

文章编号: 1672-8785(2024)10-0045-12

基于太赫兹精密波谱的麻精药物 艾司唑仑定性定量分析测定

张佳敏¹ 王俊涛¹ 郝娅婷¹ 王宁一¹ 黄星皓¹吴 旭¹ 曾令高^{2*} 彭 滢^{1*}

(1. 上海理工大学太赫兹技术创新研究院, 太赫兹频谱与成像技术协同创新中心, 上海市现代光学系统重点实验室, 上海理工大学光电信息与计算机工程学院, 上海 200093;

2. 重庆市食品药品检验检测研究院, 国家药品监督管理局麻醉精神

药品质量监测重点实验室, 重庆 401121)

摘 要: 开发了一种基于太赫兹波谱技术的艾司唑仑快速检测方法, 以期解决高效液相色谱法(High Performance Liquid Chromatography, HPLC)在检测过程中的前处理繁琐、样品损耗、耗时长和成本高等问题。首先, 采用密度泛函理论(Density Functional Theory, DFT)方法, 预测了艾司唑仑在 0.9~20.0 THz 频段内的分子振动模式, 建立了定性分析方法。其次, 通过分析艾司唑仑含量与太赫兹特征峰的相关性, 构建了高拟合度的定量分析模型。然后对 22 例不同厂家生产的艾司唑仑药品进行了太赫兹波谱解析, 并结合支持向量回归(Support Vector Regression, SVR)算法, 实现了与 HPLC 方法相比平均绝对误差(Mean Absolute Error, MAE)仅为 1.12% 的准确定量。该方法具有操作简便、分析速度快、成本低廉和无损检测等优点, 为特殊药品的质量控制提供了一种新的分析策略, 具有推广至其他麻精药物质量评价研究的潜力。

关键词: 波谱学; 药物定性定量分析; 太赫兹波谱技术; 定量工作曲线; 支持向量回归

中图分类号: O433.1 文献标志码: A DOI: 10.3969/j.issn.1672-8785.2024.10.007

Qualitative and Quantitative Analytical Determination of Estazolam Based on Terahertz Precision Spectroscopy

ZHANG Jia-min¹, WANG Jun-tao¹, HAO Ya-ting¹, WANG Ning-yi¹,HUANG Xing-hao¹, WU Xu¹, ZENG Ling-gao^{2*}, PENG Yan^{1*}

收稿日期: 2024-04-22

基金项目: 国家自然科学基金项目(61988102; 62335012); 国家重点研发计划项目(2022YFA1404004)

作者简介: 张佳敏(2002-), 女, 安徽黄山人, 本科生, 主要研究方向为太赫兹波谱技术在生物医学方面的应用。

*通讯作者: E-mail: zenglinggao@cqifdc.org.cn(曾令高); py@usst.edu.cn(彭滢)

- (1. Terahertz Technology Innovation Research Institute, Terahertz Spectrum and Imaging Technology Cooperative Innovation Center, Shanghai Key Lab of Modern Optical System, School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China;
2. NMPA Key Laboratory of Quality Monitoring of Anesthetic and Psychotropic Substances, Chongqing Institute of Food and Drug Control, Chongqing 401121, China)

Abstract: In this paper, a rapid detection method for estazolam based on terahertz spectroscopy technology is developed to solve the problems of cumbersome pre-processing, sample loss, long time and high cost in the detection process by HPLC. Firstly, DFT is used to predict the molecular vibration mode of estazolam in 0.9–20.0 THz frequency band, and a qualitative analysis method is established. Secondly, by analyzing the correlation between the content of estazolam and the characteristic peak of terahertz, a quantitative analysis model with a high fitting degree is established. Then, the terahertz spectrum analysis of 22 estazolam products from different manufacturers is performed, and the SVR algorithm is combined to achieve an accurate quantification with MAE of 1.12% compared with the HPLC method. The method has the advantages of simple operation, fast analysis speed, low cost and non-destructive testing, which provides a new analytical strategy for the quality control of special drugs, and has the potential to be extended to the quality evaluation of other anesthetic drugs.

Key words: spectroscopy; qualitative and quantitative analysis of medicine; terahertz wave spectroscopy technique; quantitative working curve; support vector regression

0 引言

艾司唑仑是一种被广泛使用的苯二氮䓬类抗焦虑药，其主要作用机制为促进 γ -氨基丁酸分泌^[1-2]。它具有较强的镇静、安眠、抗惊厥、抗焦虑效果以及较弱的中枢性肌松作用和抗胆碱作用，在临床上主要用于治疗焦虑和失眠症状^[3-8]。该药品自 2023 年被列入国家药品抽检目录，并按照第二类精神药品管控。目前，艾司唑仑的质量评价主要依靠 HPLC 法^[9]。而该方法在每次检测时需要标准品进行标定，且存在前处理繁琐、样品损耗、耗时长、成本高等问题。因此，需要开发一种能够快速、准确、低损耗地对艾司唑仑进行精密分析的定性定量分析方法。

太赫兹波通常定义为波长在 0.03~3 mm 范围内的电磁波。由于具有不同于其他波段的特性，太赫兹波在生物医学检测领域拥有显著优势^[10-11]。首先，太赫兹波的光子能量在毫电子伏特量级，较弱的光子能量不会破坏物质本身的属性，可实现无损检测；其次，太赫兹脉冲的典型脉宽在皮秒量级，能够捕捉物质在极短时间尺度上的动态变化^[12]；此外，许多生

物分子的转动和振动能级位于太赫兹频段^[13]，意味着太赫兹波谱能反映分子种类及其结构的微小差异，赋予其“分子指纹”的独特属性^[14]。

基于以上特性，太赫兹波谱技术在许多研究领域都有潜在的应用，包括物理学^[15-16]、材料学^[15]和生物医学^[17-21]。在药物检测方面，太赫兹波谱技术已经实现了对多种药物（如格列喹酮、格列吡嗪、华法林钠等西药和黄郁金、绿郁金、三七、西洋参等中草药^[22-30]）的检测。2003 年，Taday 等最早报道了通过太赫兹脉冲光谱区分盐酸雷尼替丁的两种晶型；2019 年，Wu Z 等利用太赫兹光谱技术在 0.5~4.5 THz 对吡咯烷酸的异构体和异构体混合物进行分子晶体结构的研究^[31]；2021 年，Yang Q 等使用太赫兹技术，结合化学计量学方法对复方药物中的主要成分咖啡因进行定量分析，得到药物主成分含量^[32]；2024 年，顾虹宇等人基于太赫兹光谱和卷积神经网络模型，实现了对三七原产地的快速、非破坏性鉴别，准确率达到 98.02%^[33]。

在本研究中，我们基于太赫兹波谱技术对艾司唑仑这一麻精药物进行了深入的定性定量

量分析。主要研究工作如下: 首先, 对艾司唑仑的分子模型进行理论计算, 预测其在 0.9~20.0 THz 波段内的特征吸收峰; 其次, 通过实验测试确定艾司唑仑在太赫兹波段的特征峰, 并探讨艾司唑仑含量与特征峰的吸光度及峰面积之间的关系, 以建立定量分析模型; 最后, 结合 SVR 算法, 实现了与 HPLC 方法相比 MAE 值仅为 1.12% 的准确预测。

1 材料和方法

1.1 实验药品

在本研究中, 所用的艾司唑仑纯品及实际药物样品均由重庆市食品药品检验检测研究院提供。共收集了 22 份艾司唑仑药物样品, 涵盖了 11 家不同生产厂家的两个生产批次。所有样品的主要成分均为艾司唑仑, 而具体的配方和制作工艺则属于药厂的保密信息。鉴于直接对样品进行太赫兹波谱检测可能遇到吸收饱和的问题, 我们采用了环烯烃类共聚物 (Copolymers of Cycloolefin, COC) 粉末作为辅助太赫兹检测的稀释剂。COC 粉末的粒径小于 60 μm , 具有超过 96% 的太赫兹波高透射率, 能够有效地解决吸收饱和问题。所有待测样本在使用前均经过 COC 粉末稀释处理, 而原材料并未进行进一步的提纯。实验中所使用的水为经过纯水机处理的二次纯化水, 其电阻率大于 18.4 $\text{M}\Omega\cdot\text{cm}$ 。

1.2 太赫兹波谱测试的样品制备

艾司唑仑纯品样本的制备过程如下: 将 COC 粉末与艾司唑仑纯品按不同比例混合, 制备了 8 组不同含量的样本。具体来说, 将 1.5~12 mg 艾司唑仑粉末, 以 1.5 mg 为增量, 分别与 60 mg COC 粉末混合, 制成含量为 1.5 mg、3.0 mg、4.5 mg、6.0 mg、7.5 mg、9.0 mg、10.5 mg、12.0 mg 的样本。为了精准分析, 我们在后续艾司唑仑吸收光谱的处理中也以等量的 COC 光谱作为参考信号进行了扣除。

艾司唑仑实际药片样本的制备过程如下: 在实际药物测试中, 将药物样品先粉碎研磨,

选取质量范围为 4~15 mg 的药物粉末与 60 mg COC 粉末混合配比后压片, 以获取不同含量的艾司唑仑药片混合物样本。

压片制备过程如下: 将样本混合研磨后过筛得到的直径约为 60 μm 的颗粒放入模具中, 使用压片机在 3000 kg 压力下压制 3 min, 从而制得直径为 13 mm 的样品薄片。在整个制备过程中, 样本损耗率控制在 1% 以下。样品厚度控制在 500~580 μm 之间。为了计算单位厚度下的样品太赫兹吸收率, 在后续分析中将样品的吸光度除以样品厚度。

1.3 HPLC 分析的样品制备

含量测定: 参照 HPLC 标准流程测定。

供试品溶液制备过程如下: 取本品 20 片, 研成细粉 (约相当于艾司唑仑 1 mg), 置于 100 mL 量瓶中, 加适量流动相并充分振摇使艾司唑仑溶解。用流动相稀释至刻度, 摇匀均匀后滤过, 再取续滤液作为供试品溶液。

对照品溶液的制备过程如下: 取艾司唑仑对照品并精密称定, 加流动相溶解并定量稀释制成每 1 mL 中约含 10 μg 的溶液, 将其作为对照品溶液。

线性溶液: 5 个浓度水平 (50%、80%、100%、150%、200%) 的线性范围, 即 5 $\mu\text{g}/\text{mL}$ 、8 $\mu\text{g}/\text{mL}$ 、10 $\mu\text{g}/\text{mL}$ 、15 $\mu\text{g}/\text{mL}$ 、20 $\mu\text{g}/\text{mL}$ 。

色谱条件: 将十八烷基硅烷键合硅胶作为填充剂 (SunFire@ C18, 4.6 $\times$ 250 mm, 5 μm 或效能相当的色谱柱); 以乙腈-水 (2:3) 为流动相; 柱温为 30 $^{\circ}\text{C}$; 检测波长为 223 nm; 进样体积为 20 μL 。

1.4 实验仪器

本实验中采用傅里叶变换红外光谱仪 (Fourier Transform Infrared Spectrometer, FT-IR) 对所有样本进行扫描。采用水冷汞灯作为光源, 并选用 DLaTGS/聚乙烯检测器。光谱区域有效覆盖 0.9~20.0 THz, 信噪比优于 10000:1。本文中, 我们以 4 cm^{-1} 的分辨率、32 的扫描次数和 5 kHz 的扫描速度测得所有光谱。每个样本测试 4 次并取平均值。所有光

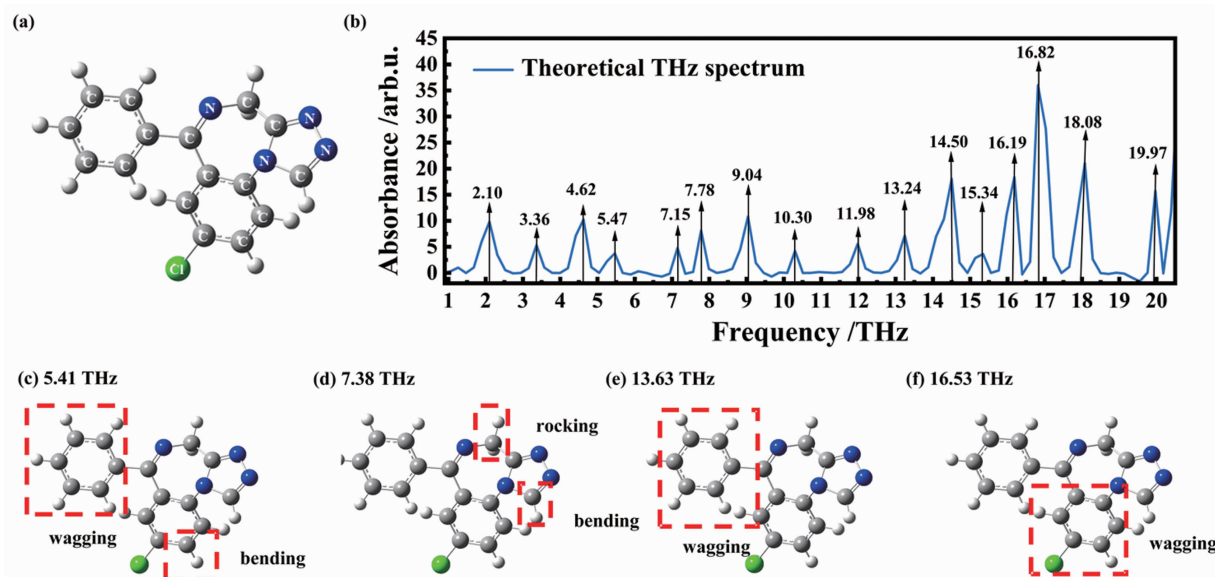


图 1 艾司唑仑的分子构型及理论波谱图：(a)分子构型图；(b)分子在 0.9~20.0 THz 波段的理论波谱；(c)~(f)经 DFT 计算的艾司唑仑分子分别在 5.41 THz、7.38 THz、13.63 THz 和 16.53 THz 处的振动模式图

谱均在室温(约 22℃)的真空环境下测量,以减少水蒸气对实验的影响。

用于 HPLC 的仪器为 Ultimate 3000 液相色谱仪、Vanquish Core 高效液相色谱仪、UV 检测器。色谱柱如下:十八烷基硅烷键合硅胶为填充剂,乙腈为色谱纯,并采用二次纯化水。

2 结果与讨论

2.1 理论仿真

太赫兹吸收峰的形成基于太赫兹波与有机物分子之间的共振吸收作用。基于分子几何优化和振动光谱属性的预测方法,我们可以在理论上预测艾司唑仑分子在太赫兹波段的特征吸收峰。DFT 可用于分子能量和结构、分子轨道、振动频率的理论计算。在本研究中计算的艾司唑仑分子结构参数,来源于有机小分子生物活性数据库(PubChem, CID: 3261),分子式为 $C_{16}H_{11}ClN_4$, 相对分子质量为 294.74 g/mol。我们借助量子化学软件 Gaussian 09W 对艾司唑仑分子进行理论计算。选取 DFT 理论 B3LYP 方法,在 6-31G(d) 基组下进行几何优化和分子振动频率计算。优化后的分子结构式如图 1(a)所示,计算结构如图 1(b)所示。分

子结构包含一个由碳氮原子掺杂的七元环、两个六元环(苯环和氯苯)以及一个由碳氮原子掺杂的五元环(称为三氮掺杂五元环)。

为了得到能够用于艾司唑仑分子结构定性识别的特征峰频率,我们统计了艾司唑仑在 0.9~20.0 THz 波段内的理论峰位、实验峰位以及两者之间的偏差(正偏差表示实验峰位比理论峰位高,负偏差表示实验峰位比理论峰位低)。峰位偏差的大小和符号可以反映实验条件、样品制备或分析过程中可能存在的不同类型的系统性误差,如表 1 所示。

可以看出,在 5.0~11.0 THz 和 13.0~17.0 THz 波段内,理论峰位与实验结果的误差均不超过 0.39 THz,表明理论计算结果与实验测试结果之间匹配度较高。因此,这两个波段可作为特征峰选择的重点关注区域。其中,5.41 THz 实验峰位与 5.47 THz 理论峰位之间仅有 0.06 THz 的偏差,可作为特征峰之一进行下一步分析。而对于一些出现较大偏差的频点,如 12.80 THz 实验峰位与 11.98 THz 理论峰位之间存在 0.82 THz 的偏差,推测可能的影响因素为理论计算时只考虑 0 K 条件下的单分子情况,而实验所测的是室温下的多晶

表 1 理论峰位与实验峰位的对比

理论峰位/THz	实验峰位/THz	偏差/THz
3.36	2.89	-0.47
4.62	3.95	-0.67
5.47	5.41	-0.06
7.15	7.38	0.23
7.78	7.82	0.04
9.04	9.18	0.14
10.30	10.22	-0.08
11.98	12.80	0.82
13.24	13.63	0.39
14.50	14.39	-0.11
15.34	15.40	0.06
16.19	16.53	0.34
16.82	17.35	0.53
18.08	18.62	0.54

材料, 计算模型差异影响了高频分子运动计算精度。

进一步, 我们分析了艾司唑仑在 0.9~20.0 THz 波段内特征峰对应的分子运动模式, 下面选取四个吸收峰作为代表, 分析它们对应的分子运动模式, 如图 1(c)~图 1(f) 所示。5.47 THz 处的理论峰与 5.41 THz 处的实验峰相对应, 主要源于苯环的面外摇摆振动和氯苯上的 C-H 键的面外弯曲振动。7.15 THz 处的理论峰与 7.38 THz 处的实验峰相对应, 主要源于三氮掺杂五元环的面外弯曲振动和亚甲基的面内摇摆振动。13.24 THz 处的理论峰与 13.63 THz 处的实验峰相对应, 主要源于苯环面外摇摆振动带动氯苯面内集体分子振动。16.19 THz 处的理论峰与 16.53 THz 处的实验峰相对应, 主要源于氯苯的面外摇摆振动。以上四个特征峰分别对应于艾司唑仑分子中不同官能团的低频运动, 因此联合分析可以更准确地对艾司唑仑进行识别。

2.2 建立艾司唑仑药物的定量分析方法

明确艾司唑仑在 0.9~20.0 THz 波段内有清晰的特征峰之后, 我们通过分析艾司唑仑含量与其特征峰之间的相关性, 分别建立吸光度、峰下面积与艾司唑仑含量之间的定量模型, 从而实现对该药物的定量检测。我们制备

了不同含量的艾司唑仑样品, 并使用 FTIR/THz 器件进行检测。如图 2(a) 所示, 随着艾司唑仑含量的增加, 样本在各个频点的吸光度也逐渐增加。

为了进一步量化, 我们选取了 5.41 THz、7.38 THz、13.63 THz、16.53 THz 四个特征峰来分析吸光度、峰下面积与样本含量之间的函数关系。在使用 Origin 软件的峰值分析功能计算吸光度和峰下面积时, 注意到具体的结果与分析时采用的吸收峰频率范围有关。本研究选用波谱一阶导数的极小值点来决定峰的频率范围, 对应的四个特征峰下的频率范围分别为 5.27~5.65 THz、7.24~7.56 THz、13.23~13.81 THz 和 15.84~16.88 THz。以吸收峰两侧最低点的切线为基线进行扣除, 并以特征峰频率范围内峰值最高点为吸光度, 取特征峰频率范围内的吸光度进行积分, 从而获得峰下面积。在选取的四个特征峰处, 为准确描述不同含量样品的吸光度和峰下面积的相关性, 我们对上述数据进行线性拟合, 结果如图 2(b)~图 2(i) 所示。相应的线性拟合函数表达式和决定系数(R^2)均标注在图上。可以看到, 不同指标对应的工作曲线中至少有一条对应的 R^2 超过 0.995, 很好地符合了朗伯-比尔定律。

进一步地, 我们想要知道这些指标在艾司唑仑含量检测中的定量检测性能, 即灵敏度 (Sensitivity, S) 和最低检测限 (Limit of Detection, LOD)。LOD 的计算公式为

$$LOD = \frac{3.3\delta}{S} \quad (1)$$

式中, δ 为响应值的偏差, 取标准曲线的剩余标准偏差; S 取标准曲线的斜率。表 2 列出了具体计算结果。可以看到, 同一个峰在不同指标(吸光度和峰下面积)下, LOD 和 S 也不相同。例如, 对于 5.41 THz 处的峰, 基于峰高指标的 S 相对较低, 基于峰下面积的 S 却相对较高。推测其原因应该是该峰幅度相对较小, 随艾司唑仑含量的变化不明显; 而峰宽较大, 导致峰下面积较大, 随艾司唑仑含量的变化明

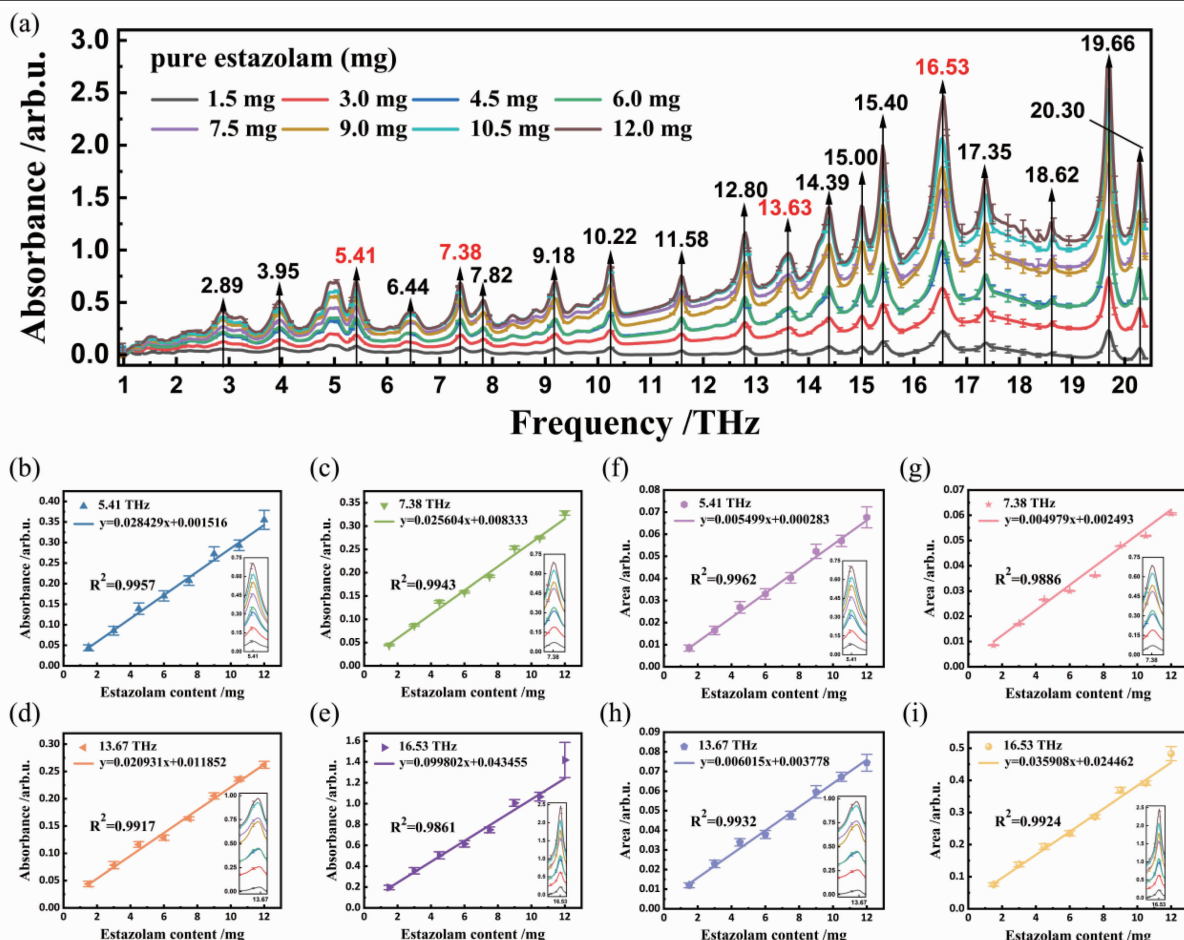


图 2 不同含量样品的波谱图及建立的定量分析曲线: (a)不同含量样品的波谱图; (b)~(e) 5.41 THz、7.38 THz、13.63 THz 和 16.53 THz 处的吸光度分析; (f)~(i) 5.41 THz、7.38 THz、13.63 THz 和 16.53 THz 处的峰下面积分析

表 2 基于太赫兹波谱技术的多指标艾司唑仑定量方法的检测性能

峰位/THz	R^2		LOD/mg		S/(arb. u.·mg ⁻¹)	
	吸光度	峰下面积	吸光度	峰下面积	吸光度	峰下面积
5.41	0.9957	0.9962	1.063	0.897	0.028	0.005
7.38	0.9943	0.9886	1.224	0.064	0.026	0.095
13.67	0.9917	0.9932	1.026	1.026	0.021	0.006
16.53	0.9861	0.9924	2.224	1.271	0.100	0.036

显。基于这四个峰,我们可以对艾司唑仑含量进行多指标联合判定,LOD 达到 0.064 mg(以 7.38 THz 处峰下面积为指标);S 达到 0.100 arb. u./mg(以 16.53 THz 处峰下面积为指标)。鉴于此,我们证明了太赫兹波谱技术对艾司唑仑药物含量进行高精度定量检测的可行性,并建立了多指标的艾司唑仑检测定量分析方法。

2.3 HPLC 检测的对比分析

HPLC 是现阶段麻醉药物艾司唑仑检测的

“金标准”。在 HPLC 色谱图中,艾司唑仑在不同浓度下展现出一致的保留时间,峰形对称且分离度好,说明了 HPLC 方法的选择性和重现性,如图 3(a)所示。根据 HPLC 含量测定方法,我们选择 8.6 min 处的峰,评价峰高、峰下面积与艾司唑仑浓度之间的关系,如图 3(b)和图 3(c)所示。校准曲线显示了艾司唑仑浓度与峰高、峰下面积之间的高线性关系($R^2=0.9997$ 与 $R^2=1$ 时),验证了 HPLC 方法的

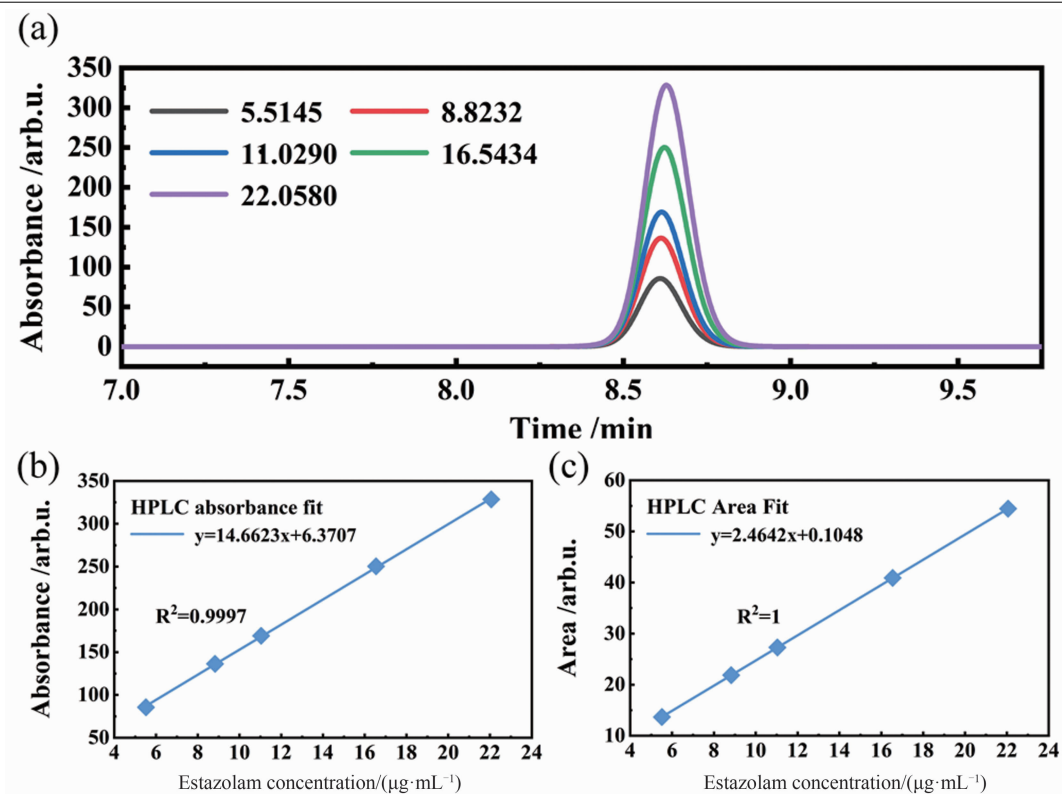


图 3 不同浓度艾司唑仑的 HPLC 色谱图及标准曲线: (a) HPLC 色谱图; (b) HPLC 峰高标准曲线; (c) HPLC 峰下面积标准曲线

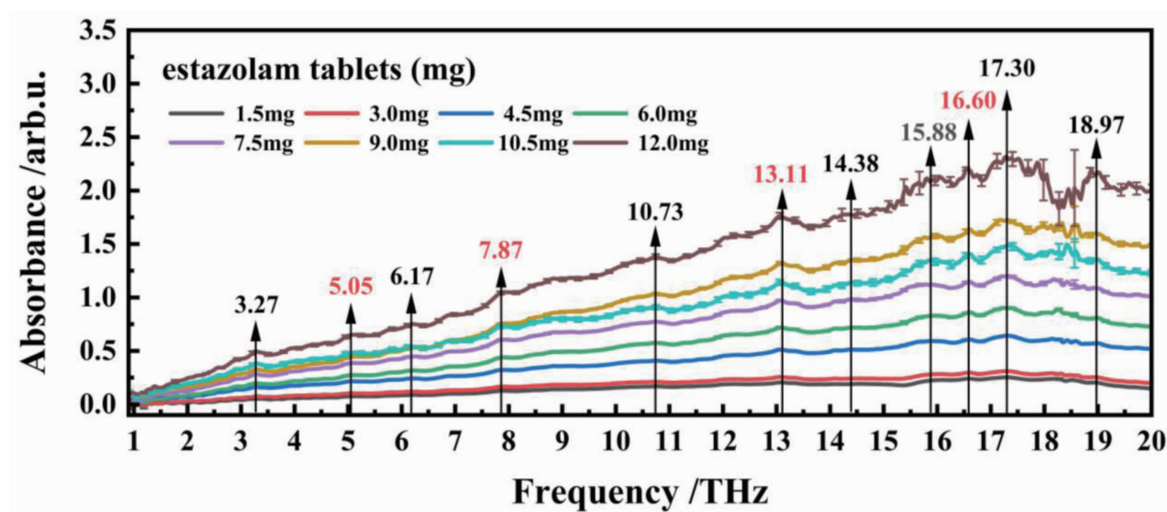


图 4 实际艾司唑仑药物在不同含量下的太赫兹波谱图

准确度和稳定性。

将两种技术的结果进行比较, 我们可以看到它们都能够对艾司唑仑进行准确的检测和分析, 且均展现出良好的线性响应。在定量分析方面, HPLC 以极高的精确度建立了标准, 但同时也面临着前处理繁琐、样品损耗、耗时长、成本高等问题。相比较而言, 太赫兹波谱

检测技术在定量分析上可凭借快速和无损检测的特性, 在快速筛查和初步分析领域展现出巨大的发展潜力。特别是太赫兹波谱技术在提升分析速度和可多指标检测上的显著优势, 不仅为药物检测和质量控制领域的快速决策提供了有力支持, 还预示了其在商业化应用中的巨大前景。

2.4 艾司唑仑实际药物的太赫兹解析

在建立基于太赫兹波谱的艾司唑仑纯品定性定量检测方法后,我们将其应用于实际药物的无损、快速解析。基于 22 例艾司唑仑实际药物,我们进行了太赫兹波谱测试。考虑到不同药厂的辅料成分差异,我们首先确定了实际药物与太赫兹辅料的最佳配比,然后考虑后续特征峰的识别和含量解析。其中一个药物的示例结果如图 4 所示。我们识别了 5.05 THz、7.87 THz、13.11 THz、16.60 THz 处的特征峰。这些峰与艾司唑仑分子的特定振动频率紧密相关,为药物核心成分的定性识别提供了可靠依据。药物波谱峰位的偏差可能源自于药物配方中的活性成分与辅料之间的相互作用。

进一步地,我们用类似的配比筛选方法,获得 22 例艾司唑仑实际药物的太赫兹波谱图(见图 5)。可以看到,图 5(b)、图 5(d)、图 5(f)、图 5(j)、图 5(k)谱线几乎重合;图 5(c)、图 5(e)、图 5(g)、图 5(i)谱线虽然在吸光度上有轻微差距,但是在峰形变化上表现较好。与其他图相比,图 5(a)表现不佳,但仍保留了重要特征峰的信息。这些结果凸显了

太赫兹波谱在识别药物和辅料特征峰方面的精确度以及对药物含量相对变化的敏感性。因此,通过分析前面重点标注的特征峰,我们可以实现对实际药物中艾司唑仑成分的快速定性识别。

由于样品与辅料之间的配比、多种成分间的相互作用等因素,传统的朗伯-比尔定律并不适用于多组分复杂样品的定量分析。在这种情况下,太赫兹波谱的信号解析和成分量化往往需要处理高维数据以及其中可能存在的非线性关系。对此,我们计划采用先进的算法来提取关键信息,减少噪声和冗余数据影响,建立一个稳健的模型来预测药物成分的含量。

尽管常见的主成分分析和偏最小二乘等方法在包含强线性关系的数据集中表现良好,但在面对波谱数据和组分含量之间存在未知或不确定的关系映射时会受到影响。因此,我们选择了 SVR 算法。该算法在处理未知关系和自动学习数据表示方面具有显著优势。其鲁棒性表现在处理噪声、轮廓线和不平衡分布数据上,同时它也具备处理线性和非线性问题时优

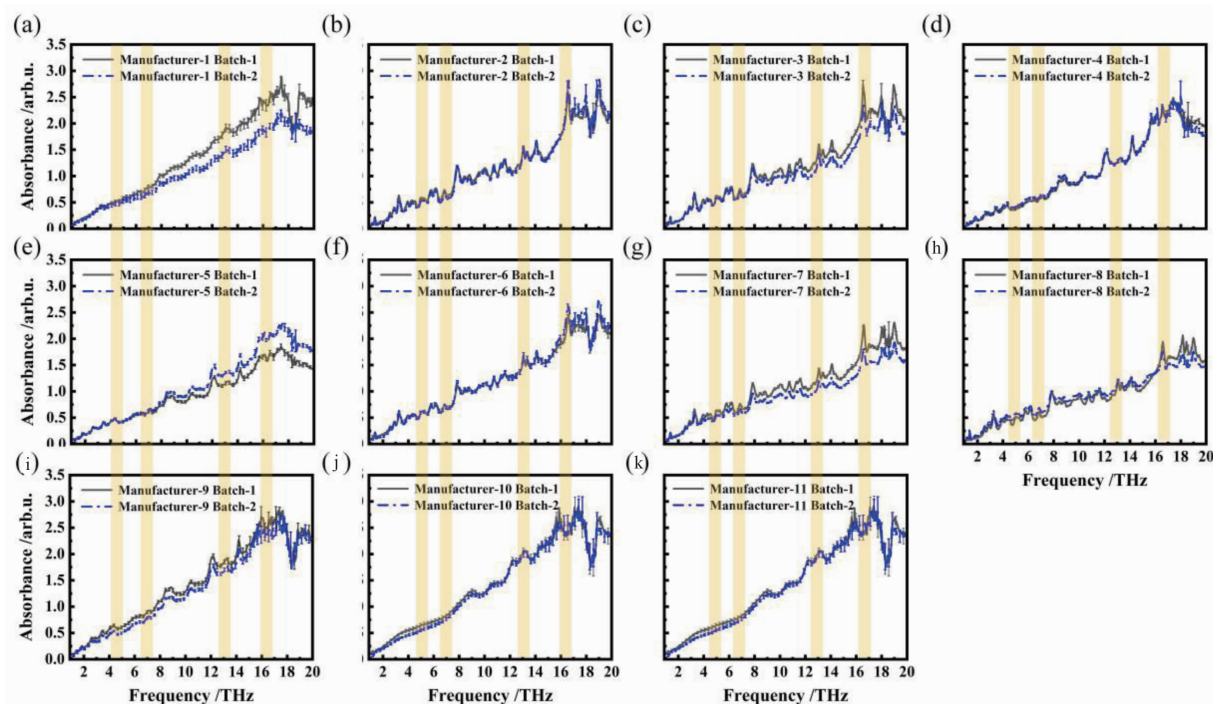


图 5 11 家药厂不同批次艾司唑仑片的太赫兹波谱

越的泛化能力。此外, 考虑到化学计量学方法(包括 SVR)通常对仪器、材料的变化敏感, 我们在实验过程中选用具有良好灵敏度和稳定性的 FTIR 设备, 并在样品制备过程中严格遵守标准化流程, 以确保数据的可重复性和可靠性。

根据 SVR 算法原理, 其目标是找到一个超平面, 使得训练数据点到这个超平面的距离尽可能小, 同时保持一定距离, 从而允许一定的误差。给定训练向量 $x_i \in R^p$, $i=1, \dots, n$ 。其中, x_i 表示实际药品的吸收波谱数据, 作为 SVR 模型的输入数据; i 是频谱数据中离散采样数据的个数。目标向量 $y \in R^p$, 表示实际药品中目标成分的含量(在本研究中就是药片中艾司唑仑的含量), 作为 SVR 模型的输出数据。SVR 能够解决以下主要问题:

(1) 最小化损失函数问题, 即

$$\min_{\omega, b, \zeta, \zeta^*} \frac{1}{2} \omega^T \omega + C \sum_{i=1}^n (\zeta_i + \zeta_i^*) \quad (2)$$

同时满足以下约束条件:

$$\omega^T \varphi(x_i) + b - y_i \leq \epsilon + \zeta_i^* \quad (3)$$

$$y_i - \omega^T \varphi(x_i) - b \leq \epsilon + \zeta_i \quad (4)$$

$$\zeta_i, \zeta_i^* \geq 0, i = 1, \dots, n \quad (5)$$

式中, ω 表示超平面的权重向量; b 是截距; ζ_i 和 ζ_i^* 是松弛变量, 作用是允许一些数据点位于间隔内或者误差范围内; C 是正则化参数, 用于控制间隔大小和进行误差的权衡。在这里, 我们拟合的样本是那些预测值偏离它们真实目标至少 ϵ 的样本。这些样本通过 ζ_i 拟合还是通过 ζ_i^* 拟合, 取决于它们的预测值是在 ϵ 的上方还是下方。

(2) 对偶问题, 即

$$\min_{\alpha, \alpha^*} \frac{1}{2} (\alpha - \alpha^*) Q (\alpha - \alpha^*) + e^T (\alpha - \alpha^*) - y^T (\alpha - \alpha^*) \quad (6)$$

满足以下约束条件:

$$e^T (\alpha - \alpha^*) = 0 \quad (7)$$

$$0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, \dots, n \quad (8)$$

式中, e 是单位向量, Q 是一个 $n \times n$ 的半正定

矩阵, $Q_{ij} = K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$ 是核函数。训练向量通过函数 φ 隐式映射到一个可能无限的高维空间。

(3) 得到预测结果:

$$\sum_{i \in SV} (\alpha - \alpha_i^*) K(x_i, x) + b \quad (9)$$

SVR 模型通过最小化损失函数来寻找超平面的最优解, 从而进行预测。

由于实际药物中除艾司唑仑之外的其他辅料的波谱未知, 我们需要建立并训练一个 SVR 模型来寻找样品波谱与目标成分含量之间的规律。我们的方法是使用 Pycharm 来构建 SVR 模型。为了拟合实验得到的数据, 在考虑到 SVR 模型的复杂性和错误率的基础上, 用一个最小偏离值 ϵ 的给定余量来接近最佳值, 从而确定实际药品中艾司唑仑的含量。

搭建和分析艾司唑仑的 SVR 定量预测模型的主要步骤包括预处理波谱数据、定义 SVR 模型、评估模型的预测性能。其中, 为了保证模型性能, 设置合适的核函数、正则化参数 C 和 ϵ 范围, 根据数据特征采用网格搜索方法确定最优参数, 从而利用这些参数建立 SVR 预测模型来预测艾司唑仑含量。为了遍历所有数据特征, 使用 5 折交叉验证法, 将所有样本按照 1:3 的比例随机划分为预测集和训练集。其中一部分作为预测集来验证模型, 另外三部分作为训练集来构建模型。最后, 对得到的预测集结果进行汇总。使用 MAE 指标来评估模型的准确性:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y'_i - y_i| \quad (10)$$

式中, n 为样本数量, 代表 22 家不同批次的实际药品编号, 在图中作为横坐标; y'_i 为实际结果, 在图中用黑色方块表示, 代表使用 HPLC 对同批次药品中艾司唑仑含量检测的结果; y_i 为预测结果, 在图中用红色圆形表示, 代表基于 SVR 模型对输入的太赫兹波谱数据进行预测得到的结果; 中间 1.0 mg 处的虚线表示实际药品的规格(每一片制剂中含有主药的含量)。实际结果和预测结果之间的比较如

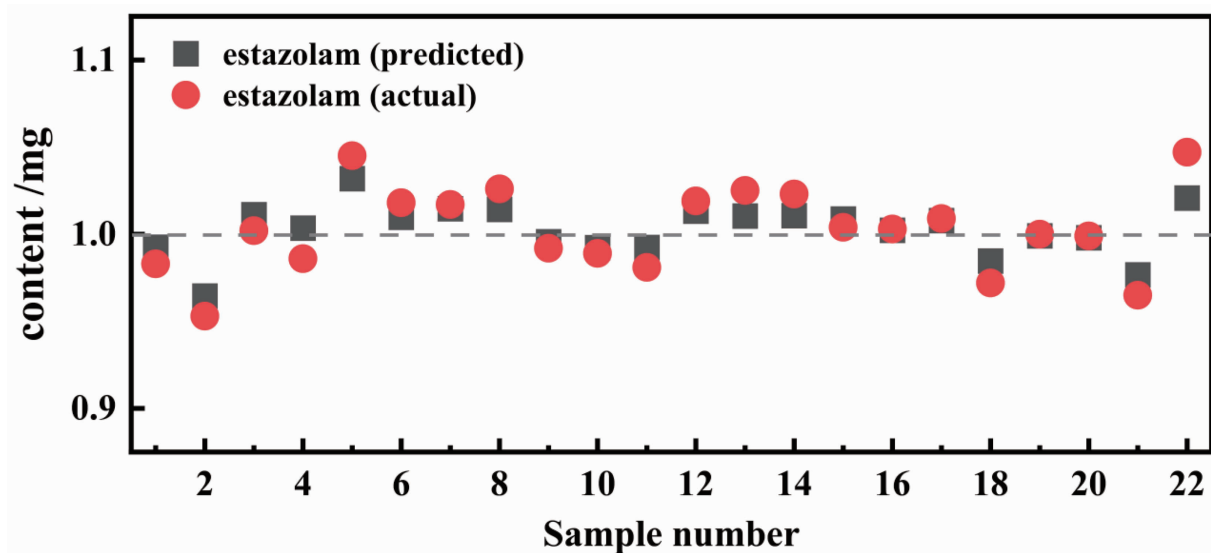


图 6 22 个实际样本中实际值和预测值的对比图

表 3 22 个实际药物中艾司唑仑的实际含量和预测含量

实际质量/mg	预测质量/mg	误差	实际质量/mg	预测质量/mg	误差	实际质量/mg	预测质量/mg	误差
0.983	0.993	0.0095	0.953	0.965	0.0116	1.002	1.011	0.0093
0.986	1.004	0.0175	1.045	1.032	0.0132	1.018	1.010	0.0083
1.017	1.015	0.0024	1.026	1.014	0.0117	0.992	0.996	0.0037
0.989	0.992	0.0033	0.981	0.992	0.0114	1.019	1.013	0.0057
1.025	1.010	0.0146	1.023	1.011	0.0124	1.004	1.009	0.0047
1.003	1.002	0.0007	1.009	1.008	0.0012	0.972	0.985	0.0126
1.000	0.999	0.0010	0.999	0.998	0.0010	0.965	0.977	0.0116
1.047	1.021	0.0262						

图 6 所示。表 3 列出了实际值、预测值以及预测值相对实际值的误差。

可以看到,相对于实验值,定性鉴别准确,艾司唑仑的 MAE 值为 1.12%。与本阶段 HPLC 检测药物中艾司唑仑的合格范围(90.00%~110.00%)相比,太赫兹检测结合 SVR 模型可有效应用于艾司唑仑药物的定性与定量检测。该方法也适用于所有具有太赫兹特征峰的物质。可通过 FTIR/THz 器件进行检测,也可推广到其他关键物质组分的分析。

3 结束语

本文成功建立了一种基于太赫兹波谱技术的麻精药物艾司唑仑的分析方法,可实现对此类药物的快速、准确且无损的定性定量评估。

首先,以艾司唑仑分子结构为基础,结合 DFT 计算与太赫兹波谱实验结果,我们精确识别了艾司唑仑在太赫兹频域(0.9~20.0 THz)内的特征频点,实现了对艾司唑仑的无损定性识别。接着选取四个太赫兹特征峰,并建立吸光度与峰下面积对艾司唑仑含量的高拟合度工作曲线,为药物定量提供了可靠依据。之后,结合 SVR 算法对实际药物样品含量进行了预测。与 HPLC 数据相比,MAE 值为 1.12%。检测结果的一致性和可重复性证实了所提方法的有效性和可靠性。本研究成果可推广到其他麻精药物质量评价研究中,对特殊药品的快检分析具有重要研究价值。

参考文献

- [1] Zhang X, Shi J, Huang X, et al. Formation and occurrence of disinfection byproducts of benzodiazepine drug estazolam in drinking water of Beijing [J]. *Science of The Total Environment*, 2022, **804**(7): 150028.
- [2] Goldschen-Ohm M P. Benzodiazepine modulation of GABAA receptors: A mechanistic perspective [J]. *Biomolecules*, 2022, **12**(12): 1784.
- [3] 包纪盛, 林志, 陈亚萍. 艾司唑仑联合尼莫地平改善老年人睡眠质量的临床观察 [J]. *中国老年学杂志*, 2011, **31**(3): 518-519.
- [4] 冯慧, 刘义, 徐虹, 等. 针刺与艾司唑仑对慢性失眠患者情景记忆及睡眠结构的影响: 随机对照研究 [J]. *中国针灸*, 2020, **40**(7): 707-712.
- [5] 林雪华. 慢性失眠症应用右佐匹克隆与艾司唑仑治疗的临床分析 [J]. *世界睡眠医学杂志*, 2019, **6**(5): 557-559.
- [6] 苏红梅, 金伟国, 汤洁. 加味黄连温胆汤治疗痰瘀交阻型慢性失眠疗效观察 [J]. *上海中医药杂志*, 2013, **47**(4): 40-42.
- [7] 王晓彤, 林海雄, 金远林. 艾司唑仑治疗失眠症临床疗效及安全性的网状 Meta 分析 [J]. *中国现代应用药学*, 2018, **35**(5): 735-740.
- [8] 张修红, 陈金雄, 郑关毅, 等. 甘麦苡仙磁石汤联合艾司唑仑片治疗老年慢性失眠心悸两虚证临床观察 [J]. *中国中西医结合杂志*, 2020, **40**(6): 666-671.
- [9] 中国药典委员会. *中华人民共和国药典(二部)* [M]. 北京: 中国医药科技出版社, 2020.
- [10] 杨航, 赵红卫, 张建兵, 等. 生物组织脱水过程的太赫兹时域光谱 [J]. *红外与毫米波学报*, 2014, **33**(3): 263-267.
- [11] 段瑞鑫, 赵红卫, 朱亦鸣. 基于太赫兹技术的药物检测 [J]. *物理*, 2013, **42**(11): 781-787.
- [12] Maleki A, Singh A, Jaber A, et al. Metamaterial-based octave-wide terahertz bandpass filters [J]. *Photonics Research*, 2023, **11**(4): 526-532.
- [13] Xu Z, Zeng Y, Zhao X, et al. High photoelectric conversion efficiency and fast relaxation time of $\text{FA}_{0.4}\text{MA}_{0.6}\text{PbI}_3$ applied in ultrafast modulation of terahertz waves [J]. *Journal of Energy Chemistry*, 2024, **91**(4): 299-305.
- [14] Liang M, Weinan S, Fei F, et al. Terahertz polarization sensing, chirality enhancement, and specific binding based on metasurface sensors for biochemical detection: a review [J]. *Chinese Optics Letters*, 2023, **21**(11): 14-31.
- [15] Zhao H, Wang X, Liu S, et al. Highly efficient vectorial field manipulation using a transmitted trilayer metasurface in the terahertz band [J]. *Optoelectronic Advances*, 2023, **6**(2): 31-42.
- [16] Huang J, Hu B, Wang G, et al. BICs-enhanced active terahertz wavefront modulator enabled by laser-cut graphene ribbons [J]. *Photonics Research*, 2023, **11**(7): 1185-1195.
- [17] 范丽洁. 太赫兹技术在医药领域的应用研究 [D]. 北京: 北京交通大学, 2015.
- [18] 彭滢, 施辰君, 朱亦鸣, 等. 太赫兹光谱技术在生物医学检测中的定性定量分析算法 [J]. *中国激光*, 2019, **46**(6): 17-24.
- [19] 彭滢. 太赫兹成像技术在肿瘤检测中的应用 [J]. *光电工程*, 2020, **47**(5): 69-78.
- [20] 彭滢. 生物医学检测中太赫兹光谱技术的算法研究 [J]. *光学学报*, 2021, **41**(1): 406-415.
- [21] 彭滢. 基于太赫兹超材料芯片的生物混合物定量检测研究 [J]. *中国激光*, 2021, **48**(23): 181-189.
- [22] 张琪, 方虹霞, 秦丹, 等. 太赫兹时域光谱技术定性定量分析抗结核药物中的吡嗪酰胺和异烟肼 [J]. *药物分析杂志*, 2016, **36**(6): 1082-1088.
- [23] 吴静, 吴旭, 黄星皓, 等. 基于太赫兹光谱的华法林钠快速定性定量检测 [J]. *中国激光*, 2024, **51**(3): 177-184.
- [24] 迟爽, 吕琴, 刘晓亭, 等. 太赫兹技术在中医学领域的应用与展望 [J]. *时珍国医国药*, 2024, **35**(2): 425-428.
- [25] 曹丙花, 侯迪波, 颜志刚, 等. 基于太赫兹时域光谱技术的农药残留检测方法 [J]. *红外与毫米波学报*, 2008, **27**(6): 429-432.
- [26] 杜勇, 夏焱, 汤文建, 等. 基于太赫兹时域光谱技术的磺胺甲噁唑多晶型现象 [J]. *红外与激光工程*, 2014, **43**(9): 2919-2924.

- [27] 袁婷婷, 吴靖文, 薄艳华, 等. 基于太赫兹超表面传感器的硝基呋喃类药物痕量检测 [J]. *光学学报*, 2023, **43**(7): 171–179.
- [28] 罗洁, 寇天一, 施辰君, 等. 神经细胞中五种常见物质的太赫兹波谱的推定 [J]. *太赫兹科学与电子信息学报*, 2019, **17**(1): 35–39.
- [29] 寇天一. 基于太赫兹光谱的人参和西洋参鉴别 [J]. *光学仪器*, 2020, **42**(5): 27–32.
- [30] 徐鸣谦, 寇天一, 彭滢, 等. 基于多个太赫兹时域光谱系统的物质识别方法 [J]. *光学仪器*, 2019, **41**(2): 28–33.
- [31] Wu Z, Zhu Z, Cheng C, et al. Terahertz spectroscopy of enantiomeric and racemic pyroglutamic acid [J]. *Spectrochim Acta A: Mol Biomol Spectroscopy*, 2020, **225**: 117509.
- [32] Yang Q, Wu L, Shi C, et al. Qualitative and quantitative analysis of caffeine in medicines by terahertz spectroscopy using machine learning method [J]. *IEEE Access*, 2021, **9**: 140008–140021.
- [33] Gu H Y, Wang S F, Hu S Y, et al. Identification of Panax notoginseng origin using terahertz precision spectroscopy and neural network algorithm [J]. *Talanta*, 2024, **274**: 125968.