{T}} + \boldsymbol{E}_4 = \boldsymbol{t}_1\boldsymbol{p}_1^{\mathrm{T}} + \boldsymbol{t}_2\boldsymbol{p}_2^{\mathrm{T}} + \boldsymbol{t}_3\boldsymbol{p}_3^{\mathrm{T}} + \boldsymbol{t}_4\boldsymbol{p}_4^{\mathrm{T}} + \boldsymbol{E}_4 \quad (16)$$

$$[\boldsymbol{X}_{\mathrm{CH_4}}, \boldsymbol{S}_{\mathrm{C_2}}] = \boldsymbol{U}\boldsymbol{Q}^{\mathrm{T}} + \boldsymbol{F}_4 = \boldsymbol{u}_1\boldsymbol{q}_1^{\mathrm{T}} + \boldsymbol{u}_2\boldsymbol{q}_2^{\mathrm{T}} + \boldsymbol{u}_3\boldsymbol{q}_3^{\mathrm{T}} + \boldsymbol{u}_4\boldsymbol{q}_4^{\mathrm{T}} + \boldsymbol{F}_4 \quad (17)$$

Inner model:

$$\boldsymbol{U} = \boldsymbol{N}(\boldsymbol{T}) + \boldsymbol{R} = [\boldsymbol{N}_1(\boldsymbol{t}_1), \boldsymbol{N}_2(\boldsymbol{t}_2), \boldsymbol{N}_3(\boldsymbol{t}_3), \boldsymbol{N}_4(\boldsymbol{t}_4)] + [\boldsymbol{r}_1 + \boldsymbol{r}_2 + \boldsymbol{r}_3 + \boldsymbol{r}_4] \quad (18)$$

where $\boldsymbol{x}_1$ to $\boldsymbol{x}_6$ are molecular contents of six elements in catalysts, $\boldsymbol{X}_{\mathrm{CH_4}}$ and $\boldsymbol{S}_{\mathrm{C_2}}$ are methane conversion and C_2 hydrocarbons selectivity, respectively.

In order to compare with linear PLS regression, the comparison in performance between NNPLS and LPLS is shown in Tab.1.

Tab.1 Comparison in performance between NNPLS and LPLS

Factors	NNPLS		LPLS	
	$\lVert \boldsymbol{E}_h - \hat{\boldsymbol{E}}_h \rVert$	$\lVert \boldsymbol{E}_h^{\mathrm{test}} - \hat{\boldsymbol{E}}_h^{\mathrm{test}} \rVert$	$\lVert \boldsymbol{E}_h - \hat{\boldsymbol{E}}_h \rVert$	$\lVert \boldsymbol{E}_h^{\mathrm{test}} - \hat{\boldsymbol{E}}_h^{\mathrm{test}} \rVert$
1	0.030 0	0.884 6	2.632 1	3.648 9
2	0.015 0	0.462 7	1.873 4	2.657 2
3	0.010 0	0.143 5	1.267 2	2.075 1
4	0.007 5	0.115 6	0.973 6	1.654 7
5	0.006 0	0.156 8	0.875 6	1.596 6
6	0.005 0	0.272 3	0.859 7	1.595 7

In Tab. 1, $\lVert \boldsymbol{E}_h - \hat{\boldsymbol{E}}_h \rVert$ is the modulus of the training set, and $\lVert \boldsymbol{E}_h^{\mathrm{test}} - \hat{\boldsymbol{E}}_h^{\mathrm{test}} \rVert$ is the modulus of the testing set. From Tab.1, it is known that NNPLS succeeds in establishing the catalyst model, which is a complex nonlinear input-output relationship, and cannot be explained simply by linear PLS regression.

Compared with the direct network model of multi-component catalyst, which will be discussed in another paper, the NNPLS model can be conveniently adjusted by experimental data and the calculation of the model is simpler and faster than that of the direct network model.

4 Conclusion

In this paper, NNPLS is used to establish a robust reaction model for a multi-component catalyst of methane oxidative coupling. The details, including the learning algorithm, the number of hidden units of the inner network, activation function, initialization of the network weights and the principal components, are discussed. The results show that the structural organizations of the inner neural network are 1-10-5-1, 1-8-4-1, 1-8-5-1, 1-7-4-1, 1-8-4-1, 1-8-6-1, respectively. The Levenberg-Marquardt method is used as the learning algorithm, and the central sigmoidal function is the activation function. Calculation results show that four principal components are convenient for the multi-component catalyst modeling of methane oxidative coupling. Therefore a robust reaction model expressed by NNPLS succeeds in correlating the relations between elements in the catalyst and catalytic reaction results.

References

[1] Wold H. Nonlinear estimation by iterative least squares procedures [A]. In: David F, ed. *Research Papers in Statistics* [C]. New York: John Wiley & Sons, 1966.411-444.

[2] Werbos P J. Beyond regression: new tools for prediction and analysis in the behavioral science [D]. Cambridge, MA: Harvard University, 1974.
[3] Rumelhart D, Hinton G, Williams R. *Parallel distributed processing* [M]. Cambridge, MA: MIT Press, 1986.
[4] Hornik K M, Stinchcombe M, White H. Multilayer feed-forward neural networks are universal approximators [J]. *Neural Networks*, **1989, 2**(2): 359－366.
[5] Qin S J, McAvoy T J. Nonlinear PLS modeling using neural networks [J]. *Computers Chem Engng*, **1992, 16** (4): 379－391.
[6] Holcomb T R, Morari M. PLS/neural networks [J]. *Computers Chem Engng*, **1992, 16**(4): 393－411.
[7] Huang K, Chen F Q, Lü D W. Artificial neural network-aided design of a multi-component catalyst for methane oxidative coupling [J]. *Appl Catal A*, **2001, 219**: 61－68.
[8] Huang K. Research on computer-aided catalyst design and an application in multi-component catalyst design of methane oxidative coupling [D]. Hangzhou: Department of Chemical Engineering, Zhejiang University, 2001.
[9] Geladi P, Kowalski B R. An example of 2-block predictive partial least-squares regression with simulated data [J]. *Analyt Chim Acta*, **1986, 185**: 19－32.

非线性偏最小二乘法在催化剂建模中的应用

黄 凯[1,3] 罗正鸿[2] 陈丰秋[3] 吕德伟[3]

([1] 东南大学化学化工系, 南京 210096)
([2] 厦门大学化学工程系, 厦门 361005)
([3] 浙江大学化学工程系, 杭州 310027)

摘要: 神经网络偏最小二乘法(NNPLS)被应用于一种甲烷氧化偶联多组分催化剂的鲁棒反应模型的建立.重点研究了内层神经网络学习算法、激活函数、网络结构(包括隐含节点数、隐含层)、网络权值初始化及主元的选取原则等.研究表明, 内层神经网络分别采用 1-10-5-1,1-8-4-1,1-8-5-1, 1-7-4-1, 1-8-4-1, 1-8-6-1 的拓扑结构是合适的; Levenberg-Marquardt 方法被用于网络的学习算法可以加快学习速度; 同时采用了 sigmoid 函数为激活函数.计算结果显示, 四主元可以满足建模的需要.与单纯的神经网络催化剂模型相比, NNPLS 方法压缩分解了变量, 减少了计算量, 同时使模型的推广能力得到提高, 有效地改善了直接神经网络建模过程中催化剂模型泛化能力较差的缺点.

关键词: 偏最小二乘法; 催化剂; 甲烷氧化偶联; 神经网络; 建模
中图分类号: TQ426.6