

T/GXAS

团 体 标 准

T/GXAS 783—2024

糖料甘蔗互联网数据采集与处理规范

Specification for internet data acquisition and processing of sugarcane

2024 - 07 - 23 发布

2024 - 07 - 29 实施

广西标准化协会 发布

前 言

本文件参照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。

本文件由广西壮族自治区农业科学院提出并宣贯。

本文件由广西标准化协会归口。

本文件起草单位：广西壮族自治区农业科学院、广西壮族自治区农业科学院农业科技信息研究所、广西甘蔗生产服务有限公司、广西壮族自治区农业科学院甘蔗研究所、广西壮族自治区自然资源信息中心、南宁师范大学、广西南亚热带农业科学研究所、广西壮族自治区气象科学研究所、广西财经学院、广西泛糖科技有限公司、广西前沿智能科技有限公司、广西大数据产业发展有限公司。

本文件主要起草人：黄启厅、杨绍锷、谢国雪、莫彬、王泽平、覃泽林、谢意、温国泉、黄清平、潘桂颖、梁永检、陈燕丽、李宇翔、唐利球、张家玫、林垚君、彭崇、马灿达、张秀龙、杨妮、朱明、阮晓静、苏秋群、段炼、孙明、何新洁、梁存穗、张小秋、马瑞升、凌玉荣、蓝奕丹、阮慧琳、刘书田、宁夏、邓树林、李敏、申垚阳、颜梅新、黄海荣、张洋洋、陈智、鄢忠波、刘琦、赖容、周燕新、闫青云、刘世平。

糖料甘蔗互联网数据采集与处理规范

1 范围

本文件界定了糖料甘蔗的术语和定义、缩略语，给出了糖料甘蔗互联网数据采集与处理的基本原则，规定了数据采集、数据处理、数据存储和数据安全保护的要求。

本文件适用于糖料甘蔗及其加工产品互联网数据的采集与处理。

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 36625.3 智慧城市 数据融合 第3部分：数据采集规范

YD/T 3865 工业互联网数据安全保护要求

3 术语和定义

下列术语和定义适用于本文件。

3.1

糖料甘蔗 sugarcane

供糖厂制糖用的原料甘蔗。

[来源：GB/T 10498—2010，3.1]

4 缩略语

下列缩略语适用于本文件。

GDP：国内生产总值（Gross Domestic Product）

5 基本原则

5.1 一致性

所采集数据与信息源提供的数据完全一致，并保持采集数据的原始完整性。

5.2 可靠性

对采集到的数据进行核验，确保数据的准确性和可信度。

5.3 合法性

数据采集过程合法合规。

5.4 安全性

对信息数据配套相应的管理措施，保证信息数据保存及传输过程中的安全。

6 数据采集

6.1 数据来源

所采集的数据应来自互联网的著作、期刊文献，以及中国气象网、国家统计局、农业农村部门和中国糖业协会等相关官方网站。

6.2 采集内容

应从糖料甘蔗生长过程及其所需的环境条件、糖料甘蔗砍收和加工产品等环节进行采集，糖料甘蔗数据采集内容包括但不限于以下参数：

——生长气象；

- 温度；
- 降雨量；
- 空气相对湿度；
- 日照时长；
- 光照强度。

——土壤；

- 土壤含水量；
- 土壤pH值；
- 土壤有机质含量；
- 土壤含氮（全氮）量；
- 土壤含磷（有效磷）量；
- 土壤含钾（速效钾）量。

——投入品；

- 蔗种；
- 肥料；
- 农药；
- 地膜
- 生物制剂。

——种植面积；

- 新植蔗；
- 宿根蔗。

——生长状况；

- 苗数；
- 有效株数；
- 株高；
- 茎径；
- 有效茎数；
- 田间锤度。

——生物灾害发生情况；

- 病害；
- 虫害；
- 草害。

——气象灾害发生情况；

- 旱害；
- 寒冻害；
- 洪涝；
- 台风。

——产量与GDP；

- 糖料甘蔗产量；
- 蔗农人数；
- 该地区总人数；
- 该地区GDP。

——食用糖产量与价格；

- 白砂糖产量；
- 红糖产量；

- 白砂糖价格；
- 红糖价格。
- 制糖副产物产量；
 - 蔗渣；
 - 浆纸；
 - 滤泥；
 - 糖蜜；
 - 酒精。
- 食用糖国内市场；
 - 食用糖国内消费量；
 - 食用糖出口量；
 - 食用糖进口量；
 - 食用糖库存量。
- 食用糖消费结构；
 - 食用糖工业消费量；
 - 食用糖民用消费量。

6.3 采集方法

采用但不限于以下方法：

- 人工采集：通过人工填报或导入等方式，从互联网采集数据；
- 网络自动获取：通过自动化程序从互联网获取数据；
- 数据接口：通过接口程序与应用程序、数据库、网络等数据源对接，获取数据。

7 数据处理

7.1 数据结构化

对采集的文本、图片、视频等非结构化或半结构化数据进行提取，并转换为结构化的格式、类型或数据库，糖料甘蔗互联网采集数据类型参数说明见附录A。

7.2 数据清洗

按GB/T 36625.3的规定执行。

7.3 数据统计

对采集到的数据，除特定说明外，均以某省（自治区）中糖料甘蔗主产区的市/县（区）为区域单位，按照规定格式汇总统计全年各项数据内容。糖料甘蔗互联网采集数据汇总见附录B。

8 数据存储

对处理后的数据，应做好数据存储及备份措施，满足以下要求：

- 应采用硬盘、光盘、磁带、磁盘阵列等大容量存储介质；
- 建立完善的数据备份管理制度，规范备份数据的等级、标识、归档工作，同时做好备份策略和恢复计划。

9 数据安全保护

按 YD/T 3865 的规定执行。

附录 A

(资料性)

糖料甘蔗互联网采集数据类型说明

表A.1给出了糖料甘蔗互联网采集数据类型说明。

表A.1 糖料甘蔗互联网采集数据类型说明

序号	采集数据类型	采集数据说明		
		数据类型	精度	单位
1	地区	char	-	-
2	年份	char	-	-
3	日期	char	-	-
4	温度	float	一位小数	°C
5	降雨量	float	一位小数	mm
6	空气相对湿度	int	零位小数	%
7	日照时长	float	一位小数	h
8	光照强度	int	零位小数	Lux
9	土壤含水量	float	一位小数	%
10	土壤 pH 值	float	一位小数	-
11	土壤有机质含量	float	两位小数	g/kg
12	土壤含氮量 (全氮)	float	两位小数	g/kg
13	土壤含磷量 (有效磷)	float	两位小数	mg/kg
14	土壤含钾量 (速效钾)	float	两位小数	mg/kg
15	蔗种	float	两位小数	$\times 10^4$ t
16	肥料	float	两位小数	$\times 10^4$ t
17	农药	float	两位小数	t
18	地膜	float	两位小数	t
19	生物制剂	float	两位小数	t
20	新植蔗种植面积	float	两位小数	(667×10^4) m ²
21	宿根蔗种植面积	float	两位小数	(667×10^4) m ²
22	苗数	int	零位小数	株
23	有效株数	int	零位小数	株
24	株高	float	一位小数	cm
25	茎径	float	两位小数	mm
26	有效茎数	int	零位小数	条
27	田间锤度	float	两位小数	%
28	病害	char	-	-
29	虫害	char	-	-
30	草害	char	-	-
31	旱害	char	-	-
32	寒冻害	char	-	-
33	洪涝	char	-	-
34	台风	char	-	-
35	气象灾害发生的起始时间	date	-	-
36	气象灾害发生的终止时间	date	-	-
37	气象灾害发生的地点	char	-	-
38	受灾面积	float	两位小数	(667×10^4) m ²
39	灾害等级	char	-	-
40	糖料甘蔗产量	float	两位小数	$\times 10^4$ t
41	蔗农人数	float	两位小数	$\times 10^4$ 人
42	该地区总人数	float	两位小数	$\times 10^4$ 人
43	该地区 GDP	float	两位小数	$\times 10^8$ 元

表 A.1 糖料甘蔗互联网采集数据类型说明(续)

序号	采集数据类型	采集数据说明		
		数据类型	精度	单位
44	白砂糖产量	float	两位小数	$\times 10^4$ t
45	红糖产量	float	两位小数	$\times 10^4$ t
46	白砂糖价格	int	零位小数	元/t
47	红糖价格	int	零位小数	元/t
48	蔗渣产量	float	两位小数	$\times 10^4$ t
49	浆纸	float	两位小数	$\times 10^4$ t
50	滤泥产量	float	两位小数	$\times 10^4$ t
51	糖蜜产量	float	两位小数	$\times 10^4$ t
52	酒精产量	float	两位小数	$\times 10^4$ t
53	食用糖国内消费量	float	两位小数	$\times 10^4$ t
54	食用糖出口量	float	两位小数	$\times 10^4$ t
55	食用糖进口量	float	两位小数	$\times 10^4$ t
56	食用糖库存量	float	两位小数	$\times 10^4$ t
57	食用糖工业消费量	float	两位小数	$\times 10^4$ t
58	食用糖民用消费量	float	两位小数	$\times 10^4$ t
59	食用糖国内消费量前三地区	char	-	-
60	食用糖出口量前三国家	char	-	-
61	食用糖进口量前三国家	char	-	-
62	数据来源及形式	char	-	-

注：若出现某类数据类型无相关数据，则在表格中用“-”表示。

GXAS
广西标准化协会

附录 B
(资料性)
糖料甘蔗互联网数据采集汇总

B.1 糖料甘蔗生长气象参数

表B.1给出了糖料甘蔗生长气象参数汇总表。

表B.1 糖料甘蔗生长气候参数汇总表

地区	日期	温度(°C)	降雨量(mm)	空气相对湿度(%)	日照时长(h)	光照强度(Lux)	数据来源及形式

注：以天为时间单位进行统计。地区填写市/县(区)。

B.2 糖料甘蔗生长环境参数

表B.2给出了糖料甘蔗生长环境参数汇总表。

表B.2 糖料甘蔗生长环境参数汇总表

地区	日期	土壤含水量(%)	土壤pH值	土壤有机质含量(g/kg)	土壤含氮(全氮)量(g/kg)	土壤含磷(有效磷)量(mg/kg)	土壤含钾(速效钾)量(mg/kg)	数据来源及形式

注：以天为时间单位进行统计。地区填写市/县(区)。

B.3 糖料甘蔗投入品使用情况

表B.3给出了糖料甘蔗投入品使用情况汇总表。

表B.3 糖料甘蔗投入品使用情况汇总表

地区	日期	蔗种用量($\times 10^4$ t)	肥料用量($\times 10^4$ t)	农药用量(t)	地膜用量(t)	生物制剂用量(t)	数据来源及形式

注：地区填写市/县(区)。

B.7 糖料甘蔗气象灾害发生情况

表B.7给出了糖料甘蔗气象灾害发生情况汇总表。

表B.7 糖料甘蔗气象灾害发生情况汇总表

自然灾害类型	地区	起始日期	终止日期	受灾面积 (667×10^4) m ²	灾害等级	数据来源及形式
干旱						
寒冻害						
洪涝						
台风						
注：地区填写市/县（区）。						

B.8 糖料甘蔗产量与 GDP

表B.8给出了糖料甘蔗产量与GDP汇总表。

表B.8 糖料甘蔗产量与 GDP 汇总表

地区	日期	糖料甘蔗产量 ($\times 10^4$ t)	蔗农人数 ($\times 10^4$ 人)	该地区总人数 ($\times 10^4$ 人)	该地区 GDP	数据来源及形式
注：地区填写市/县（区）。						

B.9 食用糖与制糖副产物产量

表B.9给出了食用糖与制糖副产物产量汇总表。

表B.9 食用糖与制糖副产物产量汇总表

地区	日期	白砂糖产量 ($\times 10^4$ t)	红糖产量 ($\times 10^4$ t)	蔗渣 ($\times 10^4$ t)	浆纸 ($\times 10^4$ t)	滤泥 ($\times 10^4$ t)	糖蜜 ($\times 10^4$ t)	酒精 ($\times 10^4$ t)	数据来源及形式
注：地区填写市/县（区）。									

B.10 食用糖日价格报表

表B.10给出了食用糖日价格报表。

表B.10 成品糖日价格报表

地区	日期	白砂糖价格 (元/t)	红糖价格 (元/t)
注：地区填写市/县（区）。			

参 考 文 献

- [1] GB/T 10498—2010 糖料甘蔗
 - [2] T/GXAS 588—2023 数字蔗田 数据分类与编码规范
-

中华人民共和国团体标准

糖料甘蔗互联网数据采集与处理规范

T/GXAS 783—2024

广西标准化协会统一印制

版权专有 侵权必究