

李 慧, 马德新. 农业大数据应用发展现状及其对策研究[J]. 江苏农业科学, 2021, 49(16): 48–52.
doi:10.15889/j.issn.1002-1302.2021.16.008

农业大数据应用发展现状及其对策研究

李 慧¹, 马德新^{1,2,3}

[1. 青岛农业大学动漫与传媒学院, 山东青岛 266109; 2. 齐鲁工业大学(山东省科学院)/
山东省计算中心(国家超级计算济南中心), 山东济南 250014; 3. 山东大学计算机科学与技术学院, 山东青岛 266237]

摘要: 农业精准高效发展是我国现代化发展的重要战略目标, 是新时期进行农业新旧动能转化及实施乡村振兴战略的必然要求。农业领域具有较大的数据基础, 为农业大数据发展提供了无尽源泉, 且随着智能信息技术广泛在农业上的应用, 利用大数据在农业方面进行相关研究, 具有重要的现实意义。为此, 使用文献调查法, 阐述大数据的发展背景与研究现状, 及其关键技术, 并分析大数据在农业上的需求、主要应用在哪些领域及其在农业方面发挥的重要作用。同时针对我国大数据在农业的应用发展中存在的问题, 提出对策建议, 以期能为农业大数据的发展提供参考依据。

关键词: 智慧农业; 大数据; 应用领域; 云计算; 对策

中图分类号: S126 **文献标志码:** A **文章编号:** 1002-1302(2021)16-0048-05

“三农”问题始终是民生大计, 是我国经济的基础, 中央一号文件提出乡村振兴战略, 将解决“三农”问题作为政府工作的重中之重。当前, 我国农业正处于传统农业向现代化农业转型发展的关键期与过渡期, 农业发展面临着诸多挑战, 包括农产品价格与生产成本、农业产业结构、农业资源环境、农业发展质量等一系列问题制约着农业产业的进步。随着每年用新的信息技术收集有关农作物和环境及其相互作用的信息, 农业数据正在迅速增长并趋向于大数据的规模, 迫切需要其他强大的工具来计算分析和处理^[1]。所以须要将传统农业与大

数据、物联网、云计算等相关网络信息技术进行有机结合, 促进传统农业产业向现代化、智慧化、高效化转型发展。

现代化信息科学技术代表着当前生产力的前进方向, 随着社会的发展, 其在不断的改良与更新, 向着智能化发展并逐渐应用到各行各业^[2]。大数据技术、农业物联网技术、精细农作技术、信息服务技术等农业中的应用范围越来越大, 组成了我国农业信息化技术^[3]。大数据技术可以应用在农业物联网、云计算等技术中, 用来更全面地分析获得有价值的农业数据信息。中美贸易战的持续升级对我国农业具有较大的冲击力, 所以要充分发挥农业大数据的作用, 加强大数据在农业各领域方面的应用, 进一步推进农业向着稳定化、智慧化、现代化的方向发展。

1 研究现状分析

1.1 农业大数据

随着大数据在商业及一些互联网产业应用的

收稿日期: 2020-12-30

基金项目: 山东省重点研发计划(编号: 2019GNC106001); 青岛市民生科技计划(编号: 18-6-1-112-nsh); 山东省高等学校科技计划(编号: J17KA154); 淄博市重点研发计划(编号: 2019gy010101)。

作者简介: 李 慧(1997—), 女, 山东淄博人, 硕士研究生, 研究方向为农业大数据。E-mail: 237900605@qq.com。

通信作者: 马德新, 博士, 教授, 研究方向为农业大数据及智能控制系统。E-mail: madexin@163.com。

[123] 曾文丹, 严华兵, 谢向誉, 等. 木薯间作套种不同作物栽培模式及经济效益研究概况[J]. 农学学报, 2016, 6(12): 11–15.

[124] 吕丰娟, 肖运萍, 汪瑞清, 等. 15 个芝麻品种(品系)红壤连作耐性分析[J]. 湖南农业大学学报(自然科学版), 2018, 44(2): 130–135.

[125] 任永权, 张国辉, 周江菊. 土壤热处理对连作太子参产量的影响[J]. 北方园艺, 2016(7): 140–143.

[126] 王 峰, 吕 艺, 刘 洋, 等. 氯化苦土壤消毒对三七种植后细

菌群多样性及烤烟轮作的影响[J]. 昆明理工大学学报(自然科学版), 2019, 44(4): 90–96.

[127] Berendsen R L, Pieterse C M, Bakker P A. The rhizosphere microbiome and plant health[J]. Trends in Plant Science, 2012, 17(8): 478–486.

[128] Zhang R F, Vivanco J M, Shen Q R. The unseen rhizosphere root-soil-microbe interactions for crop production[J]. Current Opinion in Microbiology, 2017, 37: 8–14.

成功,其他行业也纷纷采用大数据解决行业所存在的问题,大数据也随之在农业领域进行应用。农业大数据是从大数据领域中延伸出来的一个分支,通常指的就是将大数据的技术、思想、相关理念应用在农业中^[4]。所以说农业大数据通常也具有以下“5V”特性:数据量巨大(volume)、准确性高(veracity)、计算速度快(velocity)、数据类型多(variety)、价值量大(value)^[4-5]。大数据可以应用在农业物联网与智慧农业上,使农业更加智慧化、精准化、高效化、批量化,让农业产生更高的经济附加值,带给农业智能型的业态新模式。更进一方面来说,在农业中应用大数据技术结合其他信息技术,能够实现对农业海量基础数据的查询、处理、计算、存储、共享等功能,全面获得更多有价值的信息,为现代化农业的可持续发展提供重要指导依据^[6]。因此,大数据在农业方面具有很大的利用价值,有必要加深其在农业发展中的理论研究与实践探索。

1.2 农业大数据国外研究现状

2009 年奥巴马政府将农业、商业、教育、财富、消费等一系列数据通过平台进行公开,在此平台上可利用这些数据进行科研研究、开发设计等操作^[7]。2013 年 5 月,八国集团(G8)领导人在国际论坛上进行了讨论,提出了公布或共享农业数据库的方案,这为农业数据的应用发展提供了平台支持^[8]。2013 年英国政府颁发了《英国农业技术战略》,其中特别提到了大数据,利用大数据和信息技术促进农业产业的创新发展,打造英国现代农业产业优势^[9]。近年来,随着经济的快速发展,大数据也随之不断发展,一些国家将大数据的研究发展作为国家战略,并且国外的一些学者在农业大数据方面取得了重要的研究成果,他们将农业大数据与精准农业结合使用,利用大数据高效发展智慧农业,已经有较多的成功案例,产生了更高的经济效益^[8]。美国公司 The Climate Corporation 收集了大量的气象因子数据、作物生存环境因子数据及作物生长本质特征因素等,通过大数据等相关信息技术为农民提供农作物自然灾害保险服务。

1.3 农业大数据国内研究现状

我国是传统的农业大国,人口众多,农业现代化、信息化发展对我国具有较大的现实意义。从 2013 年开始至今,大数据被广泛应用在农业、医疗、交通、旅游、军事等诸多方面,并取得了一定成果。

针对大数据在农业领域的应用,国内外相关学者进行了研究探索,并提出了相关创新技术,截至 2019 年 6 月 4 日,在中国知网上检索关键词“大数据 农业”,2013—2018 年期刊、会议、报纸、博硕论文的相关文章数量逐渐增长,上升趋势稳定。

2 农业大数据技术

2.1 农业大数据技术应用体系架构

将大数据及其相关技术加入农业领域,实现对农业数据的收集、存储、计算、应用,并提供信息挖掘是传统农业向智慧农业转型的关键,农业数据挖掘是指从庞大的农业基础数据源中提取有用的信息,应用到农业生产发展服务的过程^[10]。本研究通过分析大数据在农业方面的应用过程,并结合国内外相关专家的研究分析,将农业大数据技术应用体系分为以下 4 个部分:数据资源基础层、数据技术处理层、用户应用层、平台管理层(图 1)。数据资源基础层主要包括作物本身相关信息及环境因子数据,数据技术处理层主要是运用大数据、云计算等相关技术进行分析处理,它们都属于农业大数据的基础投入部分。用户应用层属于农业大数据技术的可视化产物,方便用户使用操作,属于大数据的产出部分。平台管理层主要针对管理人员,通过设计、处理、优化和建设各应用系统,使服务更加的智能化和智慧化^[11]。

2.2 农业大数据关键处理技术

2.2.1 农业大数据采集 农业生产具有海量的数据,数据采集是农业大数据后续处理计算工作的前提,是一项关键的基础技术^[12]。互联网、物联网、大数据之间是层层递进的关系,物联网主要采用射频识别(RFID)技术与传感器网络等进行数据收集,而大数据则在此基础上通过移动网络数据及社交网络交互数据和政府部门所公开的数据平台,具有更大的广泛性^[12-14]。农业大数据主要有原本已有数据接入、数据实时采集接入及农业相关文件、音频、视频、图片、消息等常规数据接入等采集方式,以及 Kafka、ActiveMQ、ZeroMQ、Flume、Sqoop、Socket(Mina、Netty)、ftp/sftp 等常用技术^[15-16]。所以农业大数据采集信息丰富且具有时效性,并且不受时空地域的影响。

2.2.2 农业大数据预处理 农业大数据的数据类型极其复杂,属于结构化数据、半结构化数据和非结构化数据的集合,所以须对数据进行预处理^[11]。

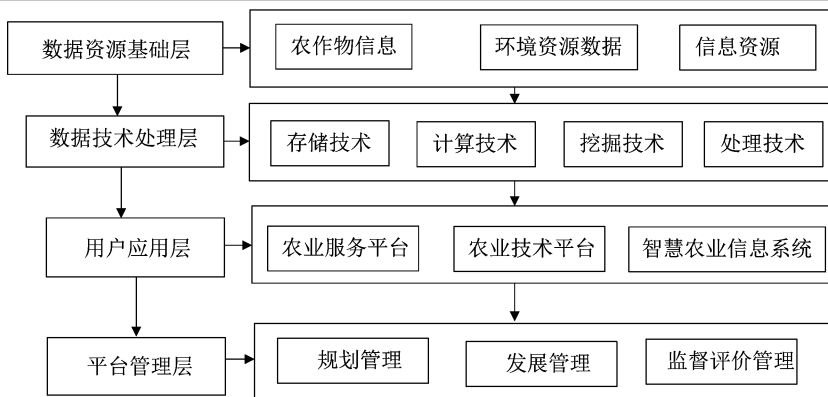


图1 农业大数据技术应用体系框架

预处理就是通过关联与聚合进行数据结构整合,采用同一方式进行数据存储处理。在对农业相关数据进行处理时首先要剔除一些没有价值的信息,排除错误数据的干扰,通过去噪技术来获取相关数据^[17]。大数据处理技术在不断发展,先后出现了基于XML技术、CORBA、P2P技术及Web Services技术的数据集成等^[18-20]。

2.2.3 农业大数据管理与模型构造 农业大数据管理主要运用Hadoop、Map reduce、HBase技术构架技术处理平台,研发大数据农业管理系统^[21]。通过管理农业大数据平台对大数据进行分析处理和展示,使其更好地应用在农业生产的各个方面。建设农业大数据相关模型是运用大数据解决农业问题的主要方法,国际标准化组织构建了网络开放参考模型(open system interconnecton, OSI)^[22], workflow管理联盟构建了工作流参考模型(work flow magement coalition, WFMC)^[23]。结合实际农业生产的应用,杨波等结合大数据技术对山东省二代玉米螟发生程度构建了预测模型,通过试验对比分析其预测结果明显高于其他同类预测^[24]。肖丽平等基于大数据对农业品牌信息数据集模型进行了研究,结果表明通过构建农业大数据模型可以有效解决大数据背景下农业品牌化的信息服务问题^[25]。赵志明等通过借鉴国外大数据参考模型,结合我国农业生产实际情况提出了农业大数据参考模型^[26],为我国农业大数据发展提供了参考依据。

2.2.4 农业大数据计算分析 农业大数据计算分析主要与云计算技术相结合使用,云计算^[27]可以调用虚拟化数据资源平台,通过农业大数据计算对数据进行挖掘分析,使得农业基础数据彰显出巨大的使用价值,达到最优化的利用。计算分析可实现农业精准化、智能化与可视化,并可与电脑端或手机

端结合进行远程操作。

3 农业大数据应用研究分析

由于时间和空间的差异,农业生产具有丰富多样的类型,并且具有较大的农业基础数据,适合大数据在农业领域方面的应用^[12]。本研究在中国知网上检索关键词“大数据 农业”,对引载量排名前10的文章进行分析,结果见表1。由表1可知,在农业大数据相关文献中综述类文章较多,应当加强大数据在农业生产各个方面和环节的应用,提高农业大数据架构平台的建设数量,增大其在农业方面发挥的作用。结合相关文献进行整理,当前大数据在农业的主要应用见表2。

表1 农业大数据相关文献按被引量 and 下载量的各前20名分析

类别	被引量文章数 (篇)	下载量文章数 (篇)
架构平台建设	4	4
农业各领域应用研究	6	7
发展综述	10	9

表2 农业大数据平台应用的主要重点领域及优势

应用领域	相比较于传统农业生产方式的优势
农产品市场监测预警 ^[28]	分析决策性高
农业生产决策 ^[29]	智慧精准决策
农业栽培育种 ^[14]	时间短、准确度高、成活率高
质量安全追溯 ^[31]	安全性高、保障性好
农业病虫害预警 ^[32]	预防、监控和干涉,提高质量
农业遥感气象 ^[33]	精准分析、提高效率
农业综合信息服务 ^[29]	整合共享资源,提供综合服务功能

在大数据技术的带动下,农产品市场监测预警的工作模式和方式发生了巨大的改变,它所分析的研究对象及内容变得更加智能化和精准化^[28,34]。

精准农业生产决策的执行离不开大数据技术的支持,有学者使用 C4.5 算法分别建立雨季和旱季的决策树模型,并且在泰国进行试验研究取得了成功的经验^[29]。国内学者齐鲁根据寿光地区土壤环境状况,使用企业级 WebGIS 构建 ArcGIS Server 平台供农户和企业使用,极大地提高了农业生产决策效果^[30]。农业大数据结合栽培育种,通过监测追踪种子的种植、栽培、生长状况、环境因子等因素为农民提供相关信息,有助于提高农业栽培效率和农作物产量^[14]。农产品质量安全追溯对我们来说极为重要,通过大数据技术可以追溯农产品从生产到加工的过程,保证农产品安全^[31]。通过大数据技术进行农作物生长环境及自身因素分析,并且改进病虫害诊断预警的数据挖掘方法,最终构建出农作物病虫害预警系统,这主要运用决策树算法在 Hadoop 中的并行化计算^[32]。作为遥感科学的分支,农业大数据同遥感技术相结合,可以广泛地应用在农业领域,如灾害、资源环境、作物表型、农业定量等方面^[33]。在农业综合信息服务平台方面,山东农业大学建立了农业大数据应用云平台,可以发布多种共享数据,提供综合农业服务^[29]。

4 农业大数据应用存在的问题及解决对策

4.1 农业大数据应用存在的问题

4.1.1 异构数据源处理技术 在农业生产中数据的类型不仅包括文本、音频等多媒体数据资源,同时还有许多种植经验、农业生产经营模式等非客观因素,这其中缺乏相应标准及数据结构,所以在数据查询和存储方面存在一定难度^[35]。同时,农业生产受环境影响因素较大,数据存在短时的波动性和不稳定性,且有效性偏低。

4.1.2 共享技术支撑存在不足 在数据格式、数据类型、存储类型、存储标准等方面存在较多的问题^[36]。第一是缺乏相应共享标准,在农业生产的各个领域针对农业多源异构、结构化、半结构化数据的标准化技术和海量数据管理的模型技术相对较少。第二是农业大数据平台较少,且存在的大数据平台也存在技术良莠不齐、无法相互沟通共享等问题。

4.1.3 农业大数据流存在错位问题 在实际生产过程中,由于主观理念及客观因素影响,农业大数据存在数据真实性、数据分类、数据共享和数据决策问题,可能会形成农业产业闭合回路^[37]。如果存

在人为操作失误就可能会出现数据失真、错位等问题。

4.1.4 农业大数据推广难度大 首先,与农业大数据相关的手机应用不多,而且实现农业大数据的可视化操作技术较为复杂。其次,农民学习农业大数据技术的能力不足,知识储备不多,须要进行一定的培训与长时间的适应。再次,农业大数据相关设备成本较高。最后,相关农业推广部门消极怠工,没有充分发挥带头和表率作用。

4.2 解决对策

开放公共数据,加强数据共享顶层设计。应当增大数据的共享性和开放融合的特性,农业相关政府部门可建立统一的数据共性资源平台。从思维层面来说,大数据非常看重数据价值,通过共享数据的互联互通,实现对大量基础数据的收集、存储、分析、挖掘等,打造透明农业和智慧农业。

结合数学建立相关模型。农业生产过程存在农作物、环境、区域等因素的影响,以农业大数据作为实际应用场景,结合数学及农作物自身情况建立适合的模型。通过加强构建模型可以使其有效地应用在某种或某一类作物上,使得作物更好地生长。所以要加强数学建模在农业大数据上的应用。

加大农业大数据基础设施建设,提升农业体系平台应用。获取农业基础数据是农业大数据技术的关键基础,加强大数据基础设施建设可以更好地减少农业数据流错位问题。

加大推广力度,增强农民、企业参与度。应当理论联系实际,通过当地农业推广部门增强本身推广人员的技术能力,引导企业做好大数据基础建设的同时,增大在农业实际生产中的应用,帮助农民在种植过程中更好地享受大数据所带来的好处。

5 结语

农业大数据是当前发展的热点问题,同时也是农业发展的重要应用技术,正确利用大数据技术促进农业智慧发展具有重大的现实意义。虽然当前我国农业大数据迅速发展,但还存在着采集与处理数据难度大、周期长、涉及范围广等不足,结合本研究的研究成果与我国农业大数据发展现状,应当加大对农业大数据关键技术的研究,要有效利用云计算、物联网等新兴技术,与大数据技术的融合发展,针对农业生产的实际,建立相应的大数据应用平台与相关模型设计,有效解决大数据在我国农业发展

应用中存在的问题。同时要充分发挥农业大数据在农业生产经营领域的作用,为智慧农业发展提供发展动力与支撑,以此助力推进传统农业转型升级,加快乡村振兴战略的实施。

参考文献:

- [1] Shrivastava S, Marshall - Colon A. Big data in agriculture and their analyses[J]. Encyclopedia of Food Security and Sustainability, 2019, 1:233 - 237.
- [2] 孙凌山. 智能化信号信息处理技术的发展与展望[J]. 数字通信世界, 2019(2): 58.
- [3] 张 泽, 吕 新, 侯彤瑜. 浅析中国农业信息化技术发展现状及存在的问题[J]. 教育教学论坛, 2019(14): 267 - 268.
- [4] Gandomi A, Haider M. Beyond the hype: big data concepts, methods, and analytics[J]. International Journal of Information Management, 2015, 35, 137 - 144.
- [5] Kamilaris A, Kartakoullis A, Prenafeta - Boldú, et al. A review on the practice of big data analysis in agriculture [J]. Computers and Electronics in Agriculture, 2017, 143: 23 - 37.
- [6] 梁柱福. 大数据在农业物联网中的应用[J]. 电子技术与软件工程, 2019(9): 156.
- [7] 郭雷风. 面向农业领域的大数据关键技术研究[D]. 北京: 中国农业科学院, 2016.
- [8] Mayer - Schönberger V, Cukier K. Big data: a revolution that will transform how we live, work, and think [M]. Boston: Houghton Mifflin Harcourt, 2013.
- [9] Government U K. A UK strategy for agricultural technologies [Z]. 2013.
- [10] 宋长青, 温孚江, 李俊清, 等. 农业大数据研究应用进展与展望[J]. 农业与技术, 2018, 38(22): 153 - 156.
- [11] 孟祥宝, 谢秋波, 刘海峰, 等. 农业大数据应用体系架构和平台建设[J]. 广东农业科学, 2014, 41(14): 173 - 178.
- [12] 张浩然, 李中良, 邹腾飞, 等. 农业大数据综述[J]. 计算机科学, 2014, 41(11A): 387 - 392.
- [13] 赵 冰, 毛克彪, 蔡玉林, 等. 农业大数据关键技术及应用进展[J]. 中国农业信息, 2018, 30(6): 25 - 34.
- [14] 孙晓勇, 刘子玮, 孙 涛, 等. 大数据在农业研究领域中的应用与发展[J]. 中国蔬菜, 2015(10): 1 - 5.
- [15] Romanowski A. Big data - driven contextual processing methods for electrical capacitance tomography [J]. IEEE Transactions on Industrial Informatics, 2019, 15(3): 1609 - 1618.
- [16] Bendre M R, Thool R C, Thool V R. Big data in precision agriculture: weather forecasting for future farming [C]//2015 1st International Conference on Next Generation Computing Technologies (NGCT), 2015: 744 - 750.
- [17] 孙忠富, 杜克明, 郑飞翔, 等. 大数据在智慧农业中研究与应用展望[J]. 中国农业科技导报, 2013, 15(6): 63 - 71.
- [18] Lehti P, Fankhauser P. XML data integration with OWL: experiences and challenges [C]//2004 International Symposium on Applications and the Internet, 2004: 160 - 167.
- [19] Passi K, Lane L, Madria S, et al. A model for XML schema integration [J]. Lecture Notes in Computer Science, 2002, 2455: 193 - 202.
- [20] Ng W S, Ooi B C, Tan K L, et al. PeerDB: a P2P - based system for distributed data sharing [C]//Proceedings 19th International Conference on Data Engineering (Cat. No. 03CH37405). IEEE, 2003.
- [21] 郭 平, 刘 波, 沈 岳. 农业云大数据自组织推送关键技术综述[J]. 软件, 2013, 34(3): 1 - 6.
- [22] International Organization for Standardization, International Electrotechnical Commission. Information technology - open systems interconnection - basic reference model: ISO - IEC - 7498 - 1 [S]. Geneva: ISO, 1996.
- [23] WFMC——工作流参考模型[Z/OL]. 张敬波, 韩 伟, 译. https://download.csdn.net/download/weixin_41718123/10581529?utm_source=iteye_new.
- [24] 杨 波, 刘 勇, 牟少敏, 等. 大数据背景下山东省二代玉米螟发生程度预测模型的构建[J]. 计算机研究与发展, 2014, 51(增刊 II): 160 - 165.
- [25] 肖丽平, 胡 春, 王学东. 基于大数据的农业品牌信息数据集模型研究[J]. 情报科学, 2019, 37(5): 13 - 18.
- [26] 赵志明, 廖晓锋, 王晓玲, 等. 农业大数据基础设施开发的参考模型方法[J]. 华东师范大学学报(自然科学版), 2019(2): 77 - 96.
- [27] Vaquero L M, Roderomerino L, Caceres J, et al. A break in the clouds: towards a cloud definition [J]. ACM Sigcomm Computer Communication Review, 2008, 39(1): 50 - 55.
- [28] 许世卫. 农业大数据与农产品监测预警[J]. 中国农业科技导报, 2014, 16(5): 14 - 20.
- [29] 王丽娟, 信丽媛, 贾宝红, 等. 农业大数据平台的研究进展与应用现状[J]. 天津农业科学, 2018, 24(10): 10 - 12, 21.
- [30] 齐 鲁. 基于 WebGIS 的寿光蔬菜安排生产决策支持系统的设计与实现[J]. 长江蔬菜, 2008(10): 82 - 85.
- [31] 郑开涛, 刘世洪, 胡海燕. 农产品质量安全溯源多边平台的追溯机制研究[J]. 江苏农业科学, 2018, 46(10): 221 - 223.
- [32] 杜 涛. 面向柑橘病虫害预警的 Hadoop 数据挖掘技术研究[D]. 南昌: 华东交通大学, 2018.
- [33] 唐华俊. 农业遥感研究进展与展望[J]. 农学学报, 2018, 8(1): 167 - 171.
- [34] 张保岩, 赵丽华, 辛北军. 天津市“三农”大数据平台构建[J]. 江苏农业科学, 2020, 48(17): 245 - 249.
- [35] 王佳方. 大数据在现代化农业发展中的应用[J]. 开放导报, 2018(4): 78 - 81.
- [36] 王东杰, 李哲敏, 张建华, 等. 农业大数据共享现状分析与对策研究[J]. 中国农业科技导报, 2016, 18(3): 1 - 6.
- [37] 唐光海, 曾 慧. 农业大数据应用需求与数据来源错位问题及其对策研究[J]. 安徽农业科学, 2018, 46(9): 174 - 176.