# 原油总碳含量的粒子群优化集成神经网络预测模型

贺婷婷,陆 军,丁进良<sup>†</sup>,刘长鑫

(东北大学 流程工业综合自动化国家重点实验室, 辽宁 沈阳 110819)

摘要:原油评价新技术的研究和应用成为目前世界石油炼制企业致力发展的方向,也是今后发展的必然趋势.本 文采用核磁共振(nuclear magnetic resonance, NMR)光谱技术和粒子群优化集成神经网络(particle swarm optimiza tion-ensemble neural network, PSO-ERNN)建立了一种快速评价原油总碳物性指标预测模型.该模型以随机向量函 数连接网络(random vector functional link network, RVFL)作为基本模型,采用正则化负相关学习策略集成基本模型, 并采用粒子群优化算法优化各基本模型的最优隐含层节点数(*L*)以及集成规模的最佳集成个数(*M*),最后利用在线 学习方法对模型进行更新.实例验证表明,所提出的模型显著提高了预报精度,避免了随机选择*L*和*M*对模型精度 的影响,对提高原油评价精度与效率和及时满足加工炼制要求具有应用价值.

关键词: 原油总碳含量; 预测模型; 粒子群; 集成学习; 神经网络; 核磁共振

引用格式: 贺婷婷, 陆军, 丁进良, 等. 原油总碳含量的粒子群优化集成神经网络预测模型. 控制理论与应用, 2019. 36(2): 192-198

中图分类号: TP183 文献标识码: A

DOI: 10.7641/CTA.2018.70779

## Prediction model of total carbon content of crude oil using ensemble random weights neural network optimized by particle swarm optimization

#### HE Ting-ting, LU Jun, DING Jin-liang<sup>†</sup>, LIU Chang-xin

(State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang Liaoning 110819, China)

Abstract: The research and application of the new technology of crude oil evaluation has become the trend of the world petroleum refining enterprises, and it is also the inevitable trend of development in the future. In this paper, a new method to rapidly evaluate total carbon of crude oil was established by using nuclear magnetic resonance (NMR) spectroscopy and ensemble neural network with random weights (ERNN) model which is optimized by particle swarm. The model uses the random vector functional link network (RVFL) as the basic model. A regularized negative correlation learning strategy is used to integrate the basic model. The particle swarm optimization (PSO) algorithm is used to optimize the hidden layer number (L) of each basic model and the number of integrated scale (M). Finally, the online learning method is used to update the model. Experiment results show that the proposed model significantly improves the prediction accuracy and efficiency of crude oil evaluation and it has a practical value to meet the requirements of oil refining in time.

**Key words:** total carbon of crude oil; prediction model; PSO; ensemble learning; neural network; nuclear magnetic resonance (NMR)

**Citation:** HE Tingting, LU Jun, DING Jinliang, et al. Prediction model of total carbon content of crude oil using ensemble random weights neural network optimized by particle swarm optimization. *Control Theory & Applications*, 2019, 36(2): 192 – 198

### 1 引言

原油是烷烃、环烷烃、芳香烃和烯烃等多种液态 烃的混合物,主要成分是碳和氢两种元素,分别占 83%~87%和11%~14%.因此总碳是原油物性中重要 的指标,是评价原油品质的主要物性之一.随着世界 经济的快速发展,国内的炼油和石化企业加工产量不 断增加并且加工难度不断增大.为了降低成本,企业 采用原油调和技术,对多种原油进行合理混炼.因而,

收稿日期: 2017-10-26; 录用日期: 2018-05-07.

<sup>&</sup>lt;sup>†</sup>通信作者. E-mail: jlding@mail.neu.edu.cn; Tel.: +86 24-83684245.

本文责任编委: 李少远.

国家自然科学基金项目(61590922, 61525302, 61621004)资助.

Supported by the National Natural Science Foundation of China (61590922, 61525302, 61621004).

在炼油前对单一或者混兑原油的性质进行快速分析, 即原油物性快速分析评价日益得到重视<sup>[1-2]</sup>.

原油评价的核心任务就是在实验室通过仪器分析 技术对原油及馏分油的物理性质、化学性质、烃类和 非烃类化合物的组成进行分析,根据得到的数据对原 油性质、加工性质以及炼制性能进行综合评价,为炼 厂常减压和二次加工装置提供基础数据和加工方案. 目前常见的原油物性检测技术主要有近红外光谱 (near infrared, NIR)<sup>[3-4]</sup>、中红外光谱(middle infrared, MIR)<sup>[5]</sup>、核磁共振波谱 (nuclear magnetic resonance, NMR)<sup>[6]</sup>及其他一些分析方法.

近红外最早应用于汽油辛烷值的测定. 陈瀑、褚 小立以及邱腾等人<sup>[7–9]</sup>在NIR光谱的基础上, 利用偏 最小二乘法(partial least squares, PLS)方法对原油诸 多物性进行了预报; 栾郭宏等人<sup>[10]</sup>运用人工神经网络 建立了辛烷值的近红外光谱模型. 由于近红外光谱采 用拍照片法, 对原油质量要求较高, 需要较长的光程, 灵敏度相对较低, 限制了其使用的可扩展性.

随着计算机技术的飞速发展,在线核磁共振分析 系统变得更完善,成为当前最先进和最有前途的过程 分析技术之一. NMR与先进的控制技术结合,可明显 地提高工业生产效率.目前很多国际大型石油公司都 已采用了在线NMR分析系统<sup>[11]</sup>.核磁共振技术在应 用于原油快速评价中,样品的纯度、粘度以及是否含 有气泡等因素不会影响测试的精度,因此具有很好的 应用前景.

但是目前基于核磁共振光谱数据与物性建模的方 法并不完善,大都是采用PLS<sup>[12]</sup>、人工神经网络<sup>[10]</sup>等 建立模型,然而,由于PLS是线性模型,在处理非线性 问题上不占优势<sup>[13]</sup>.传统的人工神经网络算法诸如后 向传播算法(back-propagation, BP)等收敛速度慢而且 容易陷入局部最优.因此精准的预测模型是原油快速 评价急需解决的问题.随机权神经网络以其学习速度 快以及良好的泛化性等优势广泛应用于许多实际问 题的解决,经典的集成学习算法如随机森林算法、 bagging以及boosting等通过使用不同的训练数据来获 得模型的多样性<sup>[14-15]</sup>.

针对上述问题,本文提出了一种基于粒子群优化 集成随机权神经网络的原油总碳物性预报模型.集成 学习通过构建并结合多个学习器来完成学习任务,常 可获得比单一学习器更好的的泛化性能,而且集成学 习考虑到集成模型中个体模型间的相关性和模型的 复杂性.在线学习策略改进了直接在集成模型的基础 上更新学习的传统方式,而是以个体网络为基础实现 在线学习更新模型.同时本文采用粒子群优化算法对 随机向量函数链接网络模型中最优隐含层节点数(L) 以及集成规模中最佳集成个数(*M*)进行优化选择,避免了随机选择带来的不确定性.通过实际炼厂数据进行了实验验证,实验结果表明本文模型能准确有效地预报原油总碳含量.

### 2 核磁共振氢谱的获取

采用Aspect Imaging AI-60在线核磁共振分析仪 进行原油物性分析,该核磁共振分析系统包括:单通 道NMR主分析仪器、工业PC机、通讯软件和用户界 面软件包.图1为NMR原油快速评价系统的示意图.



图 1 ASPECT在线NMR原油快速评价系统示意图 Fig. 1 ASPECT online NMR crude oil rapid evaluation system

原油样品经过简单过滤和加热预处理,然后通氮 气动力将样品罐中样品注入单通道NMR主分析仪器, 通信系统将分析结果上传至工业PC机,经PC机处理 以及快速傅里叶变换,进而得到原油NMR图谱.其中 NMR图谱是以四甲基硅烷(TMS)为化学位移参比物. 原油NMR图谱如图2所示.



Fig. 2 NMR spectrum of crude oil

### 3 原油总碳含量预测模型

本文提出的原油物性预报模型的建模过程主要由两大部分:集成随机权神经网络 (ensemble neural network with random weights, ERNN) 建模,粒子群优化 参数.预测模型的输入数据集为核磁共振图谱,输出 为原油重要物性总碳含量,对应的数据集为

 $D = \{ (X_{\text{spec}}^n, Y_{\text{spec}}^n) \}_{n=1}^N.$ 

首先通过NMR分析仪对原油样品进行分析,得到 NMR图谱以及相应的实验室化验物性数据.利用上 述数据对集成模型(ERNN)进行训练,同时采用粒子 群优化方法(particle swarm optimization, PSO)优化集 成网络的集成规模和隐层节点个数.当模型参数选定 后,对新采集的原油核磁共振图谱数据,利用本文所 建模型进行建模分析,可得到原油物性指标数据.本 文模型结构如图3所示.





Fig. 3 Structure of PSO-ERNN online learning model

#### 3.1 集成随机权神经网络(ERNN)预测模型

在集成网络模型中每个个体网络均是相同的单隐 藏层前馈神经网络,其表示为

$$f(X_{\text{spec}}^n) = \sum_{j=1}^L \beta_j G(\omega_j^{\text{T}} \cdot X_{\text{spec}}^n + b_j), \qquad (1)$$

其中:  $\beta = [\beta_1 \ \beta_2 \cdots \beta_L]$ 表示输出权值; *L*表示隐藏 层节点的数目;  $X_{\text{spec}}^n = [x_{n1} \ x_{n2} \cdots x_{n700}]$ 表示第*n* 个原油图谱数据作为集成模型的每个个体网络模型 的输入变量;  $\omega_j$ 表示输入权值,  $b_j \in \mathbb{R}$ 表示隐藏层节点 阀值;  $G(\cdot)$ 表示隐藏层节点激励函数. 通过最小二乘 方法求得输出权值 $\beta$ .

传统的ERNN各模型的最优隐含层节点数L以及 最佳集成规模M一般为随机选择,这种方法不但操作 事倍功半,而且对实验的精度有很大的影响.因此本 文提出了基于粒子群优化正则化负相关集成随机权 神经网络的在线学习方法.此模型不但采用粒子群离 线优化模型中的L和M,而且还结合神经网络集成学 习<sup>[16]</sup>、负相关(negative correlation learning, NCL)学 习方法<sup>[17]</sup>,以及为了避免传统神经网络模型过拟合而 引入附加正则化项的方法. Hansen等人<sup>[15]</sup>提出的集 成学习是将*M*个基本神经网络*f<sub>m</sub>*(*X*<sup>n</sup><sub>spec</sub>)组合成集成 神经网络

$$f_{\rm ens}(X_{\rm spec}^n) = \omega_m \sum_{m=1}^M f_m(X_{\rm spec}^n).$$
(2)

考虑到上述集成学习中各基本模型之间的交互性,因此本文引入了集成负相关学习方法,每个基本模型 $f_m(X_{\text{spec}}^n)$ 的代价函数 $e_m$ 定义为

$$e_m = \frac{1}{2} \sum_{n=1}^{N} \left( f_m(X_{\text{spec}}^n) - Y_{\text{spec}}^n \right)^2 + \lambda p_i, \quad (3)$$

其中: $p_i$ 为相关性惩罚项, $\lambda$ 为惩罚项参数.

通过引入正则化项能避免模型过拟合,建立正则 化负相关学习集成神经网络模型.其中个体模型的代 价函数可以表示为

$$e_{m} = \frac{1}{2} \sum_{n=1}^{N} (f_{m}(X_{\text{spec}}^{n}) - Y_{\text{pct}}^{n})^{2} - \lambda \sum_{n=1}^{N} (f_{m}(X_{\text{spec}}^{n}) - f_{\text{ens}}(X_{\text{spec}}^{n}))^{2} + \alpha_{m} \|\beta_{m}\|_{2}^{2} = \frac{1}{2} \|H_{m}\beta_{m} - Y_{\text{pct}}^{\text{target}}\|_{2}^{2} + \alpha_{m} \|\beta_{m}\|_{2}^{2} - \lambda \|H_{m}\beta_{m} - \frac{1}{M} \sum_{j=1}^{M} H_{j}\beta_{j}\|_{2}^{2}, \qquad (4)$$

其中:  $H_m$ 表示第m个基本神经网络模型的隐藏层输 出矩阵,  $Y_{\text{pct}}^{\text{target}} = [Y_{\text{pct}}^1 Y_{\text{pct}}^2 \cdots Y_{\text{pct}}^N]^{\text{T}}$ 表示目标矩 阵,  $\beta_m = [\beta_{m1} \cdots \beta_{mL}]^{\text{T}}$ 为基本模型m的输出权值 向量,  $\lambda$ 为惩罚项参数,  $\alpha_m$ 为正则化参数. 当代价函数  $e_m$ 对 $\beta_m$ 的导数为零时, 集成模型有最好的泛化性能, 由 $\frac{\partial e_m}{\partial \alpha} = 0$ 化简可得

$$\partial \beta_m = O(I) H_0^{\rm T} H_0^{\rm T}$$
  
 $H_m^{\rm T} Y_{\rm pct}^{\rm target} = (C_1 H_m^{\rm T} H_m + \alpha_m I) \beta_m + \alpha_m I$ 

$$C_2 \sum_{j=1, j \neq m}^{M} H_m^{\mathrm{T}} H_j \beta_j, \qquad (5)$$

其中:参数C<sub>1</sub>和C<sub>2</sub>表示正则化参数λ和集成规模参数 M之间的关系表达式,C<sub>1</sub>和C<sub>2</sub>可以表示为

$$C_1 = 1 - 2\lambda (1 - \frac{1}{M})^2, \tag{6}$$

$$C_2 = 2\frac{\lambda}{M}\left(1 - \frac{1}{M}\right).\tag{7}$$

由参数*C*<sub>1</sub>和*C*<sub>2</sub>可知,式(5)可以表示为线性等式,即隐藏层到输出层可以表示成如下矩阵形式:

$$H_{\rm corr}\beta_{\rm ens} = T_{\rm h},$$
 (8)

其中: βens为全局输出权值, Hcorr称为隐藏层相关矩

#### 阵, Th称为目标矩阵,

$$H_{\text{corr}} = \begin{bmatrix} C_1 H_1^{\mathrm{T}} H_1 + 2\alpha_1 & \cdots & C_2 H_1^{\mathrm{T}} H_M \\ C_2 H_2^{\mathrm{T}} H_1 & \cdots & C_2 H_2^{\mathrm{T}} H_M \\ \vdots & \ddots & \vdots \\ C_2 H_M^{\mathrm{T}} H_1 & \cdots & C_1 H_M^{\mathrm{T}} H_M + 2\alpha_M \end{bmatrix},$$
(9)

$$T_{\rm h} = \begin{bmatrix} H_1^{\rm T} Y_{\rm pct}^{\rm target} & \cdots & H_M^{\rm T} Y_{\rm pct}^{\rm target} \end{bmatrix}_{ML \times 1}^{\rm T}, \quad (10)$$

其中H为隐藏层输出矩阵:

$$H = \begin{bmatrix} G(X_{\text{spec}}^{1}; w_{1}, b_{1}) \cdots G(X_{\text{spec}}^{1}; w_{L}, b_{L}) \\ \vdots & \ddots & \vdots \\ G(X_{\text{spec}}^{N}; w_{1}, b_{1}) \cdots G(X_{\text{spec}}^{N}; w_{L}, b_{L}) \end{bmatrix}_{N \times L}$$
(11)

通过最小二乘法求得全局输出权值为

$$\beta_{\rm ens} = H_{\rm corr}^{\dagger} T_{\rm h}.$$
 (12)

传统机器学习中分类和回归算法中大都是在批学 习模式下设计,也就是通过训练所给样本,一次学习 得到最终预报模型,但是在实际应用中,数据是在不 断地更新,因此本文采用了在线学习<sup>[18]</sup>.令输入输出 集表示为{(X<sub>spec</sub>,Y<sub>pct</sub>)<sup>old</sup>,(X<sub>spec</sub>,Y<sub>pct</sub>)<sup>new</sup>},根据基 于正则化负相关的集成随机权神经网络的隐藏层相 关矩阵,求取新的集成相关矩阵

$$H_{\text{corr}} = \begin{bmatrix} C_1 H_1^{\text{old}^{\mathrm{T}}} H_1^{\text{old}} + 2\alpha \cdots & C_2 H_1^{\text{old}^{\mathrm{T}}} H_M^{\text{old}} \\ \vdots & \ddots & \vdots \\ C_2 H_M^{\text{old}^{\mathrm{T}}} H_1^{\text{old}} & \cdots & C_1 H_M^{\text{old}^{\mathrm{T}}} H_M^{\text{old}} + 2\alpha_M \end{bmatrix} + \begin{bmatrix} C_1 H_1^{\text{new}^{\mathrm{T}}} H_1^{\text{new}} + 2\alpha \cdots & C_2 H_1^{\text{new}^{\mathrm{T}}} H_M^{\text{new}} \\ \vdots & \ddots & \vdots \\ C_2 H_M^{\text{new}^{\mathrm{T}}} H_1^{\text{new}} & \cdots & C_1 H_M^{\text{new}^{\mathrm{T}}} H_M^{\text{new}} + 2\alpha_M \end{bmatrix} .$$

$$(13)$$

根据式(13)以及式(9)-(10)可得,在没有接收到新数据时,求出隐藏相关矩阵和目标矩阵;当有新的数据时,只需要利用该数据求得相应的集成模型的隐藏 层相关矩阵和目标矩阵,让其与历史隐藏相关矩阵和 目标矩阵求和,便可在更新所有数据的隐藏层相关矩 阵和目标矩阵基础上,求得集成模型的全局输出权值, 实现对模型进行更新,即

$$H_{\rm corr} = H_{\rm corr}^{\rm old} + H_{\rm corr}^{\rm new}, \tag{14}$$

$$T_{\rm h} = T_{\rm h}^{\rm old} + T_{\rm h}^{\rm new}.$$
 (15)

#### 3.2 粒子群优化算法

本文采用启发式粒子群算法对集成随机权神经网

络中基本模型的最优隐含层节点数L以及集成规模的 最佳集成个数M进行离线优化.首先需要设置最大的 速度空间,并在速度空间和搜索空间上随机初始化速 度和位置,定义群体规模为m,V<sub>L</sub>和V<sub>M</sub>分别为优化参 数L和M的粒子初始化速度,X<sub>L</sub>和X<sub>M</sub>分别为优和M 的粒子初始化位置,P<sup>L</sup><sub>best</sub>和P<sup>M</sup><sub>best</sub>分别为L和M的当前 粒子最优适应度,G<sup>L</sup><sub>best</sub>和G<sup>M</sup><sub>best</sub>分别为L和M的种群 最优适应度.在优化参数L和M的过程中,粒子群中 的所有粒子根据自己找到的当前种群极值P<sup>L</sup><sub>best</sub>和 P<sup>M</sup><sub>best</sub>和整体粒子群共享的当前全局最优解G<sup>L</sup><sub>best</sub>和 G<sup>M</sup><sub>best</sub>来调整自己的速度和位置.迭代过程中粒子速度 的更新方式为

$$\begin{bmatrix} V_{L,i} \\ V_{M,i} \end{bmatrix} = \omega_1 \begin{bmatrix} V_{L,i} \\ V_{M,i} \end{bmatrix} + C_1 \cdot \operatorname{rand} \cdot \left( \begin{bmatrix} P_{\text{best}}^L \\ P_{\text{best}}^M \end{bmatrix} - \begin{bmatrix} X_{L,i} \\ X_{M,i} \end{bmatrix} \right) + C_2 \cdot \operatorname{rand} \cdot \left( \begin{bmatrix} G_{\text{best}}^L \\ G_{\text{best}}^M \end{bmatrix} - \begin{bmatrix} X_{L,i} \\ X_{M,i} \end{bmatrix} \right).$$
(16)

粒子位置的更新方式为

$$\begin{bmatrix} X_{L,i} \\ X_{M,i} \end{bmatrix} = \begin{bmatrix} X_{L,i} \\ X_{M,i} \end{bmatrix} + \omega_2 \begin{bmatrix} V_{L,i} \\ V_{M,i} \end{bmatrix}, \quad (17)$$

其中:  $\omega_1 \pi \omega_2$ 表示惯性因子,  $C_1 \pi C_2$ 表示加速常数, 一般取 $C_1 = C_2 \in [0, 4]$ , rand表示区间[0, 1]上的随 机数.

本文采用集成随机权神经网络的评价指标均方根 误差 (root mean square error, RMSE) 作为目标函数, 其表达式为

RMSE =  

$$\sqrt{\frac{1}{N} \sum_{n=1}^{N} \left(\frac{1}{M} \sum_{M=1}^{M} f_m(X_{\text{spec}}^n) - Y_{\text{pct}}^{\text{target}}\right)^2}.$$
(18)

#### 4 实验结果与分析

#### 4.1 实验设置

模型输入数据为:现有原油NMR分析图谱数据. 模型输出数据为:与图谱相对应的物性总碳含量. 输入数据和输出数据479组均取自某炼厂的数据库平 台,训练数据为399组,测试数据为80组.

模型参数选择为: 粒子群优化算法中迭代次数为 200, 种群规模为30, 经优化后, 集成随机权神经网络 的基本模型RVFL隐藏层节点为L = 50, 集成规模为 M = 7, 对于RVFL激励函数采用反对称的tanh函数, 并且函数中的参数使用遗传算法进行优化, 结果为a = 1.716, b = 0.666.

#### 实验结果对比分析 4.2

采用本文提出的粒子群优化集成随机权神经网络 (particle swarm optimization-ensemble neural network, PSO-ERNN)模型,对原油物性中的总碳含量(质量百 分比, w%) 进行仿真实验, 并与在线随机权神经网 络(online-decorrelated neural network ensembles, DNNE)<sup>[18]</sup>、偏最小二乘 (partial least square, PLS)<sup>[19]</sup> 以及人工选择集成规模和隐藏层节点个数两个参数 的随机权神经网络(ERNN)建立的预测模型进行对比.

图4和图5分别是本文模型预报结果曲线和模型预 报误差绝对值曲线.图6-9分别为本文所建预测模型 与PLS, Online-DNNE, ERNN以及基于粒子群优化的 离线集成随机权神经网络 (offline-particle swarm optimization-ensemble neural network with random weights, offline-PSO-ERNN)ERNN模型的对比曲线. 图10为本文模型以及上述4种模型的预测值与实际值 的误差对比曲线.

从图4和图5可以看出,本文所建预测模型可以较 好的拟合原油总碳数据,预报误差较小.由图6可以看 出,相对于传统的线性PLS模型,本文所建预测模型预 报效果更好.



图 4 本文模型预测值与总碳真实值对比 Fig. 4 Comparison of the prediction results of the proposed model and the actual values of total carbon content











由图7和图8可以看出,与典型的非线性模型即 ERNN和Online-DNNE相比,本文模型同样优于其他 两种模型. 图9表明, 与离线 PSO-ERNN 相比, 在线 PSO-ERNN模型可以学习新增数据信息,从而实现模 型的更新,提高模型预测精度.



图 7 本文模型与Online-DNNE模型预测结果曲线

Fig. 7 Comparison of prediction results of PSO-ERNN and Online-DNNE



图 8 本文模型与ERNN模型预测结果曲线

Fig. 8 Comparison of prediction results of PSO-ERNN and **ERNN** 





从图10也可以看出,本文模型的预报误差绝对值 在整体上小于PLS, ERNN, Offline-PSO-ERNN以及 Online-DNNE.为进一步说明基于粒子群优化集成神 经网络参数所建预测模型的有效性,统计文中包含 的5个模型的均方根误差(root mean square error, RMSE)、平均百分比误差(mean absolute percentage error, MAPE)以及最大绝对误差(maximum absolute error, MAE),如表1所示.由表1可以看出,对于表中统 计的5种模型评价指标,本文模型均好于PLS, ERNN, Online-DNNE以及离线PSO-ERNN.



图 10 本文模型与Offline-PSO-ERNN, ERNN, PLS以及 Online-DNNE预测误差结果曲线

Fig. 10 Comparison of prediction errors of PSO-ERNN, Offline-PSO-ERNN, ERNN, PLS and Online-DNNE

衣	1 本又模型与现有模型预报结果对比
Table 1	Results comparison between the proposed
	algorithm and the existing algorithms

十二世间上四十世间元旧从田山山

模型	RMSE	MAPE	MAE
PLS	0.29862	0.25697	1.17959
Online-DNNE	0.25973	0.23735	0.66389
Offline-PSO-ERNN	0.22331	0.16572	0.59235
ERNN	0.15866	0.13121	0.53216
本文模型	0.13319	0.11981	0.48284

通过以上分析可知,对于原油物性指标快速评价 建模,采用非线性算法建立的预报模型的效果比传统 线性算法建立的模型效果好,而集成神经网络强学习 算法建立的模型比传统的弱学习算法建模效果好,通 过PSO优化集成网络参数比经验选择效果好,在线学 习比离线学习效果好,证明了本文所提算法的有效性.

#### 5 结论

原油评价得到的数据为炼厂常减压和二次加工装 置提供数据和加工方案,是石油炼制工业的重要环节. 本文采用粒子群算法优化集成随机权神经网络中最 优隐含层节点数L以及集成规模中最佳集成个数M, 建立可靠有效的原油物性总碳质量预报模型.本文与 偏最小二乘模型、随机选择集成参数的集成随机权神 经网络模型、离线集成随机权神经网络以及在线去相 关神经网络模型进行比较,实验结果表明本文方法提 高了建模预报的准确性,可以为原油加工炼制过程提 供数据支撑.

#### 参考文献:

- BARUNIK J, MANLINSKA B. Forecasting the term structure of crude oil futures prices with neural networks. *Applied Energy*, 2016, 164(2): 366 – 379.
- [2] CHATRATH A, MIAO H, RAMCHANDER S. An examination of the flow characteristics of crude oil: evidence from risk-neutral moments. *Energy Economics*, 2016, 54(1): 213 – 223.
- [3] HE K X, QIAN F, CHENG H, et al. A novel adaptive algorithm with near-infrared spectroscopy and its application in online gasoline blending processes. *Chemometrics and Intelligent Laboratory Systems*, 2015, 140(1): 117 – 125.
- [4] QIU Teng, YUAN Hongfu, SUN Jia, et al. The rapid evaluation of crude petroleum by near infrared spectroscopy (NIR). Spectroscopy and Spectral Analysis, 2008, 28(10): 131 132.
  (邱藤, 袁洪福, 孙甲, 等. 近红外光谱快速原油评价研究. 光谱学与 光谱分析, 2008, 28(10): 131 132.)
- [5] LI Danting, ZHANG Changjiang, WANG Jin. Infrared spectrum (IR) analysis based on wavelet transform. *Spectroscopy and Spectral Analysis*, 2006, 26(11): 2024 – 2026.
  (李丹婷,张长江,汪劲,等. 基于小波变换的红外光谱分析. 光谱学 与光谱分析, 2006, 26(11): 2024 – 2026.)
- [6] DANIEL M, URIBE U, MURGICH J. Correlations between SARA fractions and physicochemical properties with 1HNMR spectra of vacuum residues from Colombian crude oils. *Fuel*, 2010, 89(1): 185 – 192.
- [7] CHEN Pu, CHU Xiaoli. Development of rapid analytical technologies in crude and heavy oil. *Journal of Instrumental Analysis*, 2012, 31(9): 1191-1198.
  (陈瀑,褚小立. 原油及重油的快速分析技术进展. 分析测试学报, 2012, 31(9): 1191-1198.)
- [8] CHU Xiaoli, TIAN Songbai, XU Yupeng, et al. A study of crude oil rapid assay by near infrared spectroscopy. *Petroleum Procession and Petrochemicals*, 2012, 43(1): 72 77.
  (褚小立,田松柏,许育鹏,等. 近红外光谱用于原油快速评价的研究. 石油炼制与化工, 2012, 43(1): 72 – 77.)
- [9] LUAN Guohong, HE Kaixun, CHENG Hui, et al. Octane model based on neural network by near-infrared spectroscopy and its application. *Computers and Applied Chemistry*, 2014, 31(1): 63 – 68.

(栾郭宏, 贺凯迅, 程辉, 等. 基于神经网络的近红外光谱辛烷值模型的研究及应用. 计算机与应用化学, 2014, 31(1): 63 – 68.)

- [10] LIN Limin, CHEN Jian, JIN Jiajian. Nuclear magnetic resonance online analysis technology and its application in petrochemical unit. Automation in Petro-chemical Industry, 2004, 40(3): 55 – 59.
  (林立敏, 陈建, 金加剑. 核磁共振在线分析技术及其在炼油和化工 装置中的应用. 石油化工自动化, 2004, 40(3): 55 – 59.)
- [11] BREIMAN L. Bagging predictors. *Maching Learning*, 1996, 24(2): 123 – 140.
- [12] MOLINA V, URIBE U, MURGICH J. Partial least-squares (PLS) correlation between refined product yields and physicochemical properties with the 1H nuclear magnetic resonance (NMR) spectra of colombian crude oils. *Energy and Fuels*, 2007, 21(3): 1674 – 1680.
- [13] WANG Wei, CHAI Tianyou, ZHAO Lijie. Dynamic partial least squares modeling with recurrent neural networks of stable learning. *Control Theory & Applications*, 2012, 29(3): 337 341.
  (王魏, 柴天佑, 赵立杰. 带有稳定学习的递归神经网络动态偏最小 二乘建模. 控制理论与应用, 2012, 29(3): 337 341.)
- [14] CHENG W J, DING J L, KONG W J, et al. An adaptive chaotic PSO for parameter optimization and feature extraction of LS–SVM based modeling. 2011 American Control Conference. Rochester, NY, USA: IEEE, 2011: 3263 – 3268.
- [15] HANSEN L K, SALAMON P. Neural network ensembles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, 12(10): 993 1001.
- [16] LIU Y, YAO X. Ensemble learning via negative correlation. *Neural Networks*, 1999, 12(10): 1399 1404.

- [17] CHEN H H, YAO X. Regularized negative correlation learning for neural network ensembles. *IEEE Transactions on Neural Networks*, 2009, 20(12): 1962 – 1979.
- [18] DING J L, WANG H T, LI CH B, et al. An online learning neural network ensembles with random weights for regression of sequential data stream. *Soft Computing*, 2017, 21(20): 5919 – 5937.
- [19] MASILI A, PULIGHEDDU S, SASSU L, et al. Prediction of physical-chemical properties of crude oils by <sup>1</sup>H<sub>O</sub> NMR analysis of neat samples and chemometrics. *Magnetic Resonance in Chemistry*, 2012, 50(11): 729 – 738.

#### 作者简介:

**贺婷婷**硕士研究生,主要研究方向为复杂工业过程智能建模, E-mail: 2657485147@qq.com;

陆 军 博士研究生,主要研究方向为复杂工业过程智能建模,

E-mail: junluchn@hotmail.com;

**丁进良** 教授,博士生导师,主要研究方向为复杂工业过程智能建模与智能优化与控制、生产全流程运行优化、计算智能及其应用, E-mail: jlding@mail.neu.edu.cn;

**刘长鑫** 讲师,博士研究生,主要研究方向为基于大数据的复杂工业过程建模、优化结构与方法及其应用研究,E-mail: cxliu@mail.neu.edu.cn.