基于相关性分析的卷烟吸阻预测方法

Linsheng Chen[a]，Zhonghua Yu[a]，*Bo Zhang[b]，Qiang Zhu[b]，Hu Fan[b]，Yucan Qiu[ba] 浙江大学机械工程学院，杭州310030，中国[b] 浙江中烟工业有限责任公司技术中心，杭州 310008，中国

摘要：

卷烟吸阻监测方法尚不完善且较为单一，缺乏相关性分析和预防性建模，导致不合格卷烟流入市场。为在不增加硬件成本的前提下解决这一问题，本文采用多指标相关性分析预测卷烟吸阻。首先，基于现有质量控制框架分析吸阻监测过程，并提出优化思路。此外，针对三个生产单元——切丝供给（VE）、烟草卷制（SE）和卷烟成型（MAX），基于线性和非线性相关性分析，探究与吸阻相关的直接因素和潜在因素。随后，将吸阻的相关因素作为机器学习模型的输入，将吸阻预测值作为输出。最后，本研究还创新性地验证了吸阻预测的实际应用价值：基于预测结果分析不合格卷烟的分布特征，确定不合格卷烟的生产时间间隔，推导不合格卷烟被抽样的概率模型，并通过实例进一步验证预测结果的可靠性。结果表明，基于相关性分析的预测模型在三个月的实际生产中表现良好。

关键词：农产品、卷烟卷制、吸阻监测、相关性分析、机器学习。

1. 引言

美国的“先进制造业领导战略”、德国的“国家工业战略2030”、日本的“机器人战略”、英国的“制造业的未来：英国机遇与挑战的新时代”以及中国的“中国制造2025”等国家政策开启了一个全新的数字化时代。在此背景下，国家烟草专卖局提出建设数字化、智能化等现代化工厂的规划，卷烟制造企业积极响应，投身于企业数字化建设，尤其致力于企业资源计划、制造执行系统、计算机辅助设计/制造/测试和产品数据管理系统的开发与应用[1, 2]，，为后续的智能化建设奠定了良好基础。

目前，中国烟草在生产中配备了大量传感器和质量监测设备，每天采集超过 20T 的时间序列数据，但存在数据挖掘深度和广度不足的问题，导致海量数据的潜在值未得到充分挖掘[3]。尤其是在

吸阻质量控制方面，在线监测仅依赖卷烟检测系统-CIS，成品检验仍处于随机抽样的初级阶段[4]。为了丰富质量监测方法，并减少 CIS 漏检的不合格卷烟数量，需要开展无需额外传感器的吸阻质量改进数据驱动研究。

在国内外，烟草制造业数字化质量控制领域的学术研究较多，而系统化应用较少，因此仍处于发展初期。现有的吸阻质量控制主要采用统计指标[5]，，存在离散度高、时间序列碎片化和预警滞后等问题。传统的香烟质量控制框架尚未充分挖掘数据的应用价值，例如模型适应性差[6]，、缺乏数据质量评估[7]，，以及未对多个指标进行相关性分析[8]。因此，学者们尝试研究卷烟生产过程，并分析大量采集的数据。

在香烟加工过程中，学者们研究了烟叶、切丝烟草和卷烟卷制的质量控制方法。以烟叶为研究对象的研究主要关注原材料的质量[9-11]以及理化表征[12-14]。

对切丝烟草的研究主要集中在理化分析[15, 16]。Condorf[15] 利用图像处理技术研究温度、空气湿度、烟草重量损失与烟草图像之间的映射关系，以实现对烘烤过程中烟草重量的实时监测。Martinez[17] 利用人工神经网络构建温度和湿度的非线性拟合系统，以预测烟叶干燥过程中的湿度。

卷烟卷制是香烟生产中最后也是最关键的过程，大量监测技术被用于评估香烟的最终质量。Yin[18] 提出了一种基于机器视觉技术的新型截面几何参数测量方法，用于测量卷烟滤棒的深度、沟槽数量和面积比。Cao[19] 基于工业相机和视觉信息完成了烟支圆周值的在线测量。

总之，上述方法仅适用于对各工段分别进行质量控制。在实际生产中，传递至后续工段的在制品均符合生产要求，但最终卷制工段仍会出现因质量不达标而被剔除的香烟，这表明卷烟质量控制技术仍有待进一步改进。

卷烟吸阻直接影响消费者口味，吸阻过大会导致吸食困难，吸阻过小则会导致香气不足，因此吸阻是卷烟卷制过程中需要控制的关键质量指标之一。烟草公司的现有质量控制系统由三部分组成：

- 第1部分：基于卷烟检测系统 CIS 的所有香烟在线检测。如图 1 所示，$_1\mu$ 为卷烟吸阻的规定值，X 需要满足 $\in[\ _1 - \sigma_1,\ _1 + \sigma_1]\ X\ u\ n\ n\mu$，否则将被判定为不合格卷烟，并由设备在线剔除。吸阻本质上是

压力值的一种转换。在测量阶段，CIS系统通过一对检测帽夹住卷烟，并喷出恒定的 1.7 kPa 压缩空气。 压缩空气通过卷烟的点火端，到达卷烟滤嘴，检测帽将压力信号转换为电信号，进而计算出吸阻值。

- 第 2 部分：生产过程中，按固定间隔随机抽取极少量香烟进行人工测量。具体而言，质检员每天从卷烟机中随机抽取 20 支香烟，抽取三次，并送至检测室进行人工测量。

- 第 3 部分：从仓库随机抽取卷烟，送至成品检验室进行人工测量。具体而言，质检员从一个生产批次中随机抽取 50 支卷烟（从 1000 件中随机抽取 5 件，每件随机抽取 1 箱，每箱随机抽取 1 支，每箱 10 支卷烟，共计 50 支），并送至检测室人工测量吸阻；如果偏差在允许范围内，则判定为合格。

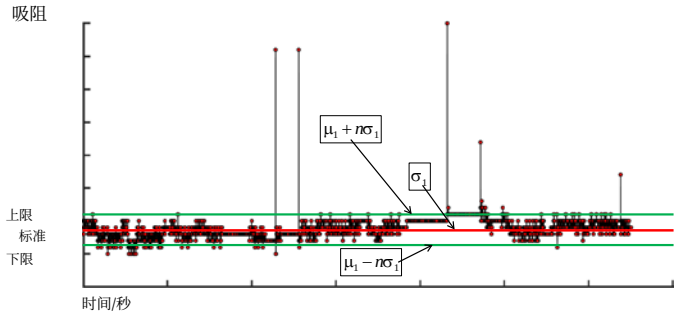


图1 吸阻达到标准的依据

卷烟质量监测的组成部分，检测精度为96%，但仍存在以下问题：

第1部分的CIS在线检测是最重要的，之

- CIS 用于直接测量吸阻，由精密仪器组成。如果发生机械故障、传感器磨损、数据传输不稳定等不确定情况，吸阻检测便失去意义。因此，工作人员必须定期对 CIS 进行繁琐的校准。

- 在线监测方法单一，缺乏相关性建模；此外，监测结果无法用于指导第2和第3部分的采样。因此，无法进一步定位故障时段。第2和第3部分的人工测量方法远不如第1部分重要，且存在以下问题：

- 人力成本。质量检验员需要同时对多台卷烟机进行采样，此外，每天还要往返于车间和检测室。

- 时间成本。在检测室进行人工检测速度较慢，而香烟生产速度很快，因此检测结果无法及时反馈给生产环节。
- 硬件的成本。检测室需要承担维修费用，并

检测仪器的更换。

- 检测结果的可靠性较差。由于质检员的专业素养、技能水平和环境等因素存在不确定性，不同质检员检测同一支卷烟时，可能会得到不同的结果。

总之，普遍认为，在第2—3部分进行随机抽样对吸阻监测的贡献较小[20]。因此，亟需开发一种能够辅助第1部分并指导第2—3部分的方法，即实现卷烟吸阻的相关性建模。

针对这些问题，一些研究采用回归分析监测吸阻。王[21, 22] 和赵[23] 通过线性网络构建了吸阻预测模型，误差小于10%，但仅考虑了线性因素，且测量这些因素耗时较长。为实现在线监测，常[24] 和陈[25] 完成了吸阻相关性建模，预测准确率达到96%，其中考虑的因素较少，且这些因素均为实时测量。然而，他们仅考虑了影响吸阻的线性相关性因素，如填充值、含水率、卷烟重量、周长等。此外，上述研究未对原始数据进行清理，也未能实现真正的在线实时预测。为解决这些问题，Pang[26]基于多传感器融合实现了卷烟吸阻在线预测，预测准确率为87%，但较大的时间粒度（1分钟）导致无法实现异常定位。Tang [27] 采用循环神经网络，以更细的时间粒度（2秒）实现了吸阻在线预测，但由于缺乏专家经验，忽略了传感器温度等重要因素，也未能指导人工抽样（即第2—3部分）。

鉴于上述情况，本文提出了一种在不增加硬件成本的前提下进行卷烟吸阻监测的方法，如图 2 所示，该方法由三部分组成：

- 步骤1——探索卷烟卷制各单元中影响吸阻的因素。对原材料、VE、SE 和 MAX 单元中的工艺指标及控制指标进行预分析，初步研究影响卷烟吸阻的直接或潜在指标。然后，采用斯皮尔曼系数和随机森林评估这些指标与吸阻的相关性，并据此进一步探索相关指标。

- 步骤2——吸阻预测。基于多指标相关性分析预测卷烟吸阻，以识别 CIS 漏检的不合格卷烟。

- 步骤3——验证卷烟监测的有效性。针对给定的生产班次，分析不合格卷烟数量的分布特征，拟合分布曲线，然后构建不合格卷烟被抽样的概率模型，进一步凸显预测模型的实用性。

本文各章节安排如下：第1章讨论当前研究现状，并提出一种改进的卷烟吸阻监测方法；

第2章，分析了各项指标与吸阻之间在三个单元（VE、SE 和 MAX）上的直接或潜在相关性，确定了吸阻的相关因素，以构建吸阻预测模型；第3章，为进一步验证卷烟吸阻预测模型的有效性，推导了抽样不合格卷烟的概率模型；第4章，介绍了实验及其结果，包括数据来源、数据预处理和实验参数设置；第5章，讨论了实验结果和未来工作。

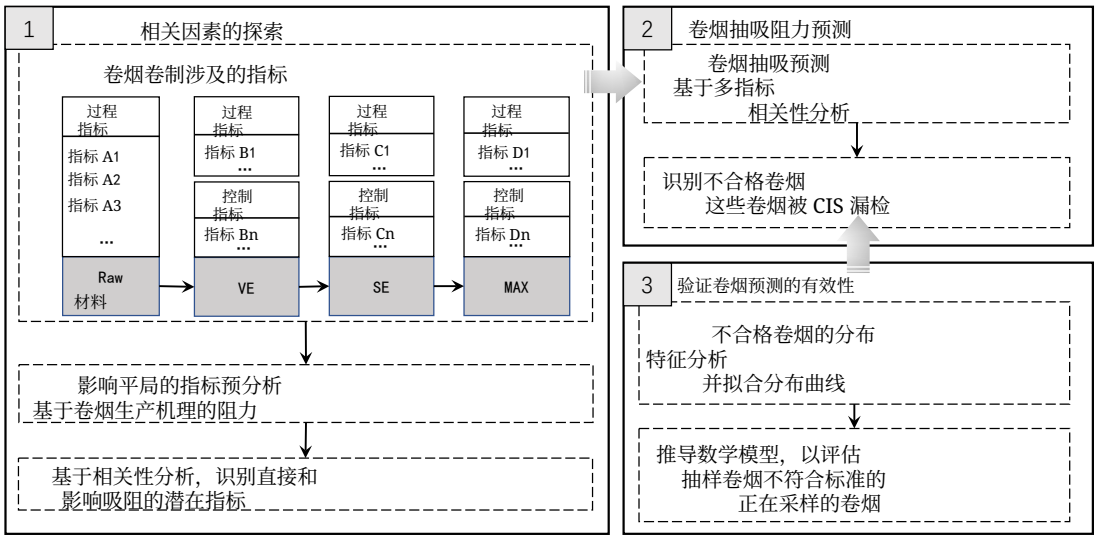


图 2 基于预测的卷烟吸阻监测方法 2. 基于预测的卷烟吸阻监测。本文提出的卷烟吸阻质量监测方法

如图2所示，本文提出了一种方法，包括对相关因素的探索、卷烟抽吸阻力预测，并验证了卷烟预测的有效性。2.1 相关因素的探索

2.1.1 基于卷烟卷制生产机理的影响吸阻指标预分析

卷烟卷制的生产过程如图3所示。卷烟卷制的目的是将切丝烟草加工成完整的香烟。基于对卷烟卷制过程中卷烟成形机理的分析，探究原材料、切丝供给（VE）、卷制烟束（SE）和卷烟成型（MAX）的相关指标，以建立吸阻预测模型，从而丰富吸阻监测方法或辅助CIS。卷烟卷制流程如图4所示。

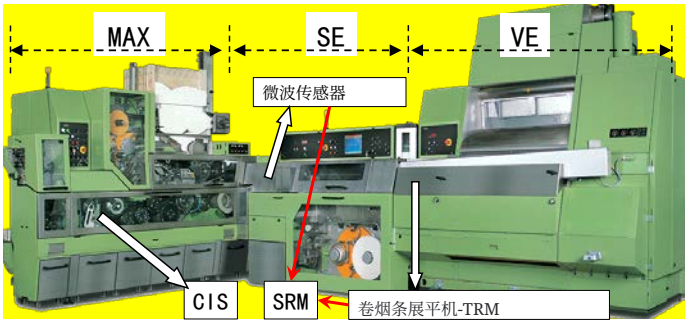


图 3 卷烟机（包括 VE、SE 和 MAX）

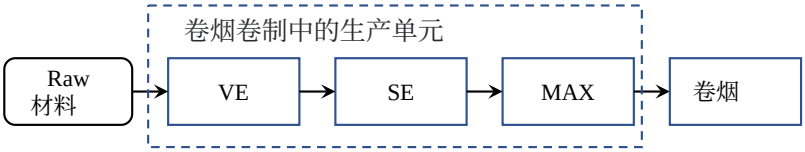
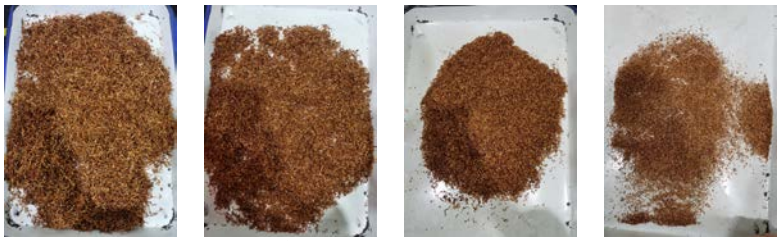


图4 卷烟卷制流程

2.1.1.1 相关因素　原材料

原材料主要是预接丝纸主要关注其温度。切丝烟草的物理指标包括：使用前含水率、填充力、长丝率、泡丝率、短丝率和碎丝率。切丝烟草需要始终保持标准含水率，以获得最佳的烟草填充力[28-30]。填充力用于衡量卷烟内部的填充程度，它直接影响卷烟的重量，并间接影响卷烟的吸阻[31]。长丝率、泡丝率、短丝率和碎丝率直接影响卷烟的硬度和周长。如果卷烟重量保持不变，短丝率增加会导致卷烟硬度和直径下降[32]。烟草丝的不同结构如图5所示。



（a）长丝　(b)短纤维　（c）断裂的灯丝　(d)泡沫丝状物

图5 烟草丝的不同结构

2.1.1.2 因素　参与 VE 的内容

在VE加工单元中，切丝供料器将切丝烟草制成烟束并输送至SE，图6展示了VE中的切丝供料器。

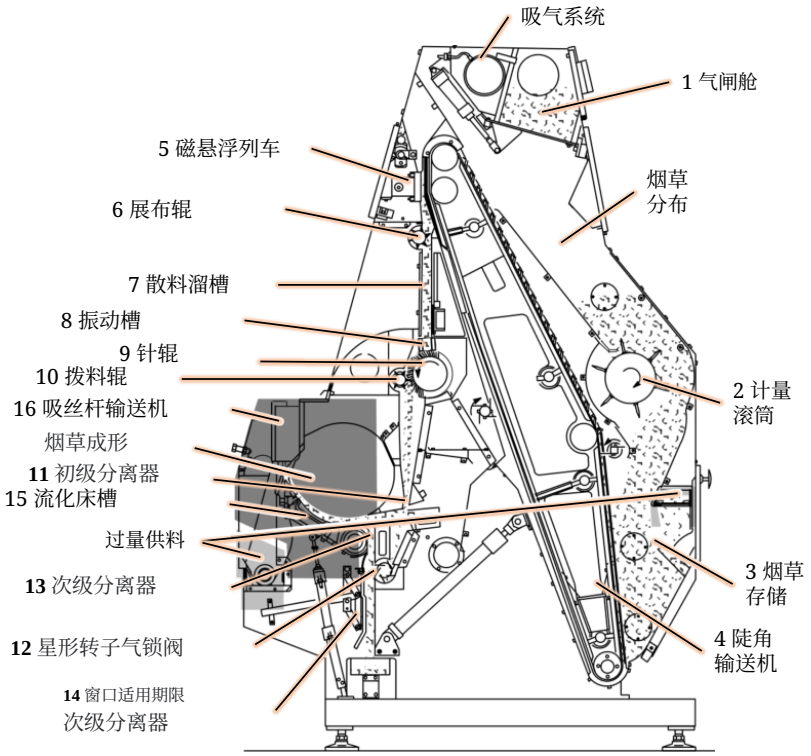


图6 VE工艺中的切丝供料器

在图6中，（1）至（13）后，烟草送入（11）和（13）的气流作用下被送至（16）。最后，（16）上的烟草在气流作用下被送入（17），并吸附在吸风带上，随后进入SE的烟枪。在烟草到达烟枪之前，它会经过修剪盘（图7-E），从烟束上切除多余的烟草，以确保烟草的重量，并间接控制吸阻，这是一种系统反馈控制。

烟草输送依赖于以蒸汽分配箱为核心单元的空气循环系统，耗气单元为(11)、(13) 和 (16)，其中 (11) 和 (13) 的作用是去除烟梗。(15) 的上部压力直接决定吸附烟草的量。烟草通过吸杆输送机真空压力吸附在吸风带上，张力压力过高时，吸风带会被撕裂。(7) 中的传感器用于测量烟草温度，而烟草温度直接影响后续 SE 单元中烟草密度的计算，进而影响烟草重量和吸阻的计算。

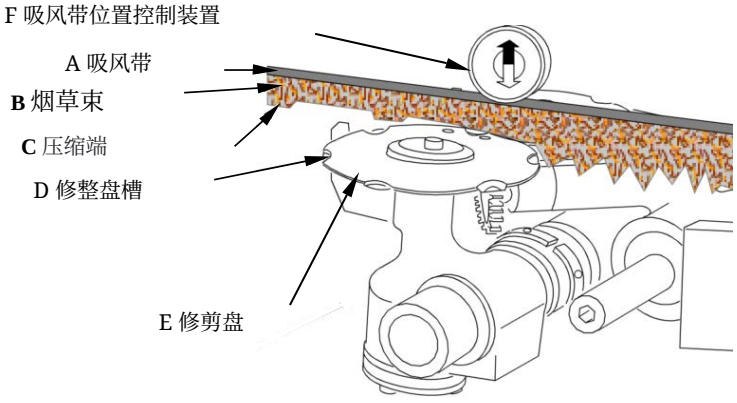


图 7 修剪器（TRM）

修剪器、所示，t.....是卷烟修剪系统（TRM）的主要设备。卷烟重量由（A）的位置决定，而（A）的位置由（F）控制，具体如下：

吸风带位置	烟草量	卷烟重量 吸阻	
Up	增加	增加	增加
Down	减少	减少	减少

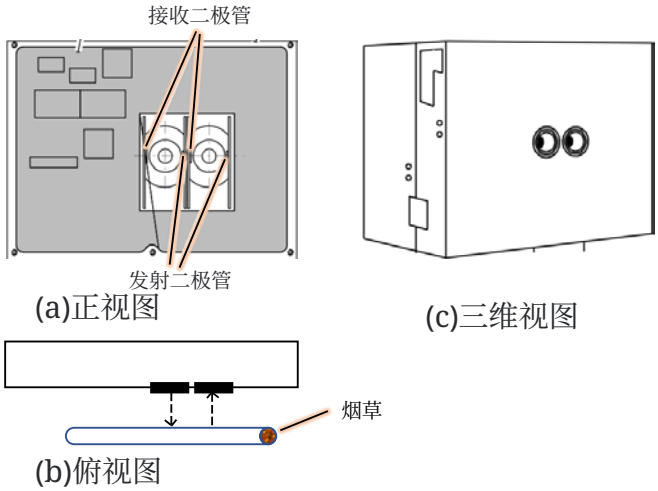
在 E 上设计 D 的目的是确保烟条末端具有较高的烟草密度，以补偿运输过程中烟草的松散和晃动。

总之，直接和间接影响 VE工艺中吸阻的因素包括：空气分配箱压力、烟草储存中的切丝温度、初级分离器压力、流化床槽上部压力、吸丝杆输送机真空压力、吸风带张力压力、吸风带位置、目标重量、卷烟重量以及卷烟压实端的位置。

2.1.1.3 SE 中涉及的因素 SE 处理这些捆绑包

将 VE 单元输送的烟草加工成烟支。烟草用纸包裹并切成条状，测量烟草的重量后，再送入 MAX滤嘴装填机。烟草用纸包裹并切成条状，测量条状烟草的重量后，再送入 MAX。

SE 的核心部分是 SRM，它通过微波（MW）测量头（MIDAS）测量卷烟重量（图 8）。在日常生产中，通过调整卷烟重量来改变卷烟吸阻，因此，与卷烟重量相关的指标也是影响吸阻的潜在因素[31-33]。卷烟圆周与压板的位置有关。



(a)正视图

(c)三维视图

(b)俯视图

图 8 微波-MIDAS

微波-MIDAS 提供 MW-MIDAS 中微波谐振器的电压随烟草密度的变化而变化。

电气特性容易受到温度的影响，因此必须考虑周围硬件（例如微波组件）的温度[34]。

此外，如果烙铁温度过低，卷烟纸粘合处的缝隙会变大，直接影响卷烟密封性。SE皮带通过烟枪带动卷烟纸和烟束，因此必须尽可能绷紧皮带，否则卷烟纸会晃动。

总之，在SE单元中直接或间接影响吸阻的因素包括：烟支圆周、压板位置、超重烟支数量、轻质烟支数量、软烟支数量、硬烟支数量、烙铁温度、电源温度、电路板温度、外壳温度、谐振器温度、二极管温度以及SE皮带张力等。值得注意的是，根据专家经验，电子元件的温度是影响测量精度的潜在因素之一，这也是本文与Pang[26]的不同之处。

2.1.1.4 MAX滤嘴接装机中涉及的因素

MAX（图 9）将滤嘴与 SE 送来的卷烟拼接，组成整支卷烟。此外，CIS 还会测量每支卷烟的一些指标，例如漏气率、通风强度和吸阻。

如图 9 所示，“无滤嘴卷烟供应”从 SE 接收卷烟，“滤棒供应”提供滤棒，此外，“接装纸供应”提供接装纸。随后，在“滤嘴卷烟生产系统”中，卷烟和滤棒均由接装纸包裹，从而形成完整的香烟。后续的 CIS 在线监测卷烟的气密性和密度。在监测卷烟气密性的过程中，恒压气流从卷烟末端流向滤嘴处的压力传感器。MAX 中的压缩空气压力由“进气压力控制”进行控制。

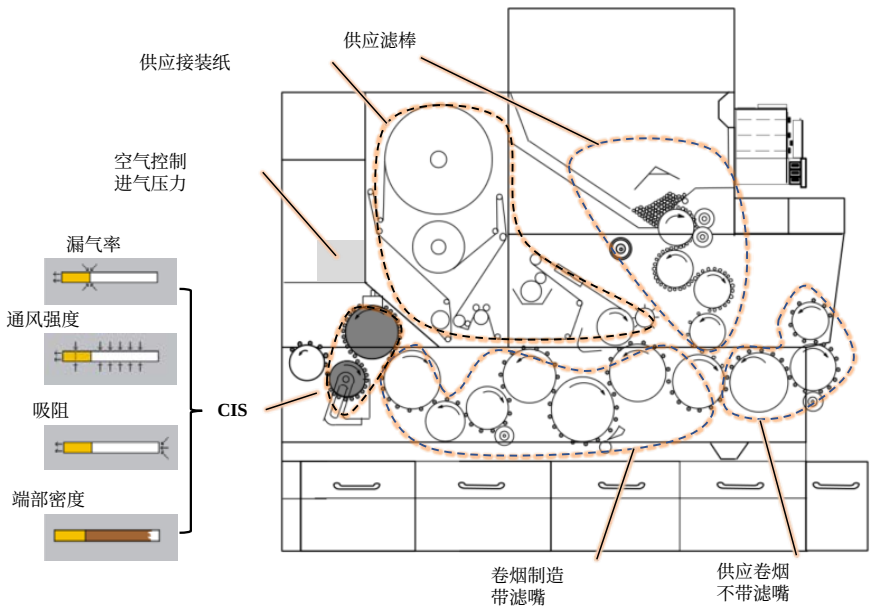


图9 MAX滤嘴接装机

卷烟密封性差的原因：
1. 搓板和接装纸温度的变化。2. 压缩空气变化导致的传感器性能变化。3. 卷烟重量的变化。

综上，直接或间接影响 MAX 单元吸阻的因素包括：搓板温度、进气压力、气源压力、烟支端部密度、卷烟重量、漏气率、通风强度等。2.1.1.5 影响指标的预分析结果

结合 VE、SE 和 MAX 单元的预分析结果，基于卷烟生产机理挖掘卷烟吸阻的影响指标，结果如图10所示。预分析共保留了 82 个指标，包括线性相关性和非线性相关性。由于吸阻与卷烟重量密切相关[24]，，本文还将充分考虑与卷烟重量相关的指标，以预测吸阻。对于与吸阻间接相关的指标，例如与气压和硬件温度相关的指标，也予以考虑。不过，MAX 中的气源压力和进气压力几乎恒定，因此不用于吸阻预测。

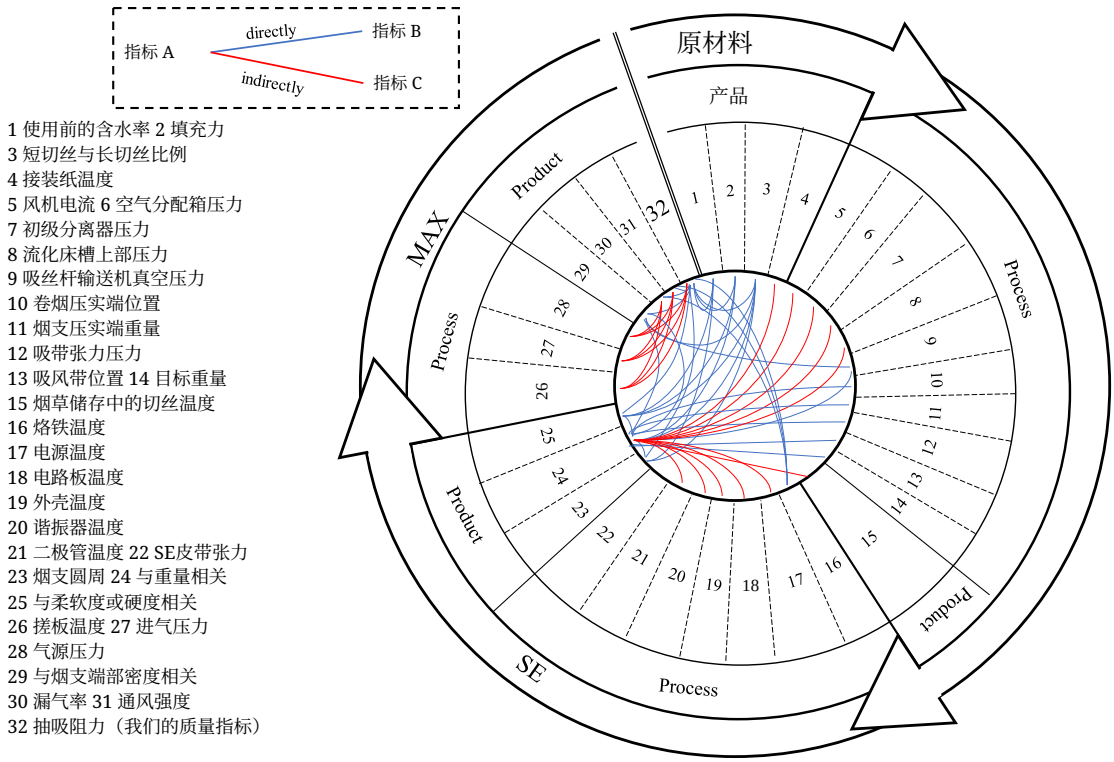


图 10 影响指标预分析结果 2.1.2 基于相关性分析挖掘影响因素

卷烟卷制生产涉及的指标规模较大

如图 10 所示，若直接将其用于预测模型，将无法准确挖掘出最优特征。相关性分析能够进行特征筛选，避免高维数据中的不重要特征。当前的卷烟吸阻相关性分析仅使用皮尔逊系数等简单的线性相关方法评估相关程度，缺乏对.....的考量。

非线性潜在指标。目前，在卷烟吸阻相关性分析中，仅使用皮尔逊系数等简单线性相关方法来评估相关性，缺乏对非线性潜在指标的考量。基于预分析（第2.1.1节），采用斯皮尔曼系数和随机森林，分别计算影响吸阻的直接相关性（线性相关）和潜在相关性（非线性相关）。

基于斯皮尔曼系数评估各指标的线性相关性。将各指标的重要性取绝对值，结果如图11所示。随机森林（RF）的参数设置为：树的数量为20，叶片数量为5，如图12所示。基于随机森林（RF）评估各指标的非线性相关性，结果如图13所示。

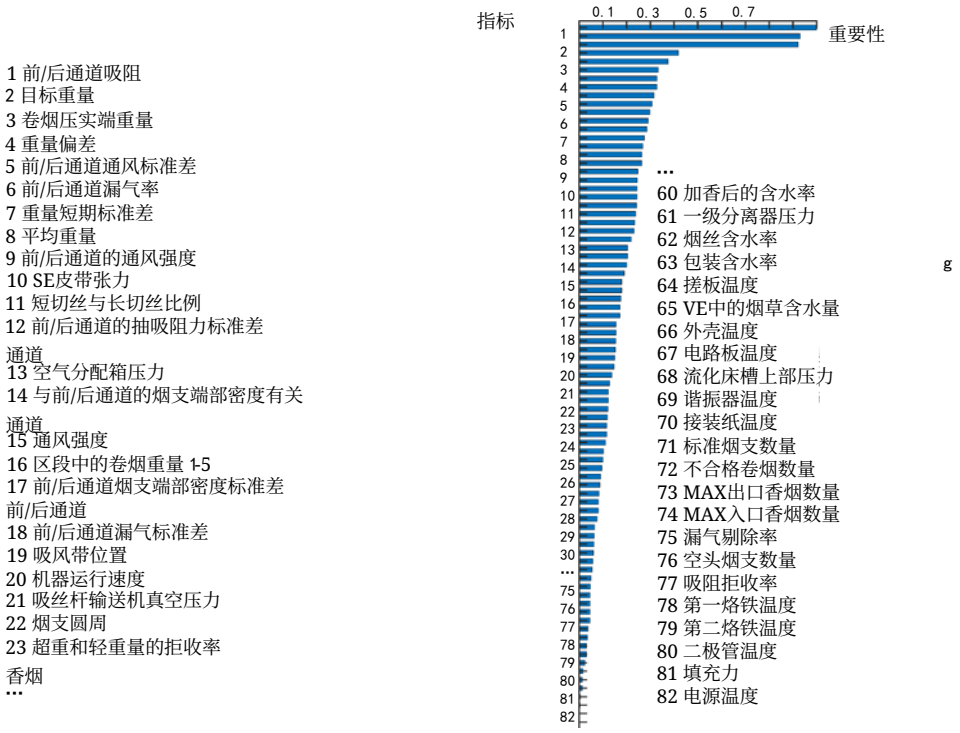


图11 基于斯皮尔曼系数的直接影响因素（线性相关）

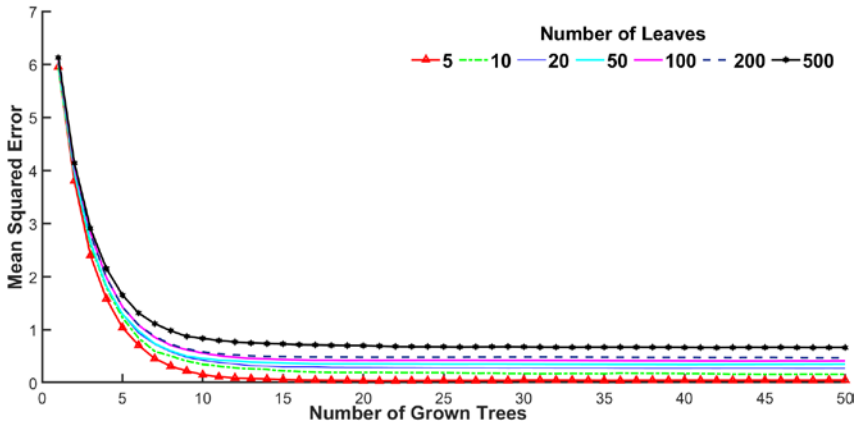


图12 确定随机森林的最优参数（= 20, = 5 树叶节点）

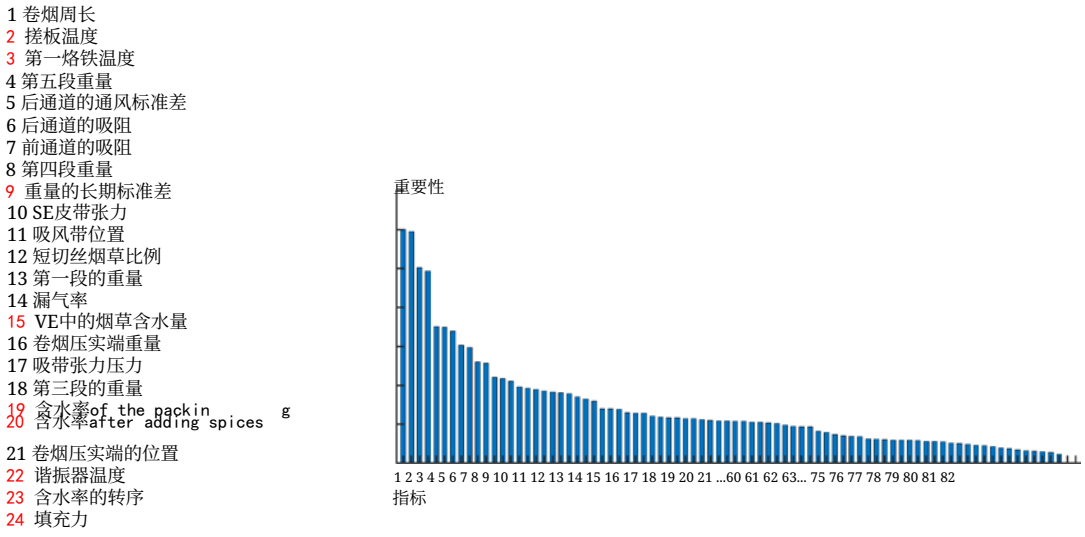


图13 基于随机森林的间接影响因素（非线性相关）因素选择的关键是识别那些影响吸阻的因素

性。受计算能力或模型拟合能力的限制，传统方法仅依赖专家经验来选择线性相关性较强的因素，例如 [21], [23], [24], [25], ，且只考虑通风强度、填充值、含水率和卷烟重量。此外，实验所用样本较少，导致无法实现在线预测，预测准确率难以提高，模型泛化能力较差。卷烟卷制过程的质量监控涉及约200个因素，每天产生2T数据，需要充分利用这些数据进行相关性建模，以挖掘数据价值。随着深度学习的发展和计算能力的提升，人们开始使用拟合能力更强的模型进行吸阻预测，例如 [26],[27], ，但由于缺乏对生产过程的系统分析，温度被忽略，最终限制了吸阻预测的准确率。如图11所示，本文将线性相关性指标中排名前30%（25个因素）的指标视为直接影响因素。如图13所示，在非线性相关性分析中，部分潜在指标的重要性显著提高，因此本文采用非线性相关性中排名前24%（14个因素）的指标（剔除直接影响因素后）。最终用于预测卷烟吸阻的指标如表5所示。

2.2 基于多指标的吸阻预测模型

基于对卷烟吸阻相关因素的分析，构建了吸阻预测模型（图13）。

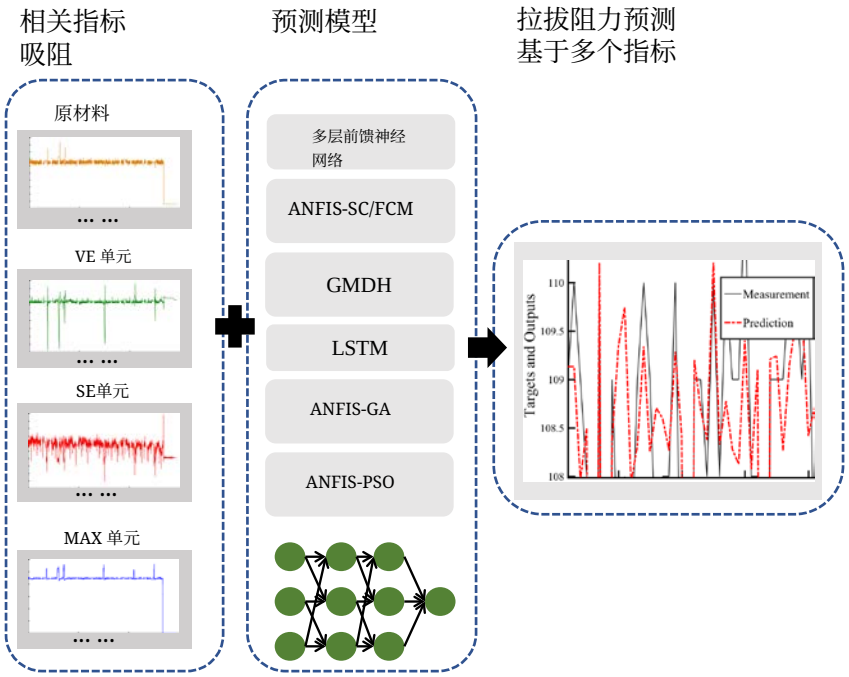


图13 基于多个指标的吸阻预测 2.2.1 多层前馈神经网络-多层前馈神经网络 多层前馈神经网络是最常用的

人工神经网络，包括感知机网络、BP网络和RBF网络。 本文所使用的BP网络是反向传播学习算法在多层前馈神经网络（MLFFNN）中的典型应用[35]，，并采用多种训练函数进行对比分析，包括梯度下降训练函数（traingd）、贝叶斯正则化训练函数（traingbr）、Levenberg-Marquardt算法（trainlm）、Plwell-Beale算法（traincgb）和共轭梯度算法（trainscg）。2.2.2 自适应神经模糊推理系统-ANFIS

自适应神经模糊系统（ANFIS）结合了人工神经网络（ANN）的学习算法和模糊推理（FIS）的简洁形式，通过误差反向传播和最小二乘法调整参数，可将模型的训练过程简化为参数调整过程。ANFIS 存在两种推理方法：Mamdani 和 Sugeno，本文采用的 Sugeno 方法更适用于优化和自适应。ANFIS 包含三个主要部分：模糊化、模糊推理和解模糊[36]。图14展示了一个具有两个输入和一个输出的简单 ANFIS 模型，其中使用了两条模糊规则，如公式（1）和（2）所示。

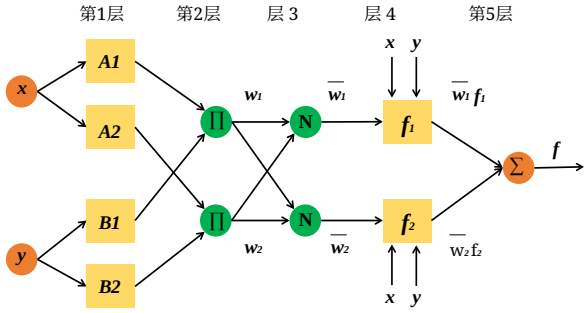


图14 ANFIS模型的简单结构

规则1：如果x是 A1 且y 是 B1，则$_{1\ 1\ 1}$ f p x q y r= + + (1)

规则2：如果 x是 A2 且 y 是 B2，那么$_{2\ 2\ 2}$ f p x q y r= + + (2)

如图14所示，ANFIS 的结构共有五层，其中第1层和第2层的方形节点需要进行参数学习，这是模型的核心部分。各层的功能如下所述：

第1层（模糊化层）：该层是模糊化层，$_i$ $_iA$ B、 为模糊集，用于对输入变量进行模糊化，并将模糊集的隶属度输出至第2层。该层的输出（^{1} $_iO$）表示x和y 分别属于$_iA$ 和$_iB$ 的程度，如公式（3）和（4）所示。$_{iA}\mu$ 和$_{iB}\mu$ 为高斯函数，如公式（5）和（6）所示。

$$O_i^1 = \mu_{A_i}(x) \quad i = 1, 2 \qquad (3)$$

$$O_i^1 = \mu_{B_i}(y) \quad i = 1, 2 \qquad (4)$$

$$\mu_{A_i} = \exp\left\{-\left[\left(\frac{x - c_i}{a_i}\right)^2\right]^{b_i}\right\} \qquad (5)$$

$$\mu_{B_i} = \exp\left\{-\left[\left(\frac{y - c_i}{a_i}\right)^2\right]^{b_i}\right\} \qquad (6)$$

where x y 变量 and $\{\ \}$ are input 、 为前件参数，需要进行训练。

第2层（规则层）：该层中的节点负责对输入信号进行乘法运算，每个节点输入的隶属度表示一条模糊规则的触发强度。

$$^i\ _{iAB}\ \mathrm{O}\ \mathrm{w}\ \mathrm{x}\ \mathrm{y}\ \mu\ (\ =\ )\ =2\ \ \mathrm{i}\ \mathrm{i}\ \bullet\ \ =1,2\mathrm{i} \qquad (7)$$

第3层（归一化层）：该层对第2层的输出进行归一化，并描述规则触发在整个规则库中的比例（以概率表示）。

$$O_3^i = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad i = 1, 2 \qquad (8)$$

第4层（解模糊化层）用于计算每条规则的输出。

$$O_4^i = \bar{w}_i f = \bar{w}_i (p_i x + q_i y + r_i) \quad i = 1, 2 \qquad (10)$$

其中 $\{\ \}$ $_i$ $_i$ $_i$ p q r、 、 为后件参数，需要进行训练。

层 5（求和层）：计算所有输入信号的总输出。

$$O_5^i = \sum \bar{w}_i f_i = \sum w_i f_i / \sum w_i \quad i = 1, 2 \qquad (11)$$

通过粒子群优化（PSO）[37]，和遗传算法（GA）[38]实现。ANFIS模型的训练过程是通过寻找最优前件参数（式（5）—（6））和后件参数（见式（10））来降低网络误差。粒子群优化和遗传算法有助于找到上述参数的全局最优解[39]。

ANFIS model can be impro

2.2.3 长短期记忆网络-长短期记忆网络

长短期记忆网络属于一种用于时间序列预测的循环神经网络，能够处理多个输入和多个输出。长短期记忆网络

在交通流控制[40], 气候预测[41, 42], 运动姿态识别 [43], 和视频处理[44]等具有显著时空特性的应用中表现出色。

2.2.4 数据处理组合法-GMDH GMDH 会自动找出各项指标之间的相关性

同一层中的神经元两两组合，生成中间模型（内部准则）。随后根据外部准则对中间模型进行筛选，满足条件的模型成为下一层的神经元，并重复上述过程，直到获得网络的最佳复杂度[45]。图15展示了 GMDH 的简单结构。

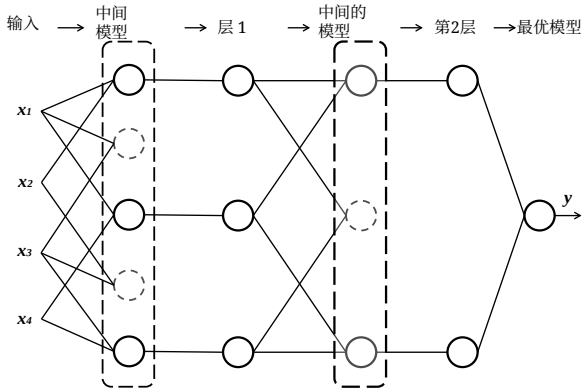


图15 GMDH模型的简单结构

3. 构建一个专业模型 概率模型

为进一步突出预测模型的实用性，本文创新性地构建了一个概率模型。基于预测结果，分析不合格卷烟的分布特征，确定不合格卷烟集中的时段，并推导出不合格卷烟被抽样的概率模型。

如果显著提高不合格卷烟被抽样的概率，则进一步验证了预测模型的实用性。（值得注意的是，该方法具有很强的通用性，对卷烟质量监测的其他领域，如溯源和重点抽样，也具有重要的启发和指导意义。）

有许多方法可用于评估预测模型的实际值，本文以抽检到不合格卷烟的概率P为例:

$$P=1-\frac{C_m^0 * C_{Y-m}^Z}{C_Y^Z}=\frac{A_{Y-m}^Z}{A_Y^Z} \quad (12)$$

其中，P为被采样的概率。Y为参与采样的香烟总数，Z为采样的香烟数量。如果Y满足，$Y\ m\ Z\gg$：

$$P\approx 1-\left(\frac{Y-m}{Y}\right)^Z \quad (13)$$

现有的采样方法需要了解分布的特征，因此采用随机抽样，默认分布为随机均匀分布:

Y：一个工作班次内生产的香烟总数，设定为 $_1Y$ ；

$$Y_1 = \sum_{id=1}^{j_1} y_{id} \quad (14)$$

其中，id 是班次中某一时段的序号。y_{id} 是序号为 id 的时段内生产的香烟数量。j_1 是班次中包含的时段数量。每个时段的长度相同，每个时段内生产的香烟数量为 y_s。

$$Y_1 = j_1 * y_s \quad (15)$$

Z：由已溶解的总数量确定，是一个固定常数；m：班次期间所有不合格卷烟的数量设为 m_1。

总之，抽检到不合格卷烟的概率为：

$$P_{old} = 1 - \frac{C_{m_1}^0 * C_{Y_1 - m_1}^{Z}}{C_{Y_1}^{Z_1}} = 1 - \frac{A_{Y_1 - m_1}^{Z_1}}{A_{Y_1}^{Z_1}} \approx 1 - \left(\frac{Y_1 - m_1}{Y_1}\right)^{Z_1} = 1 - \left(\frac{j_1 * y_s - m_1}{j_1 * y_s}\right)^{Z_1} \quad (16)$$

不合格卷烟的基本分布特征可以直接获得。潜在的分布特征包括均匀分布、正态分布和 t 分布等。本文仅考虑正态分布。如图 16 所示，横坐标为各时段的序号（无严格的时间顺序），纵坐标为不合格卷烟的数量。不合格卷烟的数量拟合为一条正态曲线，且所有时段的序号均满足 $ID \sim N(\mu, \sigma^2)$。标准差为 σ，均值 μ 为不合格卷烟数量最多的时段的序号。

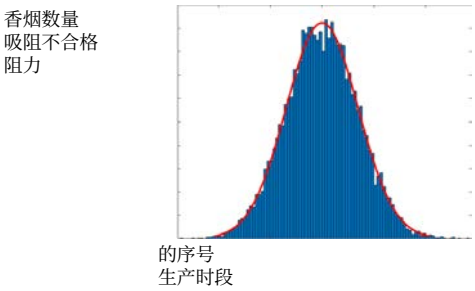


图 16 不合格卷烟的分布特征

总之，结论是，公式 (12) 和 (13) 中的 Y 不再将班次内的所有香烟都纳入考虑，而是仅考虑满足 $ID \in (n\mu - \sigma, n\mu + \sigma)$ 的时段内的香烟。因此，公式 (12) 和 (13) 优化如下：Y：满足 $ID \in (n\mu - \sigma, n\mu + \sigma)$ 的时段内生产的香烟总数，并设为 Y_2；

$$Y_2 = \sum_{id=1}^{ID} y_{id} \quad (17)$$

其中，id 是满足 $ID \in (n\mu - \sigma, n\mu + \sigma)$ 的序号 ID 的子集，所包含的时段数为 j_2。

$$j_2 = 2n\sigma, \exists n \in N^* \Rightarrow j_2 < j_1 \quad (18)$$

Z：由总数决定，是一个固定常数，$Z = Z_1$

m：时段X内生产的不合格卷烟数量，满足) (,X n $n\mu\in$ $-\sigma$ $+\sigma\mu$。

$$m_2 = \Phi(n) * m_1 \quad (19)$$

$$\Phi(n) = P(\mu - n\sigma < \mu < \mu + n\sigma) \quad (20)$$

$$P(a < x \le b) = \int_a^b \frac{1}{\sqrt{2\pi} * \sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad (21)$$

总之，本文中不合格卷烟被抽样的优化概率新P 如下:

$$P_{new} = 1 - \frac{C_{m_2}^0 * C_{Y_2-m_2}^Z}{C_{Y_2}^{Z_2}} = 1 - \frac{A_{Y_2-m_2}^{Z_2}}{A_{Y_2}^{Z_2}} \approx 1 - \left(\frac{Y_2 - m_2}{Y_2}\right)^{Z_2} = 1 - \left(\frac{j_2 * y_s - m_2}{j_2 * y_s}\right)^{Z_2} \quad (22)$$

The 专业bability of substandard cigarettes being sampled is raised by :

$$\Delta P = P_{new} - P_{old} > 0 \quad (23)$$

根据公式（16）、（19）和 (20) and (22), 可得Eq (23) is

$$\Delta P \approx \left(\frac{j_2 * y_s - m_2}{j_2 * y_s}\right)^{Z_2} - \left(\frac{j_1 * y_s - m_1}{j_1 * y_s}\right)^{Z_1} = \left(\frac{2n\sigma * y_s - \Phi(x) m_1}{2n\sigma * y_s}\right)^{Z_1} - \left(\frac{j_1 * y_s - m_1}{j_1 * y_s}\right)^{Z_1} \quad (24)$$

4. 实验 ment and discussion

首先介绍原始数据的来源，其次对数据进行预处理，最后训练预测模型并验证其准确率。

4.1 数据来源

在 VE、SE 和 MAX处理单元中，采集指标数量超过 100 个，其中包括工艺指标、产品指标等。每个指标的存储频率为 2s。为降低实际存储过程中的存储成本，只有当指标的值发生变化时才进行存储。因此，不同指标的存储时刻各不相同，如表1所示。

表1 数据库中的原始数据

指标 A		指标 B	
时间戳	值	时间戳	值
2022年2月7日 10:51:30.000	7002	2022年2月7日 10:51:30.000	58
2022年2月7日 10:51:32.000	8209	2022年2月7日 10:52:06.000	59
2022年2月7日 10:51:36.000	9490	2022年2月7日 10:52:16.000	57
2022年2月7日 10:51:44.000	9492	2022年2月7日 10:52:18.000	60
2022年2月7日 10:51:46.000	9498	2022年2月7日 10:53:44.000	63
2022年2月7日 10:51:54.000	9496	2022年2月7日 10:53:46.000	62
…	…	…	…
2022年/04/29 14:14:02.000	9496	2022年/04/29 15:28:32.000	63
2022年/04/29 15:30:34.000	9798	2022年/04/29 15:30:34.000	61

2022年2月7日至4月29日，共收集了一百多项指标，采集点分布在 VE、SE 和 MAX 生产单元。本论文重点关注与卷烟吸阻相关的指标（见第 2.1 节）。

4.2 数据预处理

由于烟草公司数据库中的存储规则不同，各指标的数据量并不相同（见表1），因此首先对数据进行填充，以确保不同指标的数据量相同，且在同一时间点采集。填充后，每个指标包含3554194个值，整个样本量为 [3554194 × 个指标]，。第一个值记录于“2022年2月7日 10:51:30”，最后一个值记录于“2022年4月29日 15:30:34”。填充工作如表2所示。

表2 数据填充工作

指标 A- 原始数据		指标 A- 填充数据	
时间戳	值	时间戳	值
2022年2月7日 10:51:30.000	7002	2022年2月7日 10:51:30.000	7002
2022年2月7日 10:51:32.000	8209	2022年2月7日 10:51:32.000	8209
2022年2月7日 10:51:36.000	9492	2022年2月7日 10:51:34.000	8209
2022年2月7日 10:51:44.000	9494	2022年2月7日 10:51:36.000	9492
2022年2月7日 10:51:46.000	9498	2022年2月7日 10:51:38.000	9492
2022年2月7日 10:51:54.000	9498	2022年2月7日 10:51:40.000	9492
2022年2月7日 10:51:56.000	9497	2022年2月7日 10:51:42.000	9492
2022年2月7日 10:52:04.000	9497	2022年2月7日 10:51:44.000	9494
…	…	…	…

在数据预处理中，我们发现原始数据存在一个严重问题：原始数据可能有误。例如，记录每班累计工作小时数的指标.....

工人在一个班次内累计工作的时长。每天安排两个生产班次，早班为“上午7:00至下午3:30”，中班为“下午3:30至晚上11:30”，且值仅在工作时间内不为零。然而，原始数据显示有人在“03:25:46”开始工作，这显然是不符合实际的错误数据，如表3所示。

此外，在数据预处理阶段，其他指标也存在类似错误 （例如指标“每班生产的香烟总数”：理论上，卷烟生产率由机器性能决定，且相对稳定，但记录的原始数据却显示香烟数量突然增加）。

还有许多其他类似的数据质量问题，本文未一一列出。因此，“烟支吸阻”指标也可能存在数据质量低等问题。数据质量低的原因可能是通信设备或 CIS 等硬件不稳定，因此，基于多指标预测吸阻值，对于丰富吸阻监测方法具有重要意义。

表3 原始数据-指标“每班累计工作小时数”

每班累计工作小时数	
时间戳	值/秒
2022年/02/10 23:35:50.000	23448
2022年/02/11 03:25:40.000	23451
2022年/02/11 03:25:42.000	23453

2022/02/11 03:25:46.000	6
2022/02/11 03:25:48.000	8
2022/02/11 03:25:50.000	10
2022/02/11 03:25:52.000	12
2022/02/11 03:25:54.000	14

预处理的最后一步是从填充数据中获取实验样本。除工作日早上 7:30 至晚上 11:30 外，其余时间均不生产，因此需要从填充数据中截取生产时段。“每班生产的香烟总数”指标记录了一个班次生产的香烟总数（机器每 60 秒约生产 10,000 支香烟）。香烟总数会突然从数万变为零，这表示工作人员下班了。而在两秒内从零增加到 2,000 以内，则表示工作人员正在上班。

通过设置约束条件（例如，将香烟数量跳变数万的错误数据设为0）及其他处理，填充数据最终被划分为100个班次（见表4）。 表4中的通勤时间段用于截取其余指标的数据，作为实验数据。 值得一提的是，工人有提前开始或结束工作的习惯，而我们的方法划分出的工作时间即为实际工作时间。

表4 准确的通勤时间

上班	下班
'2022-02-08 06:41:16'	'2022-02-08 15:26:10'
'2022-02-08 15:26:50'	'2022年-02-08 23:36:10'
'2022年-02-09 06:44:36'	'2022年-02-09 15:26:10'
'2022年-02-09 15:26:16'	'2022年-02-09 23:30:10'
.....	.....
'2022年-04-25 06:45:34'	'2022年-04-25 15:27:50'
'2022年-04-25 15:27:58'	'2022年-04-25 23:40:44'
'2022年-04-26 06:45:50'	'2022年-04-26 08:47:28'
'2022-04-27 06:46:12'	2022年-04-27 15:27:48
'2022-04-27 15:30:00'	2022-04-27 23:37:46
2022-04-28 06:44:24	'2022-04-28 14:06:10'

4.3 实验数据 以制品为例

以一个班次的生产数据为例，使用滑动窗口，窗口大小设为1，以尽可能多地包含原始特征信息，单个样本由相关指标和卷烟吸阻组成。 相关指标包括$_1$ n 个原材料指标、$_2n$ 个VE指标、$_3n$ 个SE指标、$_4n$ 个MAX指标，以及当前时刻的1个卷烟抽吸阻力值。训练样本的形式为[1, $_1$ 1, $_2$ 1, $_3$ 1, $_4$ …, , …, , …, , … ,] A An B Bn C Cn D Dn Y ，样本结构详情见图 14。

本实验的样本涵盖3个月内的100个班次，其中85%的数据用于训练和验证，15%作为测试集。实验中使用的指标如表5所示。

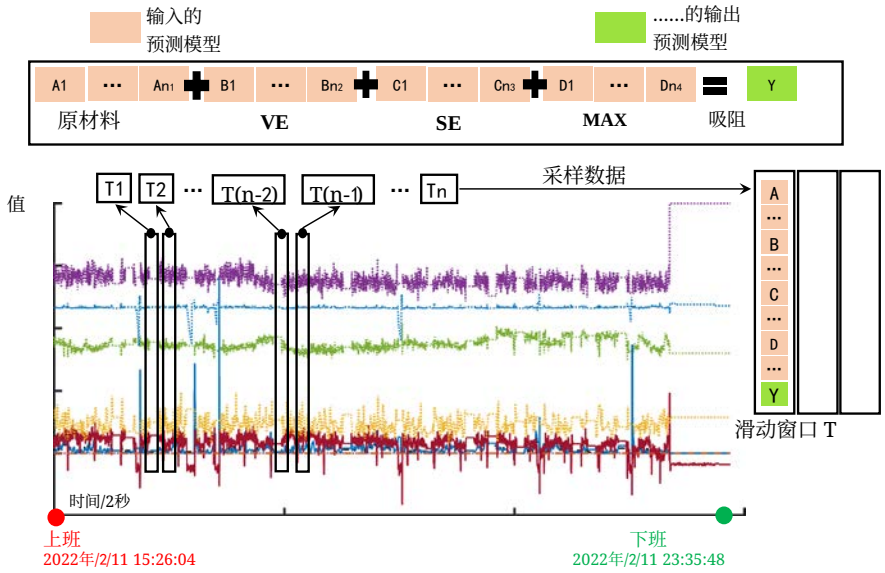


图14 样本数据

4.4 实验参数

MLFFNN模型是一种三层前馈神经网络，包含一个具有40个节点的隐藏层。考虑了五种训练函数（traingd、trainbr、trainlm、traincgb、trainscg），并选出性能最佳的训练函数。

对于 GMDH模型，网络层数 =5*L* 、每层神经元的最大数量 =50*N* 以及选择压力 =0.6*P* 。

对于深度学习模型长短期记忆网络，隐藏层的层数 =1*L* 、隐藏层维度数 =128*N* 、dropoutLayer 参数为 0.2，最大迭代次数为 100。

对于 ANFIS模型，采用两种方法（SC、FCM）构建模型。ANFIS(SC) 的核心参数为影响半径（ *_I R*），取值为 0.5 或 0.6；ANFIS(FCM) 的核心参数为聚类数量（ *_NC*）和划分矩阵指数（ *PME*），二者均取值为 2。

对于 ANFIS-GA，使用遗传算法优化 ANFIS模型（此处的 ANFIS 指 ANFIS-FCM），聚类数量为 2。此外，遗传算法的最大迭代次数、交叉概率、变异概率和变异率分别为 2000、0.3、0.6 和 0.1。种群数量（*N* ）考虑取 10/20/25/50。

对于 ANFIS-PSO，ANFIS 的聚类数为 2，PSO 参数设置如下：种群数量 *N* 分别为 10、20、25 和 50，惯性权重为 0.9，个体学习系数为 0.9，全局学习系数为 1.8，最大迭代次数为 2000。

实验统计指标包括均方误差 RMSE、均方根误差 MSE 和相关系数 R。本实验重点关注测试集的相关系数 R。

4.5 实验结果

实验结果如表6所示。仅考虑潜在/非线性相关指标的影响时，测试集相关系数仅在0.5左右，其中GMDH的表现相对较好（=10.4759均方误差 ,=3.5360 均方根误差 ,=0.5540*R* ）。

仅考虑线性指标的影响，ANFIS(FCM)、GMDH、长短期记忆网络、ANFIS-GA 和 ANFIS-PSO 模型的测试集相关系数均大于 0.9，其中 ANFIS-GA 的表现最佳。

将线性和非线性指标结合起来预测卷烟吸阻，所有模型在测试集上的相关系数均大于 0.94，其中 ANFIS-GA 的性能最佳（=1.4357均方误差，=1.1982均方根误差，=0.9684相关系数）

如表6所示，尽管仅考虑直接指标时模型也能取得良好表现，但同时考虑直接指标和潜在指标时，模型表现最佳。

如图15所示，吸阻骤降的情况可以通过预测模型准确预测。相关性分析的结果有助于CIS进行抽吸阻力监测，并识别出CIS漏检的不合格卷烟。本文以不合格卷烟被抽样的概率为例，进一步凸显该预测方法的实用价值。

一个工作班次分为 100 个时段，$_{1}=100j$。一台卷烟机每个时段约生产 43200 支卷烟，$_{s}\approx 43200j$。每个班次采样 200 支卷烟，$_{1}\approx 200Z$。根据预测结果，不合格卷烟的数量为 1055 支，$_{1}=1055m$。结合公式 (16)，$_{old}=0\ 047.P$ 。

为使采样涵盖大多数不合格卷烟，本文将99%的不合格卷烟明确定位到具体的生产时段内，$_{2}=0\ 99\ _{1}$。m m

根据预测结果，不合格卷烟的数量为 $_{1}=1055m$，随后拟合正态曲线，标准差为 $\sigma=1\ 16$。，因此 $_{2}\approx 6\ j$。每个班次采样 200 支香烟，$_{2}=200Z$。结合公式 (22)-(24)，$_{new}=0\ 558.P$，抽到不合格卷烟的概率提高了 $0.511P\Delta$ $\approx$。在实际生产中，随机选取了五个班次进行验证。由于传统方法不具备不合格卷烟分布的特征，因此无法定位不合格卷烟。然而，本文提出的基于相关性分析的预测方法能够定位不合格卷烟，并提高抽到不合格卷烟的概率。 如表 7 所示，基于本文的方法，可以找到 2022-03-14 中班、2022-03-29 中班和 2022-03-31 中班的不合格卷烟，这表明基于相关性分析的吸阻预测方法表现良好，且具有较强的实用价值，因为它能够定位不合格卷烟的生产时段。

表5 实验中使用的指标

原材料	VE 单元	SE单元	MAX 单元
直接/线性相关			
长丝率	吸杆输送机真空压力	烟支圆周	通风强度
泡丝率	吸风带位置	平均重量	前通道通风强度
短丝率	烟支压实端重量	重量偏差	前通道通风标准差
潜在/非线性相关性	空气分配箱压力	重量短期标准差	漏气率
接装纸温度	吸带张力压力	第1段烟支重量	前通道漏风
包装含水率	目标重量	第5区段中的卷烟重量	前通道漏气标准差
填充力	潜在/非线性相关性	SE皮带张力	与前通道的卷烟端部密度相关
	VE中的烟草含水量	机器运行速度	与烟支端部密度标准差相关的
	流化床槽上部压力	潜在/非线性相关性	前通道
	一级分离器压力	第一烙铁温度	潜在/非线性相关性
	烟草储存中的烟草温度	第二烙铁温度	搓板温度
		重量的长期标准差	
		外壳温度	
		二极管温度	
		谐振器温度	

表6 实验结果

方法	参数	MSE		RMSE		R	
		训练	Test	训练	Test	训练	Test
输入指标	非线性相关性						
多层前馈神经网络	rainbr	16.8604	21.0568	4.1061	4.5888	0.2597	0.234
ANFIS(SC)	I_R=0.5	14.9296	15.1718	3.8639	3.8951	0.4466	0.4521
ANFIS(FCM)	N_C=2,PME=2	14.5184	12.8724	3.8103	3.5878	0.4838	0.5071
GMDH	N=50,L=2,P=0.5	14.7841	10.4759	3.845	3.2366	0.4856	0.554
LSTM	L=1,N=128,丢弃率=0.2	14.5832	12.5036	3.8188	3.536	0.4803	0.5273
ANFIS-GA	N=10	14.7469	10.5974	3.8402	3.2554	0.4876	0.5462
ANFIS-PSO	N=20	14.4006	13.5418	3.7948	3.6799	0.4831	0.5094
输入指标	线性相关						
多层前馈神经网络	Levenberg-Marquardt训练算法	5.5118	7.2611	2.3477	2.6946	0.8394	0.7917
ANFIS(SC)	I_R=0.6	5.4993	8.4394	2.3451	2.9051	0.7748	0.8362
ANFIS(FCM)	N_C=2,PME=2	2.9341	3.7448	1.7129	1.9351	0.9157	0.9121
GMDH	N=50,L=5,P=0.5	2.9119	3.4157	1.7064	1.8482	0.9165	0.9188
LSTM	L=1,N=128,丢弃率=0.2	2.3967	2.4287	1.5481	1.5587	0.9348	0.9265
ANFIS-GA	N=50	2.3696	2.1674	1.5394	1.4722	0.9354	0.9351
ANFIS-PSO	N=50	2.2085	2.5867	1.4861	1.6083	0.9398	0.9237
输入指标	线性和非线性相关性						
多层前馈神经网络	Levenberg-Marquardt训练算法	1.8242	2.0402	1.3507	1.4284	0.9497	0.9455
ANFIS(SC)	I_R=0.6	1.9874	1.929	1.4098	1.3889	0.9447	0.9507
ANFIS(FCM)	N_C=2,PME=2	1.8016	1.8817	1.3422	1.3718	0.9498	0.9531
GMDH	N=50,L=7,P=0.5	1.829	1.7268	1.3524	1.3141	0.9491	0.9564
LSTM	L=1,N=128,丢弃率=0.2	1.8409	1.478	1.3687	1.2157	0.9484	0.9641
ANFIS-GA	N=50	1.6853	1.4357	1.2982	1.1982	0.9614	0.9684
ANFIS-PSO	N=50	1.6605	1.6811	1.2886	1.2966	0.954	0.9571

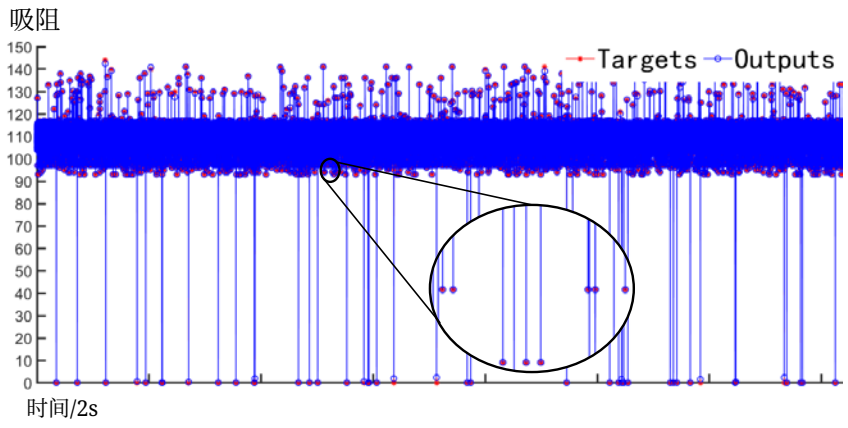


图15 卷烟吸阻预测结果

表 7 进一步凸显预测模型的实用价值

工作班次	是否发现不合格卷烟	
（早班-Mor，中班-Mid）	旧方法	我们的方法
2022-03-14中期	No	Yes
2022-03-29Mor	Yes	Yes
2022-03-29中期	No	Yes
2022-03-31年中	No	Yes
2022-04-13中期	No	No

将本研究的部分内容与本文进行比较，结果如表8所示。在吸阻监测准确率方面，CIS、[25]，和本文均达到了96%的准确率。CIS采用直接测量，而[25] 仅使用少量离线数据进行建模，无法在线预测。在所用因素方面，[26]-[27] 同时考虑了线性因素和非线性因素，但未结合专家经验进行机理分析。例如，[26] 忽略了温度对测量的影响。对于后续采样的指导，本文基于吸阻预测结果构建了概率模型，为采样提供了理论依据，并使异常定位成为可能。 综上所述，本文提出的卷烟吸阻预测方法得益于对原始数据的严格清洗以及对生产过程的全面分析。

表8 本文与其他研究的比较结果

方法	准确率	离线或在线（时间粒度）	**Data** 清理	线性或非线性因素	采样是引导的
CIS	>96%	在线	No	/	/
王[21]	>90%	离线	No	线性(7)	No
赵[23]	>95%	离线	No	线性(10)	No
陈[25]	>96%	在线（尚未实现）	No	线性(9)	No
常[24]	>92%	在线（尚未实现）	No	线性(5)	No
Pang[26]	>95%	在线（1分钟）	Yes	线性(30)+非线性(11)（不足）	No
Tang[27]	>94%	在线（2秒）	Yes	线性(12)+非线性(6)（不足）	No
本文	>96%	在线(2秒)	Yes	线性(25)+非线性(14)	是

5. 讨论与未来工作 在本文中，为了丰富

在不增加硬件成本的情况下监测吸阻的方法，本文提出了基于相关性分析的卷烟抽吸阻力预测方案。该方案包含基于机理层面对影响因素进行预分析（原材料、VE、SE 和 MAX 单元），并进一步基于线性和非线性识别指标。

相关性分析。此外，采用机器学习模型（多层前馈神经网络、ANFIS，35改进的ANFIS、GMDH、长短期记忆网络）实现抽吸阻力预测36（预测准确率提升至96.84%，这得益于额外考虑了37潜在因素）。最后，创新性地推导出不合格卷烟抽样的数学模型，并通过38案例研究进一步验证了预测模型的实用价值。总之，本文提出的基于多指标相关性分析的卷烟抽吸39阻力预测方法可丰富吸阻监测方法，并为质量追溯和异常定位提供理论40支持。414243

在下一步中，我们将继续探索更可靠且耗时更少的44不合格卷烟监测方法，例如基于现有指标构建图数据库，45并使用图神经网络（GNN）进行卷烟质量46监测。47

48

参考文献

49[1] Li X、Liu W、Zhang X. 基于烟草 CIMS 的集成质量体系研究与实现[J]. 《中国机械工程》，2002，13(13)：1125-1125. 51[2] Gang S、Xiaoming W、Junfeng W 等. 基于模糊层次分析法的卷烟企业制造过程信息管理系统分析 [J]. 《智能与模糊系统杂志》，2021，40(4)：8257-8267. 54[3] Losso K、Cardini J、Huber S 等. 利用直接实时分析质谱法和液相色谱-质谱法快速鉴别和质量控制烟草制品[J]. 《塔兰塔》，2022年，238：123057. 57[4] Wang X、Wang X、Wu J 等. 将 PDCA 质量循环与基于模糊分类的信息策略相结合的卷烟质量溯源研究[J]. 《智能与模糊系统杂志》，2021，40(4)：8217-8226. 60[5] Xin L、Haibo S、P F. 面向烟草行业的 MES 产品溯源[C], 第29届中国控制会议论文集，2010年。 62[6] Wu Y-m、Chen L-s、Li S-b 等. 一种用于处理数据流演化和奇异性的自适应算法[J]. 《信息科学》，2021年，545：312-330. 64[7] Chen S-Y、Zhang G. 基于空间分布的边坡位移实测数据质量评价[J]. 《岩土工程学报》，2022年，44(5)：845-850. 67[8] Halim Z、Ali O、Ghufran Khan M. 关于将数据集高效表示为图以挖掘最大频繁项集[J]. 《IEEE知识与数据工程汇刊》，2021年，33(4)：1674-1691. 70[9] Chen Q L、Cai L、Wang H C 等. 烟叶烘烤期间叶际真菌的组成与多样性[J]. 《微生物学前沿》，2020年，11：554051. 72[10] Jianqiang Z、Yan L、Yufeng H 等. 利用手持式数字光投影近红外光谱仪和极限学习机算法表征烟叶深绿感染[J]. 《分析快报》，2020年，53(14)：2266-2277. 75[11] Xiang L G、Wang H C、Wang F 等. 烟叶叶际微生物及潜在病原体分析[J]. 《微生物学前沿》，2022年，13：843389. 77[12] Brudzewski K、Osowski S、Golembiecka A. 差分电子鼻与支持向量机

用于快速识别烟草的[J]。《专家系统与应用》，2012，39(10)：9886-9891。79[13] 陈 J、何 X、张 X 等。不同烟草类型对加热不燃烧烟草制品的适用性 80[J]。《工业作物与产品》，2021，168：113579。81[14] 朱 W、陈 L、王 B 等。基于低能X射线成像的烟叶梗生物质副产品分离过程在线检测 82
83 [ ] 2017, J。《废弃物与生物质增值》，9(8)：1451-1458。84[15] Condorí M、Albesa F、Altobelli F 等。用于监测散叶烘烤炉中烟草烘烤过程的图像处理 85[J]。《农业中的计算机与电子学》，2020，168：105113。86[16] 张 Y、何 Q、曹 Y 等。共热解和共燃烧过程中烟丝与其他烟草基材料之间的相互作用 87[J]。《热分析与量热学杂志》，2018，136(4)：1711- 88 1721。89[17] Martinez-Martinez V、Baladron C、Gomez-Gil J 等。利用人工神经网络估算和预测烟草干燥过程中的温度和相对湿度 90
91 [ ] 2012, J。《传感器（巴塞尔）》，12(10)：14004-14021。92[18] 尹 X、尹 M、谢 L 等。基于机器视觉的复合滤棒截面几何参数自动测量系统 93 [J]。《IEEE仪器与测量汇刊》，2020，69(7)：5037-5050。95[19] 曹 J-L、李 J-F、卢 T-D。基于视觉信息的卷烟圆度测量 96 [J]。《制造业进展》，2017，5(2)：177-181。97[20] 戴 X、吴 C、赵 Z 等。测试台吸阻/压降测量单元的改进 98 [J]。《烟草科技》，2014，47(08)：18-21，DOI：10.3969/j.issn.1002- 99 0861.2014.3908.3004。100[21] 王 L、邓 G、吴 Z 等。基于线性网络模型的卷烟逐口动态吸阻与通风率关系 101[J]。《烟草科技》，2022，55(05)：102 66-72。Doi：10.16135/j.issn11002-10861.12022.10019。103[22] 王 L、YOU M、崔 X 等。基于线性网络模型的卷烟吸阻一致性评价 104 [J]。《烟草科技》，2017，50(12)：85-89。DOI：105 10.16135/j.issn11002-10861.12017.10184。106[23] 赵 W、李 W、张 T 等。基于线性网络模型的卷烟吸阻一致性评价 107[J]。2019，52(12)：86-92。DOI：10.16135/j.issn11002-10861.12019.10069。108[24] 常 Y、石 X、陈 P 等。一种基于物理指标的雪茄吸阻控制方法 109[J]。2019，52(03)：92-96。DOI：10.16135/j.issn11002-10861.12018.10010。110[25] 陈 X、魏 X、孟 X 等。梗丝填充值与卷烟吸阻关系数学模型的建立与应用 111[J]。《烟草科技》，2010，42(08)：18-21。DOI：10.3969/j.issn.1002-0861.2010.3908.3003。113[26] 庞 S、胡 R、郭 B 等。烟草制品关键属性的数据驱动预测 113[C], 载于2021年第三届大数据工程国际会议（BDE 2021），美国纽约州纽约市，115 2021，25-33，DOI：https://doi.org/10.1145/3468920.3468924。计算机协会。116[27] 唐 C。吸阻控制过程中智能优化关键技术研究 117 [D]。中国机械科学研究总院，2021。DOI：118 https://doi.org/10.27161/d.cnki.gshcs.2021.000026。119[28] 李 Q、李 Y F、张 Y 等。“差速式”实验室旋转干燥机中不规则纤维材料的干燥动力学研究：切丝烟草案例 120[J]。《干燥技术》，2017，36(5)：523-536。121[29] 梁 W、刘 Z、金 L 等。案例研究：切丝烟草含水率的统计监测 122

烟草 [J].《工业与生产工程杂志》，2017，34(8)：551-557。123[30] Sakamoto K，Katsuoka T. 在空气或过热蒸汽流中烟草切丝填充床的穿流干燥模型 [J].《食品科学与技术研究》，2012，18(5)：623-629。125[31] Poorolajal J，Assari M J，Mohammadi Y，等. 卷烟滤嘴长度和直径对卷烟烟气排放的影响 [J].《临床流行病学与全球健康》，2020，8(2)：377-382。127[32] McAdam K，Eldridge A，Fearon I M，等. 卷烟圆周对烟气化学、生物活性和吸烟行为的影响 [J].《调节毒理学与药理学》，2016，82：111-126。129[33] 吴 Z，Li L，Li D，等. 卷烟重量、吸阻、通风率与卷烟感官评价之间的相关性分析 [J]. 2010，31(2)：49-53。131[34] Yu K，Tian X，Wang X，等. 通过原位温度补偿提高钯镍合金氢传感器的准确率 [J].《传感器与执行器B：化学》，2019，299：126989。133[35] Zhou S，Liu N，Shen C，等. 一种基于反向传播（BP）神经网络的自适应卡尔曼滤波算法，用于同时检测呼出一氧化碳和一氧化二氮 [J].《光谱学报A：分子与生物分子光谱学》，2019，223：117332。136[36] Petković D，Pavlović N T，Ćojbašić Ž. 基于自适应神经模糊策略的风电场效率 [J].《国际电力与能源系统杂志》，2016，81：215-221。138[37] Poddar S，Kumar A. 用于运行中和现场惯性传感器校准的无标度粒子群优化 [J].《测量》，2019，147：106849。140[38] Qiu G，Gu Y，陈 J. 结合遗传算法和威布尔比例风险模型，对轴承集成剩余使用寿命进行预测的选择性健康指标 [J].《测量》，2020，150：107097。143[39] Khosravi A，Nunes R O，Assad M E H，等. 人工智能方法在日全球太阳辐射估算中的比较 [J].《清洁生产杂志》，2018，194：342-358。145[40] Ma X，Zhong H，Li Y，等. 使用嵌套长短期记忆模型的深度胶囊网络预测交通网络速度 [J].《IEEE智能交通系统汇刊》，2021，22(8)：4813-4824。148[41] Elmaz F，Eyckerman R，Casteels W，等. 用于室内温度预测建模的卷积神经网络-长短期记忆网络架构 [J].《建筑与环境》，2021，206：108327。150[42] Yan R，Liao J，Yang J，等. 使用CNN、长短期记忆网络、卷积神经网络-长短期记忆网络以及时空聚类方法预测北京多小时、多站点空气质量指数

[ ] 2021, J .《专家系统与应用》，169：114513。153[43] Shu X，Zhang L，Sun Y，等. Host-Parasite：用于群体活动识别的图长短期记忆网络嵌套长短期记忆网络 [J].《IEEE神经网络与学习系统汇刊》，2021，32(2)：663-674。155[44] Yang Y Z J，Ai J，等. 基于对抗长短期记忆网络的视频描述生成 [J].《IEEE图像处理汇刊》，2018，27(11)：5600-5611。157[45] Ahmadi M，Ahmadi M-A，Mehrpooya M，等. 使用群体方法数据处理神经网络对斯特林发动机的功率和扭矩进行建模 [J].《可持续发展》，2015，7(2)：2243-2255。159