

基于表面增强拉曼光谱技术和 GWO-SVR 算法 快速实现水中总氮总磷含量检测

张燕君, 康成龙, 柳雅倩, 付兴虎*, 张金霄, 王明学, 杨刘震

燕山大学信息科学与工程学院, 河北省特种光纤与光纤传感重点实验室, 河北 秦皇岛 066004

摘要 提出了一种将表面增强拉曼光谱技术(SERS)和基于灰狼优化(GWO)算法的支持向量回归(SVR)相结合快速定量检测水中总氮(TN)、总磷(TP)含量的定量分析方法。传统的 TN、TP 检测方法不但过程繁杂, 实验环境要求高, 而且耗时较长, 不能实现快速检测。而 SERS 技术操作简单, 耗时短, 将其与 GWO-SVR 算法相结合可以实现快速精确检测。以实验室配制的银溶胶作为拉曼光谱增强基底, 不同浓度梯度 TN、TP 溶液为研究对象, 分别配制 TN、TP 样本溶液 26 组和 23 组, 其中 TN 溶液选取 8 组作为测试集, TP 溶液选取 7 组作为测试集, 剩余样本溶液作为训练集。根据待测溶液与银溶胶不同体积配比确定最佳实验方案, 将 TN、TP 分别与银溶胶进行 1:1, 1:2, 1:3, 2:1 和 3:1 的体积比混合, 结果表明当待测溶液与银溶胶以 2:1 比例混合时增强效果最佳。采集光谱信息并对特征峰进行归属, 然后采用暗电流扣除、背景扣除(基线校正)和平滑处理对原始光谱数据进行预处理。经光谱分析结果可知, 由于不同浓度溶液官能团浓度差异, 光谱特征峰强度随溶液浓度变化而变化。以训练集样本溶液光谱特征峰强度和溶液浓度值作为回归预测模型的输入值和输出值, 建立 GWO-SVR 定量分析模型。通过测试集样本溶液的相关系数(r)和均方误差(MSE)对模型的预测能力进行分析, 并将 GWO-SVR 模型和其他两种模型进行对比。结果表明, GWO-SVR 模型对 TN 溶液预测的相关系数为 0.999 5, 均方误差为 0.005 8, 高于人工蜂群算法优化支持向量回归(ABC-SVR)和粒子群算法优化神经网络(PSO-BP)的 0.993 8, 0.052 7 和 0.998 3, 0.022 7。对 TP 溶液预测的相关系数为 0.998 5, 均方误差为 0.037 6, 也均高于另外两种模型。而且与 ABC-SVR 和 PSO-BP 模型相比, GWO-SVR 定量分析输入参数更少, 收敛速度更快, 更容易找到全局最优解。因此, 该方法可以实现对水中 TN、TP 含量的快速准确检测, 为水质检测提供了新方法。

关键词 表面增强拉曼光谱; 灰狼优化; 支持向量回归; 总氮; 总磷

中图分类号: TN247 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2021)10-3147-06

引言

随着科技不断地发展, 人类活动范围不断地扩大, 生活污水、工业和农业废水的排放量激增, 从而对环境的影响也越来越严重, 其中水质污染尤其严重。水中总氮(total nitrogen, TN)、总磷(total phosphorus, TP)是评价水体受到污染严重程度的关键性指标^[1]之一。当水域中 TN、TP 含量过高时, 会造成水体富营养化严重, 致使藻类和浮游生物过度繁殖, 水体透明度下降, 水质甚至会恶化至有害的程度, 最终

对人类的生活产生危害, 因此快速准确测定水体中 TN、TP 的含量十分必要。

目前对于 TN、TP 测定的常规方法有很多, 比如国家标准法即碱性过硫酸钾氧化-紫外分光光度法^[2]和钼酸铵分光光度法。但是这些方法适用范围较小, 操作繁琐, 对实验环境要求高且检测时间长。近年来光谱技术在水质检测领域得到广泛的应用, 在检测重金属离子浓度^[3]和 COD 含量^[4]方面较为成熟, 但在水中氮磷含量检测^[5]方面少有报道。而 SERS 技术^[6-7]能够实现快速检测, 操作简单, 相较于普通拉曼光谱技术, SERS 技术通过制备金属纳米粒子(如 Au, Ag

收稿日期: 2020-09-11, 修订日期: 2021-01-15

基金项目: 国家自然科学基金项目(61675176), 国家海洋局多功能海洋风电安装平台创新示范项目, 河北省重点研发计划项目(18273302D)资助

作者简介: 张燕君, 女, 1973 年生, 燕山大学信息科学与工程学院教授 e-mail: yjzhang@ysu.edu.cn

* 通讯作者 e-mail: fuxinghu@ysu.edu.cn

和 Cu 等)作为增强基底来吸附待溶液中某些分子,实现拉曼信号增强,增强效果甚至可以达到 $10^4 \sim 10^6$ 倍,并且 SERS 频道窄,水溶液对光谱信号干扰弱,可检测到的光谱信息丰富,非常适用于水质的在线快速检测^[8]。为了满足水中氮磷含量检测的高精度、高效率等要求,本文提出应用 SERS 技术结合灰狼优化支持向量回归(grey wolf optimizer-support vector regression, GWO-SVR)算法建立水中氮磷浓度预测模型。SVR 在处理非线性回归问题上是一种比较高效优越的算法,将其与其他智能算法相结合会使得预测模型更加准确高效。GWO 通过优化 SVR 的惩罚因子 C 和核参数 g 建立训练模型,避免了陷入过拟合和局部最优等问题,提高了模型的预测精度。

1 算法原理

1.1 支持向量回归算法

支持向量回归(SVR)是在支持向量机(support vector machine, SVM)基础上扩展出来的一种监督性机器学习算法,用于处理数据回归预测方面的问题^[9]。对于给定的样本数据 $\{(x_i, y_i), i=1, 2, \dots, n\}$ $x_i \in R_n, y_i \in R_n$ 这里 $x_i \in R_n$ 是第 i 个样本的输入值, $y_i \in R_n$ 是第 i 个样本的输出值, n 为训练样本的数目,假定回归函数为

$$f(x) = \omega\varphi(x) + b \quad (1)$$

式(1)中, $\varphi(x)$ 为非线性映射函数, ω 和 b 为要求解的参数。

对于 SVR 来说,预测模型其样本点和输出值之间都会有一定误差,引入不敏感系数 $g > 0$, 当样本点和回归模型足够近,将误差控制在一定范围内,这时该样本不计入损失。引入松弛变量 $\xi_i, \xi_i^* \geq 0$, 则要解决的问题就转换为

$$\begin{cases} \min(\omega, b) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ \text{s. t. } y_i - \omega\varphi(x) - b \leq g + \xi_i \\ \text{s. t. } \omega\varphi(x) - y_i + b \leq g + \xi_i^* \\ \text{s. t. } \xi_i \geq 0, \xi_i^* \geq 0 \quad i = 1, \dots, n \end{cases} \quad (2)$$

式(2)中, $\|\omega\|^2$ 为惩罚函数; C 为惩罚因子,训练误差受 C 影响; g 则影响着回归函数的精度。因此 C 和 g 是评价函数的重要指标,也是 GWO 算法要进行寻优的两个参数。

通过引入拉格朗日因子 a_i 和 a_i^* 构造拉格朗日函数,可以将上式优化问题转化为求解对偶问题

$$\begin{cases} \max_{a, a^*} \left[\sum_{i=1}^n (a_i - a_i^*) y_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (a_i - a_i^*) \cdot \right. \\ \left. (a_j - a_j^*) K(x_i, x_j) - \sum_{i=1}^n (a_i + a_i^*) g \right] \\ \text{s. t. } \sum_{i=1}^n (a_i - a_i^*) = 0 \\ \text{s. t. } 0 \leq a_i \leq C, 0 \leq a_i^* \leq C \end{cases} \quad (3)$$

式(3)中, $a_i, a_i^* \in [0, C]$; $K(x_i, x_j)$ 为核函数。与其他核函数相比,径向基函数在处理非线性问题时预测误差小、拟合精度高^[10]。所以本文采用径向基函数,最后得到 SVM 回

归估计模型为

$$f(x) = \sum_{i=1}^n (a_i - a_i^*) K(x_i, x_j) + b \quad (4)$$

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (5)$$

1.2 灰狼优化算法

灰狼优化算法(GWO)是一种群体智能优化算法,该算法模拟了自然界中灰狼的狩猎机制和领导等级^[11]。在建立灰狼等级模型时,将具有最优适应度值的狼群个体依次标记为 α, β, δ , 剩下的标记为 ω 。在寻优过程中,首先由三只头狼 α, β, δ 搜寻猎物位置,逐渐接近猎物,然后带领 ω 包围目标。

$$D = |CX_p(t) - X(t)| \quad (6)$$

$$X(t+1) = X_p(t) - AD \quad (7)$$

式(6)和式(7)中, t 为当前迭代次数, A 和 C 为协同系数向量; X_p 是猎物位置; X 表示当前灰狼位置。当狼群开始包围猎物时,随着目标猎物位置发生变化, α, β, δ 的位置也在不断的迭代变化,为获取最优值更新公式为

$$\begin{cases} D_\alpha = |C_1 X_\alpha(t) - X(t)| \\ D_\beta = |C_2 X_\beta(t) - X(t)| \\ D_\delta = |C_3 X_\delta(t) - X(t)| \end{cases} \quad (8)$$

$$\begin{cases} X_1 = X_\alpha(t) - A_1 D_\alpha \\ X_2 = X_\beta(t) - A_2 D_\beta \\ X_3 = X_\delta(t) - A_3 D_\delta \end{cases} \quad (9)$$

$$X(t+1) = \frac{X_1 + X_2 + X_3}{3} \quad (10)$$

式(8)一式(10)中, $X_\alpha(t), X_\beta(t)$ 和 $X_\delta(t)$ 分别是当前种群迭代 t 次后 α, β, δ 位置向量; $D_\alpha, D_\beta, D_\delta$ 分别表示迭代 t 次后当前候选灰狼与最优三条狼的之间距离; $X(t+1)$ 为迭代 $t+1$ 次后灰狼的位置。灰狼群通过迭代过程不断地更新位置,逐步的逼近猎物,直到算法迭代完成,成功捕获猎物获取最优解。

1.3 GWO-SVR 模型的建立

GWO-SVR 算法^[12]的核心思想是利用 GWO 的快速搜索能力,实现对 SVR 参数的自动优化,以提升模型的预测效果^[13]。首先根据所建模型需求确定输入数据和输出数据并对数据集进行预处理。然后对 GWO 算法设置初始化参数,初始化狼群数量 $M=20$, 迭代次数 $T=10$ 。根据要优化的模型和数据集的需求设置对应 SVR 模型的可调节参数数目和参数的取值范围,本文 SVR 所要优化的参数为 C 和 g , 取值范围为 $0.01 \sim 100$ 。

适应度函数是评价参数性能的重要指标,选取 SVR 算法训练集样本的均方误差(MSE)作为适应度函数, MSE 的数值越小则认为目标参数值越优,适应度函数式为

$$\text{fitness} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (11)$$

式(11)中, y_i 为实际值, \hat{y}_i 为预测值, N 是测试集样本数量,当 fitness 到达最小值时得到最优参数。利用 GWO 算法对 SVR 参数寻优的建模过程如图 1 所示。

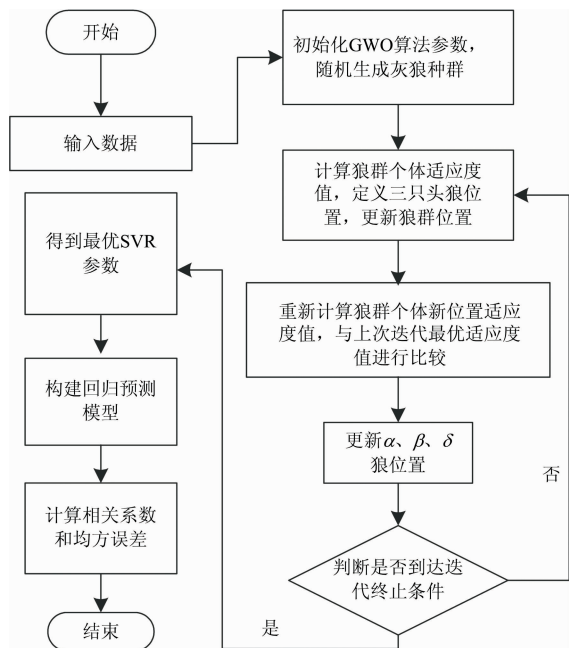


图 1 GWO-SVR 算法流程图

Fig. 1 GWO-SVR algorithm flow chart

2 实验部分

2.1 仪器和样本

实验使用的仪器是必达泰克公司生产的 BWS465-785S 便携式拉曼光谱仪, 选取激发波长 785 nm 作为激发光源, 激发功率设置为 300 mW, 在光谱范围为 175~3 200 cm^{-1} 范围内对待测样本进行测量, 积分时间设置为 10 000 ms, 每个样本检测 3 次取平均值。配合光谱仪自带软件 BWRam4TM 进行光谱预处理, 然后采集光谱数据获取光谱特征值建立模型, 测试其预测功能并分析。

实验采用的 TN 和 TP 标准液为环境保护部标准样品, 用 18 兆欧超纯水将 TN 和 TP 样本溶液分别稀释配制不同浓度梯度的待测溶液。TN 和 TP 浓度范围分为 1.5~15 和 2~15 $\mu\text{g} \cdot \text{mL}^{-1}$, 样本数为 26 和 23。对每个配制样本充分混合均匀, 放在 10 mL 离心管中密封避光保存。

2.2 纳米银溶胶的制备

硝酸银 (AgNO_3) 和柠檬酸钠 ($\text{C}_7\text{H}_5\text{Na}_3\text{O}_7$) 购置于国药集团化学试剂有限公司。取 1 mL 浓度为 1% 的硝酸银溶液于锥形瓶中并用超纯水定容至 200 mL, 进行加热并不断搅拌, 使硝酸银充分溶解。当水沸腾后迅速加入 3 mL 浓度为 1% 的柠檬酸钠溶液, 继续加热搅拌 60 min, 直到变化为乳白色即银溶胶制备完成。将制备好的银溶胶放置高硼硅试剂瓶中冷却至室温后, 4 $^{\circ}\text{C}$ 冷藏保存。

2.3 方法

用一次性滴管吸取配制好的样本溶液 2 mL 滴入试管中, 然后吸取配制好的银溶胶 1 mL 缓慢滴入试管中, 轻轻摇匀, 避光静置 30 min, 使其充分混合后进行检测。实验时, 打开激光拉曼光谱仪预热 15 min, 然后将配制好的加入了银溶胶的待测溶液用一次性滴管取 3 mL 放入比色皿中, 盖上

遮光罩制造黑暗环境进行检测。为使 SERS 基底增强效果理想化, 现通过对比实验探究基底与待测溶液的最佳体积混合比。将 TN 和 TP 分别与银溶胶进行 1:1, 1:2, 1:3, 2:1, 3:1 的体积比混合, 静置 30 min 后观察其拉曼图谱。图 2 (a)和(b)分别为不同比例 TN 和 TP 的 SERS 光谱图。由图可知在相同条件下, TN 和 TP 溶液均与银溶胶以 2:1 混合时增强效果最好。

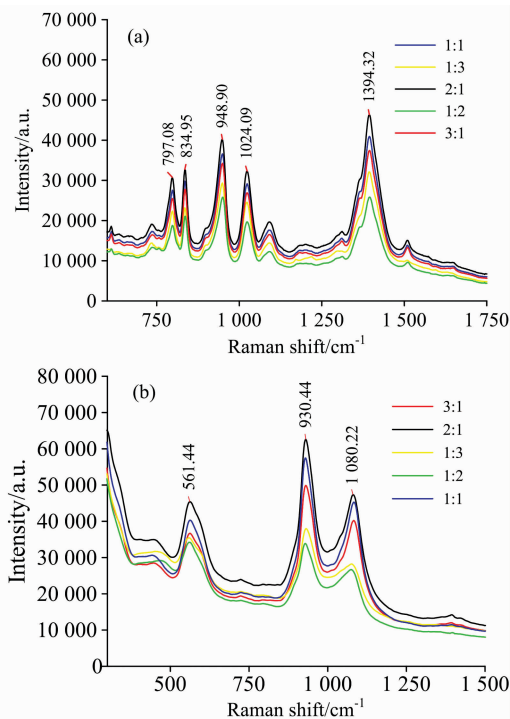


图 2 表面增强拉曼光谱图

(a): TN; (b): TP

Fig. 2 Surface enhanced Raman spectrum

(a): TN; (b): TP

TN 和 TP 的主要拉曼特征峰的相关信息见表 1 和表 2。在检测浓度范围内光谱中特征峰的峰值强度会随着分子式基团浓度的改变而改变, 峰值强度越高表示此处物质含量越高。通过配制不同浓度梯度溶液加入银溶胶进行检测, 得到 TN 和 TP 溶液能检测到特征峰的最低浓度分别为 1.5 和 2 $\mu\text{g} \cdot \text{mL}^{-1}$, 当继续降低待测溶液浓度时并不能检测到明显的特征峰。实验检测选取稳定性较好的 TN 溶液 948.90 cm^{-1} 处特征峰和 TP 溶液 930.44 cm^{-1} 处特征峰的峰值强度作为训练模型的输入值, 以其对应的浓度值作为模型的输出值。

2.4 光谱预处理

在实验过程中, 由于受到样本的荧光背景和仪器的系统噪声影响, 光谱图出现基线漂移和峰值淹没导致光谱信噪比降低, 所以需对光谱进行预处理。本文采用暗电流扣除、光谱背景扣除(基线校正)和光谱平滑对原始光谱进行预处理, 各光谱在测量之前用 BWRam4TM 软件去除暗电流。图 3(a)为原始光谱和进行光谱背景扣除(基线校正)后的光谱。图 3 (b)为光谱平滑处理, 选用 Savitzky-Golay Filters 平滑法进行

光谱平滑处理, 窗口尺寸大小选择 5 作为光谱平滑的数据点。

表 1 TN 的拉曼光谱特征峰

Table 1 The characteristic Raman spectral peak of TN

特征峰位/ cm^{-1}	官能团	振动方式
797.08	C=C—R1	伸缩振动
834.95	—CH/—CO	平面弯曲振动
948.90	CH ₃ /CH	平面摇摆振动/平面外弯曲振动
1 024.09	邻-双取代苯	伸缩振动
1 394.32	N—H	平面内弯曲振动

表 2 TP 的拉曼光谱特征峰

Table 2 The characteristic Raman spectral peak of TP

特征峰位/ cm^{-1}	官能团	振动方式
561.44	O—P—O	反对称角振动
930.44	P—O	对称伸缩振动
1 080.22	P—O	反对称角振动

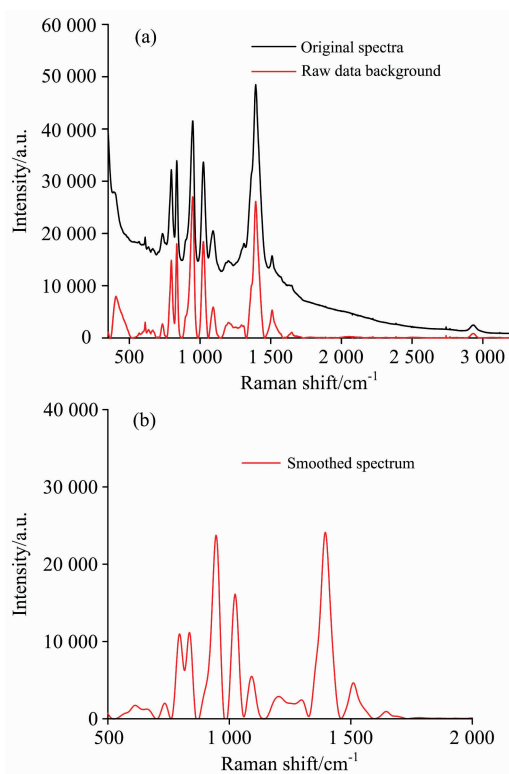


图 3 光谱预处理

(a): 背景扣除; (b): 平滑处理

Fig. 3 Spectra pretreatment

(a): Background removal; (b): Smoothed

3 结果与讨论

分别从配制好的 TN 和 TP 样本溶液中选取 8 组和 7 组作为测试集, 剩余溶液作为训练集。将处理过的训练集峰值强度作为模型的输入, TN 和 TP 溶液浓度为输出建立 GWO-SVR 浓度预测模型。

GWO-SVR 定量分析模型对 TN 和 TP 测试集预测结果模型评价如表 3 所示。基于模型的实际值和预测值之间的相关度结果分析如图 4 所示。

表 3 GWO-SVR 模型评价结果

Table 3 GWO-SVR model evaluation results

	R^2	MSE	Elapsed time/s
TN	0.999 5	0.005 8	1.467
TP	0.998 5	0.037 6	1.646

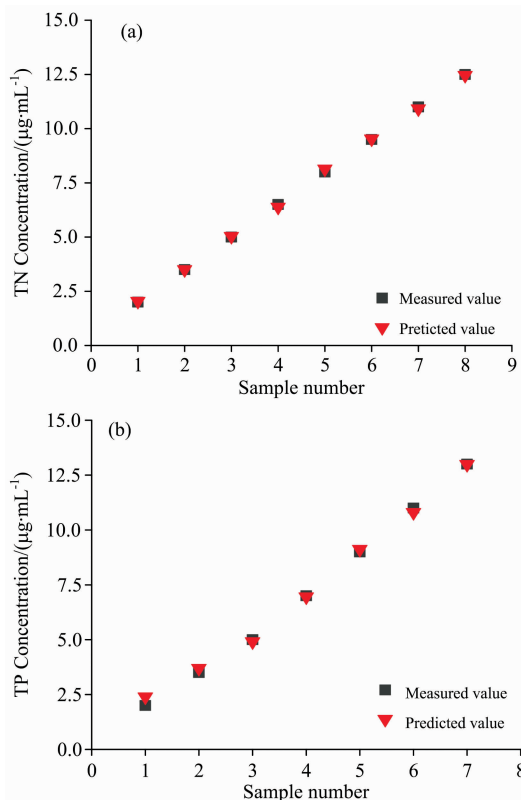


图 4 预测值和真实值相关性比较

(a): TN; (b): TP

Fig. 4 Correlation between predicted value and true value

(a): TN; (b): TP

由表 3 可以看出, GWO-SVR 定量模型得到的 TN 和 TP 溶液的相关度均高于 0.998, MSE 均低于 0.04, 具有较好的精确度。从图 4 可以看出, 模型的相关度极高, 预测值浓度和真实值浓度基本都落在同一点上, 而且建模时间均在 2 s 内, 具有良好的实时性。

将同组数据通过人工蜂群算法优化支持向量回归(ABC-SVR)和粒子群算法优化神经网络(PSO-BP)分别建模进行预测, 其中 ABC-SVR 模型初始蜂群规模 $M=20$, 迭代次数 $T=10$, 个体最大更新次数 $G=100$, 蜜源数量为 $M/2$; PSO-BP 模型初始参数设置为粒子群规模 $M=200$, 学习因子 C_1 和 C_2 均为 1, 隐藏层节点数为 13, 最大训练次数为 40 000。模型评价结果见表 4 和表 5。不同建模程序的相关度曲线见图 5。

表 4 不同模型预测结果对比(TN)

Table 4 Comparison of prediction results of different models (TN)

Model	R^2	MSE	Elapsed time/s
PSO-BP	0.998 3	0.022 7	13.94
ABC-SVR	0.993 8	0.052 7	2.29
GWO-SVR	0.999 5	0.005 8	1.47

表 5 不同模型预测结果对比(TP)

Table 5 Comparison of prediction results of different models(TP)

Model	R^2	MSE	Elapsed time/s
PSO-BP	0.996	0.042 6	19.67
ABC-SVR	0.979 2	0.299 2	1.85
GWO-SVR	0.998 5	0.037 6	1.65

从表 4 和表 5 评价结果可知, ABC-SVR 模型运行时间虽控制在 2 s 左右, 但其相关系数和均方误差评价结果均低于 GWO-SVR 模型和 PSO-BP 模型; 同样 PSO-BP 模型相关度均高于 0.996, MSE 均低于 0.05, 评价结果略低于 GWO-SVR 模型, 但运行时间和 GWO-SVR 模型相差数倍。结合图 5 分析来看, GWO-SVR 定量分析模型对 TN 和 TP 浓度的预测有着更好的相关性。综上所述, GWO-SVR 模型既能保证较好的预测精度又能将运行时间控制在 2 s 内, 具有良好实时性。而且相较于另外两种模型来说, 该模型有着更少的参数输入, 只需设置种群数量 M 和迭代次数 T 两个参数。GWO 对 SVR 进行优化, 可以在短时间内找到最优的全局最优参数组合 (C, g), 缩短了模型的运行时间, 提高了预测精度, 进而提升了模型的工作效率。由此可见 GWO-SVR 建立的定量分析模型结合拉曼光谱可以很好地实现对不同浓度梯度的 TN 和 TP 溶液含量的预测, 为以后水质监测提供了新方法。

4 结 论

提出了一种以纳米银溶胶作为基底的 SERS 技术结合 GWO-SVR 混合优化算法实现对不同浓度 TN 和 TP 溶液的快速定量检测方法。通过以测得溶液光谱强度作为输入值,

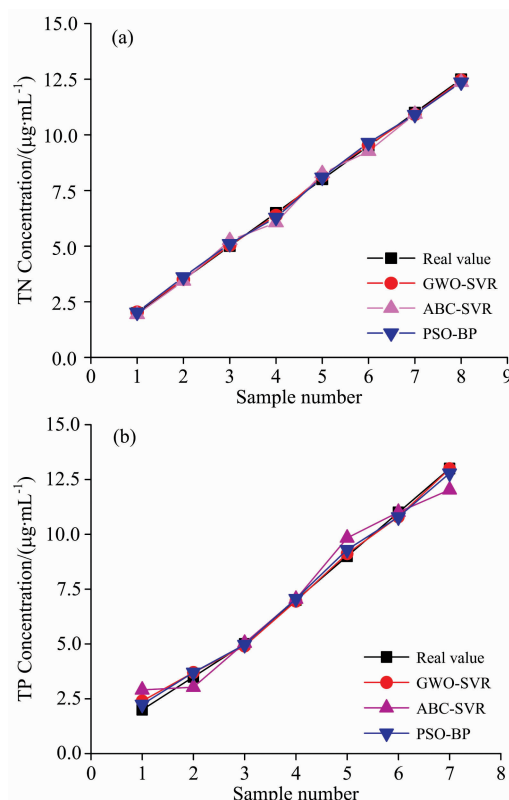


图 5 不同模型相关性曲线

(a): TN; (b): TP

Fig. 5 Different model correlation curves

(a): TN; (b): TP

溶液浓度作为输出建立模型。经过暗电流扣除、基线校正和光谱平滑三个步骤对原始数据进行预处理, 选取 GWO 优化算法对 SVR 算法的参数 C 和 g 进行寻优, 避免了局部最优和过拟合问题, 提升了收敛速度。为做对比将同组数据代入 ABC-SVR 和 PSO-BP 模型中, 以相关系数、均方误差和运行时间作为指标进行对比。预测结果表明, GWO-SVR 定量分析模型相关系数、均方误差以及建模时间均优于其他两种模型, 相关系数高达 0.998 5, 均方误差也控制在 0.04 以内, 说明该模型能较好的完成快速定量分析, 具备良好的实时性和更高的检测精度。因此, 此方法可以适用于水中 TN 和 TP 含量的快速定量检测, 为今后水质检测提供了理论依据。

References

- [1] ZHANG Yan, LI Chang, SHEN Hung Tao, et al(张岩, 李畅, SHEN Hung Tao, 等). Advances in Water Science(水科学进展), 2013, 24(5): 728.
- [2] Lin Kunning, Pei Junxian, Li Peicong, et al. Talanta, 2018, 185: 419.
- [3] Havva Tumay Temiz, Ismail Hakki Boyaci, Ivo Grabchev, et al. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2013, 116: 339.
- [4] Mathieu Lepot, Andres Torres, Thomas Hofer, et al. Water Research, 2016, 101: 519.
- [5] LI Xin-xing, ZHOU Jing, TANG Hong, et al(李鑫星, 周婧, 唐红, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2020, 40(1): 195.
- [6] Shintaro Pang, Tianxi Yang, Lili He. Trends in Analytical Chemistry, 2016, 85: 73.

- [7] Layla A Wali, Khulood K Hasan, Alwan M Alwan. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2019, 206: 31.
- [8] Bodelón Gustavo, Pastoriza-Santos Isabel. *Frontiers in Chemistry*, 2020, 8: 478.
- [9] Mathieu Wauters, Mario Vanhoucke. *Automation in Construction*, 2014, 47: 92.
- [10] WANG Xia, WANG Zhan-qi, JIN Gui, et al(王霞, 王占岐, 金贵, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2014, 30(4): 204.
- [11] Seyedali Mirjalili, Seyed Mohammad Mirjalili, Andrew Lewis. *Advances in Engineering Software*, 2014, 69: 46.
- [12] Cong Sunli, Sun Jun, Mao Hanping, et al. *Journal of the Science of Food and Agriculture*, 2018, 98(4): 1453.
- [13] CHEN Ying, ZHANG Can, XIAO Chun-yan, et al(陈颖, 张灿, 肖春艳, 等). *Acta Optica Sinica(光学学报)*, 2020, 40(10): 1030002.

Rapidly Detection of Total Nitrogen and Phosphorus Content in Water by Surface Enhanced Raman Spectroscopy and GWO-SVR Algorithm

ZHANG Yan-jun, KANG Cheng-long, LIU Ya-qian, FU Xing-hu*, ZHANG Jin-xiao, WANG Ming-xue, YANG Liu-zhen
School of Information Science and Engineering, The Key Laboratory for Special Fiber and Fiber Sensor of Hebei Province,
Yanshan University, Qinhuangdao 066004, China

Abstract A new quantitative analysis method was proposed, which combined surface-enhanced Raman spectroscopy (SERS) and support vector regression (SVR) based on Grey Wolf Optimization (GWO) algorithm to quickly and quantitatively detect the total nitrogen (TN) and total phosphorus (TP) content in water. The traditional TN and TP detection methods are complicated in process and time-consuming in the experimental environment. Therefore, rapid detection cannot be realized. However, SERS technology is easy to operate and time consuming, so combining it with the GWO-SVR algorithm can realize fast and accurate detection. With laboratory silver sol as the Raman enhanced substrate and TN, TP solutions with different concentration gradients as the research objects. TN and TP sample solutions were allocated to 26 and 23 groups respectively, in which 8 groups were selected as the test set for TN solution, 7 groups as the test set for TP solution, and the remaining sample solutions as the training set. The optimal experimental scheme was determined according to the different volume ratios of the tested solution and the silver sol. TN, TP were mixed with silver sol for 1 : 1, 1 : 2, 1 : 3, 2 : 1, 3 : 1, respectively. The results showed that the enhancement effect was the best when the solution and the silver sol were mixed at a ratio of 2 : 1. Spectral information was collected, and characteristic peaks were assigned. The original spectral data were preprocessed by dark current deduction, background deduction (baseline correction) and smoothing processing. The spectral analysis results show that the intensity of characteristic spectral peak varies with the concentration of solution due to the difference of functional group concentration in different concentrations of solution. The GWO-SVR quantitative analysis model was established by taking the spectral characteristic peak strength and solution concentration of the training set sample as the input and output values of the regression prediction model. The model's prediction ability was analyzed by correlation coefficient (r) and mean square error (MSE) of the sample solution of the test set, and the GWO-SVR model was compared with the other two models. The results showed that the GWO-SVR model predicted the TN solution with a correlation coefficient of 0.9995 and a mean square error of 0.0058, which were higher than the 0.9938, 0.0527 and 0.9983, 0.0227 of the artificial bee colony algorithm optimization support vector regression (ABC-SVR) and particle swarm optimization neural network (PSO-BP). The correlation coefficient of TP solution prediction was 0.9985, and the mean square error was 0.0376, which was also higher than the other two models. Moreover, compared with ABC-SVR and PSO-BP models, GWO-SVR has fewer input parameters, faster convergence speed, and easier to find the optimal global solution. Therefore, this method can realize the rapid and accurate detection of TN and TP content in water and provides a new method for water quality detection.

Keywords Surface-enhanced Raman spectrum; Gray Wolf optimization; Support vector regression; Total nitrogen; Total phosphorus