

Database Design of National Water Resources Monitoring Capacity Construction Project

Sihua Lei, Yongxiang Wu, Gaoxu Wang

State Key Laboratory of Hydrology Water Resources and Hydraulic Engineering, Nanjing Hydraulic Research Institute, Nanjing Jiangsu
Email: shlei@nhri.cn, yxwu@nhri.cn

Received: Jul. 12th, 2020; accepted: Jul. 27th, 2020; published: Aug. 3rd, 2020

Abstract

Database design technology is obtained through many efforts and many years of construction practice. It is an important part of a series of standards formed by the national water resources monitoring capacity construction project. All relevant standards have been applied in water resources management departments at all levels, providing a key technical support for water resources information exchange and sharing. This paper tries to concisely introduce the overall design scheme and main technical features of the database from the aspects of data division, database model, relational table design, etc., and give an effective data archiving and storage scheme according to the current demand for accurate management and data security.

Keywords

National Water Resources Monitoring Capacity Construction Project, Database Design, Object Relation Table, Data Archiving Management

国家水资源监控能力建设项目数据库设计

雷四华, 吴永祥, 王高旭

南京水利科学研究院, 水利部水文水资源工程技术研究中心, 江苏 南京
Email: shlei@nhri.cn, yxwu@nhri.cn

收稿日期: 2020年7月12日; 录用日期: 2020年7月27日; 发布日期: 2020年8月3日

摘要

该数据库设计是在多方努力、历经多年建设实践中取得, 是国家水资源监控能力建设项目形成的一系列标准规范。作者简介: 雷四华(1973-), 男, 江西南昌人, 教授级高工, 研究方向: 水文水资源及水利信息化。

范的重要组成部分；各相关标准规范已在各级水资源管理部门得到应用，为水资源信息交换共享提供了关键性技术支撑。本文力求简练，从数据划分、数据库模型、关系表设计等方面介绍该数据库整体设计方案及主要技术特点；并根据当前对精准管理及数据安全需求，给出了一种有效的数据归档存储方案。

关键词

国家水资源监控能力建设项目，数据库设计，对象关系表，数据归档管理

Copyright © 2020 by author(s) and Wuhan University.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 项目背景

为贯彻落实 2011 年中央一号文件《关于加快水利改革发展的决定》(中发[2011]1 号)和《国务院关于实行最严格水资源管理制度的意见》(国发[2012]3 号)精神，水利部、财政部于 2012 年 9 月批复了“国家水资源监控能力建设项目(2012~2014 年)实施方案”，确定 2012~2014 年初步建成包括水利部、流域机构、省(自治区、直辖市、建设兵团)等共有 40 个节点的国家水资源监控能力建设项目；为了进一步提高重要取水户、重要水功能区和大江大河主要省界断面等在线监测能力，尤其是大中型灌区用水、大型饮用水水源地水质、重要流域水资源管理主要对象的在线监测能力，进一步完善水利部、流域机构、省(自治区、直辖市、建设兵团)等三级信息平台，强化信息资源整合和水资源业务深度开发等，在 2012~2014 年项目建设基础上，2016~2018 年水利部实施了国家水资源监控能力建设二期项目。

根据统一的标准规范建设国家水资源监控能力建设项目，是保证系统顺利建设运行的基础条件和关键因素之一，是确保信息共享、互联互通的重要基础。为此，作为国家水资源监控能力建设项目重要组成内容，按照基础类、信息采集与传输类、数据资源类、空间表达类、应用支撑类、管理类等拟定了项目建设运行维护管理的一系列标准规范[1]。数据库设计规范为该系列中数据资源类规范主要内容，分别有《基础数据库表结构及标识符》(SZY301)、《监测数据库表结构及标识符》(SZY302)、《业务数据库表结构及标识符》(SZY303)、《空间数据库表结构及标识符》(SZY304)、《多媒体数据库表结构及标识符》(SZY305)等标准规范。本文侧重介绍该数据库设计技术特点及在国家水资源监控能力建设项目中的应用。

2. 数据库整体设计

大量的数据是水资源信息化管理基础，水资源管理业务内容众多、数据繁杂，有必要涉及的数据进行适当划分，促使各类数据之间关系清晰表达，提供数据资源建设管理效率。

数据分类划分可有多种方式，所涉及的数据从业务角度划分，可分别为：监测监控业务数据、行政审批业务数据、调度业务数据、应急管理业务数据等内容；按时效性划分：在线监测数据、业务流转实时服务数据、历史资料数据等内容；若从数据自身结构特征划分：数字数据、文本资料数据、视音数据、地图数据等。

本数据库从更新频次、数据来源及其所反映的对象属性特征进行划分，分为基础数据、监测数据、业务数据、空间数据、多媒体数据，如图 1 所示，该划分方式建立的数据库体系，有利于避免数据冗余，促进水资源数据统一管理，日常数据维护管理更便捷，有利于数据库稳定安全运行。

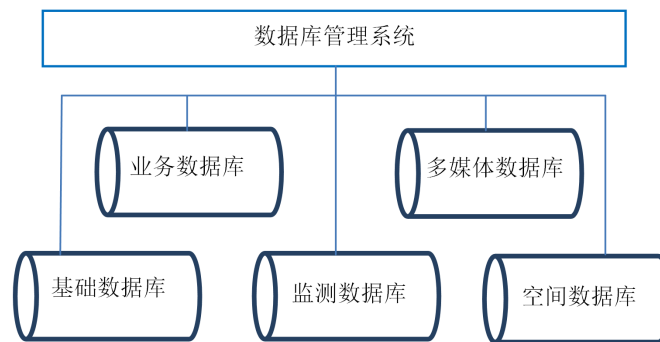


Figure 1. The databases of water resource monitor capacity project
图 1. 水资源监控能力建设项目数据库组成

基础数据库存储水资源管理中涉及的水利基础类、水资源专题类等多种类型实体对象，如河流、湖泊、取用水户、监测站等基本属性信息，这些实体对象具有自身的自然属性特点，其状态一般相对稳定、数据更新维护频度低。如水源地、水功能区、监测站、取用水户、入河排污口、水利组织机构等对象基本信息均存储于基础数据库。

监测数据库内容包括取排水监测数据、行政区断面监测数据、水源地监测数据、水功能区监测数据等为水资源管理服务的专题数据；并引入水文及防汛部门已建成的实时雨水情数据库、水质监测数据库等相关内容。

水资源管理各业务流程的属性信息、资料文档、申报登记、审批准评价、批准通知等业务过程数据存储于业务数据库，这些信息经结构化数据处理后，可分别存入相应数据库表。按业务分类有取水许可审批业务、调度业务、排水行业管理业务、规划管理业务、入河排污口审批业务、水资源费征收使用管理业务以及应急处理业务等数据。

自然空间分布的水利相关对象矢量或栅格数据存储于空间数据库，包括居民地、交通、境界线等基础地理数据，水利工程、水系、流域等基础水利地理数据，及地表水水源地、地表水取水口、地下水取水井、水功能区等水资源专题空间数据。

水资源管理涉及的影像、图片、视频、声音、文档等多媒体数据，对存储索引方式、数据格式进行规定，建立多媒体数据库管理。

3. 面向对象数据库表设计

对象-关系数据库是在关系数据库的基础上加以扩展，增加面向对象的特性，把面向对象技术和关系数据相结合起来。国家水资源监控能力建设项目数据库表设计采用面向对象的思想，根据对象-关系模型的特点，将国家水资源管理系统所用到的众多繁杂数据转换为简单的对象(Feature)、属性(Attribute)及关系(Relation)三类要素进行存储，对象的空间特征通过空间数据库存储，对象的非空间特征通过属性数据库(即基础数据库、监测数据库、业务数据库和多媒体数据库)进行存储，在表设计中分别对应空间对象表、业务属性表、关系表(包括空间关系表、业务关系表)等，其中关系表又可分为同类对象之间关系，如河流与河流之间关系；以及不同类对象关系，如河流与水库之间关系。

4. 对象编码方式

将对象引入数据库中后，每个对象在一定作用范围内应有唯一不变的代码。为了给对象赋予一定代码，制定编码原则与方法是其基本前提。编码是将事物或概念(编码对象)赋予有一定规律性的、易于被人和计算机识别与处理的符号[2]。

国家水资源监控能力建设项目涉及的对象众多，类别复杂，系统性对其进行合理的编码是一项重要的基础

工作。依据国家水资源监控能力建设项目建设需求,在分析已有编码成果的基础上,研究提出国家水资源监控能力建设项目对象的编码方案。根据对象稳定性及依赖特点,确定了3种编码方式[3],即分别以行政区划代码、单位组织机构代码及流域代码等为基础进行编码。

5. 对象关系表设计

从水资源管理对象关联角度考虑,对象之间存在多种关系,有对象间的实体依赖关系、业务逻辑关系、空间位置拓扑关系,如河道断面与行政区界断面之间为实体依赖关系,业务逻辑关系有水利工程与管理单位之间的管理关系、水文测站与水功能区之间的监测关系等,河流与水库之间存在相交空间关系,为了满足数据对象之间的依赖关系、业务逻辑关系、空间关系的管理要求,该数据库设计分别应用参照完整性约束、对象代码关系表、空间拓扑进行了规范。参照完整性约束设计详见参考文献[4],此处主要介绍应用业务逻辑对象关系表及空间位置拓扑关系。

业务逻辑关系一般是因水资源管理工作需要而建立起来的关系,例如取用水户与持有取水许可证的关系。水资源监控管理涉及各对象之间有多种业务关系,除管理关系外,还有监测关系、取水关系、服务关系、从属关系等,对于存在多对多关系的对象之间,其业务逻辑关系,通过关系表实现对象之间映射关系。此处所指的对象关系表主要特点是存储对象代码建立映射关系,一般不存储对象其他属性。

空间位置拓扑关系是空间对象之间现实存在的空间拓扑结构产生的,例如流域与湖泊之间的包含关系,在空间上湖泊处于流域内。按照水资源管理对象特点,将对象空间关系分为流向关系、包含关系、压盖关系、衔接关系、相交关系、跨越关系、依附关系等7种拓扑关系。在空间数据库设计中,对水利基础对象、水资源专题对象等逐一分析,给每一类对象分别建立空间关系表,如水系轴线(河流)与水电站建立压盖关系,与水闸工程、大坝工程等分别建立相交关系。

在业务逻辑关系中,存在传递且较复杂的关系表是基础数据库中取用水的监测关系,其关系描述及应用可参见文献[4]。在数据库表实例建设中,可采用视图方式便利信息检索与统计[5]。

6. 对象属性项设计

6.1. 属性字段定义与描述

在该数据库结构设计中,依据《水利信息数据库表结构及标识符编制规范》(SL478-2010),详细设计了对象各属性的类型及长度、主键、标识符等内容,对各属性存储设计在满足数据库系统要求的同时,力求可便利进行统计与比较,为管理提供可量化分析服务。如入河排污口对象的排污口管径属性,考虑到排污口有圆管、方管、水槽、明渠等多种输送方式,则管径属性不便于一概使用数字填写,为此,增加排污口横截面积属性,达到可兼容排污口各输送方式。

另外,该数据库设计中,详细界定了对象属性含义。对象属性是其特征内容,从数据库表设计至应用系统开发、系统使用、信息发布等均应在同一语义下表述,为了准确清晰界定对象属性语义,在该数据库设计文档,对每一属性字段均给定了描述内容,用于对属性填写内容进行界定,便利在各环节可准确使用。与此同时,在国家水资源监控能力建设项目中建立了《数据字典》规范,软件开发人员可将其应用于功能提示信息,业务管理人员可从其检索信息含义,该规范为各环节用户掌握对象属性语义提供了方便。

6.2. 枚举数据项单纬度划分

在水资源管理中按属性类型划分对象普遍存在,如工程等级、水源类型、排污方式等,为了便于对象分类统计,该类属性一般采用枚举类型存储。在该数据库设计中,对象属性的枚举数据项定义遵循以下几点基本原则。

首先，以类型、类别、规模等为对象属性名称可分类划分的属性，以枚举类型定义为其首选方式。

其次，枚举类型的各数据项内容不应存在交叉，即数据项语义对应的实体不存在交集。如取用水户的取用类别，分为“既取又用”、“只取不用”、“只用不取”等 3 个枚举项。

再次，不漏项，各枚举项所指对象的并集应是该类对象的全集，即全覆盖该属性类型。如取用水户的用户性质属性分为“拥有注册法人的单位”、“法人单位的分支机构”、“自然人”、“其他”等 4 个枚举项。

根据以上原则，在该数据库中设计了 230 余个枚举类型对象属性，其中基础数据库 81 个、业务数据库 143 个、空间数据库 5 个、多媒体数据库 7 个。

7. 各数据库设计与应用要点

7.1. 基础数据库

水资源管理基本对象信息可分为水利基础信息类、水资源专用信息类和监测设备基本信息类等三个大类，水利基础信息类分为自然类、管理类和工程设施类；基本对象的社会自然属性均在该数据库设计。纳入基础数据库设计的对象主要特点有：

- 1) 实体对象具有自身的固有属性，且是其他业务基础的数据，一般不依赖其他业务对象存在；
- 2) 实体对象状态一般相对稳定、数据更新维护不频繁。

按照《国家水资源监控能力建设项目基础数据库表结构及标识符》(SZY301-2018)数据库表存储内容划分，主要包括对象基本属性信息表、对象关系表，共计 140 余张。对象关系表是连接对象的纽带，对象之间的管理关系、监测关系、隶属关系等均需通过关系表建立。

鉴于基础数据库存储内容特点，在基础数据库应用方面须把握以下几方面：

- 1) 数据一致性约束应用。基础数据库中关系表占有较大比例，各关系表均采用外键与对象基本信息表关联，当建立对象关系时，各前置表中应首先存在该对象基本信息。
- 2) 对象属性枚举项应用。基础数据库各对象信息的枚举项是分类统计主要方式，各对象信息入库时其枚举项不宜缺项，以便分类统计分析。
- 3) 数据库视图应用。为了满足水资源管理信息服务需要，每一业务信息描述常需要多个数据库表支持，定制数据库视图是一种有效应对常用的业务查询统计需求的方式[5]。

7.2. 监测数据库

根据水资源管理的业务需求，需要进行监测的水资源管理内容主要是地表水水源地、取用水户、地表水功能区、行政区界断面、地下水取水井等管理对象的状态或行为，《国家水资源监控能力建设项目监测要素》(SZY201-2016)给定监测要素主要包括水质、水量、水位、流量等内容，因监测而产生的数据特点是更新频次较高、数据量大，并且同时需要记录状态行为发生的时间。

监测数据库表分为取用及退排水监测信息类、雨水情监测信息类、水质监测信息类、测站设备工况监测信息类等，共 4 类 36 张库表。

监测数据库中各数据库表一般按照各监测要素分别建立，如取用及退排水的水位、流量监测，分别建立监测点水位监测信息表、监测点流量监测信息表，在同一测站若既有水位又有流量监测，则需分别获取水位、流量信息；同时，取水点水位、退水点水位均在监测点水位信息表存储，测点的取水或退水属性由其基本信息表确定。采用该方式建立数据库表，提高了数据存储的紧密度，在单要素检索方面效率较高。

7.3. 业务数据库

业务数据库表主要存储水资源业务实现过程中由人机交互产生的业务数据，以及业务成果数据。依照最严

格水资源管理制度要求,及“三条红线”内容,可将水资源管理涉及的各项业务分别归类为用水总量控制管理、用水效率控制管理、水功能区限制纳污管理、水资源管理监督考核和水资源支撑保障等五大业务,依此建立各相应主题数据库。

水资源管理各业务涉及的资料信息内容繁杂,结构化数据与非结构化数据常常是焦灼混合,大量的结构化数据需从各非结构化数据中分离提取,如取水许可管理业务中,取水许可申请的地表水取水信息、地下水取水信息、取水标的信息等等,都需从取水许可申请书中获取,进行结构化处理。经结构化处理的数据采用二维表结构存储,并按照数据库表设计要求进行设计。

因水资源管理中的非结构化数据需要纳入业务数据库统一管理,因此,在数据库设计中,对于各非结构化数据需定义基本属性信息,并以数据库表存储,方便检索、比较、排序这些非结构化对象,基本属性信息一般包括方案(报告书)代码、方案(报告书)名称、发布年份、发布单位代码、编制说明等。

同结构化数据一样,非结构化数据也是支撑水资源管理工作的重要内容,基本属性信息表存储于数据库中,非结构化数据本身以多媒体文件存储,并建立相应的多媒体关系表,用于检索多媒体文件。

根据《国家水资源监控能力建设项目业务数据库表结构及标识符》(SZY303-2018)最新发布的业务数据库规范,其内容包括用水总量控制管理类库表 221 张,用水效率控制管理类库表 116 张,水功能区限制纳污管理类库表 101 张,水资源监督考核类库表 22 张,水资源管理支撑保障类库表 104 张。

7.4. 空间数据库

空间数据库是以特定的信息结构和数据模型表达、存储和管理从地理空间中获取的水资源空间信息,以满足对空间信息需求的数据库。水资源监控系统涉及对象的地理空间信息在空间数据库存储,其存储方式可通过数据文件或数据库管理系统实现,空间数据包括空间对象空间位置信息及对象之间的关系信息,均在空间数据库设计,所涉及的空间数据分为国家基础地理信息、水利基础信息、水资源专用信息三大类,《国家水资源监控能力建设项目空间数据库表结构及标识符》(SZY304-2018)包括 52 张空间对象信息表和 123 张空间关系信息表。

空间数据库、属性数据库的表标识符采用一致的命名规则,且同一对象表的主体标识一致;空间数据库与属性数据库表中的字段名称相同时,则标识符也相同。因此,在空间信息管理平台,通过对象代码可在空间数据库中获取空间信息,并能与属性数据库表关联,获得对象其它属性信息。

在水资源管理数据库设计中,对象之间的空间位置关系均通过空间关系表实现。在数据库系统应用中,存在空间位置关系的对象,由空间对象根据拓扑规则自动生成数据行,并存储于空间数据库关系表。为了顺利对空间信息存储管理,空间数据库应与属性数据库应同步建设。

7.5. 多媒体数据库

多媒体数据库主要是对各多媒体数据的索引方式、各类多媒体数据格式进行规定,并建立数据库进行管理。在该数据库设计中,多媒体数据采用计算机文件形式存储,多媒体基本信息及扩展信息在多媒体数据库存储,数据库系统与文件存储系统通过文件名进行关联。

多媒体文件基本信息的内容包括文件标题、关键词、文件类型、文件路径、文件大小等内容,扩展信息则是存储特定类型多媒体资料的个性信息。在建立多媒体基本信息表基础上,按照对象分别建立对象与多媒体的关系表,存储对象与多媒体文件关系。关系表存储方式多媒体数据库示意图如图 2 所示。该设计方案优点在于不同的对象可以和同一个多媒体基本信息建立关系表,而不同类型的多媒体信息可通过扩展信息去查找,相比于逐个对象建立多媒体数据库表或各类型多媒体分别建立基本信息表,提高了检索效率,并减少了数据冗余。

《国家水资源监控能力建设项目多媒体数据库表结构及标识符》(SZY305-2018)包括多媒体资料基本信息表、文档资料信息表、图片资料信息表、视音频资料信息表等 4 张。

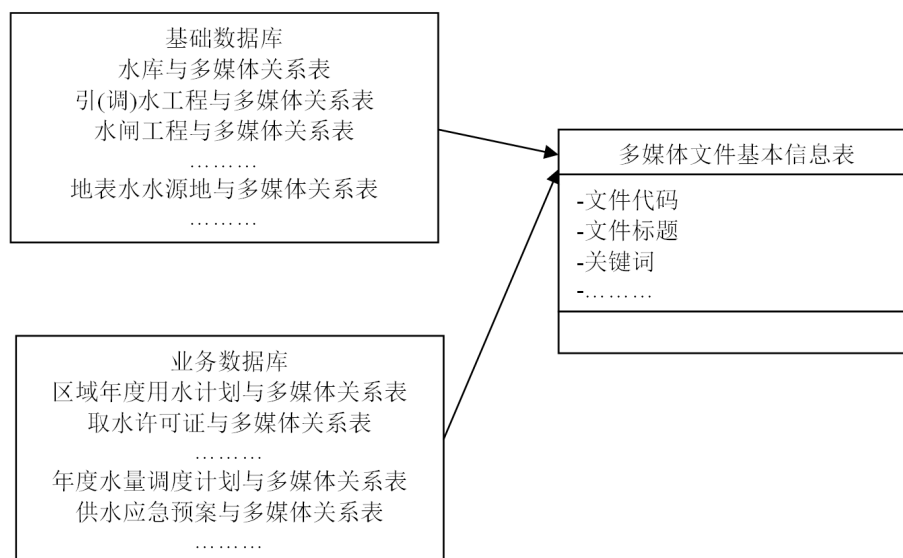


Figure 2. Diagram of multimedia database with relation table storage mode

图 2. 关系表存储方式多媒体数据库示意图

在水资源监控系统建设时,通过以下几种方式可充分利用该多媒体数据库优点。

- 1) 当建立对象基本信息表时,基础、业务等数据库只需再增加一张与多媒体文件的关系表,不需在多媒体数据库增加新的数据库表;
- 2) 在水资源管理系统应用过程,建立业务对象时,建立该对象与任一已存在的多媒体文件关系,可对多媒体文件与业务信息统一管理;
- 3) 对于具体多媒体文件,可归属于多个对象的多媒体,只需建立对象与多媒体文件关系即可,而不必重复存储多个文件副本。

8. 各数据库存档设计

在水资源监控系统运行过程,基础数据、业务数据等将不断有维护更新,若采用数据日志管理模式,其信息量及便利性均存在不足,不能满足对历史数据及时全面回溯。充分利用存储资源优势,对被更新的原数据进行归档保存,已成为经济合理、技术可行的选择;而且,从数据安全考虑,对各数据库表建立相应的归档数据库表(或称为历史数据库)已是普遍要求。

针对该数据库,各归档数据库表结构可按以下方式设计:

- 1) 在本数据库表基础上增加“入归档数据库时间”字段,并作为主键;该时间为数据行写入归档数据库时间,也是数据更新时间。因各对象为授权维护,使用计算机系统时间可满足要求,可保证其唯一性。
- 2) 在本数据库表基础上,去掉外键约束,避免参照完整性对归档数据前置表的要求。
- 3) 当对数据进行修改、删除等更新操作,需同时把被更新的数据写入归档数据库,应在系统实现时一并完成相关代码编写。

该方案主要优点有:①被更新的数据均写入归档数据库,现状有效数据与被更新的数据分库存储,数据存储清晰,查询统计等应用操作便利;②采用入归档数据库时间可简便地追溯数据历史变更过程。

9. 应用与展望

该数据库采用面向对象统一数据模型的思想设计,梳理水资源管理数据库与已建的实时雨水情数据库、基础水文数据库、水质数据库之间的关系,总结归纳已建水利行业数据库标准规范在各级水利部门的应用,提出

了基础数据库、监测数据库、业务数据库、空间数据库、多媒体数据库的库表结构、标识符命名规则、数据类型约定、主键及索引以及数据库物理设计,同时完成了数据库的元数据和数据字典设计,解决了水资源数据库的高效存储检索、方便应用以及与外部数据库的数据共享问题。

该数据库已成功应用于水利部、流域、省和地市各级水资源监控能力建设项目中,对国家水资源监控系统三级平台之间实现互联互通、信息共享和业务协调发挥了关键作用[6]。经实践证明,根据该技术建设的数据库系统运行良好,具备冗余性小、可扩展性强、空间数据与业务数据相互独立且无缝链接等方面的技术优势。该成果可为其他水利信息化系统的数据储存和应用提供借鉴,具有广阔的推广应用前景。

另外,当前信息技术日新月异,甚至可称为信息技术爆炸时代,大数据、云服务、智慧水利等新技术与名词受到技术研发及信息服务与管理人员竞相追捧与探索,技术发展不仅激烈地改变人们生活习惯,同时对经济、文化等社会方方面面都产生了深远影响。本数据库设计以传统关系数据库理论方法为基础,兼顾了对空间数据、多媒体数据管理。在该数据库基础上,还有许多水资源管理面临的困难不能解决,如:在有效监测不足的现状下,农业用水量计算统计;在数据资源不足的情况下,水资源制度落实效果、水资源节约保护成效等考核依据不充分。而当前蓬勃发展的大数据技术在本设计方案中还未得到应用,实践是新技术最终价值体现的唯一途径,探索使用大数据的收集、存储、挖掘技术解决水资源管理面临的问题具有较好的现实意义。

基金项目

国家重点研发计划课题(2016YFA0601703, 2016YFC0401005)。

参考文献

- [1] 国家水资源监控能力建设项目办公室. 国家水资源监控能力建设项目标准制修订情况[EB/OL]. <http://szyjk.mwr.gov.cn/szhjnl/>, 2019-4-29.
National Water Resources Monitoring Capacity Building Project Office. National water resources monitoring capacity building project standard preparation and revision. <http://szyjk.mwr.gov.cn/szhjnl/>, 2019-4-29. (in Chinese)
- [2] 中国标准研究中心. 分类与编码通用术语: GB/T 10113-2003 [S]. 北京: 中国标准研究中心, 2003: 1-2.
China Standards Research Center. General terms of classification and coding: GB/T 10113-2003. Beijing: China Standards Research Center, 2003: 1-2. (in Chinese)
- [3] 王高旭, 吴永祥, 雷四华, 等. 国家水资源监控能力建设项目对象编码研究[J]. 水利信息化, 2018(6): 35-38.
WANG Gaoxu, WU Yongxiang, LEI Sihua, et al. Study on object coding of national water resources monitoring capacity construction project. Water Conservancy Informatization, 2018(6): 35-38. (in Chinese)
- [4] 雷四华, 韩晴. 国家水资源监控能力建设项目数据库数据完整性约束设计分析[J]. 水利信息化, 2018(3): 28-31.
LEI Sihua, HAN Qing. Data integrity constraint design analysis of national water resources monitoring capacity construction project. Water Conservancy Informatization, 2018(3): 28-31. (in Chinese)
- [5] 金喜来, 徐瑶. 国家水资源监控能力建设项目数据库标准规范应用实践[J]. 中国水利, 2015(19): 59-61.
JIN Xilai, XU Yao. Application practice of national water resources monitoring capacity construction project database standard. China Water Resources, 2015(19): 59-61. (in Chinese)
- [6] 张丽丽, 李臣民, 刘圆圆. 国家水资源监控能力建设项目标准应用情况分析[J]. 水利信息化, 2016(5): 39-44.
ZHANG Lili, LI Chenmin, LIU Yuanyuan. Analysis on application of national water resources monitoring capacity construction project standard. Water Conservancy Informatization, 2016(5): 39-44. (in Chinese)