

紫外光谱法水体 COD 测量的线性-指数分段数学模型

田广军¹, 徐光耀¹, 田青²

1. 燕山大学电气工程学院, 河北 秦皇岛 066004

2. Department of Electrical and Computer Engineering, McGill University, QC H3A 2A7, Canada

摘要 利用紫外光谱法分析水体 COD 时, 当样品达到一定浓度之后(仍处于朗伯-比尔定律通常被适用的浓度范围内), 其溶液紫外吸光度与溶液 COD 数值会明显偏离线性关系, 此现象在许多学者发表的紫外光谱论文中都曾有所提及。对此, 选用海洋光学公司的 S2000 微型光纤光谱仪作为光谱测量仪器, 选用 PX-2 脉冲氙灯作为激发光源, 在温度为 20 °C(±0.5 °C)及湿度为 35%(±5%)的暗室中进行光谱实验, 测量了 COD 值为 40~680 mg·L⁻¹ 的 34 组邻苯二甲酸氢钾溶液样品的紫外吸收光谱, 并基于样品紫外吸收特性进行了分析建模。采用相关系数法选取优势波段, 通过对比样品紫外吸收光谱的第二特征峰和水质 COD 分析中常用波长处的 COD-吸光度关系曲线动态特性, 选定优势波长为 275 nm。采用逐点延伸的方式, 在较低浓度段利用鲁棒线性回归、较高浓度段利用非线性最小二乘回归, 反复拟合线性或指数方程, 滑动预测下一个数据点, 根据均方根误差和相对误差判断预设低浓度临界点和较高浓度临界点, 确定了低浓度段和较高浓度段 COD-吸光度关系模型的分段点分别为 300 和 560 mg·L⁻¹, 得到低浓度段模型和较高浓度段模型。通过在优势波长处进行低浓度段、较高浓度段和全浓度范围的鲁棒线性回归和非线性最小二乘法回归等不同模型的拟合精度比较, 表明 40~300 和 300~560 mg·L⁻¹ 范围内 COD-吸光度关系的线性-指数分段数学模型, 不仅拟合精度最高, 而且预测效果好, 低浓度段的预测均方根误差为 4.944 9, 较高浓度段的预测均方根误差为 6.768 9, 整体预测均方根误差为 5.664 7。研究结果对紫外光谱应用于较高 COD 的水质测量和分析具有一定的参考价值。

关键词 紫外吸收光谱; COD 测量; 模型临界点; 线性-指数分段模型

中图分类号: O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)06-1741-06

引言

化学需氧量(chemical oxygen demand, COD)作为衡量水体有机物相对含量的指标, 反映了水体受还原性物质污染的程度, 是水质监测的一个重要参数指标。目前, 国标法检测水质仍以常规化学方法为主, 应用也较为广泛, 但这类测量方法过程较为复杂, 测量时间较长, 测量结果不能及时反应水质的变化情况^[1-2]。紫外光谱分析法是依据光谱数据对污染物进行定性定量分析的纯物理光学检测方法, 相比于常规化学方法, 紫外光谱法由于设备结构简单、无二次试剂污染、实时响应速度快、易于实现在线定量分析等优点, 得到了广泛关注和深入研究^[3-5]。

紫外光谱法利用朗伯-比尔定律分析 COD 与吸光度关系时常选用邻苯二甲酸氢钾(分子量 M 为 204.22)溶液^[7-8], 在

实际溶液紫外光谱分析中, 在朗伯-比尔定律适用范围内达到某浓度值之后, 其 COD 与吸光度明显偏离了线性关系, 这一点在许多学者的研究文献中被论及和描述。李玉春在 254 和 280 nm 处测量的 10~450 mg·L⁻¹ COD 与吸光度关系曲线中, 254 nm 的数据在 COD 大于 200 mg·L⁻¹ 后开始上翘, 280 nm 的数据在 COD 大于 250 mg·L⁻¹ 后开始上翘; 李家琛^[6]在测量 25~666 mg·L⁻¹ 的九种不同浓度邻苯二甲酸氢钾溶液在 263 和 280 nm 处的吸光度时, 邻苯二甲酸氢钾溶液在 225 mg·L⁻¹ 后的浓度与吸光度关系曲线走势发生明显变化。

为了拓展光谱法溶液浓度测量范围上限, 在更宽范围内建立精确的溶液吸光度与 COD 关系模型, 本文研究了 COD 为 40~560 mg·L⁻¹ 的邻苯二甲酸氢钾溶液的紫外吸收光谱, 结合相关系数法、一元鲁棒线性回归和非线性最小二乘回归建立了吸光度-COD 的线性-指数分段数学模型, 实现了

收稿日期: 2019-05-14, 修订日期: 2019-09-18

基金项目: 国家自然科学基金项目(20577038), 河北省自然科学基金项目(E2019203524)资助

作者简介: 田广军, 1963 年生, 燕山大学电气工程学院教授 e-mail: tgj@ysu.edu.cn

邻苯二甲酸氢钾溶液 COD 的准确预测。

1 实验部分

1.1 材料与仪器

实验材料选用 $0.8502 \text{ mg} \cdot \text{mL}^{-1}$ 邻苯二甲酸氢钾标准溶液 (COD 值为 $1000 \text{ mg} \cdot \text{L}^{-1}$) 和重蒸水, 采用标准溶液稀释步骤配置得到 COD 值为 $40 \sim 680 \text{ mg} \cdot \text{L}^{-1}$ 的 34 个不同浓度的溶液样品。激发光源选用 Ocean Optics 公司的光谱仪专用脉冲氙灯 PX-2, 具有 $220 \sim 750 \text{ nm}$ 的稳定连续光谱输出; 光谱测量仪器选用 Ocean Optics 公司的 S2000 微型光纤光谱仪, 其光谱响应范围为 $200 \sim 1100 \text{ nm}$, 光学分辨率为 $0.3 \sim 10 \text{ nm FWHM}$; 透射支架选用 Ocean Optics 公司的 CUV-ALL-UV 4-WAY 比色皿支架, 实验用样品池均为紫外透射率 85% 以上的 10 mm 石英玻璃比色皿。

1.2 方法

实验在温度 $20 \text{ }^\circ\text{C} (\pm 0.5 \text{ }^\circ\text{C})$ 湿度 $35\% (\pm 5\%)$ 的暗室中进行, 实验系统方框图如图 1 所示。

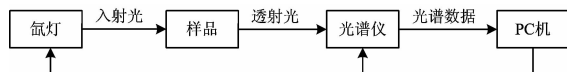


图 1 实验系统原理方框图

Fig. 1 Block diagram of experimental system

在光谱实验数据采集前, 先将光源打开预热 30 min 。设置积分时间为 10 ms 、平均光谱参数为 300 、Boxcar 平滑参数为 3 , 选择暗电流校正。按以下步骤使用 OOIBase32 光谱软件进行吸收光谱测量: 单击工具栏按钮“Scope Mode”, 检查并调整光信号强度, 使参考光信号的峰值强度为 3500 counts ; 单击工具栏上的“Store Reference Spectrum”按钮, 将当前参考光谱记录到计算机光谱实验文件; 遮挡光路并单击“Store Dark Spectrum”按钮, 保存实验系统的暗电流光谱文件以备暗电流校正用; 将样品放置在比色皿支架中, 单击“Absorbance Mode”按钮进入吸光度测量模式, 单击“Save”按钮并选择文件路径后, 将吸收光谱保存到计算机磁盘进行分析处理。

2 结果与讨论

分别配置 COD 值为 $40, 80, 160, 180, 190, 200, 210, 220, 240, 260, 280, 300, 320, 340, 360, 380, 400, 440, 480, 520, 560, 600, 640$ 和 $680 \text{ mg} \cdot \text{L}^{-1}$ 的 24 份样品作为校正集, 通过对校正集样品和邻苯二甲酸氢钾标准溶液原液的紫外吸收光谱扫描, 得到了较宽浓度范围的校正集样品 COD-吸光度紫外吸收光谱, 如图 2 所示。利用光谱仪系统软件内置的平滑滤波功能, 结合光谱观察分析, 除选取暗电流校正外, 还设置了适当大的时间窗口和波长窗口参数, 省去了附加滤波, 避免了实验室条件下对原始数据的过度预处理。

从图 2 可见, 样品的吸光度有明显的波长选择性, 即样品在 $220 \sim 300 \text{ nm}$ 之间有明显的吸收特性, 样品的吸光度在

320 nm 之后几乎为零。样品紫外吸收光谱存在两个较大的特征峰, 第一个特征峰在 233 nm 左右, 但随着样品 COD 浓度的增大, 第一个特征峰发生红移, 向着第二个特征峰的方向逐渐偏移和靠拢, 第二个特征峰的位置相对固定为 275 nm 。

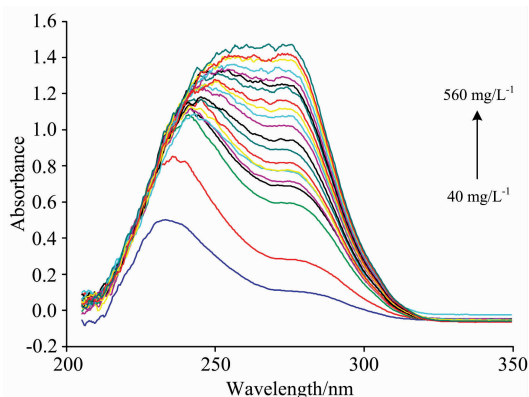


图 2 校正集样品 COD-紫外吸收光谱图

Fig. 2 COD-UV absorption spectra of samples in calibration set

2.1 选取优势波长

通过相关系数法计算样品 COD 与吸光度两个向量之间关系的相关程度。相关系数 R 由式(1)计算

$$R_{\lambda} = \frac{\sum_{i=1}^n (x_{i,\lambda} - \bar{x}_{\lambda})(y_i - \bar{y})}{\sum_{i=1}^n \sqrt{(x_{i,\lambda} - \bar{x}_{\lambda})^2 (y_i - \bar{y})^2}} \quad (1)$$

其中, $\bar{x}_{\lambda} = \frac{\sum_{i=1}^n x_{i,\lambda}}{n}$, $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$, $\lambda = 1, 2, \dots, m$, m 为波长数; $i = 1, 2, \dots, n$, n 为样品数。

用 Matlab 软件编程计算出样品 COD 值与吸光度的相关系数随波长变化曲线如图 3 所示。

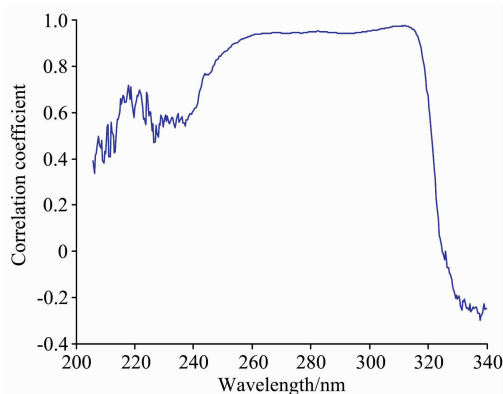


图 3 样品 COD 与吸光度的相关系数随波长变化曲线

Fig. 3 Correlation coefficient between absorbance and COD of sample

由图 3 可见, 样品 COD 与吸光度的相关系数是随波长变化的。对比图 2 与图 3 的波长变化范围可知, 虽然相关系数的变化波段比吸光度的变化波段有所扩展, 但可测范围主

要还是在 220~300 nm 之间。由于相关系数最大时的波长超出了图 2 中的吸收范围,选择相关系数大于 0.8 的波段为优势波段,对应的波长范围为 251.35~319.91 nm。由于第一个特征峰的红移现象,峰值波长不固定,且低浓度时第一个特征峰不在优势波段范围内,所以选择第二个特征峰 275 nm 作为备选优势波长。

由于紫外吸收光谱法常用 254 和 280 nm 分析水质 COD^[7-8],对比 254, 275 和 280 nm 三处 COD-吸光度数据走势如图 4 所示,同一样品在不同波长下吸光度不同,但是随着样品浓度增加,三个波长下数据变化趋势基本相同,低浓度阶段符合朗伯-比尔定律,浓度较高时样品 COD 与吸光度会偏离正比例关系,呈现类指数规律变化,由于 275 nm 波长对应的关系曲线具有较好的吸光度动态特性,所以选用实验所测的第二个特征峰波长 275 nm 作为优势波长进行分析。

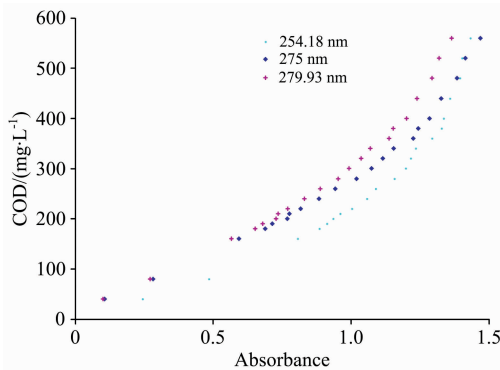


图 4 254.18, 275 和 279.93nm 波长的 COD-吸光度曲线走势对比

Fig. 4 Trend comparison of COD-absorbance data at 254.18, 275 and 279.93 nm

2.2 建立线性-指数分段回归模型

本文利用一元鲁棒线性回归方法建立低浓度段 COD 与吸光度关系的数学模型,对较高浓度段采用一元非线性最小二乘回归方法。回归算法分别采用 Matlab 工具箱中的 robustfit 函数和 nlinfit 函数。基于理论和实验分析,验证浓度较低时 COD 与吸光度成正比例关系,符合朗伯-比尔定律;确定浓度较高时吸光度与 COD 呈现的指数关系。

首先,计算各拟合模型对各假设模型临界点的下一个数据点的预测相对误差 R ,以 R 连续三次大于 5% 或小于 -5% 为判断条件,取第一次符合判断条件时对应的假设模型临界点为分段点,然后根据所确定的模型临界点对各段数据进行分段回归以获得更精确的模型。

2.2.1 低浓度段模型

以 COD 值 $40 \text{ mg} \cdot \text{L}^{-1}$ 为起点、 $190 \sim 300 \text{ mg} \cdot \text{L}^{-1}$ 的 8 个数据点为假设模型临界点分别进行一元鲁棒线性回归,得到各个回归模型的均方根误差 RMSE、对下一个数据点的预测 COD、预测相对误差 R 等参数如表 1 所示。

由表 1 可见,假设模型临界点为 $300 \text{ mg} \cdot \text{L}^{-1}$ 之后的预测相对误差 R 均小于 -5%,且偏差越来越大,即低浓度段中,COD 大于 $300 \text{ mg} \cdot \text{L}^{-1}$ 之后的溶液吸光度与 COD 关系,

已经不再符合线性回归方程。所以,选择 COD 为 $300 \text{ mg} \cdot \text{L}^{-1}$ 的数据点为低浓度段数据模型临界点。

表 1 各模型临界点的回归方程参数
Table 1 Regression equation parameters of critical point for each model

假设模型临界点/ ($\text{mg} \cdot \text{L}^{-1}$)	下一数据点 COD/ ($\text{mg} \cdot \text{L}^{-1}$)	预测 COD/ ($\text{mg} \cdot \text{L}^{-1}$)	$R/\%$	RMSE
190	200	197.777 1	-1.111	1.403 5
200	210	208.439 7	-0.743	1.734 1
210	220	217.612 6	-1.085	1.809 1
220	240	235.185 8	-2.006	1.968 7
240	260	250.938 0	-3.485	2.413 6
260	280	271.232 2	-3.131	3.398 8
280	300	286.033 0	-4.656	3.886 4
300	320	303.170 4	-5.259	5.130 7
320	340	320.596 7	-5.707	6.314 9
340	360	335.207 4	-6.887	7.566 9

在波长 275 nm 处对 COD 为 $40 \sim 300 \text{ mg} \cdot \text{L}^{-1}$ 的数据建立一元鲁棒线性回归模型,得到样品 COD 值 y 和吸光度 x 关系的回归方程为

$$y = 0.524 4 + 271.431 4x \quad (2)$$

回归方程曲线如图 5 所示,用均方根误差作为评价指标,校正模型的均方根误差为 5.130 7,对预测集的均方根误差 5.158 3。

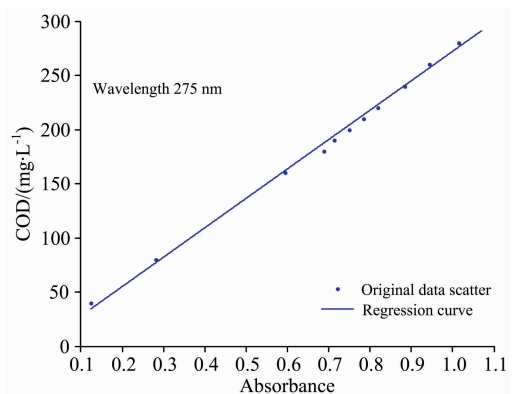


图 5 低浓度段样品 COD-吸光度关系回归曲线

Fig. 5 Regression curve of sample COD and its absorbance in low concentration range

2.2.2 较高浓度段模型

以 COD 值 $300 \text{ mg} \cdot \text{L}^{-1}$ 为起点、 $480 \sim 600 \text{ mg} \cdot \text{L}^{-1}$ 的 4 个数据点为假设模型临界点分别进行一元非线性最小二乘回归,得到各个回归模型的均方根误差 RMSE、对下一个数据点的预测 COD、预测相对误差 R 等参数如表 2 所示。

由表 2 可见,假设模型临界点为 $560 \text{ mg} \cdot \text{L}^{-1}$ 之后的预测相对误差 R 均小于 -5%,且偏差越来越大,即较高浓度段中,用 COD 大于 $560 \text{ mg} \cdot \text{L}^{-1}$ 之后的数据点进行一元非线性最小二乘回归的回归方程精度不理想,数据点再次发生

偏离。所以,选择 COD 为 $560 \text{ mg} \cdot \text{L}^{-1}$ 时的数据点为较高浓度段数据模型临界点。

表 2 各模型临界点的回归方程参数
Table 2 Regression equation parameters of critical point for each model

假设模型 临界点 COD/ ($\text{mg} \cdot \text{L}^{-1}$)	下一个数据 点 COD/ ($\text{mg} \cdot \text{L}^{-1}$)	预测 COD/ ($\text{mg} \cdot \text{L}^{-1}$)	R/%	RMSE
480	520	514.327 7	-1.091	7.175 9
520	560	548.158 0	-2.115	6.883 2
560	600	561.363 2	-6.439	7.117 5
600	640	582.097 9	-9.047	12.423 6
640	680	579.912 1	-14.719	22.224 0

在波长 275 nm 处对 COD 为 $300 \sim 560 \text{ mg} \cdot \text{L}^{-1}$ 的数据建立一元非线性最小二乘回归模型,得到样品 COD 值 y 和吸光度 x 关系的回归方程为

$$y = 2.182 5e^{1.595 2x+3.192 1} \quad (3)$$

回归方程曲线如图 6 所示,用均方根误差作为评价指标,校正模型的均方根误差为 7.117 5,对预测集的均方根误差 7.856 3。

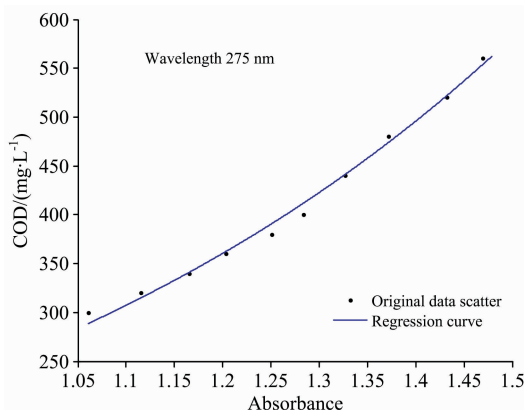


图 6 较高浓度段样品 COD-吸光度关系回归曲线

Fig. 6 Regression curve of sample COD and its absorbance in higher concentration range

令式(2)和式(3)相等,得到最终模型转折点的 COD 值 $y=288.350 3$,吸光度 $x=1.060 4$,则获得 $40 \sim 560 \text{ mg} \cdot \text{L}^{-1}$ 段样品 COD-吸光度关系的回归曲线如图 7 所示。

同时,由表 2 可见,在模型临界点 $560 \text{ mg} \cdot \text{L}^{-1}$ 后所建立的回归方程对下一个数据点 COD 的预测误差越来越大,进一步分析 COD-吸光度数据变化趋势可知,275 nm 处 COD 为 $560 \sim 1 000 \text{ mg} \cdot \text{L}^{-1}$ 的邻苯二甲酸氢钾溶液的吸光度为 $1.469 \sim 1.535$,只增大了 0.066,溶液浓度过高导致溶液吸光度接近饱和。

2.3 模型拟合精度对比

对校正集样品在 COD 值为 $300 \sim 560$, $40 \sim 560$ 以及 $40 \sim 560 \text{ mg} \cdot \text{L}^{-1}$ 的全范围内分别进行紫外吸收光谱的数据一元

鲁棒线性回归和一元非线性最小二乘回归,计算回归模型的均方根误差,得到各个模型拟合精度如表 3 所示,其中 RMSE 是校正集均方根误差。

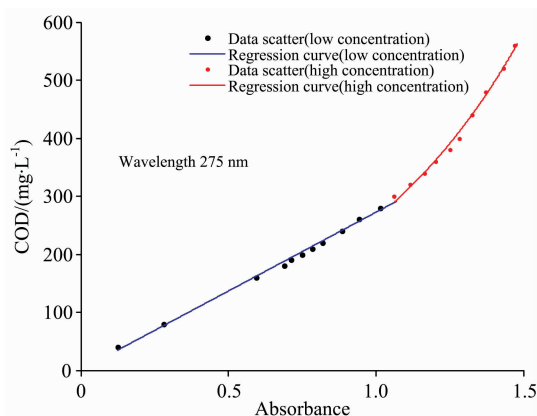


图 7 样品 COD-吸光度关系分段回归曲线

Fig. 7 Regression curve of absorbance and COD

表 3 不同模型的拟合精度

Table 3 Fitting accuracies of each model

样品 COD 值范围/ ($\text{mg} \cdot \text{L}^{-1}$)	建模类型	RMSE
40~300	鲁棒线性回归模型	5.130 7
	非线性最小二乘回归模型	10.085 8
300~560	鲁棒线性回归模型	14.109 8
	非线性最小二乘回归模型	7.117 5
40~560	鲁棒线性回归模型	41.175 0
	非线性最小二乘回归模型	12.373 1
	线性-指数分段模型	5.734 3

由表 3 可知,对全浓度段校正集样品紫外吸收光谱数据建模时,线性模型性能很差, RMSE 达到了 41.175 0;非线性模型性能也不好,其拟合数据在低浓度段时的 RMSE 达到了 10.085 8,约为线性-指数分段模型的两倍,所以不能直接使用非线性模型对水体 COD 进行测定。对较高浓度段校正集样品紫外吸收光谱数据建模时,线性模型的性能依然不好, RMSE 达到了 14.109 8,为线性-指数分段模型的两倍多,所以也不能采取两段线性模型对水体 COD 进行测定。线性-指数分段模型的拟合精度最高。

2.4 线性-指数分段模型预测效果

重新配置 COD 值为 $40 \sim 560 \text{ mg} \cdot \text{L}^{-1}$ 的 31 份样品作为预测集,用所建线性-指数分段模型对预测集样品的预测数据如表 4 所示,预测效果如图 8 所示。

用均方根误差作为评价指标,所得测量模型对预测集样本的预测效果如表 5 所示,其中 RMSE 是校正模型的均方根误差, RMSEP 是预测均方根误差。

由表 5 可见,所建线性-指数分段模型对预测集 275 nm 处 COD 范围为 $40 \sim 560 \text{ mg} \cdot \text{L}^{-1}$ 的溶液的预测效果较好,预测均方根误差仅为 5.664 7。

表 4 预测集各样品对应的预测数据
Table 4 Predictive data corresponding to the prediction set of each sample

预测集样品数据		COD 预测值/ (mg · L ⁻¹)	相对误差 R/%
COD 值/ (mg · L ⁻¹)	吸光度		
40	0.124	34.181 9	-14.545
60	0.221	60.510 7	0.851
80	0.292	79.782 4	-0.272
100	0.377	102.854 0	2.854
120	0.450	122.668 5	2.224
140	0.511	139.225 8	-0.553
160	0.597	162.568 9	1.606
180	0.681	185.369 2	2.983
190	0.715	194.597 9	2.420
200	0.746	203.012 2	1.506
210	0.777	211.426 6	0.679
220	0.823	223.912 4	1.778
240	0.883	240.198 3	0.083
244	0.864	235.041 1	-3.672
260	0.943	256.484 2	-1.352
280	1.020	277.384 4	-0.934
300	1.065	290.479 1	-3.174
320	1.118	316.106 7	-1.217
340	1.165	340.718 4	0.211
344	1.168	342.359 2	-0.477
360	1.226	375.540 5	4.317
380	1.231	378.547 8	-0.382
400	1.286	413.261 7	3.315
420	1.303	424.622 3	1.101
440	1.328	441.899 0	0.432
460	1.353	459.878 6	-0.026
480	1.384	483.192 5	0.665
500	1.397	493.317 6	-1.336
520	1.423	514.208 9	-1.114
540	1.455	541.139 7	0.211
560	1.469	553.361 2	-1.186

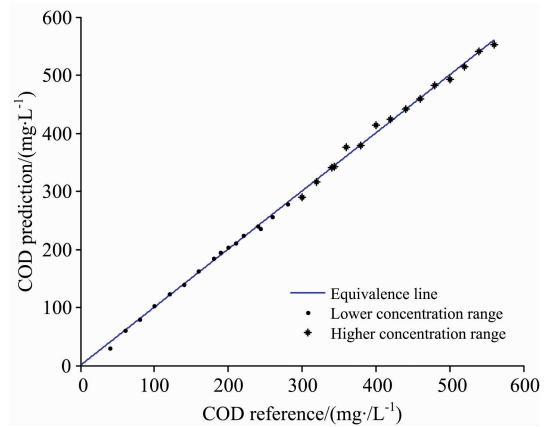


图 8 模型预测效果图

Fig. 8 Prediction effect diagram

表 5 测量模型预测效果

Table 5 Prediction effect of model

样本 COD 范围/ (mg · L ⁻¹)	优势波长/ nm	RMSE	RMSEP
40~300	275	5.130 7	4.944 9
300~560	275	7.117 5	6.768 9
40~560	275	5.734 3	5.664 7

3 结 论

邻苯二甲酸氢钾溶液的紫外吸收光谱实验结果分析表明,当溶液浓度增大到一定程度时,吸光度与 COD 值会明显偏离线性关系,低浓度段吸光度-COD 关系模型的临界点位于 COD 值 300 mg · L⁻¹ 附近,此临界点之后的吸光度-COD 关系已经不再符合朗伯-比尔定律,在 300~560 mg · L⁻¹ 较高浓度范围内 COD 与吸光度的关系呈指数变化规律。模型变化的内在机理及转折点的影响因素有待进一步探讨。本工作所建线性-指数分段测定模型对预测集样品的预测均方根误差仅为 5.664 7,预测效果较好。线性-指数分段数学模型的建立,提高了溶液 COD 紫外光谱测量方法的泛化能力,对于具有较高 COD 浓度的水体质量参数测量提供了可行性实验及理论依据。

References

[1] Storey M V, Van d G B, Burns B P. Water Research, 2011, 45(2): 741.
 [2] Kokkali V, Delft W V. TrAC Trends in Analytical Chemistry, 2014, 61: 133.
 [3] Guan L, Tong Y, Li J, et al. Optik, 2018, 164: 277.
 [4] Figueiró C S M, Bastos de Oliveira D, Russo M R, et al. Aquaculture, 2018, 490: 91.
 [5] Carré E, Pérot J, Jauzein V, et al. Water Science & Technology A Journal of the International Association on Water Pollution Research, 2017, 76(3): 633.
 [6] Li J C, Huang P J, Hou D B, et al. Applied Mechanics & Materials, 2013, 316-317(1): 606.
 [7] Guo X J, Xi B D, Yu H B, et al. Water Science and Technology, 2011, 63(5): 1010.
 [8] Chen H, Zheng B. Water Science and Technology, 2012, 65(5): 962.

A Piecewise Mathematical Model for COD Measurement of Water by UV Spectrometry

TIAN Guang-jun¹, XU Guang-yao¹, TIAN Qing²

1. School of Electrical Engineering, Yanshan University, Qinhuangdao 066004, China

2. Department of Electrical and Computer Engineering, McGill University, QC H3A 2A7, Canada

Abstract The relationship between the UV absorbance and the COD significantly deviate from the linear relationship when the samples reach a certain concentration (still within the concentration range where Lambert-Beer's law is often applied) when using UV spectroscopy to analyze COD in water, this phenomenon has been mentioned in many scholars' literature on UV spectroscopy. In this regard, the S2000 micro-fiber spectrometer and the PX-2 pulsed xenon lamp of Ocean Optics were selected for experiments in a dark room with a temperature of 20 °C (± 0.5 °C) and humidity of 35% ($\pm 5\%$), and UV absorption spectra of 34 groups of potassium hydrogen phthalate solutions with COD values of 40~680 mg · L⁻¹ were measured for analytical modeling. The dominant wavelength band was selected by correlation coefficient method combine with the UV absorption characteristics of the sample, and the dominant wavelength was determined to be 275 nm after compared the dynamic characteristics of the COD-absorbance curves at the second characteristic peak of the ultraviolet absorption spectrum of the sample with the wavelength commonly used in water quality COD analysis. A point-by-point extension method is used, which select robust linear regression and the unary nonlinear least squares regression in low concentration segment and higher concentration segment, respectively, repeated fitting of linear or exponential equations and sliding prediction of the next data point, and the segmentation points of the COD-absorbance relationship model of the low concentration segment and the higher concentration segment are determined to be 300 and 560 mg · L⁻¹, respectively, according to the root mean square error and the relative error, and a low concentration segment model and a higher concentration segment model were obtained. The comparison of the fitting accuracy of the linear or nonlinear model in a low concentration range, higher concentration range and the full concentration range at the dominant wavelength indicates that the linear-exponential segmentation model for the relationship between absorbance and COD of water quality in the range of 40 to 300 mg · L⁻¹ and 300 to 560 mg · L⁻¹ has the highest precision, and the prediction effect on the prediction set samples is better: the prediction root mean square error of the model at low concentration is 4.944 9, 6.768 9 at higher concentration and 5.664 7 overall, the prediction effect is ideal for the prediction set. The research result provides a reference value for the measurement and analysis of water with higher COD by UV spectroscopy.

Keywords Ultraviolet absorption spectrum; COD measurement; Model critical point; Linear-exponential piecewise model

(Received May 14, 2019; accepted Sep. 18, 2019)