

基于太赫兹时域光谱技术的诺氟沙星浓度检测研究

白军朋^{1,2}, 李斌^{1*}, 张淑娟², 陈怡每¹

1. 北京农业信息技术研究中心, 北京 100097

2. 山西农业大学工学院, 山西 太谷 030801

摘要 畜禽养殖中抗生素的不合理使用导致畜禽产品中抗生素残留问题时有发生, 进而通过食物链影响食品安全, 威胁人类健康。准确、快速检测出抗生素药物的含量对保障食品安全具有重要意义。以残留较为常见的喹诺酮类诺氟沙星抗菌药为研究对象, 开展基于太赫兹光谱技术的诺氟沙星较大梯度和较小梯度的全浓度含量检测研究。在较大梯度诺氟沙星样本中, 首先, 在1%~100%范围内设置了11个浓度, 完成较大梯度压片样本制备; 然后经太赫兹时域光谱系统扫描获取其时域光谱, 提取样本的吸收系数, 用卷积平滑S-G二项式拟合滤波去除噪声、平滑样本光谱数据。发现纯净的诺氟沙星在1.205 THz处存在一个强烈吸收峰, 在0.816 THz处存在一个弱吸收峰。最后, 利用逐步回归和连续投影法(SPA)选择变量并结合特征吸收峰进行多元线性回归建模预测分析。进一步在较小梯度诺氟沙星样本研究中, 首先在浓度 $100 \mu\text{g} \cdot \text{mL}^{-1}$ (0.01%)以下设置了29个浓度系列, 完成较小梯度溶液样本制备; 然后获得其太赫兹时域光谱, 用S-G二项式拟合滤波进行数据预处理, 发现各浓度的吸收光谱未出现明显差异, 最后利用逐步回归和连续投影法(SPA)选择变量结合特征吸收峰实现多元线性回归建模预测分析。结果表明: 在较大梯度诺氟沙星样本中逐步回归选择变量的多元线性回归达到了模型最优($R_p=0.962$, $\text{RMSEP}=2.74\%$), 准确率优于当前已有最优预测模型($R_p=0.867$, $\text{RMSEP}=16.6\%$); 小梯度诺氟沙星样本的逐步回归选择变量的多元线性回归模型最优($R_p=0.728$, $\text{RMSEP}=18.79 \mu\text{g} \cdot \text{mL}^{-1}$), 该方法有一定的预测能力, 但是准确率有待提升。利用太赫兹光谱技术实现了较全浓度诺氟沙星的检测, 为后续诺氟沙星检测限的探索等进一步研究提供了一定的研究基础。

关键词 太赫兹时域光谱技术; 诺氟沙星; 浓度检测; 喹诺酮; 吸收系数; 抗生素残留

中图分类号: O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2021)09-2710-07

引言

喹诺酮类抗生素是一种重要抗生素, 常用于治疗或预防人和动物的细胞感染^[1], 随着抗生素的过度使用, 抗生素会残留在人或动物体内, 导致细菌的耐药性增加。残留在动物源性食品中的抗生素会通过食物链直接进入人的体内, 对人类的健康造成危害^[2]。食用产品中的抗生素残留也逐渐成为消费者关注的食品安全问题^[3]。因此大量的动物源食品需要做抗生素残留分析, 检验任务既繁重又时间紧迫。当前传统的检测方法主要有微生物学检测方法^[4]、免疫分析法^[5], 这些方法特异性强、灵敏度高, 但存在过程复杂、耗时长、有

破坏性等缺点。因此, 发展高效、灵敏的抗生素残留检测方法对于保障国家的食品安全和人体健康具有重大意义。

太赫兹指的是频率0.1~10 THz(1 THz= 10^{12} Hz)范围内的电磁辐射的统称^[6], 通常也被称为太赫兹辐射、T射线等, 对应的波长为0.3~3 mm、波数在 $3.3 \sim 330 \text{ cm}^{-1}$ 之间^[7]。由于其具有透射性、安全性、指纹性等独特的性质, 可识别大分子物质。朱思源^[8]等以四种常用青霉素为研究对象, 获得其太赫兹光谱吸收曲线。发现它们在一定的频率范围内都存在明显的光谱特性。吉特^[9]等采用太赫兹时域光谱技术对青霉素进行研究, 发现该技术可以对青霉素进行定量检测。Long等^[10]将诺氟沙星在0~100%分了12个梯度, 测量纯净的诺氟沙星样本发现在0.825和1.187 THz处存在

收稿日期: 2020-08-10, 修订日期: 2020-12-18

基金项目: 北京市农林科学院创新能力建设专项(KJXC20180421), 北京市自然科学基金项目(6182012), 国家自然科学基金项目(61601034)资助

作者简介: 白军朋, 1995年生, 山西农业大学工学院硕士研究生 e-mail: 1131458347@qq.com

* 通讯作者 e-mail: agrithz@163.com

两个吸收峰，并且逐步回归选择的变量结合特征峰的多元线性回归预测模型最优，预测相关系数 R_p 达到了 0.867。这表明太赫兹时域光谱技术可以识别抗生素这种大分子物质，在定量测量方面也表现出了一定的潜力，但是模型精度有待提高。中国国家标准(GB 31650—2019)规定的喹诺酮抗菌类药物在动物源性食品中的残留限为 $0.01 \sim 1.9 \mu\text{g} \cdot \text{mL}^{-1}$ 。现有有关诺氟沙星太赫兹光谱检测的文献中涉及到浓度检测的研究不足以满足实际检测需求，不同浓度梯度间隔下的太赫兹光谱检测分辨能力有待系统性的开展研究。

为此，本研究探讨了较全浓度的诺氟沙星含量预测建模方法，分别设计了较大梯度诺氟沙星(浓度系列和梯度间隔均大于 $10^4 \mu\text{g} \cdot \text{mL}^{-1}$ ，即 1%)和较小梯度诺氟沙星(浓度系列和梯度间隔都小于 $100 \mu\text{g} \cdot \text{mL}^{-1}$ ，即 0.01%)并展开了探

索性检测研究。主要的研究内容如图 1 所示：首先对浓度呈较大梯度间隔的诺氟沙星固体压片进行检测研究，采集太赫兹光谱数据，采用多种建模方法预测诺氟沙星含量，尝试在检测准确率方面得到提升；其次对呈较小梯度间隔的诺氟沙星溶液样本含量检测的可行性进行探索，尝试用不同建模方法对诺氟沙星含量进行定量预测，寻找较优的定量检测方法。需要说明的是：在制备呈较小梯度浓度间隔样本时，由于诺氟沙星固体颗粒在浓度小于 5% (即 $5 \times 10^4 \mu\text{g} \cdot \text{mL}^{-1}$) 很难检测，无法满足较小梯度样本的制备要求，并且其在水中的溶解度也有限，本研究采用乙醇溶剂的诺氟沙星溶液残留标准物质($100 \mu\text{g} \cdot \text{mL}^{-1}$)为母液进行较小梯度诺氟沙星制备研究，为后续检测限的探索提供研究基础。

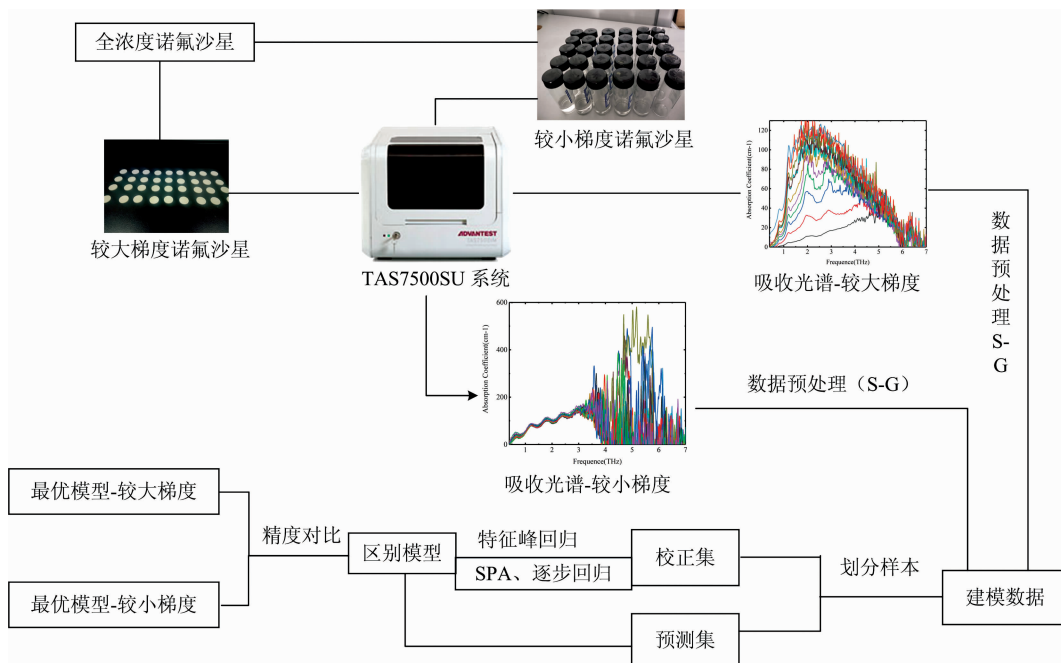


图 1 实验流程图

Fig. 1 Flowchart of the experimental process

1 实验部分

1.1 试剂和材料

实验所用的诺氟沙星(CAS 70458-96-7)从生物工程(上海)股份有限公司购买。聚乙烯粉末(CAS 9002-88-4, 粒度 $40 \sim 48 \mu\text{m}$)购买于西格玛奥德里奇(上海)贸易有限公司。用于较小梯度诺氟沙星样本制备的标准溶液购买于农业农村部环境保护科研检测所(中国天津), 是乙醇中诺氟沙星溶液残留标准物质(GSB05-3338-2016), 浓度为 $100 \mu\text{g} \cdot \text{mL}^{-1}$, 规格是 $1 \text{ mL} \cdot \text{支}^{-1}$ 。比色皿样品池光程规格为 0.2 mm 。

1.2 样本制备

1.2.1 较大梯度样本的制备

称取部分诺氟沙星和聚乙烯粉末在研钵中充分研磨之后用压片机及其模具(Specac GS15011, 英国)压片, 压力大小

是 3.5 t , 压片时间为 $3 \text{ min} \cdot \text{片}^{-1}$, 压片直径约 13 mm , 厚度在 $1 \sim 2 \text{ mm}$ 之间, 样本总质量 200 mg 。分别制备浓度 5%, 15%, 25%, 35%, 45%, 55%, 65%, 75%, 85%, 95% 和 100% 的压片样本, 为减少操作误差, 每个含量的混合物样品制备 9 个压片, 共 99 个较大梯度样本。按 2:1 划分成校正集和预测集。

1.2.2 较小梯度样本的制备

采用乙醇中诺氟沙星溶液残留标准物质(GSB05-3338-2016)为母液, 稀释得到 29 个小梯度浓度(0.001, 0.002, 0.004, 0.006, 0.008, 0.01, 0.02, 0.04, 0.06, 0.08, 0.1, 0.2, 0.4, 0.6, 0.8, 1, 2, 4, 6, 8, 10, 20, 30, 40, 50, 60, 70, 80 和 $90 \mu\text{g} \cdot \text{mL}^{-1}$)。用移液枪移取少量的诺氟沙星乙醇溶液滴加到比色皿样品池中即完成样本制备, 依次完成所有样本制作。每个浓度制作 9 个样本, 共 261 个较小梯度样本, 按 2:1 划分成校正集和预测集。

1.3 太赫兹光谱测量

使用日本 ADVANTEST 公司生产的太赫兹时域光谱系统 TAS7500SU, 使用透射工作模式。中心波长为 1 550 nm, 重复频率为 50 MHz, 脉冲宽度为 50 fs, 光纤飞秒激光器的输出功率为 20 mW 的。可测量频谱范围为 0.1~7 THz, 频谱分辨率为 7.6 GHz, 系统动态信号范围在 60 dB 以上。实验环境温度为 25 °C (± 1 °C), 相对湿度低于 3%。

将制备好的较大梯度压片样本和不同浓度的待测试较小梯度溶液样本放入透射模式的样品腔中, 依次完成所有样本的太赫兹光谱测量。每次测量前测量一次参考信号, 每个样本测量一次光谱数据, 较大梯度的参考信号为空扫信号, 较小梯度的参考信号是空扫的比色皿样品池信号。

1.4 参数提取与数据建模

时域信号在 30 ps 被截断置零, 用以去除从衬底背面反射的信号。对时域数据进行快速傅里叶变换(FFT)得到样本的频域光谱。根据 Dorney^[11] 和 Duvillaret^[12] 等提出的方法计算各样本的吸收系数。

卷积平滑(S-G)作为一种数字滤波器, 可以平滑数据并增加数据信噪比(SNR)。对光谱数据进行 S-G 平滑, 以减少噪声干扰和粒子散射。

逐步回归通过一步步引入变量, 筛选剔除多余信息, 得到有效变量; 连续投影算法(SPA)^[13] 作为一种可以使矢量空间共线性最小化的前向变量选择算法, 具有消除原始光谱矩阵中的冗余信息的优势, 可以筛选出有效的信息, 提取全频段的几个特征频率。多元线性回归进行定量建模预测分析。实验数据处理与绘图全部在 Matlab R2018a 和 Origin 2017 中完成。

1.5 模型评价

利用相关系数 r 、校准均方根误差(RMSEC)和预测均方根误差(RMSEP)来评价模型的性能。其中相关系数 r 包含 R_c (校正相关系数)和 R_p (预测相关系数)。预测诺氟沙星浓度模型中最理想的模型应具有高的校准相关系数(R_c)和预测相关系数(R_p), 低的校准均方根误差(RMSEC)和预测均方根误差(RMSEP)。为保证模型的稳定性, 还应该保证 RMSEC 和 RMSEP 有较小的差异。

2 结果与讨论

2.1 各样品的太赫兹光谱吸收曲线及分析

2.1.1 较大梯度诺氟沙星测量结果

如图 2(a)所示是较大梯度诺氟沙星中不同浓度的样本的吸收系数平均光谱, 因为低于 0.4 THz 和高于 1.5 THz 的信号信噪比(SNR)很低, 所以只选 0.4~1.5 THz 的数据进行后续建模分析。本文发现纯净的诺氟沙星样本在 0.816 和 1.205 THz 处有两个吸收峰, 与 Long 等^[10] 检测的 0.825 和 1.187 THz 及 Redo-Sanchez 等^[14] 检测的 0.79 和 1.19 THz 相近。一方面, 峰位轻微的移动可能是由样品的颗粒大小引起的; 另一方面, 也可能因为诺氟沙星的种类差异和测量仪器不同造成了这种峰位的微小移动。通过观察吸收系数图, 可以发现随着诺氟沙星浓度的增加, 吸收系数增加, 并且峰值也变的更加明显。这说明诺氟沙星的吸收系数变化和浓度有直接关系。但是在诺氟沙星浓度低于 5% 的时候, 很难发现吸收峰的存在, 原因可能是聚乙烯中诺氟沙星浓度较低的时候, 样本对太赫兹光谱的吸收变少, 从而接收到的太赫兹时域光谱携带的信息量少所致。

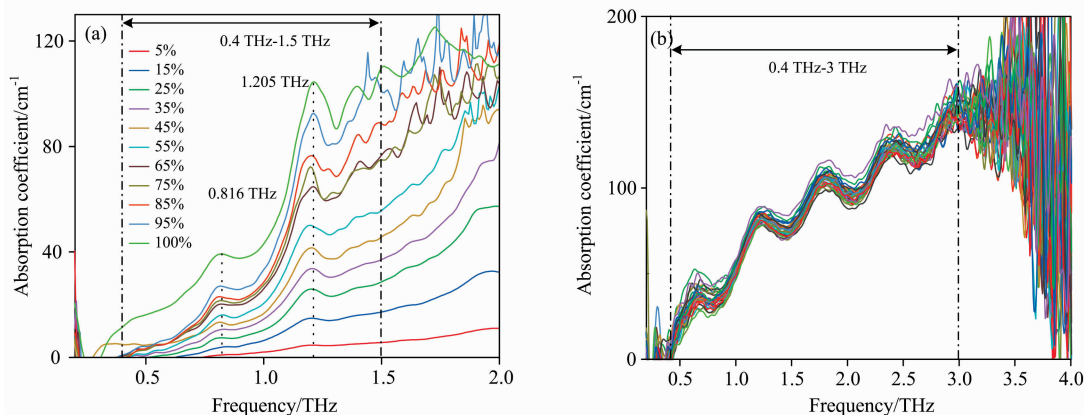


图 2 太赫兹吸收系数平均光谱

(a): 较大梯度诺氟沙星; (b): 较小梯度诺氟沙星

Fig. 2 Average spectrum of terahertz absorption coefficient

(a): Larger gradient Norfloxacin; (b): Smaller gradient Norfloxacin

2.1.2 较小梯度诺氟沙星测量结果

较小浓度诺氟沙星溶液样本的吸收系数如图 2(b)所示, 不同浓度的诺氟沙星乙醇溶液的吸收系数曲线基本重合, 未表现出明显差异。可能是因为当诺氟沙星溶液的浓度比较低时, 对太赫兹光谱影响不大, 所以表现为小梯度诺氟沙星溶

液的吸收系数没有直观上的差异。尝试用不同的化学计量学方法寻找与浓度有关的特征变量, 本文选取信噪比较高的 0.4~3 THz 范围内的光谱数据进行后续建模分析。

2.2 光谱预处理

为了提高吸收光谱的信噪比, 本文使用 S-G 滤波器对原

始光谱进行预处理,对相邻 11 个数据点进行二项式拟合。处理后的光谱对比如图 3 所示,可以看出预处理后的数据更加平滑和规范化。

2.3 较大梯度诺氟沙星含量差异样本定量建模研究

本文对较大梯度的诺氟沙星样本进行太赫兹光谱测量,获得其太赫兹吸收谱,原始光谱如图 3(a)所示,S-G 滤波后的图 3(b)所示。预处理后的数据更加平滑和规范化。利用诺氟沙星的两个特征吸收峰(0.816 和 1.205 THz)进行多元线性回归建模与预测,得到预测相关系数 R_p 和预测均方根误差 RMSEP 分别为 0.941 和 10.84%。这时发现预测相关系数较高,达到了 0.9 以上,但是 RMSEP 比较高。为了进一步提高较大梯度诺氟沙星的预测效果,本研究尝试对整个波长范围的数据采用逐步回归方法选取特征变量,这里选取 11 个特征变量(0.412, 0.443, 0.557, 0.633, 0.786, 1.106,

1.221, 1.251, 1.411, 1.450 和 1.500 THz),这里的特征变量消除了部分多重共线性,使得到的光谱数据包含更多的信息。然后进行多元线性回归分析,预测相关系数 R_c 和预测均方根误差 RMSEP 分别为 0.962 和 2.74%。显然,用逐步回归选取变量的多元线性回归建立的模型比用特征峰多元线性回归建立的模型精度更高, R_c 得到了提高, RMSEP 降低。这说明逐步线性回归方法可以从全波段中选取更多有效的信息。逐步回归是从整个波长中搜索变量,其中容易出现光谱数据过拟合,为了减轻数据过拟合,选取变量时严格的对每个波长进行逐一添加或删除。为了提高较大梯度诺氟沙星模型的预测精度,进一步用特征吸收峰结合逐步线性回归选择变量的多元线性回归建模预测较大梯度诺氟沙星的浓度,这时 $R_c=0.989$, RMSEP=5.72%,虽然预测相关系数提高了,同时预测均方根误差也增加了。

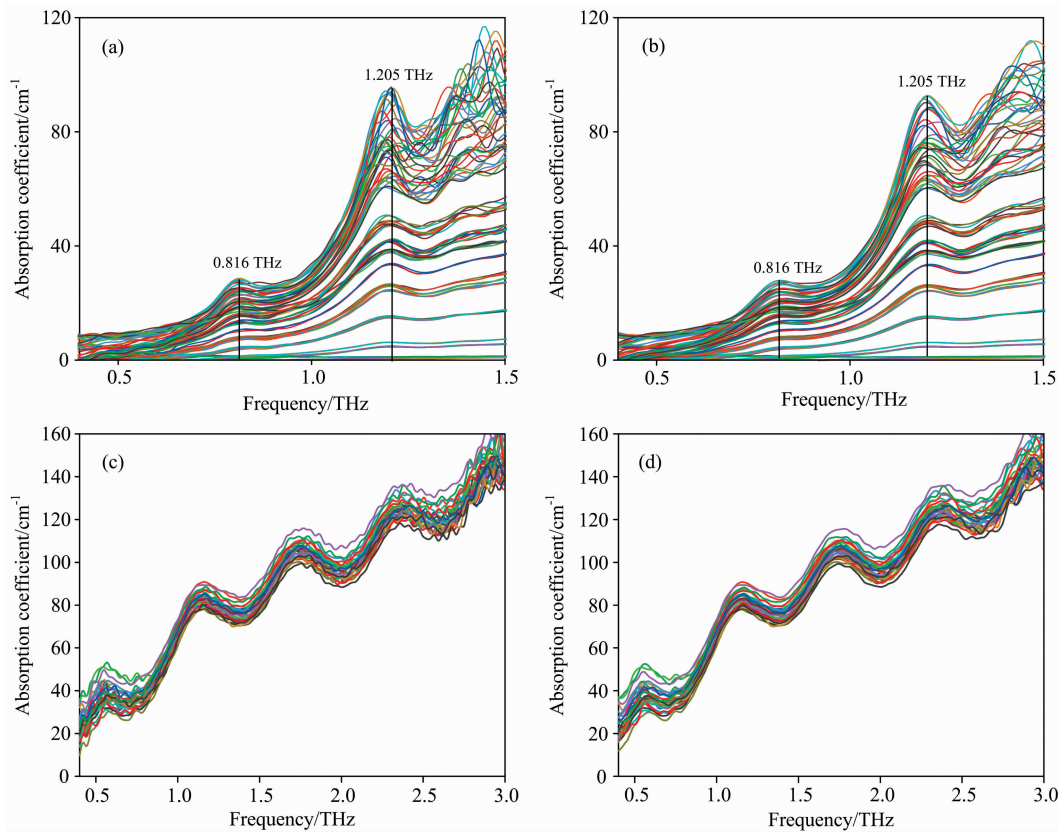


图 3 太赫兹吸收系数光谱

(a): 较大梯度诺氟沙星(原始数据); (b): 较大梯度诺氟沙星(S-G 滤波后);
(c): 较小梯度诺氟沙星(原始数据); (d): 较小梯度诺氟沙星(S-G 滤波后)

Fig. 3 Terahertz absorption coefficient spectrum

(a): Larger gradient Norfloxacin (raw data); (b): Larger gradient Norfloxacin (after S-G filtering);
(c): Smaller gradient Norfloxacin (raw data); (d): Smaller gradient Norfloxacin (after S-G filtering)

利用 SPA 选择的变量对较大梯度诺氟沙星进行建模与预测分析。选取变量过程中,当运行 SPA 程序时, SPA_MLR 程序会计算一系列的误差均方根误差(RMSE),这个过程保证选取的变量个数伴随着一个理想的 RMSE,并且这个 RMSE 不会比最小的 RMSE 明显大。图 4(a)显示了通过 SPA 选择变量的过程,从图中可以看出,当 SPA 选择前 5 个

变量时, RMSE 值显著的降低。随着选择变量个数增加, RMSE 值逐渐减小。当选择 14 个变量时, RMSE 达到了理想值(RMSE=1.66%)。选择变量过程中,尽管 RMSE 曲线存在一些波动,但是整体呈现随着变量的增加 RMSE 曲线下降的趋势。在较大梯度诺氟沙星样本中, SPA 选择的 14 个变量如图 4(b)所示, SPA 选取的变量分别用小正方形标记

(0.764, 1.000, 1.053, 1.098, 1.144, 1.311, 1.326, 1.349, 1.379, 1.402, 1.425, 1.455, 1.478 和 1.493 THz)。我们发现在 1.205 THz 以下有 5 个变量, 1.205 THz 以上存在 9 个变量, 这说明有效信息的频率大多分布在较大的频率

范围。SPA 选择的变量的多元线性回归结果为 $R_p = 0.992$, $RMSEP = 5.52\%$, 虽然模型精度提高了, 但是预测均方根误差也变大了。

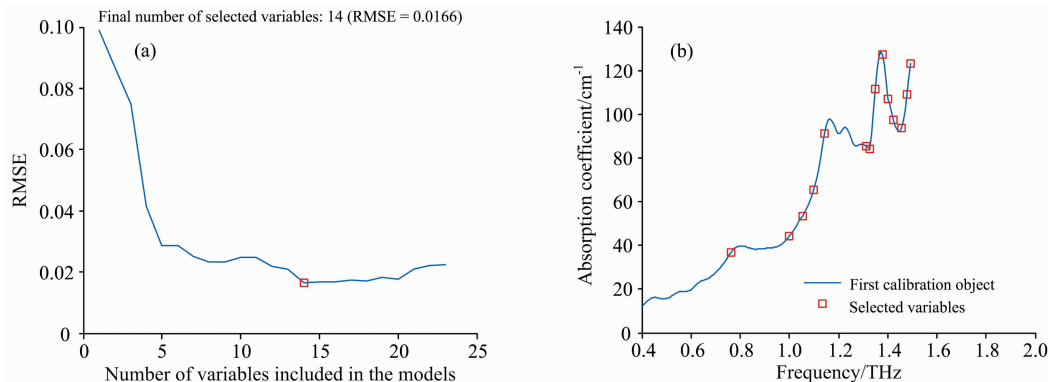


图 4 SPA 选择变量(较大梯度诺氟沙星)

(a): 选取变量数的 RMSE; (b): 选择的变量位置

Fig. 4 Variable selection SPA (Larger gradient Norfloxacin)

(a): RMSE graph of the number of variables selected; (b) Location of selected variables

较大梯度诺氟沙星的建模结果见表 1, 逐步回归选择变量后的多元线性回归达到模型效果最优, 预测相关系数 $R_p = 0.962$, $RMSEP = 2.74\%$, 该模型在保证预测相关系数的前提下, 同时保证 $RMSEC$ 和 $RMSEP$ 的差值最小, 使模型具有更高的稳定性。这说明太赫兹时域光谱技术可以检测较大梯度诺氟沙星, 预测精度相比已有检测的建模结果 ($R_p = 0.867$, $RMSEP = 16.6\%$) 得到了很大改善。

表 1 较大梯度诺氟沙星的 MLR 建模结果

Table 1 MLR modeling results of larger gradient Norfloxacin

变量选择方法	变量个数	R_c	$RMSEC / \%$	R_p	$RMSEP / \%$
特征峰	2	0.995	3.44	0.941	10.84
逐步回归	11	0.996	1.62	0.962	2.74
特征峰+逐步回归	13	0.999	1.76	0.989	5.72
SPA	14	0.999	1.58	0.992	5.52

2.4 较小梯度诺氟沙星含量差异样本定量建模研究

为了进一步研究较小梯度间隔诺氟沙星含量的定量检测问题, 本文制备了浓度系列和梯度均小于 $100 \mu\text{g} \cdot \text{mL}^{-1}$ 的 29 个较小梯度浓度, 对较小梯度的诺氟沙星样本进行了测量, 得到原始光谱如图 3(c) 所示。经 S-G 滤波后如图 3(d) 所示, 预处理后的数据更加平滑和规范化, 后续对其进行建模分析。首先, 用纯净的诺氟沙星样本的特征峰 (0.816 和 1.205 THz) 进行多元线性回归建模预测小梯度诺氟沙星无水乙醇溶液的浓度, 发现得到 R_c 和 R_p 分别是 0.590 和 0.307, 建模效果很不理想, 可能是由于较小梯度诺氟沙星溶液中含诺氟沙星少, 所以光谱所包含的特征峰信息也比较少, 从而导致特征峰建模效果不理想。之后用 SPA 算法选取

变量进行建模分析, SPA 选择了两个有效变量 (0.496 和 1.77 THz), 利用这两个变量进行多元线性回归建模预测较小梯度诺氟沙星溶液的浓度, 我们得到 R_c 和 R_p 分别是 0.485 和 0.472。发现预测相关系数 R_p 只有 0.472, 模型效果仍然不理想, 这可能是因为 SPA 算法不适合用于较小梯度诺氟沙星溶液来提取变量, 也即 SPA 找到的变量含有有限的浓度信息。最后利用逐步线性回归选取变量进行多元线性回归建模, 逐步回归找到了 10 个特征变量 (0.565, 0.610, 0.641, 1.335, 1.762, 1.793, 1.862, 1.892, 2.152 和 2.228 THz), 建立模型的 R_c 和 R_p 分别是 0.859 和 0.728。较小梯度诺氟沙星的建模结果如表 2 所示。可以看出相比特征峰多元线性回归和 SPA 的多元线性回归, 逐步线性回归的多元线性回归模型效果达到了最优, 但是预测相关系数为 0.728, 预测均方根误差为 $18.79 \mu\text{g} \cdot \text{mL}^{-1}$ 。这说明太赫兹时域光谱技术在检测较小梯度诺氟沙星溶液方面能力较弱, 需要继续寻求提升的手段。

表 2 较小梯度诺氟沙星的 MLR 建模结果

Table 2 MLR modeling results of smaller gradient Norfloxacin

变量选择方法	变量个数	R_c	$RMSEC / (\mu\text{g} \cdot \text{mL}^{-1})$	R_p	$RMSEP / (\mu\text{g} \cdot \text{mL}^{-1})$
特征峰	2	0.590	34.5	0.307	10.85
SPA	2	0.485	23.50	0.472	24.07
逐步回归	10	0.859	15.18	0.728	18.79

总体来说, 太赫兹时域光谱技术作为一种新型的无损检测技术, 因其具有独特的指纹特性, 且样品预处理简单等优势, 近年来被广泛应用于抗生素的定性定量检测研究探索。存在的问题主要有:

(1) 对于较高浓度检测,一般具有较好的精度;(2) 对于较低浓度检测,检测精度较差,特别是当检测浓度低于中国国家标准(GB 31650—2019)时,检测变得异常困难。近些年来,国内外相关研究人员也在不断探索新型探测方法,如超材料^[15]、平行平板金属波导^[16]等用于痕量检测研究。这也成为本研究的未来方向之一。

3 结 论

分别以较大梯度(浓度系列和浓度间隔均大于 $10^4 \mu\text{g} \cdot \text{mL}^{-1}$, 即 1%)诺氟沙星和等于或者低于国家规定检测限的较小梯度(浓度系列和浓度间隔均小于 $100 \mu\text{g} \cdot \text{mL}^{-1}$, 即 0.01%)诺氟沙星为研究对象,通过设计梯度实验,探索太赫兹时域光谱技术用于较全浓度诺氟沙星的检测潜力。较大梯度诺氟沙星样本的检测分析发现:纯净的诺氟沙星样本存

在两个吸收峰,峰位与已有的研究结果基本一致。逐步线性回归选择变量进行多元线性回归分析,预测集的相关系数 $R_c=0.962$, 预测均方根误差 $\text{RMSEP}=2.74\%$ 。检测结果相比已有检测的最优模型($R_c=0.886$, $\text{RMSEP}=16.6\%$)得到了提升;进一步的较小梯度诺氟沙星溶液样本检测分析发现:逐步线性回归选择变量的多元线性回归达到了模型效果最优,预测集相关系数 $R_c=0.728$, $\text{RMSEP}=18.79 \mu\text{g} \cdot \text{mL}^{-1}$, 该模型相比大梯度诺氟沙星的预测能力明显下降。说明太赫兹时域光谱技术在预测小梯度诺氟沙星含量方面能力不足,需要继续寻找提升手段。总体来说,太赫兹时域光谱技术可以比较准确的预测较大梯度的诺氟沙星,并在较小梯度诺氟沙星的预测方面也表现出一定的潜力,但检测能力较弱。本文为诺氟沙星浓度的进一步检测研究提供了理论基础。

References

- [1] Li S W, Wang D, Liu H B, et al. *Journal of Applied Ichthyology*, 2013, 29(6): 1204.
- [2] LI Bin, ZHAO Xu-ting, ZHANG Yong-zhen, et al(李斌, 赵旭婷, 张永珍, 等). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2019, 39(12): 3659.
- [3] Chafer-Pericas C, Maquieira A, Puchades R. *TrAC-Trends in Analytical Chemistry*, 2010, 29(9): 1038.
- [4] FAN Wei, GAO Xiao-yue, LI He-nan(范维, 高晓月, 李贺楠). *Food Science(食品科学)*, 2017, 38(16): 239.
- [5] Chen Y Q, Chen Q, Han M M, et al. *Biosens Bioelectron.*, 2016, 79(4): 430.
- [6] XIE Li-juan, XU Wen-dao, YING Yi-bin, et al(谢丽娟, 徐文道, 应义斌, 等). *Transactions of the Chinese Society for Agricultural Machinery(农业机械学报)*, 2013, 44(7): 246.
- [7] LI Bin, LONG Yuan, LIU Huan, et al(李斌, 龙园, 刘欢, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2018, 34(2): 1.
- [8] ZHU Si-yuan, ZHANG Man, SHEN Jing-ling(朱思源, 张曼, 沈京玲). *Infrared and Laser Engineering(红外与激光工程)*, 2013, 42(3): 626.
- [9] JI Te, ZHAO Hong-wei, ZHANG Zeng-yan, et al(吉特, 赵红卫, 张增艳, 等). *Acta Physico-Chimica Sinica(物理化学学报)*, 2006, 22(9): 1159.
- [10] Long Y, Li B, Liu H. *Applied Optics*, 2018, 57(3): 544.
- [11] Dorney T D, Baraniuk R G, Mittleman D M. *Journal of the Optical Society of America A-Optics Image Science and Vision*, 2001, 18(7): 1562.
- [12] Duvillaret L, Garet F, Coutaz J L. *Applied Optics*, 1999, 38(2): 409.
- [13] Li J, Zhao C, Huang W, et al. *Anal. Methods*, 2014, 6(7): 2170.
- [14] Redo-Sanchez A, Salvatell G, Galceran G, et al. *Analyst*, 2011, 136, 1733.
- [15] Qin J Y, Xie L J, Ying Y B. *Food Chemistry*, 2016, 211: 300.
- [16] Zhang J Q, Grischkowsky D. *Optics Letters*, 2004, 29(14): 1617.

Study on Norfloxacin Concentration Detection Based on Terahertz Time Domain Spectroscopy

BAI Jun-peng^{1, 2}, LI Bin^{1*}, ZHANG Shu-juan², CHEN Yi-mei¹

1. Beijing Research Center for Information Technology in Agriculture, Beijing 100097, China

2. College of Engineering, Shanxi Agricultural University, Taigu 030801, China

Abstract The unreasonable use of antibiotics in poultry breeding has led to the frequent occurrence of antibiotic residues in poultry products, affecting food safety and posing a threat to human health. Therefore, accurate and rapid detection of antibiotic drug content is of great significance. This study successfully carried out larger and smaller gradient whole content detection of norfloxacin. This study successfully carried out larger out based on terahertz spectroscopy. First, in the samples of higher-gradient norfloxacin, 11 concentrations were set in the range of 1% ~ 100% to complete the preparation of high-gradient samples. Then the time spectrum was scanned by the terahertz time spectrum system, the absorption coefficient of the sample was extracted, and the Savitzk-Golay(S-G) binomial fitting filter was used to remove the noise and smooth the spectral data of the sample. It was found that the absorption coefficient of pure norfloxacin had a strong absorption peak at 1.205 THz and a weak absorption peak at 0.816 THz. Finally, stepwise regression and successive projections algorithm (SPA) is used to select variables and combine the characteristic absorption peak to realize the prediction analysis of multiple linear regression modeling. Further, in the study of smaller gradient norfloxacin samples, we set 29 concentration series below $100 \mu\text{g} \cdot \text{mL}^{-1}$ (0.01%) to complete the preparation of smaller gradient samples. Then the terahertz time domain spectrum was obtained, the S-G binomial fitting filter was used for data preprocessing, and no significant difference was found in the absorption spectrum of each concentration. Finally, stepwise regression and SPA were used to select variables and combine the characteristic absorption peak to realize multiple linear regression modeling and prediction analysis. The results showed that the multiple linear regression of stepwise regression selection variable in the higher-gradient norfloxacin sample reached the optimal model ($R_p=0.962$, $\text{RMSEP}=2.74\%$), and the accuracy was better than the existing optimal prediction model ($R_p=0.867$, $\text{RMSEP}=16.6\%$). The multiple linear regression model of the stepwise regression selection variable of the small gradient norfloxacin sample is optimal ($R_p=0.728$, $\text{RMSEP}=18.79 \mu\text{g} \cdot \text{mL}^{-1}$). This method has a certain prediction ability, but its accuracy needs to be improved. In this study, terahertz spectroscopy was used to detect norfloxacin in full concentration, which provided a certain research basis for exploring norfloxacin detection limit and other further studies.

Keywords Terahertz time domain spectroscopy; Norfloxacin; Concentration detection; Quinolone; Absorption coefficient; Antibiotic residue

(Received Aug. 10, 2020; accepted Dec. 18, 2020)

* Corresponding author