

基于大数据分析的水资源调度优化方法与技术路径

杨秀辉

(桂林市水利电力勘测设计研究院, 广西 桂林 541000)

摘要 本文以大数据背景下水资源调度优化为研究对象, 首先从数据挖掘角度出发, 对基于大数据的水文预测模型进行了研究; 其次建立了实时动态调度方案生成模型; 最后通过对预测模型的效果评估和调度优化结果对比, 说明大数据分析对于水资源调度的作用, 并提出该类问题的技术路径, 以期为进一步提高水资源配置水平和用水效益提供参考。结果显示, 在线监测和预警平台的应用可使水利工程运行更加稳定可靠, 且能及时发现隐患并采取措​​施, 保证水库的安全运行。

关键词 水利工程; 大数据分析; 水资源调度; 水文预测模型; 水资源配置

中图分类号: TP3; TV67

文献标志码: A

DOI: 10.3969/j.issn.2097-3365.2026.08.009

0 引言

随着人口的增长和社会经济的发展, 水资源日益紧缺已成为一个普遍的问题。如何解决水资源不足这一问题成为人们共同关注的重大课题之一。而合理的水资源调度则是提高水资源利用率的关键环节, 也是保证供水可靠性和用水合理性的重要措施。传统的水资源调度方式主要是依靠经验和简单的数学模型来确定水量分配方案或调度参数, 很难适应复杂的水资源系统及其变化情况。本文探索以大数据的理念应用于水资源联合调度管理的新思路和新途径, 力求做到创新性工作的同时也能取得实际意义的应用效果。

1 大数据分析在水资源调度中的价值

1.1 提高数据处理的准确性和效率

水资源调度涉及大量的数据, 包括气象数据、水文数据、用水需求数据等。传统数据处理方法在处理这些海量、复杂的数据时往往效率低下且准确性不高。大数据分析技术能够快速、准确地处理和分析这些数据, 挖掘数据背后的潜在规律和关系, 为水资源调度提供更精确的信息。

1.2 实现实时监测和动态调度

通过大数据分析技术, 可以实时收集和分析水资源系统的各种数据, 及时掌握水资源的动态变化情况。基于实时监测数据, 水资源调度部门可以根据实际情况及时调整调度方案, 实现水资源的动态调度, 提高水资源的利用效率。

1.3 支持多目标决策

水资源调度通常需要考虑多个目标, 如满足用水需求、保障生态流量、提高供水可靠性等。大数据分析技术可以对多个目标进行综合分析和评估, 为决策者提供全面的决策信息, 帮助决策者制定科学合理的调度方案, 实现多目标的优化。

1.4 预测和预警功能

利用大数据分析技术可以对水资源系统的未来发展趋势进行预测, 如预测降雨量、河流水位、用水需求等。通过建立预测模型, 可以提前发现可能出现的水资源短缺、洪涝灾害等问题, 并及时发出预警, 为水资源调度和防灾减灾提供有力支持。

2 水资源调度优化模型构建

针对不同来源数据存在的时间-空间分辨率不一致及质量控制等问题, 要对各种原始数据进行适当的预处理以达到如下目的: 第一, 解决各组分间缺损现象; 第二, 实现多尺度间的有效耦合; 第三, 消除冗余量并保持数据的一致性。

该体系主要从两方面入手。首先, 对与水文学相关的主要变量进行观测和采集, 包括雨强(降水量)、蒸发量、径流量等反映当前水文状态的重要参数, 并且要求其具有较高的准确性; 然后, 收集并整合过去的一些关于流域的相关水文资料信息, 如河道断面面积、水库库容及蓄积量等这些相对稳定的数据。建模过程中使用 LSTM (Long Short-Term Memory) 来搭建

作者简介: 杨秀辉 (1993-), 男, 本科, 工程师, 研究方向: 水文与水资源工程。

模型架构。由于 LSTM 是一种专门用于处理时序性数据的深度学习方 法，在 LSTM 的内部有一个特殊的存储机制称为 memory cell (记忆单元)，能够提取出隐藏层中每个时刻的信息并将之保存下来。设定水文预测模型的输入变量为前 n 个时刻的水文观测值序列 $X(t-n+1), X(t-n+2), \dots, X(t)$ ，预测目标为未来 m 个时刻的水文值 $Y(t+1), Y(t+2), \dots, Y(t+m)$ 。LSTM 模型的隐藏状态更新公式为：

$$f_t = \sigma(W_f \cdot [ht-1, xt] + bf) \quad i_t = \mathcal{O}(W_i \cdot [ht-1, xt] + bi) \quad Ct = f_t * c_{t-1} + i_t * \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (1)$$

式(1)中： f_t 为遗忘门， i_t 为输入门， C_t 为细胞状态， σ 为 sigmoid 激活函数， W 和 b 分别为权重矩阵与偏置向量^[1]。

3 调度优化实验测试与分析

3.1 实验环境与数据准备

针对典型的山区河谷型小流域，选取其 20 年来较为完整的水文气象资料作为本次实验的数据集。其中包括日降水量(雨和雪)、径流深、平均水位、实测洪峰流量等多个重要指标，总计有约 73 000 条数据组成。其中部分数据来源于国家基本水文站，还有部分来自民间水文网站及网络爬虫获取的数据源。整个项目使用到的硬件环境是基于 Hadoop 的 MapReduce 框架构建而成的一个小型集群，所有节点均为阿里云提供的 ECS 实例。每台 ECS 实例均配置 8 个 CPU 内核，且内存为 32GByte，并安装 Ubuntu 18.04 操作系统。每个 ECS 节点中部署了一个 TensorFlow 服务进程用于训练神经网络模型。而为了应对大量数据带来的读写压力问题，则采用了开源的 HDFS (Hadoop Distributed File System)，它能够提供高容错性并且高效的存储海量结构化或非结构化的数据块的能力^[2]。

3.2 预测模型性能评估

为证明本文提出的改进型 LSTM 模型的有效性和可行性，以比较各种预测算法的效果来达到此目的。通过反复推敲后选择了 RMSE、MAE 和 R 三个评价标准对模型进行了测试(其中 RMSE 越低说明拟合越好，R 越接近于 1 则表示相关性越高)，并与基于 LSTM 的 RNN 预测模型相比，分别将其与其他两种预测模型 ARIMA 及传统的回归模型相比较(见表 1)。

结果显示，利用 LSTM 神经网络建立预测模型比传统的预测方法均方根误差(RMSE)下降约 39.8%，平均绝对误差(MAE)减少约 42.9%，并且相关性提高到 0.18，说明基于 LSTM 神经网络建立的预测模型更准确、稳定；通过将注意力机制及多尺度特征融合技术加入 LSTM 模型中进一步提升预测精度后，得到的误差更加集中且

峰值显著减小，表明经过改进的 LSTM 模型有更好的鲁棒性和稳定性；特别是在汛期高水位下的预测应用场景下，本文提出的模型可以有效将预测偏差控制在 5% 以内，为防洪调度决策提供了高质量的数据支持。

表 1 不同预测模型性能对比表

模型类型	RMSE	MAE	R
线性回归	12.34	9.87	0.73
ARIMA	10.52	8.23	0.78
LSTM	8.16	6.12	0.87
改进 LSTM	7.43	5.64	0.91

3.3 调度优化效果分析

选择近三年的调度运行资料进行对比研究，并以“水资源开发利用程度”“水生态环境质量变化情况”的相关评价指标为依据，主要对三个方面的具体效益进行了比较：一是从供水安全保障的角度来看，通过对近年来实际运行中各水库多年平均日保证率与基于模型计算结果进行比较发现，采用本方案所得到的结果较以往计算值提高了约 4%~5%；二是从工程防洪减灾的效果看，通过近几年汛期调度运行实践表明，由于采用了较为准确的预报成果及实时在线控制方式，不仅有效减少了洪水灾害的发生概率，而且对于可能发生的超标准洪水事件也能够采取相应的措施予以应对；三是从水环境治理效果上看，通过对近几年的库区水质监测数据分析可以得知，采用本方案后的调蓄利用过程更加合理化，其调控水平明显高于其他方案或常规模式下的调蓄状态，从而使得入库径流含沙量得到有效降低，入河泥沙总量大幅度减少^[3]。

从图 1 中可以看出，传统的调度方式的平均响应时

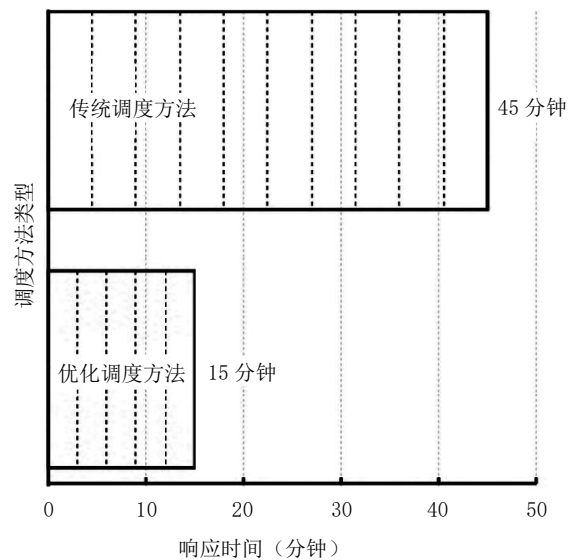


图 1 调度响应时间对比分析

间为 45 分钟左右；而采用大数据分析优化后的方法，则可以达到小于 15 分钟，显著提升了调度的及时性和有效性，在面对突发事件的时候，反应速度越快就越能更好地防范和减少损失。另外，优化后的调度方法可以使生态用水保证率达到 92.3%，比传统方法高出了 7.8%。这有利于维持河流生态系统平衡发展，并且在确保生活生产的同时兼顾了生态用水需求。

4 技术实现路径

4.1 构建水资源调度大数据平台

建立一个具有“全链路”的水务大数据平台。首先，建立数据层，数据层中对所有数据按照类型划分成三类：原生数据、中间处理数据和加工后数据，并且按照不同的组织方式存放在对应的数据集中。原生数据主要为业务系统产生的各类数据文件，在某些情况下也可能来自外部数据源；中间处理数据指由数据源经过清洗转换等操作得到的数据集或者部分数据集；加工后数据则是根据具体需求利用各种算法计算出来的新的数据集或者是直接展示给用户的数据表单。其次，在数据层之上搭建处理与分析层。这个层面包含了两部分内容：一是批处理，二是流式处理。前者主要用于一些长时间运行的任务，比如历史数据回放、模型训练以及未来场景模拟；后者则更多地应用于对当前时刻正在发生事件的检测、预警和近期水文情势变化的趋势预测。再次，面向应用的应用层，它负责向用户提供调度辅助决策、情景仿真、风险评估等功能模块，帮助管理者做出科学合理的判断。最后，服务与治理，主要包括权限管理、数据安全、元数据管理、接口规范化、协同机制等^[4]。

4.2 多源数据采集与整合

首先，集成多类型数据采集设备：包括城市供水管网中的各类型的计量仪表、阀门井盖监测装置、雨量站点、水质检测点等。其次，统一数据接口规范和技术标准：打破传统“烟囱式”的建设模式带来的数据壁垒效应；通过对不同来源数据的标准化转换，消除异构数据库之间的差异性，实现跨部门、跨地域的信息资源共享。最后，在硬件层面上，可以使用物联网技术来连接不同的感知节点从而构成无线传感网，并将这些传感器所获取的各种信号，通过网络传送到相应的服务器上并最终保存至云平台上。

4.3 模型开发与集成

利用大数据技术可以基于水资源系统的特性及调度的目标，建立相应的优化调度模型；同时将描述型分析、预测型分析与评价型分析等不同类型的模型融

合到一个大数据平台上，通过对模型不断的训练和迭代来提升其准确性以及可靠性的方法。例如：以机器学习为理论基础，结合已有的大量历史信息数据（包括气象条件、河道流量、水库蓄水量等），运用机器学习的方法构建出一套针对区域供水量的需求预测模型并将其嵌入调度决策支持系统之中，从而能够及时有效地指导具体的生产实践工作开展。

4.4 可视化与决策支持

利用大数据可视化的原理，把海量且复杂的资源转换成易于理解的信息图及动态图像的方式呈现给管理者，并对相关管理人员进行决策辅助服务。通过构建具有交互功能的可视化仪表盘，使决策人员能够直接获取当前水情概况、调度方案执行情况等相关内容并及时作出判断与调整；同时结合 FineBI 等主流大数据可视化工具实现大屏展现以及多元数据分析，提升决策水平^[5]。

5 结束语

本文针对现有水资源调度存在的问题，在国内外已有研究成果的基础上，以“互联网+”理念为基础，提出了一种基于大数据的水资源调度优化方法和技术路线，并进行了相应的理论推导及数学建模；通过实验测试验证其可行性并取得了良好的效果。试验结果显示，基于大数据的水资源调度优化方法能有效地提升水资源的利用率，确保供水安全性，给实际的水资源管理工作提供了科学的技术支撑。未来的研究工作需要继续扩展大数据在水资源调度领域的应用场景，并进一步探索大数据与物联网相结合的方式，利用物联网技术和设备对水利基础设施实施远程监测和自动控制。

参考文献：

- [1] 吴瑶,夏正豪,胡杨颂,等.基于数字化技术共建“和而不同”动态能力:2011~2020年索菲亚与经销商的纵向案例研究[J].管理世界,2022,38(01):144-163.
- [2] 张振刚,许亚敏,罗秦晔.大数据时代企业动态能力对价值链重构路径的影响:基于格力电器的案例研究[J].管理评论,2021,33(03):339-352.
- [3] 李延东,余博文,赵雪亮.基于数据特征的边缘云资源调度策略分析[J].广东通信技术,2025,45(01):21-24.
- [4] 周娜,李述,陈鹏,等.面向智慧水务的北京市水资源调度业务体系[J].中国水利,2023(14):61-65.
- [5] 谭凌照.水利工程设计中基于GIS技术的水资源调度与管理研究[J].水上安全,2025(02):169-171.