

毛先才,张红燕,刘亚文,等. 基于文献计量的农业信息化研究特征阐释与热点分析[J]. 江苏农业科学,2020,48(15):64-70.
doi:10.15889/j.issn.1002-1302.2020.15.011

基于文献计量的农业信息化研究特征阐释与热点分析

毛先才,张红燕,刘亚文,阳灵燕

(湖南农业大学信息与智能科学技术学院,湖南长沙 410128)

摘要:对国内外农业信息化领域的研究特征与研究热点进行归纳和分析,为我国农业信息化的后续研究提供参考。采用文献计量法和共词聚类法,借助 R 语言编程工具,针对 CNKI 和 Web of Science 数据库中与农业信息化相关文献,分别从时间分布、核心作者、来源期刊、研究热点和研究趋势进行定量统计和定性分析。结果表明,国内研究更关注农业信息服务、农村信息化建设与现代农业发展等,而国际研究更关注与农业生产和管理相关的气候变化监测、土地利用管理、生态服务等。因此,我国农业信息化的研究应更加深入,并更多关注现代信息技术在具体农业生产实践中的应用。

关键词:农业信息化;文献计量;发表时间;核心著者;来源;高频关键词;共词聚类分析;R 语言;研究热点

中图分类号: G250.252;S126 **文献标志码:** A **文章编号:** 1002-1302(2020)15-0064-07

农业信息化是一个涵盖多学科、多领域的综合系统工程^[1],它是应用信息技术对农业生产领域、农业流通领域、农业科技领域、农业教育领域以及农村文化生活领域进行提升和改造的一种活动^[2]。自 2005 年以来,中央一号文件连续 15 次强调农业农村信息化建设,从中不难看出我国对农业信息化建设的重视程度和投入力度,农业信息化也成为众多国内外科研工作者关注的焦点。农业信息化领域的研究文献客观记录了该领域的发展概貌,开展相关文献分析以了解农业信息化领域的研究现状和研究热点,对把握其未来的发展趋势意义重大。目前,已有学者对农业农村信息化领域进行相关研究,如杨杰等基于 CNKI 数据库对国内农村信息化研究热点与主题演进进行可视化分析^[3]。曹利红采用文献计量学方法结合 CiteSpace 软件,分析我国农业信息化领域的研究现状及发展趋势^[4]。桑静利用知识图谱对我国农业信息化研究历史和现状进行全面的梳理^[1]。丁恩俊等基于文献计量对国内外农业信息化领域的发文量、发文机构和作者合作等关系进行分析^[5]。但基于 CNKI 中文数据库和

Web of Science 外文数据库对比分析国内外农业信息化领域研究特征和热点的相关报道却极其鲜见。本试验以国内外农业信息化领域相关文献为研究对象,利用 R 编程工具对文献进行计量分析、高频关键词共词聚类分析,并以图文展示国内和国际农业信息化领域的发展现状、研究热点与研究趋势,以为农业信息化领域科研人员的后续研究提供参考。

1 数据来源与研究方法

1.1 数据来源

国内文献以“农业信息”为主题词,在中国知网 CNKI 中进行检索,选择全部中文数据库,出版年限为 1979—2018 年,去除其中征文启示、学术辑刊等非学术文献,共获得有效文献 16 472 篇。国内核心期刊文献仍以“农业信息”为主题词检索,选择科学引文索引 (Science Citation Index, SCI)、工程索引 (EI)、核心、中文社会科学引文索引 (Chinese Social Sciences Citation Index, CSSCI)、中国科学引文数据库 (Chinese Science Citation Database, CSCD) 期刊作为来源期刊,出版年限为 1992—2018 年 (相关核心期刊论文最早发表于 1992 年),筛选出有效论文 1 828 篇。国外文献以“agricultural information”为主题词,在 Web of Science 数据库中检索,选择其核心合集,出版年设定为 2007—2018 年 (受版权限制),共检索到 12 127 篇。

收稿日期:2019-09-03

基金项目:国家重点研发计划 (编号:2017YFD0301506);湖南省长沙市工业科技特派员项目 (编号:CKS201845)。

作者简介:毛先才 (1992—),男,湖南岳阳人,硕士研究生,主要从事农业信息化研究。E-mail:1018610379@qq.com。

通信作者:张红燕,博士,教授,主要从事农业信息化研究。E-mail: hongyan_zhang@hunau.edu.cn。

1.2 研究方法

文献计量法是以数学、统计学等计量方法,对文献情报的分布结构、数量关系以及变化规律进行定量分析研究的一种方法^[6]。本研究运用 R 语言编程实现文献的汇总、筛选和计量,并从时间分布、核心作者、文献来源 3 个方面对国内外农业信息领域文献的形式特征进行分析。

共词聚类分析法采用聚类的计算方法,对文章中共现的关键词对的关联性进行运算,将关系密切的词聚集归类,进行信息挖掘^[7]。通常词对共现的频次越高,表明它们的关系越密切。因此,共词聚类分析常被用于探索该领域潜在主题和学科间的发展联系及学术趋势演化^[8]。本研究运用 R 语言

编程及 igraph 等软件包,实现关键词的分离、高频关键词的提取、相近关键词的合并与高频关键词的共词聚类分析,并梳理归纳国内外农业信息化领域的研究热点和研究趋势。

2 农业信息化文献的形式特征分析

2.1 文献发表时间分布

时间分布的研究可以间接反映某一学科情况的总体趋势,展现其研究历程与现状。本研究分别对农业信息领域的 CNKI 全部中文文献、CNKI 核心期刊论文、Web of Science 核心集国际文献的年文献量进行统计,并基于统计结果利用 R 软件中的 plot 函数绘制相应的文献发表时间分布图(图 1、图 2)。

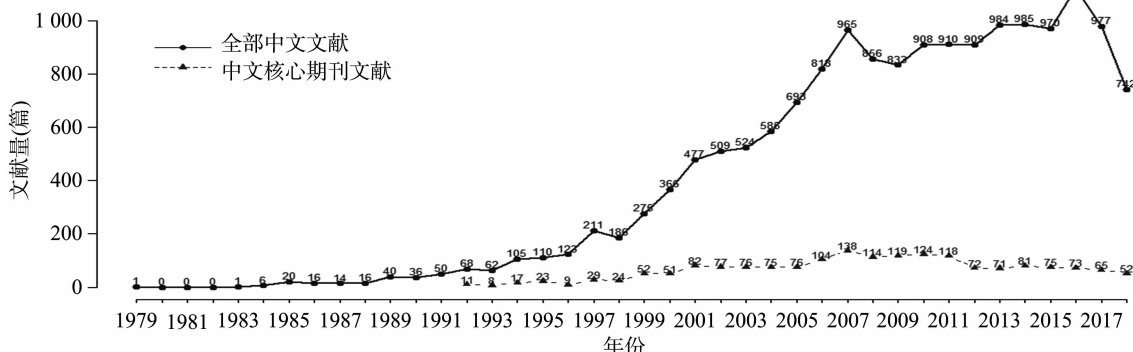


图1 1979—2018 年国内农业信息化领域文献发表情况

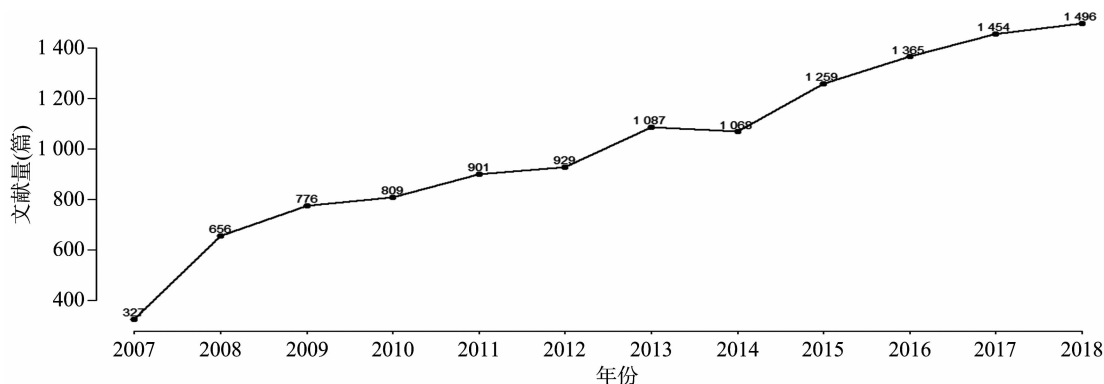


图2 2017—2018 年国际农业信息化领域文献发表情况

由图 1 可知,农业信息化相关研究始于 1979 年,在接下来的 20 年(1979—1998 年)整体发文量较少,但仍呈缓慢增长趋势;1999—2007 年发文量急剧增加,并于 2007 年达到高峰期;2008—2018 年发文量相对恒定。从中文核心期刊论文发表量的时间分布来看,1992 年开始有农业信息化方面的研究论文在核心期刊发表,直至 1998 年,每年发文量均不超过 30 篇。1999—2007 年发文量稳步增长,于 2007 年达到顶峰,到 2012 年发文量又有所回落。其中,2006—2011 年农业信息化领域核心期刊论文

量显著,这与 2005 年中央 1 号文件中首次提出“加强农业信息化建设”不无关系。由图 1 还可知,国内农业信息化研究走过了萌芽期(1979—1998 年)、成长期(1999—2007 年),现在正处于成熟期(2008 至今)。由图 2 可知,2007—2018 年国际农业信息领域的文献量总体呈上升趋势,说明全球对农业信息的关注度越来越高。

2.2 文献的核心著者分析

对农业信息化研究中文文献作者进行统计,结果显示发文量最多的是中国农业科学院农业信息

研究所的许世卫,共发表相关论文 46 篇。根据普赖斯的核心著者最低发文量计算公式 $m \approx 0.749 \sqrt{n_{\max}}$ (其中: n 为发文最多著者的论文发表量, m 为核心著者最低发文量)^[9],可得 $m \approx 5.07$,按向上取整原

则,发文 6 篇及以上者为核心著者。据此统计核心作者人数多达 711 人,由此可见农业信息化已成为学术界关注的热点(表 1)。

表 1 国内农业信息化文献核心著者(前 10 名)

序号	作者	第一作者发文量(篇)	发文量(篇)	作者单位
1	胡绪岚	39	39	云南农业大学
2	许世卫	32	46	中国农业科学院农业信息研究所
3	李道亮	27	42	中国农业大学
4	郭作玉	26	31	农业农村部信息中心
5	刘月姣	24	26	农产品市场周刊
6	梅方权	20	24	中国农业科学院文献信息中心
7	俞菊生	20	21	上海市农业科学院
8	刘小平	19	19	贵州师范大学
9	郑业鲁	17	33	广东省农业科学院
10	常 青	16	19	湖南农业大学

同样,通过对农业信息化研究领域中文核心期刊论文的作者情况进行统计,发现发文量最多的是北京农业信息技术研究中心的李瑾,共发表相关文献 18 篇。同上,可得核心著者最低发文量 $m \approx 3.18$,即发文 4 篇及以上者为核心著者。据此统计核心著者人数仅为 76 人(表 2)。与上述据所有中文文献统计出的核心著者人数相比,数量相对悬殊,从侧面也反映出大多数国内农业信息化科研工作者的研究不够深入、发文质量还有待提高。

表 2 CNKI 中文核心期刊农业信息化论文核心著者(前 10 名)

序号	作者	第一作者发文量(篇)	发文量(篇)	作者单位
1	李 瑾	9	18	北京农业信息技术研究中心
2	刘小平	8	8	贵州师范大学
3	刘丽伟	6	6	中共辽宁省委党校基础教研部
4	常 青	6	16	湖南农业大学
5	孙素芬	6	6	北京市农林科学院
6	王恒玉	5	5	西北师范大学经济管理学院
7	陈熙隆	5	7	四川文理学院
8	左 雄	5	5	四川省气象局
9	王 伟	5	5	东北林业大学机电工程学院
10	蒋 勇	5	7	华中农业大学公共管理学院

通过对农业信息化相关国际文献作者情况进行统计,发现发文量最多的是 Beltsville Agricultural Research Center 的 Martha C Anderson,共发表相关文献 21 篇。计算可得其核心著者最低发文量 $m \approx 3.43$,即发文 4 篇及以上者为核心著者,据此统计全

球核心作者人数为 570 人(表 3)。

2.3 文献来源分析

通过对农业信息化领域中文文献的来源期刊和硕博论文进行统计,发现发文量居多的期刊主要有《农业图书情报学刊》《农业网络信息》《安徽农业科学》《农业信息探索》《农机化研究》《中国农业信息》《农产品市场周刊》《计算机与农业》《现代农业科技》《农业经济》等期刊,而发文量居多的硕博论文主要来自中国农业科学院、湖南农业大学、西北农林科技大学、华中师范大学、南京农业大学等机构。同样对中文核心期刊论文来源的统计结果显示《农业经济》《农机化研究》《世界农业》《情报杂志》《农村经济》等核心期刊发文量居多。

通过对国际文献的来源期刊进行统计,发现发文 100 篇以上的有 9 种期刊,分别是《Science of the Total Environment》《PLoS One、Computers and Electronics in Agriculture》《Agriculture Ecosystems & Environment》《Remote Sensing》《Sustainability》《Environmental Monitoring and Assessment》《Ecological Indicators》《Journal of Environmental Management》《Geoderma》等。

3 农业信息化研究热点可视化分析

3.1 国内农业信息化研究热点可视化分析

某关键词在某学科出现越频繁,表明其受相关学者的关注度越高,也可理解为该学科的研究热点^[10]。通过对农业信息化领域中文核心期刊论文

表 3 国际农业信息化文献核心著者(前 10 名)

序号	作者	第一作者发文量 (篇)	发文量 (篇)	作者单位
1	Martha C Anderson	8	21	Beltsville Agricultural Research Center
2	Fukuda Shinji	6	6	Kyushu University Faculty of Law
3	Luo Zhongkui	5	13	Commonwealth Scientific and Industrial Research Organization
4	Ouyang Wei	5	9	Beijing Normal University
5	Lee Kang Oh	5	7	Kyushu University
6	Vadas Peter A	5	5	USDA Institute of Agriculture
7	Triantafilis J	4	9	University of New South Wales
8	Ortiz Oscar	4	5	International Potato Center
9	Smalling Kelly L	4	5	United States Geological Survey
10	Garbach Kelly	4	5	University Of California

的关键词进行统计分析,筛选出其中频次≥18 的关键词,并剔除“对策”“问题”“综述”等无明确内涵的关键词,共得到 39 个高频关键词(表 4)。

表 4 核心期刊文献高频关键词

序号	关键词	频次	序号	关键词	频次	序号	关键词	频次
1	农业信息化	337	14	农业信息技术	39	27	精准农业	22
2	农业	298	15	农业信息服务	34	28	经济	21
3	农业信息	99	16	农业现代化	31	29	农业科技信息	21
4	信息化	84	17	信息资源	31	30	农业信息工作	21
5	农民	79	18	农业信息系统	30	31	网络	21
6	农村	65	19	美国	27	32	图书馆	20
7	信息服务	65	20	数字农业	27	33	新农村	20
8	农村信息化	59	21	信息化建设	27	34	中华人民共和国	20
9	农业信息资源	57	22	农业产业化	26	35	农业网站	18
10	劳动者	56	23	新农村建设	25	36	市场经济	18
11	现代农业	54	24	信息需求	24	37	数据库	18
12	信息技术	49	25	产业化	23	38	物联网	18
13	农产品	41	26	电子商务	23	39	信息网	18

将 39 个高频关键词两两配对,形成高频关键词对。统计关键词对出现在同一文献中的频次,构建高频关键词共现矩阵。基于共现矩阵,运用 R 软件绘制高频关键词共现网络,并进行聚类分析(图 3)。

从关键词的共现聚类结果来看,39 个关键词共聚为 4 类:第 1 类包含 14 个关键词,其中农业信息资源、农业信息系统、农产品、劳动者等关键词中心度较高。第 2 类包含 9 个关键词,其中信息服务、农村信息化、新农村等关键词中心度较高。第 3 类包含 8 个关键词,其中农业、农业信息、图书馆等关键词中心度较高。第 4 类包含 8 个关键词,其中信息技术、物联网等关键词中心度较高,但这些仅是技术支撑,其具体研究对象中心度较高的则是现代农业和精准农业。

根据关键词频次、聚类结果和中心度情况,再结合具体文献和专业知 识,将国内农业信息化领域的研究热点概括为以下 6 个主要方面:第一,农业信息资源。农业信息资源是农业自然资源和农业经济技术资源的信息化。农业信息资源研究主要围绕天气信息、土壤信息、水分信息、农作物生产信息、病虫害信息等自然信息,以及农产品市场信息、法律法规与科技教育信息等经济技术信息的收集与分析。第二,农业信息服务。农业信息服务研究主要围绕农业信息的传播形式、农业知识库的建立、农业信息服务平台的构建、农业信息咨询和检索服务等。第三,农村信息化。农村信息化是信息技术和通信技术在农村生产、生活和社会管理中实现普遍应用和推广的过程。农村信息化研究主要

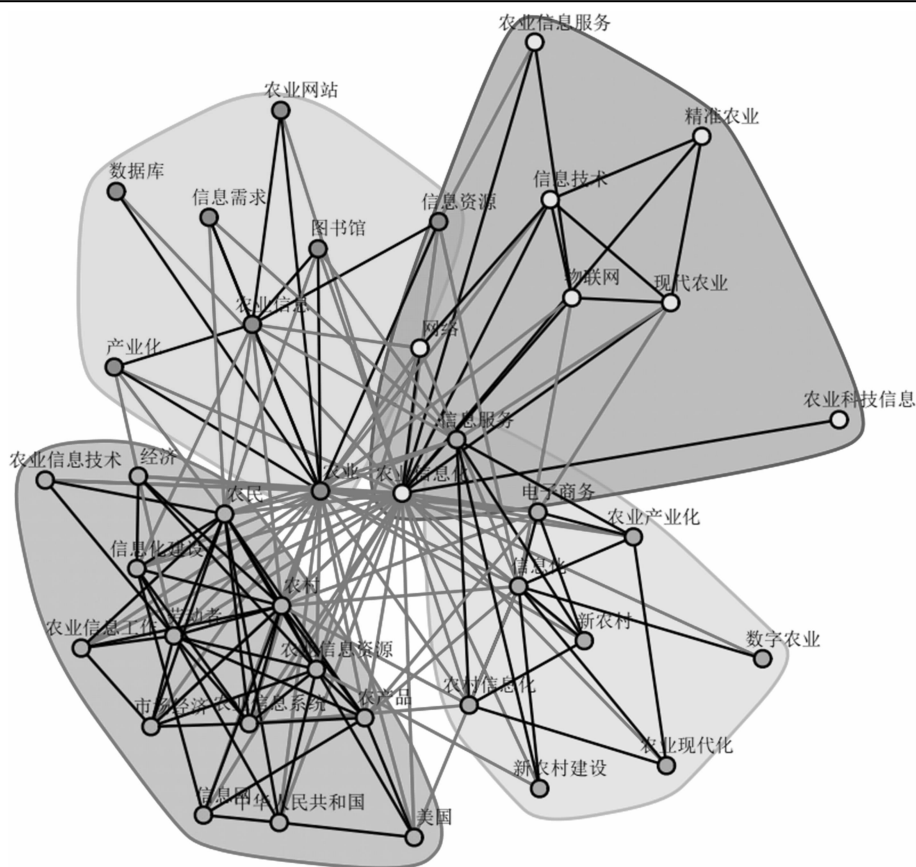


图3 国内农业信息化文献高频关键词共现网络

包括农村电信、互联网等基础设施建设、农村经济文化的新农村建设、农业电子商务及数字农业的推广等。第四,现代农业。现代农业研究主要围绕农业的现代化和产业化展开,一方面是农业设备和技术的现代化;另一方面是农业管理方面的专业化和企业化。另外,还包括现代生态农业和旅游农业的研究。第五,信息技术。信息技术是农业信息化的支柱,国内农业信息技术研究主要包括计算机、互联网、云计算、大数据、物联网等在农业生产、管理、营销等各方面的应用。第六,精准农业。精准农业是以信息技术为支撑,根据空间变异,定位、定时、定量地实施一整套现代化农事操作与管理的系统^[11]。精准农业研究主要包括“3S”技术(GIS、RS、GPS)、农业自动化技术^[12]、互联网技术、物联网技术,以及农业生产中的精准灌溉、精准施肥、精准喷药等精细化操作。

考虑到国内文献跨越年限较长,进一步参照“2.1”节分析得到的3个时期按上述方法分阶段进行高频关键词共词分析以深入了解国内研究热点和研究趋势的变化。结果表明,在萌芽期(1979—1998年),研究主要集中在市场经济体制、农业信息

系统建设、农业信息资源建设等方面;在成长期(1999—2007年),对农业信息资源、现代农业、农村信息化、信息技术、农业电子商务等的关注度较高;在发展成熟期(2008—2018年),农业信息化领域的研究更广,主要包括农业信息服务、智慧农业、精准农业、现代农业、现代信息技术(互联网+、物联网、大数据、云计算)等多个研究领域,呈现出现代化、自动化和多元化的发展趋势。

3.2 国际农业信息化研究热点可视化分析

对 Web of Science 外文数据库中农业信息化领域文献的关键词进行统计分析,并合并其中的同义关键词,如 GIS、geographic information system、geographical information system、land use change、land-use change 等。对合并后的关键词重新进行频次统计,筛选出其中频次 ≥ 60 的关键词(共 30 个)作为高频关键词(表 5)。

同上,利用这 30 个高频关键词构建共现网络,并进行聚类分析(图 4)。

由图4可知,30个关键词共聚为4类:第1类包含9个关键词,其中GIS、soil、water quality等关键词中心度较高。第2类包含8个关键词,其中

表 5 Web of Science 文献关键词

序号	关键词	频次	序号	关键词	频次	序号	关键词	频次
1	GIS	648	11	China	105	21	adaptation	85
2	agriculture	506	12	conservation	100	22	heavy metals	80
3	remote sensing	347	13	irrigation	99	23	precision agriculture	79
4	climate change	239	14	drought	91	24	modis	75
5	land use	217	15	ecosystem services	91	25	risk assessment	74
6	pesticides	162	16	soil erosion	90	26	uncertainty	73
7	water quality	154	17	sustainability	90	27	nitrate	71
8	land use change	121	18	food security	89	28	nitrogen	71
9	biodiversity	115	19	soil moisture	88	29	classification	68
10	groundwater	112	20	soil	86	30	evapotranspiration	64

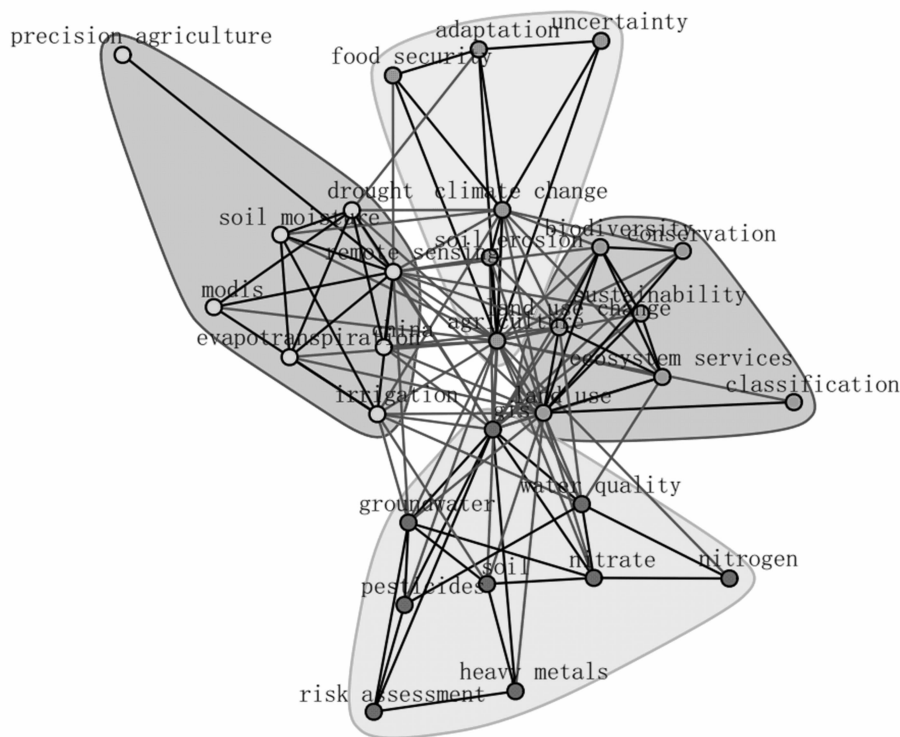


图4 国际农业信息化文献高频关键词共现网络

remote sensing、drought、irrigation 等关键词中心度较高。第 3 类包含 6 个关键词,其中 climate change、soil erosion、food security 等关键词中心度较高。第 4 类包含 7 个关键词,其中 land use、biodiversity、ecosystem services 等关键词中心度较高。

依据关键词的频次、共词聚类结果及中心度情况,结合具体文献和专业知 识,将国际农业信息化领域的研究热点概括为以下 6 个主要方面:第一,地理信息系统(Graphic Information System, GIS)。GIS 是用于采集、存储、管理、分析、表达空间数据的信息系统,农业信息化领域中的 GIS 研究主要包括 GIS 在植物病虫害预警中的应用, GIS 在地下水、土

地利用管理中的应用^[13], GIS 在水质、土壤重金属污染等环境监测中的应用等。第二,遥感技术(remote sensing)。农业遥感主要研究利用遥感技术监测土地利用情况、土壤干旱情况、作物蒸散量和需水量以及农作物的生长情况等,并结合 GIS 和 GPS 为精准农业服务。第三,土地利用(land - use)。土地利用是人类对土地进行有计划的利用,相关研究主要包括土地利用变化监测、土地分类、生态系统服务等。第四,生物多样性(biodiversity)。生物多样性是生物及其环境形成的生态复合体以及与此相关的各种生态过程的综合^[14],其研究主要包括生物多样性保护、水土保持、气候变化调节、生

态系统稳定等。第五,气候变化(climate - change)。通常以不同时期的温度、降水量等气候要素的差异来反映气候的变化。在农业信息领域主要研究气候变化对农业生产的影响,并通过建立相关模型进行预测分析。第六,粮食保障(food security)。该方向的研究主要包括与遥感技术结合的粮食产量估计,气候变化对粮食保障的影响等。

4 结论

本试验选取 CNKI 中文数据库和 SCI 外文数据库中与农业信息化相关的国内外文献作为研究对象,采用文献计量学和共词聚类分析法,借助 R 语言编程工具,从时间分布、核心作者、来源期刊、研究热点对国内和国际农业信息化领域的研究进行统计与分析。首先,从文献量的时间分布来看,国内农业信息化领域的研究经历了萌芽期、成长期,现在正处于成熟期,呈现稳定发展的趋势;而国际研究成果逐年增长,对农业信息领域的关注度越来越高。其次,从核心作者来看,在农业信息领域发文量较高的国内核心作者有胡绪岚、许世卫、李道亮、郭作玉、刘月姣、梅方权等,在中文核心期刊发文量较高的作者是李瑾、刘小平、刘丽伟、常青、孙素芬等。国际农业信息化领域的核心作者有 Martha C Anderson、Shinji Fukuda、Zhongkui Luo、Wei Ouyang、Kang Oh Lee 等。其中 Wei Ouyang 是北京师范大学教师。事实上,国内不少学者的高质量研究都发表在国际 SCI 期刊上,这也从侧面反映出我国农业信息化研究在逐步走向国际化。再次,从文献来源来看,国内农业信息领域文献主要来自于科技管理、农业、计算机等学科期刊以及农业院校的硕博论文。而国际农业信息化领域的文献主要来自《Science of the Total Environment》《PLoS One》《Computers and Electronics in Agriculture》《Agriculture Ecosystems & Environment》等期刊。最后,从研究热点来看,国内外农业信息化领域的研究有相同的部分,都关注“3S”技术与互联网技术在现代农业、精准农业中的应用研究;但也存在不同的部分,国内研究更关注农业信息服务、农村信息化建设、现代农业发展等,国际研究则更关注与农业生产和管理相关的气候变化监测、土地利用管理、生物多样性、生态服务等研究。这与我国信息化基础设施相对薄弱、经营管理方式相对落后、生产方式传统不无关系^[15]。

纵观我国农业信息化领域的研究动态,即从早期的农业信息资源和信息系统建设,到后期的农业电子商务与农村信息化建设,再到近期的农业信息服务、精准农业与智慧农业,虽然我国农业信息化已呈多元化发展,但大多仍围绕在信息化建设与信息服务上,信息技术在农业生产实践中的研究深度不够。随着我国农业信息化基础设施的逐步完善和农业信息技术的不断推进,后期我国农业信息化研究应紧跟国际农业信息化的步伐,更多关注“3S”技术、互联网技术、物联网技术等现代信息技术在具体农业管理、农业生产实践、农业信息监测中的应用研究,以大幅度提升我国农业信息化实践的质量和效益。

参考文献:

- [1] 桑 静. 基于知识图谱的我国农业信息化研究回顾与展望[D]. 武汉:华中师范大学,2014.
- [2] 张红燕. 探讨农业信息化技术专业硕士研究生的培养[J]. 农业网络信息,2010(9):122-124,130.
- [3] 杨 杰,汤 薇. 国内农村信息化研究热点与主题演进的可视化分析[J]. 图书情报导刊,2019,4(2):33-42.
- [4] 曹利红. 基于 Citespace 软件的我国农业信息化领域研究的可视化分析[J]. 贵州农业科学,2018,46(2):143-147.
- [5] 丁恩俊,谢 佳,申丽娟,等. 基于文献计量的国内外农业信息化研究态势分析[J]. 西南大学学报(自然科学版),2017,39(8):116-125.
- [6] 卢 垚,宋 敏. 基于文献计量学的我国 2006—2015 年植物分子生物学研究态势分析[J]. 中国农业科技导报,2017,19(6):10-20.
- [7] 钟伟金,李 佳,杨兴菊. 共词分析法研究(三)——共词聚类分析法的原理与特点[J]. 情报杂志,2008(7):118-120.
- [8] 宋茂海,李东方. 基于共词分析的国内生物信息学热点领域研究[J]. 生物信息学,2014,12(1):46-52.
- [9] 宗淑萍. 基于普赖斯定律和综合指数法的核心著者测评——以《中国科技期刊研究》为例[J]. 中国科技期刊研究,2016,27(12):1310-1314.
- [10] 赵 爽. 基于共词分析的我国电子商务领域研究热点探析[D]. 长春:吉林大学,2017.
- [11] 曹幸穗,柏 芸,张 苏. 大众农学史[M]. 济南:山东科学技术出版社,2015
- [12] 王 伟,张 刚,刘九庆,等. 农业信息采集机器人设计及试验研究[J]. 农机化研究,2016(9):206-209.
- [13] Marquesperez I, Moránnavarro G, Perezsalas J L, et al. Agricultural land values using geographic information systems: design location model and tools for information available by geoportal. Application to a spanish agricultural area[J]. Survey Review,2017(5):1-10.
- [14] Kok M T J, Kok K, Peterson G D, et al. Biodiversity and ecosystem services require IPBES to take novel approach to scenarios[J]. Sustainability Science,2017,12(1):177-181.
- [15] 孔繁涛,朱孟帅,韩书庆,等. 国内外农业信息化比较研究[J]. 世界农业,2016(10):10-18.