

大数据在半导体制造过程质量管控中的应用研究

方云鹏 赵山岭^{通讯作者}

润鹏半导体（深圳）有限公司，广东深圳，518000；

摘要：随着半导体制造技术的不断进步，对产品质量的要求日益严苛。大数据技术凭借其强大的数据处理和分析能力，为半导体制造过程的质量管控提供了新的解决方案。本文深入探讨大数据在半导体制造质量管控中的应用，详细阐述其在数据收集与预处理、质量特征提取、模型构建与质量预测、实际应用场景以及面临的挑战与应对策略等方面的内容。研究表明，大数据技术能够有效提升半导体制造过程的质量管控水平，增强企业竞争力，但在应用过程中也需解决数据安全、人才短缺等问题，以推动半导体产业的高质量发展。

关键词：大数据；半导体制造；质量管控；机器学习

DOI：10.64216/3080-1508.25.06.031

引言

半导体作为现代信息技术产业的核心，广泛应用于电子、通信、计算机、汽车等众多领域，其制造过程极为复杂，涉及数百甚至上千道工序，微小的工艺波动或设备异常都可能导致产品质量问题。随着半导体技术节点不断缩小，对产品质量和性能的要求达到了前所未有的高度，传统的质量管控方法已难以满足高精度、高效率的生产需求。大数据技术的崛起为半导体制造质量管控带来了新契机，通过收集、存储和分析半导体制造过程中产生的海量多源数据，能够实时监测生产状态，精准定位质量问题根源，提前预测质量风险，实现生产过程的优化和质量的提升。国内外研究中，英特尔、三星、台积电等国际企业已在实践中广泛应用并取得成效，国内企业和科研机构也在积极探索，但在应用深度和广度上仍有差距。本文通过梳理大数据在半导体制造质量管控中的应用原理、关键技术、应用场景及挑战，为相关研究和实践提供参考。

1 半导体制造过程质量管控特点与挑战

1.1 半导体制造工艺流程概述

半导体制造是一个复杂且高度精密的过程，主要包括硅片制造、芯片设计、光刻、蚀刻、离子注入、薄膜沉积、化学机械抛光等多个关键环节。首先将硅材料制成高纯度硅片，随后通过芯片设计确定电路结构，光刻工序将电路图案转移到硅片上，蚀刻去除不需要的材料以形成电路结构，离子注入引入特定杂质改变电学性能，薄膜沉积生长各类薄膜材料，化学机械抛光则用于平坦化硅片表面。这些工序相互关联、相互影响，任一环节出现问题都可能导致芯片质量下降甚至报废。

1.2 质量管控的重要性及特点

在半导体制造中，质量管控贯穿于整个生产流程。

高质量的半导体产品是保证电子设备性能稳定、可靠运行的关键，直接影响下游应用领域的产品质量和用户体验，同时能有效降低生产成本、提高生产效率和企业经济效益。其质量管控具有高精度、复杂性和实时性等特点：高精度要求对生产参数进行精确控制，任何微小偏差都可能产生重大影响；复杂性体现在涉及多学科领域和复杂工艺，质量影响因素众多且相互关联；实时性要求实时监测质量数据，及时发现并处理问题。

1.3 传统质量管控方法的局限性

传统的半导体制造质量管控方法主要依赖统计过程控制（SPC）、抽样检验等技术。SPC 通过对关键质量特性进行数据采集和统计分析判断生产状态，但基于历史数据设定控制界限，难以反映实时状态和潜在风险，且只能发现已出现的问题；抽样检验存在随机性和误差，无法全面检测产品质量，对小概率但影响重大的问题可能漏检。此外，传统方法在处理海量、多源、异构数据时能力有限，难以挖掘数据背后的质量信息和规律，已难以满足现代半导体制造对高质量、高效率的要求。

2 大数据在半导体制造质量管控中的应用原理与关键技术

2.1 大数据技术概述

大数据是指无法在一定时间范围内用常规软件工具处理的数据集合，具有数据量大、类型多样、处理速度快和价值密度低的“4V”特征。大数据技术是一系列用于采集、存储、管理、分析和可视化大数据的技术体系，包括数据采集与预处理、分布式存储、分布式计算与并行处理、数据挖掘与机器学习、数据可视化等技术。在半导体制造领域，大数据来源于生产设备传感器数据、工艺参数记录、产品检测数据、原材料质量数据等，为大数据技术应用提供了丰富资源，可通过高效处理和深

度分析挖掘质量信息和规律。

2.2 大数据在半导体制造质量管控中的应用原理

大数据在半导体制造质量管控中的应用基于数据驱动理念,通过收集、整理、分析和挖掘生产过程中的海量数据,实现对质量的全面监控、精准预测和有效控制。实时采集设备运行状态、工艺参数、环境参数等数据并预处理,去除噪声和异常值以提高质量;运用数据挖掘和机器学习算法提取与质量相关的特征和模式,建立质量预测模型;通过实时监测模型输入数据对产品质量进行预测评估,及时发现潜在问题并迅速定位根源,提供解决方案以优化生产过程和改进质量。

2.3 关键技术解析

-数据采集与预处理技术:数据采集通过在生产设备和工艺环节部署大量传感器和数据采集设备,实时采集多源数据,包括有线和无线传输方式及基于物联网的设备互联技术,确保数据准确及时传输。预处理技术对采集的数据进行清洗、去噪、标准化、归一化等处理,以提高数据质量和可用性。

-机器学习算法在质量预测中的应用:机器学习是大数据分析的核心技术,在质量预测中应用广泛,包括监督学习、无监督学习和半监督学习算法。监督学习利用已有质量数据和特征数据训练模型,学习工艺参数与质量指标的映射关系;无监督学习用于发现数据潜在模式和规律,为质量管控提供参考。

-数据挖掘技术挖掘质量相关模式与规律:数据挖掘从大量数据中发现潜在模式、关联规则和知识,在质量管控中可挖掘质量相关模式和规律,为质量改进提供决策支持,如通过关联规则挖掘发现工艺参数与质量缺陷的关联关系,通过序列模式挖掘分析质量数据随时间的变化趋势。

3 大数据在半导体制造过程中的数据处理与分析

3.1 数据收集来源与渠道

半导体制造过程中的数据来源广泛,包括生产设备数据(设备运行状态、工艺参数、报警信息、维护记录等)、工艺参数数据(各工序的关键控制参数)、产品检测数据(各环节的质量检测结果)、原材料数据(原材料质量参数等)。数据收集渠道主要有设备自带数据采集系统传输和利用物联网技术实时采集,还可通过人工录入收集无法自动采集的数据。

3.2 数据预处理方法与流程

数据预处理是大数据分析的关键环节,方法主要包

括数据清洗(去除噪声、异常值、重复数据和补充缺失值)、数据去噪(消除干扰信号,采用滤波算法、小波变换等方法)、数据标准化与归一化(将数据转换为统一格式和范围,便于分析和模型构建)、数据集成与融合(整合不同数据源数据,消除冗余和冲突,综合处理获得更全面准确的信息)。流程包括数据采集、传输、清洗、去噪、标准化与归一化、集成与融合,最终存储到数据仓库或大数据平台。

3.3 基于大数据的质量特征提取技术

质量特征提取是从原始数据中提取与产品质量密切相关的特征信息,作为质量预测模型的输入。基于大数据的提取技术包括:基于统计学的方法(计算均值、方差等统计量,提取基本特征和与质量相关的工艺参数);基于机器学习的方法(如主成分分析等降维算法,提取对质量影响大的主成分);基于深度学习的方法(如卷积神经网络、循环神经网络等,自动提取深层次、抽象的特征,适用于图像数据和时序数据处理)。

4 大数据驱动的半导体制造质量管控应用场景

4.1 实时质量监测与异常预警

大数据技术通过实时采集生产设备、工艺环节和产品检测的海量数据,构建实时质量监测系统,利用分布式计算和流处理技术快速处理分析数据。将实时数据与预设质量标准 and 历史正常数据对比,发现异常即发出预警并通知相关人员处理。结合机器学习算法不断学习历史异常数据特征,提高异常识别的准确性和及时性,能捕捉早期、微小的质量异常,为管控争取时间。

4.2 质量问题溯源与根因分析

当产品出现质量问题时,大数据技术整合生产全流程的历史数据,构建完整质量数据链条,利用关联规则分析、因果分析等数据挖掘方法,对质量问题数据与各环节数据进行深度关联,追溯问题产生路径。相比传统人工排查,该方式能快速准确地定位问题根源,为工艺优化和质量改进提供明确方向。

4.3 质量预测与工艺优化

大数据技术通过构建质量预测模型对产品质量进行前瞻性预测,基于历史生产数据选择相关特征参数,利用机器学习算法训练模型,根据实时数据预测质量指标。若预测质量不达标,系统提前提示并推荐工艺参数调整方案。同时,通过分析历史数据挖掘工艺参数最优组合范围,在保证质量的前提下降低能耗、减少原材料消耗、提高生产效率。

5 典型案例分析

5.1 国际案例：台积电基于大数据的良率提升实践

台积电构建覆盖全生产流程的大数据分析平台，实时采集数万台设备的传感器数据、工艺参数数据和产品检测数据，日均处理 PB 级数据。通过机器学习算法分析数据建立工艺参数与良率的预测模型，能提前预测参数偏差对良率的影响并自动发送调整指令，同时利用大数据进行质量问题溯源，大幅缩短溯源时间。这些措施使先进制程良率短期内提升 10% 以上，降低生产成本，巩固市场竞争力。

5.2 国内案例：中芯国际大数据质量管控系统应用

中芯国际搭建基于工业互联网的大数据平台，整合生产设备、检测系统和管理系统数据，重点应用于蚀刻和薄膜沉积工序的质量监测。通过分析参数变化趋势结合历史数据建立异常预警模型，提高异常检出率，利用数据挖掘技术优化光刻工序参数设置，提升产品合格率。其实践表明大数据技术能有效提升国内企业质量管控水平，缩小与国际先进水平的差距。

6 大数据在半导体制造质量管控中面临的挑战与对策

6.1 面临的主要挑战

大数据在半导体制造质量管控中面临多方面挑战：数据安全与隐私保护压力突出，制造数据包含企业核心技术和生产机密，泄露将造成重大损失，且涉及多方信息隐私，同时需应对不同地区数据安全法规，管控难度大；数据孤岛现象显著，企业内不同车间、设备和系统的数据格式与存储方式各异，难以整合分析，限制技术应用效果；专业人才短缺，既懂半导体制造又掌握大数据分析的复合型人才供不应求，影响系统构建与优化；数据质量参差不齐，受传感器精度、设备故障等影响，数据存在噪声、缺失值等问题，降低分析结果准确性，影响质量预测与预警可靠性。

6.2 应对策略

针对上述挑战，可采取以下应对策略：加强数据安全体系建设，建立分级分类管理制度，明确数据访问权限与处理规范，采用加密技术和安全设备保护数据，遵守相关法规并规范数据跨境流动；推动数据整合与标准化，制定统一数据标准和格式，规范不同系统接口，通过数据中台打破数据孤岛，实现全流程数据集中管理；加大人才培养与引进力度，与高校合作开设交叉学科专业培养人才，内部开展培训提升员工技能，同时通过多

种方式引进外部高端人才，构建复合型团队；优化数据采集与预处理流程，选用高精度采集设备，建立数据质量评估机制，及时处理数据问题，优化预处理算法以确保输入分析模型的数据质量。

7 结论与展望

7.1 研究结论

大数据技术为半导体制造质量管控提供了全新解决方案，通过实时质量监测与异常预警、质量问题溯源与根因分析、质量预测与工艺优化等场景，能有效提升产品质量、提高生产效率、降低生产成本。国内外企业实践验证了其有效性，但应用中仍面临数据安全、数据孤岛、人才短缺、数据质量等挑战，需通过相应策略解决。

7.2 未来展望

未来，随着相关技术发展，大数据在半导体制造质量管控中的应用将呈现以下趋势：智能化水平进一步提升，实现从“被动响应”到“主动预防”；全产业链数据融合应用，提升整个产业链质量水平；边缘计算与云计算协同，提高管控实时性和智能化；数字孪生技术深度融合，为质量问题诊断和工艺优化提供更直观精准的支持。大数据技术应用前景广阔，将推动半导体产业向更高质量、更高效率方向发展。

参考文献

- [1] 张明, 李华. 基于深度学习的半导体检测与预测算法研究(二)——多模态融合晶圆检测系统[J]. 半导体学报, 2025, 46(2): 123-135.
- [2] 王磊, 赵静. 基于深度学习的半导体故障诊断与寿命预测算法研究——基于机器学习的半导体探测器健康状态监测及寿命预测方法[J]. 微电子学与计算机, 2025, 42(5): 45-56.
- [3] 刘阳, 陈晨. 基于云大数据湖的半导体制造先进分析应用[J]. 计算机工程与应用, 2023, 59(15): 25-35.
- [4] 周强, 吴悦. 贝叶斯 AEWMA 控制图在半导体制造过程监测中的应用[J]. 工业工程与管理, 2023, 28(3): 112-120.
- [5] 陈辉, 李萌. 多模态融合晶圆检测系统在半导体生产中的效能研究[J]. 现代制造工程, 2025, 45(4): 45-56.
- [6] 赵宇, 孙浩. 基于机器学习的半导体探测器健康状态监测及寿命预测[J]. 传感器与微系统, 2025, 44(6): 35-45.