

中红外光谱对甲醇汽油甲醇含量检测研究

刘燕德, 胡军, 唐天义, 张宇, 欧阳玉平, 欧阳爱国

华东交通大学机电与车辆工程学院, 江西 南昌 330013

摘要 甲醇汽油是一种清洁能源, 甲醇汽油中甲醇的含量决定了汽油的性能。通过中红外光谱对甲醇汽油中甲醇含量进行定量检测和分析。首先, 对采集的甲醇汽油原始中红外光谱进行平滑处理(smoothing)、多元散射校正(MSC)、基线校正(baseline)、归一化(normalization)等预处理, 再建立 PLS 模型, 对比选择最佳预处理方法, 结果表明: 在多元散射校正(MSC)处理后建立的 PLS 模型效果最好, 模型的预测集相关系数 r 为 0.918, 预测均方根误差 RMSEP 为 2.107。为进一步简化模型, 提高预测精度, 采用无信息变量消除(uninformative variable elimination, UVE)方法对波长进行筛选, 将 UVE 波段筛选之后的作为模型的输入变量, 采用偏最小二乘法(partial least squares, PLS)、主成分回归(principal components regression, PCR)和最小二乘支持向量机(least square support vector machine, LSSVM)三种方法分别建立甲醇汽油中甲醇含量的定量预测模型, 并比较不同模型的预测效果和结果。结果表明, 使用无信息变量消除可以较好提高数据的运算速度, 其中, UVE-PLS 模型建模效果最好, r 和 RMSEP 分别为 0.923 和 2.075。该实验表明中红外光谱检测甲醇汽油中甲醇含量是可行的并可以得到较好的效果; UVE 是一种对甲醇汽油的中红外光谱非常有效的波段筛选方法, 该模型的建立对石油化工领域具有较为重要的意义。

关键词 中红外光谱; 甲醇汽油; 无信息变量消除; 偏最小二乘支持向量机

中图分类号: O657.6 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2019)02-0459-06

引言

随着汽车工业的发展、石油消耗量越来越大, 其储量减少并伴随着汽油燃烧对环境的污染, 汽油和柴油等传统燃料的使用面临着巨大挑战。由于甲醇具有清洁可再生的优点, 使得甲醇汽油成为了一种可替代的能源, 受到了越来越多的重视^[1-3]。

甲醇汽油的使用可以提高汽油中的含氧量, 使汽油更充分燃烧, 此外能够有效降低一氧化碳和有毒气体排放。目前市场上低浓度的甲醇汽油可以直接替代汽油使用, 例如 M15 汽油中甲醇含量约为 15%; 当比例达到 85% 以上, 则称为高浓度甲醇汽油^[4], 若需要使用高浓度甲醇汽油, 则需要添加助溶剂或者改变发动机结构, 针对目前市场而言, 使用更多的还是低浓度甲醇汽油^[5-7]。

甲醇汽油作为燃料也存在一些不足, 比如甲醇汽油的热值较低, 当掺入过量时会导致发动机动力不足。此外, 过量的甲醇燃烧容易产生气阻, 影响系统供油, 从而对发动机的

气密性产生不良影响, 所以甲醇含量是甲醇汽油品质的一个重要衡量指标。寻找一种高效的准确的方法来快速检测甲醇汽油中甲醇的含量显得尤为重要^[8-10]。

在过去的二十年里, 国内外学者做了大量关于汽油的光谱分析研究, 光谱检测方法分析石油炼制和石油化工产品比一般技术(ASTM)更快, 呈现良好的准确度和精密度, 无需破坏样品, 可用于远程质量监控, 在食品、医药、石化、农业等多个行业领域的定性定量检测分析中被广泛应用^[11]。中红外光谱(mid infrared spectroscopy, MIR)的波数范围在 $4\ 000\sim 400\text{ cm}^{-1}$ 之间, 反映的是分子中原子的伸缩和变形振动运动。张瑞华^[12]利用近红外和中红外对汽油的调和比进行测定, 相关系数大于 0.9。毛佳伟等^[13]利用中红外快速测定汽油中的甲苯含量, 得到了较好的效果。欧阳爱国^[14]等采用了近红外和拉曼光谱仪两种光谱仪器, 建立了甲醇含量的 PLS 模型。实验结果表明, 同样是 PLS 模型, 应用近红外光谱获得的结果更好。姚捷^[15]等采用拉曼光谱检测甲醇汽油, 针对拉曼特征峰建立甲醇含量的一元线性回归模型, 实验结果表明, 使用拉曼特征峰所需样本数更少且模型具有良

收稿日期: 2016-12-23, 修订日期: 2017-05-04

基金项目: 国家“十二五”(863)计划项目(SS2012AA101306), 江西省优势科技创新团队建设计划项目(20153BCB24002), 南方山地果园智能化管理技术与装备协同创新中心项目(赣教高字[2014]60号), 江西省研究生创新资金项目(YC2016-S253)资助

作者简介: 刘燕德, 女, 1967年生, 华东交通大学机电与车辆工程学院教授 e-mail: jxliuyd@163.com

好的精度和稳健性。

应用中红外光谱技术对甲醇汽油含量的检测,目前还鲜有报道。本工作采用中红外光谱,建立对甲醇汽油不同预处理后的甲醇汽油甲醇含量的 PLS 模型,为进一步优化模型,结合 UVE 波段筛选算法,建立了 UVE-PLS, UVE-PCR 和 UVE-LSSVM 三种模型,并对三种建模效果进行评价,最终确定最优甲醇汽油中甲醇含量定量检测模型。

1 实验部分

1.1 材料

实验所用的汽油从学校附近某加油站购买,汽油型号为 93#, 实验用的无水甲醇的分析纯度大于或等于 99.7%。样品在配制时,按照 0.6% 的浓度梯度进行配制,配比浓度范围从 0.6%~18.6%。先按照相应梯度,计算不同比例下甲醇的体积,然后将对应体积的甲醇加入到容量瓶,再用汽油进行滴定至 50 mL,振荡均匀,直至液体不分层为止。按照相同方法,配制 30 组不同浓度的甲醇汽油溶液。每个浓度配制 4 个样品,共得到 120 个甲醇汽油样品,建模集和预测集数量比为 3:1,即建模集样本 90 个,预测集样本 30 个。利用建模集建立模型,利用预测集对模型进行验证,所配制的 30 组甲醇汽油浓度表如表 1 所示。表为建模集和预测集的甲醇汽油中甲醇含量真实值分布统计结果。

表 1 30 组不同浓度样品的体积比(V/V)
Table 1 The composition of methanol-gasoline mixtures with different ratios

序号	浓度 /%	序号	浓度 /%	序号	浓度 /%	序号	浓度 /%
1	0.6	9	5.6	17	10.6	25	15.6
2	1.2	10	6.2	18	11.2	26	16.2
3	1.8	11	6.8	19	11.8	27	16.8
4	2.4	12	7.4	20	12.4	28	17.4
5	3.0	13	8.0	21	13.0	29	18.0
6	3.6	14	8.6	22	13.6	30	18.6
7	4.2	15	9.2	23	14.2		
8	5.0	16	10.0	24	15.0		

表 2 建模集和预测集的甲醇汽油中甲醇含量真实值分布统计结果
Table 2 Statistic information of methanol contents ($\text{g} \cdot \text{mL}^{-1}$)

样品集	数量	最小值 /%	最大值 /%	平均值 /%	标准差
全部样本	120	0.6	18.6	9.60	5.403
建模集	90	0.6	18.6	9.61	5.414
预测集	30	0.6	18.6	9.59	5.458

1.2 光谱采集

采用 WQF-510A 傅里叶变换红外光谱仪,与 OPUS 光谱采集软件配套使用,获得甲醇汽油中红外光谱。实验的环

境温度控制恒温 20 °C 左右,湿度控制在 35% 以下,设置光谱仪参数,测量范围为:4 000~600 cm^{-1} ,分辨率优于 0.85 cm^{-1} ,扫描时间 32 s,开机后预热 30 min 待仪器稳定后方可使用。样品测定采用水平衰减全反射(attenuated total reflection, ATR)附件(包括硒化锌晶体片)。每次采集光谱时,取约 1 mL 甲醇汽油溶液均匀滴在硒化锌晶体片上,关闭仪器舱门,进行光谱采集。数据处理在 Unscrambler (version. 10.1; CAMO)化学计量学分析软件中进行,所有程序均在 MATLAB2012a(The MathWork, Inc, USA)中编辑与运行。为减少因样品挥发等因素对实验结果的影响,每个样品分别采集 5 条光谱,并求出平均中红外光谱用于建立模型。

1.3 模型评价标准

偏最小二乘法是光谱技术检测分析中最成熟的一种算法,首先对中外光谱进行预处理,建立 PLS 模型评价预处理方法的优劣。然后利用无信息消除算法(UVE)对其进行波段筛选,再分别采用偏最小二乘(PLS)、主成分回归(PCR)和偏最小二乘支持向量机(LS-LVM)建立光谱定量检测模型。建模集和预测集数量比约为 3:1,建模集样品用于建立模型,验证集样品用于对模型的验证,预测相关系数(correlation coefficient, r)和预测均方根误差(root mean square error of prediction, RMSEP)作为本模型的评价指标。采用 Matlab 软件处理数据和建立模型。甲醇汽油甲醇含量的 PLS 模型建模效果好坏由预测集的 r 和 RMSEP 共同决定,PLS 模型的相关系数越高且预测均方根误差越小则模型的精度就越高,并且 RMSEC 与 RMSEP 两者差值越小,模型的稳健性越好。如图 1 为甲醇汽油中甲醇含量检测模型建立的流程图。

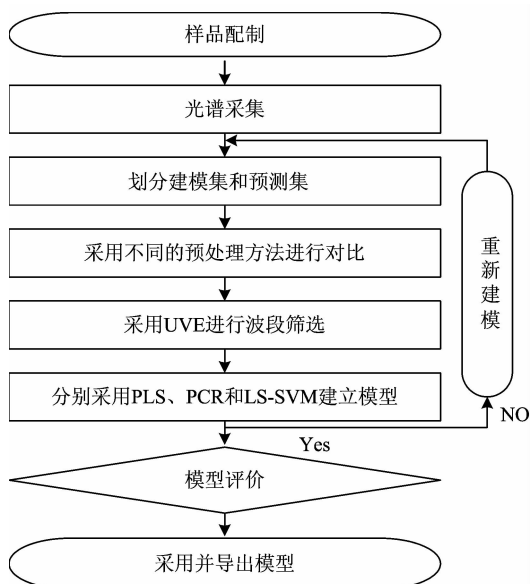


图 1 甲醇汽油甲醇含量检测模型建立的流程图

Fig. 1 Flow chart of model establishing

2 结果与讨论

2.1 甲醇汽油的中红外光谱图及甲醇含量分析

图 2(a) 为未添加甲醇的汽油原始中红外光谱图,图 2

(b)为无水甲醇样品的原始中红外光谱图,图 3 为不同浓度甲醇汽油的中红外光谱图,如图 3 所示,在 1 500~2 750 和 3 050~3 500 cm^{-1} 范围内,谱强与甲醇含量有较好的线性关系。而在 1 000~1 250 以及 3 500~4 000 cm^{-1} 范围内光谱数据与甲醇含量的线性相关性较差,这是由于汽油中烃类物质的官能团吸收峰重叠导致的。为方便数据处理,选择中红外光谱 1 250~3 500 cm^{-1} 波段进行建模。

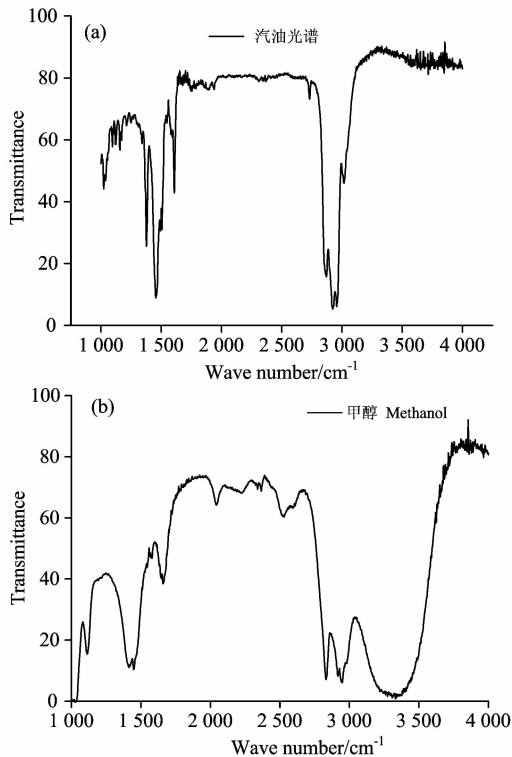


图 2 纯品汽油和甲醇样品的原始中红外光谱图对比

(a): 汽油原始光谱; (b): 无水甲醇样品的原始光谱

Fig. 2 Mid infrared spectra of gasoline and methanol

(a): Gasoline original spectrum; (b): Methanol original spectrum

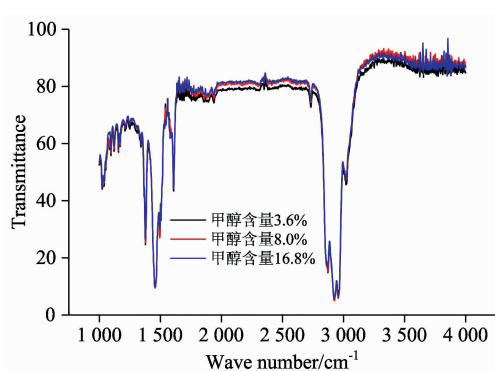


图 3 不同比例甲醇汽油的中红外光谱图

Fig. 3 Mid infrared spectra of mixtures of methanol and gasoline

2.2 甲醇汽油甲醇含量光谱预处理

实验仪器的振动、噪声以及温度的变化都会使测得的原

始中红外光谱产生噪音、基线上下漂移、光散射等影响。为减小外界条件干扰,对采集的中红外光谱进行适当预处理非常有必要,通过合适的预处理方法可以获得较好的建模效果。主要采用了以下预处理手段,如平滑(smoothing)、多元散射校正(MSC)、基线校正(baseline)、归一化(normalization)等,利用预处理后的甲醇汽油光谱数据分别建立偏最小二乘模型,并评价不同预处理后数据的建模效果。

将采集到的甲醇汽油中红外光谱经预处理(smoothing, MSC, baseline, normalization)后得到的数据作为模型输入变量 X, 甲醇浓度作为输出变量 Y, 建立偏最小二乘模型。表 3 为甲醇汽油中红外光谱经过不同预处理建立的 PLS 定量检测模型。比较可得经多元散射校正处理后的 PLS 模型效果最好,模型的预测集相关系数 r 为 0.918, 预测均方根误差 RMSEP 为 2.107。而经过基线校正和归一化处理后的建模效果均低于用原始中红外光谱建模效果,此时预处理对建模结果不但没有优化反而降低了模型的预测精度。图 4 为经过多元散射校正预处理后的不同浓度甲醇汽油中红外光谱图,通过该预处理方法消除了测试液体样品因涂抹不均产生的散射影响。

表 3 甲醇含量不同预处理的 PLS 模型

Table 3 Results of PLS model for ethanol content of ethanol diesel oil

Model	Preprocessing	主成份因子数	建模集		预测集	
			r_c	RMSEC	r_p	RMSEP
PLS	原始光谱	10	0.985	0.928	0.916	2.166
	Baseline	8	0.974	1.217	0.908	2.253
	MSC	9	0.989	0.795	0.918	2.107
	Normalization	8	0.974	1.215	0.902	2.326

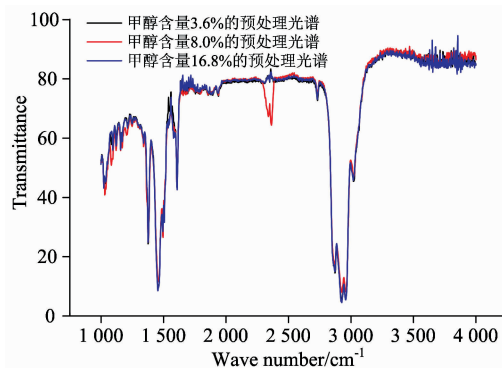


图 4 经过 MSC 预处理后的不同浓度甲醇汽油中红外光谱图

Fig. 4 Mid infrared spectra of methanol and gasoline mixtures with different ratios by MSC preprocessing

2.3 甲醇汽油甲醇含量中红外光谱变量筛选

波段筛选是一种非常有效的光谱处理方法,能够大大优化模型。中红外光谱数据量较大,以全波段中红外光谱作为输入变量,数据处理耗时长。选择一种适当的波长筛选方法,不仅可以缩短建模时间,还能提高建模的预测精度和稳

定性。为避免光谱仪噪声的影响以及一些波长对外界环境因素变化敏感的干扰,去掉冗余信息,从而提高数据的处理速度。

本工作采用无信息变量消除(UVE)方法,其原理是基于 PLS 回归系数建立的波段筛选方法,将回归系数作为重要衡量指标,首先将一定数目的随机变量矩阵加入光谱矩阵中,再通过交互验证建立 PLS 模型,通过计算 PLS 与回归系数的平均值与标准偏差比值来选取有效光谱信息,该方法将噪声信号和浓度信息集于一体,方便和直观。

在全谱范围内采用 UVE 方法对不同浓度甲醇汽油的中红外光谱进行变量优选,图 5 为 UVE 变量稳定性的分析结果,垂直竖实线为波长变量和随机噪声变量的分割线,左边为波长变量,右边为噪声变量。T value1 和 T value2 分别代表阈值筛选的上下限,在 T value1 和 T value2 外部对应的变量被保留,变量稳定性值位于两阈值之间的波长变量为无用信息变量,被剔除,在最后的模型中,虚线内部对应的变量不用于建模。通过 UVE 波段筛选变量后,共有 103 个波长变量被选择。以筛选的波长点所含信息作为输入变量建立更加简化的甲醇汽油甲醇含量定量检测模型。

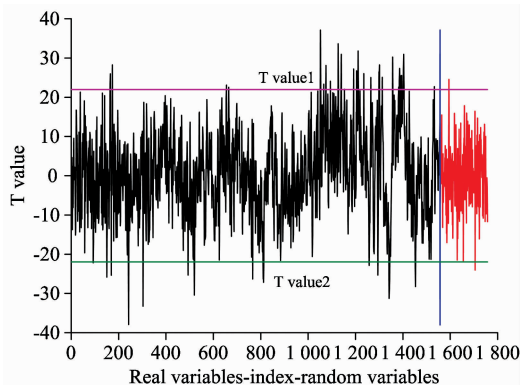


图 5 全谱范围内采用 UVE 方法进行变量优

Fig. 5 Stability of each variable in the UVE analysis

2.4 甲醇汽油甲醇含量光谱定量模型对比

2.4.1 UVE-PLS 模型的建立

偏最小二乘是近年来提出的统计方法,它可充分利用甲醇汽油光谱数据与甲醇含量之间的线性关系,建立最优检测模型。图 6 为甲醇含量建模和预测集均方根误差随主成分因子数变化曲线图,建模时所选主成分因子数(PC)过低,可能忽略部分有效光谱信息,导致模型精度降低,出现“欠拟合”现象;而当所选 PC 过高时,易出现“过拟合”现象,例如建模 PC 为 8, RMSEC 值更低,但 RMSEP 值较 PC=7 时略高。因此,本实验建立甲醇汽油中甲醇含量的 PLS 模型中,最佳主成分因子数为 7。图 7 为 PLS 模型预测甲醇含量的散点图, UVE-PLS 模型的预测效果为:预测集 r 和 RMSEP 分别为 0.923 和 2.075。

2.4.2 UVE-PCR 模型的建立

主成分回归(PCR)法是采用多元统计中的主成分分析(PCA)法, UVE 筛选的波长变量作为模型输入变量 X , 甲醇汽油中甲醇浓度作为输出变量 Y 。首先对甲醇汽油甲醇含量

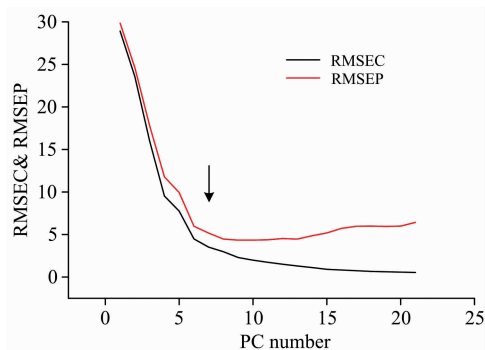


图 6 PLS 建模预测集均方根误差随主成分因子数变化曲线图

Fig. 6 RMSEC and RMSEP vs PC number for PLS modeling

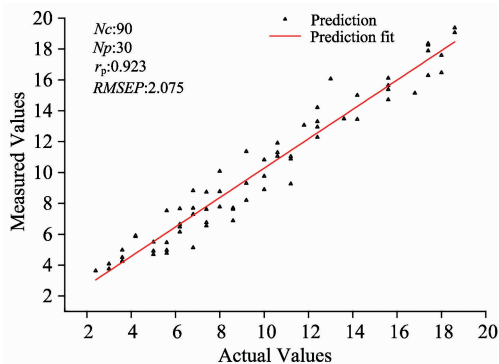


图 7 PLS 预测甲醇含量的散点图

Fig. 7 Scatter plots of the methanol content prediction with PLS method

光谱矩阵进行分解,然后选取主成分得分来进行线性回归运算,从而得到定量模型。图 8 为 PCR 预测甲醇含量的散点图 UVE-PCR 模型的 r 和 RMSEP 分别为 0.916 和 2.159。

2.4.3 UVE-LSSVM 模型的建立

在 UVE-LSSVM 建模中,采用 Lin 核函数和 RBF 核函数分别建立支持向量机模型,比较两种核函数分别的建模效果,结果表明:采用 Lin 核函数时, $\gamma=4.559$, 其预测集 r 和 RMSEP 分别为 0.873 和 2.624, 采用 RBF 核函数时,该模型有两个重要参数: γ 和 σ^2 , 这两个未知参数在一定意义上影响模型泛化能力和建模效果。通过多次尝试,确定了模型的两个参数,此时 UVE-LSSVM 模型参数的最佳组合为: $\gamma=1.2071 \times 10^3$, $\sigma=392.379$ 。模型的预测结果如表 4 所示,其预测集 R 和 RMSEP 分别为 0.913 和 2.252。通过对比,发现选用 RBF 核函数更加有利于模型的建立。图 9 为 LS-SVM 预测甲醇含量的散点图。

2.4.4 UVE-PLS, UVE-PCR 和 UVE-LSSVM 模型对比

表 4 可以得出,在三种不同建模方法中, UVE-PLS 取得最好的结果。通过对比, UVE-PLS 模型的预测集 r 和 RMSEP 分别为 0.923 和 2.075, UVE-PCR 模型的预测集 r 和 RMSEP 分别为 0.916 和 2.159, UVE-LSSVM 模型的预测集 r 和 RMSEP 分别为 0.913 和 2.252。表明:在甲醇汽油

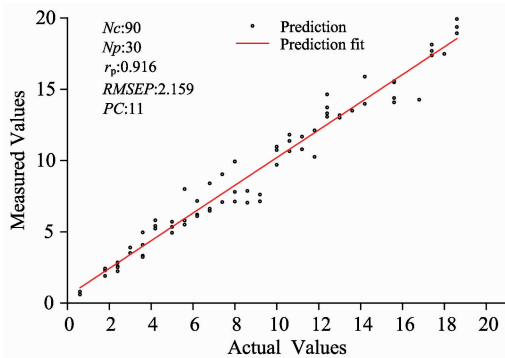


图 8 PCR 模型预测甲醇含量的散点图

Fig. 8 Scatter plots of methanol content prediction with PCR method

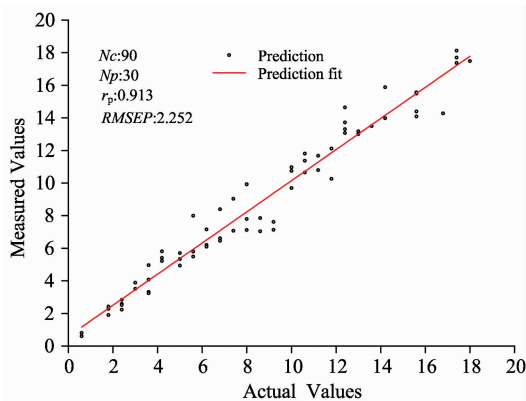


图 9 LS-SVM 预测甲醇汽油中甲醇含量的散点图

Fig. 9 Scatter plots of methanol content prediction with LS-SVM method

甲醇含量检测方面,应用中红外光谱结合 UVE-PLS 组合模型可以快速准确的预测甲醇含量,为进一步研究甲醇汽油其他指标奠定了基础。

表 4 UVE-PLS, UVE-PCR 和 UVE-LSSVM 建模的预测效果对比

Table 4 Prediction results by UVE-PLS, UVE-PCR and UVE-LSSVM

Model	Prediction	
	r	RMSEP
UVE-PLS	0.923	2.075
UVE-PCR	0.916	2.159
UVE-LSSVM	0.913	2.252

3 结 论

甲醇汽油是汽油替代燃料的一种,在汽油中添加甲醇能有效减少发动机污染物的排放,甲醇含量是衡量甲醇汽油品质的一个重要指标,所以对甲醇汽油中甲醇含量的检测显得尤为重要。应用中红外光谱建立甲醇汽油甲醇含量检测的 PLS 模型,不仅实现了对甲醇汽油中甲醇含量的快速无损检测,而且有较好的预测效果。在此基础上,利用无信息变量消除算法筛选得到的有效波长分别建立了 UVE-PLS, UVE-PCR 和 UVE-LSSVM 三种模型,最终获得甲醇汽油甲醇含量的最优模型为 UVE-PLS 模型。研究表明 UVE 是一种行之有效的波长筛选方法。中红外光谱在甲醇汽油甲醇含量快速无损检测是有效的,且模型的预测效果较好。该方法能够为甲醇汽油生产厂家提供一种新的检测方法,具有较强的现实意义。

References

- [1] LI Yan-ru(李雁如). Shanxi Chemical Industry(山西化工), 2014, 4: 38.
- [2] JIN Pan-pan, ZHANG Teng, WANG Min, et al(金盼盼, 张 腾, 王 珉, 等). Energy and Energy Conservation(能源与节能), 2012, 5: 3.
- [3] OUYANG Ai-guo, LIU Jun, WANG Ya-ping(欧阳爱国, 刘 军, 王亚平). Laser & Infrared(激光与红外), 2012, 8: 901.
- [4] YAO Jie, DAI Lian-kui, LIN Yi-ling(姚 捷, 戴连奎, 林艺玲). Chinese Journal of Light Scattering(光散射学报), 2013, 1: 59.
- [5] LI Yan-ru, SUN Rui-qing, WANG Yong-miao(李雁如, 孙瑞卿, 王永苗). Modern Chemical Industry(现代化工), 2013, 2: 113.
- [6] DING Hong-yuan, HUANG Rong-hua, WANG Zhao-wen, et al(丁红元, 黄荣华, 王兆文, 等). Journal of Huazhong University of Science and Technology · Natural Science Edition(华中科技大学学报·自然科学版), 2012, 6: 113.
- [7] JIN Pan-pan, ZHANG Teng, YANG Xiao-ping, et al(金盼盼, 张 腾, 杨晓平, 等). Auto Sci-Tech(汽车科技), 2012, 5: 66.
- [8] LIU Yan-de, WU Ming-ming, SUN Xu-dong, et al(刘燕德, 吴明明, 孙旭东, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2016, 6: 289.
- [9] ZU Qin, ZHANG Shui-fa, CAO Yang, et al(祖 琴, 张水发, 曹 阳, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2015, 35(2): 479.
- [10] HAN Zhong-zhi, WAN Jian-hua, LIU Kang-wei(韩仲志, 万剑华, 刘康伟). Chinese Journal of Analysis Laboratory(分析实验室), 2015, 11: 1268.
- [11] OUYANG Ai-guo, TANG Tian-yi, LIU Yan-de, et al(欧阳爱国, 唐天义, 刘燕德, 等). Chinese Journal of Luminescence(发光学报), 2016, 10: 1253.
- [12] ZHANG Rui-hua(张瑞华). Journal of Science and Technology Information(科技信息), 2014, (15): 298.
- [13] MAO Jia-wei, ZOU Yong, ZENG Ling-xi, et al(毛佳伟, 邹 勇, 曾令羲, 等). Shandong Chemical Industry(山东化工), 2015, (3): 85.

- [14] OUYANG Ai-guo, LIU Jun(欧阳爱国, 刘 军). Journal of Southwest China Normal University • Natural Science Edition(西南师范大学学报 • 自然科学版), 2012, 9: 98.
- [15] YAO Jie, DAI Lian-kui, LIN Yi-ling(姚 捷, 戴连奎, 林艺玲). Chinese Journal of Light Scattering(光散射学报), 2013, 1: 59.

Methanol Content Determination in Methanol Gasoline with Mid Infrared Spectroscopy Analysis

LIU Yan-de, HU Jun, TANG Tian-yi, ZHANG Yu, OUYANG Yu-ping, OUYANG Ai-guo
School of Mechatronics Engineering, East China Jiaotong University, Nanchang 330013, China

Abstract Methanol gasoline is a new type of environmentally friendly diesel fuel, and the performance and quality of methanol-gasoline are based on the diesel methanol content. In this work, mid-infrared spectroscopy was successfully used to evaluate the methanol contents in methanol-gasoline samples with the aid of chemometric approaches. First, the mid-infrared spectral data obtained were pre-processed by smoothing standard normal, multiple scatter correction (MSC), baseline correction and normalization, and the partial least-square (PLS) quantitative calibration models were established, and the best pre-processed method was found. It was found that the PLS model pre-processed by MSC was much better than others, and the r and RMSEP evaluated were 0.918 and 2.107, respectively. In order to simplify the model and improve the prediction accuracy, uninformative variable elimination (UVE) was used to select the optimal wavelengths, and the experimental results showed that the prediction ability was greatly improved. Different quantitative calibration models by UVE selected wavelength, such as partial least-square (PLS), principal component regression (PCR) and least square-support vector machine (LS-SVM) for measuring methanol content were established and their prediction results were compared. It was found that the UVE-PLS model was much better than others, and the r and RMSEP evaluated were 0.923 and 2.075. It suggested that infrared spectroscopy in the detection of methanol content in the methanol gasoline is feasible and can bring good prediction results. UVE is an effective method for methanol gasoline in the infrared spectrum of band selection method, which is significant for the development of oil chemical industry.

Keywords Mid-infrared spectroscopy; Ethanol gasoline; UVE; LS-SVM

(Received Dec. 23, 2016; accepted May 4, 2017)