

荒漠化和沙化监测数据库框架设计与应用

孙 伟^{1,2}, 侯瑞霞^{1*}, 庞丽峰¹, 唐小明¹

(1. 中国林业科学研究院 资源信息研究所, 北京 100091; 2. 新疆农业大学 计算机与信息工程学院, 新疆 乌鲁木齐 830052)

摘 要:针对我国荒漠化和沙化监测数据组织管理存在的体系不健全、标准难统一、时效性不强这 3 个问题,通过对监测调查数据的分析,结合分布式空间数据库的分片设计思想,设计了荒漠化和沙化的分布式多专题监测数据库框架,探讨了监测数据库的综合集成模式结构和管理部署模式,进而基于全国第 3 期和第 4 期监测数据,利用 ArcSDE 在 Oracle 数据库中建立且优化了国家级和省 级监测数据库,并在荒漠化和沙化监测信息管理系统和林业资源监管服务系统进行了应用验证。结果表明,设计的数据库框架是可行的,为我国荒漠化和沙化海量监测数据的科学管理提供了有力的技术支撑和解决方案。

关键词:荒漠化监测;沙化监测;分布式空间数据库

中图分类号:S771.8 **文献标志码:**A **文章编号:**1001-7461(2012)06-0139-07

Framework Design and Application of Desertification and Sandification Monitoring Database

SUN Wei^{1,2}, HOU Rui-xia^{1*}, PANG Li-feng¹, TANG Xiao-ming¹

(1. Research Institute of Forest Resources Information Techniques, Chinese Academy of Forestry, Beijing 100091, China;
2. College of Computer and Information Engineering, Xinjiang Agricultural University, Urumqi, Xingjiang 830052, China)

Abstract: To solve the three problems existed in the desertification and sandification monitoring data organization and management, including: unsound system, difficult to uniform standards and weak timeliness, based on the analysis of the monitoring survey data and combined with distributed spatial database slice design ideas, a desertification and sandification monitoring data distributed and multi-thematic database framework was designed and the comprehensive integrated model structure and management and deployment model of the monitoring database were explored further. The national and provincial monitoring databases, which were constructed and optimized in Oracle DBMS through ArcSDE by utilizing the third and fourth monitoring data in China, were verified both in the build-up of desertification and sandification monitoring information management systems and in the implementation of the forest resources monitoring services system. The results showed that the database framework was feasible and could provide strong technical support and solutions for the scientific management of massive desertification and sandification monitoring data in China.

Key words: desertification monitoring; sandification monitoring; distributed spatial database

我国是世界上受荒漠化危害最为严重的国家之一,也是世界上唯一开展国家级荒漠化和沙化定期监测的国家^[1-2]。我国已初步建立了由宏观监测、专

题监测和典型定位监测 3 个层次组成的监测管理体系,开展了 4 期全国荒漠化和沙化监测,积累了丰富的历史和现势监测数据资源^[3],但是,在监测数据组

收稿日期:2012-02-16 修回日期:2012-04-06
基金项目:“十一·五”国家科技支撑计划项目(2006BAD23B02);中央级公益性科研院所基本科研业务费专项资金资助项目(CAFYBB2008002)。
作者简介:孙伟,男,在读博士,研究方向:林业 GIS 开发与应用。E-mail:maplesunw@gmail.com
*通信作者:侯瑞霞,女,助理研究员,研究方向:GIS 在林业中的应用、数据管理与建模。E-mail:hrx2004126@126.com

织与管理方面目前尚存在 3 个主要问题:①数据管理体系不健全。目前尚未形成科学合理的监测数据资源组织管理框架,不利于国家、省、县等不同管理层次之间的数据资源组织管理和协同共享;②监测指标难以保持一致性。我国各期监测在技术标准等方面基本保持了一致,但由于监测体系本身不断地完善和发展,监测区范围大,目视解译误差,现地核查人为因素等原因,实际操作中很难确保监测指标的一致性,不利于各期监测数据的一体化管理与动态分析比较;③监测数据时效性不强。从目前的监测数据采集、汇集、处理、分析与利用过程来看,未能高效地将各地监测数据及时提交到国家,也无法有效保证地方与国家的同步更新,监测数据的提供时间与需求时间不吻合,不利于国家林业有关部门及时了解荒漠化和沙化发展现状与动态变化趋势。本文设计荒漠化和沙化监测数据库框架,为各层管理部门准确、及时、全面地掌握各层次荒漠化和沙化土地消长变化信息提供有力的数据支撑。

1 监测数据分析

我国的荒漠化和沙化监测以 5 a 为一个周期,已对全国范围内的荒漠化和沙化土地展开了 4 期调查。20 世纪 90 年代初,出于防沙治沙工程和三北防护林体系建设的实际需要,于 1994 年组织开展了第 1 期全国荒漠化和沙化土地普查和监测,采用地面区划调查法,第 1 次全面系统地查清了我国沙化土地资源状况及土地退化趋势;第 2 期监测始于 1999 年,采用抽样调查法;第 3 期监测始于 2004 年,调查手段由地形图调绘技术向“3S”技术转变,调查方法由抽样调查向全面调查转变,信息传输由人员传送向计算机网络化传输转变,成果制作由纸质成果向电子化成果转变;第 4 期监测于 2009 年开始,沿用并扩展第 3 期调查技术与方法,开始关注动态变化^[4-8]。显然,第 3、第 4 期监测技术与方法较为稳定,且具有良好的延续性,其监测区划系统为省(区、市)、县(旗)、乡(苏木)、图斑(地块)4 级,调查监测范围内所有土地,并以矢量图斑为基本单位。第 3 期监测图斑 502 万个,获取各类信息 1.56 亿条记录^[9];第 4 期监测图斑 592 万个,获取各类监测数据 2.5 亿条记录^[10]。本研究着重分析这 2 期监测数据,归纳出监测数据资源分类(图 1)。

2 监测数据库框架

2.1 多区域分片

我国的荒漠化和沙化监测体系分为 3 个层级:

国家级由国家林业局统一部署,防沙治沙办公室具体组织实施,在国家林业局成立全国荒漠化监测中心,建立全局统一的空间数据模型;省级由省级荒漠化监测中心负责,由林业部门牵头,农业、国土和环保等部门配合,在省级范围内确立局部空间数据模型,标准化整合市县上报的监测数据;市县级林业主管部门内部成立监测管理机构,分级对口,层层管理,具体实施监测调查(图 2)。

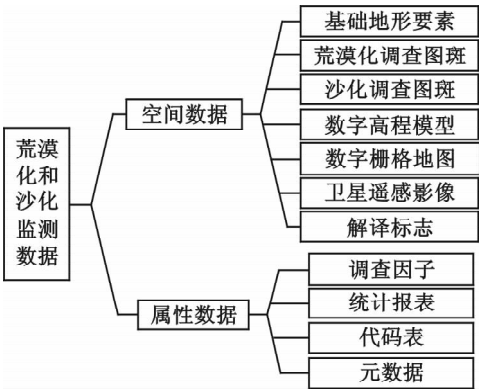


图 1 荒漠化和沙化监测数据分类

Fig. 1 Classification of desertification and sandification monitoring data

可见,荒漠化和沙化监测数据本身在地理空间范围上自然地呈现水平方向的分布特征,而区域分片也称为空间分割分片或横向分片,是把同一地理覆盖的空间信息分割存放在不同存储位置的同构数据库表,是空间数据水平方向的分布,是由地理信息的区域性所决定的^[11]。因此,本文按照空间数据分片的扩展原则^[12]:空间聚集性、空间对象的不分割性、逻辑无缝保持性,在空间尺度上对数据进行多区域分片设计,由省级监测数据库和国家级监测数据库构成。其中,省级荒漠化监测数据库包括 4 个部分:省级监测数据库,省级监测数据库与国家规范化的监测数据库之间的逻辑关系映射数据库,省级基础地理数据库和省级支撑数据库;国家级监测数据库包括 4 个部分:国家级逻辑集成数据库,省级逻辑集成数据库,国家级基础地理数据库和支撑数据库。

在多区域分片模式下,省级负责监测数据采集、汇集、集成和维护更新,可通过数据服务实时提供省级监测数据,并可利用省级逻辑关系映射数据库形成国家直接调用的数据,从而形成国家逻辑映射数据库,保证国家级能获取省级最新的荒漠化和沙化监测数据,可有效解决监测数据时效性不强的问题。

必须考虑各期监测指标之间的映射与转换,即,需要在时间尺度上考虑分片设计,进而需要在时空尺度上进行综合分片设计,这样势必带来分片设计

的复杂性,本研究采用“以空间换时间”的思路,通过在国家级和省级支撑数据库中建立各期监测指标之间的对应关系,实现各期监测数据之间有效的映射与转换,对各期监测异构数据进行整合与集成,形成多期监测一致性数据,可有效解决监测指标难以保持一致性问题。

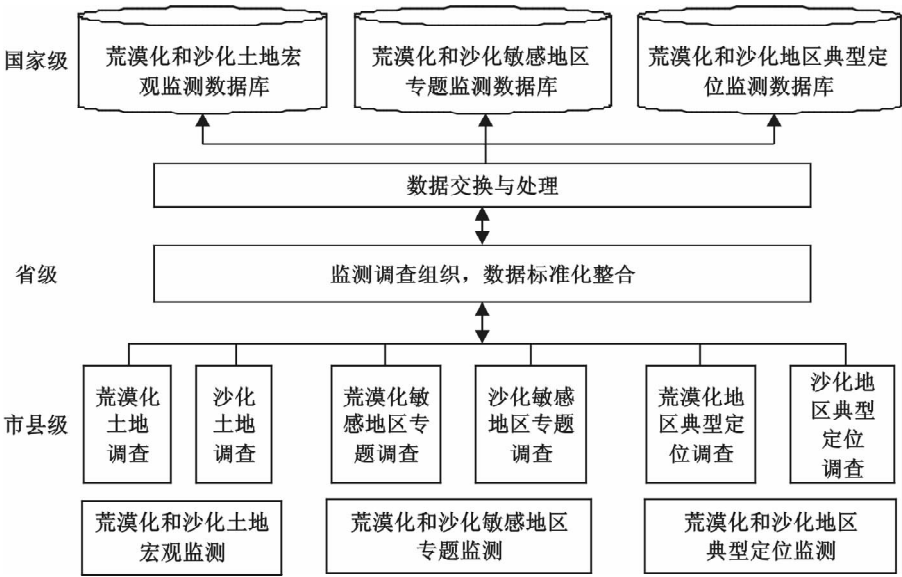


图 2 荒漠化和沙化土地监测体系

Fig. 2 Desertification and sandification monitoring system

2.2 多专题分片

专题分片也称为层分片或纵向分片,是指同一地理范围内不同主题层的空间信息存放在不同的存储位置,是空间数据垂直方向的分布,是由地理信息的专题性决定的^[11],而荒漠化和沙化不同部门收集和维护各自业务领域数据,本研究设计了 6 个专题分片。

2.2.1 基础地理专题 涉及各种比例尺的各期监测基础地理数据,主要包括境界线、行政区划、林业管理区划、气候类型界线及社会情况等,以及数字化地形图、DEM 和 DRG 等。

2.2.2 荒漠化监测专题 包括 4 类:荒漠化土地类型及程度、荒漠化监测区土地利用类型、荒漠化监测区植被盖度和荒漠化土地利用类型。

2.2.3 荒漠化动态专题 包括 4 类:荒漠化土地类型动态、荒漠化土地程度动态、荒漠化土地利用类型动态和荒漠化监测区植被盖度动态。

2.2.4 沙化监测专题 包括 7 类:沙化土地类型、沙化土地程度、具有明显沙化趋势的土地、沙化监测区土地利用类型、沙化土地利用类型、沙化监测区植被盖度和沙化土地植被盖度。

2.2.5 沙化动态专题 包括 5 类:沙化土地类型动态、沙化土地程度动态、沙化土地利用类型动态、沙化监测区土地利用类型动态和沙化土地植被盖度动态。

2.2.6 辅助专题 包括代码、元数据模型及数据字典、统计报表模型、制图模板和逻辑检查映射等数据,以及数据库模式、监测指标、空间转换等一系列关系的映射和链接等。

2.3 分布式多专题监测数据库框架

基于上述多区域分片设计和多专题分片设计,针对海量监测数据,采用混合分片的方式,设计荒漠化和沙化分布式多专题监测数据库框架,构建以省级监测数据库为支撑的全国荒漠化和沙化分布式多专题监测数据库,由 18 个省级荒漠化监测数据库、30 个省级沙化监测数据库和 1 个国家级整体集成的监测数据库构成,省级和国家级均涉及 6 类专题信息,进而形成物理分散、逻辑集中、分布式处理的荒漠化和沙化监测数据库框架模式,确保国家级和省级数据库的无缝集成与互操作,如图 3 所示。

其中,省级数据库的管理、更新和维护可与省级荒漠化和沙化信息系统融为一体,为省级业务用户服务;国家级数据库可利用国家林业专网通过各类数据服务获得省级数据和服务。对省级而言,是全省统一集中的物理数据库,对国家级而言,省级数据库是全国数据库组成部分,国家级与各省数据库通过信息交换系统实现数据交换与共享,既能确保数据库的逻辑独立性,亦能体现全局的逻辑数据独立于物理数据的特征。

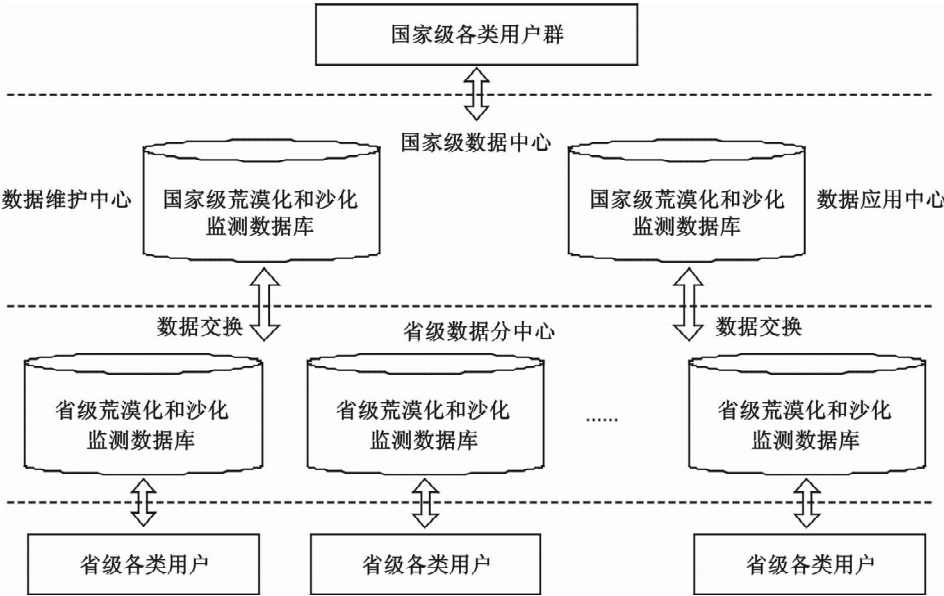


图 3 分布式多专题监测数据库框架

Fig. 3 Distributed and multi-thematic monitoring database framework

3 监测数据库集成与部署模式

3.1 综合集成模式结构

采用“异构同化,同构整体化,分布透明性”的思路和原则^[13-14],国家级监测数据库可由物理分布在省级数据分中心的省级监测数据库综合集成,以省

级监测数据为分区单位划分多个不相交的分片以形成分片模式,采用统一的空间元数据模型和统一的全局空间数据模型^[15],通过模式映射规则从分片模式映射到全局概念模式,构建全局共享的、涵盖全国荒漠化和沙化监测数据的综合集成模式结构。如图 4 所示。

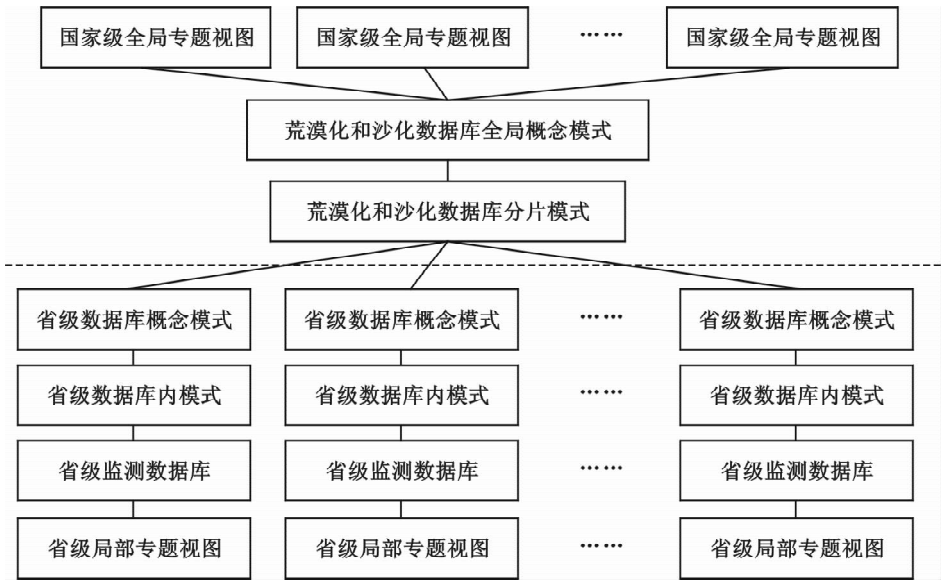


图 4 监测数据库综合集成模式结构

Fig. 4 Monitoring database integration model structure

图 4 中,上半部分是分布式多专题数据库系统全局模式,下半部分是集中式空间数据库系统局部模式。国家级用户利用全局概念模式描述的虚拟数据获取全国范围内的各类业务专题信息,而全局概念模式可利用国家级与省级数据库之间的映射机制,把省级监测数据分片集成为逻辑上为体的国

家级监测数据库,也即,国家级用户可以将这些数据片段集成为全局虚拟数据库中的实际数据,同时也避免了省级数据库异构性导致的模式冲突。省级监测数据库不论是概念模式、逻辑模式还是物理模式,均完全独立,且具有自治性,为省级各类业务用户提供局部视图,而且可通过国家林业专网连接起来。

3.2 管理与部署模式

根据监测数据库框架,结合各级业务部门实际应用需求和管理特点,管理与部署模式可以采用集中与分布相结合的方式。国家级数据库集中部署在

国家级数据中心,省级数据库可采用 2 种模式:省级托管集中式管理模式、省、市、县分级管理的集中与分布相结合模式,则可按照分布式管理原则,避免负载沉重和单点故障等问题(图 5)。

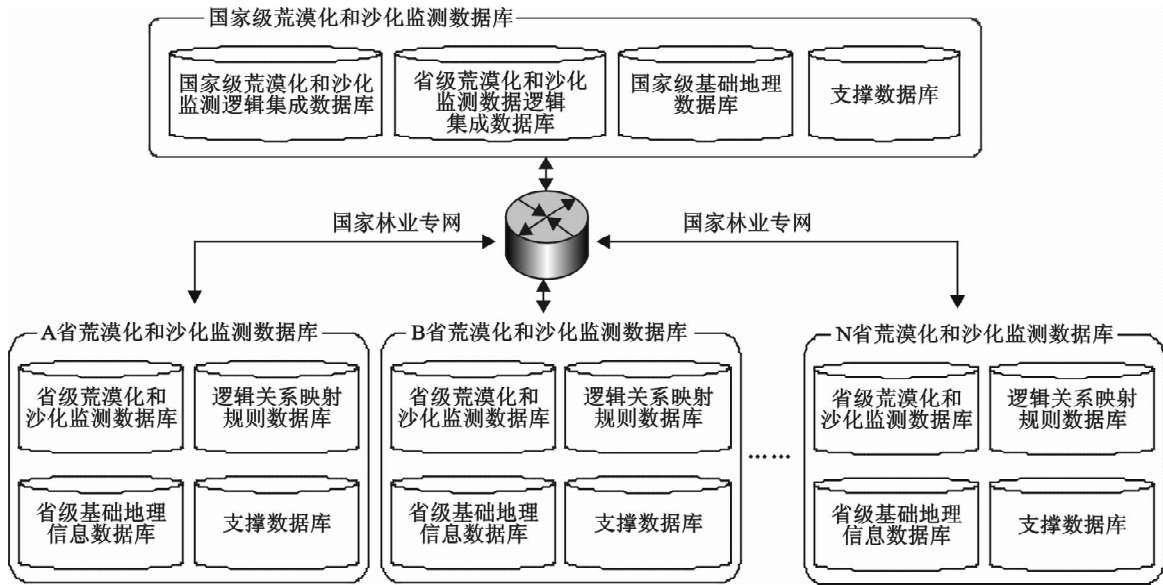


图 5 监测数据库管理与部署模式

Fig. 5 Monitoring database management and deployment models

其中,国家级负责逻辑集成的全局数据库管理,省级负责本省区物理数据库管理,根据各省区不同的网络环境,采用共享交换体系策略,采用空间数据的分布式存储策略^[16],自下而上地进行数据汇集与维护,并利用分配模型^[17]实现数据片段到省级物理数据库中的存储映射,保持数据的一致性与实时主动更新,同时,结合目录体系和信息交换体系,分配和建立各级数据应用权限,确保数据应用不越权操作。

4 监测数据库实现与应用

4.1 监测数据库构建

基于监测数据库框架,结合监测数据资源分类,经过空间数据库的信息建模、语义建模和实体及联系建模过程,利用全国第 3 期和第 4 期监测数据,通过 ArcSDE 9.3.1 在 Oracle 10.2.0.4 g 中构建并整合了监测数据库,设计了 6 个专题数据库:1)基础地理数据库,整合全国和各省的多期监测基础地理数据,确保各期监测数据在空间检索和统计上的唯一性;2)荒漠化监测专题数据库,包括 4 类监测专题数据,在统一的数据库标准和模式下,建立省级荒漠化监测专题数据库,并进行数据校验和规范化整合;3)荒漠化动态专题数据库,包括 4 类动态专题数据,利用 Model Builder 建立多期监测数据综合集成模型,增加动态分析字段,形成荒漠化多期监测数据关

联动态数据库;4)沙化监测专题数据库,包括 7 类监测专题数据;5)沙化动态专题数据库,包括 5 类动态专题数据;6)辅助专题数据库,国家级辅助数据包括:代码、元数据及数据字典、数据交换模型、用户信息、统计报表模型、制图模板和逻辑检查映射等数据;省级辅助数据除此之外还包括:省级与国家之间的逻辑映射规则数据,涉及数据库模式、数据结构、交换方式、时空尺度等一系列映射和链接。

4.2 数据库优化策略

由于监测数据库业务数据总量达 110 GB,涉及矢量数据集 62 个,栅格数据集 20 个,国家级专题图 20 幅,省级专题图 720 幅等内容,需要高效、稳定和可靠的数据访问能力,因此,针对荒漠化和沙化监测海量时空数据的特点,基于 Geodatabase 模型,利用 Oracle 优化器和相关工具,对监测数据库整体进行了物理组织优化和逻辑组织优化。

物理组织优化是合理调整文件物理位置和参数等,以提高系统性能,采用的优化策略有:1)物理文件优化:把 Oracle 的 5 种文件存储在不同磁盘上,避免磁盘 I/O 竞争;2)Oracle 参数调优,调整 Oracle 相关参数,减少 I/O 次数;3)SDE 参数调优,针对实际访问统计数据,基于 I/O 分析,调整 ArcSDE 参数,增强空间数据并发访问性能;4)CACHE 策略,调整缓存参数并缓存索引表,提升吞吐率;5)冗余存储,对数据存储设定了适当冗余,提高本地数据

访问率,同时起到异地备份的效果。

逻辑组织优化是在逻辑上有效组织数据,以达到数据库效率最优,主要采用的优化策略有:1)SQL 语句优化,利用 Oracle 优化器分析和调整造成瓶颈的 SQL 语句,提升数据访问效率;2)索引机制,有针对性地建立属性索引,并在 GBDT 中对空间索引进行优化调整,确保空间格网索引效率;3)多级比例尺制图综合,融合常用比例尺级别数据,转换为地图切片,确保小比例尺级别专题地图的快速显示;4)栅格优化,调整影像压缩格式与压缩比、切片尺寸、金字塔和分幅参数等,提升影像数据访问效率。

通过上述数据组织优化,针对监测数据的应用特点调整其存储和结构关系,有效均衡了各级数据

库服务器之间的负载,确保了海量监测数据的高效存取与平滑扩展,极大地提高了荒漠化和沙化监测数据库系统的整体效率。备份和恢复策略与机制的建立,确保数据库无缝迁移和灾难恢复。

4.3 实际应用

设计并构建的荒漠化和沙化监测数据库在荒漠化和沙化监测信息管理系统的构建过程中得到了有效应用,取得了良好效果(图 6~图 10)。图 6 是由国家级监测数据库全局视图查询出来的荒漠化土地类型及程度专题,图 7 是由青海省和宁夏回族自治区的省级监测数据库局部视图共同构成的两省荒漠化土地类型及程度专题。

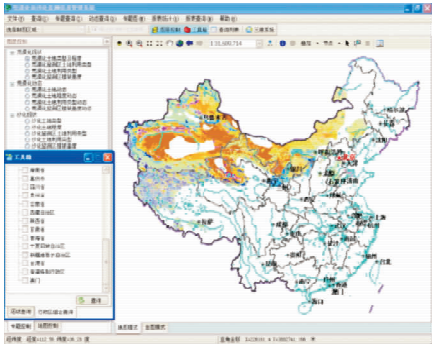


图 6 国家级荒漠化土地类型及程度专题

Fig. 6 National desertification type and extent thematic

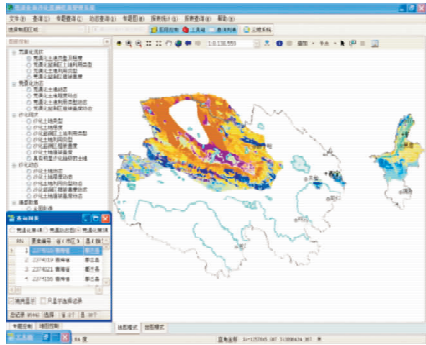


图 7 省级荒漠化土地类型及程度专题

Fig. 7 Provincial desertification type and extent thematic



图 8 全国荒漠化和沙化监测数据服务

Fig. 8 National desertification and sandification monitoring data services



图 9 省级荒漠化土地类型及程度专题服务

Fig. 9 Provincial desertification type and extent thematic services

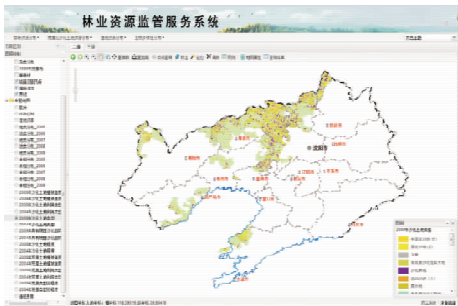


图 10 省级沙化土地类型专题服务

Fig. 10 Provincial desertification type thematic services
监测数据库框架在林业资源监管服务系统——

荒漠化和沙化监管服务子系统的构建中的实际应用,为各级林业资源监管部门提供了良好的分布式多专题监测数据支持。图 8 是部署在国家级数据中心的全国荒漠化和沙化监测数据库提供的数据服务,图 9 和图 10 是由辽宁省数据中心的监测数据库提供的 2009 年荒漠化土地类型及程度专题服务和 2009 年沙化土地类型专题服务。

5 结论与讨论

本研究设计了荒漠化和沙化监测数据库框架,实现并验证了分布式多专题组织管理体系和综合集

成模式,形成了国家级和省级的分布与集中相结合的部署模式,能够使不同管理层次合理地组织和协同共享监测数据资源,能够满足多期监测数据复杂的空间性和动态性管理需求,也能够有效保证省级与国家级监测数据的同步更新,进而为荒漠化和沙化监测体系的各级业务用户高效、全面地管理多期监测数据提供有力的技术支撑,也为各级监测和管理部门全面掌握荒漠化和沙化土地时空分布状态和动态变化格局提供良好的数据服务。此外,本文提出的数据库框架可供林业其它类资源监测数据库构建提供借鉴。

进一步研究可以在以下两个方面继续完善和提升荒漠化和沙化监测数据组织与管理能力。1)在监测数据内容层面上,需要结合治理工程、环境和社会经济数据,利用遥感定量提取栅格结果图^[18],形成综合监测数据库体系,提升防沙治沙决策指挥的科学合理性,亦可进行综合生态系统管理效果评价^[19];2)在监测数据框架层面上,可借鉴云计算思想,把数据资源作为一种服务,在本文提出的框架基础上进一步探讨针对荒漠化和沙化监测数据的云计算服务构建技术。

参考文献:

[1] 林进,周卫东. 中国荒漠化监测技术综述[J]. 世界林业研究, 1998, 11(5): 58-63.
LIN J, ZHOU W D, A review desertification monitoring in China[J]. World Forestry Research, 1998, 11(5): 58-63. (in Chinese)

[2] 张广军,赵晓光. 水土流失及荒漠化监测与评价[M]. 北京:中国水利水电出版社, 2005: 11.

[3] 祝列克. 中国荒漠化和沙化动态研究[M]. 北京:中国农业出版社, 2006: 4-11.

[4] 国家林业局. 第四次全国荒漠化和沙化监测技术规定[Z]. 2008: 1-2.

[5] 国家林业局. 第四次全国荒漠化和沙化监测实施方案[Z]. 2009: 1-7.

[6] 国家林业局. 第三次全国荒漠化和沙化监测技术规定[Z]. 2003: 1-2.

[7] 国家林业局. 全国荒漠化监测主要技术规定(试行)[Z]. 1997: 1-6.

[8] 杨晓晖,张克斌,慈龙骏. 中国荒漠化评价的现状、问题及其解决途径[J]. 中国水土保持科学, 2004, 3(2): 22-27.

YANG X H, ZHANG K B, CI L J, Desertification assessment in China: the state-of-art, problems and solution measures[J]. Science of Soil and Water Conservation, 2004, 3(2): 22-27. (in Chinese)

[9] 国家林业局. 第三期中国荒漠化和沙化状况公报[Z]. 2005.

[10] 国家林业局. 第四期中国荒漠化和沙化状况公报[Z]. 2011.

[11] 邬伦,张毅. 分布式多空间数据库系统的集成技术[J]. 地理学与国土研究, 2002, 18(1): 6-10.
WU L, ZHANG Y. The integrated framework on distributed multi-spatial database system[J]. Geography and Territorial Research, 2002, 18(1): 6-10. (in Chinese)

[12] 朱欣焰,周春辉,芮维. 分布式空间数据分片与跨边界拓扑连接优化方法[J]. 软件学报, 2011, 22(2): 269-284.
ZHU X Y, ZHOU C H, GUO W, *et al.* Distributed spatial data fragmentation and cross-border topological join optimization[J]. Journal of Software, 2011, 22(2): 269-284. (in Chinese)

[13] 邵培英. 分布式数据库系统及其应用[M]. 2 版. 北京: 科学出版社, 2005: 25-26.

[14] 崔铁军. 地理空间数据库原理[M]. 北京: 科学出版社, 2007.

[15] 叶圣涛,张新长. 分布式空间数据库的体系结构研究[J]. 地理信息世界, 2005, 3(3): 47-51.

[16] 孟令奎,张文, WANG F Z. 基于层次化 P2P 协议的网格空间数据库系统模型[J]. 武汉大学学报:信息科学版, 2008, 33(12): 1233-1236.
MENG L K, ZHANG W, WANG F Z. A grid geo-spatial database system model based on layered P2P protocol[J]. Geomatics and Information Science of Wuhan University, 2008, 33(12): 1233-1236. (in Chinese)

[17] 申德荣,于戈. 分布式数据库系统原理与应用[M]. 北京:机械工业出版社, 2011: 49-53

[18] 滑永春,彭道黎,陈鹏飞. 基于 MODIS NDVI 的京津风沙源工程治理区荒漠化动态监测[J]. 西北林学院学报, 2010, 25(6): 210-215
HUA Y C, PENG D L, CHEN P F. MODIS NDVI based dynamic monitoring of desertification in the sandstorm source control project area around Beijing and Tianjin[J]. Journal of Northwest Forestry University, 2010, 25(6): 210-215. (in Chinese)

[19] 王鸣远,杨素堂. 中国荒漠化防治与综合生态系统管理[J]. 西北林学院学报, 2005, 20(2): 1-6.
WANG M Y, YANG S T, Prevention of desertification through integrated management of ecosystems in China[J]. Journal of Northwest Forestry University, 2005, 20(2): 1-6. (in Chinese)