

# 大数据时代国土变更调查的数据处理与优化

胡杰 林航

DOI:10.12238/gmsm.v7i12.2050

**[摘要]** 大数据时代, 国土变更调查面临数据量急剧增加、数据类型复杂化和更新频率不断提高的挑战。传统的数据处理和分析手段已经很难适应国土资源高效精确管理的需要。本文以大数据为背景, 研究国土变更调查数据处理和优化的关键对策, 包括数据质量提升、存储效率优化、协同处理强化、动态维护机制和成果应用, 采用分布式存储、智能算法、动态监测等方法, 对国土空间规划数据处理过程进行优化, 提高数据的科学性与时效性, 为国土空间规划编制和实施提供科学依据, 促进国土资源管理向智能化、精细化方向发展。

**[关键词]** 大数据时代; 国土变更调查; 数据处理; 优化

中图分类号: P628+.4 文献标识码: A

Data processing and optimization of land change investigation in the era of big data

Jie Hu Hang Lin

**[Abstract]** In the era of big data, land change surveys face challenges such as a sharp increase in data volume, complexity of data types, and increasing frequency of updates. Traditional data processing and analysis methods are no longer suitable for the needs of efficient and precise management of land and resources. This article takes big data as the background to study the key strategies for processing and optimizing land change survey data, including improving data quality, optimizing storage efficiency, strengthening collaborative processing, dynamic maintenance mechanism, and application of results. Distributed storage, intelligent algorithms, dynamic monitoring and other methods are adopted to optimize the data processing process of land spatial planning, improve the scientificity and timeliness of data, provide scientific basis for the preparation and implementation of land spatial planning, and promote the development of land and resource management towards intelligence and refinement.

**[Key words]** big data era; Land change investigation; Data processing; optimize

## 前言

海量的地理空间数据, 遥感影像, 土地利用现状等信息, 为准确把握土地资源动态提供了可能; 再加上数据具有复杂多样、更新速度快等特点, 对传统的数据处理和分析方法提出了更高的要求。目前, 国土变更调查数据存在重、错、存储量低、协同处理不充分等问题, 严重影响国土变更调查的科学性与时效性。如何优化数据处理流程, 提高数据质量和使用效率, 是国土变更调查急需解决的关键科学问题。

## 1 数据处理与优化在国土变更调查中的意义

### 1.1 通过高效的数据处理手段, 加速调查进程

通过对原始数据的清洗, 剔除重复、无效或错误信息, 可以有效地降低后续分析过程中的扰动, 保证数据质量, 为快速、准确地调查打下基础。例如, 在土地利用现状调查中, 剔除因设备故障或人为因素造成的误报数据, 可以避免重复调查、核实, 节约大量的人力物力。借助机器学习等先进算法与工具, 实现国土

变更信息的自动提取与分类, 不仅可以提高数据处理速度, 而且可以降低人为干预带来的误差, 提高调查效率。本文以地表覆盖变化监测为例, 采用机器学习算法自动解译遥感图像, 实现大范围土地分类和变化识别。国土变更调查涉及到跨部门、跨区域的数据采集和分析。通过搭建一个统一的数据平台, 实现数据实时共享和协同处理, 实现各个部门之间的工作同步, 避免在数据传递过程中出现的延迟和重复劳动。

### 1.2 为土地资源管理提供科学依据

通过剔除重复的、无效的和错误的数据, 对数据进行规范化处理, 可以准确地反映出土地利用的类型、面积、分布等关键信息, 为我国土地资源合理分配、合理规划提供科学依据, 有助于决策者对土地资源优势与劣势进行精准识别, 进而制定更具前瞻性、适应性的土地利用策略。例如, 在城市扩展过程中, 准确获取土地变更数据, 可以为新增建设用地选址提供科学依据, 避免盲目开发, 造成资源浪费。通过历史数据与现状数据的比较分

析,可使管理者对土地资源利用效率及生态环境状况作出准确的评价。例如,利用土地质量变化监测数据,可以及时发现土地退化、违法占用等问题,及时采取对策。本文的实施将为我国国土资源保护与修复提供有力支撑。通过大数据分析,可以发现土地资源管理中存在的问题及发展趋势,为政府决策提供科学依据<sup>[1]</sup>。

## 2 数据时代国土变更调查的数据处理要点

### 2.1 去除重复、无效或错误数据

国土变更调查所涉及的数据来源广泛而复杂,既有卫星遥感影像数据,也有野外调查记录和历史档案数据。在采集、录入过程中,这些数据极易产生重复记录。为此,本文拟采用数据清洗工具与算法,结合地块编号、地理坐标等唯一标识符,快速定位并剔除重复数据,提高数据精度与一致性。在国土变更调查过程中,通过设置土地面积合理性范围和土地利用类型合法性等数据校验规则,逐一验证数据的正确性。对于不合规的记录,系统将自动打上无效的标签,并将其隔离。在此基础上,结合人工审核方法,对可能出现的特殊情形(如特殊用途用地面积超标等)作出判断与修正,以保证数据真实有效。不正确的数据来源有采集设备故障、人工录入错误或数据传输过程中的干扰等。在地面变化调查中,通过异常检测等机器学习算法,可以自动识别异常点。在土地质量评价中,由于仪器的故障,一些土壤指标会出现异常值。采用算法检测和现场复查相结合的方法,可以对错误数据的精确定位和校正<sup>[2]</sup>。

### 2.2 实施数据标准化,提高数据可比性

国土变更调查是一项综合性的研究课题,这些数据在采集、存储过程中存在格式上的差异,通过建立统一的坐标体系、数据编码规则、数据字段格式等数据格式标准,可以把零散的数据转换成标准化的数据结构,为后续数据的分析处理提供统一的依据。在国土变更调查过程中,不同区域、不同部门在定义上可能会有不同的含义。建立统一的术语及分类标准,明确各数据字段的含义与范围,可消除因语义差异而引起的误会与混乱。如在土地利用类型划分上,“耕地”统一划分为水田、旱地等具体类别,明确了它们的定义与范围,从而实现了不同区域、不同时段的数据的直接对比与分析。这一语义规范化不仅增强了数据的可比性,而且也方便了跨地域、跨部门的数据共享与协同工作。采用数据管理系统及自动化工具,实现数据输入、存储、更新等全过程的规范控制。比如,在DBMS中引入约束机制,保证数据的格式与内容满足预先设定的标准;使用数据质量检查工具,定期扫描、检查数据,发现不合格的数据记录,及时纠正<sup>[3]</sup>。

### 2.3 挖掘数据间的潜在关联与规律

运用聚类分析、决策树等方法,对地形、土壤、气候等地理要素与土地利用类型的关系进行分析。这一分析结果表明,在一定的地理条件下,水稻田往往分布在低地和水源充足的地区。研究结果不仅可以为土地利用规划提供科学依据,而且可以指导土地整治、生态保护,保障土地利用与自然环境的协调发展。结合机器学习算法,识别不同时段土地利用类型的演变模式及驱

动因子。例如,在城市化进程中,耕地转化为建设用地的速度、方式及其与经济发展、人口增长的关系等。利用关联分析方法挖掘土地质量、农业产出、生态环境等数据间的潜在关联。通过对跨部门数据的集成和分析,可以打破“信息孤岛”,促进不同部门间的协作,提高国土资源管理工作的整体效率<sup>[4]</sup>。

### 2.4 将复杂数据以图表、地图等形式直观呈现

GIS平台可以将不同时期、不同地域的土地数据叠加到地理底图上,以颜色、符号等直观的方式标示出土地类型、面积变化和用途转换等信息。该地图形式可使管理人员快速把握该地区土地利用的总体布局及变化趋势。例如,通过编制土地利用变更图,可以清楚地反映过去若干年来某一地区耕地减少、建设用地增加的具体位置及范围,从而为区域国土规划的编制提供直观的依据。运用统计图分析并展示调查数据,可以更好地揭示数据背后的规律性;如用柱状图、折线图等直观地展示了不同类型用地面积的变化趋势,便于管理人员迅速地识别出重点问题区及关键变化节点。饼状图可以用来分析各类型用地所占的比重,并能直观地反映各类型土地资源所占的比重。本研究不仅 can 将复杂的数据简单化,而且可以帮助管理者更好地进行深度分析,为政策制定、土地资源配置等提供定量的决策支持。

### 2.5 利用智能算法优化数据处理流程与结果

国土变更调查的数据规模巨大,涉及到多源异构地理信息,土地利用类型,地物特性等。传统的数据处理方法费时费力,而且容易出错,以神经网络、决策树等机器学习算法为代表的机器学习算法能够快速识别出数据中的离群点、错误记录和重复信息,实现对数据的自动修正和筛选。通过对数据处理过程的建模和分析,使算法能够根据实际需要动态地调整处理步骤和参数配置。例如,在筛选阶段,可以利用聚类算法自动地将数据分类,并将其优先级较高或较重要的片段作为优先级;在数据融合阶段,采用集成学习算法对多源数据进行融合处理,保证结果的可靠性和一致性。利用时间序列分析方法,预测未来土地利用变化趋势,为土地规划与政策调整提供科学依据。

## 3 数据时代国土变更调查的数据优化对策

### 3.1 建立完善的数据质量检查与评估机制

在数据收集阶段,使用自动工具对原始数据进行初步检验,检查数据格式是否规范,关键字段有没有遗漏。例如,在输入土地面积与座标数据时,设定数据范围限制及逻辑检查规则,以保证数据的合理性;在数据集成阶段,对多源数据进行重复、一致的比对,及时发现和纠正数据冲突。建立清晰的数据质量评价指标,包括数据的准确性、完整性比例和更新及时性等,实现数据质量的定量评价。例如,对土地利用类型数据的设定精度要达到95%以上,而土地面积数据的误差幅度要控制在一定的范围内。这些量化指标既可以为数据质量评价提供清晰的度量指标,又可以方便地对数据质量进行持续监测与改善。定期收集用户反馈信息,及时发现应用过程中出现的问题。比如,当土地资源管理决策时,如果数据不能满足分析要求或者有偏差,可以利用反馈机制快速找到问题数据并加以纠正。同时,根据数据质量评

价结果,对数据处理过程进行定期优化调整。

### 3.2 采用分布式存储、云存储等技术,提高数据存储效率

国土变更调查数据量大、增长快,涉及到海量的地理空间数据、遥感影像和土地利用变化记录。分布式存储体系结构可以根据数据访问的频度和重要性对存储资源进行合理配置,从而实现对数据的快速读写和有效管理。以 Hadoop 分布式文件系统 (HDFS) 为例,将海量数据进行分块存储,并采用冗余备份机制保证数据的安全可靠,从而提高数据存储的柔性与可扩展性。云存储平台可以根据数据量的变化动态调整存储资源,有效地解决了传统存储系统因容量不足或过剩而造成的资源浪费问题。针对国土变更调查数据量随季节变化或特殊任务变化而变化的特点,云存储灵活地应对变化,按需分配存储空间,减少存储开销。与此同时,云存储平台通常具有高效的数据备份与恢复功能,能够保证数据在遭受非预期破坏或丢失情况下的快速恢复,进一步提升数据存储的安全性。

### 3.3 加强数据协同处理,提升整体调查效率

国土变更调查涉及到自然资源、生态、农业等多个部门,不同部门掌握的数据类型不同,重点也不同。通过建立一个统一的数据共享平台,各部门可将各自收集到的土地利用现状、生态环境监测、耕地质量评价等数据上传到平台上,实现实时更新。通过整合自然资源部门土地利用变化数据和生态环境部门的生态红线数据,可以更全面地反映生态保护红线变化情况。这种跨部门数据共享机制,在提高数据使用效率的同时,也可避免重复收集、不一致性等问题,大幅提高调查效率。国土变更调查数据规模巨大,传统的单一模式很难满足其高效率处理要求,我们可以将复杂计算任务分解成多个子任务,并在不同计算节点上进行并行执行。例如,分布式计算能够同时提取多块影像的特征、分类以及变化检测,从而极大地提高了数据处理的速度。这种并行处理方法不但提高了数据处理速度,而且增强了系统的扩展性,使之能适应日益增长的数据量。

### 3.4 对数据进行持续维护,处理新增或变更的数据

随着土地利用类型的改变、基础设施的建设和生态保护工程的实施,国土调查数据也在不断地更新。在此基础上,结合卫星遥感、无人机测绘和野外调查,建立基于时间序列的土地资源动态监测体系。例如,利用高分辨率卫星影像对城市扩展区土地利用变化进行定期监测,一旦发现新的建设用地和耕地减少,系统会自动启动更新程序,并及时补充最新数据。这一动态更新机制保证了数据能够及时地反映土地资源的实际情况,为国土资源管理提供及时的决策依据。在数据更新过程中,由于人为因素

或人为因素,可能会造成数据质量问题。建立数据质量追溯机制,记录数据更新过程及结果,及时发现错误并加以改正。例如,在土地利用类型更新与实地调查结果不一致的情况下,利用回溯机制查找问题根源,确定是否是数据收集失误或处理算法失误,并有针对性地修正。

### 3.5 利用数据分析结果,指导国土空间规划的制定与实施

结合地理信息系统(GIS)、遥感等技术,分析土地利用格局演变趋势,明确城镇扩张对周边耕地的占用情况及生态脆弱区的退化程度,科学地划定生态保护红线、永久基本农田及城镇发展边界,实现土地资源合理布局和高效利用。这种以数据为基础的规划方法可以有效地避免决策的盲目性,提高规划的科学性与前瞻性。在实施国土空间规划时,由于受多种因素影响,可能出现与规划目标不符的状况。在规划实施过程中,通过建立实时监测系统,并结合数据分析技术,实现对土地利用变化的动态追踪。比如,对耕地质量进行时间序列分析,对耕地质量进行动态监测,一旦发现耕地质量下降,及时调整整治方案,加大生态修复力度,保证规划目标的实现。

## 4 结束语

大数据时代国土变更调查数据处理与优化是促进国土资源科学管理与可持续发展的重要一环,随着大数据技术的发展,土地变更调查的应用场景不断扩大,土地变更调查向智能化和精细化方向发展,我们要进一步加强技术创新,深化多部门协作,推进国土变更调查工作与时俱进,为国土资源高效利用、生态保护目标的实现提供技术支持。

## 参考文献

- [1] 云超.土地资源利用变更调查中3S技术的应用分析[J].科技资讯,2024,22(24):154-156.
- [2] 冯文利,吴海平,曾珏.“空天地网”自然资源调查监测技术体系构建的业务实践[J].自然资源情报,2024,(09):1-7.
- [3] 吴相焱,许秋成,杜永乐,等.基于微服务架构的第三次全国国土调查数据库管理与共享服务平台设计与实现[J].测绘与空间地理信息,2023,46(02):114-117+120.
- [4] 尹莹.基于数据中台的国土信息平台框架体系研究[J].电脑与电信,2021,(09):55-58.

## 作者简介:

胡杰(1990--),本科,从事土地利用调查工作,毕业于浙江农林大学。

林航(1995--),本科,从事国土调查相关工作,毕业于东华理工大学。