

# 生化废水 COD 的近红外光谱法测定<sup>①</sup>

栗晖<sup>a</sup> 姚志湘<sup>②a</sup> 黎国梁<sup>a,b</sup> 兰月来<sup>a</sup>

<sup>a</sup>(广西工学院生物与化学工程系 广西柳州市城中区东环大道 268 号 545006)

<sup>b</sup>(广西大学化学化工学院 南宁市 530004)

**摘要** 采用近红外光谱测定在紫外光谱区表现为弱吸收响应的生化废水的化学需氧量(COD)。通过主成分分析-欧氏空间距离聚类筛选出具有代表性的样本,从系统独立变量数的角度判断了主成分数,结合变量标准化-基线扣除的预处理方法消除光谱信号背景干扰,建立了偏最小二乘校正模型。在系统主成分数为 6 时,水样的 COD 实测值与预测值之间的相关系数为 0.9348,预测标准差(RMSEC)为 61.89mg/L,模型有较好的预测能力。

**关键词** 化学需氧量;近红外光谱;偏最小二乘法

**中图分类号:** X 502; O 657.33 **文献标识码:** A **文章编号:** 1004-8138(2011)06-3040-04

## 1 引言

在水质监测中化学需氧量(COD)是最常使用的评价反映有机物的污染程度指标之一。废水中 COD 含量的国标法测定存在种种限制<sup>[1]</sup>,测定 COD 的方法操作复杂,测定时间长,无法及时了解水体污染的动态过程。近红外光谱测量分析技术已经在石化、制药等众多领域得到广泛的应用<sup>[2]</sup>,在水质监测方面,史永刚等<sup>[3]</sup>采用近红外光谱对水中有机污染物进行了快速鉴别,何金成等<sup>[4]</sup>探讨了近红外光谱分析测定废水 COD/生化需氧量(BOD)指标的可行性。本研究采用均匀设计得到乙醇、葡萄糖、蔗糖、谷氨酸和柠檬酸 5 种组分混合溶液的模拟生化废水样本,通过分析样本的近红外谱图中不同波长处的光密度大小及分布,采用偏最小二乘法建立了模拟废水中 COD 含量测定的定量分析模型,探讨在紫外区表现为弱吸收响应的物质的近红外建模方法。

## 2 实验部分

### 2.1 仪器与试剂

NIR 256 光纤光谱仪(美国海洋光学公司, InGaAs 检测器;波长范围:800—2500nm); Maya 2000 Pro 光纤光谱仪(美国海洋光学公司,波长范围:190—1100nm)。

重铬酸钾标准溶液:  $C(1/6K_2Cr_2O_7) = 0.500\text{mol/L}$ ; 硫酸银-硫酸溶液:  $\rho(Ag_2SO_4) = 10\text{g/L}$ ; 硫酸汞溶液:  $\rho(HgSO_4) = 24\text{g/L}$ ; COD 标准溶液为邻苯二甲酸氢钾溶液(100—1000mg/L)。葡萄糖、蔗糖、柠檬酸、谷氨酸、乙醇以上试剂均为分析纯。实验用水为二次蒸馏水。

① 广西自然科学基金资助项目(桂科自 0832064); 广西教育厅桂教科研资助项目(200708MS067)

② 联系人,电话:(0772) 2686670; E-mail: zxyao@gxut.edu.cn

作者简介:栗晖(1971—),女,广西临桂县人,高级实验师,硕士,主要从事生化过程分析技术研究工作。

## 2.2 实验方法

### 2.2.1 校准曲线绘制

在消解管中加入 1.00mL 重铬酸钾标准溶液和 6.00mL 硫酸银-硫酸溶液, 再分别加入 100、200、400、600、800、1000mg/L 的 6 种浓度的标准 COD 溶液 3.00mL, 拧紧管盖、摇匀溶液, 恒温  $165 \pm 2^\circ\text{C}$ , 计时加热 15min。冷却至室温, 以水为参比, 在  $600 \pm 20\text{nm}$  波长处测定吸光度, 绘制校准曲线。

### 2.2.2 模拟生化废水样本的配制

配制浓度分别为 2.5、2.0、2.0、1.0、0.79mg/mL 的葡萄糖、蔗糖、柠檬酸、谷氨酸、乙醇溶液。按  $U_{120}(120^\circ)$  表的设计, 分别量取各组分溶液于 25mL 容量瓶中, 用二次蒸馏水定容混合得到模拟废水样本。

### 2.2.3 样本的测定

每一水样分成两组: 一组测其近红外光谱, 另一组测其 COD, 把该化学值作为标定该水样光谱的“真值”。COD 值的测定参照 HJ/T 399-2007 环境行业标准<sup>[5]</sup>进行实验, 每个样本平行测定 3 次, 取平均值。近红外光谱采用透射光谱法采集, 用 10mm 玻璃比色皿作样品池, 空气作参比, 测量波长区间为 800—2500nm。

## 3 结果与讨论

### 3.1 实验样本的近红外光谱分析

图 1 为实验样本的 5 组分混合近红外光谱图。从图中可以看出, 在 900—1000nm 出现一个较小的吸收峰, 推断为—OH 伸缩振动三级倍频产生的吸收, 在 1000nm 附近为—CH 伸缩振动和变形振动的合频吸收。在 1200nm 的吸收峰可能是由—CH 基团伸缩振动的二级倍频产生, 而—NH 伸缩振动的二级倍频一般出现在 1570nm 左右。水分子在近红外谱区的吸收较强, 吸收带发布离散, 其中一些特性很强的合频吸收带(其他各级倍频相对较弱)分布在 1900—2100nm 和 2400nm 附近。由于水中的有机物对近红外光吸收很弱, 常被水的吸收峰掩盖, 或使水的吸收峰变宽, 或变窄, 不易从原始光谱图中直接观察到。故整个近红外谱区表现为水分子的光谱廓线, 仅在个别区域存在细微的差别, 即由水溶液中的有机物的弱吸收产生。

在近红外区域, 若以基频吸收判断特征吸收峰的常规方法进行分析是十分困难的, 因为在某一波长点可能有多种组分光谱的重叠, 难以区分不同组分。任意两种组分在许多波长点或者在近红外的全谱区内一般不可能完全相同, 另外, 由于各组分间的相互影响, 也可能会发生谱峰的漂移或变形。

### 3.2 建模样本的筛选

由于近红外光谱不仅包含了待测物质的吸收光谱信息, 而且还包含了背景、噪声等干扰信息, 光谱干扰的存在将严重影响样本选择的准确性。为提高样本选择的准确性和代表性, 光谱干扰应当预先被扣除。本研究运用主成分分析(PCA)降噪, 结合欧氏空间距离判断样本的分布状况, 进行建模样本的筛选。

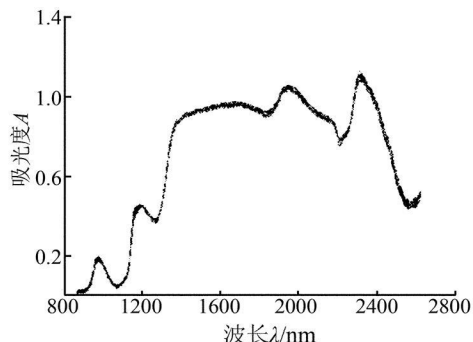


图 1 样本近红外光谱

本点。从图中可见,先采用 PCA 压缩数据,可达到降噪的效果,而且选出的样本点具有代表性,基本涵盖整个样本空间边界,只需少量的样本就足以代表整个空间的分布。

### 3.3 PLS 模型的建立和预测结果

运用筛选样本点算法取 67 个样本作为校正集,10 个作为预测集。建模采用校正集每个样本的各测量波长点的吸光度值构成吸收光谱矩阵  $X$ ,对应 COD 化学方法测量值作为因变量矩阵  $Y$ 。由  $X \rightarrow Y$  建立偏最小二乘(PLS)函数映射,回归模型的精度以均方根误差 RMSE 和线性相关系数  $r$  来评价。

吸光度矩阵用变量标准-基线扣除(SNV + Detrending)算法进行数据预处理。图 3 为经过预处理后的光谱图,这种预处理方法可在一定程度上排除背景干扰,提高光谱信噪比。

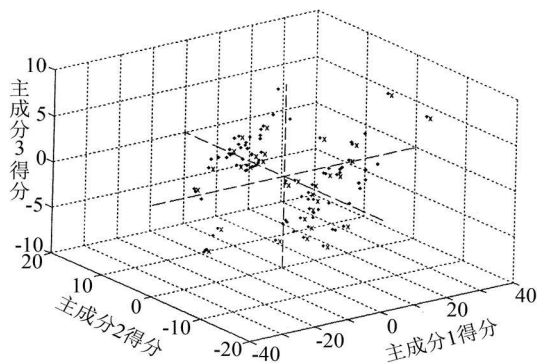


图 2 样本点空间分布

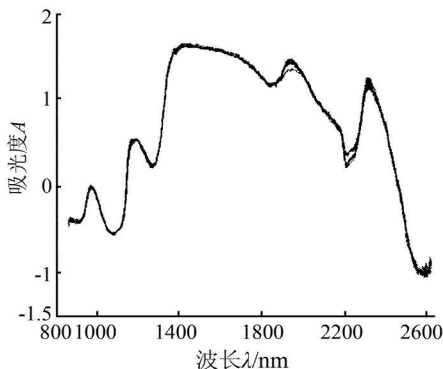


图 3 预处理后光谱

根据交互检验判据确定最佳主成分数,从图 4 看出,最佳主成分数为 6。从实际情况分析,实验测定是以空比色皿为参比,体系除 5 组分混合物外,还包括了溶剂水,水对近红外光吸收响应较敏感,周围环境的变化均会影响近红外光谱的测定。将预处理后的吸光度矩阵与 COD 值关联,建立 PLS 校正模型。从图 5 可知,预测值与实测值的线性相关系数为 0.9348,预测标准差(RMSEC)为 61.89mg/L。从预测结果可以看出,模型有较好的预测能力。

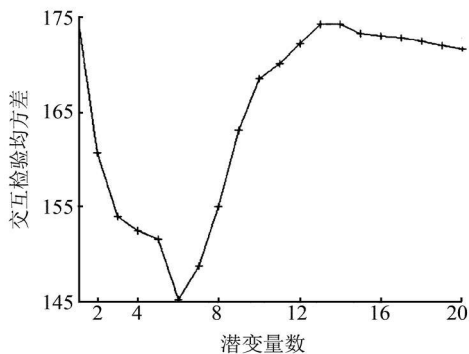


图 4 交互检验均方差与主成分数关系图

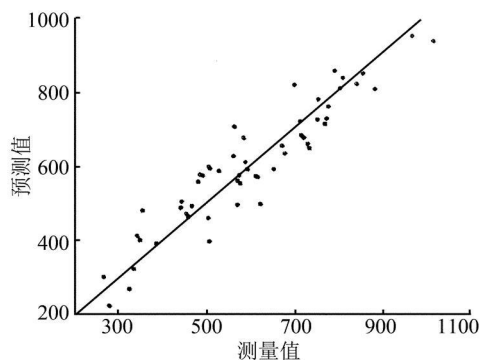


图 5 校正集模型的实测值与预测值关系

用此校正模型对预测集中的 10 个样本进行预测验证,结果如表 1 所示。由表 1 可见,利用模型计算的预测值与实测值之间的相对误差大部分在  $\pm 5\%$  以内,但也出现了偏差较大的点 - 9.21%。究其原因可能是由于近红外光谱响应较复杂,而本文所运用的 PLS 只是建立在全谱范围的初步计算,从而引入了一定程度的冗余信息和噪声干扰,导致某些测量样本点出现了较大误差。针对目前预测效果不够理想的情况,将对建模样本和计算的波段优选进行进一步的研究。

表 1 样本预测结果

样本序号	实测 COD 值(mg/L)	预测 COD 值(mg/L)	相对误差(%)
1	412.00	405.28	- 1.63
2	455.64	430.75	- 5.46
3	467.84	472.23	0.94
4	530.30	498.25	- 6.04
5	569.70	545.03	- 4.33
6	274.34	249.08	- 9.21
7	631.91	652.63	3.28
8	658.87	678.36	2.96
9	696.19	706.43	1.47
10	872.46	902.92	3.49

## 4 结论

饱和链烃小分子有机物在紫外光谱区表现为弱吸收响应,通过模拟这类废水的的样本,建立了近红外光谱法测定其 COD 值的方法,模型有较好的预测能力,满足分析要求。生化废水来源多种多样,且易受各外界条件的影响,要建立实用的 COD 近红外光谱法预测模型,需要收集足够数量的、有代表性的废水样品,同时在提高模型精度、改善不同类型废水的模型的自适应能力、保证恶劣环境下取样有效性等方面开展研究。

## 参考文献

- [1] 时红,孙新忠,范建华.水和废水测定方法[M].第4版.北京:中国环境科学出版社,2002.200—226.
- [2] 陆婉珍,袁洪福,徐广通.现代近红外光谱分析技术[M].北京:中国石化出版社,2000.14—36.
- [3] 史永刚,冯新沪,孙萍.水中有机污染物的近红外光谱快速鉴别[J].光谱实验室,2005,22(3):575—577.
- [4] 何金成,杨祥龙,王立人.近红外光谱法测定废水化学需氧量[J].浙江大学学报(工学版),2007,4(5):752—755.
- [5] 中国环境科学出版社编.水质 化学需氧量的测定 快速消解分光光度法[S].HJ/T 399-2007.北京:中国环境科学出版社,2008.

# Determination of COD in Biological and Chemical Wastewater by Near Infrared Spectrometry

SU Hui<sup>a</sup> YAO Zhi-Xiang<sup>a</sup> LI Guo-Liang<sup>a,b</sup> LAN Yue-Lai<sup>a</sup>

<sup>a</sup>(Department of Biology and Chemical Engineering, Guangxi University of Technology, Liuzhou, Guangxi 545006, P. R. China)

<sup>b</sup>(College of Chemistry and Chemical Engineering, Guangxi University, Nanning 530004, P. R. China)

**Abstract** The COD in biological and chemical wastewater that performed to weak absorption in the UV spectral region was determined by NIR spectrometry. Principal component analysis-Euclidean distance cluster was applied to select the representative sample. From the aspect of the number of independent variables, the number of principal components was determined. Combined with the SNV-Detrending preprocessing, the background signal of spectra was eliminated. The partial least squares model was built and the predictions were satisfied. When main component in the system amounted to 6, the linear correlation coefficient of measured values and predicted values was 0.9348 and the RMSEC was 61.89mg/L, the model had a better predictive ability.

**Key words** Chemical Oxygen Demand; Near-Infrared Spectrometry; PLS