Journal of Water Resources Research 水资源研究, 2025, 14(1), 1-11
Published Online February 2025 in Hans. https://www.hanspub.org/journal/jwrr
https://doi.org/10.12677/jwrr.2025.141001

水库调度模型与优化求解方法研究进展

谢雨祚，郭生练*，钟斯睿，李承龙，孙博凯

武汉大学水资源工程与调度全国重点实验室，湖北 武汉

收稿日期：2025年1月24日；录用日期：2025年2月4日；发布日期：2025年2月27日

摘 要

水库调度是兴利避害、发挥水资源综合利用效益的一项重要的非工程措施。随着大批水库电站的建成投运，我国水利水电事业已经进入了由规划设计到管理运行的关键转型期。本文根据水库与决策者之间的交互作用，把水库调度模型归纳为三类，即无策略调度模型、开环策略调度模型、闭环策略调度模型；综述了基于最优控制理论的水库模拟调度和优化求解方法。合理的建模与优化方法能够适应多种目标导向、对不同时空间尺度的水库调度进行模拟仿真，具有重要的研究和应用价值。

关键词

水库调度，模型分类，最优控制，动态规划，优化求解

Research Progress on Reservoir Operation Models and Optimal Solution Methods

Yuzuo Xie, Shenglian Guo*, Sirui Zhong, Chenglong Li, Bokai Sun

State Key Laboratory of Water Resources Engineering and Management, Wuhan University, Wuhan Hubei

Received: Jan. 24th, 2025; accepted: Feb. 4th, 2025; published: Feb. 27th, 2025

Abstract

Reservoir operation is a crucial non-structural measure for optimizing the benefits of water resource utilization and mitigating adverse impacts. With the completion and commissioning of numerous reservoirs and hydropower stations, a critical transition phase from construction to operational management is going in China. The reservoir operation models are classified into three categories based on the interaction between the reservoir and its managers, *i.e.*, policy-free operation models, open-loop policy operation models, and closed-loop policy operation models. The simulation and optimization techniques based on optimal control

作者简介：谢雨祚(1998-)，男，湖北宜城人，博士研究生，主要从事水文水资源方面研究。Email: yuzuoxie@whu.edu.cn
*通讯作者简介：郭生练(1957-)，男，教授，挪威工程院外籍院士，主要从事水文水资源研究。Email: slguo@whu.edu.cn

文章引用: 谢雨祚, 郭生练, 钟斯睿, 李承龙, 孙博凯. 水库调度模型与优化求解方法研究进展[J]. 水资源研究, 2025, 14(1): 1-11. DOI: 10.12677/jwrr.2025.141001

theory are systematically reviewed for each model type. These modeling and optimization approaches are capable of generalizing operational behaviors under varying objectives and spatiotemporal scales, which has significant research and application value.

Keywords

Reservoir Operation, Model Class, Optimal Control, Dynamic Programming, Optimal Solution

Copyright © 2025 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. 引言

水库调度(也称水库控制运用)是指根据水库开发任务和综合利用需求，按水库来水和蓄水情况，对入库径流进行重新分配或确定水库蓄放水量的过程，是水库工程管理的基本任务。在确保水库大坝和防洪安全的前提下，采用合理适当的蓄放水策略，以保证防洪、电力、供水等多个行业部门的用水需求，充分发挥水库的综合利用效益[1]。为确保上述基本任务能够足量完成，可将决策者(即水库管理者)实际指导水库调度的行为抽象为决策者与决策者之外所有与其相互作用的环境之间的持续交互，以时间离散化的形式构建水库调度模型进行具体探讨研究。其中，决策者指能进行学习、根据策略进行一系列蓄放水操作决策的智能体；若将环境简化为水库，考虑决策者与水库之间的交互作用，则待研究的水库调度的行为可概化为水库调度模型。

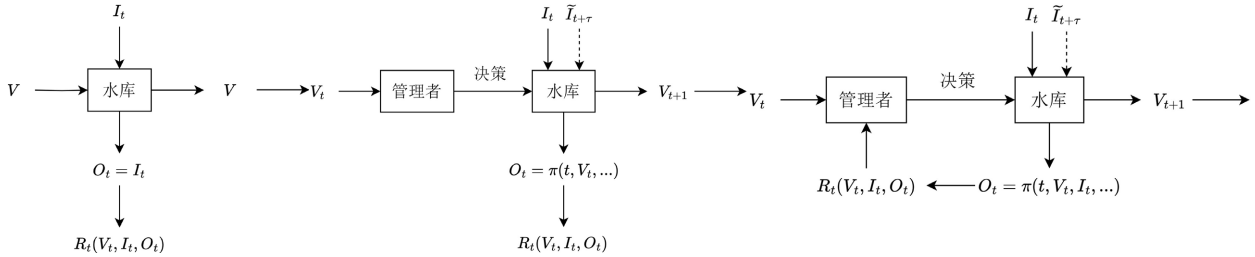
水库优化调度属于最优控制理论[2]的范畴，而最优控制理论是应用数学中最优化的分支，着重研究使控制系统的指标达到最优化的条件和方法[3]。“最优控制”最早出现于 20 世纪 50 年代用于描述设计控制器，使动态系统随时间变化的某种度量最小化或最大化。1954 年，钱学森编著的 *Engineering Cybernetics* 促进了最优控制理论的发展。1955 年的庞特里亚金最大化原理(Pontryagin’s Maximum Principle)、1957 年的贝尔曼方程(Bellman Equation)等控制理论的出现则标志着最优控制理论的形成。最优控制理论在科学、工程等研究中均有广泛应用。在水库优化调度领域，其目的是最大化水库调度所产生的综合利用效益。效益可概化为目标函数，是决策者在采取某种策略下，水库于调度期限 T 内所获得收益的整体性表达，也是决策者需要优化的定量指标，如发电量、供水量等。随着水资源综合利用观念的加强，水库调度已成为综合考虑水资源经济、环境、生态和社会和谐发展的多目标调度问题[4]，可定义多目标函数为[5]：

$$J^b = \Psi_{e\in(1,\cdots,E)}\left(\Phi_{t\in(1,\cdots,T)}\left(R_t^b\left(V_t, I_t, O_t\right)\right)\right) \tag{1}$$

式中：J^b 表示第 b 个目标，$b=1,\cdots,B$；$R_t^b(\bullet)$ 为水库在第 t 阶段中对应于第 b 个目标的收益函数为发电量、供水量等与社会经济相关的指标，$t=1,2,\cdots,T$；$\Phi_{t\in(1,\cdots,T)}(\bullet)$ 是对调度期限 T 进行统筹考虑的运算符，如累加运算、取平均运算等；$\Psi_{e\in(1,\cdots,E)}(\bullet)$ 为过滤由 E 个入库径流集合的干扰所产生的噪声的运算符[6] [7]，如最不利情况、最大效益最大化、期望值等，反映了决策者处理影响水库系统的不确定性时的态度；V_t，I_t，O_t 分别为水库当前时段的库容，来水和出流。

决策者探究使水库达到效益最优时所对应的蓄放水策略，并以此为基础指导水库调度运行。为与其决策过程相匹配，连续时间内的水库调度过程被离散为多阶段决策(又称序贯决策)过程，同时假设 Δt 等于决策者实际对水库进行蓄放水操作的时间间隔。而调度期限 T 通常是有限的，但也可以假设 $T\to\infty$，在该假设下，为了保证水库的效益价值函数能够收敛，可采用贴现因子对未来的回报进行贴现，或者估计平均收益[8]。

就某一离散时段 $T = t$ 时的调度过程，可根据水库管理系统的交互情况、水库的特征参数和水力联系等条件的不同，将水库调度模型分为无策略调度模型(图 1(a))、开环策略调度模型(图 1(b))和闭环策略调度模型(图 1(c))，对应着不同的建模需求、模型结构和优化理论。



(a) 无策略调度模型　　(b) 开环策略调度模型　　(c) 闭环策略调度模型

图 1. 水库调度模型与离散时段 $T = t$ 时的概化图

2. 无策略调度模型

见图 1(a)为无策略调度模型于离散时段 $T = t$ 时的概化图，适用于以下情况：

1) 决策者没有制定水库调度策略、不受水库调度的收益回馈；或决策者不以某些收益为主要目的，如水库汛期的防洪调度必须预留一定的防洪库容。

2) 水库为径流式发电形式，几乎没有调节径流的能力[9]；径流式水库常作为调峰或基荷电厂，其库容一般较小、水力发电几乎不需要水的储存，或仅需要较小的储水设施。

3) 水库具有一定的调节能力，但是调度模型的时间离散时段步长Δt 大于水库调节能力对应的时间尺度；如某水库为日调节水库，当以旬为调度时段步长时，其调节能力与模型离散时段步长不匹配，一个旬的入库水量远大于兴利库容，因此调度过程不够精确。

无策略调度模型中，决策者与水库或几乎不存在交互作用、或交互过程难以被精细化表达、或须预留一定的防洪库容等，水库以当时段来水 I_t 控制泄流 O_t，即 $O_t = I_t$，可认为水库在调度期限 T 内任何离散时段的状态(即库容 V)不变。

3. 开环策略调度模型

图 1(b)为开环策略调度模型于离散时段 $T = t$ 时的概化图，该模型为每个调度时刻指定了先验的操作决策，而不考虑可能发生的条件，适用于决策者已经制定蓄放水策略并要求水库调度过程遵照相关规则执行，但水库的状态与收益并不反馈给决策者的情况，例如常规调度图、风险对冲规则和模型预测控制等。

3.1. 常规调度图

常规调度图是在横轴为时间、纵轴为水库水位的坐标图上，由水库坝址径流实测资料计算和绘制一组调度线，将调度图划分为若干个调度区，以反映水库在年内不同时段与蓄水状态下的调度策略[10]。常规调度图是基于现状需求和站点水文规律制定的[11]，可分为兴利调度图、防洪调度图、防洪兴利联合调度图等类型。而调度线可根据重要性分为(上、下)基本调度线和附加调度线，上、下基本调度线是在包含年内分配的各种设计枯水年来水情况下，水库电站按保证出力工作时，水库在年内各个时段的最高和最低蓄水指示线，介于上、下基本调度线的区域为保证出力区，以保证水库电站的正常发电运作[1]。尽管如此，调度图基于典型年来水情况制定，对水库所处的水文气候等边界条件具有一致性的强假设[12] [13]。

从数学映射的角度来看，常规调度图可将出库流量 O_t 表示为水库当前时段 t 和库容 V_t 的函数，即 $O_t = \pi_{\mathrm{RC}}(t, V_t)$，式中 $\pi_{\mathrm{RC}}(\bullet)$ 为决策者进行蓄放水决策的动作函数，此处为常规调度图的函数表达。

3.2. 风险对冲规则

风险对冲规则的“风险”代指水库调度可能存在的损失，来源于入库流量与损失拟合函数的不确定性；而“对冲”作为金融学术语，本意为同时进行两笔行情相关、盈亏相反的交易时，特意降低另一项投资的风险投资。在水库调度范畴内，风险对冲规则指水库在放水的同时，通过一部分水量的存蓄以冲销未来可能存在的缺水风险、以平衡当前与未来效益[14]。简化运行策略[1]是风险对冲规则的最基本形式，其在水库兴利调节计算中，以用水部门规划的额定供水量为约束、以水库面临时段可用水量为基础，根据决策者既定的、以可用水量和出库流量为坐标轴的调度指示曲线，推求当前时刻水库出库流量[15]。

从数学映射的角度来看，风险对冲规则可将出库流量 O_t 表示为水库当前库容 V_t 和入库流量 I_t 的函数，即 $O_t = \pi_{\mathrm{HR}}(V_t, I_t)$，式中 $\pi_{\mathrm{HR}}(\bullet)$ 为风险对冲规则的函数表达。

3.3. 模型预测控制

模型预测控制(Model Predictive Control, MPC)依时间顺序滚动、在线解决在一个有限的时间范围内(有效预见期)定义的多个开环控制问题[16]，经常用于运行水位动态控制[17]的研究中，也称预报预泄调度。基于模型对系统未来动态行为的预测，MPC 具有显式处理约束的能力，通过把约束加到未来的输入、输出或状态变量上，可以把约束显式表示在一个在线求解的数学规划问题中[18]。例如运行水位动态控制可以在保证防洪安全的前提下打破不超过汛限水位的硬性约束。MPC 分为模型、预测和控制三个步骤，t 时刻在预见期 τ 内，模型预报[t, $t+\tau$]内的入库流量，然后根据预报值进行当前出库流量 O_t 的决策，接着进入 $t+1$ 时刻并重复上述过程，是一个具有实时性、根据随时间产生的入库流量数据动态更新的模型。尽管 MPC 的滚动优化引入了闭环行为，但这是通过不断重复开环优化实现的，理论上每次在当前时刻生成的解，仍然是独立的开环解。

从数学映射的角度来看，MPC 可将出库流量 O_t 表示为水库当前库容 V_t、入库流量 I_t、预见期 τ 和预见期内入库流量预报值 $\tilde{I}_{t+\tau}$ 的函数，即 $O_t = \pi_{\mathrm{MPC}}\left(V_t, I_t, \tau, \tilde{I}_{t+\tau}\right)$，其中 $\pi_{\mathrm{MPC}}(\bullet)$ 为预见期内调度规则的函数表达，也可以是求解数学规划问题的方法函数[19]。

开环策略调度模型中，水库管理者或是基于历史水文气候规律、典型年水情特征制定离线策略，如常规调度图和风险对冲规则；或是采用 MPC，在有限的有效预见期内，根据实时入库流量的预报情况及其他有利于目标的实时外生信息[20]，对出库流量进行在线的滚动控制。然而，当水库受到水文随机性、气候变化和人类活动等综合影响时[21] [22]，以常规调度图和风险对冲规则作为策略指导会逐渐偏离其设计时所假设的典型条件，如典型枯水年的来水频率或量级已经发生了变化，采用初设调度规则会失之偏颇。MPC 虽然可以借助一定的外生信息，但是由于较高预报精度的要求，其优化是局部的，仅限于有效预见期内；同时又是贪婪的，仅能在当前状态下达到最优，难以在整个调度期限上实现最优化且优化目标单一[23]，仅适用于实时或短期水库调度。因此，作为缺乏信息反馈的调度方式，决策者并不根据反馈的收益来改善调度策略，开环策略调度模型仅在水库所处水文气象等环境条件具有强假设、水文气象预报精度较高的情况下才具有较好的效益，更适合作为决策参考而不是制定为指导策略[24]。

4. 闭环策略调度模型

与开环策略调度模型的先验决策不同，闭环策略调度模型(见图 1(c))根据系统的状态和观测信息(如入库流量)确定每个调度时刻的决策，这些决策依赖于前一个时刻的状态，因此可引入递归循环，将决策者的实际决策与调度过程建模为(有限)马尔科夫决策过程(Markov Decision Process, MDP)，决策者为水库在每个时间段依序制定相应决策且受水库的收益回馈，能够根据水库在整个调度期限 T 上的收益优化策略。得名于俄国数学家 Andrey Andreyevich Markov，MDP 早在 20 世纪 50 年代就已经为人所知，1960 年由 Howard 所著的 *Dynamic Programming and Markov Processes* 详细记载了关于 MDP 的核心研究。MDP 的状态转移满足马尔可夫性，又称无后效性，即

当前状态一旦确定，$t+1$时段及以后的状态转移过程与$t-1$时段及以前的状态与所采取的决策无关[25]。对水库调度而言V_t即为状态变量，如式(2)所示，水库的水量平衡方程即为 MDP 的状态转移方程：

$$V_{t+1}=V_t+\left(I_t-O_t\right)\cdot\Delta t \tag{2}$$

从式(2)所示状态转移方程可知，MDP 对环境(即水库)建立了模型，在给定一个状态和决策的情况下，该模型可以预测水库的下一个状态和收益，从而允许决策者作规划，即在真正经历之前考虑未来可能的各种情境预先决定采取何种决策，属于有模型的策略学习方法[25]。又因为 MDP 属于多阶段决策过程，因此常采用动态规划算法等最优控制策略[26]。虽然线性规划、非线性规划、混合整数规划[27]等与时间无关的静态规划亦可以人为地引入时间因素、将其视作一系列前后有关联的单阶段决策问题、并逐个解决，从而得出整个调度期内的最优策略[28]，然而多阶段静态规划在进行决策时难以考虑未来的状态与收益，亦属于贪婪的方法，且不易扩展至具有大量离散时段数量的中长期调度问题。而网络流方法[29]虽然可以考虑多阶段问题，但是该方法基于图论，仅能解决阶段较少、拓扑结构图方便绘制的简单规划问题，难以应用至梯级水库多阶段优化。此外，近年来逐渐涌现出的次优策略控制方法同样具有足够性能，如近似动态规划，在人工智能领域又称强化学习。

4.1. 动态规划

动态规划(Dynamic Programming, DP)包括确定动态规划(Deterministic Dynamic Programming, DDP)和随机动态规划(Stochastic Dynamic Programming, SDP)两类[30]。

4.1.1. 确定动态规划

确定动态规划在每个时刻基于 MDP 依次决策，这些决策能够产生一定的收益并影响水库系统的下一个状态，从而影响后续的所有收益。DDP 将 MDP 表述为一个单目标问题，其状态转移方程见式(2)，可通过贝尔曼方程计算 t 时刻系统每种可能状态下的蓄放水策略的长期收益：

$$H_t\left(V_t,I_t\right)=\underset{u_t}{\text{opt}}\ \Psi_e\left(\Phi_t\left(R_t\left(V_t,I_t,O_t\right),H_{t+1}\left(V_{t+1},I_{t+1}\right)\right)\right)\quad t=T,T-1,\cdots,1 \tag{3}$$

其边界条件为：

$$H_{T+1}\left(V_{T+1},I_{T+1}\right)=0 \tag{4}$$

式中：$R_t\left(\bullet\right)$为水库在第 t 阶段的收益函数，在 DP 的相关研究中又称指标函数、效用函数或瞬时代价函数等；$H_t\left(\bullet\right)$为水库在第 t 阶段的状态价值函数，又称最优值函数、间接效用函数或未来代价函数。目标函数与状态价值函数有一定区别，在 DP 中，式(1)所示的水库调度的目标与状态价值函数的关系可以表示为$J=H_1\left(V_1\right)$，需要注意的是，式(3)仅限于$\Psi_e\left(\bullet\right)$和$\Phi_t\left(\bullet\right)$两函数的特定、单一组合，例如，当$\Phi_t\left(\bullet\right)$为取平均的函数时，入库流量干扰的不确定性须通过求期望的方式进行过滤，即$\Psi_e\left(\bullet\right)=\mathbb{E}_e\left(\bullet\right)$。

根据动态规划最优性定理可知，该问题存在最优策略[28]；而在最优策略已知的前提下，由贝尔曼最优化原理[26]可知，最优策略的子策略也是子过程的最优策略，因此可以对式(3)进行逆序求解，以得到最优策略

$$\pi^*\left(t,V_t,I_t\right)=\arg\underset{\pi}{\text{opt}}\ H_t^*\left(V_t,I_t\right) \tag{5}$$

式中：$H_t^*\left(V_t,I_t\right)$为水库在第 t 阶段的最优状态价值函数；arg 为$H_t^*\left(V_t,I_t\right)$最优时变量或参数的取值。

4.1.2. 随机动态规划

随机动态规划(SDP)的基本方程与式(3)大致相符，所不同的是以概率分布的形式显式地考虑了径流的不确定性，而不是直接以确定性的径流系列为输入，因此又被称为显随机优化(Explicit stochastic optimization, ESO)[31]方法。相比 DDP，SDP 能够用于随机条件下的水库调度决策优化[32]，贝尔曼方程可表示为：

$$H_t\left(V_t,I_t\right)=\underset{u_t}{\text{opt}}\ \Psi_e\left(\Phi_t\left(R_t\left(V_t,I_t,O_t\right),\sum_{I_t}p_{t+1}\left(\tilde{I}_{t+1}\mid I_t\right)H_{t+1}\left(V_{t+1},\tilde{I}_{t+1}\right)\right)\right)\quad t=T,T-1,\cdots,1 \tag{6}$$

式中：$p_{t+1}\left(\tilde{I}_{t+1}\mid I_t\right)$为$t+1$时刻不确定未来径流$\tilde{I}_{t+1}$ t 时刻确定性径流I_t的条件概率密度函数，反映了预报的不确定性。

然而，采用式(3)或式(6)计算状态价值函数，限制了 DP 在复杂高维条件下的应用。Giuliani 等[33]归纳 DP 在梯级水库多目标优化调度中存在的三种“灾难”：①“维数灾难”[26]：状态价值函数的计算时间和存储需求随着梯级水库数量的增加呈指数级增长[34]。当梯级水库数量大于 3 时，DP 可能面临着计算成本较高等问题[35]，若梯级水库上下游水力联系中存在顶托影响，则水库调度不满足无后效性[36]，无法采用贝尔曼最优原理将问题分解为递归形式、无法实现动态规划。②“建模灾难”[37]：在不增加“维数灾难”的前提下，对可能会改善水库调度性能的外生信息进行建模。为了进行每个阶段的顺序决策过程，必须对 DP 框架中包含的任何信息(如状态、径流、收益等)进行显式建模，以充分预测下一个状态转移并估计价值函数。因此，能观测到但不受水库管理者决策影响的诸如流量、降水等外生信息可能会改善水库调度的性能[24]。然而，除非为每个外生信息构建动态模型，否则无法在 DP 中考虑这些外生信息，此举亦会加剧“维数灾难”。③“多目标灾难”[38]：DP 仅能解决单目标优化问题，虽然可以通过算法改进以获得一系列单目标近优解集[39]，但是难以直接求解多目标问题并提供一系列可构成替代方案的 Pareto 最优解集[40]。此外，由于式(3)仅限于$\Psi_e(\bullet)$和$\Phi_t(\bullet)$两函数的特定、单一组合，因此难以探索反映不同目标的不同风险态度的竞争问题[7]。

鉴于上述“灾难”导致的应用局限性，学者们从多个角度对 DP 系列算法进行了一系列改进：① 针对 DDP 提出了逐步优化算法[41]、离散微分动态规划法[42]，逐次逼近动态规划法[43]；针对 SDP 也提出了逐次逼近随机动态规划法[44]等改进算法，此外基于大系统控制理论中的分解协调技术能够进行目标或策略分解以达到在宏观上分层降维的目的[10] [45]，使用多种计算资源的并行计算[46] [47]能够提高计算效率。除了并行计算外，上述改进算法和方法虽然降低了计算复杂度，缓解了“维数灾难”问题、能够适应有后效性模型，但却在一定程度上牺牲了解的最优性使其变为局部最优解[48]。② 抽样随机动态规划和替代随机动态规划以引入情景集合的方式描述了径流的转移概率，在一定程度上减轻了“建模灾难”[49]。③ DP 求解多目标问题有约束扰动法和标量函数法两种方法，前者是先选定一个主要目标，将其余目标处理成约束条件，然后针对各个约束条件的可行域优化[50]；后者虽然可以通过某标量函数(如凸组合函数、非线性 Chebyshev 标量函数等)将目标空间的维度降成单目标[51]。但是当目标数量大于等于 3 时，这些方法对计算成本的要求非常高，且考虑到标量化单目标与向量化多目标的非线性关系，Pareto 前沿的逼近精度可能会降低[52]。综上所述，虽然 DP 的改进方法从不同角度提高了水库调度的效率，但是仍未能从根本上攻克这三种“灾难”。

4.2. 近似动态规划

近似动态规划又称强化学习(Reinforcement Learning, RL)，是最优控制领域兴起的一种近似最优方法，“近似”包括值空间近似(Approximation in Value Space, AVS) [53]和策略空间近似(Approximation in Policy Space, APS) [54]两类。

4.2.1. 值空间近似

对于一个简单的有限 MDP，其状态和决策空间小到可以用数组或表格形式表示状态价值函数，称为表格型求解方法。当状态空间足够大时，无论是表格内存还是精确地填写表格所需时间和数据均是求解 MDP 所面临的挑战。AVS 的一种做法是采用连续函数(如多项式函数、样条函数、人工神经网络等)逼近的方式对水库调度的相似状态进行泛化以近似表达状态价值函数，属于有监督学习的范畴。随机对偶动态规划[55]就是采用连续函数逼近的 AVS 方法之一，对应的贝尔曼方程可表示为

$$\tilde{H}_t\left(V_t,I_t\right)=\underset{u_t}{\text{opt}}\ \Psi_e\left(\Phi_t\left(R_t\left(V_t,I_t,O_t\right),\tilde{H}_{t+1}\left(V_{t+1},I_{t+1}\right)\right)\right)\quad t=T,T-1,\cdots,1 \tag{7}$$

式中：$\tilde{H}_{t+1}\left(V_{t+1},I_{t+1}\right)$ 为基于拟合的连续函数和状态－价值对 $\left[V_{t+1}^i,I_{t+1}^i,\tilde{H}_{t+1}\left(V_{t+1}^i,I_{t+1}^i\right)\right]$ 进行插值得到的近似状态价值函数值，$i=1,2,\cdots,\tilde{N}_V$；$\tilde{N}_V$ 和 N_V 分别为较少数量和原问题中的状态变量离散域，满足 $\tilde{N}_V<N_V$。

除了采用连续函数直接逼近状态价值函数外，RL 还提供了一种不依赖水库调度模型、而是直接对水库调度过程进行仿真并观察得到状态以计算状态价值函数的 Q-learning 算法[56]，该算法通过在学习阶段的增量更新来优化贝尔曼方程的动作价值函数，其优化并没有探索整个决策空间，而是在通过一定的实验设计测试生成的备选决策序列中，对状态的访问选择采取某一项具体决策。此外，还有将上述直接逼近状态价值函数及从仿真经验中学习的动作价值函数逼近的思想相结合的拟合 Q 迭代(Fitted Q-iteration, FQI)算法[34] [57]，FQI 将贝尔曼方程的递归解替换为样本数据集上的非线性回归序列。相比较而言，Q-learning 只有在增量更新动作价值函数时才收敛，而 FQI 则能够对完整的数据集进行批处理。综上所述，虽然 AVS 能够减轻 MDP 的“维数灾难”问题，但是在水库调度计算模型已经较为完备且成熟的情况下，对动作价值函数的近似可能会引起调度过程的失真从而导致较大的误差。

4.2.2. 策略空间近似

前述 EDP (AVS)先学习调度过程的状态价值(动作价值)函数，然后基于精确(近似)的价值函数选择决策，如果没有价值函数的计算，策略也就不会存在。而策略空间近似(APS)不再依赖于价值函数进行决策的选择，而是直接对参数化的策略进行学习。APS 可分为直接策略搜索(Direct Policy Search, DPS)和隐随机优化(Implicit Stochastic Optimization, ISO)两种类型。

1) 直接策略搜索

参数化－模拟－优化，又称策略梯度方法，是 APS 的主要途径，DPS 通过将调度规则参数化为某一类函数，并在函数的参数空间内采用启发式算法[58]进行搜索，以优化目标函数 $\mathbf{J}(\boldsymbol{\theta})$：

$$\underset{\pi}{\text{opt}}\ \mathbf{J}(\boldsymbol{\theta})=\left[J^1(\boldsymbol{\theta}),J^2(\boldsymbol{\theta}),\cdots,J^B(\boldsymbol{\theta})\right]=\left[H_{\pi_0}^1\left(V_0,I_0\right),H_{\pi_0}^2\left(V_0,I_0\right),\cdots,H_{\pi_0}^B\left(V_0,I_0\right)\right] \tag{8}$$

式中：$\boldsymbol{\theta}$ 为策略的参数集。

DPS 相对 AVS 具有重要的理论优势，对于连续的策略参数化，前者选择的决策作为待优化参数的函数会平滑变化，而后者的变化则不一定平滑，即使近似动作价值函数只发生了任意小的变化，决策的选择也可能会发生较大变化[59]，因此，DPS 相比 AVS 具有更好的收敛性和精确度。此外由于调度策略的参数化，选择合适的函数类型对于发现高性能的优化解集至关重要[60]，不少对于开环策略调度模型的闭环优化改进方案均属于 DPS 的范畴，如基于传统调度图改进的优化调度图[10] [61]，基于风险对冲规则改进的优化风险对冲规则等，上述直观定义参数搜索空间的函数类型统称为规则曲线[62] [63]，此外还包括其它线性或非线性的函数类型[33] [64]。

另一方面，DPS 从根本上克服了对 EDP 和 AVS 影响深远的三个“灾难”：① 对“维数灾难”，由于避免了状态价值函数的计算，DPS 能够在中长期调度中允许处理三个及以上数量的梯级水库群[47]。维度增加难免造成计算成本的上升，亦可以采用高性能启发式算法通过并行计算缓解[65]；② DPS 能够直接在调度规则中纳入外生信息(如水文预报等)，而不需要将外生信息耦合进 MDP 的闭环决策中，可以有效解决“模型灾难”[66] [67]。此外，区别于 MPC 需要较为精确的预报来确保水库调度的风险可控，DPS 具备能够从有偏差的外生信息中挖掘到提升调度规则性能的能力[68] [69]；③ 将 DPS 与多目标进化算法(Multi-Objective Evolutionary Algorithms, MOEAs)相结合，可以生成 Pareto 前沿，从而扩大目标的数量和复杂性[70] [71]，攻克了“多目标灾难”的难题，此外还能够通过竞争框架探索多种竞争问题的表述[7]。然而，DPS 无法保证解的最优性，且难以判别迭代收敛的程度。

2) 隐随机优化

隐随机优化[72]又称调度规则提取，是另一种 APS 的实现途径，分两步进行：首先解决一个长系列确定性问题，并找出能够代表该流域水文条件下的最优蓄放水策略；随后将最优策略集合作为某函数回归问题的目标，推求近优解的参数化策略。为了使近优解尽可能接近最优解，确定性水库调度问题常常采用 DP 及其改进算法求解[73] [74]，而调度规则的提取中，回归函数从最早的多元线性回归发展至多元非线性函数回归以及机器学习相关函数，如径向基函数、神经网络、决策树等[74] [75]。ISO 的第二步主要解决了“建模灾难”，而第一步若采用 DP 系列方法求解则同样面临着“维数灾难”和“多目标灾难”的问题。

闭环策略调度模型以 MDP 为基础，管理者对水库采取一定的调度策略，并接收水库的收益回馈，水库调度与收益计算的完备流程使得管理者能够事先对水库调度策略进行规划，因此常采用 DP 或 ADP 方法求解。DP 及其改进方法在固定的离散时间尺度下理论上可以获得最优解，但是依然会面临“维数灾难”、“建模灾难”和“多目标灾难”的制约；而 ADP 舍弃了解的最优性，以缓解或解决各种“灾难”问题，其中 DPS 从根本上攻克了三个“灾难”，成为目前最流行的梯级水库多目标优化调度研究方法[5]。

5. 结语

水库调度作为水资源管理一项重要的非工程措施，能够对重新分配入库径流、进行水库蓄放水操作，以达到兴利除害的目的，对于发展水利新质生产力、实现可持续发展具有重要作用。经过多年的理论发展与技术完善，水库调度已经形成了较为完整的模拟与优化理论，本文根据水库与水库管理者的交互情况，归纳了三种最常见的调度模型：无策略调度模型、开环策略调度模型和闭环策略调度模型，并根据最优控制理论，针对不同模型综述了模拟或优化方法，能够涵盖绝大部分工程实践与应用条件下的建模与优化需求，对水库调度研究的实际应用与发展均有借鉴意义。

基金项目

国家自然科学基金青年学生基础研究项目(524B2128)和国家自然科学基金长江联合基金(U2340205)资助。

参考文献

[1] 万俊. 水资源开发利用[M]. 第二版. 武汉: 武汉大学出版社, 2008.

[2] KIRK, D. E. Optimal control theory: An introduction. North Chelmsford: Courier Corporation, 2004.

[3] 高桂革. 最优控制理论的发展与展望[J]. 上海电机学院学报, 2005(3): 33-35+39.

[4] 郭生练, 陈炯宏, 刘攀, 等. 水库群联合优化调度研究进展与展望[J]. 水科学进展, 2010, 21(4): 496-503.

[5] GIULIANI, M., LAMONTAGNE, J. R., REED, P. M., et al. A state-of-the-art review of optimal reservoir control for managing conflicting demands in a changing world. Water Resources Research, 2021, 57(12): e2021WR029927.

[6] GIULIANI, M., CASTELLETTI, A. Is robustness really robust? How different definitions of robustness impact decision-making under climate change. Climatic Change, 2016, 135(3): 409-424.

[7] QUINN, J. D., REED, P. M., GIULIANI, M., et al. Rival framings: A framework for discovering how problem formulation uncertainties shape risk management trade-offs in water resources systems. Water Resources Research, 2017, 53(8): 7208-7233.

[8] SONCINI-SESSA, R., WEBER, E. and CASTELLETTI, A. Integrated and participatory water resources management—Theory. Amsterdam: Elsevier Science, 2007.

[9] 张英贵. 径流式水电站调蓄库容的利用[J]. 水力发电学报, 1992(1): 47-50.

[10] 刘攀, 郭生练, 郭富强, 等. 清江梯级水库群联合优化调度图研究[J]. 华中科技大学学报(自然科学版), 2008, 36(7): 63-66.

[11] TU, M. Y., HSU, N. S. and YEH, W. W. G. Optimization of reservoir management and operation with hedging rules. Journal of Water Resources Planning and Management, 2003, 129(2): 86-97.

[12] BAHRAMI, M., BOZORG-HADDAD, O. and CHU, X. Application of cat swarm optimization algorithm for optimal reservoir operation. Journal of Irrigation and Drainage Engineering, 2018, 144(1): 04017057.

[13] HADDAD, O. B., HOSSEINI-MOGHARI, S. M. and LOÁICIGA, H. A. Biogeography-based optimization algorithm for optimal operation of reservoir systems. Journal of Water Resources Planning and Management, 2016, 142(1): 04015034.

[14] 万文华. 变化环境下的干旱演变与水库适应性调度策略[D]: [博士学位论文]. 北京: 清华大学, 2018.

[15] 冯尚友. 水资源持续利用与管理导论[M]. 北京: 科学出版社, 2000.

[16] SCATTOLINI, R. Architectures for distributed and hierarchical model predictive control—A review. Journal of Process Control, 2009, 19(5): 723-731.

[17] 郭生练, 刘攀, 王俊, 等. 再论水库汛期水位动态控制的必要性和可行性[J]. 水利学报, 2023, 54(1): 1-12.

[18] 席裕庚, 李德伟, 林姝. 模型预测控制——现状与挑战[J]. 自动化学报, 2013, 39(3): 222-236.

[19] MACIEJOWSKI, J. M. Predictive control with constraints. New York: Pearson College Div., 2000.

[20] GALELLI, S., CASTELLETTI, A. and GOEDBLOED, A. High-performance integrated control of water quality and quantity in urban water reservoirs. Water Resources Research, 2015, 51(11): 9053-9072.

[21] 尹家波, 郭生练, 顾磊, 等. 中国极端降水对气候变化的热力学响应机理及洪水效应[J]. 科学通报, 2021, 66(33): 4315-4325.

[22] 尹家波, 郭生练, 杨妍, 等. 基于陆地水储量异常预估中国干旱及其社会经济暴露度[J]. 中国科学: 地球科学, 2022, 52(10): 2061-2076.

[23] GIULIANI, M., CASTELLETTI, A. Assessing the value of cooperation and information exchange in large water resources systems by agent-based optimization. Water Resources Research, 2013, 49(7): 3912-3926.

[24] FABER, B. A., STEDINGER, J. R. Reservoir optimization using sampling SDP with ensemble streamflow prediction (ESP) forecasts. Journal of Hydrology, 2001, 249(1): 113-133.

[25] SUTTON, R. S., BARTO, A. G. Reinforcement learning: An introduction. Second Edition. Cambridge: Bradford Books, 1998.

[26] BELLMAN, R. Dynamic programming. Princeton University Press, 1957.

[27] 吴宏宇, 管晓宏, 翟桥柱, 等. 水火电联合短期调度的混合整数规划方法[J]. 中国电机工程学报, 2009, 29(28): 82-88.

[28] 《运筹学》教材编写组. 运筹学[M]. 第 4 版. 北京: 清华大学出版社, 2012.

[29] 罗强, 宋朝红, 雷声隆. 水库群系统非线性网络流规划法[J]. 武汉大学学报(工学版), 2001(3): 22-26.

[30] BERTSEKAS, D. P. Dynamic programming: deterministic and stochastic models. Englewood Cliffs: Prentice Hall, 1987.

[31] CELESTE, A. B., BILLIB, M. Evaluation of stochastic reservoir operation optimization models. Advances in Water Resources, 2009, 32(9): 1429-1443.

[32] VICUNA, S., DRACUP, J. A., LUND, J. R., et al. Basin-scale water system operations with uncertain future climate conditions: Methodology and case studies. Water Resources Research, 2010, 46(4): W04505.

[33] GIULIANI, M., CASTELLETTI, A., PIANOSI, F., et al. Curses, tradeoffs, and scalable management: Advancing evolutionary multiobjective direct policy search to improve water reservoir operations. Journal of Water Resources Planning and Management, 2016, 142(2): 04015050.

[34] CASTELLETTI, A., GALELLI, S., RESTELLI, M., et al. Tree-based reinforcement learning for optimal water reservoir operation. Water Resources Research, 2010, 46(9): W09507.

[35] LOUCKS, D. P., VAN BEEK, E., STEDINGER, J. R., et al. Water resources systems planning and management: An introduction to methods, models and applications. Paris: UNESCO, 2005.

[36] 梅亚东. 梯级水库防洪优化调度的动态规划模型及解法[J]. 武汉水利电力大学学报, 1999(5): 10-12+91.

[37] TSITSIKLIS, J. N., VAN ROY, B. Feature-based methods for large scale dynamic programming. Boston: Springer US, 1996: 59-94.

[38] PARDALOS, P. M. Approximate dynamic programming: Solving the curses of dimensionality. Optimization Methods and Software, 2009, 24(1): 155-155.

[39] 杨光, 郭生练, 李立平. 考虑生态流量的梯级水库柔性决策研究[J]. 华中科技大学学报(自然科学版), 2015, 43(9): 114-116+122.

[40] COHON, J. L., MARKS, D. H. A review and evaluation of multiobjective programing techniques. Water Resources Research, 1975, 11(2): 208-220.

[41] ZHA, G., ZHOU, J., YANG, X., et al. Modeling and solving of joint flood control operation of large-scale reservoirs: A case study in the middle and upper Yangtze River in China. Water, 2021, 13(1): 41.

[42] 支悦, 艾学山, 董祚, 等. 水库发电优化调度模型的快速求解算法及应用[J]. 水力发电学报, 2020, 39(6): 49-61.

[43] 艾学山, 郭佳俊, 穆振宇, 等. 梯级水库群多目标优化调度模型及 CPF-DPSA 算法研究[J]. 水利学报, 2023, 54(1): 68-78.

[44] 王金文, 王仁权, 张勇传, 等. 逐次逼近随机动态规划及库群优化调度[J]. 人民长江, 2002(11): 45-47+54.

[45] 郭生练, 何绍坤, 陈柯兵, 等. 长江上游巨型水库群联合蓄水调度研究[J]. 人民长江, 2020, 51(1): 6-10+35.

[46] LU, Q., ZHONG, P., XU, B., et al. Stochastic programming for floodwater utilization of a complex multi-reservoir system considering risk constraints. Journal of Hydrology, 2021, 599: 126388.

[47] HE, S., GUO, S., YIN, J., et al. A novel impoundment framework for a mega reservoir system in the upper Yangtze River basin. Applied Energy, 2022, 305: 117792.

[48] 冯仲恺, 牛文静, 程春田, 等. 大规模水电系统优化调度降维方法理论分析研究[J]. 水利学报, 2017, 48(2): 146-156.

[49] KIM, Y. O., EUM, H. I., LEE, E. G., et al. Optimizing operational policies of a Korean multireservoir system using sampling stochastic dynamic programming with ensemble streamflow prediction. Journal of Water Resources Planning and Management, 2007, 133(1): 4-14.

[50] TAUXE, G. W., INMAN, R. R. and MADES, D. M. Multiobjective dynamic programing with application to a reservoir. Water Resources Research, 1979, 15(6): 1403-1408.

[51] 刘攀, 郭生练, 张文选, 等. 梯级水库群联合优化调度函数研究[J]. 水科学进展, 2007, 18(6): 816-822.

[52] FLEMING, P. J., PURSHOUSE, R. C. and LYGOE, R. J. Many-objective optimization: An engineering design perspective. Berlin: Springer, 2005: 14-32.

[53] BERTSEKAS, D. P. Dynamic programming and suboptimal control: A survey from ADP to MPC. European Journal of Control, 2005, 11(4): 310-334.

[54] DEISENROTH, M. P., NEUMANN, G. and PETERS, J. A survey on policy search for robotics. Foundations and Trends in Robotics, 2013, 2(1-2): 1-142.

[55] PEREIRA, M. V. F., PINTO, L. M. V. G. Multi-stage stochastic optimization applied to energy planning. Mathematical Programming, 1991, 52(1): 359-375.

[56] WATKINS, C. J. C. H., DAYAN, P. Technical note: Q-learning. Machine Learning, 1992, 8(3): 279-292.

[57] ERNST, D., GEURTS, P. and WEHENKEL, L. Tree-based batch mode reinforcement learning. Journal of Machine Learning Research, 2005, 6(18): 503-556.

[58] REED, P. M., HADKA, D., HERMAN, J. D., et al. Evolutionary multiobjective optimization in water resources: The past, present, and future. Advances in Water Resources, 2013, 51: 438-456.

[59] SUTTON, R. S., MCALLESTER, D., SINGH, S., et al. Policy gradient methods for reinforcement learning with function approximation. In Advances in Neural Information Processing Systems: Vol. 12. Cambridge: MIT Press, 1999: 1057-1063.

[60] LUND, J. R., GUZMAN, J. Derived operating rules for reservoirs in series or in parallel. Journal of Water Resources Planning and Management, 1999, 125(3): 143-153.

[61] HE, S., GUO, S., YANG, G., et al. Optimizing operation rules of cascade reservoirs for adapting climate change. Water Resources Management, 2020, 34(1): 101-120.

[62] WAN, W., ZHAO, J. and WANG, J. Revisiting water supply rule curves with hedging theory for climate change adaptation. Sustainability, 2019, 11(7): 1827.

[63] LI, X., LIU, P., WANG, Y., et al. Derivation of operating rule curves for cascade hydropower reservoirs considering the spot market: A case study of the China's Qing River cascade-reservoir system. Renewable Energy, 2022, 182: 1028-1038.

[64] YANG, G., GUO, S., LI, L., et al. Multi-objective operating rules for Danjiangkou reservoir under climate change. Water Resources Management, 2016, 30(3): 1183-1202.

[65] HADKA, D., REED, P. Large-scale parallelization of the Borg multiobjective evolutionary algorithm to enhance the management of complex environmental systems. Environmental Modelling & Software, 2015, 69: 353-369.

[66] YANG, G., ZAITCHIK, B., BADR, H., et al. A Bayesian adaptive reservoir operation framework incorporating streamflow nonstationarity. Journal of Hydrology, 2021, 594: 125959.

[67] HE, S., GUO, S., ZHANG, J., et al. Multi-objective operation of cascade reservoirs based on short-term ensemble streamflow prediction. Journal of Hydrology, 2022, 610: 127936.

[68] YANG, G., GUO, S., LIU, P., et al. Integration and evaluation of forecast-informed multiobjective reservoir operations. Journal of Water Resources Planning and Management, 2020, 146(6): 04020038.

[69] GIULIANI, M., CROCHEMORE, L., PECHLIVANIDIS, I., et al. From skill to value: Isolating the influence of end user behavior on seasonal forecast assessment. Hydrology and Earth System Sciences, 2020, 24(12): 5891-5902.

[70] GIULIANI, M., HERMAN, J. D., CASTELLETTI, A., et al. Many-objective reservoir policy identification and refinement to reduce policy inertia and myopia in water management. Water Resources Research, 2014, 50(4): 3355-3377.

[71] GIULIANI, M., MASON, E., CASTELLETTI, A., et al. Universal approximators for direct policy search in multi-purpose water

reservoir management: A comparative analysis. IFAC Proceedings Volumes, 2014, 47(3): 6234-6239.

[72] YOUNG, G. K. Finding reservoir operating rules. Journal of the Hydraulics Division, 1967, 93(6): 297-322.

[73] YANG, G., GUO, S., LIU, P., et al. Multiobjective reservoir operating rules based on cascade reservoir input variable selection method. Water Resources Research, 2017, 53(4): 3446-3463.

[74] 钟斯睿, 何彦锋, 郭生练, 等. 雅砻江梯级水库优化调度规则提取研究[J]. 水力发电学报, 2023, 42(10): 50-59.

[75] 郭旭宁, 秦韬, 雷晓辉, 等. 水库群联合调度规则提取方法研究进展[J]. 水力发电学报, 2016, 35(1): 19-27.