

8-20-2020

Prediction and Estimation of 95% Distillation Point Diesel in Delayed Coking Unit Based on Kernel Partial Least Squares

Cheng Jian

Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China;

Lingbo Zhang

Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China;

Xingsheng Gu

Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research](#), [Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Prediction and Estimation of 95% Distillation Point Diesel in Delayed Coking Unit Based on Kernel Partial Least Squares

Abstract

Abstract: In order to save labor cost and time of 95% distillation point of diesel analysis, and improve the efficiency of adjusting the device, an algorithm was proposed to predict 95% distillation point of diesel based on Kernel Partial Least Squares(KPLS), on the basis of studying the process of delayed coking unit and analyzing the main factors related to 95% distillation point of diesel. The experiment results show that this algorithm can deal with the non-linear relationship among the variables effectively and quickly. Furthermore, the results also show that the model based on KPLS is better than that based on partial least squares. So this algorithm can be used to guide production and improve the profit of enterprise in practice.

Keywords

95% distillation point diesel, fractionating tower, partial least squares, kernel partial least squares

Recommended Citation

Cheng Jian, Zhang Lingbo, Gu Xingsheng. Prediction and Estimation of 95% Distillation Point Diesel in Delayed Coking Unit Based on Kernel Partial Least Squares[J]. Journal of System Simulation, 2015, 27(3): 598-602.

基于 KPLS 的延迟焦化柴油 95% 点的预测与估计

程剑, 张凌波, 顾幸生

(华东理工大学化工过程先进控制和优化技术教育部重点实验室, 上海 200237)

摘要: 为了节约人工分析柴油 95% 点的成本和时间、提高延迟焦化装置调整的效率, 在研究延迟焦化装置工艺流程和分析柴油 95% 点影响因素的基础上, 提出了一种基于核偏最小二乘的柴油 95% 点的预测方法。仿真结果表明该方法能够有效、快速地处理各变量之间的非线性关系, 并且预测结果也明显优于偏最小二乘法等方法。所提出的方法可用于指导实际生产过程, 提高企业效益。

关键词: 柴油 95% 点; 分馏塔; 偏最小二乘法; 核偏最小二乘法

中图分类号: TP301

文献标识码: A

文章编号: 1004-731X (2015) 03-0598-05

Prediction and Estimation of 95% Distillation Point Diesel in Delayed Coking Unit Based on Kernel Partial Least Squares

Cheng Jian, Zhang Lingbo, Gu Xingsheng

(Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China)

Abstract: In order to save labor cost and time of 95% distillation point of diesel analysis, and improve the efficiency of adjusting the device, an algorithm was proposed to predict 95% distillation point of diesel based on Kernel Partial Least Squares (KPLS), on the basis of studying the process of delayed coking unit and analyzing the main factors related to 95% distillation point of diesel. The experiment results show that this algorithm can deal with the non-linear relationship among the variables effectively and quickly. Furthermore, the results also show that the model based on KPLS is better than that based on partial least squares. So this algorithm can be used to guide production and improve the profit of enterprise in practice.

Keywords: 95% distillation point diesel; fractionating tower; partial least squares; kernel partial least squares

引言

当代, 轻质原油资源的消耗, 导致炼油企业越来越趋向于原油重质化, 产品轻质化。为了充分利用有限的石油资源, 满足轻质油快速增长的需求和企业经济效益的追求, 重油加工技术发展迅速, 其中以延迟焦化工艺最为明显。延迟焦化装置投资抵、费用少, 最大的优点是能够将各种重质渣油提

炼成为特种石油焦、液、气体等经济产品, 使汽油和柴油产量大大提升。因而如何使延迟焦化技术更加成熟是我们以后所要努力研究的一个重要方面。

柴油干点是化工企业一项重要的安全和经济指标, 能够反映整个分馏过程的合理性。在未使用软测量技术之前, 粗柴油 95% 点是通过人工定时采样化验分析获得。车间工艺根据化验结果, 对相关工艺参数进行合理调整, 保证柴油 95% 点维持在要求的控制范围之内。但是, 所需化验时间偏长, 导致控制存在时延性, 因而柴油收率将会下降。为了减少柴油 95% 点操作裕量, 可以通过软测量技术实时性的优点, 减少时间损耗。从而实现真正意义上



收稿日期: 2014-04-16 修回日期: 2014-10-30;
基金项目: 中央高校基本科研业务费专项基金; 上海市重点学科项目(B504);
作者简介: 程剑(1989-), 男, 浙江杭州人, 硕士生, 研究方向为系统建模与优化; 张凌波(1974-), 男, 湖南永州人, 副教授, 研究方向为过程控制。

的卡边生产, 获取充分的经济和环保效益。

在化工生产与研究中, 物质与物质之间有些是线性关系, 或者通过变换可以表示为近似线性关系。当关系较为简单时, 研究其相关性可以通过偏最小二乘^[1-2](partial least square, PLS)来分析和推断。PLS模型能够实现数据降维, 提取有用信息。把数据中尽可能多的信息用来建模, 实现分析结果的合理性^[3]。但是, PLS模型对于一些非线性问题或者较为复杂的问题则不适合。为了能够处理非线性问题, Trejo和Rosipal教授^[4]将核函数引入到PLS模型中, 提出了核偏最小二乘(kernel partial least square, KPLS)。该方法不但能够处理非线性问题, 而且能够充分利用样本信息把物质之间联系起来建立关系模型。KPLS模型已在医药和电力等方面得到了很好的应用^[5-6]。同时在化工领域也慢慢开始受到重视。

针对影响柴油干点的因素存在着复杂的非线性, 为了提高预测结果的精度。利用KPLS模型能够很好且快速地处理非线性问题的能力, 通过PLS算法和KPLS算法建模的结果比较。提出了基于核的偏最小二乘的延迟焦化分馏塔95%点的预测与估计。通过实例仿真, 表明核偏最小二乘模型在预测柴油95%点上具有较高的精度。

1 核偏最小二乘方法

1.1 核偏最小二乘算法基本原理

一般而言, 预测模型最常用的有多元回归、主元回归和偏最小二乘回归等算法。它们的本质其实就是寻找变量之间的联系。多元回归能够得到自变量和应变量之间的关系, 是其他预测模型的基础。主元回归和偏最小二乘回归都可以从多元事物中解析出主要的影响因素, 将高维数据投影到低维空间, 简化复杂问题, 找出事物本质。

如果说基于核函数支持向量机(support vector machine, SVM)^[7]的成功是人们引入核函数的动力, 那么因为核函数的引用而产生的KPLS模型就是核函数引入的代表, 它给非线性建模问题带来了

生机。它拥有典型相关分析, 多元线性回归分析和主成分分析的基本功能。KPLS的原理是通过非线性函数 $\phi()$ 将输入空间映射到高维特征空间。然后在高维特征空间中运用PLS模型。从而可以得到输入空间和输出空间之间的非线性关系^[8-10], KPLS的输入空间到高维空间原理示意图如图1所示。

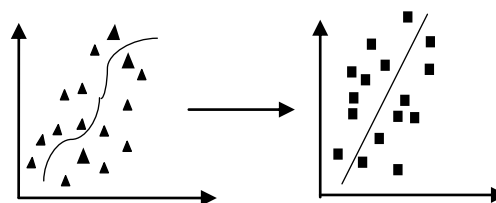


图 1 输入空间到高维特征空间投影图

核偏最小二乘的最大优点是不必纠结于非线性函数 $\phi()$ 的具体形式。在满足Mercer条件^[11]的情况下, 可以通过输入空间的点积运算直接得到核函数, 从而绕过了前面所提的非线性函数, 进而直接得到核函数用于下一步分析, 核函数可描述为:

$$\mathbf{K}(x_i, x_j) = \phi(x_i)^T \cdot f(x_j) \quad (1)$$

式中, $\mathbf{K}(x_i, x_j)$ 为核函数。

因为核函数的重要性, 所以对其研究也开始逐渐深入, 如核函数的选择以及参数的调整^[12]。常见的核函数有以下3种:

(1) Sigmoid核函数, 也叫双层神经网络核函数。表示为:

$$\mathbf{K}(x_i, x_j) = \tanh(w * (x_i \cdot x_j) + coef) \quad (2)$$

式中: w 为核函数的宽度; $coef$ 为偏置系数。

(2) Polynomial核函数, 表示为:

$$\mathbf{K}(x_i, x_j) = \exp(-w * \|x_i - x_j\|^2) \quad (3)$$

式中: w 为核函数的宽度; $coef$ 为偏置系数; d 为多项式阶数。

(3) 高斯核函数, 表示为:

$$\mathbf{K}(x_i, x_j) = \exp(-w * \|x_i - x_j\|^2) \quad (4)$$

式中: w 为核函数的宽度。

本文采用了高斯核函数, 因为高斯核函数能够把输入空间变量投射到无限维的空间上, 即有足够

的维数, 这样在高维特征空间上就可以选择合适的线性建模方法, 实现线性建模。这样不但理解简单, 而且实现方便。

1.2 核偏最小二乘算法基本步骤

核偏最小二乘算法首先需要进行核函数的选择, 又因为KPLS算法是在映射数据均值为零的基础上得到的, 所以需对核矩阵 K 进行中心化处理, 对于 $N \times N$ 维的核矩阵 K , 中心化过程如下式所示:

$$K = K - \frac{1}{N} O_N K - \frac{1}{N} K O_N + \frac{1}{N^2} O_N K O_N \quad (5)$$

式中: O_N 为元素全为 1 的 $N \times N$ 矩阵。

当选取完核函数, 以及对输入空间变量矩阵 X 和核矩阵中心化以后。可以通过以下步骤完成KPLS算法。

(1) 随机选取输出空间变量的 Y 的一初始列 u 。

(2) 计算输出空间变量的得分向量。

$$t_i = K_i u_i \quad (6)$$

(3) 单位化 t_i 向量。

$$t_i = t_i / \|t_i\| \quad (7)$$

(4) 计算输出空间变量的得分向量的权重 w_i 。

$$w_i = Y^T t_i \quad (8)$$

(5) 循环计算输入空间变量得分向量 u_i 。

$$u_i = Y_i w_i \quad (9)$$

并对其单位化,

$$u_i = u_i / \|u_i\| \quad (10)$$

(6) 重复步骤(2)~(5), 直至收敛。

(7) 依据下列公式更新矩阵 K 和 Y 。

$$K_{(i+1)} = (I - t_i t_i^T) K_i (I - t_i t_i^T) \quad (11)$$

$$Y_{i+1} = Y_i - t_i t_i^T Y_i \quad (12)$$

$$\text{式中, } I = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix} \quad (13)$$

(8) 转至步骤(2)直到计算出所有的得分向量。

建立KPLS的回归模型和预测模型。

$$\hat{Y} = K U (T^T K U)^{-1} T^T Y \quad (14)$$

$$\hat{Y}_i = K_i U (T^T K U)^{-1} T^T Y \quad (15)$$

式中: U 和 T 分别由成分 u_i 和 t_i ($i=1, 2, \dots, n$), Y 与 K 为训练输出与核矩阵, \hat{Y}_i 与 K_i 为预测输出与核矩阵。

2 延迟焦化工艺研究

2.1 问题描述

本文所介绍的延迟焦化装置采用一塔二炉的设计方案, 以减压渣油为原料。延迟焦化与其他形式焦化的不同之处, 即为了防止热渣油在炉管进行反应, 渣油以很高的流速流过加热炉的炉管, 几乎瞬间加热到反应所需的 $500 \sim 550$ °C, 然后进入焦炭塔。由于在焦炭塔时间较长, 产生缩合与裂解等一系列反应, 生焦于焦炭塔。生产的油气通过焦炭塔顶进入分馏塔进行分馏, 最终生成富气、汽油、柴油、蜡油等产品^[13-14]。

延迟焦化装置主要由加热炉、焦炭塔、分馏塔、气压机和汽轮机等装置组成。粗柴油 95% 点是在分馏塔内产生, 根据所要建立模型的实际特点。所以有必要对分馏塔进行分析, 分馏塔的主要作用为换热和分馏。换热作用是在分馏塔底与高温油气换热, 减轻加热炉的热负荷和提高装置的热利用率。分馏作用是根据各种油气的挥发度不同切割成不同沸点范围的石油产品, 如图 2 所示。因为分馏塔内成分复杂, 无法通过确切的机理方法进行建模。而人工采样分析时间间隔长, 误差大, 不能实时对装置进行调整。通过使用软测量技术可以从根本上改变这种现状。

2.2 过程因素选定

根据对分馏塔的分析和研究, 石油炼制的基本原理以及对延迟焦化的实际操作过程的了解, 仔细分析筛选后得到影响柴油 95% 点的因素如表 1 所示。

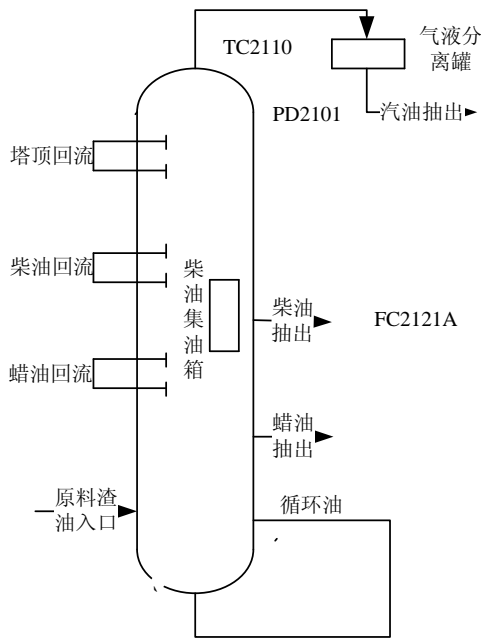


图 2 延迟焦化分馏塔工艺流程示意图

表 1 延迟焦化柴油 95%点软测量辅助因素

| 位号 | 描述 |
|---------|-----------|
| FC2121A | 柴油抽出量 |
| PD2101 | 分馏塔顶压力 |
| TC2110 | 分馏塔顶温度 |
| TD2116 | 柴油集油箱温度 |
| TC2107 | 柴油上返塔冷流流控 |
| FC2105 | 柴油上返塔热流流控 |
| TC2111 | 柴油下返塔温控 |
| TC2109 | 柴油下返塔流控 |
| TD2117 | 柴油抽出温度 |
| FC2106 | 柴油上返塔冷流流控 |

因此, 表 1 中所列的影响柴油 95%点的因素, 柴油 95%点可用下式表示:

$$y(k) = f[x_1(k), x_2(k) \cdots x_{10}(k)] \quad (16)$$

式中: $x_1(k), x_2(k) \cdots x_{10}(k)$ 为表 1 中的各因素。

3 延迟焦化柴油95%点软测量

根据上面所建立的模型。采集了276组延迟焦化柴油95%点的人工分析值和与之对应的现场操作各影响因素数据作为样本数据集。经过数据处理后, 并以180组为训练值, 76组为预测值, 分别用PLS和KPLS算法建立柴油95%点的软测量模型。

3.1 基于 PLS 的柴油 95%点软测量模型

PLS 算法可以很好的消除变量之间的线性关系, 但是它对一些非线性模型可能会出现较大的偏差。PLS 模型预测仿真如图 3 所示。

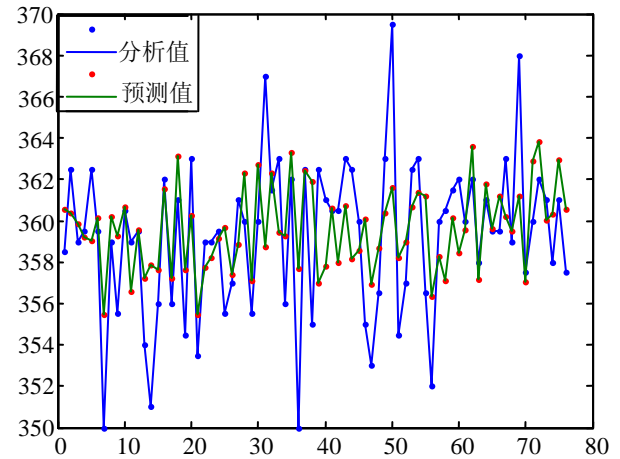


图 3 PLS 模型预测仿真曲线

3.2 基于 KPLS 的柴油 95%点软测量模型

KPLS 算法能够克服非线性因素对预测模型的不利影响。通过特征提取后, 主元信息之间具有高阶相关性。KPLS 预测仿真结果如图 4 所示。

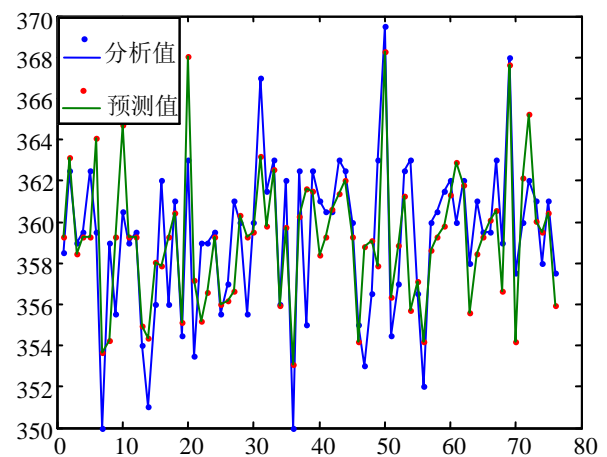


图 4 KPLS 模型预测仿真曲线

从以上PLS和KPLS模型的仿真曲线可以看出, PLS算法能够较好的对一些数值变化比较平稳的柴油95%点进行预测, 但是对于一些数值变化较大的柴油95%点则明显预测能力不足, 分馏塔内复

杂的成分导致了各影响因素之间存在着非线性关系，这已经超出了PLS算法所适用的范围。相反KPLS算法对一些变化较大的柴油95%点的预测效果要好。

为了更加直观的说明PLS算法和KPLS算法对于柴油95%干点建模的适应性，分别列出对应的均方差、相对误差以及绝对误差如表2所示，通过各个误差的对应比较我们可以看出无论是均方差、相对误差还是绝对误差，KPLS算法与PLS算法相比都具有较大优势。

表 2 PLS 和 KPLS 预测拟合结果误差比较

| 误差 | PLS | KPLS |
|------|--------------|-------------|
| 均方差 | 10.229 970 3 | 7.0795 080 |
| 相对误差 | 0.704 30 % | 0.579 62 % |
| 绝对误差 | 2.527 646 9 | 2.079 760 9 |

4 结论

为了解决实际生产过程中柴油95%的点检验，缩短装置调整的时间，使生产过程更加人性化。本文以某化工厂的实际化工数据为例，选择能够实现快速计算柴油干点的KPLS算法，对延迟焦化柴油95%点进行建模，取得了很好的效果。KPLS算法所需运行时间极短，从精度上优于只能处理线性问题的PLS算法，因而能够快速及时指导实际生产。KPLS算法能够很好的搭建输入和输出之间关系的桥梁，给工业建模带来了新的思路。

参考文献:

- [1] 王惠文. 偏最小二乘回归方法及其运用 [M]. 北京: 人民邮电出版社, 2001: 80-120.
- [2] Geladi P, Kowalski B R. Partial least-squares regression: a tutorial [J]. *Analytica chimica acta* (S0003-2670), 1986, 185(86): 1-17.
- [3] Fortuna L, Graziani S, Rizzo A, *et al.* Soft sensors for monitoring and control of industrial processes [M]. Germany: Springer, 2007.
- [4] Rosipal R, Trejo L J. Kernel partial least squares regression in reproducing kernel hilbert space [J]. *The Journal of Machine Learning Research* (S1533-7928), 2001, 2(2): 97-123.
- [5] 蒋红卫, 夏结来, 张春霞, 等. 核偏最小二乘回归及其在医学中的应用 [J]. *中国卫生统计*, 2007, 24(3): 239-242.
- [6] 张曦, 陈世和, 陈锐民, 等. 基于核偏最小二乘的电厂热力参数预测与估计 [J]. *南方电网技术*, 2011, 5(1): 127-127.
- [7] 邹永杰, 端木京顺, 高海龙. 基于核偏最小二乘的支持向量机回归算法研究 [J]. *计算机工程与设计*, 2010 31(10): 2290-2293.
- [8] Roman R. Kernel partial least squares regression in reproducing kernel hillbert space [J]. *Journal of Machine Learning Research* (S1533-7928), 2000, 40(10): 25-47.
- [9] Zhang Y, Teng Y, Zhang Y. Complex process quality prediction using modified kernel partial least squares [J]. *Chemical Engineering Science* (S0009-2509), 2010, 65(6): 2153-2158.
- [10] Rosipal R, Trejo L J, Matthews B. Kernel PLS-SVC for linear and nonlinear classification [C]// *ICML*. Menlo Park, USA: AAAI Press, 2003: 640-647.
- [11] H Hochstadt. *Integral equations* [Z]. New York, USA: Wiley Classics Library. John Wiley & Sons Inc., 1989. Reprint of the 1973 original, A Wiley-Interscience Publication.
- [12] 武优西, 郭磊, 柴欣, 等. 基于优化算法的核函数参数选择的研究 [J]. *计算机应用与软件*, 2010, 27(1): 137-139.
- [13] 刘立民. 延迟焦化装置软测量及先进控制研究 [D]. 北京: 中国石油大学, 2009.
- [14] 梁朝林, 沈本贤. 延迟焦化 [M]. 北京: 中国石化出版社, 2007.