

农作物品种试验数据管理平台设计与实现

于春花¹ 王玉玺² 邱 军² 曾 波² 杨 锋¹ 王晓锋¹

闫宝岐¹ 杨 硕¹ 丁欢欢¹ 葛 晓¹ 王开义¹

(¹北京市农林科学院信息技术研究中心,北京 100097;²全国农业技术推广服务中心,北京 100125)

摘要:品种试验作为国家品种审定的重要组成部分,是新品种审定与推广的基础和依据。以农作物品种试验为对象,分析研究品种试验的组织架构、试验流程、各阶段业务及需求,构建了品种试验信息化流程,将信息技术运用到品种试验中,开发了农作物品种试验数据管理平台,具有品种管理、方案管理、数据管理、数据汇总、统计监督等功能,实现了品种试验全业务环节的信息化管理。通过平台的应用,实现了流程管理信息化、试验操作标准化、数据采集即时化,构建了国内渠道权威、数据标准、管理规范品种表型数据库,为种业大数据分析利用提供了数据支撑。

关键词:农作物;品种试验;信息化;数据管理平台

国家品种审定制度是保障我国种源自主可控的重要支撑,也是农作物育种产业化应用的关键环节。品种试验作为国家品种审定的重要组成部分,是新品种审定与推广的基础和依据,有利于生产用种的安全^[1]。品种试验包括区域试验、生产试验及品种特异性、一致性、稳定性测试(《主要农作物品种审定办法》)。每年度五大农作物大约有 800 多个试验站承担公益性田间试验,从业人员包括数据采集人员、试验主持人等,流程上包括田间采集、数据保存、数据监管、数据汇总等环节,形成了复杂的数据收集体系。在数据采集层面,试验点对标准的理

解及记载方法有差异,相同性状的单位或者分级不一致,导致底层数据不统一。在电子化保存方面,试验点数据的录入容易出错,Excel 也不能对数据正确性进行主动检验;在数据汇总时,要花费大量的时间核查试验点数据,重新整理数据,并在 Excel 表中编写不同公式对多个试验点填报的性状数据进行计算、汇总,整个过程繁琐、效率低。在数据监管层面,数据分级报送方式,导致试验管理部门、试验执行部门很难查看即时数据,不利于品种试验过程的监管。因此针对以上问题,农作物品种试验数据管理平台(以下简称平台)的开发和应用,能够较好地推动品种试验信息化技术标准,提高品种试验质量,保障底层数据质量。

通信作者:王开义

4.2 提升种子质量监督管理人员素质水平 目前,种子质量检测人员的素质与能力参差不齐,工作过程中还需要进一步加强教育和培训工作,确保其综合素质能力的有效提升^[4-5]。与此同时,在种子质量检测人员的教育与培训工作中,可以引入检测技能竞赛等新型考核制度,不仅能使其专业技能与素质得以提升,还可以激发其工作积极性。另外,适当增加种子检测机构的人员编制,引进专业技术人员进入种子检验队伍;争取财政资金,购入新的仪器设备,保证检测工作稳步推进,确保农民用上安全种、放心种。

参考文献

- [1] 宋国瑞,于喜泉,张敏,董丽娜.新形势下种子检验工作的现状与发展趋势.吉林蔬菜,2020(2): 47-48
- [2] 韩海亚,郑伟才.论加强种子检验工作对农业增产增效的重要性.种子科技,2017,35(5): 29-31
- [3] 周泽宇,张力科,金石桥.认清形势 把握机遇 全面推进种子检验工作.中国农技推广,2017,33(4): 4-6
- [4] 李建红,董琳娜,邓志文,欧阳昊婷.加强农作物种子质量监管的思考.中国种业,2018(8): 12-13
- [5] 付娜.强化种子质量监管 确保农业生产用种安全.农业科技通讯,2020(11): 24-26

(收稿日期: 2022-05-18)

信息技术已经应用到农业很多方面。在育种方面, Han 等^[2-3]基于云体系结构, 研发了作物育种信息管理系统, 可以高效地管理育种材料并进行家谱跟踪, 进一步计算亲本的一般配合力和特殊配合力, 为亲本选择和组合选择提供了依据。李建新等^[4]基于 LNMP (Linux 平台 +Nginx 服务器软件 +MySQL 数据库管理系统 +PHP 超文本预处理软件), 创建了玉米品种及其亲本系谱数据库。在主要农作物品种审定、非主要农作物品种登记以及品种权保护方面, 分别有品种审定数据管理系统、非主要农作物品种登记管理系统和植物新品种保护管理系统, 实现了农作物品种保护、审定、登记的信息化管理。

但在品种试验方面, 大部分的研究还集中在试验结果分析、试验设计优化等方面, 利用信息技术实现品种试验全流程的管理很少。如刘太国等^[5]、胡学旭等^[6]、宋晓霞等^[7]利用参试品种, 在抗病性、品质变化、品种的主要系谱和生产潜力方面进行了分析研究。高辉明等^[8]、姚金保等^[9]利用多年区试中对照品种和参试品种的资料, 研究了品种表现与环境之间的关系。许乃银等^[10-11]采用 GGE 双标图方法对皮棉产量、综合评价品种选择指数的鉴别力、代表性、理想指数等进行了分析和综合评价。张毅等^[12]提出了北部冬麦区农作物品种区域试验的重复次数和试点数量的优化设计方案。刘哲等^[13]利用区试数据和调查数据, 在县域精细尺度下, 研究作物品种种植适宜性精细区划方法。许乃银等^[14]、冯勇等^[15]利用 Excel 自带的功能实现了试验点数据的规范以及汇总报告中“品种评述”的自动生成。王虎等^[16]、叶思菁等^[17]利用移动端、GIS 技术实现了作物大田测试数据、环境数据的采集。也有少部分学者利用信息技术研发了专门的系统, 分别用于“品种评述”、试验数据预处理以及数据分析等环节^[18-19], 但是未涉及品种试验全流程的管理, 普及度也不够高。

农作物品种试验数据管理平台将信息技术、数据库技术等运用到品种试验过程, 构建品种试验信息化管理流程, 实现品种试验的全流程可跟踪、数据可追溯、性状可重现, 实时上报、实时监督, 提升数据上报及时性和准确性, 规范了试验数据。

1 品种试验信息化需求分析

1.1 品种试验管理模式 品种试验由试验管理部

门组织实施, 试验点管理部门协助开展品种试验监督管理相关工作。试验方案执行部门负责协助制定试验实施方案、开展试验监督检查和技术培训, 监督指导各试点品种试验, 起草试验总结报告。试验点按照试验实施方案开展品种试验, 接受试验管理部门、试验点管理部门和试验方案执行部门的监督指导, 及时提交客观、准确的试验数据和总结报告, 做好品种试验资料留档。

1.2 品种试验信息化流程分析 根据品种试验管理模式, 平台设置 4 个角色类型, 明确各角色职责范围的业务及业务流程: (1) 试验管理部门: 审核初试品种, 制定试验实施方案, 对品种试验过程、数据及汇总数据进行监督。(2) 试验执行部门: 评级已参试品种, 制定生态区组方案, 生成试验任务下发试验点人员进行数据录入, 监管试验过程、数据并进行试验数据汇总, 形成汇总报告。(3) 试验点管理部门: 监管权限内试验点。(4) 试验点: 根据下发的试验任务种植品种, 采集性状数据、图片后上传, 遇到试验异常情况报备给试验方案执行部门, 试验结束后编写年终报告并提交。

2 品种试验数据管理平台设计

平台采用 B/S 架构, 以 Oracle 关系型数据库、Java 语言为开发工具。平台不同角色的权限不同, 登录平台后, 不同角色看到的功能以及数据由权限决定。平台整体架构见图 1, 分 4 层: (1) 数据采集层: 包括通过平台上传数据及田间采集设备 (简称 PDA)、考种设备等外接设备采集数据; (2) 数据存储层: 根据不同数据的特性分为参试品种数据、试验方案数据、品种性状数据、异常报备数据以及图片数据。(3) 数据处理层: 用户权限验证用于过滤不同角色的用户登录平台后使用的功能及数据; 品种参试验证用于品种申报时品种信息的验证; 试验方案制定、试验任务生成用于分配试验任务到具体试验点; 试验数据上报用于试验点将性状数据、图片数据上传到平台; 试验数据汇总用于多点数据的处理与分析。(4) 数据应用层: 包括试验性状和图片数据查询、试验过程监督、试验点年终报告和汇总报告的生成、试验数据的总体分析。

3 品种试验数据管理平台的实现

3.1 平台功能 在平台架构设计基础上开发了一级功能模块 8 个: (1) 品种管理: 管理参试品种的基

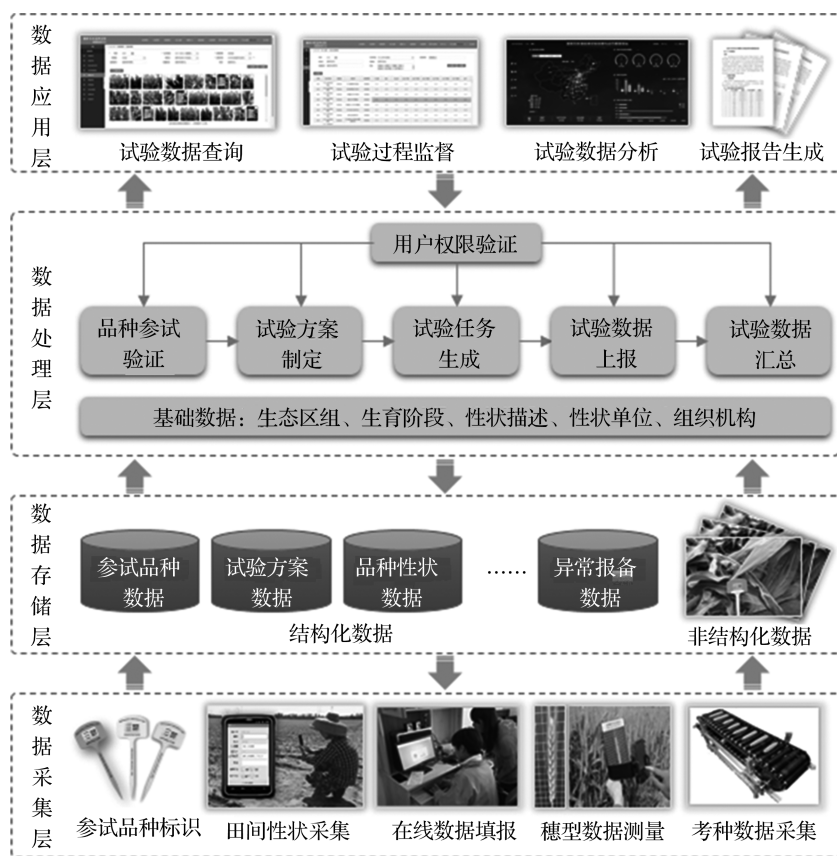


图1 平台架构

本信息、评级信息,并实现品种查询。(2)方案管理:制定品种试验实施方案,包括各生态区组试验品种分组、大田试验的试验点、需要鉴定的抗性品质项目和鉴定单位以及试验任务生成。(3)数据管理:主要用于试验过程中品种性状数据和图片录入、查看、编辑;异常情况的报备与审核,试点报告提交与审核。(4)数据汇总:由试验执行部门对所管辖生态区组所有试验进行试点/品种剔除、数据处理、分析、汇总,最终生成汇总报告。(5)统计监督:用于品种试验过程进度把握及运行数据统计。(6)基础数据:用于支撑整个平台运行的基础类数据,包括性状、生态区组、试验点、性状单位、报告模板管理等。(7)系统管理:用于角色权限配置、用户信息管理。(8)培训资料:用于各角色操作手册上传、下载。

3.2 实现了品种试验标准的数字化 品种试验的核心目标是获得多年多点品种的性状数据,以此数据来评价品种的产量、抗性、适应性等内容,因此高质量的数据是评价的保证。首先,对各作物品种试验技术规范、历史数据进行研究,梳理了各作物需要采集的性状以及不同生态区组间的操作差异,在作

物范围内对相同性状进行名称、采集标准的规范统一,并对性状的数据类型、采集阶段、多点间性状处理等属性做了详细规定。其次,统一了多年数据处理、数据汇总的方法以及试验点、品种数据的质量指标。最后,统一了年终报告以及汇总报告的体例、展示内容。以上内容的统一,大幅度提高了数据的科学性、准确性,实现了不同生态区组间、不同年份间、不同试验点间在空间、时间维度上数据的统一。

3.3 实现了试验进度的监督管理 在试验过程中,数据的及时上报与监督有助于管理层实时掌握试验情况,按照以往管理方式,方案执行部门和管理部门只有在试验点提交年终报告后才能看到试验数据。平台根据各区性状设置预警时间,在数据最佳上报时间跟踪提醒,促进试验点在各生育期采集数据后及时上报性状数据、图片以及田间施肥等栽培信息,管理层可在平台查询各生育期数据,实时掌握试验品种田间生长情况、各试验点数据的上报情况,及时给予相应指导,保证试验顺利实施。图2为数据监督查看相关功能。

3.4 构建了品种试验全流程信息化管理体系 在业务上,平台构建了从试验品种申请、品种审核、试

验方案制定、田间数据填报、田间数据审核、数据处理分析、数据汇总的信息化全流程管理(图3),明确在各业务阶段不同角色的职责范畴,并根据不同角色职责分配相应功能权限以及数据权限。在品种试

验周期上,构建了试验第1年到完成试验程序多年试验流程以及各年度品种全生育期周期的性状数据、参试信息的信息化,实现了品种试验的全流程可跟踪、数据可追溯、性状可重现。

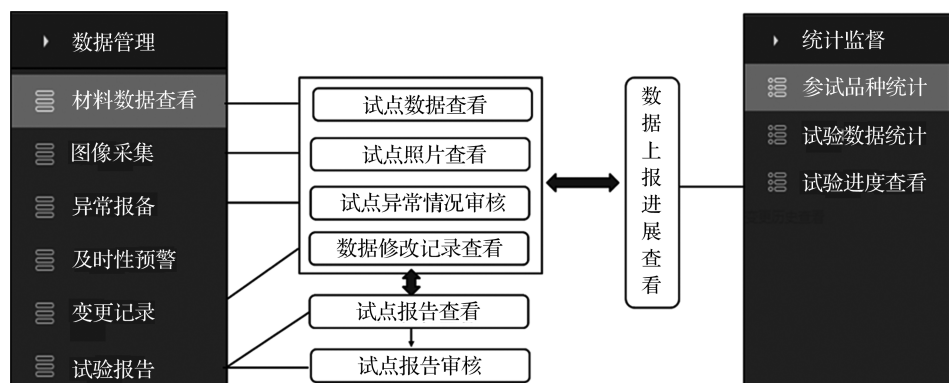


图2 数据监督查看

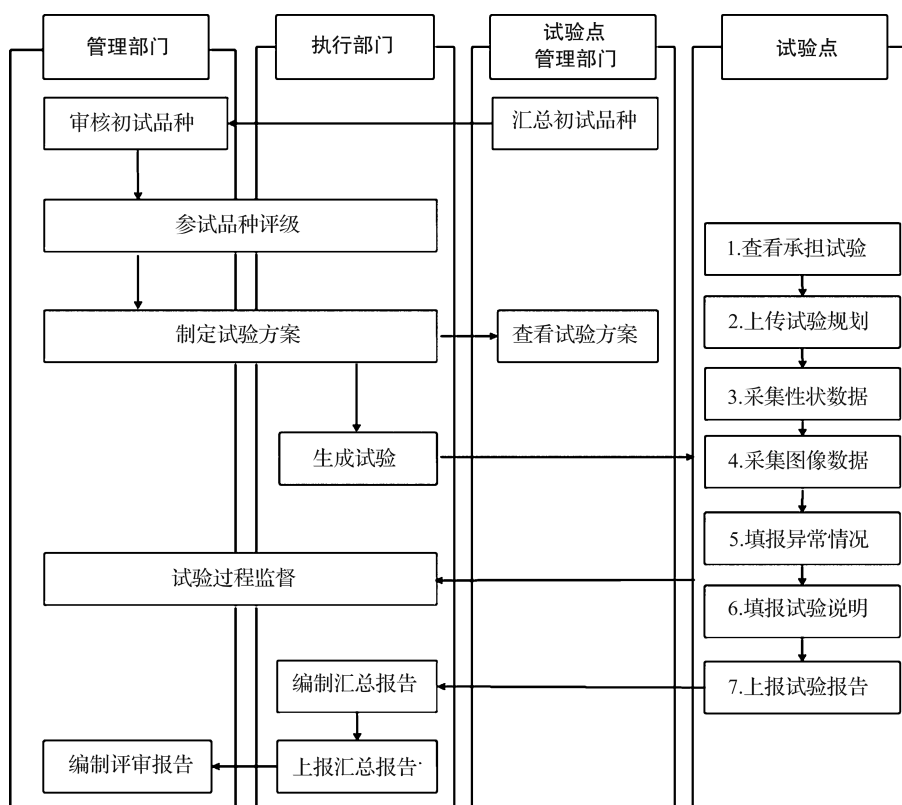


图3 信息化业务流程

4 品种试验数据管理平台的实施效果

4.1 平台运行数据 农作物品种试验数据管理平台实现了从品种到试验方案,到数据填报汇总分析,再到数据监督的规范化、标准化、信息化管理。平台从2017年开始示范实施,截止到2021年,覆盖五大

农作物全部生态区组,累计管理公益性试验参试品种5000多个、试验点1000多个,生成试验12224条、试验点年终报告6422份、汇总报告407份,采集图片数超35万张,采集性状数(包括公益性试验、联合体、绿色通道3个渠道)约2000多万条,构建了

国内渠道权威、数据标准、管理规范农作物品种表型数据库,为种业大数据的建设提供了重要支撑。

4.2 信息化流程的建立促进了业务流、数据流的优化 农作物品种试验信息化的关键是构建业务流程和数据标准的统一。首先,研究试验的业务过程及关键阶段、角色划分、各角色功能和数据权限、应用场景,建立科学合理的阶段与角色职责,实现线下业务向线上业务的转变。其次,本平台在性状采集、图片采集、报告格式及报告内容、数据处理、数据分析方面建立了作物内统一的标准,性状标准、数据处理标准、数据分析标准使得性状数据在试验点间、年度间、不同区组间都具有可比性以及更好的分析性,显著提高了数据质量。通过平台的使用,把品种试验各环节对业务及角色的要求固化到平台,促进各角色理解各自的职责范畴,使得品种试验业务更为规范流畅,形成优质的业务流,进一步产生优质数据流,试验数据更加准确规范。

4.3 作为表型数据库参与种业大数据联合分析 植物表型组学是在基因组水平上系统研究植物或细胞在不同环境条件下所有表型的学科,包括空间尺度以及整个生育期的时间维度上的信息^[20]。品种试验采用多年多点方式,在多个年度间、多个地理环境下种植,可收集到品种在不同年度、不同环境下的品种个体、群体以及整个生育期的多维度的数据,加上平台对数据规范的硬约束,保证了数据的准确性、高质量,可作为表型数据库,具有很强的可利用价值。

在数据利用上,通过平台中品种在产量、抗性、品质等性状上的变化,研究品种审定的发展趋势,为育种家制定育种目标提供参考。此外,联合试验过程中气象试验站采集的环境气象数据,可进行品种环境适应性分析,为品种适宜推广区域做参考。结合基因型、环境数据进行联合分析,对揭示作物生命科学规律、提高作物功能基因组学和分子育种研究水平等具有重大意义^[21]。

参考文献

- [1] 郭利磊,张笑晴.我国农作物品种区试审定制度的改革与发展.中国种业,2019(2):12-15
- [2] Han Y Y, Wang K Y, Liu Z Q, Pan S H, Zhao X Y, Wang S F. Golden seed breeding cloud platform for the management of crop breeding material and genealogical tracking. Computers and Electronics in Agriculture, 2018, 152: 206-214

- [3] Han Y Y, Wang K Y, Liu Z Q, Pan S H, Zhao X Y, Zhang Q, Wang S F. Research on hybrid crop breeding information management system based on combining ability analysis. Sustainability, 2020, 12: 1-16
- [4] 李建新,席蒙慧,张嘉玮,席蒙娟,田丁,鲁懿哲,陈晓阳,李卫华,张雪海,汤继华.中国玉米品种及其亲本系谱数据库的创建与利用.中国农业科学,2020,53(16):3404-3411
- [5] 刘太国,邱军,周益林,徐世昌,陈怀谷,刘艳,高利,刘博,郑传临,陈万权.中国冬小麦区域试验品种抗病性评价.中国农业科学,2015,48(15):2967-2975
- [6] 胡学旭,孙丽娟,周桂英,吴丽娜,陆伟,李为喜,王爽,杨秀兰,宋敬可,王步军.2000-2015年北部、黄淮冬麦区国家区试品种的品质特征.作物学报,2017,43(4):501-509
- [7] 宋晓霞,吉万全.黄淮南片小麦区域试验品种(系)的生产潜力及主要系谱分析.麦类作物学报,2018,38(12):1427-1436
- [8] 高辉明,张正斌,徐萍,杨引福,卫云宗,刘新月.2001-2009年中国北部冬小麦生育期和产量变化.中国农业科学,2013,46(11):2201-2210
- [9] 姚金保,张鹏,余桂红,马鸿翔,杨学周,周森平,张平平.江苏省小麦品种(系)籽粒产量基因型与环境互作分析.麦类作物学报,2021,41(2):191-202
- [10] 许乃银,张国伟,李健,周治国.基于HA-GGE双标图的长江流域棉花区域试验环境评价.作物学报,2012,38(12):2229-2236
- [11] 许乃银,李健.棉花区试中品种多性状选择的理想试验环境鉴别.作物学报,2014,40(11):1936-1945
- [12] 张毅,许乃银,郭利磊,杨子光,张笑晴,杨晓妮.我国北部冬麦区小麦区域试验重复次数和试点数量的优化设计.作物学报,2020,46(8):1166-1173
- [13] 刘哲,唐日晶,赵祖亮,李绍明,朱德海,张晓东.黄淮海地区玉米品种适宜性精细区划研究.农业机械学报,2015(8):281-288
- [14] 许乃银,李健.棉花区试总结中“品种评述”批量生成系统的构建与应用.中国棉花,2014,41(6):17-20
- [15] 冯勇,侯旭光.玉米品种试验Excel数字化报表的建立.北方农业学报,2017,45(5):17-21
- [16] 王虎,杨耀华,李绍明,刘哲,许哲,张晓东,朱德海.基于移动端作物大田测试数据采集技术研究与实现.中国农业科技导报,2013,15(4):156-162
- [17] 叶思菁,朱德海,姚晓闯,岳彦利,黄健熙,李林.基于移动GIS的作物种植环境数据采集技术.农业机械学报,2015,46(9):325-334
- [18] 王虎,李绍明,刘哲,张晓东,朱俊.作物品种试验数据预处理系统的设计与实现.中国农业科技导报,2010,12(2):138-144
- [19] 杨付新,付小琼,王秀玲,周关印.农作物品种区域试验数据分析系统专用软件在国家棉花品种区试中的应用.农业图书情报学刊,2000(2):24-27
- [20] 胡伟娟,傅向东,陈凡,杨维才.新一代植物表型组学的发展之路.植物学报,2019,54:558-568
- [21] 赵春江.植物表型组学大数据及其研究进展.农业大数据学报,2019,1(2):5-18

(收稿日期:2022-05-04)