

基于中国知网的农业研究热点分析

张小娟, 张永恒 (榆林学院信息工程学院, 陕西榆林 719000)

摘要 以中国知网数据库中收录的核心期刊文献题录为数据基础, 借助文献计量学理论方法, 以信息分析软件 SPSS 为工具, 定量分析了农业领域的 839 108 篇学术论文, 绘制了国内农业数据的知识图谱, 揭示了农业领域研究的学科领军人物、核心研究机构、研究热点等。

关键词 农业数据; 研究热点; 数据分析; 知识图谱

中图分类号 S-058 **文献标识码** A **文章编号** 0517-6611(2018)20-0193-03

Analysis of Agricultural Hot Domains Based on CNKI

ZHANG Xiao-juan, ZHANG Yong-heng (School of Information Engineering, Yulin University, Yulin, Shaanxi 719000)

Abstract This paper used information analysis software SPSS, by means of literature metrology method, quantitative analysis of 839 108 agriculture academic papers in CNKI database, drawing the domestic agricultural knowledge mapping data, revealed discipline leader, core research institutions, hot domains of agricultural research field.

Key words Agriculture data; Hot domain; Data analysis; Knowledge spectrum

农业是支撑国民经济建设与发展的基础产业, 作为农业大国的我国一直非常重视农业的发展^[1]。21 世纪是农业发展的重要阶段, 生命科学和其他最新科学技术相结合, 将使世界农业发生根本性变化。随着分子生物学的发展, 生物基因库的建成, 遗传工程的崛起, 克隆技术的广泛应用, 农业面貌将焕然一新^[2]。

国内农业领域的学者专家在各种学术期刊上发表了大量的研究论文, 产生了大量有价值的农业研究数据。笔者借助 Excel、SPSS 软件, 对这些海量的农业研究数据进行定量统计分析, 找出农业领域的研究热点、发展轨迹、学科领军机构及人物等特征, 为该领域相关研究人员追踪学科发展方向提供参考。

1 研究方法

文献计量学是对文献和文献工作进行定量研究的方法, 它采用数学语言进行定量分析, 以数学公式或图形来表现规律, 对问题的陈述及计算都采用简明的数学符号, 从而简化和加速思维过程, 具有深刻的描述性能和高度的概括能力^[3]。文献计量方法在学科发展分析、研究范式的演化等方面得到了广泛的应用, 能够帮助探索不同学科领域的研究主题、研究热点及其演变发展规律^[4]。

笔者利用文献计量学的理论及方法, 借助 SPSS 统计软件对农业领域的海量数据进行了定量统计分析。

2 数据来源及处理

2.1 数据来源 中国知网(CNKI)是全球最大的知识门户网站, 收录文献范围全, 是集期刊、博士硕士学位论文、报纸、会议论文、年鉴、图书、多媒体教育教学素材为一体的知识服务网站。中国知网数据库中的期刊数据库是目前世界上最大的连续动态更新的中国期刊全文数据库, 目前收录 7 600 多

种重要期刊, 内容覆盖自然科学、工程技术、农业、哲学、医学、人文社会科学等各个领域。

该研究统计分析的原始数据源自中国知网期刊数据库。对中国知网收录的农业相关文献 839 108 篇核心期刊文献进行统计分析, 这些数据具有较高的准确性和权威性, 是有价值的研究数据。核心期刊是指某学科(或某领域)的核心期刊, 是指那些发表该学科(或该领域)论文较多、使用率(含被引率、摘转率和流通率)较高、学术影响较大的期刊^[5]。由于核心期刊是期刊中学术水平较高的刊物, 对核心期刊中收录的农业领域研究文献进行检索分析, 可以更准确地获得农业领域的研究热点。

2.2 数据获取及预处理

2.2.1 数据检索。 进入中国知网(CNKI)期刊数据库, 在学科专辑导航中选择“农业科技”, 来源类别选择“核心期刊”, 共检索到农业领域的相关核心文献 839 108 篇。

2.2.2 数据预处理。 通过数据检索, 检索出农业领域的核心期刊文献之后, 利用知网的文献题录导出功能, 把文献的题名、作者、单位、关键词等题录数据导出, 将所有的题录数据进行汇总, 共得到 839 108 条题录数据, 对导出的原始数据进行整合, 便于之后对数据分析。

3 数据统计与分析

3.1 文献年度分布 科技论文的主要功能是记录、总结科研成果, 是科学研究的重要手段。作为科学研究成果的科技论文通常在专业刊物上发表, 科技论文数量与质量可以在一定程度上反映学科发展和学术研究状况^[6]。

从图 1 可以看出, 2000 年以前的 10 年, 农业方面的文献数量波动不大, 说明我国农业研究处于一个相对平稳的发展期, 2000—2010 年, 农业文献呈现急剧上升的态势, 说明我国农业研究在这 10 年内极其活跃, 处于农业活动的爆发期, 2010 年以后, 农业活动又逐渐进入平稳发展期。

3.2 基于作者单位统计的机构科研能力排名分析 机构的科研能力可以从所发学术论文, 特别是核心论文这一角度来进行测评^[7]。该研究将导出的题录信息按作者单位进行分

基金项目 陕西省农业科技创新与公关项目(2016NY141); 陕西省教育厅 2018 年度科学研究项目“基于云计算的农业大数据服务平台集成关键技术研究”。

作者简介 张小娟(1981—), 女, 山东淄博人, 讲师, 硕士, 从事信息组织与信息检索、数据分析研究。

收稿日期 2018-02-13

类汇总,进而统计每个科研机构的发文数量,将作者单位按发文数量从多到少进行排序,此排序结果能够从一定程度上反映农业领域各科研机构的科研能力。从图2可以看出,排在前面的科研单位基本上是我国以农学为特色的全国重点高校。前10名科研机构依次为西北农林科技大学、中国农业大学、南京农业大学、华南农业大学、华中农业大学、东北农业大学、沈阳农业大学、山东农业大学、河北农业大学、北京林业大学。



图1 文献年度分布

Fig.1 Annual literature analysis

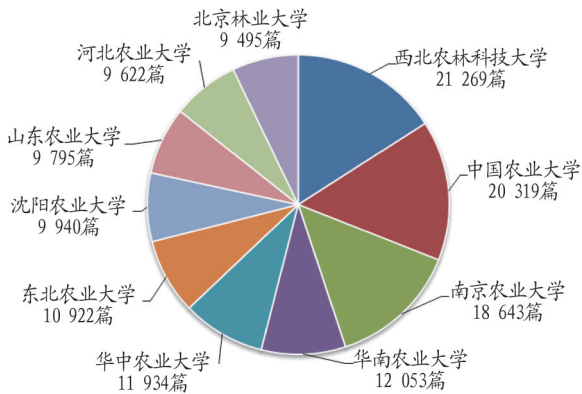


图2 机构发文数量

Fig.2 Published papers by the organization

3.3 基于作者统计的农业领域领军人物分析 统计农业领域作者的发文量,可以分析得出文献作者相对于农学领域的研究贡献,发文量较高的作者一般为该领域的领军人物^[8]。使用获取到的839 108条题录数据,选取每篇文献中的作者,利用Excel软件对文献作者进行排列,按作者汇总统计,进而统计每位作者的出现频次,排序前10名的作者依次为张福锁、李生秀、沈志强、陈国宏、刘秀梵、曹卫星、张玉龙、沈其荣、鲁剑巍、陈新军(表1)。

排名前3位的作者简介如下:张福锁,中国工程院院士,中国农业大学资源环境与粮食安全研究中心主任、教授、民盟中央委员、民盟中央农业委员会副主任、中国农业大学委员会主任、农业部科技委委员、教育部科技委农林学部副主任、全国测土配方施肥技术专家组组长、国家环境特约检察员、北京市人大代表、教育部长江学者特聘教授、国家自然科学基金创新群体和国家科技部973项目首席科学家;李生秀,西北农林科技大学教授、兰州大学兼职教授、博士生导

师;沈志强,山东省滨州畜牧兽医研究院院长兼书记、山东省畜禽蜂胶疫苗研究开发推广中心主任、中国科技开发与人才研究中心研究员、中国农业大学和山东农业大学等多所高校博士、硕士研究生导师。

表1 农业领域高产作者发文数量

Table 1 Published papers produced by high-yield authors in the field of agriculture

序号 No.	作者 Author	发文数量 Published papers	序号 No.	作者 Author	发文数量 Published papers
1	张福锁	336	6	曹卫星	248
2	李生秀	279	7	张玉龙	225
3	沈志强	267	8	沈其荣	224
4	陈国宏	266	9	鲁剑巍	213
5	刘秀梵	254	10	陈新军	213

3.4 基于单项关键词的热点研究领域分析 每篇文献都能够概括文献内容的关键词。关键词通常是著者构思论文内容时经过高度概括产生的^[9]。每篇文献的关键词虽然仅有几个,但可以较充分且较恰当反映出论文的主要内容,通过文献标引的关键词可以较集中地了解到论文的整体概况^[10]。笔者对农业领域的研究文献进行关键词统计分析,进一步得出我国农业热点研究领域。

从图3可以看出,农业热点研究领域主要集中在产量、水稻、玉米、小麦、品质、生长、遗传多样性等方面。文献中关键词出现频次最高的是“产量”,表明农业研究热点还是集中在农作物产量方面;关键词“水稻”“玉米”“小麦”出现频次仅次于“产量”,表明水稻、玉米、小麦这3种主要农作物依然是农业领域的关注点;“品质”“遗传多样性”出现频次也较高,表明人们生活水平的提高,促使农业领域开始关注农作物的品质及其改良。

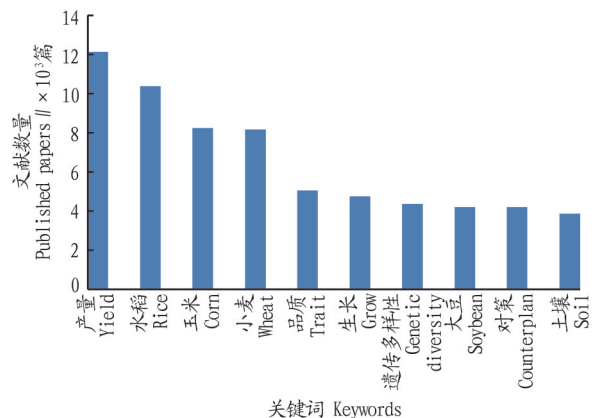


图3 关键词对应文献数量

Fig.3 Published papers corresponding to keywords

3.5 基于关键词共现矩阵的农业领域研究热点分析 学术论文的关键词是论文内容的提示符,是作者学术思想及学术观点的凝结^[11]。2个或更多关键词在同一篇文献中同时出现称为关键词共现^[12]。共现矩阵可反映研究领域的某些方向,进行关键词共现矩阵统计分析可得出农业领域的研究热点。

对“3.4”中所统计的前20个高频关键词进行共词分析,

列出了关键词共现矩阵,进一步绘制出关键词共现网络图。

图 4 为统计所得的关键词共现网络图。

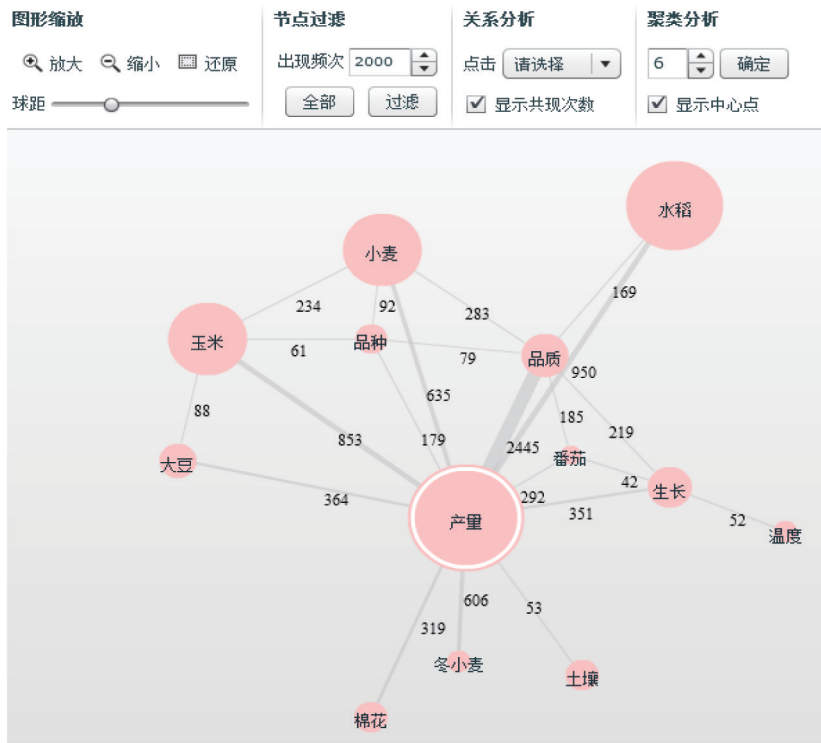


图 4 关键词共现网络

Fig.4 Keywords co-occurrence network

通过关键词共现网络图,可以获得农学领域的热点研究方向;关键词共现次数较高的有“水稻”与“产量”、“玉米”与“产量”、“小麦”与“产量”、“克隆”与“序列分析”等,由此可推出:水稻、玉米、小麦这 3 种传统主农作物的产量问题一直是农业的关注热点,并且由图 4 可知,冬小麦的产量也是小麦产量问题中的重点研究方向;如何将克隆技术更好地应用在农业中,研究热点主要集中在克隆技术的序列分析上。由关键词共现矩阵还可得出:农业领域除了关注三大主农作物产量外,大豆、棉花、番茄等农作物产量问题、小麦品质问题、番茄品质问题、品种对农作物产量的影响等也成为农业领域的关注点。

4 小结

通过对农业领域的 839 108 篇学术文献的计量学实证研究得出:我国农业发展呈现 10 年一个阶段的特点;农业领域的主要研究机构集中于我国的重点本科院校;研究领域围绕农作物产量展开,研究方向包括水稻、玉米、小麦这 3 种传统主农作物产量问题,大豆、棉花、番茄等农作物产量问题、小麦品质问题、番茄品质问题、品种对农作物产量的影响;如何将克隆技术更好地应用在农业中。

参考文献

- [1] 祁豪荣,孙强强,刘明,等.基于知识图谱和 PSR 框架的北京区域土壤重金属研究现状分析[J].中国农业大学学报,2017,22(12):211-220.
- [2] 吴靖瑶,吴克宁,李晨曦.中国土地评价研究知识图谱:基于 CiteSpace 的计量分析[J].土壤通报,2017(6):1513-1522.
- [3] 谭雪兰,欧阳巧玲,于思远,等.基于 CiteSpace 中国乡村功能研究的知识图谱分析[J].经济地理,2017,37(10):181-187.
- [4] 杨景宁,沈禹颖,王彦荣.基于文献计量的国际草地农业生态学研究发展态势分析[J].草业学报,2017,26(5):224-233.
- [5] 师丽娟,杨敏丽.基于科学计量的中国农业工程研究热点探析[J].农业工程学报,2016,32(S2):430-438.
- [6] 闫鹏,陈源泉,隋鹏.基于文献计量的华北地区农业用水问题研究态势分析[J].中国农业大学学报,2016,21(9):206-214.
- [7] 邱均平,温芳芳.近年来图书情报学研究热点与前沿的可视化分析:基于 13 种高影响力外文源刊的计量研究[J].中国图书馆学报,2011,37(2):51-60.
- [8] 郭祥云,刘杰,徐龙琴,等.基于 Citespace 的我国农业物联网文献计量分析[J].仲恺农业工程学院学报,2017,30(4):32-39,45.
- [9] 刘彬,邓秀新.基于文献计量的园艺学基础研究发展状况分析[J].中国农业科学,2015,48(17):3504-3514.
- [10] 李致庆,胡卉.近年来我国节水灌溉技术研究进展:基于文献计量的科学知识图谱分析[J].节水灌溉,2015(8):104-109.
- [11] 徐良培,陶建平.2000~2013 年中国农业保险问题的研究动态及展望:基于文献计量方法[J].保险研究,2014(4):3-15,62.
- [12] 张会巍,李启正,徐石勇.基于 CiteSpaceⅢ 的我国服装数字化技术文献知识图谱[J].浙江理工大学学报(社会科学版),2016,36(4):354-360.

科技论文写作规范——讨论

着重于研究中新的发现和重要方面,以及从中得出的结论。不必重复在结果中已评述过的资料,也不要模棱两可的语言,或随意扩大范围,讨论与文中无多大关联的内容。