

团 体 标 准

T/CBJ 2211—2024

白酒智能酿造 投配料应用指南

Baijiu intelligent brewing—Input material ratio application guide

2024-04-08 发布

2024-05-08 实施

中国酒业协会 发布
中国标准出版社 出版

中国标准出版社

目 次

前言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 技术架构	1
5 数据采集要求	2
6 模型建立方法	3
7 模型应用规范	5

中国标准出版社

中国标准出版社

前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第 1 部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国酒业协会提出并归口。

本文件起草单位：中国食品发酵工业研究院有限公司、中国酒业协会、泸州老窖股份有限公司、江苏今世缘酒业股份有限公司、中粮营养健康研究院有限公司、贵州习酒股份有限公司、江苏洋河酒厂股份有限公司、四川威斯派克科技有限公司、四川轻化工大学、济南趵突泉酿酒有限责任公司、泸州智通自动化设备有限公司、北京中医药大学、普瑞特机械制造股份有限公司、中国农业大学、山东景芝白酒有限公司、安徽宣酒集团股份有限公司、河南仰韶酒业有限公司、北京红星股份有限公司、宜宾南溪酒业有限公司。

本文件主要起草人：郭新光、杜小威、王健、王旭亮、程亚娟、丁子元、蔡小波、汪地强、黄越、罗惠波、吴志生、王莞、苏宁、赵亮亮、金绍武、孙云权、郑森、孙伟、赵婷、陈杰、李宗朋、李子文、杜静怡、李明磊、熊雅婷、周靖、侯琦、王小伟、乔新建、汪家胜。

中国标准出版社

中国标准出版社

白酒智能酿造 投配料应用指南

1 范围

本文件规定了白酒智能酿造过程中智能投配料的技术架构、数据采集要求、模型建立方法以及模型应用规范。

本文件适用于指导白酒智能酿造投配料生产过程的建设决策。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 5492 粮油检验 粮食、油料的色泽、气味、口味鉴定

GB/T 10345 白酒分析方法

GB/T 15686 高粱 单宁含量的测定

GB/T 33404 白酒感官品评导则

QB/T 4257 酿酒大曲通用分析方法

DB34/T 2264 固态发酵酒醅分析方法

3 术语和定义

下列术语和定义适用于本文件。

3.1

白酒智能投配料 **intelligent feeding and proportioning of materials in baijiu**

根据原辅料的品质、曲的各项检测指标、酒醅的各项理化指标以及温度、生产基本物理数据等，再结合粮、辅料、醅、曲、基酒的数据建立智能投配料模型，用模型指导生产，从而实现白酒生产投配料过程的精准化、智能化、数字化。

4 技术架构

白酒智能投配料的技术架构见图 1。

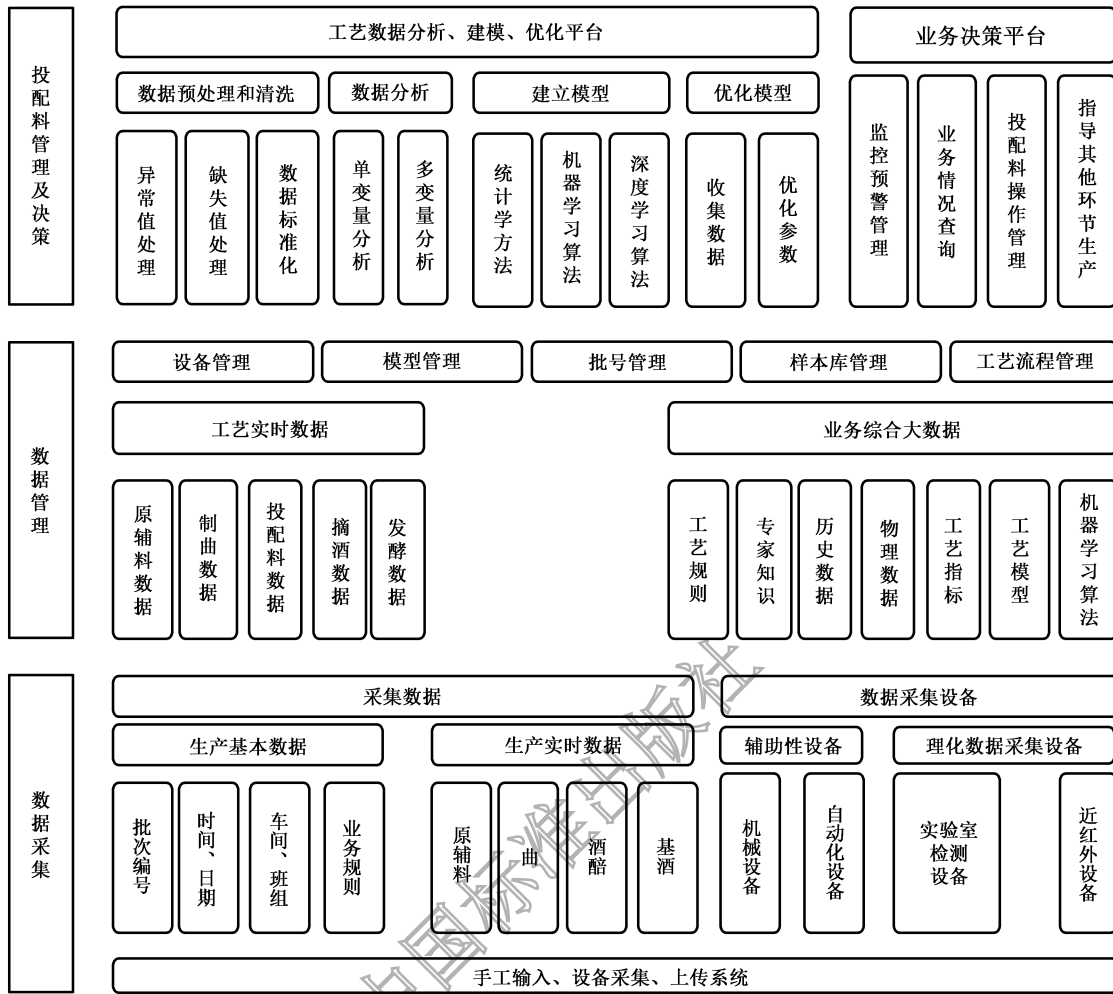


图 1 白酒智能投配料的技术架构图

5 数据采集要求

5.1 明确需要收集的数据和指标

按环节对采集的感官、理化指标和基础信息对象进行较全面的举例说明,见表 1。但是由于工艺不同,采集的数据指标也会有所差异。

注: 酒醅又称糟醅、酒糟;酒精度又称酒度和酒分;入池又称入窖;出池又称出窖。

表 1 各环节采集数据指标表

序号	生产环节	采集对象	感官/理化指标	基础信息
1	原粮采购	原料	淀粉、水分、单宁、脂肪酸值、支链淀粉、直链淀粉、色泽、气味、外观	入库批次编号、产地、种类、入库时间、入库数量
		辅料	水分、淀粉、果胶质、松紧度、色泽、外观	入库批次编号、产地、种类、入库时间、入库数量
	原料处理	粮粉	水分、破碎度	产地、种类、配比

表 1 各环节采集数据指标表（续）

序号	生产环节	采集对象	感官/理化指标	基础信息
2	制曲	曲块	水分、酸度、淀粉、糖化力、液化力、发酵力、酯化力	成品曲块批次编号、制曲原料信息；成曲时间、存储环境条件（温度、湿度等）
3	出池配醅	酒醅	酸度、淀粉、水分、酒精度、还原糖、容重/重量	窖池编号、酒醅批次编号、酒醅层数/类型；原料的用量、谷壳辅料用量
4	蒸馏/蒸料	酒醅/原料	蒸粮后糊精（淀粉）	窖池编号、酒醅批次编号、甑次编号；上甑花费时间、馏酒气压、装甑气压、蒸煮时间
5	摘酒	基酒	酒精度、总酸、总酯、己酸乙酯、乳酸乙酯、乙酸乙酯	窖池编号、酒醅批次编号、甑次编号；馏酒速度、馏酒温度、馏酒数量和类别（酒头、中段、酒尾各级别）
6	摊晾加曲	粉碎大曲	水分、酸度、淀粉、糖化力、液化力、发酵力、微生物	窖池编号、酒醅批次编号、甑次编号、粉碎曲块基础信息；加曲时间、曲的用量、量水用量、量水温度、摊晾时间
7	高温堆积	糖化堆	糖化堆发酵温度（顶温、中温、底温）	窖池编号、糖化堆批次编号、堆积发酵时间
8	入池发酵	酒醅	酸度、淀粉、水分、酒精度、还原糖	窖池编号、酒醅批次编号、酒醅层数/类型；入池温度、窖池发酵期间温度；环境温度、湿度、风速
9	基酒定级	基酒	基酒色泽和外观、香气、口味口感	窖池编号、基酒类别、基酒定级类别

5.2 建立各环节指标数据库

根据生产流程明确所需数据和指标后，需考虑采用 GB/T 5492、GB/T 15686、QB/T 4257、DB34/T 2264、GB/T 33404、GB/T 10345 等相关标准，在环节中选择合适的采集方式和设备，收集原辅料、曲、酒醅、基酒等样品的理化数据和基础信息数据。在确定数据采集频率和存储方式时，要结合建立智能投配料模型的分析需求和生产环节特点，例如：每小时采集一次或每天采集一次，并采用数据库、云端存储等方式进行存储。

6 模型建立方法

6.1 模型建立流程

原辅料的品质、曲的各项检测指标、酒醅的各项理化指标以及温度等都影响投配料决策中物料的比例。建立模型的前提是及时获取粮、辅料、醅、曲的质量参数，通过数据预处理等手段进行数据处理，保证数据的有效性。其次，根据以往基酒的情况，结合实际生产业务规则，进行单变量分析和变量间关联分析。最后，利用统计学、机器学习、深度学习等算法建立数学建模，获得投配料模型，从而指导粮醅比、曲醅比、糠醅比等工艺过程。

6.2 数据预处理和清洗

6.2.1 查找数据集中的缺失值与异常值

为了检测数据集中是否存在缺失、无效数据，常见的查找缺失值的方法如下。

- a) 汇总统计信息:利用 Excel 或其他类似工具,可以生成数据的汇总统计信息,比如计数、平均值、标准差等,从而快速发现缺失值。
- b) 编程语言:在使用编程语言(如 Python、R 等)进行数据分析时,可以使用相应的库和函数来识别缺失值。
- c) 数据可视化:通过绘制直方图、散点图等图表,可以发现数据中的缺失值。缺失值通常在图表中呈现为“空洞”。

异常值指白酒酿造过程的数据集中与业务规则明显不符的数值,如理化数据检测错误、数据录入错误等。这些值通常需要结合酿酒专家的经验来确定。在白酒酿造生产中常用的检测异常值的方法主要有以下几种。

- a) 统计学方法:通过统计学方法(例如标准差、中位数等)来衡量一个数值是否远离其他值,并将其定义为异常值。
- b) 可视化方法:利用可视化工具(例如箱线图、散点图等)展示数据分布情况,通过直观感受来确定哪些数值可能是异常值。
- c) 基于专家经验的:根据酿酒领域的专家知识和经验,确定哪些数值或者超过某些数值范围的数据是异常值。

6.2.2 数据清洗

数据清洗是建立智能投配料模型过程中必不可少的一步,它可以提高数据质量和可靠性,从而保证模型的准确度。数据清洗主要包括缺失值处理、异常值处理、数据标准化等流程。当发现数据中存在缺失值和异常值时,需要根据实际情况进行补充或处理。常见的处理方法包括删除含有缺失值或异常值的记录、使用统计方法进行插值、利用机器学习算法进行填补或替换等。另外,为了便于比较和分析,将数据按照一定规则进行标准化处理非常必要。例如,在白酒生产中,水分、还原糖等指标数值差异较大,可以将其标准化到一定区间范围内。通过数据清洗,可以使数据更加规范和规整,为后续分析和建模提供更加可靠的基础。

6.3 数据分析

6.3.1 单变量分析

6.3.1.1 描述性统计分析

进行描述性统计分析可以对指标的数据集进行总结和分析,有助于理解数据集的基本特征和潜在规律,包括中心位置、离散程度等,为后续的进一步分析和建模提供基础。在投配料分析中常见的描述性统计分析方法包括以下几个方面。

- a) 中心位置:反映指标数据集的平均水平,例如均值、中位数、众数等。
- b) 离散程度:反映指标数据集内部差异程度,例如标准差、方差等。

6.3.1.2 指标分布的可视化展示

指标分布的可视化展示是对数据集中每个指标的分布情况进行直观展示和比较。常见于投配料分析中的指标分布可视化方法如下。

- a) 直方图:用于显示数值型变量的分布情况,将数值范围划分为若干个区间,统计每个区间内的数据个数,并以条形图的形式展示,可用于出入池的水分、淀粉等指标。
- b) 箱线图:用于显示数值型变量的分布情况和离群点情况,将数据按照四分位数分成四个箱体,并展示出异常值,可通过该方法对指标的数据集分类。
- c) 饼图:用于显示分类变量的相对比例,将各个类别的占比以扇形图的形式展示,从而找到核心

类别或者核心环节的核心指标。

6.3.2 多变量分析

在投配料环节中,制酒全流程生产指标和理化指标之间存在相互影响和关联。为了描述并探究它们之间的关联关系,可以使用相关系数来衡量定量变量之间的相关性。由于制酒工艺的复杂性,还可以考虑用建模的方法探究它们之间的相关性。除了使用相关系数,还可以使用散点图和回归分析来探究多个变量之间的关联关系。此外,通过交叉分析比较制酒过程中不同环节的差异,也可以发现变量之间的联系,进一步深入理解数据集的特征和规律。通过这些探索性分析方法,可以更好地理解数据集,并为后续的建模和预测提供基础。

6.4 建立智能投配料模型

在收集到的数据批次间有对应关系,量化智能投配料模型目标的前提下,可以根据数据集,运用统计学、机器学习、深度学习等方法,对数据进行深入挖掘和分析,获得智能投配料模型,定量刻画自变量原辅料、曲、酒醅等指标与预测变量基酒之间的关系。通过获得的投配料预测模型函数表达式,实现对关联关系的解释与预测,动态调整配料量,得到最优的目标值,从而实现白酒智能投配料。

7 模型应用规范

7.1 投配料指导

在投配料环节中实现了投配料工艺的精细化和定量化控制,指导粮、辅料、曲、量水的投入量,确保白酒生产工艺执行的稳定,保证白酒的产量与质量。

7.2 优化出酒率、优级酒率

精准投配比可以在不增加成本的情况下,精准控制各环节生产情况,将出酒量和优级酒量最大化,从而提高生产效益,帮助企业实现对整个酿造过程的智能化控制。

7.3 模型持续升级

随着生产数据的不断积累和采集指标的不断完善,能够建立更全面、完善、准确的数据库。基于生产工艺特点,在数据挖掘技术的帮助下,对历史生产数据进行深入分析,并结合先进的人工智能算法,考虑多个相关因素对原料、酒醅、基酒产量和质量波动的影响,从而发掘数据中蕴含的规律,不断优化智能投配料模型,提高投配料的准确度和稳定性。

中国标准出版社

中国标准出版社

中国标准出版社

中国标准出版社