

文章编号:1001-9014(2007)04-0317-04

## 近红外光谱透射法测量废水化学需氧量 (COD)的光程选择

何金成<sup>1,2</sup>, 杨祥龙<sup>1</sup>, 王立人<sup>1</sup>

(1. 浙江大学 生物系统工程与食品科学学院, 浙江 杭州 310029;  
2. 福建农林大学 机电工程学院, 福建 福州 350002)

**摘要:**在用近红外光谱法测量废水化学需氧量(COD)时,水是一个强干扰因素,如何选择合适的光程长以提高测量灵敏度,减少测量误差很重要.本文通过在傅里叶变换近红外光谱仪上用透射光谱法测量废水COD实验来讨论废水测量的光程选择.采集了60个养猪废水水样,分别用1mm、2mm、5mm和10mm的样品池作对比研究,用PLS回归建模,比较不同光程长对废水COD预测的效果,结果显示在800~1350nm谱区选用10mm的光程,在1350~2500nm谱区选用1mm或2mm的光程测量,使用双光程组合方法建模可以提高模型在800~2500nm谱区的预测性能.该光程组合方法对其它复杂组分的水溶液的近红外定量测量也具有一定的参考意义.

**关键词:**近红外光谱;光程;废水;化学需氧量

**中图分类号:**S;TH744.1 **文献标识码:**A

## PATHLENGTH SELECTION OF DETERMINING THE CHEMICAL OXYGEN DEMAND (COD) IN WASTEWATER BY USING NEAR-IR TRANSMISSION SPECTRA

HE Jin-Cheng<sup>1,2</sup>, YANG Xiang-Long<sup>1</sup>, WANG Li-Ren<sup>1</sup>

(1. College of Biosystem Engineering and Food Science, Zhejiang University, Hangzhou 310029, China;  
2. College of Mechanical and Electronic Engineering, Fujian Agriculture and Forestry University,  
Fuzhou 350002, China)

**Abstract:** The chemical oxygen demand (COD) in wastewater can be determined by near-infrared spectra (NIRs), however it is strongly influenced by the water absorbance. It is then very important to choose proper pathlength for improving measurement sensitivity and thus minimize the measurement error. In this study, we presented a series of experiments to determine the optimum pathlength for COD determination using a Fourier transform near infrared instrument with transmission measurement. Sixty wastewater samples for this study were collected from hoggery. Partial least squares (PLS) models were developed based on the spectral data which had been collected by using the cell pathlength of 1mm, 2mm, 5mm, and 10mm. The performances of the different models were compared. The results show that the pathlength of 10mm develops the optimal performance in the spectral range 800-1350nm and that of 1mm or 2mm of pathlength can be obtained in the spectral range of 1350~2500nm. In addition, our results suggest that two pathlength combination can improve the reliability and accuracy of the model in the region of 800~2500nm. This method of two pathlength combination may also be applied to the quantitative analysis of complex components in aqueous solutions.

**Key words:** near infrared spectroscopy; pathlength; wastewater; chemical oxygen demand (COD)

### 引言

化学需氧量(COD)是反映水体受有机物污染程度的最常用、最重要的综合指标之一.以近红外光

谱分析法代替传统的化学分析法具有测定迅速,时间短,便于在线、动态监测废水有机污染物的变化,测量操作简便,不消耗化学试剂,不产生二次污染等优点.近红外光谱测量废水COD的原理是通过用近

收稿日期:2006-10-12,修回日期:2007-01-09

Received date: 2006-10-12, revised date: 2007-01-09

基金项目:福建省自然科学基金(Z0513014)和浙江省教育厅(G20060346)资助项目

作者简介:何金成(1973-),男,福建建宁人,博士生,主要从事农业生物环境工程研究.

红外光谱测量废水的有机物总量,以此建立与 COD 的关系从而实现 COD 的测量<sup>[1]</sup>.

近红外光谱谱带宽,强度较弱而且谱峰重叠,要想从中提取微弱的信号进行分析,就必须选择合适的测量条件来提高测量灵敏度,即提高单位被测组分浓度变化所引起的被测光谱信号强度的变化,以减少仪器噪声引起的浓度测量误差.选择合适的光程长是提高测量灵敏度的重要方法.最佳光程长通常是通过吸光度的最佳值而间接获取.目前对最佳吸光度较为深入的研究主要集中在单波长光谱分析中,如 H. Willard 等<sup>[2]</sup>提出了忽略参考光谱噪声下的最佳吸光度为 0.434, Howard L. Mark 等<sup>[3]</sup>提出了考虑参考光谱噪声下的最佳吸光度为 0.4816.对于多波长的情况,目前多采用单一光程长进行测量, D. E. Honigs 与 T. Hirschfeld 等<sup>[4,5]</sup>在这方面作了一些初步的探讨.

对于水溶液,最佳光程长不依赖于溶质的性质,而是与溶剂(水)的摩尔吸收系数成反比例<sup>[6]</sup>.在短波近红外区(800~1100nm)水对近红外光的吸收较弱,为提高测量精度,必须使用长光程样品池,在长波近红外区(1100~2500nm),吸收强度大,须使用短光程样品池.不同的有机物在近红外短波区和长波区吸收带分布不一样,如酰氯、酯、酸酐、酰胺等的吸收带主要分布在长波区,羧酸、酚、醇、烃等在短波区和长波区均有吸收带<sup>[7]</sup>.废水中有机物成分复杂,要建立有效的近红外光谱的 COD 预测模型,必须充分利用近红外全谱区(800~2500nm)信息.本文就在 800~2500nm 谱区废水近红外光谱测量的光程长的选择作一个探讨.

## 1 材料与方法

### 1.1 仪器

使用 Nicolet 公司生产的傅里叶变换红外光谱仪(Nexus 870);分别采用 10mm、5mm、2mm 和 1mm 光程的石英比色皿作样品池,InGaAs 检测器.

### 1.2 水样

实验用的水样采集自杭州市某养猪厂的排出的废水,收集经生物塘处理后水样 60 个.

### 1.3 实验方法

现场采集来的水样稀释后经定性滤纸(中速,型号为 102)过滤后分为 2 组,一组测量其 COD 化学值,一组采集其近红外光谱. COD 化学值参照 GB11914-89(水质—化学需氧量的测定—重铬酸盐法)测定.近红外光谱采用透射光谱法采集.分别采

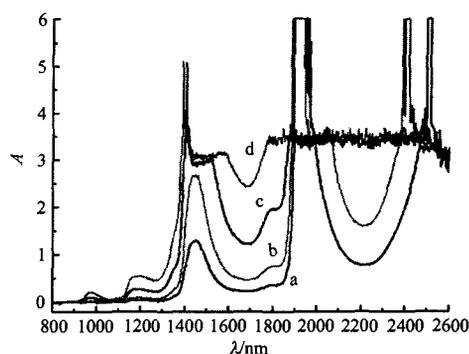


图 1 同一水样 4 种光程的近红外光谱图  
Fig. 1 Near infrared spectra of a wastewater sample obtained in various pathlength.

用 10mm、5mm、2mm 和 1mm 的石英比色皿作样品池,以空气为参比,分辨率为  $3.86 \text{ cm}^{-1}$ ,扫描区间 800~2500 nm ( $12500 \sim 4000 \text{ cm}^{-1}$ ),扫描次数为 32 次.

### 1.4 光谱数据的预处理及建模方法

对原始光谱数据采用窗口宽度为 5 个点的 Savitzky-Golay 滤波器滤波,去除高频噪音对信号的干扰,并选择基线线性去除及二阶微分来消除仪器背景或漂移对信号的影响.采用 Nicolet 公司的光谱定量分析软件 TQ Analyst v6 建立废水样品的偏小二乘法(PLS)回归预测模型.

## 2 结果与分析

### 2.1 水样化学值

重铬酸盐法测得 60 个水样 COD 值分布在 80~500mg/L,平均值为 210mg/L.

### 2.2 近红外光谱

图 1 为同一水样在 4 种光程长下采集到的近红外光谱,图中,a、b、c 和 d 分别为 1mm、2mm、5mm 和 10mm 光程长的水样近红外光谱.从图中可以看出,各波长处的吸光度随着光程长的增大而增强.10mm、5mm 光程长的光谱在 1800nm 以后出现全吸收带,1mm、2mm 光程长的在 1960nm 和 2500nm 附近也出现全吸收带.四张光谱具有相似的吸收峰,即在 960nm、1200nm 和 1410nm 附近有较明显的吸收峰,属于液态水的特征吸收,分别为水 O-H 伸缩振动的二级倍频、合频和一级倍频的吸收带.水中有机物对近红外光吸收很弱,常被水的强吸收峰掩盖,不易从原始光谱图中直接观察到.因此,很有必要通过选择合适的光程长来提高水中有机物单位被测组分浓度变化所引起的被测光谱信号强度的变化<sup>[2,3]</sup>,减少水的影响,从而提高透射测量灵敏度.

表1 800~1350nm 谱区吸光度频度分布

Table 1 Absorbance distribution of spectra from 800 to 1350nm

光程长	吸光度分段频度 (%)			
	<0.01	0.01~1.5	1.5~2.4	>2.4
1mm	72.8	27.2	0.0	0.0
2mm	55.4	44.6	0.0	0.0
5mm	30.9	69.1	0.0	0.0
10mm	7.3	92.7	0.0	0.0

表2 1350nm~2500nm 谱区吸光度频度分布

Table 2 Absorbance distribution of spectra from 1350nm to 2600nm

光程长	吸光度分段频度 (%)			
	<0.01	0.01~1.5	1.5~2.4	>2.4
1mm	0.0	75.1	7.4	17.6
2mm	0.0	42.8	21.2	36.0
5mm	0.0	18.1	17.3	64.6
10mm	0.0	0.8	3.4	95.9

表3 不同校正模型及其预测结果的比较

Table 3 Calibration and prediction results of several models for the wastewater samples

校正方法	主成分数	校正集		预测集	
		RMSECV (mg/L)	$r_c$	SEP (mg/L)	$r_p$
SMLR	/	67.6	0.62331	71.2	0.61101
PCR	4	48.2	0.77475	57.0	0.75327
PLS	3	40.3	0.84692	37.4	0.83882

表4 单一光程 PLS 建模结果

Table 4 Results of PLS model for single pathlength method

光程长	主成分数	校正集		预测集	
		RMSECV (mg/L)	$r_c$	SEP (mg/L)	$r_p$
1mm	2	38.3	0.83908	40.1	0.80109
2mm	2	37.5	0.82485	39.2	0.81121
5mm	2	36.2	0.84932	42.0	0.73882
10mm	3	40.3	0.84692	37.4	0.83882

注意到图1中1350nm附近是吸光度急剧变化的转折处,把光谱分为800~1350nm和1350~2500nm 2个区,对全部水样在4种光程长下光谱吸光度大小进行频度统计,统计结果如表1和表2所列。

表1中800~1350nm谱区4种光程下最大吸光度均小于1.5,其中1mm和2mm光程下的吸光度值主要分布在0.01以下,10mm和5mm光程下的吸光度值主要集中在0.01~1.5。表2中在1350~2500nm谱区,1mm和2mm光程下的吸光度值主要分布在0.01~1.5,而10mm和5mm光程下的吸光度值主要分布在大于2.4的区间。根据傅里叶变换

红外光谱仪厂商提供的参数,吸光度以0.01~1.5为佳。因此在800~1350nm谱区,可以选用10mm或5mm的光程,在1350~2500nm谱区可以选用1mm或2mm的光程。

### 2.3 各光程长近红外光谱 PLS 建模对比

在近红外区,吸收带宽,谱峰重叠,谱带精确归属困难,使用近红外光谱测量,必须借助多元校正方法把复杂的光谱数据转换为浓度数据。常用定量建模方法有多元线性回归(MLR)、主成分回归(PCR)、最小偏二乘法回归(PLS)等。在本课题的废水测量中,比较了多种预处理方法配合MLR、PCR和PLS等方法建模(表3为3种方法对滤波平滑后的二阶微分光谱建模结果),发现使用PLS对滤波平滑后的二阶微分光谱进行建模,校正、预测结果最好。因此,在本文中均采用滤波平滑后的二阶微分光谱用PLS法建立各光程下校正模型。

#### 2.3.1 单一光程建模

在60个样品中随机挑选40个样品作为校正集,采用“留一法”交叉验证选取主成分数建立PLS校正模型,其余20个样品作为预测集,用于检验模型的预测能力。4种光程下的建模参数和预测结果如表4,COD实测值与PLS模型预测值关系如图2所示。比较4种光程下校正模型的交叉验证校正标准差(RMSECV)和校正相关系数 $r_c$ ,以及预测结果——预测标准差(SEP)和预测相关系数 $r_p$ ,发现建模效果和预测能力并没有显著差异。

因此,在800~2500nm谱区如果使用单一光程来测量废水,可以选用1mm、2mm、5mm和10mm 4种光程的任意一种。考虑到样品池光程越短,相对光程误差越大,而且,光程越短,样品池清洗越困难,建议选用10mm光程。

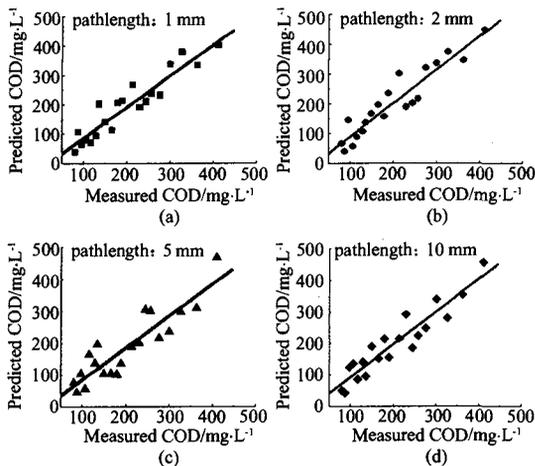
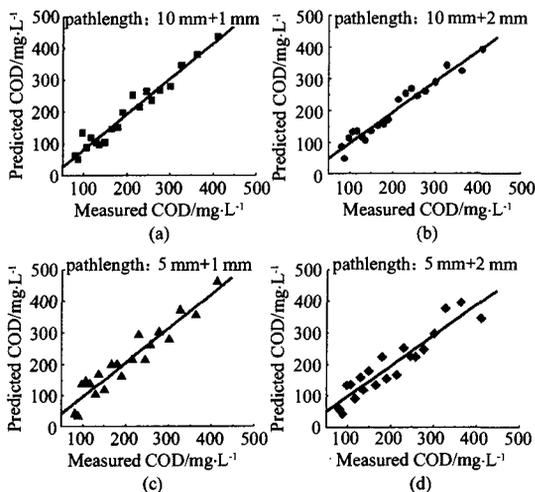
#### 2.3.2 双光程组合建模

校正集、预测集的选择方法以及主成分数的确定方法都与单一光程建模相同。根据表1和表2的吸光度频度计算结果,在800~1350nm谱区,选10mm或5mm的光程,在1350~2500nm谱区选1mm或2mm的光程,形成4种组合方式,即10mm和1mm、10mm和2mm、5mm和2mm、5mm和12mm光程的组合。采用自编软件把每一个水样800~1350nm谱区和1350~2500nm谱区两部分光谱数据组合成一张800~2500nm光谱图,然后导入TQ Analyst。建模计算结果如表5所列,COD实测值与PLS模型预测值关系如图3所示。

表5 双光程组合 PLS 建模结果

Table 5 Results of PLS model for the method of two path-length combination

光程组合方式	主成分分数	校正集		预测集	
		RMSECV (mg/L)	$r_c$	SEP (mg/L)	$r_p$
10mm + 1mm	5	20.4	0.95505	20.7	0.94976
10mm + 2mm	5	21.0	0.95101	21.1	0.93954
5mm + 1mm	3	31.6	0.90660	32.0	0.86694
5mm + 2mm	4	32.3	0.89018	31.2	0.86204

图2 四种光程下 COD 实测值与 PLS 预测模型预测值关系  
Fig.2 Measured COD versus predictive values at the pathlength of (a) 1mm (b) 2mm (c) 5mm and (d) 10mm for the PLS mode图3 双光程组合方式下 COD 实测值与 PLS 预测模型预测值关系  
Fig.3 Measured COD versus predictive values of PLS mode for the method of two pathlength combination

4种组合方式下,10mm和1mm光程组合方式效果最好,其次是10mm和2mm光程的组合.10mm和1mm光程的组合与10mm和2mm光程的组合建模预测能力相近,校正标准差(RMSEC)校正相关系数 $r_c$ 分别

为20.4mg/L, 21.0 mg/L和0.95505,0.95101;预测标准差(SEP)和预测相关系数 $r_p$ 分别为20.7mg/L, 21.1mg/L和0.94976, 0.93954.5mm和1mm光程的组合与5mm和2mm光程的组合建模预测能力相近,比较表4和表5,无论是校正集还是在预测集,双光程的4种组合方式下均比单一光程下性能更优越.

### 3 结语

在近红外全谱区(800~2500nm)由于水的摩尔吸光系数变化大,对于已知组分水溶液的测量可以根据组分的官能团所对应近红外吸收区选择相应的光程,但是测量如废水这类有机物成分复杂、未知的水样,单一光程测量难于保证测量灵敏度.通过对废水水样1mm、2mm、5mm和10mm等4种光程下的近红外光谱吸光度的频度分析及建模性能比较可知,在800~1350nm谱区选用10mm的光程,在1350~2500nm谱区选用1mm或2mm的光程测量,使用双光程组合方法建模,效果最好.10mm和1mm光程的组合与10mm和2mm光程的组合建模预测能力相近,考虑到光程误差、样品池的清洗等因素,建议采用10mm和2mm光程的组合建模方式.该光程组合方法简单、实用、经济,对其它复杂组分的水溶液的近红外测量具有一定的借鉴意义.

### REFERENCES

- [1] HE Jin-Cheng, YANG Xiang-Long, WANG Li-Ren, et al. Determination of chemical oxygen demand in wastewater by near-infrared spectroscopy[J]. *Journal of Zhejiang University (Engineering Science)* (何金成,杨祥龙,王立人,等.近红外光谱法测定废水化学需氧量的可行性研究.浙江大学学报(工学版)), 2007, 41(5): 50—54.
- [2] Willard H, Merrit L, Dean J. *Instrumental Methods of Analysis* [M]. New York: van Nostrand Co., 1983: 73—75.
- [3] Howard L, Mark Peter, Griffiths R. Analysis of noise in Fourier transform infrared spectra [J]. *Applied Spectroscopy*, 2002, 56(5): 633—639.
- [4] Hirschfeld T, Honigs D E, Hieftje G. Optimization of sample absorbance for quantitative analysis in the presence of pathlength error in the IR and NIR regions [J]. *Applied Spectroscopy*, 1985, 39(3): 430—433.
- [5] Honigs D E, Hieftje G M, Hirschfeld T. An optimization of detector-noise-limited infrared multiwavelength determinations [J]. *Applied Spectroscopy*, 1985, 39(2): 253—256.
- [6] Sergei Yu Venyaminov, Franklyn G Prendergast. Water (H<sub>2</sub>O and D<sub>2</sub>O) molar absorptivity in the 1000~4000cm<sup>-1</sup> quantitative infrared spectroscopy of aqueous solutions [J]. *Analytical Biochemistry*, 1997, 248: 234—245.
- [7] Phil Williams, Karl Norris. *Near-Infrared Technology in the Agricultural and Food Industries (Second edition)* [M]. St. Paul, Minnesota: American Association of Cereal Chemists, Inc., 2001: 25—36.