

文章编号: 2095-2163(2019)05-0273-04

中图分类号: TP181

文献标志码: A

# 基于天然气脱碳工艺的溶液循环量组合预测模型

官莉萍, 曩莹

(西安石油大学 计算机学院, 西安 710065)

**摘要:** 在天然气脱碳工艺工业生产中, 优化脱碳工艺流程可以有效提高生产效益。文章首先对脱碳工艺参数进行分析, 采用 Shapley 值法对单预测模型进行分配权重, 构建组合预测模型。实验对比表明, 基于 Shapley 值的组合预测模型可以有效提高预测模型的稳定性, 在优化脱碳工艺参数中具有一定的参考价值。

**关键词:** 天然气脱碳; Shapley 值; 组合预测

## Combined prediction model of solution circulation based on natural gas decarburization process

GUAN Liping, CUAN Ying

(School of Computer Science, Xi'an Shiyou University, Xi'an 710065, China)

**[Abstract]** In natural gas decarburization process industry production, optimization of decarburization process flow can effectively improve production efficiency. Firstly, the decarburization process parameters are analyzed, and the Shapley value method is used to assign weights to the single prediction model and build the combined prediction model. Experimental comparison shows that the combined prediction model based on Shapley value can effectively improve the stability of the prediction model, and has certain reference value in the optimization of decarburization process parameters.

**[Key words]** natural gas decarburization, Shapley value; combination prediction

## 0 引言

随着低碳经济的发展, 作为一种清洁能源, 天然气已经成为非常普遍且广泛消耗的清洁能源<sup>[1]</sup>。天然气中含有的 CO<sub>2</sub> 杂质严重影响天然气的品质, 甚至带来安全隐患, 危害环境, 威胁人体健康。考虑到天然气中 CO<sub>2</sub> 的危害, 天然气脱碳尤为重要。研究可知, 醇胺法脱碳工艺<sup>[2]</sup>是目前最常用的天然气脱碳方法。实际生产过程中要根据原料气压力、酸性气体浓度、净化气气质要求等条件进行分析, 确定并优化工艺流程, 由于大型实验装置成本高、设计复杂、调整工况耗时长、对操作人员技术要求高等特点, 导致现如今国内外关于此方面的研究不多, 相关的研究结论较少且多属于公司专有技术<sup>[3]</sup>。

时下, 组合预测模型在脱碳工艺的应用研究较少, 现阶段将机器学习相关算法应用于工业生产, 可有效提高工业生产效益。

组合预测<sup>[4-5]</sup>的思想最早是由 Bates 和 Granger (1969) 提出来的。对于同一个问题, 可以采用不同的预测方法进行预测, 不同的预测方法其预测精度往往不同, 一般来说, 没有一种预测技术可以做到零

误差, 组合预测模型在一定程度上可以充分利用单预测模型来提高预测精度。

针对醇胺法脱碳工艺存在的实际问题, 本项目在 PZ 活化 MDEA 脱碳工艺实验数据的基础上结合组合预测理论, 为天然气脱碳工艺提供理论指导。

## 1 天然气脱碳工艺参数分析

在天然气工业中应用最广泛的是醇胺法脱碳工艺, 这是一个典型的伴有化学反应的气液吸收(传质)过程<sup>[6]</sup>。针对低 H<sub>2</sub>S 含量的原料气在不同分压、不同温度下, CO<sub>2</sub> 在不同浓度的 PZ 活化 MDEA 中的溶解度进行分析, 通过脱碳后溶液循环量的大小来判断脱碳效果。溶液循环量计算步骤可分述如下。

(1) 在不同分压、不同温度下, 对 CO<sub>2</sub> 在不同浓度的 PZ 活化 MDEA 中的溶解度进行计算。

(2) 考虑不完全相平衡, 取平衡溶解度的 70%, 计算吸收塔底富液中实际酸气负荷。

(3) 根据贫液中残余酸气负荷, 计算溶液的净酸气负荷。

(4) 由脱碳装置日处理量、原料气二氧化碳含

**作者简介:** 官莉萍(1996-), 女, 硕士研究生, 主要研究方向: 数据挖掘、软件工程; 曩莹(1968-), 女, 教授, 硕士生导师, 主要研究方向: 智能计算可视化、大数据与智能处理、数字图形与图像处理。

收稿日期: 2019-07-16

哈尔滨工业大学主办 ◆ 专题设计与应用

量,计算单位时间脱除的酸气量。

(5)计算溶液摩尔流量,溶液体积流量、即溶液循环量。

## 2 基于 Shapley 值的天然气 MDEA 脱碳工艺参数组合预测模型

Shapley 值法<sup>[7]</sup>是 1953 年由 Shapley 提出的用于解决  $N$  人合作数学方法,是博弈论中的一种方法,用于分配合作组织中各成员之间的利益,这种方法具有公平性和有效性。

假设现有  $N$  种单一预测模型,记  $G = \{1, 2, 3, \dots, N\}$ 。 $a, b$  是集合  $G$  中的任一子集,其中,子集  $a, b$  表示  $N$  种预测方法中的任意一种组合,  $E(a), E(b)$  表示各自预测的误差,定义为:

(1)对集合  $G$  的任一子集  $a$  和  $b$  都有:  $E(a) + E(b) \geq E(a \cup b)$ 。

(2)将误差总和  $E$ ,按 Shapley 分配法,分配每种预测模型的误差值  $E'_i$ ,其中  $E = \sum_{i \in N} E'_i$ 。

(3)每种预测模型的权重分配公式为:

$$E'_i = \sum_{a_i \in a} \omega(|a_i|) [E(a_i) - E(a_i - \{i\})], \quad (1)$$

$$\omega(|a_i|) = \frac{(N - |a_i|)! (|a_i| - 1)}{N!}, \quad (2)$$

由上式可得到组合预测中各单一预测模型的权重分配公式为:

$$\omega_i = \frac{E - E'_i}{E(N - 1)}, \quad (3)$$

组合预测模型公式  $y = \sum w_i y_i$ ,由公式可以计算组合预测模型的预测结果。

组合预测模型的精度评估是确定预测模型优劣的标准。一般情况下,研究引用评价指标可阐释解析如下。

(1)均方误差 ( $MSE$ )。计算公式如下:

$$E_s = \frac{1}{m} \sqrt{\sum_{i=1}^m (R_i - R_{\text{实}})^2}, \quad (4)$$

(2)  $R^2$ 。取值为范围为  $[0, 1]$ ,用来评估预测结果的拟合精确度,计算公式如下:

$$R^2 = 1 - \frac{\sum_i (y'_i - y_i)^2}{\sum_i (y_i - \bar{y}_i)^2}. \quad (5)$$

## 3 实例分析

本实验的数据采用 PZ 活化 MDEA 脱碳工

艺<sup>[8-9]</sup>实验数据,数据参数主要有 MDEA 质量浓度 ( $X_1$ )、活化剂 PZ 质量浓度 ( $X_2$ )、原料气进塔压力 ( $X_3$ )、富液出塔温度 ( $X_4$ )、 $\text{CO}_2$  转化度 ( $X_5$ )、富液酸气负荷 ( $X_6$ )。通过实验的数据分析,MDEA 脱碳工艺溶液循环量在实际生产过程中具有一定的不稳定性。

### 3.1 工艺参数数据预处理

对获得的工艺参数实验数据进行归一化处理,减小各参数之间数值大小不均衡影响预测结果的准确性,对数据归一化处理能够有效提高预测结果的准确性。运用 Max-Min 方法进行归一化处理。即假设 MDEA 脱碳工艺参数为 ( $X_1, X_2, X_3, X_4, X_5, X_6$ ),当  $X_{li} \in X_1, (i = 1, 2, 3, \dots, N)$  时,将用到如下数学公式:

$$f(X_{li}) = \frac{X_{li} - X_{\min}}{X_{\max} - X_{\min}}. \quad (6)$$

其中,  $X_{\min}$  表示  $X_{li}$  中的最小值,同理  $X_{\max}$  表示最大值。当预测数据集中出现  $X'_{li} > X_{\max}$  时,取  $f(X_{li}) = 1$ ;当预测数据集中出现  $X'_{li} < X_{\min}$  时,取  $f(X_{li}) = 0$ 。

### 3.2 MDEA 脱碳工艺参数单预测模型

本实验中应用的单预测模型为贝叶斯岭回归模型、线性回归模型和支持向量机回归模型。单预测模型的评估中,运行得到的均方误差和  $R^2$  值结果见表 1。

表 1 单预测模型的均方误差和  $R^2$  值

Tab. 1 Mean square error and  $R^2$  value of single prediction mode

单预测模型	误差绝对值	均方误差	$R^2$
贝叶斯	0.948 620	0.002 59	0.948 62
线性回归	0.950 670	0.002 48	0.950 67
SVR	0.792 803	0.011 34	0.778 92

预测模型的拟合曲线分别如图 1 ~ 图 3 所示。

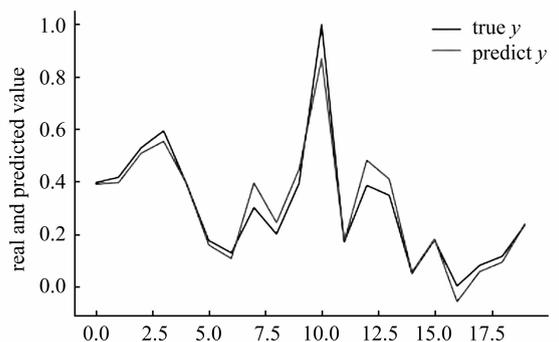


图 1 贝叶斯回归模型

Fig. 1 Bayesian regression model

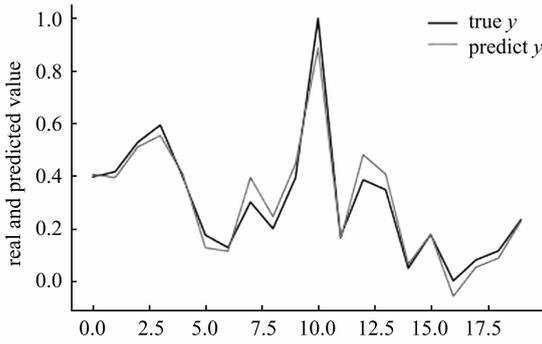


图 2 线性回归模型  
Fig. 2 Linear regression model

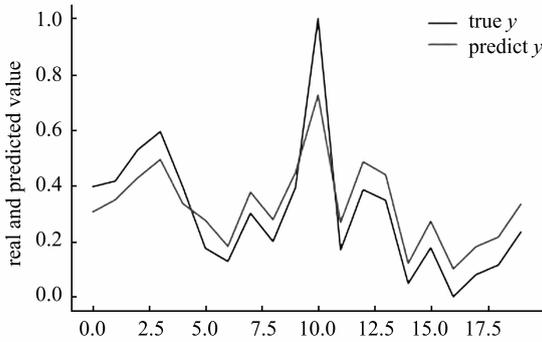


图 3 SVR 回归模型  
Fig. 3 SVR regression model

### 3.3 基于 Shapley 值的天然气 MDEA 脱碳工艺预测模型

运用 Shapley 值<sup>[9]</sup>进行权值分配,组合模型子集误差分摊见表 2。

表 2 组合模型子集误差分摊

Tab. 2 Error allocation in subset of combined model

子集	误差
$E_1$	0.002 6
$E_2$	0.002 5
$E_3$	0.011 3
$(E_1, E_2)$	0.002 5
$(E_1, E_3)$	0.007 0
$(E_2, E_3)$	0.006 9
$(E_1, E_2, E_3)$	0.005 5

根据公式(4)~(6)求取基于 Shapley 值法的误差分配及所占有的权重见表 3。

表 3 Shapley 值分配

Tab. 3 Shapley value allocation

序号	预测方法	$E_i$	$\omega_i$
1	贝叶斯	-0.000 3	0.528 78
2	GBR	-0.000 4	0.537 87
3	SVR	0.006 2	-0.066 65

由表 3 可得组合预测模型公式为:

$$y = 0.528\ 78\ y_1 + 0.537\ 87\ y_2 - 0.066\ 65\ y_3. \tag{7}$$

运用误差绝对值、均方误差和  $R^2$  对基于 Shapley 值分配的组合预测模型进行评估,最终结果见表 4。

表 4 组合预测模型评估

Tab. 4 Assessment of combination forecasting model

预测模型	误差绝对值	均方误差	$R^2$
组合预测	0.950 05	0.002 51	0.949 99

将表 3 和表 4 进行对比可知,基于 Shapley 值的组合预测的结果更具有稳定性。

预测算法的预测结果与真实值的拟合如图 4 所示。预测模型与真实值的绝对误差值情况如图 5 所示,观察可知组合预测模型在保证预测精度的同时提高了模型的稳定性。

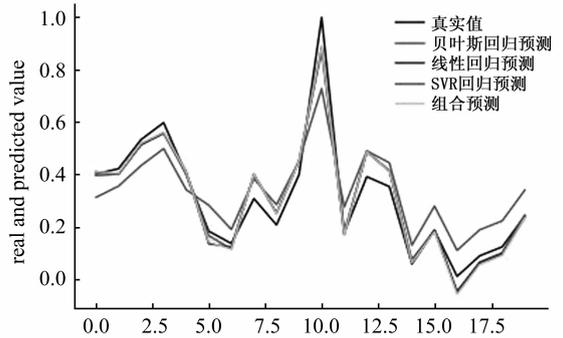


图 4 预测模型拟合结果

Fig. 4 Prediction model fitting results

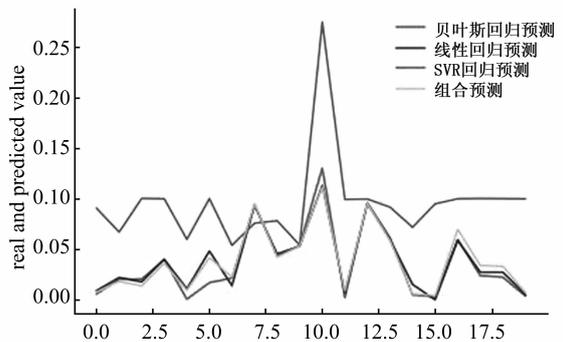


图 5 预测模型与真实值的绝对误差

Fig. 5 The absolute error between the prediction model and the real value

## 4 结束语

运用组合预测模型对天然气脱碳工艺参数预测分析可知,相较于单预测模型,基于 Shapley 值的组合预测算法可以提高预测模型的稳定性,针对脱碳工艺过程中各参数之间的非线性关系具有一定的参考价值。(下转第 278 页)