

近红外光谱技术在卷烟配方设计与相似度评估中的应用综述

王修铭, 白 钰, 王维维

红云红河烟草(集团)有限责任公司新疆卷烟厂, 新疆 乌鲁木齐

收稿日期: 2025年12月9日; 录用日期: 2026年1月2日; 发布日期: 2026年1月13日

摘要

本文系统综述了近红外光谱技术在卷烟配方设计与相似度测量中的关键方法与应用进展。在配方设计方面, 重点阐述了光谱数据的预处理、特征提取、基于遗传算法等非线性优化方法的配方回归模型, 以及叶组配方的解析与数字化设计平台构建。在相似度评价方面, 探讨了多种算法(如余弦相似度、Jaccard系数、汉明距离)的融合应用, 以及基于Fisher投影的质量判别模型。研究表明, 这些技术能够有效提升配方预测的准确性、优化效率和产品质量一致性, 并为卷烟叶组配方的配方构建、快速维护与稳定性控制提供了系统的技术框架与方法支持。

关键词

近红外光谱技术, 卷烟配方设计, 相似度评估

Review of the Application of Near-Infrared Spectroscopy Technology in Cigarette Formulation Design and Similarity Assessment

Xiuming Wang, Yu Bai, Weiwei Wang

Xinjiang Cigarette Factory, Hongyun Honghe Tobacco (Group) Co., Ltd., Urumqi Xinjiang

Received: December 9, 2025; accepted: January 2, 2026; published: January 13, 2026

Abstract

This paper systematically reviews key methods and application advances in near-infrared

文章引用: 王修铭, 白钰, 王维维. 近红外光谱技术在卷烟配方设计与相似度评估中的应用综述[J]. 交叉科学快报, 2026, 10(1): 258-263. DOI: 10.12677/isl.2026.101032

spectroscopy technology for cigarette blend formulation and similarity measurement. Regarding formulation design, it elaborates on spectral data preprocessing, feature extraction, formulation regression models based on nonlinear optimization methods such as genetic algorithms, as well as the analysis of leaf group formulations and the construction of a digital design platform. For similarity evaluation, it explores the integrated application of various algorithms (e.g., cosine similarity, Jaccard index, Hamming distance) and quality discriminant models based on Fisher projection. Research indicates that these technologies can effectively enhance the accuracy of formulation prediction, optimization efficiency, and product quality consistency. Furthermore, they provide a systematic technical framework and methodological support for formula construction, rapid maintenance, and stability control of cigarette leaf group formulations.

Keywords

Near-Infrared Spectroscopy Technology, Cigarette Formulation Design, Similarity Evaluation

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在公共卫生关切日益增加、环保政策收紧与市场竞争加剧的多重背景下，烟草行业正面临向高质量、低释放与可持续方向转型的关键阶段。卷烟产品的内在品质、感官稳定性及潜在风险控制，从根本上取决于其配方设计与工艺水平。近年来，近红外光谱技术作为一种快速、无损的分析手段，因其高效、环保及适于在线监测的特点，在烟草行业获得了广泛应用与研究关注。该技术不仅能够实现对烟叶原料化学成分的快速预测，更能为配方设计、相似度评价与生产过程控制提供客观、量化的数据支撑，从而推动卷烟产品研发从传统经验模式向数字化、可控的方向转型[1]。

当前，针对近红外光谱在卷烟领域的应用，已有诸多研究聚焦于光谱预处理、特征提取、定量预测模型构建及相似性分析等方面，并在配方回归优化、叶组结构解析等方面取得了实质性进展。然而，相关技术方法较为分散，尚未形成面向行业技术人员的系统性技术框架与应用指南。为此，本文旨在通过梳理与整合近年来的重要研究成果，系统阐述近红外光谱技术在卷烟配方设计、相似度测量及叶组配方维护等关键环节中的方法体系与应用实践。本综述系统梳理了从数据预处理到模型构建的全流程关键技术，并对多种算法在实际生产环境中的融合应用策略进行了归纳总结。研究旨在提供一个清晰、完整的方法学参考框架，以支持相关制造领域在工艺控制、质量一致性与产品开发效率方面的技术优化，从而为行业推进技术升级与生产流程改进提供理论依据与应用参考[2]。

2. 基于近红外光谱的卷烟配方设计

(一) 近红外光谱数据预处理与特征提取技术

为有效消除光谱数据中由颗粒不均、光程差异等因素引起的光散射效应和基线漂移，现有研究通常采用标准正态变量变换(SNV)与多元散射校正(MSC)相结合的方法对原始近红外光谱进行预处理[3]。该方法被证实能够显著提升光谱的信噪比与可比性，为后续分析奠定稳定的数据基础。进一步地，在相关文献中，连续波变换(CWT)常被用于滤除高频噪声，以增强关键光谱特征，提高数据的解析度[4]。在数据降维方面，概率主成分分析(PPCA)等方法可用于在保留主要光谱信息的同时降低计算复杂度，有助于避免模型过拟合。而在特征提取环节，已有工作广泛探讨了从预处理后的光谱中筛选代表性波长区间的多

种策略，其中非线性降维算法如局部线性嵌入(LLE)和等距映射(ISOMAP)受到了较多关注，学者们通过比较这些方法的性能，为不同应用场景提供了算法选择的依据[5]。整体而言，上述系列方法构成了光谱数据预处理与特征提取的一套较为完整的技术体系，为后续的建模与分析提供了重要的方法论基础。

(二) 配方回归与非线性多变量优化算法

配方回归旨在通过优化叶组配比，实现特定质量目标的精准达成。现有研究表明，可采用基于遗传算法的非线性回归模型，对叶组配方的化学组成进行预测与优化。遗传算法凭借其优秀的全局寻优能力，已被证实在处理多维、非线性的复杂配方问题中具有显著优势[6]。相关文献指出，在依据卷烟产品的感官与化学指标构建目标函数后，通过遗传算法的选择、交叉、变异等操作进行多代迭代寻优，能够逐步逼近最优解。与传统线性回归及经验调配方法相比，杨等人[7]通过基于遗传算法计算最佳出库搭配组合，与随机出库对照组相比各等级成品片烟烟碱变异系数降幅范围在10%~30%，糖碱比变异系数降幅范围在20%~50%，说明该方法在提升配方设计科学性与一次成功率方面具有可行性与有效性，为产品质量一致性控制与配方快速迭代提供了可供参考的技术路径。

(三) 叶组配方解析与设计

叶组配方的解析与设计是通过近红外光谱分析现有配方特征，进而科学指导新配方构建的过程。研究者们将非线性多变量算法应用于叶组配方的光谱解析，系统实现了对多个叶组配方结构的有效分析与重构。以某特定品种叶组为研究对象，通过对其三种主要原料进行替代性试验，相关研究显示所得新配方的关键常规化学指标与原始配方相比，绝对误差控制在0.3以内，表明二者具有较高的化学组成相似性与品质一致性。在此基础上，研究进一步开发了基于MATLAB图形用户界面的配方分析与维护集成系统，该平台实现了从光谱解析、配方比对到结果可视化的全流程整合，显著简化了配方设计操作步骤，提升了工作效率与系统可操作性，为叶组配方的快速优化与稳定性控制提供了实用的辅助工具[8]。

3. 基于近红外光谱的卷烟相似度测量

(一) 基于多种算法的相似度测量

在卷烟产品的相似度评价中，综合运用余弦相似度、Jaccard系数及汉明距离等多种测度方法，可以实现对产品间相似性的多维度精确评估。不同方法各有其适用场景与理论优势：余弦相似度适用于衡量向量间方向一致性，其计算结果不受向量模长影响，特别适用于评价以成分为基础的配方结构相似性；Jaccard系数通过计算交集与并集的比例，适用于评估离散特征集合之间的相似与差异程度，在类别型指标对比中具有良好解释性；汉明距离则用于度量等长字符串在对应位置上的差异数量，在编码化质量特征的一致性比对中效果显著。孔等人[6]在复杂多维的卷烟特性数据中，融合使用多种相似度测度方法能够显著提升评估的全面性与稳健性，避免单一方法可能带来的系统性偏差。这种多元综合评估策略不仅能够更准确地反映产品间的真实相似关系，也为后续基于相似度分析的配方优化、质量一致性控制及新产品对标开发提供了可靠的数据基础与决策依据。

(二) 联合相似度算法的构建

在卷烟配方设计的相似性评价过程中，为有效整合不同相似度度量方法的优势，研究者提出了构建联合相似度算法的解决方案。该方法通过加权平均、层次融合等多种数据融合策略，将来自不同测度模型的计算结果进行有机整合，最终形成一个具有更高代表性的综合相似度得分。这种融合方法不仅能够显著提升评估结果的数据一致性和可信度，还能更全面、准确地反映样本间的真实相似关系。具体而言，在实际应用中可先利用余弦相似度获取样本在特征空间中的方向相似性初值，再结合杰卡德系数与汉明距离分别从集合匹配度和方位差异角度对初值进行修正与加权集成，从而构建出一个层次清晰、信息互补的复合评价体系。该联合相似度算法在卷烟配方与原料的相似性分析中表现出良好的测量准确性、稳

定性与适应性，能够为配方优化、原料替代及新产品开发提供更为可靠的理论依据与决策支持，对实现卷烟产品感官质量的稳定控制具有重要的技术参考价值[9]。

(三) 基于 Fisher 投影的质量模型

基于费歇尔投影的质量模型是评估卷烟相似度的另一种有效工具。费歇尔投影是一种能把高维数据向低维空间映射，并使类之间的差别最大、类中的变化最小的一种统计方法。该方法以烟叶中 70 种常量和半微量化学成分为基础，通过引入类多项式核函数将样本映射至高维特征空间，并基于核 Fisher 分类判别算法建立烟叶质量可用性判别模型，从而实现对烟叶质量可用性的有效识别。该方法不仅通过非线性建模提高了判别准确性，而且基于成分清晰的常量与半微量化学成分，使不同质量等级烟叶的物质基础更为明确，实现了对烟叶质量可用性的快速、客观、稳健且精准的表征。在此基础上，通过计算各烟叶样本与不同质量可用性质心的距离标准差，可对烤烟质量可用性作进一步细分与数字化赋值，从而为后续卷烟叶组配方及打叶复烤模块配方的数字化设计提供可靠依据[10]。

4. 基于近红外光谱的叶组配方维护

(一) 叶组配方解析的数学模型

针对传统叶组配方设计中客观性不足、效率偏低的问题，相关研究提出了一种基于近红外光谱表征烟叶品质的配方设计方法。该方法首先运用偏最小二乘法，建立烟叶化学成分、部位及香型与近红外光谱之间的预测模型，进而实现对烟叶品质指标的快速预测。在此基础上，以各项品质指标构建多维向量，并计算其综合质量表征指数；依据“相似相替”原则，参照目标配方中的原烟等级，筛选出品质相近的替代烟叶，形成候选烟叶集合；最后通过线性规划方法求解各替代烟叶在配方中的最优比例。为验证该方法的有效性，研究对 2020 年度四川烟叶进行了配方设计与效果验证，结果表明：替代配方在烟叶部位及总糖、还原糖、总植物碱、总氮等关键化学成分含量上，与目标配方的相对误差均小于 6%；感官评价显示，除刺激性指标外，替代配方在香气量、清晰度、透发性、成团性、杂气、余味及劲头等方面，与目标配方均无显著差异，说明该方法在保持配方感官质量一致性的同时，显著提升了配方设计的客观性与效率[11]。

(二) 基于光谱相似度测量的叶组配方维护

叶组配方养护是指在维持原有配方整体质量稳定的前提下，通过对部分原料进行科学调整与替代，实现配方的可持续维护与优化。为系统指导这一过程，现有研究通过对比待调整叶组与标准配方样品的光谱特征，建立基于光谱相似度的配方维护指导体系。该体系综合运用余弦相似度、Jaccard 系数、汉明距离等多种相似性度量方法，构建多维度的联合相似度评估模型，从而全面量化新旧配方间的光谱一致性。李等人[12]发现，当新配方与原有配方组合的相似度高于 0.85 时，其关键化学指标与传统质量参数的平均相对误差可控制在 0.3 以内，说明新配方在感官与理化性质上与原配方具有高度一致性。基于该相似度体系的配方养护程序，能够在确保产品质量稳定不受影响的前提下，为配方人员提供科学、客观的调整依据，支持其在原料供应波动、成本控制或工艺优化等场景下，实现配方结构的灵活、可靠调整，从而提升配方管理的适应性与精细化水平。

(三) 叶组配方设计平台的构建与应用

甘肃烟草工业有限责任公司研发了一种用于预测叶组配方烟气常规化学成分释放量的方法与系统。该方法首先采集烤烟样品，经预处理后获取近红外光谱数据，并采用 Kennard-Stone 算法筛选代表性样品，划分为校正集与验证集。随后，将样品制成烟丝并卷制成试验烟支，依据国家标准方法测定其单位质量燃烧烟丝中常规成分的释放量。在此基础上，结合近红外光谱数据建立烟气常规化学成分释放量的预测模型，并进行系统优化。最终，通过对叶组配方中所含烤烟样品的导数光谱进行线性叠加，利用已

构建的释放量模型, 实现对卷烟配方的烟气常规化学成分释放量的准确预测。该方法为卷烟配方的数字化、定量化设计提供了重要数据支撑[13]。红云红河烟草(集团)有限责任公司公开了一项用于快速预测叶组配方模块中薄片比例的方法及相关存储介质。该方法基于近红外光谱技术, 通过建立化学成分特征与薄片比例之间的分类预测模型, 实现快速、准确的薄片掺配比例监测。其具体步骤包括: (1) 依据实际生产中的检测精度要求, 确定分类范围及组数; (2) 使用纯烟片与纯薄片配制标准样品; (3) 利用近红外光谱预测烟叶的13种化学成分, 并计算4种比例关系, 共形成17种化学成分特征; (4) 基于上述特征构建分类预测模型; (5) 以相同预处理方式对待测叶组配方样品进行光谱扫描与处理; (6) 将样品对应的17种化学成分特征输入分类模型进行计算; (7) 输出薄片比例的快速预测结果。该方法适用于生产过程中的实时监控, 能够有效识别薄片掺配异常, 为叶组配方质量稳定性控制提供了可靠的技术手段[14]。

5. 结论

当前近红外光谱技术在卷烟配方设计与相似度评价领域已形成从光谱预处理、特征提取、回归建模到多算法融合评价的完整方法体系, 显著提升了产品研发的数字化水平与质量一致性。然而, 该技术在实际应用中仍面临多重挑战: 其一, 模型可移植性与普适性不足, 受限于特定原料样本与生产条件; 其二, 多源异构数据(如感官评价、物理指标、工艺参数)的融合机制尚未成熟, 制约了综合分析效能; 其三, 现有预测模型多为“黑箱”, 缺乏对工艺参数影响的合理解释性; 其四, 在线实时监测与动态调控体系尚未形成完整闭环, 限制了其在生产现场的即时应用。展望未来, 建议重点开展以下方向研究: 发展基于迁移学习与领域自适应的稳健性模型以提升跨场景适用性; 构建融合多模态数据的综合分析框架, 实现品质多维度精准预测; 引入可解释人工智能技术增强模型决策透明度与工艺关联性; 研发集成边缘计算与实时反馈的智能监测调控系统, 推动近红外光谱技术从实验室分析向生产线全过程智能控制演进。通过上述技术路径的探索与完善, 有望推动相关制造工艺进一步向数据驱动与智能决策方向演进, 从而为行业在提升生产可控性与过程一致性方面提供系统性的方法学支持。

参考文献

- [1] 周锋, 王黎亚, 林原, 等. 烟草制品合理化布局数字化转型框架体系研究[J]. 中国市场, 2025(33): 197-200.
- [2] 刘克强, 聂红资, 郭文强, 等. 近红外光谱技术在烟草领域的研究进展及前景展望[J]. 模型世界, 2025(27): 240-242.
- [3] 何昕. 基于近红外光谱和化学计量学的植物叶片分类研究[D]: [硕士学位论文]. 安徽: 中国科学技术大学, 2024.
- [4] 谢有超, 彭黔荣, 杨敏, 等. 基于烟丝近红外光谱的卷烟品牌识别方法[J]. 烟草科技, 2021, 54(3): 72-77.
- [5] 仇逊超, 曹军, 张怡卓. 基于流形学习的红松仁脂肪近红外定量检测[J]. 浙江农业学报, 2023, 35(8): 1915-1926.
- [6] 孔波, 蔡佳校, 邹有, 等. 基于NSGA-II遗传算法的烟用香精数字化调香研究及应用[J]. 烟草科技, 2020, 53(2): 72-79.
- [7] 杨江平, 林珈夷, 易百灵, 彭云发, 李小生, 张薇薇. 基于均质化加工的烟叶生产投料方式对比研究[J]. 南方农机, 2024, 55(6): 29-35.
- [8] 谢有超. 基于近红外光谱的卷烟相似性度量及数字化叶组配方设计研究[D]: [硕士学位论文]. 贵阳: 贵州大学, 2021.
- [9] 王林, 郑明明, 王翀, 等. 基于近红外光谱的卷烟配方模块香型预测[J]. 华中农业大学学报, 2024, 43(1): 226-231.
- [10] 云南中烟工业有限责任公司. 一种基于工业大数据的卷烟制造智能排产调度方法[P]. 中国, CN202411090833.1. 2025-03-25.
- [11] 郝贤伟, 彭钰涵, 杨泽会, 等. 基于近红外光谱表征品质设计烟叶配方[J]. 湖南农业大学学报(自然科学版), 2023, 49(3): 284-290.
- [12] 李华杰, 王道铨, 朱叶梅, 等. 近红外光谱结合模式识别方法所建模型分析卷烟烟丝配方比例[J]. 理化检验-化学分册, 2022, 58(7): 760-767.

-
- [13] 甘肃烟草工业有限责任公司. 预测叶组配方烟气常规化学成分释放量的方法及系统[P]. 中国, CN202510290970.8. 2025-07-08.
 - [14] 红云红河烟草(集团)有限责任公司. 一种快速预测叶组配方模块中薄片比例的方法及存储介质[P]. 中国, CN202510194376.9. 2025-05-23.