

基于网格搜索和交叉验证的支持向量机在梯级水电系统 隐随机调度中的应用

纪昌明^{1,2}, 周 婷¹, 向腾飞², 黄海涛¹

(1. 华北电力大学 经济与管理学院, 北京 102206; 2. 华北电力大学 可再生能源学院, 北京 102206)

摘要: 将支持向量机(SVM)理论与网格搜索及交叉验证相结合,应用于梯级水电系统隐随机优化调度中,实现径流不确定条件下的梯级实际优化运行。以系统结构风险最小为 SVM 训练目标,结合参数分布规律,采用指数划分的网格搜索对模型参数进行优选;将 K-fold 交叉验证技术引入到 SVM 训练性能评价中,降低了训练样本随机性对训练模型性能的干扰,提高了模型的泛化能力。建立 VC_ 与 MATLAB 混合编程平台,对梯级水电系统隐随机优化调度运行进行仿真,结果表明基于采用最优参数 SVM 的隐随机优化调度在梯级系统发电量和发电过程方面取得了良好成果。

关键词: 支持向量机; 网格搜索; 交叉验证; 混合编程; 梯级水电系统; 优化; 水电

中图分类号: TM 732

文献标识码: A

DOI: 10.3969/j.issn.1006-6047.2014.03.021

0 引言

水电作为开发技术成熟、运行成本低廉、负荷响应快速的可再生能源,在我国能源格局中发挥着重要作用。随着“十二五”规划对我国水电开发积极有序的推动,在大规模水电系统建成后,如何在实际运行中挖掘水库水电站的发电效益空间、提高水电站实际调度水平,是一项兼具理论意义和实践价值的研究课题。

近年来,基于优化调度成果和规则提取的隐随机优化调度理论被越来越多地用于水库水电站实际运行研究中。从传统的线性回归^[1],到神经网络^[2]、遗传算法^[3-4]、蚁群算法^[5]等智能算法,以及贝叶斯网络理论^[6]等,均在隐随机优化调度中有所应用。但上述方法大多以训练误差最小为目标对调度规则进行提取,容易导致模型过度追求训练精度,用复杂的表达式去拟合有限的样本,而模型的复杂度与其推广能力往往成反比,致使模型的预测能力降低^[7]。

支持向量机 SVM(Support Vector Machine)以系统结构风险最小化为优化目标,利用惩罚参数平衡模型的学习能力和推广能力,将回归问题转化为二次规划问题,克服了局部最优的弊端,尤其适用于小样本回归问题。但另一方面,模型中惩罚参数及核函数参数的率定,以及小容量样本所带来的训练偏差是 SVM 在实际应用中的难点。

本文将 SVM 应用于梯级水电站群隐随机优化调度中,结合网格搜索对模型参数进行优化,并利用

K-fold 交叉验证方法评价模型的训练性能,降低了抽样随机性所带来的预测误差,提高了模型的推广能力。为提高成果的实用性,设计了基于 VC_ 和 MATLAB 混合编程的仿真平台,实现 SVM 指导梯级水电系统实际优化运行仿真,并对仿真结果进行了全面的评价。

1 梯级水电系统隐随机优化调度

水库水电站隐随机优化思想由美国学者 Young 提出^[8],其基本原理是:基于长系列已知径流过程,运用确定性优化方法得到调度期内最优调度过程,以此作为样本制定优化调度策略,指导径流未知条件下的实际运行。隐随机优化理论将随机优化问题分解为确定性优化调度和规则提取 2 步,避免了随机优化“维数灾”问题,因此在大型水电站群实际优化调度研究中得到了广泛的应用。

1.1 基于逐步优化算法的确定性优化调度模型

对于以发电功能为主的梯级水电系统,优化调度目标函数为调度期内发电量最大:

$$\max \sum_{i=1}^T \sum_{j=1}^J N_{i,j} \Delta t = \sum_{i=1}^T \sum_{j=1}^J k_i Q_{i,j} H_{i,j} \Delta t \quad (1)$$

其中, $N_{i,j}$ 为电站 j 在时段 i 的发电出力; k_i 为发电系数; $Q_{i,j}$ 为发电流量; $H_{i,j}$ 为发电水头; Δt 为时段长度; T 为调度期总时段数; J 为梯级系统中水电站个数。

约束条件如下。

a. 水库库容约束:

$$V_{i,\min} \leq V_{i,j} \leq V_{i,\max} \quad (2)$$

b. 下泄流量约束:

$$Q_{j,\min} \leq Q_j \leq Q_{j,\max} \quad (3)$$

c. 电站预想出力约束:

$$N_{j,\min} \leq N_j \leq N_{j,\max} \quad (4)$$

收稿日期:2013-05-01;修回日期:2014-01-14

基金项目:国家自然科学基金资助项目(51279062,41340022);
中央高校基本科研业务费专项资金资助项目(11QX52)

Project supported by the National Natural Science Foundation of China(51279062,41340022) and the Fundamental Research Funds for the Central Universities(11QX52)

d. 系统保证出力约束:

$$\text{prob}(N_j \geq N_{\text{lim},j}) \geq p_j^* \quad (5)$$

其中, $V_{i,j,\min}$ 和 $V_{i,j,\max}$ 为水库 j 在时段 i 的最小及最大库容限制; $Q_{j,\min}$ 和 $Q_{j,\max}$ 为水库 j 在时段 i 的最小及最大下泄流量约束; $N_{j,\min}$ 和 $N_{j,\max}$ 为电站 j 的最小和最大出力约束; $N_{\text{lim},j}$ 为电站 j 的保证出力; p_j^* 为 $N_{\text{lim},j}$ 对应的保证率; $\text{prob}(N_j \geq N_{\text{lim},j})$ 代表调度期内 $N_j \geq N_{\text{lim},j}$ 的频率。式(5)的含义是系统时段出力大于保证出力的频率必须满足设计保证率要求。

考虑到梯级水电站群变量众多并且可行域广阔,采用逐步优化算法 POA (Progressive Optimization Algorithm) 对模型进行求解:首先根据常规方法生成可行调度过程作为初始解,再从中选取两阶段,并固定其他阶段变量,对所选两阶段变量在发电量最大目标下进行优化。各时段依次逐步循环,直至达到目标函数不再增加为止。POA 将多阶段决策问题分解为一系列的二阶段决策问题,并已证明可收敛至全局最优^[9],极大提高了梯级电站优化问题的求解效率。

1.2 调度规则变量选取

调度规则变量包括输出变量(因变量)和输入变量(自变量),变量选取不但影响隐随机模型性能,也反映了调度规则的直观性和可解释性。一般而言,输出变量应取物理意义明确、易于操作的指标,如时段出力、下泄流量等。输入变量则要求能够准确反映水库水电站的运行状态和特征,同时应考虑调度实际,确保指标的可获取性。例如实际运行中径流预报预见期仅为当前时段,则未来时段的径流信息不可纳入输入变量中。

考虑到下泄流量较时段出力更为直观,且贴近水电站运行操作,本文选取时段下泄流量为输出变量。输入变量的选择面较宽,时段初水位和面临时段入库流量是与调度决策相关性最强的 2 个指标;入能和蓄能反映了水库的能量状态,与发电关系较为密切,因此也将其纳入;考虑水文过程的连续性,相近时段的调度决策往往互相影响,因此将上一时段的下泄流量也纳入输入变量。综上,输入变量包括时段初水位、入库流量、入能、蓄能、上一时段的下泄流量。

2 基于网格搜索和交叉验证的 SVM 回归模型

调度规则中的输出变量和输入变量构成了 SVM 回归的训练样本集 (x_1, y_1) 、 (x_2, y_2) 、 \dots 、 (x_n, y_n) , SVM 通过一个非线性映射 ϕ 将原始自变量投影到高维空间,将非线性回归转化为高维空间内的线性回归,求解过程不存在局部最优问题,具有坚实的理论基础^[10-12]。

目前常用的 SVM 参数优化方法包括梯度法^[13]、

遗传算法^[14]、粒子群算法^[15]等。这些方法具有收敛速度快的优点,但是在对 2 个以上参数进行优化时,参数之间互相影响,往往不能保证结果最优。本文选择网格搜索法对参数组合 (C, g) 同时进行寻优,规避了局部最优,并且便于并行计算,结合指数网格划分,提高了整体效率。

交叉验证(cross validation)是一种消除取样随机性所带来的训练偏差的统计学方法。常用的交叉验证方法有重复随机抽样法、 K -fold 交叉验证法、留一法等。考虑样本容量及计算效率,本文引入 K -fold 交叉验证法对模型的训练性能进行评价,结合参数优化,提高模型的稳定性和泛化能力。

2.1 SVM 回归基本原理

SVM 回归的目标函数为高维空间线性回归函数 $f(x) = w \cdot \phi(x) + b$ 的复杂度和拟合误差之和最小。假设所有样本均可以在不敏感度为 ε 、正、负松弛度分别为 ξ 、 ξ^* 水平下进行无误差拟合,则优化模型可表达为:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (6)$$

$$\begin{cases} (w \cdot \phi(x_i) + b) - y_i \leq \varepsilon + \xi_i \\ y_i - (w \cdot \phi(x_i) + b) \leq \varepsilon + \xi_i^* \end{cases}$$

其中, n 为测试样本个数。

目标函数的前半部分代表回归函数 $f(x)$ 的复杂度,即泛化(推广)能力,后半部分代表拟合误差。常数 $C > 0$ 为非负惩罚因子,起到平衡模型的泛化能力和拟合能力的作用。

模型式(6)是典型的凸二次规划问题,引入 Lagrange 函数:

$$L = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) - w \cdot \phi(x_i) + b - \sum_{i=1}^n \alpha_i [\varepsilon + \xi_i + y_i - f(x)] - \sum_{i=1}^n \alpha_i' [\varepsilon + \xi_i^* + f(x) - y_i] \quad (7)$$

其中, α_i 、 α_i' 为 Lagrange 因子。根据 Kuhn-Tucker 定理,式(7)的最优解在 w 和 b 方向上的梯度均应为 0:

$$\begin{cases} \nabla_w L(w, \alpha_i, \alpha_i', b) = \sum_{i=1}^n (\alpha_i - \alpha_i') = 0 \\ \nabla_b L(w, \alpha_i, \alpha_i', b) = w - \sum_{i=1}^n (\alpha_i - \alpha_i') x_i = 0 \end{cases} \quad (8)$$

其中, ∇ 为梯度算子。

将式(8)代回式(7),消去 w 和 b ,经整理可得原最优化问题的对偶问题:

$$\max \sum_{i=1}^n [(\alpha_i - \alpha_i') y_i - (\alpha_i - \alpha_i') \varepsilon] - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n [(\alpha_i - \alpha_i') (\alpha_j - \alpha_j') k(x_i, x_j)] \quad (9)$$

$$\text{s.t. } 0 \leq \alpha_i, \alpha_i' \leq C \quad i = 1, 2, \dots, n$$

$$\sum_{i=1}^n (\alpha_i - \alpha_i') = 0$$

其中, $k(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)$ 为核函数, 代表将 \mathbf{x}_i 与 \mathbf{x}_j 向高维空间投影后再进行内积运算。由此得到回归函数为:

$$f(\mathbf{x}) = \sum_{i=1}^n [(\alpha_i - \alpha'_i) k(\mathbf{x}, \mathbf{x}_i)] + b \quad (10)$$

式(10)表明, 只有 $\alpha_i - \alpha'_i \neq 0$ 所对应的样本对回归起到了实质性作用, 这些样本因此被称为支持向量(SV)。常用的核函数如下。

a. 多项式核函数:

$$k(\mathbf{x}_i, \mathbf{x}_j) = [(\mathbf{x}_i \cdot \mathbf{x}_j) + 1]^d \quad (11)$$

b. 径向基(RBF)核函数:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-g \|\mathbf{x}_i - \mathbf{x}_j\|^2) \quad (12)$$

c. sigmoid 核函数:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \tanh[g(\mathbf{x}_i \cdot \mathbf{x}_j) + c] \quad (13)$$

其中, d, g, c 均为核算函数参数。

2.2 基于网格搜索和交叉验证的参数优化

由 SVM 原理可知, 非负惩罚因子 C 和核函数参数 g 对模型的效果起着重要的作用。网格搜索法的优点是可以保证所得的搜索解是划定网格中的全局最优解, 避免重大误差。一般而言, 为便于计算和结果显示, 网格的划分多采用等距划分法, 但是在 SVM 模型中, 根据经验, C 和 g 一方面取值范围较广, 另一方面最优值落入 $[0, 1]$ 区间的概率又较大。若网格步长太小, 会导致计算量过大; 若步长过大, 则容易遗漏最优值。为此, 本文引入指数函数对网格进行划分。

另一方面, 由于对参数组合的性能评价受到训练数据的影响, 对于同一组 (C, g) , 当训练数据发生变化时, 拟合性能往往随之改变, 尤其对于小样本训练, 参数优选受到样本随机性的影响较大, 不利于模型的泛化和推广。鉴于此, 本文在网格搜索中应用 K -fold 交叉验证法对每组 (C, g) 的性能进行综合评价。基于网格搜索和 K -fold 交叉验证的参数优化步骤如下。

a. 建立网格坐标。令 $a = [-5, 5], b = [-5, 5]$, 步长均为 1, 取模型参数的网格点为 $C = e^a, g = e^b$, 根据指数函数性质, C 和 g 的取值可以保证非负, 且参数取值越小网格越密集, 参数越大取值越稀疏, 符合 (C, g) 的最优取值分布。网格点分布如图 1 所示。

b. 将训练数据等分为 K 个子集, K 一般取值为 4~10, 以确保训练样本容量远大于测试样本容量。

c. 对于网格中每一组 (C, g) , 以任意一个子集作为测试集, 其余 $K-1$ 个子集作为训练集, 训练模型后对测试集进行预测, 统计测试结果的均方误差:

$$\delta_{\text{MSE}} = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2 \quad (14)$$

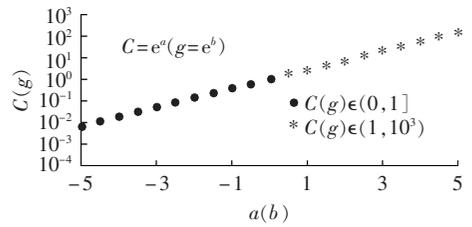


图 1 基于自然指数函数的参数网格划分
Fig.1 Parameter grid division based on natural exponential function

其中, \hat{Y}_i 为预测值; Y_i 为样本实际值。

d. 将测试集更换为另一子集, 再取剩余 $K-1$ 个子集作为训练集, 再次统计均方误差, 直至对 K 个子集都进行一次预测后, 取 K 组均方误差的平均值 $\bar{\delta}_{\text{MSE}}$ 作为该组 (C, g) 的预测误差。

e. 更换参数组合 (C, g) , 重复步骤 b—d, 顺序算出网格中各参数组合下训练模型的 $\bar{\delta}_{\text{MSE}}$ 并逐一比较, 平均误差最小的参数组合即为网格区间内的最佳参数组合。

将交叉验证与网格搜索相结合, 以参数平均误差最小化为参数优选的目标, 提高了参数优选的效率和准确性, 同时极大规避了训练样本的抽样随机性对模型性能的影响。在结果展示方面, 将 (C, g) 网格转换为对数坐标, 即对应 $(\lg C, \lg g)$, 则可呈现等间距的网格搜索结果。

3 基于混合编程的梯级水电系统跨平台仿真

基于 SVM 的梯级水电系统的运行仿真具有一定的特殊性。在大多数预测案例应用中, 各预测值之间是相互独立的, 例如径流预报、金融指数预测以及医学案例预测等, 因此预测部分与仿真部分也是相互独立的。而对于梯级水电运行而言, 一个时段的预测下泄流量(因变量)不但影响该电站下一时段的状态(自变量), 并且影响同时段下游电站的状态(自变量)。因此在梯级水电系统运行仿真中, 需要在每一时段基于当前的实际水库电站状态对调度决策进行预测, 再根据调度决策进行水电站状态转移, 实现滚动模拟。

梯级水电系统调度运行具有约束条件众多、电站间水力联系密切、状态变量相互影响的特点。VC_语言功能全面、逻辑预算效率高, 适合复杂系统设计, 是水库群调度仿真的经典平台。SVM 模型的特点是计算重复性较大, 输入输出频繁, 且主要以数据运算为主。MATLAB 语言具有良好的矩阵运算及数据处理能力, 输出形式直观多样化, 因此适合 SVM 模型训练及参数优化。为充分发挥 2 种语言的优势, 本文基于混合编程技术将 VC_ 与 MATLAB 模块进

行整合,以实现水库水电站群仿真运行,仿真设计框架如图 2 所示。

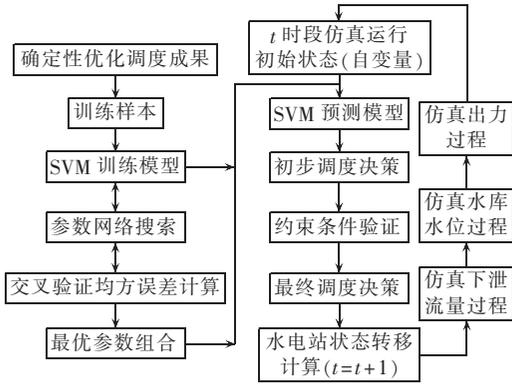


图 2 VC_与 MATLAB 混合编程仿真平台框架
Fig.2 Simulation framework of hybrid VC_ and MATLAB programming

VC_与 MATLAB 混合编程的实现方法是将 MATLAB 中 SVM 模型训练及预测模块(.m 文件)编译为能够被 VC_调用的动态库(.dll 文件),再将动态库按照一定格式添加到 VC_工程中,最后配置 VC_和 MATLAB 中对应的数据接口,以实现变量传输。文献[16]对混合编程的实现步骤进行了详细说明。

采用 VC_与 MATLAB 混合编程的仿真平台,能够充分发挥不同语言的优势,将基于 SVM 的调度规则制定部分和基于物理背景的仿真部分区分开来,使模型的结构更加清晰。另一方面,当需要调整调度规则制定方法时,只需对 MATLAB 模块进行修改;而当需要修改水电站调度物理背景时,只需改动 VC_部分,体现了良好的兼容性和可扩展性。

4 案例应用及结果评价

4.1 案例应用

某梯级水电系统由 12 座水电站组成,其中月调节性能以上电站 4 座,自上而下分别为第 1 级(年调节)、第 9 级(季调节)、第 10 级(季调节)和第 11 级(季调节)。其余电站为径流式电站,按照设计水位指导运行。系统调度期长度为 48 a,调度时段为月。

首先采用 POA 对调度期进行确定性优化调度,并取前 36 a 的确定性优化调度成果作为训练样本,运用 SVM 理论制定 4 座电站的月度调度规则,并采用网格参数优化及 4-fold 交叉验证对 SVM 性能进行优化。限于篇幅,以电站 1 的 6 月份为例,网格参数优化过程如图 3 所示,最优参数组合对应的拟合均方误差为 0.058。

对 4 座电站的各月参数优化结果进行统计,如图 4 所示。从图中参数取值分布可见,落入(0,1]区间的最优参数为 19 个,占优化参数总数的 50%,同

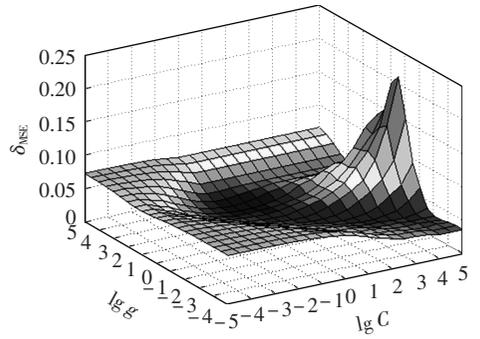


图 3 网格参数优化过程(以电站 1 的 6 月为例)
Fig.3 Process of grid parameter optimization (June, power station No.1)

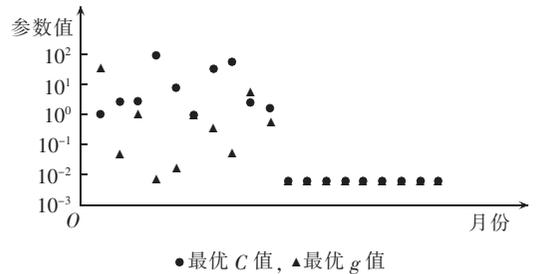


图 4 最优参数分布图
Fig.4 Distribution of optimal parameters

时参数在 $[e^{-5}, e^5]$ 区间内分布较广,参数优化结果验证了基于指数函数网格搜索设计的合理性。

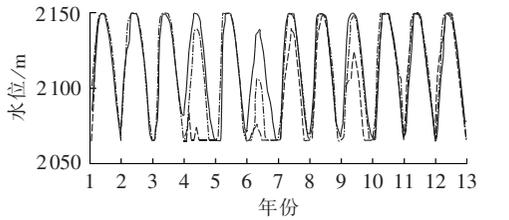
4.2 仿真运行及结果评价

基于 VC_和 MATLAB 混合编程平台,结合调度规则与水库水电站运行约束(式(2)—(5)),对调度期后 12 a 梯级水电系统运行进行仿真,并将仿真结果与同期确定性优化调度结果进行对比分析。为进一步检验参数优化的效果,另取默认参数 $C=1, g=1$,即认为模型中 C 和 g 参数均不起调节作用,以此 SVM 模型制定调度规则,并采用同样方法中对梯级运行进行仿真,作为结果评价的对照组方案。

a. 年均发电量分析。经统计,确定性优化调度、参数优化 SVM 仿真和默认参数 SVM 仿真 3 种方案下的梯级系统年平均发电量分别为 295.73 TW·h、288.75 TW·h 和 280.89 TW·h。基于参数优化的 SVM 仿真较确定性优化调度年均发电量减小 2.3%,高出无参数优化仿真结果 2.8%,说明基于 SVM 的隐随机调度在发电总量方面与确定性优化调度整体差距较小,而参数优化使年发电量效益得到了一定程度的提高。

b. 水库运行过程分析。将 4 座月调节以上电站在上述 3 种方案下的水库运行过程进行对比,如图 5—8 所示。

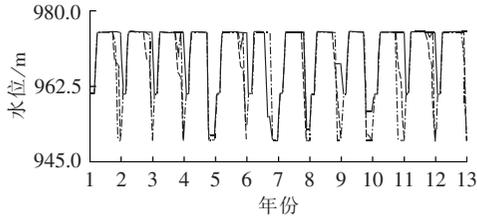
采用标准均方根误 NRMSD(Normalized Root Mean Square Deviation)对图 5—8 中 2 组仿真水位过程与确定性优化调度水位过程分别进行差异度



—确定性优化调度, ---基于默认参数的 SVM 仿真
-- 基于参数优化的 SVM 仿真

图 5 水库 1 运行过程对比

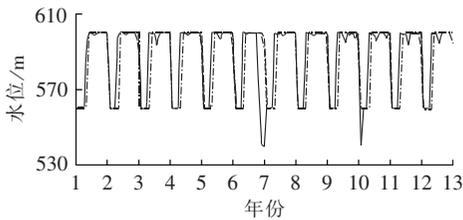
Fig.5 Comparison of operation process of reservoir No.1



—确定性优化调度, ---基于默认参数的 SVM 仿真
-- 基于参数优化的 SVM 仿真

图 6 水库 9 运行过程对比

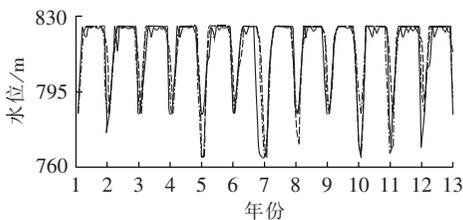
Fig.6 Comparison of operation process of reservoir No.9



—确定性优化调度, ---基于默认参数的 SVM 仿真
-- 基于参数优化的 SVM 仿真

图 7 水库 10 运行过程对比

Fig.7 Comparison of operation process of reservoir No.10



—确定性优化调度, ---基于默认参数的 SVM 仿真
-- 基于参数优化的 SVM 仿真

图 8 水库 11 运行过程对比

Fig.8 Comparison of operation process of reservoir No.11

分析:

$$\delta_{NRMSD} = \frac{\delta_{RMSD}}{y_{max} - y_{min}} = \frac{\sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / n}}{y_{max} - y_{min}} \quad (15)$$

其中, δ_{RMSD} 为均方根误差; y_{max} 和 y_{min} 分别为确定性优化运行水位最大、最小值; y_i 和 \hat{y}_i 分别为优化和仿真水位。

由表 1 可见, 基于参数优化的仿真 δ_{NRMSD} 整体

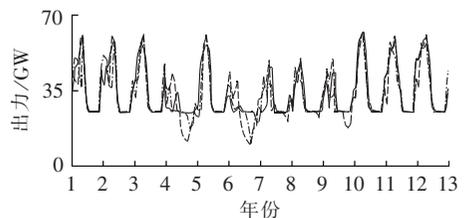
表 1 不同仿真方案下各水库标准均方根误差

Tab.1 NRMSD of reservoirs for different simulation scenarios

仿真方案	δ_{NRMSD}				平均
	水库 1	水库 9	水库 10	水库 11	
默认参数仿真	0.27	0.27	0.17	0.23	0.24
参数优化仿真	0.15	0.25	0.13	0.22	0.19

小于基于默认参数的仿真方案, 尤其对于具有年调节性能的水库 1。这说明基于参数优化的仿真过程更为接近确定优化调度过程, 在实际运行中具有更高的可控性。

c. 系统保证出力评价。图 9 展示了梯级系统在 3 种方案下的出力过程。在确定性优化调度中, 系统保证出力(25.22 GW)仅在第 6 年发生了 4 个时段破坏, 破坏深度为 20%。在参数优化仿真运行的过程中, 梯级系统保证出力得到了较好的保障, 破坏时段数为 7, 较确定性优化调度增加了 3 个时段, 历时保证率为 95%。其中第 6 年破坏深度与优化调度相比较严重, 这是由于第 6 年为特枯年份, 在仿真运行中, 调度规则未能利用径流预报进行水库预蓄等减小破坏手段, 导致供水后期发生集中破坏, 这一点在图 5—8 中各水库第 6 年相应的仿真运行中也有所反映, 水库 1 在该年份汛期蓄水明显不足。而在采用默认参数组合的仿真过程中, 系统保证出力破坏时段数增加至 18, 历时保证率为 88%, 且破坏深度明显加大。



—确定性优化调度, ---基于默认参数的 SVM 仿真
-- 基于参数优化的 SVM 仿真

图 9 优化和仿真运行梯级系统出力对比

Fig.9 Comparison of system output between optimal and simulative operations

上述分析表明, 基于参数优化的 SVM 隐随机优化调度能够较好地实现确定性优化调度在发电量、发电过程等方面的成果, 并且随着评价标准的深入, 参数优化所带来的仿真性能改进也逐步明显, 可以看出, 参数优化后的仿真出力过程更为可靠, 电能质量更高, 更有利于电力系统的整体运行。

5 结论

本文将 SVM 优化理论应用于梯级水电系统隐随机优化调度中, 结合指数网格搜索和 K-fold 交叉验证手段, 降低了模型预测误差, 提高了模型的泛化

能力。建立基于 VC_ 和 MATLAB 混合编程的仿真平台,实现了梯级水电系统的实际优化运行,具有较好的可操作性和兼容性。对梯级水电系统进行隐随机调度仿真,结果表明,基于 SVM 的隐随即优化调度运行基本能够实现确定性优化调度的各方面效益,参数优化机制显著提高了实际调度效益,为提高水电站实际运行水平提供了有效的技术支撑。

该方法应用于水库水电站实际运行中可能存在的挑战有以下 2 点。

a. 梯级电站之间的调度决策信息共享问题。由于上级电站的下泄流量是下级电站入库流量的一部分,而入库流量是 SVM 模型输入变量之一,因此下级电站在制定调度决策时须基于上级电站的决策信息,这对于隶属不同发电主体的电站之间或有难度。

b. 系统保证出力的协调问题。系统在枯期保证出力的维持需各电站(尤其是调节性电站)互相协调,而当电站间发电效益存在冲突,以及特枯年份面临破坏时,不同电站及其与系统之间的协调机制需要进一步完善。此外,鉴于当前中长期径流预报水平的提高,今后可尝试将预报径流信息作为输入变量纳入隐随机优化模型,进一步提高水电站群中长期调度水平。

参考文献:

- [1] 纪昌明,苏学灵,周婷,等. 梯级水电站群调度函数的模型与评价[J]. 电力系统自动化,2010,34(3):33-37.
JI Changming,SU Xueling,ZHOU Ting,et al. Model establishment and evaluation of operation function for cascade reservoirs[J]. Automation of Electric Power Systems,2010,34(3):33-37.
- [2] 刘攀,郭生练,庞博,等. 三峡水库运行初期蓄水调度函数的神经网络模型研究及改进[J]. 水力发电学报,2006,25(2):83-89.
LIU Pan,GUO Shenglian,PANG Bo,et al. A modified approach for deriving storage operating rules of the Three Gorges reservoir with artificial neural network[J]. Journal of Hydroelectric Engineering,2006,25(2):83-89.
- [3] LI Chen,JAMES M,WILLIAM W G Y. A diversified multiobjective GA for optimizing reservoir rule curves[J]. Advances in Water Resources,2007,30(5):1082-1093.
- [4] 张东晓,王静,范炜,等. 基于并行遗传算法的水电站群中期优化调度[J]. 电力自动化设备,2012,32(12):87-91.
ZHANG Dongxiao,WANG Jing,FAN Wei,et al. Midterm optimal operation based on parallel genetic algorithm for hydropower stations[J]. Electric Power Automation Equipment,2012,32(12):87-91.
- [5] 纪昌明,喻杉,周婷,等. 蚁群算法在水电站调度函数优化中的应用[J]. 电力系统自动化,2011,35(20):103-107.
JI Changming,YU Shan,ZHOU Ting,et al. Application of ant colony algorithm for hydropower dispatching function optimization[J]. Automation of Electric Power Systems,2011,35(20):103-107.
- [6] BAHRAM M,REZA K,BANAFSHEH Z. Developing monthly operating rules for a cascade system of reservoirs application of Bayesian networks[J]. Environmental Modeling & Software,2009,24(12):1420-1432.
- [7] 李祥泳,蔡辉,丁晶. BP 网络的学习能力与推广能力之间满足的不确定关系式[J]. 四川大学学报:工程科学版,2003,35(1):15-19.
LI Zuoyong,CAI Hui,DING Jing. Uncertainty relation between learning ability and generalization ability of BP neural network[J]. Journal of Sichuan University:Engineering Science Edition,2003,35(1):15-19.
- [8] YOUNG G K. Finding reservoir operating rules[J]. Journal of the Hydraulics Division,1967,93(6):297-322.
- [9] HOWARD R A. Dynamic programming and Markov processes[M]. Cambridge,UK:Technology Press of Massachusetts Institute of Technology,1960:70-78.
- [10] 凌武能,杭乃善,李如琦. 基于云支持向量机模型的短期风电功率预测[J]. 电力自动化设备,2013,33(7):34-38.
LING Wuneng,HANG Naishan,LI Ruqi. Short-term wind power forecasting based on cloud SVM model[J]. Electric Power Automation Equipment,2013,33(7):34-38.
- [11] 韩勇,李红梅. 基于小波分解的支持向量机母线负荷预测[J]. 电力自动化设备,2012,32(4):88-91.
HAN Yong,LI Hongmei. Bus load forecasting based on wavelet transform and SVM[J]. Electric Power Automation Equipment,2012,32(4):88-91.
- [12] VAPNIK V N. The nature of statistical learning theory[M]. 2nd ed. New York,USA:Springer-Verlag,1999:219-230.
- [13] 刘宝英,杨仁刚. 基于主成分分析的最小二乘支持向量机短期负荷预测模型[J]. 电力自动化设备,2008,28(11):13-17.
LIU Baoying,YANG Rengang. Short-term load forecasting model based on LS-SVM with PCA[J]. Electric Power Automation Equipment,2008,28(11):13-17.
- [14] 刘昌平,范明钰,王光卫,等. 基于梯度算法的支持向量机参数优化方法[J]. 控制与决策,2008,23(11):1291-1295,1300.
LIU Changping,FAN Mingyu,WANG Guangwei,et al. Optimizing parameters of support vector machine based on gradient algorithm[J]. Control and Decision,2008,23(11):1291-1295,1300.
- [15] 徐茹枝,王宇飞. 粒子群优化的支持向量回归机计算配电网理论线损方法[J]. 电力自动化设备,2012,32(5):86-89,93.
XU Ruzhi,WANG Yufei. Theoretical line loss calculation based on SVR and PSO for distribution system[J]. Control and Decision,2012,32(5):86-89,93.
- [16] 徐鑫鑫,刘涤尘,黄涌. 基于 VC++ 和 Matlab 混合编程实现电力故障再现及分析系统研究[J]. 电力自动化设备,2006,26(12):38-40,44.
XU Xinxin,LIU Dichen,HUANG Yong. Power system fault re-occurrence and analysis system based on hybrid programming of VC++ and Matlab[J]. Electric Power Automation Equipment,2006,26(12):38-40,44.

作者简介:



纪昌明

纪昌明(1956-),男,湖北洪湖人,教授,博士研究生导师,博士,研究方向为水库优化调度、洪灾风险管理;

周婷(1985-),女,安徽六安人,博士研究生,研究方向为水库群隐随机优化调度、多目标优化理论(E-mail:zhouting02@gmail.com)。

Application of support vector machine based on grid search and cross validation in implicit stochastic dispatch of cascaded hydropower stations

JI Changming^{1,2}, ZHOU Ting¹, XIANG Tengfei², HUANG Haitao¹

(1. School of Economic and Management, North China Electric Power University, Beijing 102206, China;

2. School of Renewable Energy, North China Electric Power University, Beijing 102206, China)

Abstract: The combination of SVM (Support Vector Machine) theory combined with grid search and cross validation is applied to the implicit stochastic optimal dispatch of cascaded hydropower stations under the condition of uncertain inflow. With the minimum system structure risk taken as the training objective of SVM and the parameter distribution law combined, the exponentially divided grid is searched for determining the optimal model parameters. K -fold cross validation technique is adopted in the evaluation of SVM training performance, which decreases the influence of training sample randomness on the performance of training model and improves its generalization ability. The implicit stochastic optimal dispatch of cascaded hydropower stations is simulated on the hybrid programming platform of VC_ and MATLAB, which shows that, the implicit stochastic optimal dispatch based on SVM with optimized parameters enhances both power generation capacity and process.

Key words: support vector machines; grid search; cross validation; hybrid programming; cascaded hydropower stations; optimization; hydroelectric power

(上接第 107 页 continued from page 107)

Decoupled three-phase power flow calculation for distribution network with DGs

YANG Xiong¹, WEI Zhinong¹, SUN Guoqiang¹, SUN Yonghui¹, DING Xiaohua², XU Xiaohui²

(1. Research Center for Renewable Energy Generation Engineering, Ministry of Education, Hohai University,

Nanjing 210098, China; 2. China Electric Power Research Institute (Nanjing), Nanjing 210003, China)

Abstract: A method of decoupled three-phase power flow calculation is proposed for distribution network with DGs. The three-phase load models, sequence parameter models of distribution network and the grid-connection models of different DG types are built based on the sequence component method. Based on the structure of distribution network, the three-sequence decoupling-compensation models of unsymmetrical line and the path-loop analysis method, an effective method of improved power flow calculation in sequence network is proposed for the unbalanced three-phase distribution network. The grid-connection interfaces for different DG types are classified into PQ, PQ(V), PV and PI node types, the corresponding DG models suitable for the power flow algorithms of unbalanced three-phase distribution network are built, and their iterative calculation models are derived. Case analysis verifies the effectiveness, versatility and convergence of the proposed algorithm, showing its better performance in DG node handling and DG reactive power boundary treatment.

Key words: distribution network; three-phase decoupling; power flow calculation; distributed power generation; three-phase unbalance; models