www.sciengine.com/CSD
ISSN 2096-2223
CN 11-6035/N

OPEN ACCESS
CC BY

文献 CSTR：
32001.14.11-6035.csd.2025.0066.zh
文献 DOI：
10.11922/11-6035.csd.2025.0066.zh
数据 DOI：
10.57760/sciencedb.j00001.01420

文献分类：地球科学

收稿日期：2025-04-22
修回日期：2025-07-17
录用日期：2025-08-21
发表日期：2026-06-10

* 论文通信作者
于君宝：jbyu@swfu.edu.cn
郑丽：zhenglixiao4a@163.com

专题 全球智能生态地平线计划（GIEHP）

基于文献计量的生态遥感中国知网引文数据集（2010–2024）

张凯[1,2,3]，徐洋[1,2,3]，陈晓娜[4,5]，于君宝[1,2,3*]，郑丽[1,2,3*]

1. 西南林业大学，云南省高原湿地保护修复与生态服务重点实验室，昆明 650224
2. 西南林业大学，国家高原湿地研究中心，昆明 650224
3. 国家林业和草原局，滇池湖泊生态系统云南省野外科学观测研究站，昆明 650224
4. 中国科学院地理科学与资源研究所，国家地球系统科学数据中心，北京 100101
5. 江苏省地理信息资源开发与利用协同创新中心，南京 210023

摘要：生态遥感是遥感技术与生态学交叉融合形成的前沿领域，通过卫星、无人机、地面传感器等多平台数据获取和分析，实现对生态系统结构与功能的动态监测与定量评估。本研究基于中国知网引文数据库，系统检索 2010–2024 年以“生态遥感”为主题的学术文献，构建包含 5999 篇期刊论文的引文数据集，采用文献计量学方法，结合自主开发的科学知识图谱工具 biblioCNKI，对数据进行清洗、整合、解析，得到本数据集。本数据集通过 R 语言处理，再经人工补全、整编、校准，具有较高的正确率和可信度。本数据集可为解析生态遥感领域发展趋势、发现研究热点、优化科研布局提供决策依据。

关键词：生态遥感；文献计量；中国知网；biblioCNKI；R 语言

数据库（集）基本信息简介

数据库（集）名称	基于文献计量的生态遥感中国知网引文数据集（2010–2024）
数据通信作者	徐洋（xuyang@lreis.ac.cn）
数据作者	张凯、徐洋
数据时间范围	2010–2024年
地理区域	中国
数据量	5.82 MB
数据格式	*.xlsx
数据服务系统网址	https://doi.org/10.57760/sciencedb.j00001.01420
基金项目	云南省联合基金（U2002209）
数据库（集）组成	数据集共包含4个数据文件，其中：（1）生态遥感引文原始数据集.xlsx是直接从知网自定义导出的最原始的数据集，一共包含2010–2024年5999篇引文数据的信息;（2）生态遥感清洗后引文数据集.xlsx是将生态遥感引文原始数据集.xlsx清洗处理后的数

数据库（集）组成	据集，包含5807行数据;（3）生态遥感关键词数据集.xlsx是生态遥感清洗后引文数据集.xlsx经过R语言编码，提取后的关键词时间数据，包含26387行数据;（4）生态遥感地理单位解析数据集.xlsx是由生态遥感清洗后引文数据集.xlsx经过R语言处理，结合百度地图API编码得到的数据，包含11980行数据。

引　言

生态遥感作为遥感技术与生态学深度融合的前沿交叉领域，通过遥感数据获取与分析方法，为生态系统动态监测与可持续发展目标实现提供了重要的技术支撑[1-2]。近年来，中国生态遥感研究呈现指数级增长，研究热点不断涌现，但领域内缺乏系统性的文献计量数据，并且现有文献计量工具（如 CiteSpace、VOSviewer）主要面向英文论文数据库（如 Web of Science），对中文文献的字段解析、关键词聚类支持不足，分析能力弱、灵活性不足、可视化效果差，使得对学科发展趋势、研究热点演变及科研资源分布的量化分析面临挑战，研制覆盖中国生态遥感领域核心文献的引文数据集具有重要的理论意义与实践价值[3-5]。

当前，国内外已有部分文献计量研究聚焦于生态遥感领域，但多基于 Web of Science 或 Scopus 等国际数据库，缺乏对中国本土研究成果的系统整合[6]。中国知网（CNKI）作为国内权威学术资源平台，汇聚了大量中文核心期刊文献，是分析中国生态遥感研究进展的重要数据源[7]。然而，现有研究或局限于单一时间节点，或未结合自主分析工具实现多维可视化，导致数据覆盖范围有限、分析维度单一，难以全面揭示中国生态遥感研究的时空特征与技术演化路径。

本数据集基于中国知网引文数据库，构建了 2010–2024 年以生态遥感为主题的引文数据集，其优势体现在以下三个方面：其一，数据覆盖完整，时间跨度长达 15 年，完整记录了研究发展历史轨迹；其二，数据集提供的数据信息完善，不仅对总数据进行拆分细化，也巧妙地保留了各数据集之间的联系，为研究生态遥感领域提供更加完整的数据；其三，方法创新性强，结合自主开发的 R 语言工具包 biblioCNKI，实现了发文趋势、关键词聚类、机构地理分布等多维度动态分析，弥补了传统文献计量工具对中文数据支持的不足[8]。本数据集是基于文献计量的方法结合中国知网引文数据库构建的高质量、结构化的引文数据集，不仅是解析学科发展脉络的基础，也为优化科研布局、促进多源数据共享提供关键依据。本数据集的发布，将助力深化遥感技术与生态学理论的协同创新，推动中国生态遥感领域的精细化管理和可持续发展[9-10]。

1　数据采集和处理方法

1.1　数据采集方法

本数据集数据来源于中国知网。通过专业检索得到，检索式为：

SU=((生态保护 + 生态恢复 + 生态环境 + 生物多样性 + 植被覆盖 + 湿地 + 碳汇 + 荒漠化 + 土地利用变化) * (遥感技术 + 遥感监测 + 遥感反演 + 遥感制图 + 卫星影像 + 无人机遥感 + 高光谱遥感 + 雷达遥感 + 激光雷达 + NDVI + EVI + 植被指数 + 影像分类 + 变化检测 + Landsat + Sentinel-2 + MODIS))，这个检索式是通过大模型辅助编写，经过多次检索对比，形成得检索内容最为全面，期刊数量最多的检索式，检索后中文期刊一共 1.3 万篇。

本数据集收集了 2010–2024 年北大核心、CSCD 和 WJCI 等来源中文期刊一共 5999 篇[11]。引文

数据须使用自定义格式导出，导出字段应包含“SrcDatabase-来源库”、“Title-题名”、“Author-作者”、“Organ-单位”、“Source-文献来源”、“Keyword-关键词”、“Summary-摘要”、“PubTime-发表时间”、“Year-年”、“CLC-中图分类号”、“ISSN-国际标准刊号”、“URL-网址”和“DOI-DOI”等，输出格式为 xls 格式。

1.2 数据预处理方法

知网导出的 xls 文件无法直接使用 R 语言打开，需要先使用 Excel 软件打开，另存为 xlsx 格式，然后再使用 R 语言 biblioCNKI 包处理。

1.3 数据处理方法

本数据集使用 R 语言 biblioCNKI 包处理生成，数据处理示意代码如下。实际应用时，需将下面代码中的文件位置和输出位置改为用户自定义设置。

```
# 下载安装 biblioCNKI 包
library(remotes)
install_github("zk204510/biblioCNKI")

# 加载文献发表数量时间序列分析
library(biblioCNKI)
cumulative("知网遥感引文数据的位置（xlsx）", "输出图片的格式名称")

# 加载研究热点演变和趋势分析
library(biblioCNKI)
keydumbbell ("知网遥感引文数据的位置（xlsx）", "输出图片的格式名称")

# 加载研究机构空间分布与热度分析
library(biblioCNKI)
Geoheatmap ("知网遥感引文数据的位置（xlsx）", "输出图片的格式名称")

# 加载研究数据清洗函数
library(biblioCNKI)
process_citation_data("知网遥感引文数据的位置（xlsx）", sheet = 1 , "输出清洗后数据的位置（xlsx）")

# 加载关键词提取函数
library(biblioCNKI)
process_keywords("输出清洗后数据的位置（xlsx）", "输出关键词保存位置.xlsx" )

# 加载地理解析提取函数
library(biblioCNKI)
```

```
geocode_batch_processing("输出清洗后数据的位置（xlsx）","输出地理解析保存位置.xlsx",
"高德地图 API")
```

2 数据样本描述

基于文献计量的生态遥感中国知网引文数据集（2010–2024）原始数据来源于知网，由 2010–2024 年生态遥感相关的引文数据组成，包含生态遥感引文原始数据、生态遥感清洗后引文数据、生态遥感关键词数据、生态遥感地理单位解析数据 4 个表，具体组成如表 1 所示。

表 1　数据集组成说明

Table 1　Explanation of dataset composition

序号	数据表名称	数据字段	采集说明
1	生态遥感引文原始数据	SrcDatabase-来源库、Title-题名、Author-作者、Organ-单位、Source-文献来源、Keyword-关键词、Summary-摘要、PubTime-发表时间、CLC-中图分类号、URL-网址	运用大模型自动生成的尽可能全面的生态遥感领域知网专业检索词，并选择自定义导出。
2	生态遥感清洗后引文数据	Pid、Title、Author、AuthorAddress、Keyword-关键词、Date	基于知网导出的原始数据，使用 R 语言给每行加上 Pid，作为区分论文的主键。对原始数据进行数据清洗，去除不完全的数据，保留并重新整合处理 Title、Author、AuthorAddress、Keyword-关键词、Date 等字段。
3	生态遥感关键词数据	Keyword-关键词、Pid、Date	运用 R 语言提取每个数据的关键词、Pid、Date 等字段并且每个关键词单独一行展示。
4	生态遥感地理单位解析数据	Pid、unit、Longitude、Latitude、Date	基于 R 语言，对每篇论文的发表单位进行拆分，通过调用高德 API，并对每个单位的经纬度进行解析。

这些数据可以使用 R 语言 biblioCNKI 包进行可视化，得到生态遥感文章数量趋势图、生态遥感关键词时间分布趋势图、生态遥感发文单位空间时间分布图。

生态遥感文章数量趋势图主要反映生态遥感领域在 2010–2024 年的研究趋势和发文情况，如图 1 所示。

生态遥感关键词时间分布趋势图，用来表现生态遥感在 2010–2024 年研究热点变化趋势，如图 2 所示，绿点位置代表关键词出现时间的一分位数，红点位置代表关键词出现时间的三分位数，中间蓝点位置代表关键词出现时间中值，蓝点的大小表示关键词出现频次。生态遥感领域在 2010–2024 年之间研究热点由初期的“纯遥感、遥感监测、生态环境监测”等方面到中期转变为以“归一化植被指数（NDVI）”为主要研究和发文方向，中期将遥感影像与植被覆盖度相结合，通过 NDVI 来更好地研究遥感生态监测与生态环境质量评估，后期研究热点转向“无人机、地理探测器”等方面，通过先进仪器设备对生态遥感领域实现智能化监测、细致化研究。

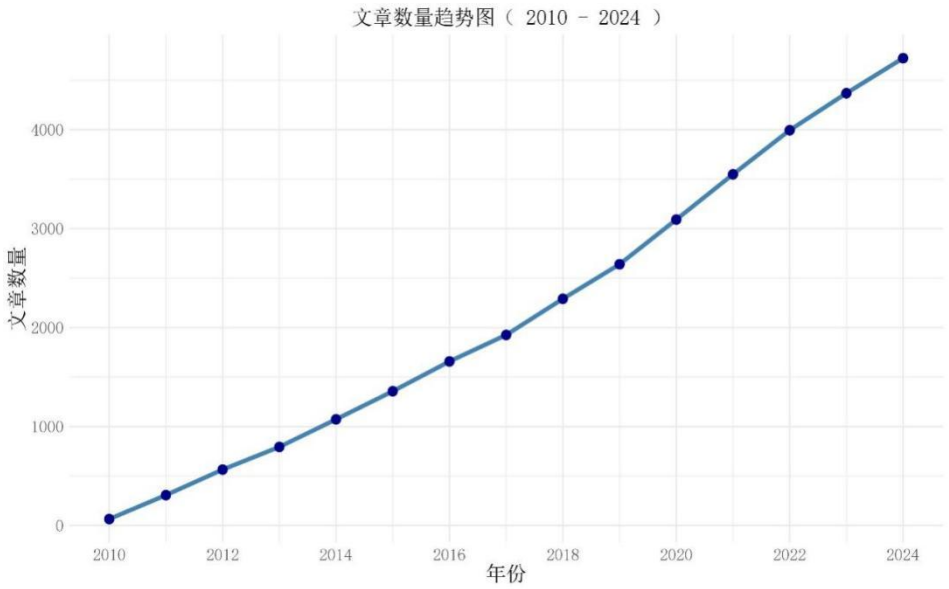


图 1　生态遥感文章数量趋势图

Figure 1　Publication trend of ecological remote sensing publications

生态遥感发文单位空间时间分布图，选取了 2010 年、2015 年、2020 年、2024 年 4 个年份来进行论文发表单位空间热力可视化，展现了论文发表单位的时空变化，如图 3 所示。在生态遥感领域的研究单位集中分布于东部沿海和中部地区，并且随时间推移，中国中部和东部沿海发文数量增加迅速。其中在北京地区从事生态遥感领域研究单位的数量最多，在 2020 年远远领先于其他地区。

3　数据质量控制和评估

本数据集由大模型辅助生成检索词，检索出引文文本经 R 语言处理生成，经人工补全、整编、校准，形成最终数据集。

在数据质量控制和评估方面，本套数据集中只有“生态遥感地理单位解析数据集”受部分单位名称变更，单位存在多处地址等问题影响，导致有少部分单位的经纬度数据未解析成功。其中对部分未解析成功的数据进行了人工解析处理，除此之外，其余数据集都完整处理。

本数据集由于采用计算机语言进行处理，暂时没有可以直接方便的校验方法和系统，所以采取人工校验。生态遥感引文原始数据集经知网导出后一共 5999 条数据，其中含有空值和不完整信息的数据 192 条，导出数据的完整正确率为 96.7%。生态遥感清洗后引文数据集有 5807 条数据，均为有效数据。对这 5807 条数据每 100 条中随机抽取 5 条，其数据信息正确。生态遥感关键词数据集经人工校验比对，数据信息正确。生态遥感地理单位解析数据集中除了未解析成功的数据之外，对解析成功的数据采用人工校验。在每一百个地理单位数据中采用随机抽样的方法选取 5 个数据，通过人工查询经纬度信息进行校验，一共抽取 250 条数据。由于部分作者单位为二级学院，或涉及保密单位、国外单位、单位名称变更等原因在地图上解析不到，因此，部分二级学院作者单位只能解析到学校一级，经纬度只能解析到小数点后 3–4 位，抽取校验的 250 条数据地址全部正确。

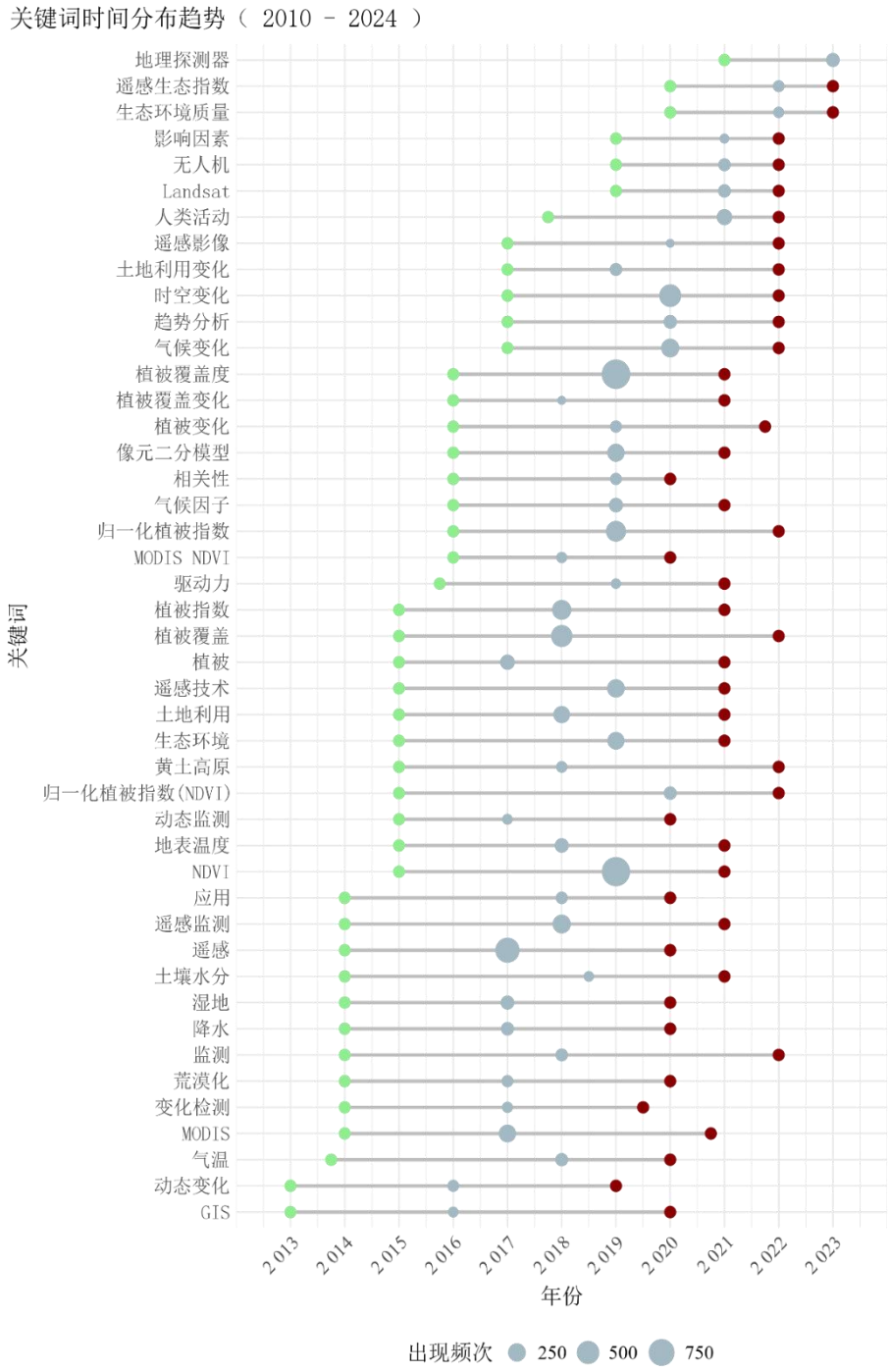


图 2　生态遥感关键词时间分布趋势图

Figure 2　Temporal distribution of keywords in ecological remote sensing

4　数据价值

本数据集是基于中国知网（CNKI）构建的生态遥感引文数据集（2010–2024 年），在数据覆盖范围、加工方法创新性及潜在应用价值等方面具有显著优势，弥补了国内外相关数据集的不足。与基于 Web of Science 或 Scopus 的全球性文献计量数据集相比，本数据集聚焦于中国生态遥感研究，现有国际数据集虽覆盖广泛，但对中国本土研究关注不足。本数据集通过系统整合中文文献，实现

了对中国生态遥感领域发展轨迹的全景式记录，为分析国内学科动态提供了数据基础。并且本数据集配套自主研发的 R 语言工具包 biblioCNKI，针对 CNKI 数据格式进行了深度适配，实现了从数据清洗、结构化存储到可视化分析的全流程自动化处理。

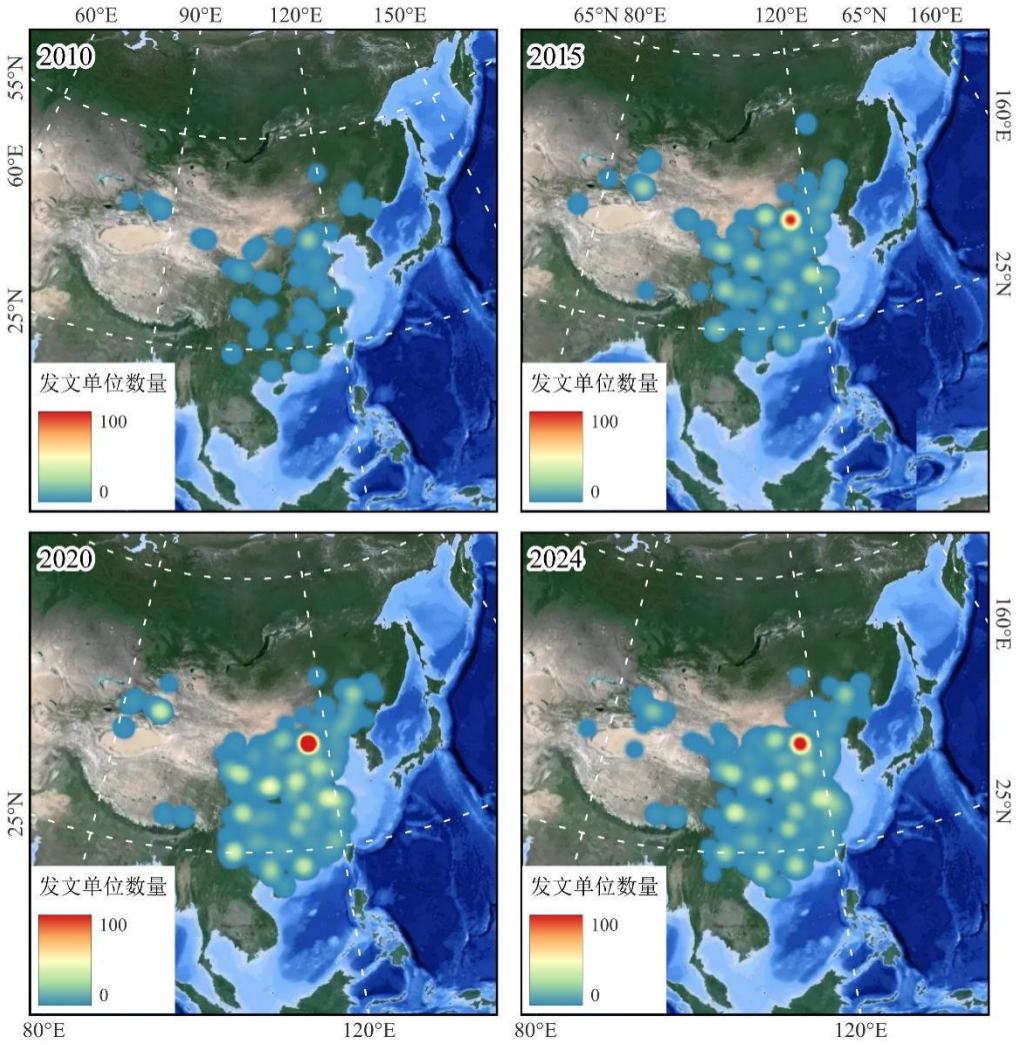


图 3　生态遥感发文单位空间时间分布图

Figure 3　Spatial-temporal distribution of publishing institutions of ecological remote sensing papers

本数据的服务对象为从事生态遥感研究的学生、科研工作者，为本、硕、博研究论文写作提供参考，方便相关领域研究人员了解国内生态遥感 2010–2024 年研究热点变迁、研究单位时空变化。

5　数据使用方法和建议

基于文献计量的生态遥感中国知网引文数据集（2010–2024），通过知网检索和生态遥感相关领域数据的导出结合 biblioCNKI 工具包（网址链接：https://github.com/zk204510/biblioCNKI）进行处理，从而实现数据的可视化，本数据集从知网导出时需要进行自定义导出，导出的引文数据内容为 SrcDatabase-来源库、Title-题名、Author-作者、Organ-单位、Source-文献来源、Keyword-关键词、Summary-摘要、PubTime-发表时间、CLC-中图分类号和 URL-网址等。由于知网导出的数据格式不允许直接读取，所以需要进行复制、另存，方可使用。然后结合 biblioCNKI 包，通过 R 语言调用

biblioCNKI 包对数据集进行处理，从而实现数据的可视化。

作者分工职责

张凯（2004—），男，河南南阳人，本科生，主要承担软件开发和论文撰写工作。
徐洋（1992—），男，河北省保定市人，讲师，研究方向为环境遥感与科学数据共享，主要承担方法设计和论文修改工作。
陈晓娜（1984—），女，河南汝州人，副研究员，主要从事地表参数反演与数据共享研究工作，主要承担方法设计与数据论文修改工作。
于君宝（1970—），男，吉林长春人，教授，主要从事湿地元素生物地球化学与生态修复研究工作，主要承担数据论文修改工作。
郑丽（1979—），女，黑龙江省鹤岗市人，副研究员，研究方向为环境遥感与生态学，主要承担方法设计和论文修改工作。

参考文献

[1] GUO Q H, HU T Y, MA Q, et al. Advances for the new remote sensing technology in ecosystem ecology research[J]. Chinese Journal of Plant Ecology, 2020, 44(4): 418–435. DOI:10.17521/cjpe.2019.0206.
[2] 高吉喜, 赵少华, 侯鹏. 中国生态环境遥感四十年[J]. 地球信息科学学报, 2020, 22(4): 705–719. [GAO J X, ZHAO S H, HOU P. Advances of remote sensing on ecology and environment in China[J]. Journal of Geo-Information Science, 2020, 22(4): 705–719.]
[3] 丁琪洵, 江文娟, 童童, 等. 基于 CiteSpace 的土壤质量遥感研究文献计量分析[J]. 土壤通报, 2023, 54(3): 721–729. DOI:10.19336/j.cnki.trtb.2022042605. [DING Q X, JIANG W J, TONG T, et al. Bibliometric analysis of soil quality remote sensing literature based on CiteSpace[J]. Chinese Journal of Soil Science, 2023, 54(3): 721–729. DOI:10.19336/j.cnki.trtb.2022042605.]
[4] 徐恒康, 逯辉, 刘灏, 等. 全球草地碳汇研究趋势与重点领域: 基于 1992—2022 年文献计量分析[J]. 草地学报, 2024, 32(7): 2169–2178. [XU H K, LU H, LIU H, et al. Research trends and focus areas on the global grassland carbon sink—a bibliometric analysis for 1992—2022[J]. Acta Agrestia Sinica, 2024, 32(7): 2169–2178.]
[5] 方喻弘, 宋丽, 肖潇, 等. 基于文献计量的中国水利遥感研究热点分析[J]. 地理空间信息, 2024, 22(8): 15–22. DOI:10.3969/j.issn.1672-4623.2024.08.004. [FANG Y H, SONG L, XIAO X, et al. Research hotspots analysis of Chinese water conservancy remote sensing based on bibliometric method[J]. Geospatial Information, 2024, 22(8): 15–22. DOI:10.3969/j.issn.1672-4623.2024.08.004.]
[6] 蔡灵, 黄雨欣, 戴娟娟, 等. 基于文献计量分析的河口生态修复与评估的现状和展望[J]. 应用海洋学学报, 2025, 44(2): 402-412.[CAI L, HUANG Y X, DAI J J, et al. Present situation and prospect of ecological restoration and effect evaluation of estuaries based on bibliometric analysis[J]. Journal of Applied Oceanography, 2025, 44(2): 402-412.]
[7] 尹慧敏, 胡秀娟, 杨立娟, 等. 城市生态遥感研究热点与发展趋势的可视化分析[J]. 遥感技术与应用, 2024, 39(3): 642–658. DOI:10.11873/j.issn.1004-0323.2024.3.0642. [YIN H M, HU X J, YANG L J,

et al. Visualization analysis of the development and hot spots of urban ecological remote sensing research[J]. Remote Sensing Technology and Application, 2024, 39(3): 642–658. DOI:10.11873/j.issn.1004-0323.2024.3.0642.]

[8] XU Y, YANG Y P, CHEN X N, et al. Bibliometric analysis of global NDVI research trends from 1985 to 2021[J]. Remote Sensing, 2022, 14(16): 3967. DOI:10.3390/rs14163967.

[9] 徐克立, 吕伟才, 汤连盟. 基于改进遥感生态指数(MRSEI)的淮河流域生态环境质量评价[J]. 中国环境监测, 2025, 41(1): 225–234.. DOI:10.19316/j.issn.1002-6002.2025.01.24 [XU K L, LYU W C, TANG L M. Ecological environment quality assessment of Huaihe River Basin based on modified remote sensing ecological index(MRSEI)[J]. Environmental Monitoring in China, 2025, 41(1): 225–234.. DOI:10.19316/j.issn.1002-6002.2025.01.24.]

[10] 潘占冬, 王轶夫, 王可月, 等. 基于改进型遥感生态指数的将乐县生态质量评估及驱动分析[J]. 环境科学, 2025, 46(10): 6419–6431. DOI:10.13227/j.hjkx.202408293. [PAN Z D, WANG Y F, WANG K Y, et al. Ecological quality assessment and driving analysis of Jiangle County based on modified remote sensing ecological index[J]. Environmental Science, 2025, 46(10): 6419–6431. DOI:10.13227/j.hjkx.202408293.]

[11] LI X B, FU X J, WANG Z G. Bibliometric analysis of Salvia miltiorrhiza research based on CNKI Chinese academic journal database[C]. 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). San Diego, CA, USA. IEEE, 2019: 2499–2502. DOI: 10.1109/BIBM47256.2019.8982991.

论文引用格式

张凯, 徐洋, 陈晓娜, 等. 基于文献计量的生态遥感中国知网引文数据集（2010–2024）[J/OL]. 中国科学数据, 2026, 期卷未定. (2026-06-10). DOI: 10.11922/11-6035.csd.2025.0066.zh.

数据引用格式

张凯, 徐洋. 基于文献计量的生态遥感中国知网引文数据集（2010–2024）[DS/OL]. V1. Science Data Bank, 2025. (2025-04-28). DOI: 10.57760/sciencedb.j00001.01420.

A dataset of CNKI citations in ecological remote sensing based on bibliometrics analysis (2010–2024)

ZHANG Kai[1,2,3], XU Yang[1,2,3], CHEN Xiaona[4,5], YU Junbao[1,2,3*], ZHENG Li[1,2,3*]

1. Yunnan Key Laboratory of Plateau Wetland Conservation, Restoration and Ecological Services, Southwest Forestry University, Kunming 650224, P.R. China

2. National Plateau Wetlands Research Center, Southwest Forestry University, Kunming 650224, P.R. China

3. Dianchi Lake Ecosystem Observation and Research Station of Yunnan Province, National Forestry and Grassland Administration, Kunming 650224, P.R. China

4. National Earth System Science Data Center, Institute of Geographic Sciences and Natural Resources

Research, Chinese Academy of Sciences (CAS), Beijing 100101, P.R. China
5. Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, P.R. China
*Email: jbyu@swfu.edu.cn (YU Junbao); zhenglixiao4a@163.com (ZHENG Li)

Abstract: Ecological remote sensing is a frontier field formed through the cross-fusion of remote sensing technology and ecology. By integrating multi-platform data sources, such as satellites, drones, and ground-based sensors, it enables dynamic monitoring and quantitative evaluation of ecosystem structure and function. Based on the citation database of the China National Knowledge Infrastructure (CNKI), this study systematically retrieved academic literature on “ecological remote sensing” published from 2010 to 2024 and constructed a citation dataset comprising 5,999 journal articles. Using bibliometric methods and a self-developed scientific knowledge mapping tool, biblioCNKI, the data were collected, cleaned, integrated and analyzed to generate the final dataset. This dataset was further processed using R language, followed by manual completion, verification, and standardization, ensuring high accuracy and reliability. This dataset provides a solid foundation for analyzing development trends in ecological remote sensing field, identifying research hotspots, and optimizing the layout of scientific research.

Keywords: ecological remote sensing; Bibliometrics; CNKI; biblioCNKI; R programming language

Dataset Profile

Title	A dataset of CNKI citations in ecological remote sensing based on bibliometrics analysis (2010–2024)
Data corresponding author	XU Yang (xuyang@lreis.ac.cn)
Data authors	ZHANG Kai, XU Yang
Time range	2010–2024
Geographical scope	China
Data volume	5.82 MB
Data format	*.xlsx
Data service system	https://doi.org/10.57760/sciencedb.j00001.01420
Source of funding	NSFC Yunnan Joint Fund (U2002209)
Dataset composition	The dataset consists of 4 data files: (1) Eco-Remote Sensing Citation Raw Dataset.xlsx is the original dataset directly exported from CNKI using customized query settings, containing 5,999 citation records from 2010 to 2024. (2) Cleaned Eco-Remote Sensing Citation Dataset.xlsx is the cleaned version of the raw dataset, comprising 5,807 records. (3) Eco-Remote Sensing Keyword Dataset.xlsx contains keyword-time data extracted from the cleaned citation dataset using the R Language, with 26,387 rows. (4) Eco-Remote Sensing Geographic Unit Parsing Dataset.xlsx is generated by processing the cleaned citation dataset with the R programming language and integrating the Baidu Map API, containing 11,980 records.