

DOI:10.16031/j.cnki.issn.1003-8035.2016.04.18

## 地质环境数据集成方法探讨及实例应用

喻孟良,任晓霞,曾青石,杨 飞,孙婉秋  
(中国地质环境监测院,北京 100081)

**摘要:**地质环境数据具有多源、异构、复杂等特点。为更好地发挥地质环境数据的作用,针对不同类型的地质环境数据,融合数据格式转换、数据互操作和直接访问三种数据整合方法,提出了一种地质环境数据的集成方法。该方法依托地质环境信息服务平台,实现了矢量空间数据、遥感栅格数据和属性数据的集成整合。最后,将该方法应用到 1:20 万区域水文地质空间数据库,并对具体集成整合过程进行说明,最终将集成后的数据入库到地质环境信息服务平台,提高了数据应用价值。

**关键词:**地质环境数据;数据集成;1:20 万区域水文地质空间数据库

中图分类号: P642.22

文献标识码: A

文章编号: 1003-8035(2016)04-0103-06

## Discussion and application of data integration method for geological environment data

YU Mengliang, REN Xiaoxia, ZENG Qingshi, YANG Fei, SUN Wanqiu  
(China Institute of Geological Environment Monitoring, Beijing 100081, China)

**Abstract:** Geological environment data have features of multi-sources, heterogeneous and complex and so on. To ensure geological environment data play better roles data integration method was summed up for different types of geological environment data using a combination of three methods, which included data format conversion, data interoperability and direct access to data. Relying on geological environment information service platform (GEISP), data integrations were finished for vector spatial data, raster data and attribute data with this method. Finally, taking 1:20 000 regional hydrogeological spatial database as an example, the specific data integration process was illustrated and data value was improved with the integrated data storage to GEISP.

**Keywords:** geological environment data; data integration; 1:20000 regional hydrogeological spatial database

### 0 引言

随着国土资源大调查和地质环境信息化工作的深入推进,我国已经积累了大量的地质灾害、地下水、矿山地质环境、地质遗迹等地质环境调查与监测数据<sup>[1-2]</sup>,建立了全国 1:20 万分幅水文地质空间数据库、1:50 万分省环境地质数据库、1:10 万县市地质灾

害调查数据库、1:5 万地质灾害详细调查数据库、国家级地下水监测数据库、矿山环境数据库等不同类型、不同比例的地质环境数据库<sup>[3-4]</sup>。

地质环境数据具有涉及领域多、覆盖面广、数据格式多样等特点<sup>[5-8]</sup>。地质环境数据根据生产来源和数据格式的不同,数据处理的平台和数据存储的格式也就不同。而随着信息化技术的不断发展,不

收稿日期: 2016-03-02; 修订日期: 2016-05-25

基金项目: 中国地质调查局项目环境地质调查数据集成及服务系统建设(1212011220353)

第一作者: 喻孟良(1978-),男,贵州六盘水人,硕士,高级工程师,主要从事水工环信息化工作。E-mail: yuml@mail.cigem.gov.cn

通讯作者: 任晓霞(1981-),女,山东安丘人,硕士,高级工程师,主要从事水工环信息化工作。E-mail: renxx@mail.cigem.gov.cn

同来源的数据越来越多,而数据处理能力远跟不上生产能力,使得大量数据未能得到有效使用。加之不同数据具有不同的数据来源和数据结构,加剧了数据的破碎化<sup>[9]</sup>,不能为资源共享、信息服务等工作发挥应有作用,严重影响地质环境数据信息化的深入发展。

数据集成技术的出现可以有效提高数据处理能力,使数据得到更加有效的应用<sup>[10-11]</sup>。通过对不同来源的数据进行数据集成,可有效消除不同来源数据的异构性,从而提高数据的利用效率。为此,本文基于地质环境信息化工作基础,以地质环境信息服务为目标,对地质环境数据集成方法进行探索,提出地质环境不同数据类型的具体集成方法。最后,以 1:20 万区域水文地质空间数据库为应用实例,对具体的集成过程进行详细说明,为集成方法提供实际应用。

## 1 地质环境数据集成概述

所谓数据集成是将异构系统中分散的数据转换成另外一个系统的数据,并在此系统上提供服务<sup>[8,10-11]</sup>。目前数据集成方法主要有数据格式转换、数据互操作和直接访问。数据格式转换是将原始数据通过转换软件实现原数据到目的数据的转换。这是当前地理信息系统数据集成与共享的主流方法。数据互操作是在遵循统一标准规范下,透明地获取对方所需的数据。一般采用数据交换格式作为代理实现不同数据的互操作。国际标准化组织(ISO)和 OGC 推出的 GML 数据格式,可以作为空间数据交换格式的标准,但这些数据格式转换标准一般都很复杂,加之目前非 OGC 的空间数据占有很大部分,互操作在数据集成方面有一定局限性。直接访问是在一个 GIS 软件中实现对其他不同数据源数据的直接读取。该方法避免了繁琐的数据转换过程,但对用户要求较高,必须充分清楚不同数据源的数据格式和数据模型,同时编写相对应数据的数据引擎。

地质环境信息平台面向地质环境业务应用,提供空间信息的一体化管理、可视化和信息服务,实现地质环境数据共享与服务,是面向地质环境业务的一个信息展示平台、工作平台和管理平台<sup>[12]</sup>。目前版本的数据架构分为原始库、操作库和产品库,空间数据格式为基于 ArcGIS 的 SHP 格式、数据库为基于 Oracle 和 SQLServer 的两类主流数据库。本文所进行的数据集成工作依托现有的地质环境信息平台,在其框架下,根据地质环境数据特点,在集成过程中,采用三

种方法相结合的方式,以最大程度保证数据集成的高效性和数据的有效利用,针对不同类型的数据分别制定集成方法,将通用的集成原则和数据处理方法实现为自动化辅助工具,形成集成方法的流程化。通过人工操作和自动化辅助工具,最终将集成的数据在地质环境信息服务平台上展示并提供服务。通过地质环境数据整合,可将各异构系统产生的多种地质环境数据在地质环境信息服务平台上提供服务,使其最大程度地发挥作用。

## 2 地质环境数据集成方法

### 2.1 集成自动化辅助工具

在数据集成时,可将常用数据处理方法实现为数据转换工具软件,以辅助数据整合的自动化处理。

数据转换工具软件由切片工具、数据处理工具、影像后处理工具以及其他工具四部分组成。具体功能包括 DEM 切片、DOM 切片、矢量切片、数据打包、数据解包、数据浏览、查看元数据、切片数据库导入、数据拷贝、影像插值、集合格式互转、生成等高线、DEM 噪声检测、生成图框、矢量图层栅格化、投影转换、创建缓冲区、SHP 文件分层、DEM 投影转换、MapGIS 格式转为 SHP 格式。

数据处理工具将需要入库的数据进行数据格式化处理,使入库的数据符合地质环境信息服务平台的数据规范,从而满足地质环境信息服务平台的整体要求(包括地图浏览流畅性、地图操作高效性、地图数据压缩、分层显示等)。同时,数据处理工具减轻了人工进行数据整合的压力,减少人工处理数据可能出现的错误,提高了数据整合的效率和数据本身的精度。

### 2.2 方法简介

从数据类型上看,地质环境数据主要包含矢量数据、栅格数据和属性数据等。矢量数据多以 MapGIS 或者 ArcGIS 的数据格式存在,其特点为按照图层存储业务数据,图层本身可能会带属性数据。栅格数据有遥感影像数据、DEM、DOM 等类型,多以图片形式存储,图片一般比较大,有的可达数 10GB。属性数据多以表格形式存储,有文本、Excel、Access、SQL Server、Oracle 等格式。

一般格式数据的集成方法较为简单不再介绍,本文着重对矢量空间数据、遥感栅格数据、属性数据进行介绍。

#### 2.2.1 矢量空间数据集成整合

矢量空间数据集成采用数据格式转换和数据互操

作两种方式。数据格式转换使用数据转换工具软件并配合 ArcGIS 提供的数据转换或者数据拼接工具来完成矢量数据到 SHP 格式的转换。数据互操作则采用 OGC 的数据格式,发布为数据服务提供数据服务以实现数据的透明获取。具体采用哪种方式视数据情况而定。数据格式转换方式的集成过程包括数据格式转换、坐标转换、数据拼接、数据入库(操作库)、数据切片和金字塔处理等步骤。集成过程涉及哪些步骤由数据最终决定,如:MapGIS 格式的原始数据使用数据转换工具软件完成到 SHP 格式转换,渲染信息等属性数据需要人工辅助完成,ArcGIS 格式的原始数据则不需要转换或者使用 ArcGIS 提供的工具完成到 SHP 文件格式的转换。

### 2.2.2 遥感栅格数据集成整合

遥感栅格数据集成使用数据格式转换方式。集成过程与矢量空间数据集成处理过程类似,使用数据转换工具软件来完成。集成过程包括数据格式转换、坐标转换、影像拼接、数据入库(操作库)、数据切片和金字塔处理,形成统一的遥感影像及 DEM 本底数据库,入库到地质环境信息服务平台进行统一管理。

### 2.2.3 属性数据集成整合

属性数据集成可采用直接访问和数据格式转换两种方式,整合过程视数据情况而定。非地理信息数据(无地理坐标信息)可采用保留原始数据格式在地质环境信息服务平台提供在线浏览或者下载等服务。含有地理信息数据需要根据不同数据类型来书写不同数据格式转换接口,利用 ODBC、ADO 等数据库驱动接口来实现源到目的数据库 SQL Server 与 Oracle 数据库的转换。

## 3 数据集成方法实例应用

下面以 1:20 万区域水文地质空间数据库为例,介绍具体的集成整合过程。

### 3.1 1:20 万区域水文地质空间数据库数据分析

1:20 万水文地质区域空间数据库的矢量数据包括基础地理数据和空间专题数据两个部分,按照标准分幅存储,共 1017 幅。数据具有以下特点:

(1)存储格式为 MapGIS,包括无坐标有渲染信息和带坐标无渲染信息两种数据。无坐标有渲染信息数据,虽然有渲染信息但是不同分幅下的数据渲染方式不同。

(2)原始数据的元数据信息按照省份进行整理,

且元数据内容并不十分详尽。

(3)图层数据文件按图幅存放,图幅的文件夹名称、图层文件名称、要素属性结构有统一的标准。

通过分析 1:20 万区域水文地质空间数据库的数据特点,采用有坐标无渲染信息数据来进行集成整合。集成整合后的 1:20 万水文地质区域空间数据库,应提供按照图幅查询的元数据信息、按照图幅任意提取原始数据或者地图上拉框选择提取原始数据功能以及地图浏览操作等基本功能。

### 3.2 集成过程

1:20 万水文地质区域空间数据库集成方法流程图见图 1。

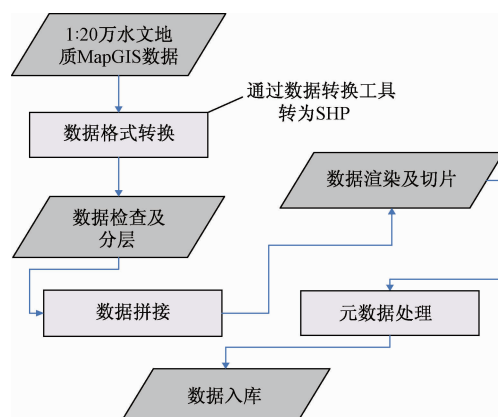


图1 1:20 万区域水文地质空间数据库数据集成方法流程图

Fig.1 Data integration flow graph of 1:200000 regional hydrogeological spatial database

(1)数据格式转换:MapGIS 转为 SHP 文件,采用数据转换工具软件来完成,实现单个文件或者文件夹下的数据格式转换。

(2)数据检查及分层:转换后的 SHP 文件需要进行数据检查,检查的内容包括坐标信息、图层数量和属性数据是否齐全等,当出现问题时对照原始 MapGIS 格式数据进行处理。数据分层处理是指将原来分散在不同文件夹下的同一图层的数据文件存放同一文件夹下。原始数据文件共包括 17776 个图层,分散存储在不同的文件夹下,文件夹以图幅号进行区分。每个图层数据文件按照“图幅号+图层主分类编码+图层子分类编码”的形式进行命名,图层主分类编码+图层子分类编码可以区分不同图层,转换后的 SHP 文件仍遵循原始命名方式,如 E49Z003S11.WP 数据文件,其中 E49Z003 是图幅号,S 代表水文地质,11 为地下水类型,所有 S11 的图层数据文件存放到一个文件夹下。合并后的图层见表 1。

表 1 合并后 1:20 万区域水文地质空间数据库图层表  
Table 1 Merged layers of 1:200000 regional hydrogeological spatial database

| 图层名称      | 图层类型<br>(带√为空<br>间图层) | 图层主<br>分类<br>编码 | 图层子<br>分类<br>编码 | 备注        |
|-----------|-----------------------|-----------------|-----------------|-----------|
| 图幅基本信息    |                       | T               |                 |           |
| 边框(内图框)   | 线元√                   |                 | 21              | 内图框       |
| 地理底图信息    |                       | L               |                 |           |
| 河流        | 线元√                   |                 | 21              |           |
| 湖泊或水体     | 面元√                   |                 | 23              |           |
| 交通        | 线元√                   |                 | 41              |           |
| 境界        | 线元√                   |                 | 61              |           |
| 等高线       | 线元√                   |                 | 71              |           |
| 高程点       | 点元√                   |                 | 72              |           |
| 地貌类型      | 面元√                   |                 | 75              |           |
| 居民地       | 面元√                   |                 | 31              | 居民地真形     |
| 城镇[政府所在地] | 点元√                   |                 | 69              |           |
| 特殊地物点     | 点元√                   |                 | 79              |           |
| 特殊地物线     | 线元√                   |                 | 80              |           |
| 基础地质      |                       | D               |                 |           |
| 地层分区      | 面元√                   |                 | 21              |           |
| 地层界线      | 线元√                   |                 | 23              |           |
| 断裂        | 线元√                   |                 | 29              |           |
| 地层产状      | 点元√                   |                 | 31              |           |
| 水文地质      |                       | S               |                 |           |
| 地下水类型     | 面元√                   |                 | 11~19           | 多层结构      |
| 地下水富水性    | 面元√                   |                 | 21~29           | 多层结构      |
| 地下水水质     | 面元                    |                 | 41              | 矿化度       |
| 地下水水质     | 点元√                   |                 | 43              | 水质超标点     |
| 地下水水质     | 面元√                   |                 | 42              | 水质超标范围    |
| 水文地质特征点   | 点元√                   |                 | 51              | 所有水文地质特征点 |
| 水文地质特征界线  | 线元√                   |                 | 61              | 所有水文地质特征线 |
| 地下水利用规划   | 面元√                   |                 | 81              |           |
| 水文地质专题图   |                       | Z               |                 |           |
| 综合水文地质柱状图 | 面元√                   |                 | 21              |           |
| 地层层序      |                       |                 |                 |           |

(3) 数据拼接:转换成 SHP 格式后,通过 ArcGIS 建立空 GDB 数据库,再将各层数据导入到 GDB 数据库中。数据拼接是在数据分层的基础上把相同图层的数据通过工具拼接成一整幅数据。可通过 ArcGIS 的 Merge 工具合并成一层数据,合并的图层命名标准为:图层主分类编码+图层子分类编码。合并后删除多余按图幅分层的小数据(小数据为原始文件)。

(4) 数据渲染:数据的渲染和符号化按照各图层不同属性通过不同标准处理,可按基础地理底图(L)、基础地质(D)和水文地质(S)三类进行分别处理。①基础地理底图可参照有关标准简单渲染。②基础地质需遵循专业相关标准进行符号化。如地层产状图层需要制作产状符号,并且产状符号应根据属性中的倾角渲染出方向。地层分区参照地质年代表按地层单位分

类,各类按不同颜色填充。③水文地质总体按照《综合水文地质图图例及色标》(GB/T 14538—1993)标准来渲染。如水文地质特征点和地下水水质按照标准规定来制作相应符号。地下水类型,采用黄、褐、蓝、红 4 种颜色代表松散岩类孔隙水含水岩体、碎屑岩类裂隙水含水岩体、可溶岩类岩溶水含水岩体、块状岩类火成岩裂隙水含水岩体。地下水富水性按 SWBFF(富水等级)分类,制作填充符号,用细条稀密程度,线条方向表示富水等级。

(5) 数据切片:数据各图层渲染后就可以做切片处理,切片分为基础地理底图、基础地质、水文地质和水文地质专题图四类。考虑到原始数据的保密等级,整合发布后的数据需要进行脱密处理,因此只保留水文地质专题图数据作为切片后的发布数据,其他数据可做切片处理但最终不进行发布。按照数据精度可做 0~9 级共 10 级数据切片。基础地理底图的各图层按渲染方式通过数据处理工具软件进行切片处理。基础地质的各图层叠加后进行切片,叠加要素按照点、线、面从上到下的顺序进行叠加。水文地质的各图层都是面层,可以单独切图。其它点层和线层按需要与面层叠加一起切图。水文地质专题图包括综合水文地质柱状图和水文地质剖面图,可进行单独切图。

(6) 元数据处理:数据标识信息类包括元数据编号、元数据名称、空间参考系、关键词、边界坐标、比例尺、密级、数据采集起止日期、摘要、质量和分类信息等内容,其中前 4 项是必填信息。由于原始数据只有按照省份的元数据,因此需要建立所有单幅元数据,描述数据名称、图幅编号、关键词等信息,以用于查询浏览、按图幅进行提取处理等需求。单幅的元数据功能见图 2。

(7) 数据入库:包括元数据入库、原始数据入库、经过集成整合后的操作数据入库等操作,入库后的数据由地质环境信息服务平台统一来管理数据。

### 3.3 集成效果

通过将 1:20 万区域水文地质空间数据库进行数据集成整合,实现了原始数据的图层合并、单幅元数据建立、数据分层切片等,集成整合后的详细结果见表 2。将整合后的数据入库,由地质环境信息服务平台统一管理,可进行地图查询、地图浏览、地图定位、元数据查询、按照地理区间快速提取数据等操作,进一步挖掘了数据利用价值,更好地发挥了数据应用效果。

通过地质环境信息服务平台,可进行浏览查询、按照图幅查询等功能。效果图见图 3。



图 2 单幅元数据功能示意图

Fig. 2 Graph of single sheet metadata function

表 2 1:20 万区域水文空间数据库集成整合结果表

Table 2 Table of 1:200000 regional hydrogeological spatial database integration results

| 处理名称    | 个数    | 备注                                                                                                   |
|---------|-------|------------------------------------------------------------------------------------------------------|
| 图幅      | 1017  |                                                                                                      |
| 需转换图层   | 17776 | MapGIS 转 SHP 后图层个数                                                                                   |
| 合并图层    | 26    | 17776 可以分成 26 个数据类别                                                                                  |
| 渲染图层    | 20    | 基础地理 (L21、L23、L31、L41、L61、L71、L72)、基础地质 (D21、D23)、水文地质 (S11、S12、S13、S14、S21、S22、S23、S24、S43、S42、S43) |
| 出图图件    | 1017  | 原始 MapGIS 数据下出图                                                                                      |
| 基础库入库数据 | 1017  | 原始 MapGIS 数据进入平台统一管理                                                                                 |
| 切片数据    | 5     | 地层分区、地下水类型分区、富水性分区 (2 个)、综合图                                                                         |
| 单幅元数据   | 1017  | 需要建立并入平台管理                                                                                           |

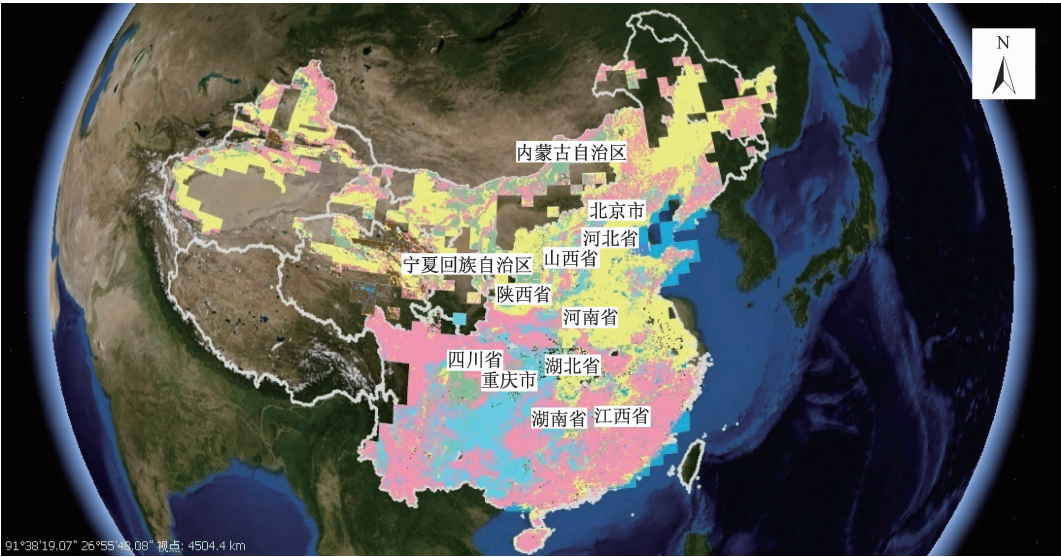


图 3 地图浏览示意图

Fig. 3 Graph of map operation of 1:200000 regional hydrogeological spatial database

4 结语

针对不同类型地质环境数据,分析其特点,依托地质环境信息服务平台,提出了一种地质环境数据集成

方法,实现了从各个异构系统产生的多种地质环境数据到操作库的数据集成,有效缓解了数据应用情况不佳的状况,为地质环境信息资源共享、地质环境信息服务等做好数据支撑。

## 参考文献:

- [1] 刘晓慧, 吴信才, 罗显刚. 面向对象的地质灾害数据模型与时空过程表达[J]. 武汉大学学报(信息科学版), 2013, 38(8): 958-962.
- LIU Xiaohui, WU Xincan, LUO Xiangang. Object-oriented geological disaster data model and spatio-temporal process expression [J]. Geomatics and Information Science of Wuhan University, 2013, 38(8): 958-962.
- [2] 张进德, 袁西龙. 全国矿山地质环境调查信息系统及应用[J]. 中国地质灾害与防治学报, 2006, 17(4): 134-136.
- ZHANG Jinde, YUAN Xilong. Information system of geo-environmental investigation and its application in mining areas in China [J]. The Chinese Journal of Geological Hazard and Control, 2006, 17(4): 134-136.
- [3] 邓青军, 唐仲华. 江汉平原水土地质环境综合监测与评价[J]. 水文地质工程地质, 2014, 41(4): 131-137.
- DENG Qingjun, TANG Zhonghua. Soil-groundwater geological environment integrated monitoring and evaluation in Jianghan plain [J]. Hydrogeology & Engineering Geology, 2014, 41(4): 131-137.
- [4] 赵汀, 赵逊, 田娇荣, 等. 地质遗迹数据库及网络电子地图系统建设——以庐山地质公园数据库和河北省地质遗迹 WEBGIS 系统为例[J]. 地球学报, 2010, 31(4): 600-604.
- ZHAO Ting, ZHAO Xun, TIAN Jiaorong, et al. The establishment of geoheritage database and WEBGIS system: case studies in Lushan geopark and Hebei geoheritages [J]. Geoscientia Sinica, 2010, 31(4): 600-604.
- [5] 候春堂, 冯翠娥, 王轶, 等. 全国 1:50 万环境地质调查信息系统开发初探[J]. 地质通报, 2003, 22(7): 540-544.
- HOU Chuntang, FENG Cuie, WANG Yi, et al. Development of the national 1:500000 environmental-geological survey information system [J]. Regional Geology of China, 2003, 22(7): 540-544.
- [6] 任晓霞, 曾青石, 喻孟良, 等. 地质环境数据交换与数据共享思路探讨[J]. 国土资源信息化, 2015(4): 17-22.
- REN Xiaoxia, ZENG Qingshi, YU Mengliang, et al. Discussion of geological environment data exchange and sharing schema idea [J]. Land and Resources Informatization, 2015(4): 17-22.
- [7] 常青, 杨希, 从常奎, 等. ArcGIS 在矿山地质环境保护与恢复治理中的应用[J]. 中国地质灾害与防治学报, 2014, 25(1): 131-134.
- CHANG Qing, YANG Xi, CONG Changkui, et al. Application of ArcGIS in mining geological environment protection and recovery [J]. The Chinese Journal of Geological Hazard and Control, 2014, 25(1): 131-134.
- [8] 张鸣之, 诸云强, 罗德利, 等. 地质环境数据集成服务及其系统实现[J]. 中国地质灾害与防治学报, 2013, 24(3): 84-89.
- ZHANG Mingzhi, ZHU Yunqiang, LUO Deli, et al. Geo-environmental data integration and service research and software system implement [J]. The Chinese Journal of Geological Hazard and Control, 2013, 24(3): 84-89.
- [9] 杨林, 杜小平, 周顺平. 基于分布式地质图空间数据库的数据交换策略[J]. 地球科学-中国地质大学学报, 2006, 31(5): 659-664.
- YANG Lin, DU Xiaoping, ZHOU Shunping. Data exchange strategy of distributed geological spatial database [J]. Earth Science-Journal of China University of Geosciences, 2006, 31(5): 659-664.
- [10] 何朝阳, 巨能攀, 范强, 等. 多源异构地质灾害监测数据集成技术研究[J]. 人民长江, 2014, 45(13): 94-98.
- HE Chaoyang, JU Nengpan, FAN Qiang, et al. Research on data integration technology of multi-source heterogeneous geo-hazard monitoring [J]. Yangtze River, 2014, 45(13): 94-98.
- [11] 占美志, 何政伟, 李程. 基于 GML 的空间数据集成技术研究[J]. 地理信息世界, 2014(2): 29-32.
- ZHAN Meizhi, HE Zhengwei, LI Cheng. Research on integration technology of spatial data based on GML [J]. Geomatics World, 2014(2): 29-32.
- [12] 诸云强, 周天墨, 喻孟良, 等. 中国地质环境信息服务平台研究[J]. 地球科学与环境学报, 2013, 35(2): 120-126.
- ZHU Yunqiang, ZHOU Tianmo, YU Mengliang, et al. Study on China geo-environmental information service platform [J]. Journal of Earth Sciences and Environment, 2013, 35(2): 120-126.