

研究论文

# 基于 MLFN-PLSR 的 PX 氧化反应组合建模方法

颜学峰

(华东理工大学自动化研究所, 上海 200237)

**摘要:** 针对对二甲苯 (*p*-xylene, PX) 氧化反应过程中影响主要副产物对羧基苯甲醛 (4-carboxybenzaldehyde, 4-CBA) 含量的因素众多且呈高度非线性的特征, 提出了多层前向型神经网络 (multi-layer feedforward network, MLFN) 与偏最小二乘回归 (partial least squares regression, PLSR) 相结合的建模方法, 建立反应产物中 4-CBA 含量关联模型。MLFN-PLSR 采用三层网络结构和尽量多的隐节点, 通过 MLFN 充分提取样本数据信息; 然后采用 PLSR 消除隐含层输出冗余信息, 建立具有良好预测精度的模型。与 MLFN 相比, 最佳性能模型的预测偏差平方和均值下降了 12.11%、模型平均预测偏差平方和均值下降了 8.37%。与 PLSR 相比, 最佳性能模型的预测偏差平方和均值下降了 70.62%。

**关键词:** 神经网络; 偏最小二乘回归; 对二甲苯; 对羧基苯甲醛

中图分类号: TP 183

文献标识码: A

文章编号: 0438-1157 (2007) 01-0149-06

## Develop *p*-xylene oxidation reaction model based on MLFN-PLSR

YAN Xuefeng

(Institute of Automation, East China University of Science and Technology, Shanghai 200237, China)

**Abstract:** Due to the fact that there exist many factors having highly-nonlinear and complex effects on the concentration of 4-carboxybenzaldehyde (4-CBA), the most important intermediate product of *p*-xylene (PX) oxidation reaction, a novel approach integrating multi-layer feedforward network (MLFN) with partial least squares regression (PLSR) was proposed to develop a model of 4-CBA concentration in the PX oxidation product. A three-layer network consisting of an input layer, a single hidden layer and an output layer was selected by MLFN-PLSR and the number of the hidden layer nodes was as large as possible. Firstly, MLFN learned from the training sample. Secondly, PLSR was used to identify PLS components from the hidden-layer node output and remove the correlation among them. And, an optimal prediction ability model with the optimal number of the latent variables was obtained according to the prediction ability of the model for the verified sample. The comparison results showed that the prediction ability of the optimal MLFN-PLSR was 12.11% higher than that of the optimal MLFN and 70.62% higher than that of the optimal PLSR, and the mean prediction ability of MLFN-PLSR was 8.37% higher than that of MLFN.

**Key words:** artificial neural network; partial least square regression; *p*-xylene; 4-carboxybenzaldehyde

2005-11-30 收到初稿, 2006-04-27 收到修改稿。

联系人及第一作者: 颜学峰 (1972-), 男, 博士, 副研究员。

基金项目: 国家自然科学基金项目 (20506003), 教育部科学技术研究重点项目 (106073), 上海科技启明星项目 (04QMX1433)。

Received date: 2005-11-30.

Corresponding author: YAN Xuefeng. E-mail: xfyan@ecust.edu.cn

Foundation item: supported by the National Natural Science Foundation of China (20506003), the Key Project of Chinese Ministry of Education (106073), the Science and Technology Phosphor of Shanghai (04QMX1433).

## 引 言

精对苯二甲酸 (PTA) 是合成聚酯纤维和塑料的重要原料, 主要用来合成聚酯的中间体苯二甲酸乙二醇酯。PTA 的生产工艺主要是以钴、锰为催化剂, 溴为促进剂, 乙酸为溶剂, 液相催化氧化对二甲苯 (PX) 生成对苯二甲酸 (TA)<sup>[1-7]</sup>。在 PX 氧化生成 TA 过程中, 对羧基苯甲醛 (4-CBA) 是氧化反应的主要副产物, 4-CBA 含量过高将增加下游精制负担, 甚至影响 PTA 产品品质; 4-CBA 含量过低, 则氧化过度, 燃烧损失严重, 物耗增大。因此各 PTA 生产厂商都采取各种途径力图将 TA 产品中 4-CBA 含量稳定在一个合适的水平。由于在 PX 氧化反应过程中, 4-CBA 含量只能离线采样分析, 当前的反应情况一般要等到 2 h 后才能得知, 这给优化控制 PX 氧化反应过程, 即 TA 产品中 4-CBA 含量带来困难。因此建立良好反映各反应因素对 4-CBA 含量影响的关联模型意义重大。

过程数学建模, 已提出了很多方法。其中神经网络 (ANN), 特别是多层前向型神经网络 (MLFN), 由于具有逼近任意非线性过程等特点得到了广泛地应用<sup>[8-11]</sup>, 但其最佳隐节点个数难以确定, 会出现过拟合等现象<sup>[11]</sup>。而各种全局多元线性回归方法<sup>[12-14]</sup>, 具有简洁明确的解析表达形式, 建模速率高, 但由于在 PX 氧化反应过程中, 影响 TA 产品中 4-CBA 含量的因素多且复杂, 且呈高度非线性特征, 由此所建的回归模型, 其预报能力往往不能令人满意。针对 MLFN 和多元线性回归方法的优缺点, 提出一种 MLFN 和偏最小二乘回归 (PLSR) 相结合的组合建模方法, 该方法融合 MLFN 和 PLSR 优点, 克服各自的缺点, 具有强的非线性表达能力, 且模型鲁棒性强。

## 1 影响 4-CBA 含量的因素分析

约 80% 的 PTA 采用 BP-AMOCO 工艺生产, 图 1 是 BP-AMOCO 工艺氧化工段流程图。新鲜 PX、新鲜 HAc、新鲜钴锰催化剂、新鲜溴促进剂、母固回收、氧化反应单元冷凝器部分凝液、高压吸收塔底洗涤液、循环母液进入配料混合罐, 混合成一定 PX 浓度的溶剂; 然后依次进入氧化反应单元、结晶单元、真空过滤机, 最后到母液罐; 含 TA 滤饼经干燥机干燥成粗 TA 出系统, 母液罐

中母液部分出系统去溶剂回收单元, 其余循环回配料混合罐。PX 首先在氧化反应单元的氧化反应器中进行一次氧化, 接着反应器的出料进入结晶单元的第一结晶器进行二次氧化。

由 PX 氧化反应机理<sup>[1-7]</sup>和工艺流程分析, 在工业 PX 氧化反应过程中, 影响 TA 产品中 4-CBA 含量 ( $C_{4-CBA}$ , %) 的主要因素有: 反应进料中的钴催化剂浓度 ( $C_{Co}$ ,  $\text{kg} \cdot \text{kg}^{-1}$ ), 锰催化剂浓度 ( $C_{Mn}$ ,  $\text{kg} \cdot \text{kg}^{-1}$ ), 溴促进剂浓度 ( $C_{Br}$ ,  $\text{kg} \cdot \text{kg}^{-1}$ ), 反应温度 ( $T_{\text{reactor}}$ ,  $^{\circ}\text{C}$ ), 反应器抽出水流量 ( $m_{\text{H}_2\text{O}}$ ,  $\text{kg} \cdot \text{h}^{-1}$ ), 氧化反应器停留时间 ( $\tau_{\text{reactor}}$ , s), 第一结晶器温度 ( $T_{\text{crystallizer}}$ ,  $^{\circ}\text{C}$ ), 第一结晶器停留时间 ( $\tau_{\text{crystallizer}}$ , s) 等因素。这些影响因素值都可以通过装置 DCS 系统直接 (或间接计算) 获得。

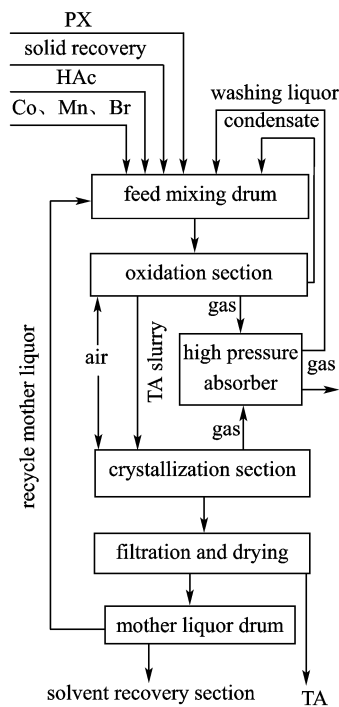


图 1 对二甲苯氧化反应工艺流程图

Fig. 1 Flow chart of PX oxidation reaction workshop section

## 2 MLFN-PLSR 组合方法

### 2.1 多层前向型神经网络

设 MLFN 由  $L+1$  层组成, 第  $j$  层有  $N_j$  个神经元, 则输入层 (第 0 层) 神经元个数为  $N_0$ , 等于输入向量  $x$  的维数, 输出层 (第  $L$  层) 神经元个数为  $N_L$ , 等于样本的期望输出向量  $y$  的维数。

设训练样本集  $S$  含  $A$  个样本, 对第  $m$  输入样本向量  $\mathbf{x}_m$  作用于网络后, 第  $L$  层神经元的目标输出  $\hat{y}_{mL}$  和期望输出  $y_m$  的偏差为

$$e_{mLk} = y_{mk} - \hat{y}_{mLk}, k = 1, 2, \dots, N_L \quad (1)$$

其中,

$$\begin{cases} \hat{y}_{mj k} = f(z_{mj k}) & j = 1, 2, 3, \dots, L \\ \hat{y}_{mj k} = x_{mk} & j = 0 \end{cases} \quad (2)$$

$$z_{mj k} = \sum_{i=0}^{N_{j-1}} \omega_{jki} \hat{y}_{m(j-1)i} \quad (3)$$

式中  $e_{mLk}$  为第  $m$  样本输入向量  $\mathbf{x}_m$  作用于网络后, 第  $L$  层 (输出层) 第  $k$  神经元函数输出的偏差;  $\hat{y}_{mLk}$  为第  $L$  层第  $k$  神经元在第  $m$  样本状态下的函数输出;  $y_{mk}$  为第  $m$  样本第  $k$  期望输出分量;  $\hat{y}_{mj k}$  为  $j$  层中第  $k$  神经元在第  $m$  样本状态下的函数输出, 其中  $\hat{y}_{mj0} = 1$ ;  $f(\cdot)$  为神经元激活函数;  $x_{mk}$  为  $\mathbf{x}_m$  的第  $k$  分量;  $\omega_{jki}$  为  $j-1$  层中第  $i$  神经元到  $j$  层中第  $k$  神经元的连接权值, 其中  $\omega_{jko}$  定义为  $j$  层第  $k$  神经元的阈值。则对第  $m$  样本学习的误差  $E_m = \frac{1}{2} \sum_{k=1}^{N_L} e_{mLk}^2$ , 对所有样本的学习误差平方和  $E$  定义为

$$E = \sum_{m=1}^A E_m \quad (4)$$

网络训练<sup>[8,10,15]</sup> 通过寻求最佳权值和阈值, 减少  $E$ 。

## 2.2 偏最小二乘回归

设样本容量为  $n$ , 自变量维数为  $p$ , 因变量维数为  $q$ , 则自变量数据矩阵  $\mathbf{X}$  为  $n \times p (n > p)$  维, 因变量数据矩阵  $\mathbf{Y}$  为  $n \times q (n > q)$  维。针对各自变量间可能存在复共线性, 可以采用 PLSR<sup>[13-14]</sup>。PLS 成分的提取可用 NIPALS 算法<sup>[14]</sup>, 设取前  $k$  个 PLS 成分形成隐变量组成矩阵  $\mathbf{T}$ , 则 PLSR 模型

$$\mathbf{Y} = \mathbf{T}\mathbf{V}^T + \mathbf{Y}_{k+1} = \mathbf{X}\mathbf{W}\mathbf{V}^T + \mathbf{Y}_{k+1} \quad (5)$$

式中  $\mathbf{V}$  是  $q \times k$  维回归系数矩阵,  $\mathbf{W}$  是  $p \times k$  维转置矩阵,  $\mathbf{Y}_{k+1}$  是  $n \times q$  维残差矩阵。在 NIPALS 算法中可以同时计算出  $\mathbf{W}$  与  $\mathbf{V}$ , 则  $q$  个因变量的预报值可用如下的回归方程计算

$$\hat{\mathbf{y}} = \mathbf{x}\mathbf{W}\mathbf{V}^T \quad (6)$$

式 (6) 的预报性能与选用的 PLS 成分个数  $k$  (即隐变量个数) 密切相关<sup>[12-13]</sup>。可以采用交叉验证算法<sup>[12-13]</sup> 或以对校验样本预测性能为指标求得最佳参与回归 PLS 成分个数  $k$  (即隐变量个数)。

## 2.3 MLFN-PLSR 组合方法步骤

由于 3 层神经网络就可以描述任何非线性过程, MLFN-PLSR 的网络结构为 3 层, 即包含输入层、隐含层和输出层。MLFN-PLSR 将采用尽可能多的隐节点, 通过网络训练, 充分提取样本数据中的信息, 建立 MLFN 模型; 然后针对隐节点个数增多, 造成隐含层输出之间存在严重的相关性, 甚至复共线性, 采用 PLSR 建立隐含层输出与输出层期望和输出之间最佳关联模型, 最终获得 MLFN-PLSR 模型。

设样本输入向量  $\mathbf{x}$  的维数为  $p$ , 样本期望输出向量  $\mathbf{y}$  的维数为  $q$ , 采集到总样本为  $\mathbf{S}_{\text{ALL}}$ , 将  $\mathbf{S}_{\text{ALL}}$  分为训练样本  $\mathbf{S}$  和校验样本  $\mathbf{S}_T$ ,  $\mathbf{S}$  的样本个数为  $A$ ,  $\mathbf{S}_T$  的样本个数为  $A_T$ , 并假设采集到的样本都经过归一化处理。则输入层神经元个数  $N_0 = p$ , 输出层神经元个数  $N_2 = q$ , 设隐含层神经元个数为  $N_1$ 。MLFN-PLSR 算法步骤如下。

(1) 采用 BP 算法或进化算法等对训练样本  $\mathbf{S}$  进行学习, 确定网络权值和阈值。

(2) 对输入向量  $\mathbf{x}_m, m=1, 2, \dots, A$ , 由式 (2)、式 (3) 求得隐含层输出  $\hat{\mathbf{y}}_{m1}$ , 即

$$\begin{cases} z_{m1i} = \sum_{j=0}^{N_0} \omega_{1ij} x_{mj} \\ \hat{y}_{m1i} = f(z_{m1i}) \end{cases} \quad i = 1, 2, \dots, N_1 \quad (7)$$

式中  $z_{m1i}$  是输入向量  $\mathbf{x}_m$  作用于 MLFN 后, 隐含层第  $i$  神经元的和输出;  $x_{mj}, j=1, 2, \dots, N_0$  是  $\mathbf{x}_m$  的第  $j$  个分量,  $x_{m0} = 1$ ;  $\omega_{1i0}$  为隐含层第  $i$  神经元的阈值;  $\hat{y}_{m1i}$  为  $\hat{\mathbf{y}}_{m1}$  的第  $i$  分量。

(3) 对隐含层输出向量  $\hat{\mathbf{y}}_{m1}$  添加阈值分量, 形成新的向量  $\bar{\mathbf{x}}_m$ , 即

$$\begin{cases} \bar{x}_{mi} = \hat{y}_{m1i} & i = 1, 2, \dots, N_1 \\ \bar{x}_{mi} = 1 & i = N_1 + 1 \end{cases} \quad (8)$$

式中  $\bar{x}_{mi}, i=1, 2, \dots, N_1+1$  是  $\bar{\mathbf{x}}_m$  的第  $i$  个分量。设由所有  $\bar{\mathbf{x}}_m$  组成非线性变换后矩阵  $\bar{\mathbf{X}}$ 。对校验样本  $\mathbf{S}_T$  中所有输入向量采用相同方法求得对应矩阵  $\bar{\mathbf{X}}_T$ 。

(4) 通过输出层神经元激活函数的逆函数, 求得训练样本  $\mathbf{S}$  中所有期望输出向量  $\mathbf{y}_m, m=1, 2, \dots, A$  的输出层期望和输出向量  $\bar{\mathbf{y}}_m$ , 即

$$\bar{y}_{mi} = f^{-1}(y_{mi}), i = 1, 2, \dots, N_2 \quad (9)$$

式中  $\bar{y}_{mi}$  是  $\bar{\mathbf{y}}_m$  的第  $i$  个分量;  $y_{mi}$  是  $\mathbf{y}_m$  的第  $i$  个分

量;  $f^{-1}(\cdot)$  是输出层神经元激活函数的逆函数。设由所有输出层期望和输出向量  $\bar{\mathbf{y}}_m$  组成矩阵  $\bar{\mathbf{Y}}$ 。对校验样本  $\mathbf{S}_T$  中所有期望输出向量采用相同方法求得对应变换矩阵  $\bar{\mathbf{Y}}_T$ 。

(5) 针对  $\bar{\mathbf{X}}$  和  $\bar{\mathbf{Y}}$  组成的样本数据, 采用 NIPALS 算法<sup>[14]</sup> 从矩阵  $\bar{\mathbf{X}}$  中提取前  $k$  个 PLS 成分, 构成  $A \times k$  维隐变量数据矩阵  $\mathbf{T}_k$ , 并求出对应的  $(N_1 + 1) \times k$  维转置矩阵  $\mathbf{W}_k$  和  $N_2 \times k$  维回归系数矩阵  $\mathbf{V}_k$ 。则回归模型如下

$$\tilde{\mathbf{y}} = \bar{\mathbf{x}} \mathbf{W}_k \mathbf{V}_k^T \quad (10)$$

式中  $\bar{\mathbf{x}}$  是回归模型的自变量向量,  $\tilde{\mathbf{y}}$  是回归模型的预测向量。则回归模型式 (10) 对变换后的校验样本  $\bar{\mathbf{X}}_T$  的预测值为

$$\tilde{\mathbf{Y}}_T = \bar{\mathbf{X}}_T \mathbf{W}_k \mathbf{V}_k^T \quad (11)$$

则当隐变量个数为  $k$  时, 对变换后的校验样本  $\bar{\mathbf{Y}}_T$  的预测偏差平方和均值为

$$\text{MSE}_k = \frac{1}{A_T} \sum_{i=1}^{A_T} \sum_{j=1}^{N_2} (\tilde{y}_{T,ij} - \bar{y}_{T,ij})^2 \quad (12)$$

式中  $\bar{y}_{T,ij}$  是  $\bar{\mathbf{Y}}_T$  的第  $i$  向量第  $j$  分量;  $\tilde{y}_{T,ij}$  是  $\bar{y}_{T,ij}$  的预测值。 $k$  依次取  $k = 1, 2, \dots, (N_1 + 1)$ , 并求得对应的  $\text{MSE}_k$ , 则最小的  $\text{MSE}_k$  所对应的  $k$  值为最佳隐变量个数, 设为  $k_{\text{opt}}$ 。

以上是 MLFN-PLSR 算法的执行步骤, 当隐含层输出与输出层期望和输出 PLSR 模型的最佳隐变量个数确定, MLFN-PLSR 模型也就建立。对于一个要预测期望输出的输入向量  $\mathbf{x}_p$ , MLFN-PLSR 模型的应用步骤如下。

(1) 通过式 (7)、式 (8) 求得  $\mathbf{x}_p$  对应的隐含层非线性变换并添加阈值分量后输出向量  $\bar{\mathbf{x}}_p$ 。

(2) 通过式 (10) 求得输出层期望和输出预测值  $\tilde{\mathbf{y}}_p$ , 其中隐变量个数取为  $k_{\text{opt}}$ 。

(3) 通过输出层神经元激活函数 [即公式 (2)] 求得期望输出向量预测值  $\hat{\mathbf{y}}_{p2}$ , 即  $\hat{y}_{p2i} = f(\tilde{y}_{pi})$ ,  $i = 1, 2, \dots, N_2$ , 其中  $\hat{y}_{p2i}$  是  $\hat{\mathbf{y}}_{p2}$  的第  $i$  分量,  $\tilde{y}_{pi}$  是  $\tilde{\mathbf{y}}_p$  的第  $i$  分量。

### 3 4-CBA 含量关联模型及性能分析

#### 3.1 试验设计

将以上分析的影响 TA 中 4-CBA 含量  $C_{4\text{-CBA}}$  的 8 个主要因素 ( $C_{\text{Co}}$ 、 $C_{\text{Mn}}$ 、 $C_{\text{Br}}$ 、 $T_{\text{reactor}}$ 、 $m_{\text{H}_2\text{O}}$ 、 $\tau_{\text{reactor}}$ 、 $T_{\text{crystallizer}}$ 、 $\tau_{\text{crystallizer}}$ ) 作为关联模型的自变

量, TA 中 4-CBA 含量  $C_{4\text{-CBA}}$  作为因变量, 则网络的输入节点个数为 8, 输出节点个数为 1。采集到 197 个具有代表性的样本数据, 随机选取其中 132 个样本作为训练样本数据  $\mathbf{S}$ , 剩余 65 个样本作为校验样本数据  $\mathbf{S}_T$ 。

为了分析 MLFN-PLSR 模型的性能, 设计如下两个试验。

试验 1: 分别采用 MLFN 和 MLFN-PLSR 建立 TA 中 4-CBA 含量关联模型, 其中 MLFN 就是 MLFN-PLSR 建模中第一步获得的模型。MLFN 和 MLFN-PLSR 隐节点个数依次取  $N_1 = 1, 2, \dots, 30$ 。为了减少偶然性影响, MLFN 和 MLFN-PLSR 在每个隐节点个数下执行 10 次建模过程, 各获得 10 个模型。

试验 2: 基于样本数据  $\mathbf{S}$ , 采用 PLSR 建立 TA 中 4-CBA 含量关联模型。

#### 3.2 结果分析

3.2.1 试验 1 结果 试验 1 中, MLFN 模型有 4 次 (隐节点个数分别为 22、27、28 和 30) 出现对校验样本数据  $\mathbf{S}_T$  的预测偏差平方和均值为 0.171049, 严重的偏离 MLFN 模型的平均性能 0.003976; 而所对应的 MLFN-PLSR 方法通过剔除干扰信息, 最终获得模型的预测偏差平方和均值分别为: 0.005331、0.003019、0.003077 和 0.003371, 模型预测性能获得极大的提高。为了减少偶然因素的影响, 以下分析中将剔除这 4 个结果和对应 MLFN-PLSR 结果。

图 2、图 3 分别是 MLFN 和 MLFN-PLSR 在不同隐含层节点个数  $N_1$  下, 各获得 10 (或 9) 个模型对校验样本数据  $\mathbf{S}_T$  的预测偏差平方和均值  $\text{MSE}_{N_1}$  分布图 (在  $N_1 = 1$  时, MLFN 和 MLFN-PLSR 模型均出现  $\text{MSE}_{N_1}$  为 0.012594 和 0.012775, 为了放大图形, 这两个点在图 2、图 3 中均未画出)。图中横坐标表示隐节点个数  $N_1$ , 纵坐标表示模型的预测偏差平方和均值  $\text{MSE}_{N_1}$ 。从图 2 中可以看出, MLFN 模型在隐节点个数比较少时, 模型的  $\text{MSE}_{N_1}$  波动比较小; 随着隐节点个数增多, 模型的  $\text{MSE}_{N_1}$  波动增大。从图 3 中可以看出, MLFN-PLSR 模型的  $\text{MSE}_{N_1}$  稳定在一定范围内, 并不随着隐节点个数增多而变化。因此, MLFN-PLSR 模型对隐节点个数的敏感度低, 模型的鲁棒性强。所有 MLFN 模型  $\text{MSE}_{N_1}$  的平均值为 0.003976, 所有 MLFN-PLSR 模型  $\text{MSE}_{N_1}$  的平

均值为 0.003643，因此，MLFN-PLSR 模型的平均预测精度比 MLFN 模型提高了 8.37%。

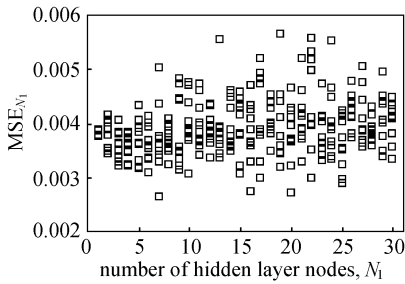


图 2 MLFN 模型的  $MSE_{N_1}$  分布图

Fig. 2 Distribution of MLFN  $MSE_{N_1}$  with  $N_1$

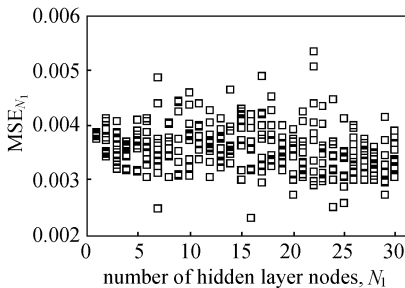


图 3 MLFN-PLSR 模型的  $MSE_{N_1}$  分布图

Fig. 3 Distribution of MLFN-PLSR  $MSE_{N_1}$  with  $N_1$

同时，MLFN-PLSR 模型在隐节点个数  $N_1 = 16$  时， $MSE_{16}^{opt} = 0.002323$  为全部 MLFN-PLSR 模型的最小值，对应的隐变量个数  $k_{opt} = 10$ ；MLFN 模型在隐节点个数  $N_1 = 7$  时， $MSE_7^{opt} = 0.002643$  为全部 MLFN 模型的最小值。因此，MLFN-PLSR 模型的  $MSE_{16}^{opt} = 0.002323$  比 MLFN 模型的  $MSE_7^{opt} = 0.002643$  下降了 12.11%。

图 4 分别是 MLFN 和 MLFN-PLSR 在不同隐含层节点个数  $N_1$  下，各获得 10（或 9）个模型对  $S_T$  的预测偏差平方和均值  $\overline{MSE_{N_1}}$  的平均值  $\overline{MSE_{N_1}}$ ，以及 MLFN 对  $S$  的拟合偏差平方和均值  $\overline{MSE_{N_1}}$  的平均值  $\overline{MSE_{N_1}}$ 。图中左边纵坐标表示模型预测  $\overline{MSE_{N_1}}$ ，右边纵坐标表示 MLFN 训练  $\overline{MSE_{N_1}}$ 。从图 4 中可以看出，随着隐节点个数增加，MLFN 模型的训练精度不断提高，预测精度则先呈急剧下降趋势，后呈上升趋势；而 MLFN-PLSR 模型的预测精度先呈急剧下降趋势，后呈缓慢下降趋势。由此可见，MLFN 隐节点个数太少，则有限的处理能力造成模型性能降低；而隐节点个数太多，则包含干扰信息，模型预测性能也降低。而 MLFN-PLSR 模型可以采用尽量多的隐节点个数，充分提出样本信息，然后通过 PLSR 消除干扰信息，获得

良好预报性能模型。从图 3 和图 4 中可以看出，MLFN-PLSR 隐节点个数取大于 10 均可获得较佳结果。

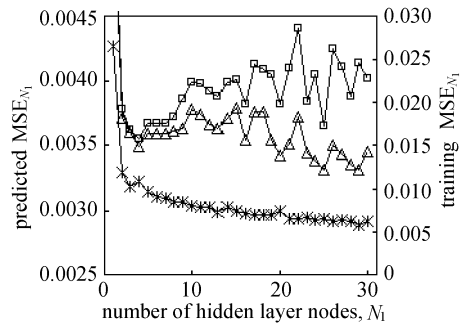


图 4  $\overline{MSE_{N_1}}$  随隐变量个数  $N_1$  变化曲线

Fig. 4 Model  $\overline{MSE_{N_1}}$  versus  $N_1$

□ MLFN; △ MLFN-PLSR; \* MLFN training

3.2.2 试验 2 结果 表 1 中给出采用不同隐变量个数  $k$  时，PLSR 模型对校验样本  $S_T$  的预测偏差平方和均值  $MSE_k$ 。从表 1 中可以看出，模型最佳隐变量个数  $k_{opt} = 8$ ，此时模型对校验样本  $S_T$  的预测偏差平方和均值  $MSE_k = 0.0124$ 。可见，最佳 MLFN-PLSR 模型的预测偏差平方和均值比 PLSR 模型下降了 70.62%，各因素对 4-CBA 含量影响呈高度非线性的特征。

表 1 不同隐变量个数  $k$  下 PLSR 模型的  $MSE_k$

Table 1 PLSR model  $MSE_k$  versus  $k$

$k$	$MSE_k$	$k$	$MSE_k$
1	0.0137	5	0.0127
2	0.0129	6	0.0124
3	0.0130	7	0.0124
4	0.0128	8	0.0124

## 4 结 论

(1) 多层前向型神经网络 (MLFN) 和偏最小二乘回归 (PLSR) 相结合的组合方法融合 MLFN 和 PLSR 方法的优点，并克服各自的缺点。与 MLFN 和 PLSR 相比，MLFN-PLSR 鲁棒性强，模型精度高。

(2) 采用 MLFN-PLSR 建立反应产物 TA 中 4-CBA 含量的关联模型，获得满意结果。

## 符 号 说 明

- A——训练样本个数
- $A_T$ ——校验样本个数
- $C_{4-CBA}$ ——TA 产品中 4-CBA 含量，%

$C_{Co}$ ——反应进料中的钴浓度,  $kg \cdot kg^{-1}$   
 $C_{Br}$ ——反应进料中的溴浓度,  $kg \cdot kg^{-1}$   
 $C_{Mn}$ ——反应进料中的锰浓度,  $kg \cdot kg^{-1}$   
 $E$ ——所有样本的学习误差平方和  
 $E_m$ ——第  $m$  样本的学习误差  
 $e_{m,k}$ ——第  $m$  样本第  $k$  分量的学习误差  
 $f(\cdot)$ ——神经元激活函数  
 $k$ ——隐变量个数  
 $k_{opt}$ ——最佳隐变量个数  
 $L$ ——神经网络层数  
MSE——偏差平方和均值  
 $m_{H_2O}$ ——反应器抽出水流量,  $kg \cdot h^{-1}$   
 $N_j$ ——第  $j$  层神经元个数  
 $n$ ——样本容量  
 $p$ ——自变量维数  
 $q$ ——因变量维数  
 $S$ ——训练样本集  
 $S_{ALL}$ ——总样本集  
 $S_T$ ——校验样本集  
 $T$ ——隐变量数据矩阵  
 $T_{crystallizer}$ ——第一结晶器的温度,  $^{\circ}C$   
 $T_{reactor}$ ——反应器的温度,  $^{\circ}C$   
 $V$ ——回归系数矩阵  
 $W$ ——转置矩阵  
 $\omega$ ——神经元的连接权值或阈值  
 $\bar{X}$ ——由  $\bar{x}_m$  组成的矩阵  
 $X$ ——自变量数据矩阵  
 $x_m$ ——第  $m$  样本向量  
 $\bar{x}_m$ —— $x_m$  经过隐含层变换后的向量  
 $x_{mk}$ —— $x_m$  的第  $k$  分量  
 $Y$ ——因变量数据矩阵  
 $\bar{Y}$ ——由  $\bar{y}_m$  组成的矩阵  
 $Y_{k+1}$ ——残差矩阵  
 $y_m$ ——第  $m$  样本的期望向量  
 $\hat{y}$ ——PLSR 模型的因变量预报值  
 $\bar{y}_m$ ——第  $m$  样本输出层期望和输出向量  
 $\hat{y}_{mj}$ —— $x_m$  作用于网络后, 第  $j$  层神经元输出向量  
 $y_{mk}$ ——第  $m$  样本第  $k$  期望输出分量  
 $\hat{y}_{mjk}$ —— $x_m$  作用于网络后, 第  $j$  层第  $k$  神经元的函数输出  
 $z_{mji}$ —— $x_m$  作用于网络后, 第  $j$  层第  $i$  神经元的和输出  
 $\tau_{crystallizer}$ ——第一结晶器的停留时间, s  
 $\tau_{reactor}$ ——氧化反应器的停留时间, s

catalytic oxidation of *p*-xylene. *Ind. Eng. Chem. Res.*, 1992, **31** (2): 453-462

[3] Yan Xuefeng, Du Wenli, Qian Feng. Development of a kinetic model for industrial oxidation of *p*-xylene by RBF-PLS and CGA. *AIChE J.*, 2004, **50** (6): 1169-1176

[4] Wang Qinbo, Li Xi, Wang Lijun, Cheng Youwei, Xie Gang. Kinetics of *p*-xylene liquid-phase catalytic oxidation to terephthalic acid. *Ind. Eng. Chem. Res.*, 2005, **44** (2): 261-266

[5] Wang Qinbo, Li Xi, Wang Lijun, Cheng Youwei, Xie Gang. Effect of water content on the kinetics of *p*-xylene liquid-phase catalytic oxidation to terephthalic acid. *Ind. Eng. Chem. Res.*, 2005, **44** (13): 4518-4522

[6] Yan Xuefeng (颜学峰), Yu Juan (余娟), Qian Feng (钱锋). An evolution algorithm with select-best and prepotency operator and parameter estimation of 4-CBA model. *Journal of Chemical Engineering of Chinese Universities* (高校化学工程学报), 2005, **19** (2): 238-243

[7] Sun Jingmin (孙静民). Polyester Technics (聚酯工艺). Beijing: Chemical Industry Press, 1985

[8] Peng Qianrong (彭黔荣), Yang Min (杨敏), Shi Yanfu (石炎福), Yu Huarui (余华瑞), Liu Zhongxiang (刘钟祥). Artificial neural network based on hybrid genetic algorithm and prediction of melting points of organic compounds. *Journal of Chemical Industry and Engineering (China)* (化工学报), 2005, **56** (10): 1922-1927

[9] E Jiaqiang, Wang Yaonan, Mei Chi. Soft-sensing model of copper liquid temperature in copper refining process and its application. *Journal of Chemical Industry and Engineering (China)*(化工学报), 2006, **57** (1): 203-209

[10] Zhou Yunlong (周云龙), Sun Bin (孙斌), Lu Jun (陆军). Application of improved BP neural network in identification of air-water two-phase flow patterns. *Journal of Chemical Industry and Engineering (China)*(化工学报), 2005, **56** (1): 110-115

[11] Zhao Weixiang (赵伟祥). Novel methods for establishing the reaction model of complex material system based on observation data [D]. Hangzhou: Zhejiang University, 2000

[12] Yan Xuefeng (颜学峰), Yu Juan (余娟), Qian Feng (钱锋). Development of a naphtha dry point soft sensor by adaptive partial least square regression. *Journal of Chemical Industry and Engineering (China)*(化工学报), 2005, **56** (8): 1511-1515

[13] Chen Dezhaoh (陈德钊). Multivariate processing (多元数据处理). Beijing: Chemical Industry Press, 1998

[14] Geladi P, Kowalski B R. Partial least-squares regression: a tutorial. *Analytical Chemical Acta*, 1986, **185** (1): 1-17

[15] Liu Fang (刘芳), Li Renhou (李人厚). The evolving artificial neural network based on genetic algorithm. *Journal of System Simulation* (系统仿真学报), 2003, **15** (10): 1431-1433

References

[1] Prengle Jr H W, Barond N. Make petrochemicals by liquid phase oxidation. *Hydrocarbon Process*, 1970, **49** (3): 106-118

[2] Raghavendrachar P, Ramachandran S. Liquid-phase