

Technical Working Party on Testing Methods and Techniques**TWM/3/19****Third Session
Beijing, China, April 28 to May 1, 2025****Original:** English
Date: April 9, 2025

DEVELOPMENT OF BIG DATA PLATFORM FOR DUS EXAMINATION*Document prepared by an expert from China**Disclaimer: this document does not represent UPOV policies or guidance*

The annex to this document contains a copy of a presentation “Development of big data platform for DUS examination”, to be made by an expert from China, at the third session of the TWM.

[Annex follows]



中国农业科学院蔬菜花卉研究所
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences



DEVELOPMENT OF BIG DATA PLATFORM FOR DUS EXAMINATION

Yang Kun
Deputy Director of Beijing Sub-center of New Plant Variety Tests, Ministry of Agriculture and Rural Affairs, China
TWM3, Beijing, April 28 to May 1, 2025

1

CONTENTS

1. Background
2. Current development
3. Future plan

2

CONTENTS

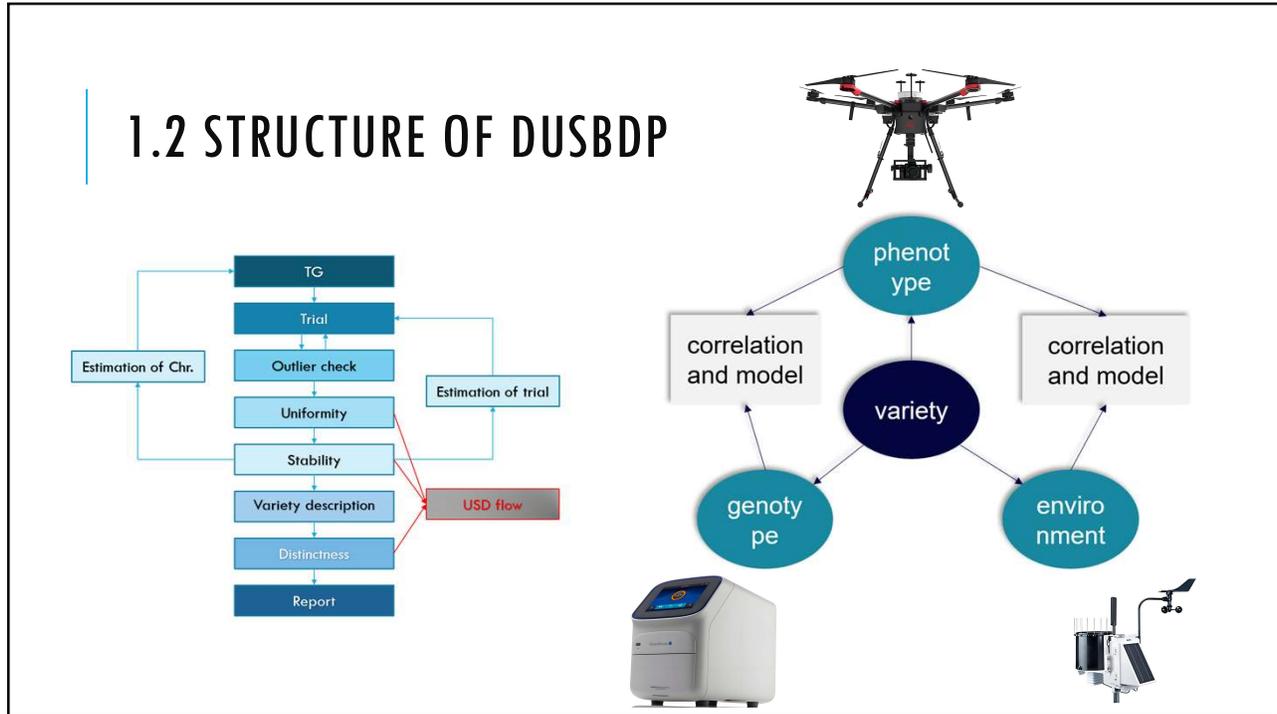
1. Background
2. Current development
3. Future plan

3

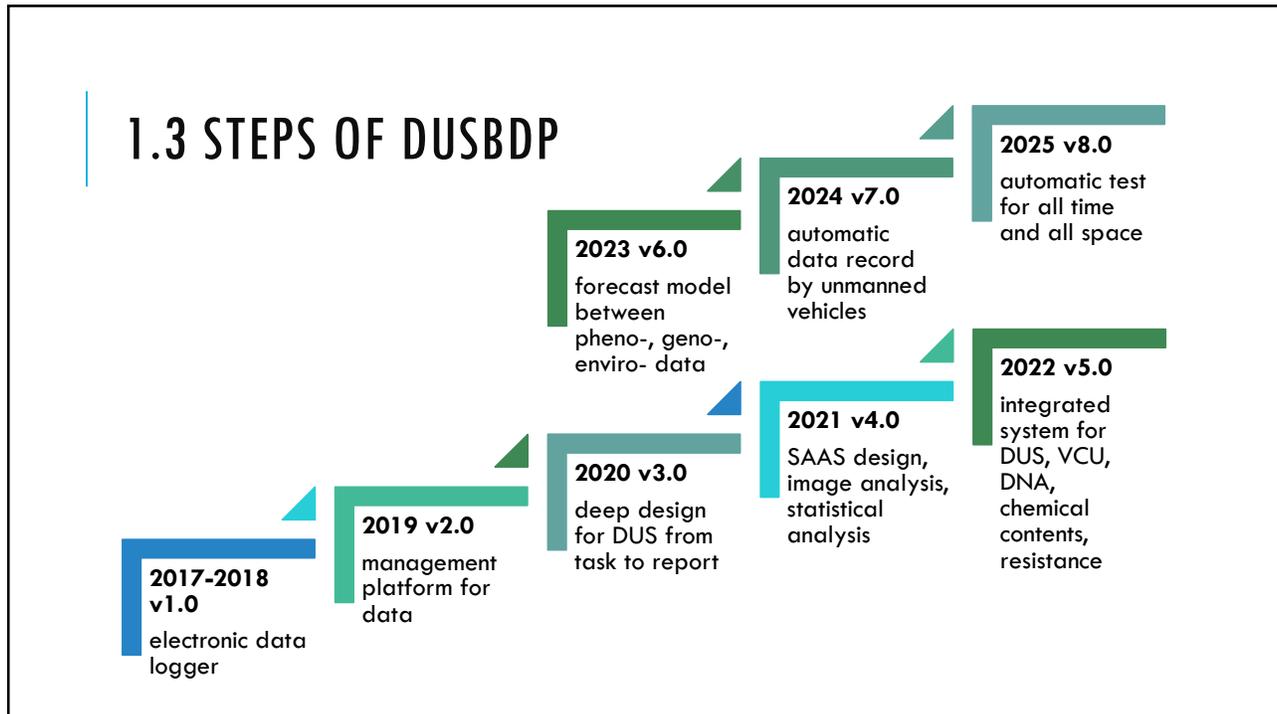
1.1 PROBLEMS IN DUS TESTING

- ① Personnel errors and environmental errors.
- ② Low efficiency of data collection and analysis.
- ③ Low linear correlation between molecular distance and morphological distance.
- ④ Imperfection of minimum distance of D, U and S.

4



5



6

1.4 HISTORY OF DUSBDP

1. 2017 We developed a datalogger which is prototype of DUSBDP.
2. 2021 We reported DUSBDP4.0 in TWC39.
3. 2023 We reported DUSBDP5.5 in TWO55.
4. 2024 We abandoned old platform and built a new one for better analysis module.

7

CONTENTS

1. Background
2. Current development
3. Future plan

8

2.1 DATA MANAGEMENT PLATFORM AND LOGGER APP



Data management platform



Data logger APP

9

2.2 MANAGEMENT OF TG PARAMETERS

性状管理

① Crop

② Test guideline

③ Characteristics

④ Note

序号	编号	名称	数据类型	单位	备注	操作
1	株高	株高	数值	厘米		编辑 删除
2	穗长	穗长	数值	厘米		编辑 删除
3	穗重	穗重	数值	克		编辑 删除
4	穗粒重	穗粒重	数值	克		编辑 删除
5	穗粒数	穗粒数	数值			编辑 删除
6	穗粒长	穗粒长	数值	厘米		编辑 删除
7	穗粒宽	穗粒宽	数值	厘米		编辑 删除
8	穗粒厚	穗粒厚	数值	厘米		编辑 删除
9	穗粒重	穗粒重	数值	克		编辑 删除
10	穗粒长	穗粒长	数值	厘米		编辑 删除
11	穗粒宽	穗粒宽	数值	厘米		编辑 删除
12	穗粒厚	穗粒厚	数值	厘米		编辑 删除

10

2.3 MANAGEMENT OF IMAGE TYPES

图像设置

序号	作物名称	图像编号	描述	是否应用于报告	图像类别	操作
1	玉米	1	幼苗	<input checked="" type="checkbox"/>	检测图片	编辑 删除
2	玉米	2	植株	<input checked="" type="checkbox"/>	检测图片	编辑 删除
3	玉米	3	花序	<input checked="" type="checkbox"/>	检测图片	编辑 删除
4	玉米	4	雄穗	<input checked="" type="checkbox"/>	检测图片	编辑 删除
5	玉米	5	花药	<input checked="" type="checkbox"/>	检测图片	编辑 删除
6	玉米	6	雄穗	<input checked="" type="checkbox"/>	检测图片	编辑 删除
7	玉米	7	籽粒果穗	<input checked="" type="checkbox"/>	检测图片	编辑 删除
8	玉米	8	小穴期	<input checked="" type="checkbox"/>	检测图片	编辑 删除
9	玉米	9	雌花	<input checked="" type="checkbox"/>	检测图片	编辑 删除
10	玉米	10	籽粒	<input checked="" type="checkbox"/>	检测图片	编辑 删除
11	玉米	11	中心轴	<input checked="" type="checkbox"/>	检测图片	编辑 删除
12	玉米	12	横断面	<input checked="" type="checkbox"/>	检测图片	编辑 删除
13	玉米	100	分枝果穗	<input checked="" type="checkbox"/>	检测图片	编辑 删除

Image without algorithm



Image with algorithm



11

2.4 DATA MANAGEMENT

数据录入

试验任务	操作	类型	品种	测试编号	试验	试验类型	总株数	种植株数	1	2	3
2024年上海农科院油菜籽制式试验	详情	标准	CA11912	20237000108Z	2023	1	120	0	3	3	2
2024年上海农科院转基因玉米性状试验	详情	申请	C20178	20235010304A	2023	1	120	0	3	3	2
2014年上海农科院玉米性状试验	详情	申请	C21962	20225010306A	2023	1	120	0	2	2	2
2023年玉米GWAS试验	详情	申请	C22614	20225010300A	2023	1	120	0	2	3	2
2022年玉米杂交种试验	详情	申请	C44726	20225010303A	2023	1	120	0	2	3	2
2022年甜糯分枝试验	详情	申请	C27729	20225010305A	2023	1	120	0	3	3	2
2022年玉米杂交种试验	详情	申请	DK223	20225010347A	2023	1	120	0	4	5	2
2022年玉米GWAS试验	详情	申请	DK226	20225010348A	2023	1	120	0	4	5	2
2021年玉米GWAS试验	详情	申请	DK228	20225010345A	2023	1	120	0	4	3	2
2021年玉米杂交种试验	详情	申请	DS2301	20225010209A	2023	1	120	0	4	5	2
2020年玉米田间试验	详情	申请	DS2302	20225010208A	2023	1	120	0	4	5	2
2019年玉米田间试验	详情	申请	DS2304	20225010210A	2023	1	120	0	4	3	2

Data from Excel

Data from APP

数据录入

试验编号: yj350

类型: 申请

总株数: 96

种植株数: 0

试验任务

12 16

12 16

12 16

12 16

保存并继续

返回

12

2.5 DATA ANALYSIS

The screenshot shows a data analysis interface. At the top, there are tabs for '数据概览' (Data Overview) and '数据分析' (Data Analysis). Below the tabs, there are filters for '当前试验组: 2022年玉米YOUSI测试组' and '对比试验组: 2021年玉米YOUSI测试组'. A green button labeled '一键数据分析' (One-click Data Analysis) is highlighted with an arrow and the text 'One click for all analysis'. Below this, a flowchart shows the process: '数据源' (Data Source) -> '数据清洗' (Data Cleaning) -> '数据建模' (Data Modeling). From '数据建模', the flow splits into two paths: '模型训练' (Model Training) leading to 'OFFLINE', '模型', '在线', and '模型'; and '模型部署' (Model Deployment) leading to '模型', '模型', and '模型'. To the right, a table titled 'CORN' displays data for various corn samples, including '品种' (Variety), '日期' (Date), '产量' (Yield), '水分' (Moisture), and '其他' (Other). The table has columns for '品种', '日期', '产量', '水分', and '其他'. The data rows show various corn varieties and their corresponding metrics.

13

2.6 IMAGE MANAGEMENT AND ANALYSIS

The screenshot shows an image management and analysis interface. On the left, there is a list of image files. In the center, there are three image thumbnails labeled 'TX-SIC01172', 'TX-SIC01172', and 'TX-SIC01172'. A red circle highlights the first thumbnail, with an arrow pointing to the text 'Data from image'. Below the thumbnails, there is a table with columns for '品种' (Variety), '日期' (Date), '产量' (Yield), '水分' (Moisture), and '其他' (Other). The table has columns for '品种', '日期', '产量', '水分', and '其他'. The data rows show various corn varieties and their corresponding metrics.

14

2.7 DUS REPORT

The screenshot displays a web-based DUS report interface. On the left, a table lists 13 entries with columns for '序号' (Serial Number), '品种名称' (Cultivar Name), '特异性分析' (Specificity Analysis), '一致性分析' (Uniformity Analysis), '稳定性分析' (Stability Analysis), '报告状态' (Report Status), and '报告管理' (Report Management). The '报告状态' column contains colored circles (orange, green, blue) indicating the status of each report. On the right, three detailed charts are shown, each with a title and a data visualization area. The top chart is titled 'distinctness', the middle one 'uniformity', and the bottom one 'stability'. Blue arrows point from the '报告管理' column of the table to these three charts, indicating that the charts provide detailed views for the reports listed in the table.

15

CONTENTS

1. Background
2. Current development
3. Future plan

16

3.1 ADVANTAGES

- ① One-click data analysis with several checks generates perfect reports for a trial.
- ② Images play an increasingly important role in DUS testing.
- ③ We have more time to seek secrets in data.

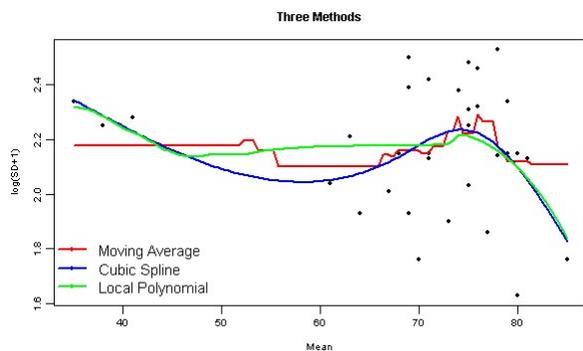
17

3.2 SHORTCOMINGS

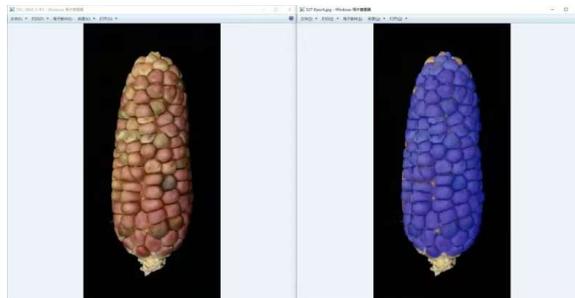
- ① There are a lot of gaps between our platform and commercial software or hardware.
- ② We need time to move molecular data and environmental data from old platform to new platform.
- ③ We need time to develop algorithms for correlation analysis between morphological data, genotypic data and environmental data.

18

3.3 OPTIMISM ABOUT THE FUTURE



Mathematic improvement



AI in image analysis

19



中国农业科学院蔬菜花卉研究所
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences





THANK YOU FOR YOUR ATTENTION!

Yang Kun
yangkun@caas.cn
+86-10-8210 5623
+86-(0)1391 1259 308

20