

# Prospect for Big Data Applications in Water Resources Management

Yuwei Liu<sup>1</sup>, Dongrun Liu<sup>2</sup>, Xianyun Chen<sup>3</sup>

<sup>1</sup>Bureau of Hydrology, Changjiang Water Resources Commission, Wuhan Hubei

<sup>2</sup>Hunan Provincial Hydrology and Water Resources Survey Bureau, Changsha Hunan

<sup>3</sup>Nanjing Hydraulic Research Institute, Nanjing Jiangsu

Email: [liuyw@cjh.com](mailto:liuyw@cjh.com), [liudongrun@163.com](mailto:liudongrun@163.com), [chenxy@nhri.cn](mailto:chenxy@nhri.cn)

Received: Sep. 29<sup>th</sup>, 2015; accepted: Oct. 14<sup>th</sup>, 2015; published: Oct. 28<sup>th</sup>, 2015

Copyright © 2015 by authors and Hans Publishers Inc.

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

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



Open Access

---

## Abstract

This paper introduces the basic concept of big data and its application in related field, analyses the existing problems both in current water resources management and its information system, based on the basic principle of big data and the development trend of information technology, borrowing big data application practice from other fields, taking the national water resources monitoring ability construction project as an example, analyses the data requirements in the water resources management and information system platform construction in the era of big data, taking the existing or possible problems in the actual work as examples, explores and prospects the application of big data in water resources management.

## Keywords

Big Data, Water Resources Management, Application

---

# 大数据在水资源管理中的应用展望

刘予伟<sup>1</sup>, 刘东润<sup>2</sup>, 陈献耘<sup>3</sup>

<sup>1</sup>长江水利委员会水文局, 湖北 武汉

<sup>2</sup>湖南省水文水资源勘测局, 湖南 长沙

<sup>3</sup>南京水利科学研究所, 江苏 南京

作者简介: 刘予伟(1959-), 男, 高工, 主要从事水文水资源工作。

文章引用: 刘予伟, 刘东润, 陈献耘. 大数据在水资源管理中的应用展望[J]. 水资源研究, 2015, 4(5): 470-476.

<http://dx.doi.org/10.12677/jwrr.2015.45058>

Email: [liuyw@cjh.com](mailto:liuyw@cjh.com), [liudongrun@163.com](mailto:liudongrun@163.com), [chenxy@nhri.cn](mailto:chenxy@nhri.cn)

收稿日期: 2015年9月29日; 录用日期: 2015年10月14日; 发布日期: 2015年10月28日

## 摘要

本文介绍了大数据的基本概念及其在相关领域的应用实例, 分析了当前水资源管理及其现有信息系统存在的问题, 依据大数据的基本原理和信息技术的发展趋势, 借鉴和参考其他领域大数据的应用实践, 以国家水资源监控能力建设项目为例, 分析了大数据时代水资源管理数据需求与信息系统平台的构建要求, 并结合实际工作中已经或可能出现的问题举例, 探讨和展望大数据在水资源管理中的应用前景。

## 关键词

大数据, 水资源管理, 应用

## 1. 引言

近年来, 气候变化和人类活动的影响, 水资源时空分布变化突出, 水极端事件频发, 尤其随着城镇化、工业化和农业现代化进程的加快, 各类用水户大量增加, 用水类型增多, 水资源利用形势越来越复杂, 水量分析和配置监控越来越困难, 加上退水增多, 退水水质混杂, 污水处理难以到位, 退水影响越来越难以准确计算和评估, 致使水资源管理难度加大, 其中重要的原因之一就是现有的技术手段很难与日益提高的水资源管理要求相适应, 水资源信息缺乏共享和综合利用, 水资源管理信息技术手段落后于当前主流发展趋势。已有的相关信息系统, 比如水文信息采集系统, 防汛信息系统, 水资源信息管理平台, 均缺乏跨部门、跨行业数据共享和综合利用的机制和手段, 基本上还没有或很少应用大数据及其相关新技术。

当前, 随着各种通信和计算机技术深度融合, 移动互联网、智能传感网、物联网和云计算等新技术发展日新月异, 渗透并改变着人类社会的各个方面, 数据信息和知识呈爆炸式增长, 大数据时代已悄然走来。相比较而言, 大数据应用已在商业和金融领域开展较为广泛, 但水资源及其相关领域还尚未开展或正在起步, 由于水资源在社会经济发展中的作用和地位越来越重要, 水资源管理所涉及的面越来越广, 难度越来越大, 尤其是国家出台最严格的水资源管理制度, 迫切需要采用大数据等新技术来创新管理, 本文试图依据大数据理论, 借鉴和参考其他领域的应用实践, 探讨和展望大数据在水资源管理中的应用前景。

## 2. 大数据概述

### 2.1. 大数据的基本概念

大数据概念的提出可以追溯到《自然》杂志 2008 年 9 月专刊中发表的文章“Big Data: Science in the Petabyte Era”。此后, 大数据概念被广泛应用和传播。维基百科对大数据的定义是: “大数据是由于规模、复杂性、实时而导致的使之无法在一定时间内用常规软件工具对其进行获取、存贮、搜索、分享、分析、可视化的数据集”。互联网数据中心将大数据定义为: 为更经济地从高频率的、大容量的、不同结构和类型的数据中获取价值而设计的新一代架构和技术[1]。

目前, 对大数据概念有多种解析和定义, 但基本一致的是所谓大数据的“4V”理论, 即所具备的 4 个基本特征: 规模化(volume)、多样性(variety)、高速率(velocity)和真实性(veracity)。但还有另一种非常相似的 4V 理论, 其前面的 3V 内容定义相同, 而第 4 项变成了有价值(value)。另外, 还有学者提出大数据呈现出“4V + 1C”的

特点,除了对“volume”、“variety”和“velocity”赋予了一些新的释义,又提出第4V——“vitality”,即数据可动态持续的不断更新与扩充,响应速度快,时效性强,有广泛应用的活力等;关于对1C——“complexity”的描述,主要是指鉴于数据结构的复杂,需要有新的技术方法,来满足异构数据统一接入和实时数据处理等方面的需求[2]。

综上所述,大数据的概念尚没有非常统一的定义和表述方式,但其核心和实质内容是基本一致的,即:数据容量大,数据形式多样化、非结构化特征明显,数据存储、处理和挖掘异常困难,其中蕴含着巨大的应用价值。多年来,人们熟悉的主要是基于结构化数据的分析与应用。有资料显示,在实际应用中有超过85%的数据属于半结构化或非结构化数据,这些数据广泛存在于社交网络、互联网、物联网、电子商务之中,如办公文档、文本、XML、HTML、各类报表、图片、图像、音频和视频信息等。获取、存储、挖掘处理上述非结构化或半结构化的数据,是大数据应用所面临的重要挑战。

## 2.2. 大数据的应用价值

进入大数据时代,人们在研究与分析某个现象时,将会更多地依赖和使用全部数据而非抽样数据;不需要一味地追求数据的精确性,而要适应数据的多样性、丰富性;应用大数据之所以存在巨大的潜力,其关键是有可能从支离破碎的、看似冗余和无序的、毫不相干的海量数据中抽炼出真知灼见,从中发现新问题,找到解决问题的新途径和新方法。人们通过对大数据的分析,更容易获得或更关注的是事件产生的结果(是什么),而不是产生的原因(为什么)。对数据之间的相关性分析,胜于对因果关系的探索,当然对相关性有足够的了解,必然促进对因果关系的认知。大数据继承了统计科学的一些特点,但又不同于传统的逻辑推理研究。大数据的应用更多的是遵循“从数据直接到价值”的思路。数据规模越大,处理难度也就越大,对其进行挖掘分析所能得到的价值也就越大。大数据既能快捷准确地提供社会经济的整体变化用于决策,也能让普通的社会公众获得直接可靠的信息服务。

在一些领域,已经衍生出许多基于大数据分析的模式,最典型的案例是谷歌公司的“流感预报”,其方法是构建一个基于每天来自全球的30多亿条搜索指令的系统,这个系统在2009年甲型流感爆发之前就开始对美国各地区成功地进行“流感预报”和“谷歌流感趋势”服务。在农业领域,硅谷有个气候公司,从美国气象局等数据库中获得几十年的天气数据,将各地降雨、气温、土壤状况与历年农作物产量的相关度做成精密图表,预测农业产量,向农户出售个性化保险。在商业领域,沃尔玛公司通过分析销售数据,了解顾客购物习惯,得出适合搭配在一起出售的商品,还可从中细分顾客群体,提供个性化服务。在金融领域,阿里巴巴公司根据淘宝网上中小企业的交易状况,筛选出财务健康和讲究诚信的企业,对他们发放无需担保的贷款[1][3]。

大数据的这种特性必然也适合在水资源管理上的应用,水资源在时空上变化大、周期长、影响因子复杂。在水资源管理上,关联领域多,涉及面广,信息资源庞大,对于出现的问题往往很难直接找到原因,管理和决策难度大,但通过大数据应用技术可获得各种相关信息,找出问题产生的主要影响因子,为水资源管理提供技术支撑。

## 3. 基于大数据应用的水资源管理数据需求分析

现有的水资源管理系统是由水利部门建设和管理,其数据需求多从水利行业出发进行分析,所建立系统受数据来源的限制,其功能、作用、时效性和服务面相应受到限制;同时,由于受当时信息技术条件的限制,无法处理和分析数据结构不同、数据来源错综复杂的庞大的信息资源,因此,基于大数据应用的水资源管理,就是要打破行业和部门界限,全面分析其数据需求。

### 3.1. 基础数据

水资源管理的目的是实现对水资源的合理开发、综合治理、优化配置、全面节约、有效保护,其首要条件

是及时详细掌握水资源及水资源开发、利用和保护状况，准确把握水资源变化特征和规律，这需要大量的基础数据支撑。这些数据主要包括：

(1) 水文及水资源监测、水环境及水生态监测数据、雨水情测报数据等，这些数据主要来源于水文部门；

(2) 用水总量相关数据，包括地表和地下水取用水监测数据、水量分配、取水许可、水资源调度、用水户及水权交易等数据，这些数据主要来源于水行政主管部门；

(3) 用水效率相关数据，包括节水指标、计划用水指标、用水定额、用水效率，非常规水等数据。这些数据主要来源于工业及其相关行业管理部门；

(4) 水功能区纳污量相关数据，包括水功能区划、水域纳污能力核定、水功能区水质监测及达标评价、入河排污口调查监测、饮用水源地保护等数据。这些数据主要来源于水利环保等部门；

(5) 水资源经济核算数据，包括耗水总量核算、排污总量核算数据，水资源费核算、分行业供水水费核算、水权转让交易核算、生态补偿标准定量核算等数据。这些数据主要来源于经济及其相关管理部门。

### 3.2. 关联数据

我国现有的水资源管理，是以水利部门为主、环保、农业、国土等其他政府部门兼有职责，要适应社会经济发展和最严格的水资源管理，必须解决数据关联和共享的问题。同时，社会其他部门、相关团体和公民也越来越多地参与或涉及水资源监督和管理，由此而产生的与水资源有关的(包括系统和非系统的、间接相关的)数据，比如通过互联网以网站、论坛、微信、微博和博客等方式发布的有关信息，可以被利用和挖掘，可以弥补已建信息监控网络未覆盖范围的信息，并可以借鉴这些数据，对常规渠道的信息进行佐证性评估分析。水资源管理链条上各个相关部门收集或发布的众多信息中，有时存在不对称或差异性，通过数据溯源和分析，可以对数据的准确性和真实性作出评估，尤其是有的突发水事件信息，往往首先是从非正规监控部门得到的，因此，对这类数据的关联、共享、挖掘和利用，是提高水资源管理效能的重要途径。

## 4. 水资源大数据应用展望

### 4.1. 数据采集与整合

上述基础数据和关联数据的集合即基本构成水资源大数据。水利环保等部门主要借助卫星遥感技术、全球定位系统、地理信息系统、无线传感网络等现代化量测技术和数据采集手段，快速有效获取并存储其各自管辖范围内的基础数据，这些数据都分散存储在各自的数据库。目前，还没有形成各行业部门之间数据共享的机制，要实现流域水量水质、供水水源地、规模以上取用水、出入境水量、地下水超采区、入河排污口的在线实时监测等数据的整合与共享，需要构建面向各级部门的水资源管理信息系统平台。

另外，对于分散在不同机构和企业、各工厂和灌溉系统管理单位、居民和工业园区等相对分散的关联数据(包括结构化、非结构化和半结构化数据)，需要借助于移动互联网、社交网、云计算等技术将其整合到一个灵活、开放的高性能平台上，加上前述的基础数据，形成一个互联互通的数据共享网络体系，即所谓水资源大数据云平台。

目前，各相关行业和部门对数据的管理多采用传统关系型数据库，其数据是用统一的结构来表示，即称之为结构化数据，但随着互联网的发展，可以把超文本文件传送到用户的浏览器里，不仅可以支持文字、图形、图像、声音、视频等多媒体信息，还可以支持一些较为复杂的对象，比如电子表格。此类信息根本无法用数字或者统一的结构表示，即称之为非结构化数据。因此，基于大数据的应用，就是要在原有关系数据库的基础上，补充构建非结构化数据库。这两类数据库将是水资源大数据云平台的基础。

通过云平台对分布在各部门和行业数据库的数据进行关联、清洗、甄别、挖掘和提炼，选取有效和有用的数据，供决策者对水资源实施动态有效的监控、调配和管理，如图1所示。

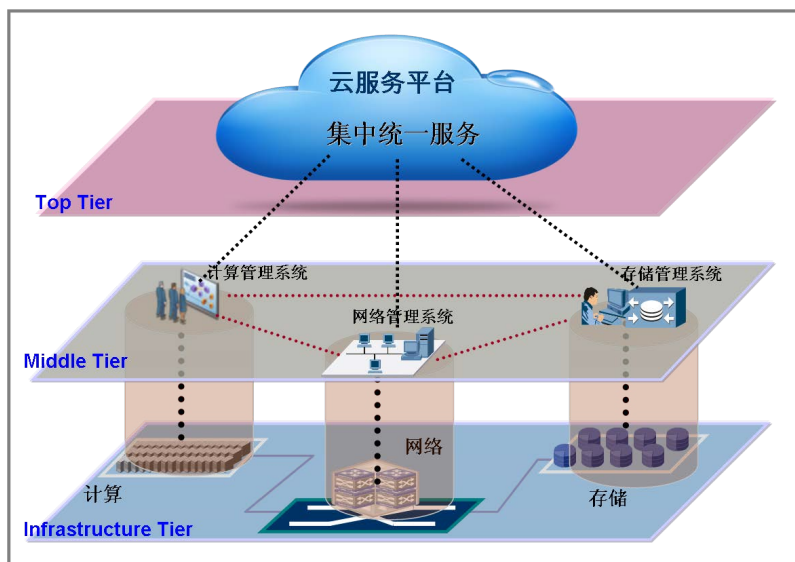


Figure 1. Schematic diagram of information system cloud platform  
图 1. 信息系统云平台示意图

## 4.2. 实例分析

正在建设中的国家水资源监控能力建设项目，其目的是初步形成与实行最严格水资源管理制度近期目标相适应的国家水资源监控能力，为支撑水资源管理定量考核工作奠定基础。

根据项目建设总体设计方案[4]，系统的业务应用主要包括：水资源信息服务、水资源业务管理、水资源应急管理、水资源调配决策四大方面。其数据范围包括基础信息、监测信息、业务信息、空间信息和多媒体信息：在纵向上实现水利部、全国 7 个流域机构、31 个省(自治区、直辖市)水利(水务)厅(局)及新疆生产建设兵团水利局的水资源数据共享交换；在横向上实现与水利防汛抗旱、电子政务系统等水利业务系统的交互关联，同时在技术上考虑实现与农业、环保、统计、气象等部门的交互，实现内外网之间的数据交换。从系统总体构架来看，该项目建设具有大数据应用的理念和方向，是目前为止我国实施水利信息化项目中应用现代信息技术较全面、较先进的信息系统，在未来的拓展应用开发中，随着海量数据的扩大，将加入移动互联网、云计算、数据挖掘等新技术的应用，从而形成基于大数据应用的水资源监控云平台。图 2 显示了该建设项目数据交换共享框架。

## 4.3. 应用展望举例

大数据应用不同于传统数据应用的最大区别就是如何通过海量数据，挖掘出所有的隐藏信息。这些数据往往是通过各种不同内、外部渠道收集的，而且很多是非结构化数据。尤其是在水资源突发事件的应急管理、水资源违法违规的监管难点和盲点、水资源调查评价、论证和水资源开发利用保护规划等方面，有着非常重要而深远的应用前景。为了理解大数据分析的价值和在大数据分析的帮助下如何解决复杂的问题。以下是一些举例分析。

(1) 环保或水文部门发现某一个地方或水域发生严重污染事件，通过水资源大数据应用平台，可马上作出污染对水环境以及周边的影响预测；水利部门和自来水公司可立即根据季节和用水量，实时调整水库蓄泄水量和供水分配；所在地环保、水利、市政等部门可以及时知晓城镇用水和调配情况，并采取相应的应急措施。

(2) 一些城镇和企业的污水处理厂长期未按设计生产能力运行、实际处理污水量未达到核算认定要求，超标排放等问题非常普遍，监管难度大，但用电量是衡量污水处理厂运行效益的一个重要指标，通过电网的关联数据，可以查到污水处理厂时段总的耗电量，从而监测评价其污水实际处理量。大数据应用平台建立，可以达到

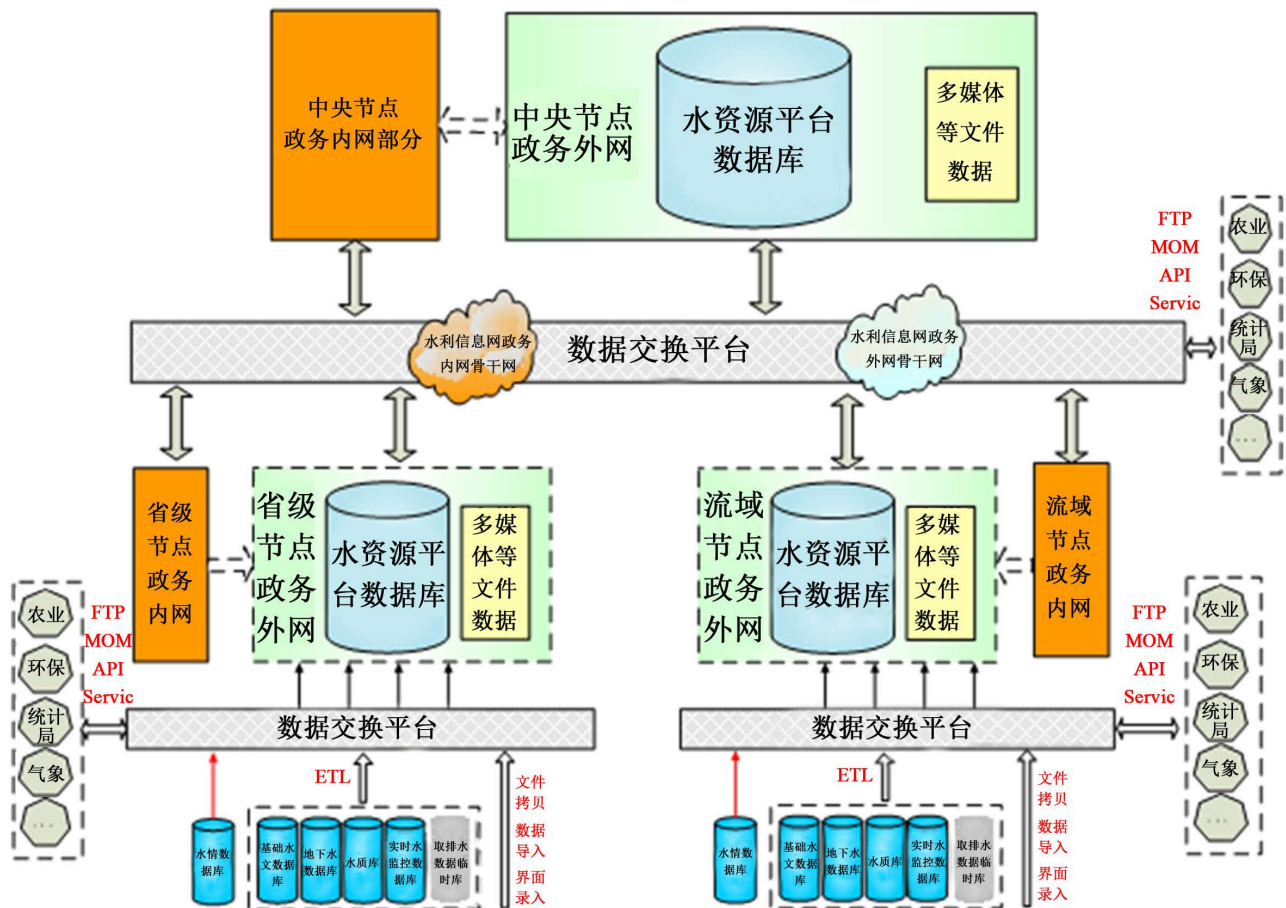


Figure 2. Data exchange and sharing framework of the national water resources monitoring ability construction project  
图 2. 国家水资源监控能力建设项目数据交换共享框架图

实时监控的目的。

(3) 一般来说，政府实施水资源管理与治理，其管理和治理的过程及效果，除了由政府自身来监督和评估外，很难有效通过市场来实现，毕竟市场带有盈利动机，甚至市场的某些取用水或排污行为恰恰是政府需要管理和治理的对象。因此，监督和评价水资源管理有关措施落实情况 and 效果，关靠政府部门监管可能存在难点和盲点，需要社会和公民参与。一些非营利组织、社会团体和公民，通过网络以微博、微信、论坛等方式发布的文字、图像、视频信息，揭露和反映一些违规现象和事件，起到了有效的监督作用，通过对这些数据和关联信息的挖掘，为水资源管理和治理提供及时有效的服务。

(4) 通过某批次大米查出重金属超标，可追溯到大米产地土壤和水的污染状况，从而溯源到当地或上游化工、冶炼、化肥、采矿等企业的超标排放以及监管漏洞等违规行为，继而监测到水域的水质状况。这些数据往往来自质检部门和新闻媒体。

## 5. 结语

随着大规模水资源监测网络的建设和运行，水资源基础数据量日趋庞大，同时，水资源管理链上的关联数据和社会化数据，也是水资源管理可利用的有价值的信息资源，信息共享和网格化管理的技术日趋成熟，为海量异构数据的预处理、存储创造了条件，在此背景下，催生了云计算、大数据应用技术。大数据在水资源管理中的应用，能更有效地对水资源数据进行甄别和质量评估，可实现多因素、多区域的水资源变化趋势分析和预

测、为各种水资源管理研究、为水资源公共事件的监测和预警、为水资源管理中的决策与政策评估提供强有力的信息和技术支撑。

## 基金项目

973 计划资助项目：长江中游通江湖泊江湖关系演变及环境生态效应与调控(项目编号：2012CB417000)。

## 参考文献 (References)

- [1] 维克托·迈尔-舍恩伯格, 肯尼思·库克耶, 著, 盛杨燕, 周涛, 译. 大数据时代: 生活、工作与思维的大变革[M]. 杭州: 浙江人民出版社, 2012: 12.  
Viktor Mayer-Schonberger, Kenneth Cukier, Write, Sheng Yangyan, Zhou Tao, Trans. Big data: A revolution that will transform how we live, work, and think. Hangzhou: Zhejiang people's Publishing House, 2012: 12. (in Chinese)
- [2] 李国杰, 程学旗. 大数据的研究现状与科学思考[J]. 中国科学院院刊, 2012, 27(6): 647-651.  
LI Guojie, CHENG Xueqi. Research status and scientific thinking of big data. Bulletin of the Chinese Academy of Sciences, 2012, 27(6): 647-651. (in Chinese)
- [3] 孟小峰, 慈祥. 大数据管理: 概念、技术与挑战[J]. 计算机研究与发展, 2013, 50(1): 146-169.  
MENG Xiaofeng, CI Xiang. Big data management: Concepts, techniques and challenges. Journal of Computer Research and Development, 2013, 50(1): 146-169. (in Chinese)
- [4] 水利部水利信息中心. 国家水资源监控管理信息平台技术设计报告[R]. 2013, 6.  
Water Conservancy Information Center of the Ministry of Water Resources. Technology Design Report for the National Water Resources Monitoring Management Information Platform. 2013, 6. (in Chinese)