

基于 DBSCAN 与 HL 相结合的晶圆成品率预测研究

姚晓童, 李林

(上海理工大学 管理学院, 上海 200093)

摘要: 对半导体成品率进行预测分析可有效控制产品成本, 提高产品质量, 而缺陷问题是导致半导体成品率损失的关键因素。因此, 考虑晶圆缺陷聚集特性和数据嵌套性, 研究了一种密度聚类与多水平逻辑回归相结合的受缺陷限制的成品率预测方法。首先采用密度聚类算法获取晶片缺陷模式类型; 将在线缺陷数据在晶片水平进行整合, 作为多水平逻辑回归模型的输入参数; 根据多层次晶圆结构, 在模型中加入嵌套变量, 在批次层、晶圆层和组别层构建随机截距效应模型; 在产品层构建非随机变化截距与斜率模型进行成品率预测; 最后, 根据回归结果分析引起成品率损失的主要因素并提出相应的改进措施。通过仿真实验表明, 多水平逻辑回归模型的预测精度优于常用的 Seed's 成品率模型与嵌套结构逻辑回归模型, 该模型具有更高的预测能力与可行性。

关键词: 成品率模型; 密度聚类; 多水平逻辑回归; 嵌套结构

中图分类号: TP 315 文献标志码: A

Wafer yield prediction based on DBSCAN and HL

YAO Xiaotong, LI Lin

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: Semiconductor yield prediction is essential to control product cost and improve product quality. Defect is the key factor leading to yield loss. Therefore, considering defect gathered characteristics and the nested structure, a yield prediction methodology was presented based on density clustering and hierarchical logit regression. A density clustering algorithm was used to obtain the pattern classification of die defects, and on-line defect data were integrated at the die level as input parameters of hierarchical logit regression model. According to the hierarchical wafer structure, nested variables were added to the model to construct random intercept effect model in lot, wafer and group layers. A non-random intercept and slope model was constructed at the product layer. Finally, according to the regression results, the main factors causing yield loss were analyzed and the corresponding improvement measures were put forward. The simulation experiment shows that the prediction accuracy of the hierarchical logit regression model is better than that of Seed's yield model and nested structure logit model, and has higher prediction ability and feasibility.

Keywords: wafer yield model; density clustering; hierarchical logit regression; nested structure

收稿日期: 2020-08-31

基金项目: 上海市软科学研究领域重点项目 (19692104000)

第一作者: 姚晓童(1996-), 女, 硕士研究生。研究方向: 产线预测。E-mail: 2431445703@qq.com

通信作者: 李林(1966-), 女, 副教授。研究方向: 工业工程、质量管理。E-mail: LI_ft@163.com

成品率是半导体制造过程中一项重要的工艺性能指标,精准预测晶圆成品率对于改进生产工艺、减少晶圆制造损失具有重要意义。

目前成品率建模大多基于影响成品率的关键参数,如缺陷数量、大小、电性参数等。文献[1]考虑缺陷概率密度分布建立了经典的泊松成品率预测模型,由于模型结构简单,应用最为广泛,但预测准确度较低。文献[2]考虑缺陷分布存在过度分散性,对文献[1]进行了改进,建立了广义泊松分布模型预测晶圆成品率,提高了成品率的预测精度。文献[3]基于缺陷类型、大小、位置等基本信息建立因果统计模型,进行成品率的预测,并提出了成品率优化策略,但由于缺陷数据量庞大,模型求解困难。文献[4-5]考虑晶圆电性测试的关键参数,以最小化测试参数和预测误差为目标,选取显著特征向量,进而预测晶圆成品率,但电气参数数据对象多,对象间关联性强,导致模型预测准确度较低。文献[6]基于成品率历史数据提出一种混合聚合-熵-共识的模糊协同智能算法,首先得到成品率预测区间,然后用反向传播网络去模糊化得到成品率的预测值,具有较好的预测精度。影响晶圆成品率因素众多,数据量大且关系复杂,但缺陷数据仍然是成品率建模最具影响力的因素^[7]。

为提高预测精度以及确定晶圆缺陷来源,部分学者考虑晶圆缺陷的聚集性建立成品率预测模型。文献[8]考虑晶圆缺陷的邻近效应,利用缺陷模式建立成品率模型,具有较高的预测精度。文献[9-10]通过分析晶圆图,识别缺陷模式,帮助管理者确定减少缺陷任务的优先级以提高晶圆成品率。以上研究表明,考虑晶圆缺陷聚集特性,不仅可以快速查找缺陷来源,且模型预测精度高。但以上成品率模型均是在晶圆水平进行建模,仅考虑晶圆缺陷的一种模式。晶圆生产过程复杂,一片晶圆通常会存在多种缺陷模式,在晶圆水平进行建模在一定程度上降低了模型的预测准确度,难以获得全面的缺陷损失来源。实际生产过程中,缺陷数据存在嵌套结构^[11],即同组晶片缺陷模式具有相似性,现有文献大多都没有考虑缺陷数据嵌套结构特性。

本文考虑缺陷聚集性和数据的嵌套结构,首先用基于噪声的密度聚类(density-based spatial clustering of applications with noise, DBSCAN)算法

识别缺陷模式,将缺陷计量数据按缺陷模式在晶片水平汇总,基于缺陷模式和缺陷数量建立多水平逻辑回归(Hierarchical Logit, HL)晶片失效概率模型,研究引起晶片失效的主要缺陷模式、关键工序,进而完成成品率的预测。

1 问题描述

在半导体行业中,计算机辅助制造会产生海量数据集,通常每个数据集不能直接用来预测成品率,需要对不同数据集进行有效整合,如何获得可行数据是成品率建模复杂的原因之一。本文所用数据集为产线缺陷监控系统数据和电气参数测量数据。缺陷监控系统有两种:一种是以缺陷为单位的缺陷记录数据库,该系统主要记录了缺陷的大小、缺陷的类型以及缺陷的位置等基本信息;另一种为缺陷分析工程师使用光学显微镜扫描晶圆得到晶圆图(wafermap),分析晶圆图可以识别不同的缺陷模式,例如线性缺陷、团状缺陷等,帮助工程师定位缺陷来源。电气参数测量数据指在晶圆加工完成后,对晶圆上每个晶片进行针测,以检查晶片的性能,根据针测结果将晶片划分为不同的bin值。考虑晶圆缺陷聚集特性,本文用DBSCAN算法对晶圆缺陷进行聚类获取晶片缺陷模式;将在线缺陷检测系统中的缺陷在晶片水平进行汇总,即获得每个晶片缺陷数量;根据晶圆针测数据集获取每个晶片bin值,最后将晶片缺陷模式、晶片缺陷数量以及晶片bin值3个数据集进行整合,得到适用于成品率建模的数据集。

基于缺陷计量数据成品率建模通常会假设缺陷数据是独立的,但在晶圆实际生产过程中,缺陷数据是呈嵌套结构存在的,例如晶片嵌套在晶圆上,晶圆组合在批次中。因此,不同批次或者不同晶圆上的晶片并不是相互独立的,晶圆嵌套结构数据如图1所示。图中: y_{nijkt} 表示第 t 个产品、第 k 个批次、第 j 个晶圆、第 i 个组别的失效晶片; n_j 表示第 j 个晶圆失效晶片的数量; N 表示失效晶片总数。在该结构中,不同组的缺陷数据呈差异性,同组缺陷数据呈聚集性,HL模型可以解决嵌套结构数据组间非独立性与组内聚集性^[12]。为此,本文建立了HL成品率预测模型,该模型还可帮助企业识别显著性批次、组别,以快速定位关键工序。

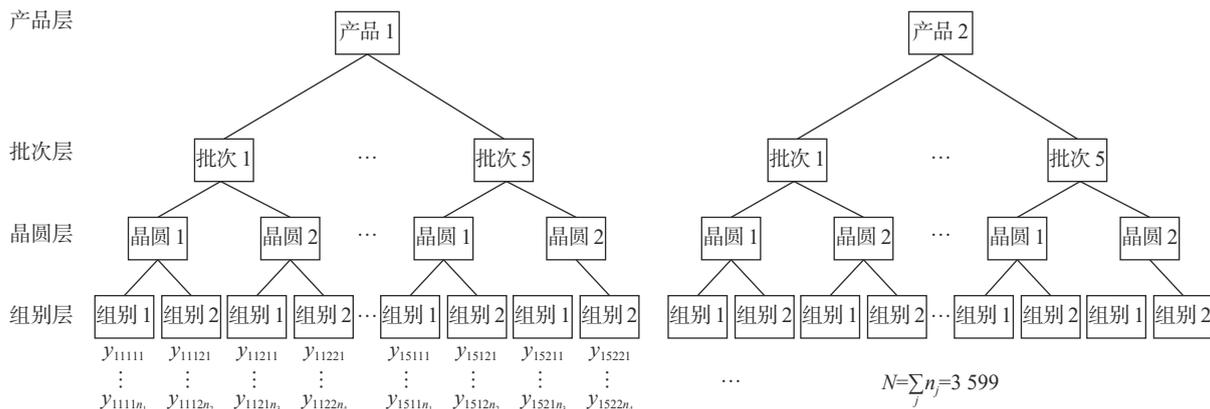


图 1 多层次晶圆结构

Fig.1 Multi-level wafer structure

2 晶圆缺陷模式识别

2.1 晶圆缺陷聚集特点及划分依据

晶圆缺陷聚集性表现为具有明显的簇、缺陷点集中和簇中点密度大。晶圆缺陷模式主要分为随机缺陷、局部团状缺陷、线性缺陷、边缘环状缺陷以及粒子缺陷。5种缺陷模式划分^[13]的主要依据如表1所示。

表 1 5种缺陷模式主要特点及划分依据

Tab.1 Main characteristics and classification basis of 5 defect modes

缺陷模式	特点及划分依据
局部团状缺陷	局部密度很大,缺陷点集中聚集,对成品率影响很大
线性缺陷	划痕缺陷,对成品率影响由划痕长短决定
边缘环状缺陷	只出现在晶圆边缘部分呈半圆或者圆形
随机缺陷	局部密度较大,但无规律性
粒子缺陷	尘埃粒子缺陷,密度小,对成品率影响较小

基于噪声的密度聚类算法可由密度可达关系找出最大密度相连的集合,不需要指定聚类个数,在有噪声的情况下可识别出任意形状的簇。

2.2 DBSCAN 聚类算法

假设邻域为 ϵ , 聚类密度阈值为 $\min Pts$, 缺陷点集为 $T = \{o_i, i = 1, 2, \dots, 6\}$, 对缺陷点进行相关定义。

a. 邻域: 缺陷点 $o_1, o_2 \in T$, 如果点 o_2 与 o_1 的距离不大于 ϵ , 则 o_1 的邻域可表示为 $N_\epsilon(o_1) = \{o_2 \in T | d(o_1, o_2) \leq \epsilon\}$, d 为缺陷点之间的距离。

b. 核心对象: 如果缺陷点 o_1 的邻域内至少存在 $\min Pts$ 个缺陷点, 则 o_1 是一个核心对象。

c. 密度直达: 如果缺陷点 o_2 在缺陷点 o_1 的邻域

内, 且 o_1 点为核心对象, 则称 o_2 由 o_1 的密度直达。

d. 密度可达: 存在缺陷点序列 $o_1, o_2, o_3, o_4, o_5, o_6$ 。在这个序列中, 每一个缺陷点都与它前一个缺陷点密度直达, 则称 o_1 与 o_6 密度可达。

e. 密度相连: 如果缺陷点 o_1 与缺陷点 o_6 密度可达, 缺陷点 o_2 与缺陷点 o_6 密度可达, 则 o_1 与 o_2 密度相连。

f. 簇: 由所有密度可达的缺陷点导出的最大密度相连的点集合为一个簇。对于不属于任何一个核心对象邻域的缺陷点标示为噪音点。

通过实验法调整邻域 ϵ 与密度阈值 $\min Pts$ 完成晶圆图所有缺陷点聚类, 形成分类簇。将分类簇根据表1划分缺陷模式(将小于密度阈值的缺陷点划分为粒子缺陷), 获得各个缺陷点分类。

3 Hierarchical Logit 模型

逻辑回归模型常用于晶圆成品率建模, 但它的基本假设为数据是相互独立的。考虑晶圆缺陷数据嵌套结构和聚集性, 本文使用多水平逻辑回归构建晶圆成品率模型。多水平逻辑回归是广义混合线性模型的一种。

广义混合线性模型主要由抽样模型、结构模型和连接函数3部分组成。抽样模型是指反应变量服从某一指数族分布, 本文晶圆成品率为二分类问题(晶片 pass/fail), 定义 $\varphi_{ni jkt}$ 为第 t 个产品、第 k 个批次、第 j 个晶圆、第 i 个组别、第 n 个晶片的失效概率, y_j 代表第 j 个晶圆上有 m_j 个缺陷晶片的失效晶片数量, 则 y_j 服从有 m_j 次实验、每次实验成功概率为 φ_n 的二项分布

$$y_j | \varphi_n \sim B(m_j, \varphi_n) \quad (1)$$

采用 Logit 连接函数,即

$$\varphi_{nijkt} = \frac{1}{1 + \exp(-\eta_{nijkt})} \quad (2)$$

式中, η_{nijkt} 为晶片失效概率的对数发生比, $\eta_{nijkt} = \lg\left(\frac{\varphi_{nijkt}}{1 - \varphi_{nijkt}}\right)$ 。 φ_{nijkt} 取值范围为 (0,1), 当 $\varphi_{nijkt} > 0.5$ 时, 晶片失效。

结构模型是解释变量与自然参数的线性表达式, 与广义线性模型不同的是, 广义混合线性模型的结构模型包括随机效应与固定效应两部分。当不定义随机效应时, 广义混合线性模型变为广义线性模型。通过晶片组合在晶圆上, 晶圆嵌套在批次上, 批次嵌套在产品层建立嵌套变量, 假设不同层次不同组别的晶片是相互独立的, 同组内晶片具有相似性。在批次层、晶圆层和组别层建立嵌套变量的随机截距模型, 在产品层引入晶片大小变量, 构建非随机变化斜率与截距模型。

多层次晶圆晶片失效概率的对数发生比综合线性模型为^[12]

$$\begin{aligned} \eta_{nijkt} = & \pi_{00000} + \pi_{00001}w_t + r_{0ijkt} + e_{00jkt} + u_{000kt} + \\ & (\pi_{10000} + \pi_{10001}w_t)x_{nijkt} + (\pi_{20000} + \pi_{20001}w_t)x_{nijkt} + \\ & (\pi_{30000} + \pi_{30001}w_t)x_{nijkt} + (\pi_{40000} + \pi_{40001}w_t)x_{nijkt} + \\ & (\pi_{50000} + \pi_{50001}w_t)x_{nijkt} \end{aligned} \quad (3)$$

式中: w_t 表示晶圆特征尺寸; x_{nijkt} 表示晶片缺陷数量; π_{00000} 为截距项, $\pi_{l0000}(l=1,2,3,4,5)$ 表示第 l 个缺陷模式的自然参数; π_{00001} 表示产品大小的自然参数; $\pi_{l0001}(l=1,2,3,4,5)$ 表示产品层与个体层交互效应的自然参数; r_{0ijkt} , e_{00jkt} , u_{000kt} 分别表示失效概率对数发生比在批次层、晶圆层、组别层的随机效应参数。

第 j 个晶圆成品率 Y_j 可表示为

$$Y_j = 1 - E(\varphi_j) \quad (4)$$

式中, $E(\varphi_j)$ 表示在第 j 个晶圆上晶片失效概率的期望值。

使用 SAS 9.1 软件 NLMIXED 过程构建模型。选用极大似然估计算法与对偶拟牛顿优化算法进行模型求解, 使用比值比(odds ratio)解释特征向量对晶片失效概率的影响, 如式(5)所示。比值比代表当晶片为第 l 个缺陷模式时, 缺陷数量增加一个单位会影响晶片失效概率的增加值。选用 Pearson 卡方统计量与 AIC 信息准则指标评价模型拟合优度^[13]。

$$Q_{OR} = \frac{odds_{x_{nijkt}+1}}{odds_{x_{nijkt}}} = e^{\pi_{l0000}} \quad (5)$$

式中: Q_{OR} 表示被检验的解释变量的比值比; $odds$ 表示晶片失效发生概率与不失效概率之比。

4 仿真实验及结果分析

本文缺陷数据样本选自某晶圆厂实际生产线监测系统, 包含 3 种不同工艺尺寸产品, 15 个批次, 30 个晶圆。3 种不同工艺尺寸产品分别可划分出 2480, 848, 533 个晶片。首先对 30 个晶圆图进行密度聚类, 共识别出 4578 个缺陷晶片。将产品 1 与产品 2 作为训练集, 产品 3 作为测试集。训练集数据共包括 3599 个缺陷晶片, 测试集数据共包含 979 个缺陷晶片。

4.1 晶圆缺陷模式识别

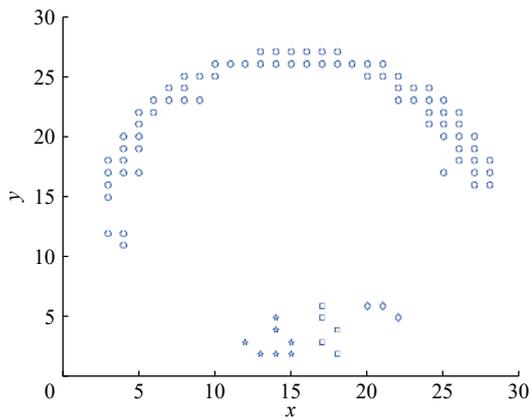
基于 DBSCAN 对每个晶圆图进行聚类, 根据实验法确定邻域 ϵ , 聚类密度阈值 $\min Pts$, 部分晶圆缺陷晶片聚类结果如坐标图图 2 所示。

每张晶圆图具有多种缺陷模式, 结合表 1 可知: 图 2(a) 属于环状缺陷与随机缺陷; 图 2(b) 属于线性缺陷; 图 2(c) 属于随机缺陷和局部团状缺陷。根据晶圆图获取每个缺陷晶片的坐标、缺陷模式、到中心的径向距离, 并根据径向距离对晶片分组, 分组界限为半径的一半。

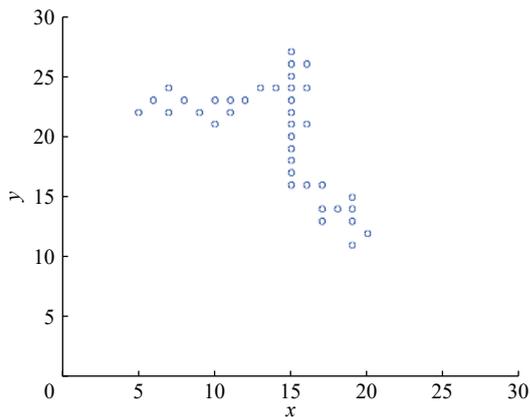
4.2 数据聚合及异常值检测

本文将在线缺陷测量数据与晶片针测数据基于晶片缺陷模式进行整合, 以获取适用于缺陷建模的数据。部分缺陷原始数据如表 2 所示, 加粗部分为位于同一晶片上的不同缺陷的数据。利用 Excel 数据透视表功能对其进行聚合, 缺陷测试数据与晶片针测数据整合后的数据如表 3 所示。根据晶片聚类结果, 对表 3 数据进行整合, 整合后的数据为 HL 模型的输入数据, 如表 4 所示。

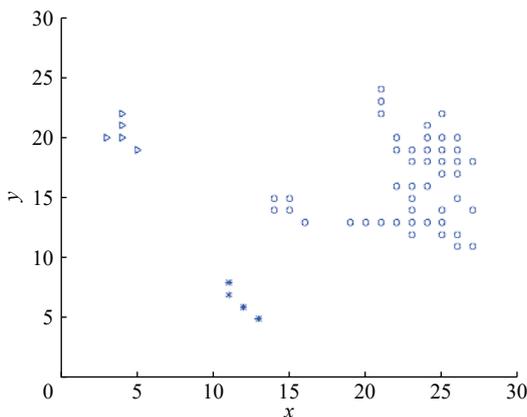
为了提高预测精度, 对数据进行分析与预处理, 将异常值在数据中剔除。对训练数据集进行统计分析, 每个晶片中存在缺陷均值为 2.8, 中位数为 1, $\frac{1}{4}$ 分位数为 1, $\frac{3}{4}$ 分位数为 2, 最小值为 1, 最大值为 144。可见, 在训练数据集中存在大量异常值。为消除异常值对模型的影响, 采用 one-side trimmed 均值异常值检测方法剔除异常值。该方法统一了数据集中的均值与中值, 其效率优于中值法异常值检测^[14]。本文设置 p 值为 5%, 根据统计分析结果, 95% 的晶片缺陷数在 5 以内, 所以将晶片缺陷数大于 5 的晶片样本在训练集中删除。



(a) 缺陷模式聚类结果 1



(b) 缺陷模式聚类结果 2



(c) 缺陷模式聚类结果 3

图 2 晶圆聚类结果

Fig.2 Wafer clustering results

表 2 缺陷原始数据

Tab.2 Defect raw data

产品编号	批次编号	晶圆编号	晶片x坐标	晶片y坐标
1	1	1	25	2
1	1	1	25	3
1	1	1	26	2
1	1	1	26	3
1	1	1	24	3
1	1	1	25	4
1	1	1	25	4
1	1	1	26	4
1	1	1	24	3
1	1	1	24	3
1	1	1	24	3

表 3 聚合后数据

Tab.3 Aggregated data

产品编号	批次编号	晶圆编号	晶片x坐标	晶片y坐标	缺陷数量	Bin箱	晶片测试值
1	1	1	25	2	1	1	0
1	1	1	25	3	1	1	0
1	1	1	26	2	1	1	0
1	1	1	26	3	1	1	0
1	1	1	24	3	5	5	1
1	1	1	25	4	2	1	0
1	1	1	26	4	1	1	0

为验证模型方法的有效性对测试集数据不作处理。

4.3 回归结果分析

4.3.1 Logit 回归

为了探讨该数据集是否需要使用多水平 Logit

回归，首先对训练集数据作一般 Logit 回归分析。模型输入参数为产品编号、批次编号、晶圆编号、组别、晶圆特征尺寸以及缺陷模式，其中晶圆特征尺寸为标准化后的数值。实验结果表明，模型拟合优度检验 Pearson 卡方 p 值为 0.009，但预测准确率 90.83%，表明模型没有充分利用样本数据，模型拟合优度表现不优，AIC 指标值为

表4 模型分析数据
Tab.4 Model analysis data

产品编号	批次编号	晶圆编号	晶圆尺寸	组别	团状缺陷	随机缺陷	线性缺陷	环状缺陷	粒子缺陷	晶片测试值
1	1	1	0.895	2	3	0	0	0	0	0
1	1	1	0.895	2	1	0	0	0	0	0
1	1	1	0.895	2	1	0	0	0	0	0
1	1	1	0.895	2	1	0	0	0	0	0
1	1	1	0.895	2	5	0	0	0	0	0
1	1	1	0.895	2	1	0	0	0	0	0
1	1	1	0.895	2	3	0	0	0	0	0
1	1	1	0.895	2	3	0	0	0	0	0
1	1	1	0.895	2	2	0	0	0	0	0
1	1	1	0.895	2	2	0	0	0	0	0
1	1	1	0.895	2	4	0	0	0	0	1

2423.61。Logit 模型显示晶圆 1、晶圆 2、组别 3、组别 7、组别 9、组别 15、组别 19、组别 23、随机缺陷模式、团状缺陷模式、线性缺陷模式、环状缺陷模式以及粒子缺陷模式因子具有显著性 ($p < 0.1$)。

4.3.2 Nested Logit 回归

考虑缺陷数据嵌套结构,在 Logit 回归模型中加入嵌套变量, Nested Logit 模型输入参数为产品编号(批次编号)、批次编号(晶圆编号)、晶圆编号(组别)、晶圆特征尺寸以及缺陷模式,输入数据集为剔除异常值后的训练集。与 Logit 模型相比, Nested Logit 模型给出了更加精确的信息,不仅识别出重要的批次、晶圆信息,同时给出产品 1 的批次 1 与其他批次有显著差异,批次 1 中的晶圆 1、晶圆 2 有显著差异。对于同一片晶圆,不同组别晶片失效概率差异较大,这些信息可以帮助企业工程师找出潜在的质量偏差,进而在制造过程中进行故障排除。在显著性水平 $\alpha = 0.1$ 下, Nested Logit 回归模型识别出产品 1(批次 1)、批次 1(晶圆 1)、批次 1(晶圆 2)、晶圆 1(组别 2)、随机缺陷、团状缺陷、线性缺陷、环状缺陷以及粒子缺陷解释变量具有显著性。Nested Logit 回归模型的 Pearson 卡方 p 值为 0.128,表明 Nested Logit 模型对数据拟合较优, AIC 指标值为 1902.191。该模型 AIC 指标值小于 Logit 回归模型的 AIC 指标值, p 值更加显著,表明 Nested Logit 回归模型优于一般 Logit 回归模型。

根据 Nested Logit 回归模型显著性因子构建晶片失效概率模型,由于产品编号、批次编号、晶圆编号、组别等变量对于每个数据样本都是不同的,因此利用随机缺陷、团状缺陷、线性缺陷、环状缺陷以及粒子缺陷因子构造晶片失效概率模型为

$$\varphi_n = \frac{1}{1 + \exp(-\eta_n)} \quad (6)$$

式中, $\eta_n = -4.432 1 + 1.419 9x_n + 1.420 3x_n + 1.377x_n + 1.156 6x_n + 0.814 9x_n$ 。

4.3.3 Hierarchical Logit 回归

Nested Logit 回归模型的基本假设为同组内的晶片是相互独立的,但在晶圆生产过程中所产生的系统缺陷通常是集群存在的,且 Nested Logit 回归模型没有考虑产品层与个体层的交互效应,在实现多产品的混合预测方面有待进一步考察。

本文所构建的 HL 回归模型的固定效应参数包括晶圆特征尺寸、缺陷模式、晶圆特征尺寸与缺陷模式交互变量,随机截距模型参数为产品编号(批次编号)、批次编号(晶圆编号)、晶圆编号(组别)。结果显示:随机效应参数中晶圆编号(组别)和产品编号(批次编号)具有显著性,协方差估计分别为(0.2584, 0.1053),表明缺陷模式斜率受产品层和组别层(晶片失效概率组均数)影响较大;但批次编号(晶圆编号)不具有显著性,表明在批次层和晶圆层晶片失效概率组间差异不大。在显著性水平 $\alpha = 0.1$ 下, HL 模型识别出产品 1(批次

1)、晶圆 1(组别 2)、晶圆 2(组别 2)、随机缺陷、团状缺陷、线性缺陷、环状缺陷以及粒子缺陷、晶圆特征尺寸、晶圆特征尺寸以及随机缺陷、晶圆特征尺寸以及团状缺陷、晶圆特征尺寸以及线性缺陷、晶圆特征尺寸以及环状缺陷解释变量具有显著性($p < 0.1$)。Pearson 卡方 p 值为 0.79, 模型的拟合优度提高, 在显著性 0.05 下, 该模型基于缺陷数据拟合晶片失效概率是合适的。AIC 指标值为 1692.5, 小于 Logit 回归模型 AIC 指标值与 Nested Logit 回归模型 AIC 指标值, 因此, HL 回归模型拟合更优。

在显著性水平 $\alpha = 0.1$ 下, 使用固定效应参数因子构建晶片失效概率, 模型为

$$\varphi_{nt} = \frac{1}{1 + \exp(-\eta_{nt})} \quad (7)$$

式中, $\eta_{nt} = -4.4155 + 0.553w_t + 1.5424x_n - 0.2498w_t \cdot x_n + 1.3376x_n - 0.2159w_t x_n + 1.3267x_n - 0.2433w_t x_n + 1.138x_n - 0.1694w_t x_n + 0.6889x_n$ 。

4.4 成品率预测及分析

根据模型(6)与模型(7)预测测试集晶片失效概率, 通过式(4)计算每个晶圆成品率。HL 回归模型、Nested Logit 回归模型与 Seed's 模型对测试集中每一个晶圆的缺陷晶片 pass 数预测值, 如图 3 所示。图 4 展示了 HL 回归模型、Nested Logit 回归模型与传统成品率预测模型 Seed's 模型^[1]对测试集中每个晶圆成品率的预测情况。

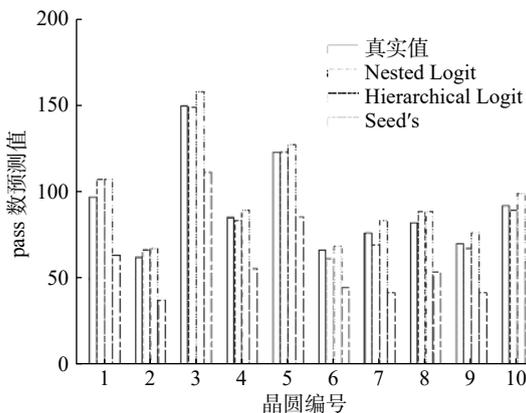


图 3 晶圆预测通过数

Fig.3 Number of wafers predicted to pass

由图 3 和图 4 可以看出 Seed's 模型成品率预测值低于成品率实际值, 而低估成品率会增加厂商缺陷分析成本。Nested Logit 回归模型的成品率预测值比真实值高, 预测偏差比 Seed's 模型低, 预测准确率较高。因此, 考虑晶圆生产过程的嵌

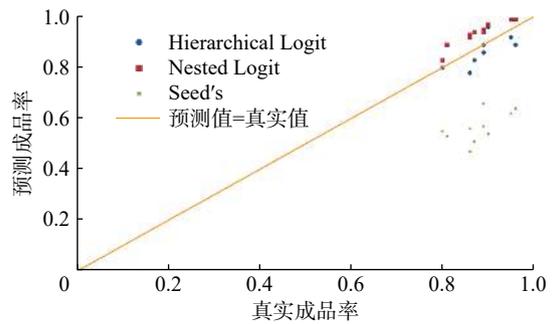


图 4 测试集晶圆成品率预测值

Fig.4 Predicted yield of the test set

套结构可以提高成品率预测精度, 同时可辅助工程师决策, 但未考虑产品层与个体层的交互效应和缺陷数据的组内聚集性, 可适用于单一产品的预测。本文提出的 HL 模型成品率预测值与实际值最为接近, 预测偏差远小于 Seed's 模型, 比 Nested Logit 回归模型预测更加精确, 可适用于不同产品成品率预测, 具有更高的预测能力与可行性。

根据 HL 模型实验结果得出, 缺陷分析工程师应重点关注产品 1 的批次 1。分析产品 1 批次 1, 晶圆缺陷模式多为团状缺陷且晶片缺陷数量多, 团状缺陷主要发生在刻蚀工艺和光刻工艺, 可检查各工艺步骤, 通过实验排除法确定缺陷来源。根据缺陷模式比值比, 团状缺陷模式对晶圆受缺陷限制的成品率影响最大, 晶片缺陷数量每增加一单位, 缺陷失效概率增加 4.27, 其次为线性缺陷模式、环状缺陷模式与粒子缺陷模式。晶圆特征尺寸与受缺陷限制的成品率呈负相关, 即晶圆特征尺寸越小, 对缺陷控制要求越高。

5 结论

基于缺陷计量数据构建晶片失效概率模型, 进而预测成品率。首先采用基于噪声的密度聚类算法对缺陷点聚类, 获得缺陷模式; 考虑缺陷聚集性与缺陷数据的嵌套结构, 构建 HL 模型预测晶片失效概率。该模型考虑了晶片水平更加具体的信息, 违背了更少的 Logit 回归模型的假设, 具有更高的预测精度。同时考虑了产品层与个体层的交互效应, 可适用于不同产品的混合预测。

下一阶段研究可以考虑更多的影响因素提高模型的预测能力, 如缺陷大小、缺陷类型等, 并针对各缺陷模式对成品率影响进行更深入的研究, 提出指导半导体生产线加工与调度决策的可行性策略, 提高芯片质量, 降低芯片成本。

参考文献:

- [1] SEEDS R B. Yield and cost analysis of bipolar LSI[J]. IEEE Transactions on Electron Devices, 1968, 15(6): 409.
- [2] YUAN T, BAE S J, KUO Y. Statistical models of overdispersed spatial defects for predicting the yield of integrated circuits[J]. IEEE Transactions on Reliability, 2020, 69(2): 510–521.
- [3] BARON M, TAKKEN A, YASHCHIN E, et al. Modeling and forecasting of defect-limited yield in semiconductor manufacturing[J]. IEEE Transactions on Semiconductor Manufacturing, 2008, 21(4): 614–624.
- [4] 吕佑龙, 许鸿伟, 郑城, 等. 基于混合式特征选择模型的晶圆允收测试关键参数识别方法 [J]. 中国机械工程, 2020, 31(16): 1978–1984, 1997.
- [5] 张秉裕, 陈琨太, 王振宇. 利用 WAT 资料建构晶圆良率预测模型之研究 [J]. 品质学报, 2011, 18(6): 519–538.
- [6] CHEN T. Forecasting the yield of a semiconductor product using a hybrid-aggregation and entropy-consensus fuzzy collaborative intelligence approach[J]. Measurement, 2019, 142: 60–67.
- [7] KUMAR N, KENNEDY K, GILDERSLEEVE K, et al. A review of yield modelling techniques for semiconductor manufacturing[J]. International Journal of Production Research, 2006, 44(23): 5019–5036.
- [8] HOCHBAUM D S, LIU S. Adjacency-clustering and its application for yield prediction in integrated circuit manufacturing[J]. Operations Research, 2018, 66(6): 1571–1585.
- [9] HESSINGER U, CHAN W, SCHAFMAN B, et al. Statistical correlation of inline defect to sort test in semiconductor manufacturing[C]//Proceedings of ASMC 2013 SEMI Advanced Semiconductor Manufacturing Conference. Saratoga Springs, NY, USA: IEEE, 2013.
- [10] 邱明辉, 曹政才, 刘民, 等. 基于 DBSCAN 与 FSVM 的半导体生产线成品率预测方法 [J]. 计算机集成制造系统, 2016, 22(11): 2594–2601.
- [11] KRUEGER D C, MONTGOMERY D C, MASTRANGELO C M. Application of generalized linear models to predict semiconductor yield using defect metrology data[J]. IEEE Transactions on Semiconductor Manufacturing, 2011, 24(1): 44–58.
- [12] 曹诗若, 苏宇楠, 田茂再. 基于分层线性模型的贝叶斯推断及其应用 [J]. 统计与决策, 2015(3): 4–8.
- [13] 卢笑蕾, 余建波. 基于混合模型与流形调节的晶圆表面缺陷识别 [J]. 计算机集成制造系统, 2018, 24(2): 302–308.
- [14] HU T M, SUNG S Y. A trimmed mean approach to finding spatial outliers[J]. Intelligent Data Analysis, 2004, 8(1): 79–95.

(编辑: 丁红艺)