

DOI: 10.13476/j.cnki.nsbdtqk.2018.0070

王迁, 杨明祥, 雷晓辉, 等. 基于 PSO-SVR 的丹江口年径流预报[J]. 南水北调与水利科技, 2018, 16(3): 65-71. WANG Q, YANG M Xi, LEI X H, et al. Annual runoff forecast for Danjiangkou based on PSO-SVR[J]. South to North Water Transfers and Water Science & Technology, 2018, 16(3): 65-71. (in Chinese)

基于 PSO-SVR 的丹江口年径流预报

王 迁¹, 杨明祥², 雷晓辉², 舒 坚¹, 孙利民³, 黄雪姝⁴

(1. 南昌航空大学 软件学院, 南昌 330063; 2. 中国水利水电科学研究院 流域水循环模拟与调控国家重点实验室, 北京 100038; 3. 中国科学院 信息工程研究所, 北京 100093; 4. 黄河水利委员会 信息中心, 郑州 450003)

摘要: 目前应用于丹江口水库年径流预报的方法主要为物理统计和人工神经网络(ANN)等方法,但这些方法普遍存在预报精度不高和稳定性不强等缺点。选择回归支持向量机(SVR)模型应用于丹江口水库年径流预报,针对惩罚系数 C 、核参数 σ 和不敏感损失系数 ϵ 三个参数在实际赋值过程中存在计算量大、难以得到最优值等问题,将粒子群优化算法(PSO)加入到SVR模型中,建立PSO-SVR模型,实现了参数的自动优选。结果表明,PSO-SVR模型较之SVR模型,提高了预报精度;较之ANN模型,稳定性更强,可信度更高。该模型具有较好的应用价值,可为南水北调中线工程调度方案制定提供一定的参考依据。

关键词: 丹江口水库; 回归支持向量机; 粒子群优化算法; 年径流预报; 预报因子

中图分类号: TV 121 **文献标志码:** A **文章编号:** 1672-1683(2018)03-0065-07

Annual runoff forecast for Danjiangkou based on PSO-SVR

WANG Qian¹, YANG Mingxiang², LEI Xiaohui², SHU Jian¹, SUN Limin³, HUANG Xueshu⁴

(1. School of Software, Nanchang Hangkong University, Nanchang 330063, China; 2. State Key Laboratory of Water Cycle Simulation and Regulation, China Academy of Water Resources and Hydropower Research, Beijing 100038, China; 3. Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China; 4. Information Center of Yellow River Conservancy Commission, Zhengzhou 450003, China)

Abstract: At present, the methods of annual runoff forecast for Danjiangkou reservoir mainly include physical statistical approach and artificial neural network (ANN). However, these methods have the disadvantages of low accuracy and low stability. In this paper, we applied the regression support vector machine (SVR) model to the annual runoff forecast for Danjiangkou Reservoir. Considering that the penalty coefficient C , the kernel parameter σ , and the insensitive loss coefficient ϵ all require a large amount of calculation and it is difficult to obtain their optimal value in the actual assignment process, we added the particle swarm optimization (PSO) algorithm to the SVR model and established a PSO-SVR model to realize the automatic optimization of parameters. The results showed that the PSO-SVR model has higher prediction accuracy compared with the SVR model, and has better stability and reliability than the ANN model. The model has a good application value, and can provide some reference for the development of the operation scheme of the middle route of the South to North Water Transfer Project.

Key words: Danjiangkou Reservoir; regression support vector machine; particle swarm optimization; annual runoff forecast; forecast factor

收稿日期: 2017-10-09 修回日期: 2017-12-27 网络出版时间: 2017-12-29

网络出版地址: <http://kns.cnki.net/kcms/detail/13.1334.TV.20171228.1811.012.html>

基金项目:“十三五”国家重点研发计划子课题“梯级水库群影响下流域水文循环演变规律研究”(2016YFC0402201-01); 中国水利水电科学研究院基本科研业务费专项项目(WR0145B212017)

Funds: Sub-topic of National Key R & D Program of the 13th Five Year Plan Period "Research on Evolution Laws of Hydrological Cycle in the Basin under the Influence of Cascade Reservoirs"(2016YFC0402201-01); Special Fund for Basic Scientific Research Expenditure of China Water Conservancy and Hydropower Research Institute (WR0145B212017)

作者简介: 王 迁(1992-), 男, 江西丰城人, 主要从事中长期降雨、径流预报研究。E-mail: 963282600@qq.com

通讯作者: 杨明祥(1986-), 男, 河南新乡人, 工程师, 博士, 主要从事气象水文预报, 水利信息化研究。E-mail: yangmx@iwhr.com

丹江口水库位于汉江中上游,南水北调中线工程水源地^[1]。丹江口水库总面积 846 km²,多年平均入库水量 394.8 亿 m³,丹江口大坝加高以后,水库正常蓄水位提高至 170 m,库容达到 290.5 亿 m³,水域面积达到 1 022.75 km²。2012 年开始向南水北调中线工程沿线地区的河南、河北、北京、天津等 4 个省市的 20 多座大中城市提供用水,有效缓解中国北方部分地区的水资源严重短缺局面^[2]。年径流预报由于具有较长的预见期,对水库的优化管理和综合调度有着重要的指导意义和经济价值^[3]。因此分析丹江口水库年入库径流特性和演变规律,准确预报水库来水,对南水北调中线工程实际调度也有着重要意义。

近年来,人工神经网络(Artificial Neural Network, ANN)被广泛应用在水文分析和水文预报中,尤其在长期径流预报中取得了不错的效果^[4],但是用 ANN 模型预测径流时,容易出现模型稳定性不好和预报误差较大的问题,且往往拟合效果好而预测效果差^[5]。针对以上存在的问题,以及考虑到实际应用中惩罚系数、核参数和不敏感损失系数的选取对模型性能有较大影响,且三个参数赋值存在计算量大、难以得到最优值等问题,本文将粒子群优化算法(PSO)加入到 SVR 模型中,建立了 PSO-SVR 模型,并将此模型应用于丹江口水库的年径流预报中,取得了较好的效果,以期期为南水北调中线工程调度方案制定提供一定的依据。

1 SVR 基本原理及其算法实现

支持向量机(Support Vector Machine, SVM)在 20 世纪 90 年代中期被 Vapnik 等人首次提出,它是一种以统计学习的 VC 维理论为理论基础,以结构风险最小化原理为目标新兴的机器学习方法^[6]。支持向量机主要用于解决分类和回归预测问题,针对回归预测问题,研究者提出了一系列回归算法,回归支持向量机(Support Vector Machine for Regression, SVR)就是其中的一种,尤其适用于小样本、非线性问题的回归预测^[7]。SVR 的基本思想是利用核函数^[8]将低维度非线性问题转换成高维度线性问题,在高维特征空间中利用线性方法解决非线性问题^[9]。通过多次试验证明,SVR 在解决回归预测问题时具有较高的可信度和良好的泛化能力。

SVR 实现回归预测的步骤如下。

给定训练集 T

$$T = \{(x_1, y_1), \dots, (x_l, y_l)\} \in (X \times Y)^l \quad (1)$$

其中, $x_i \in X = R, y_i \in Y = R, i = 1, 2, \dots, l$ 都是所

选样本, x_i 为输入, y_i 为输出。回归支持向量机的方法就是寻找一个映射 Φ^0 , 将低维空间的非线性问题映射到高维特征空间 H 中,使得训练集变成了:

$$T = \{(\Phi(x_i), y_i), \dots, (\Phi(x_l), y_l) \in (X \times Y)^l\} \quad (2)$$

其中, $\Phi(x_i) \in H, y_i \in Y = R, i = 1, 2, \dots, l$, 随后在 H 中构造回归超平面^[10]。即:

$$\left\{ \begin{aligned} & \min_{w, b, \zeta} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\zeta_i + \zeta_i^*) \\ & s. t. \begin{cases} (w \Phi(x_i)) + b - y_i \leq \varepsilon + \zeta_i \\ y_i - (w \Phi(x_i)) - b \leq \varepsilon + \zeta_i^* \\ \zeta_i, \zeta_i^* \geq 0 \end{cases} \end{aligned} \right. \quad (3)$$

其中 $\zeta^* = (\zeta_1, \zeta_1^*, \dots, \zeta_l, \zeta_l^*)^T$; ε 为不敏感损失系数,其控制的是函数的精度,且 $\varepsilon > 0$; C 是一个常数,其控制对错分样本的惩罚程度,且 $C > 0$ 。

采用对偶原理^[11]、Lagrange 乘法子法即可求得式(3)、式(4)的对偶形式分别为:

$$\min_{\alpha^*} \frac{1}{2} \sum_{i,j=1}^l (\alpha_i^* - \alpha_j^*) (\Phi(x_i) \Phi(x_j)) + \varepsilon \sum_{i=1}^l (\alpha_i^* + \alpha_i) - \sum_{i=1}^l y_i (\alpha_i^* - \alpha_i) \quad (5)$$

$$s. t. \begin{cases} \sum_{j=1}^l (\alpha_j - \alpha_j^*) = 0 \\ 0 \leq \alpha_i, \alpha_i^* \leq C \end{cases} \quad i = 1, 2, \dots, l \quad (6)$$

利用式(6)得到 α, α^* 的值,再根据 KKT 条件^[12],即:

$$y_i - \sum_{j=1}^l (\alpha_j - \alpha_j^*) (\Phi(x_i) \Phi(x_j)) - b = \varepsilon \quad (7)$$

得到参数 b 的值,最后得到最优回归超平面为:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) (\Phi(x_i) \Phi(x)) + b \quad (8)$$

式(8)中对应系数 $(\alpha_i - \alpha_i^*) \neq 0$ 的样本 (x_i, y_i) 称为支持向量^[13]。

SVR 通过引入核函数 $K(x, x')$ 来代替式(8)中的内积 $(\Phi(x_i) \Phi(x))$, 从而最优回归超平面可表示为:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x, x') + b \quad (9)$$

常用的核函数有如下几种。

- (1) 线性核函数。 $K(x, x') = x \dot{x}'$;
- (2) 多项式核函数。 $K(x, x') = [(x \dot{x}') + 1]^d$;
- (3) 径向基核函数。

$K(x, x') = \exp(-\|x - x'\|^2 / \sigma^2)$, 其中, σ 为尺度参数,其反应了函数图像的宽度。

- (4) Sigmoid 核函数。

线性核函数一般用于分类,适用于线性可分的情况,而此时得到的是样本空间中的超平面;多项式核函数的特点时将低维问题转换到高维特征空间,但是参数多,计算复杂度较高;径向基核函数可将低

维非线性问题转换成高维线性问题,降低计算复杂度,对大小样本都有较好的性能,而且其相对于多项式核函数参数要少,因此大多数情况下优先使用径向基核函数;采用 sigmoid 核函数时,支持向量机实现的就是一种多层神经网络^[14]。

径流预报为多因子影响的非线性问题,所以线性核函数在此并不适用;而选用 sigmoid 核函数的模型预报效果与神经网络的效果相似,结果并不好;对于多项式核函数,其函数复杂度高,而且随着径流历史样本数量的增多,计算复杂度可能大到无法计算。所以,基于复杂度和优先原则的考虑,本文将径向基核函数作为 SVR 模型的核函数。

核函数选定以后,SVR 模型就可以抽象表示为:

$$y = f(x|C, \varepsilon, \sigma) \quad (10)$$

2 PSO-SVR 预报模型

年径流预报的首要环节是筛选预报因子,目前筛选预报因子的常用方法有:相关概率法、相关系数法、斯皮尔曼等级相关系数法等^[15]。在 SVR 模型中,预报因子选定后,选择历史年径流数据作为历史样本,将其分为训练集和检验集两部分,并对训练集进行训练,历史径流数据和相应的预报因子值作为模型输入,径流预测结果作为模型输出,即得到年径流预测结果。

2.1 预报因子的挑选

本文采用相关关系分析法和物理成因分析法共同确定最终 SVR 模型的预报因子,具体步骤如下。

首先,从中国气象局国家气候中心(http://cmdp.ncc-cma.net/Monitoring/cn_index_130.php?ui2468)获取百项气候系统指数集(即 88 项大气环流指数、26 项海温指数和 16 项其他指数)。采用相关关系分析法确定年径流量 y 和各预报因子 x 的相关性系数。相关性系数计算公式为:

$$\rho = \frac{\sum_{i=1}^l (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^l (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^l (y_i - \bar{y})^2}} \quad (11)$$

式中: ρ 表示第 i 个因子与年径流量 y 的相关系数; \bar{x} 表示某个因子多年平均值; $\bar{x} = \frac{1}{l} \sum_{i=1}^l x_i$ 表示样本径流平均值, $\bar{y} = \frac{1}{l} \sum_{i=1}^l y_i$; l 为径流样本年数。

根据式(11),可以计算得到各个因子 x 与径流量 y 之间的相关性系数大小,将相关性系数值按从大到小的顺序排列,挑选出相关性系数位于前 20 位的因子作为初选因子,并根据式(11)对初选因子两两分析

互相关性,去掉互相关性系数大于 0.4 的两者中的一个(一般将与径流相关性系数较小的因子去掉),然后对初选因子进行物理成因分析,并剔除对研究区没有物理影响的因子,剩下的因子即作为最终的预报因子。

2.2 参数对 SVR 模型预报精度的影响

由 SVR 抽象模型表达式(式(10))可知,SVR 模型的预测效果主要由惩罚系数 C 、核函数参数 σ 以及不敏感损失系数 ε 三个参数决定^[16]。

(1) σ 对 SVR 的影响。

当 σ 取值很小时,SVR 模型对样本训练集的拟合效果较好,但是对检验集的预测效果却很差;而当 σ 取值很大时,SVR 模型对样本训练集的拟合效果不好,同时对检验集的预测效果也不好。

(2) C 对 SVR 的影响。

C 越小,对经验风险的惩罚越小,回归函数曲线越平滑,SVR 模型的复杂度越小,经验风险的值越大^[17];当 C 逐渐增大时,训练和测试的误差都将减小^[18]。同时 C 的取值也将直接影响模型的泛化能力和稳定性。

(3) ε 对 SVR 的影响。

ε 的值越大,支持向量的个数就越少,模型的拟合效果变差^[19]; ε 的值越小,函数精度越高,模型复杂度增加^[20]。

综上所述,为保证 SVR 模型各方面的良好性能,需要对以上三个参数进行优选。

2.3 PSO-SVR 模型

目前多数文献确定 C 、 σ 和 ε 三个参数普遍采用的方法是网格搜索法,该算法以网格搜索和交叉验证相结合的方式确定参数值,计算量大,且容易陷入局部最优解,从而导致预报精度不高^[21]。为此,本文用 PSO 算法代替网格搜索法进行参数的优选,与其他算法相比,PSO 算法具有实现容易、收敛速度快、全局搜索能力强等优点^[22],所以将粒子群算法应用在 SVR 模型参数寻优上具有一定的优势。

基于 PSO 的 SVR 预报模型的径流预测步骤如下(图 1)。

(1) 确定预报因子。按照本文 2.1 节中的方法得到预报因子。

(2) 把步骤(1)中筛选好的预报因子值作为 SVR 模型输入,利用公式 $y_i^* = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}}$ 对输入进行归一化处理;其中 y_i , y_{\max} , y_{\min} 分别表示选定时间段内某一时刻任意预报因子数值、选定时间段内所

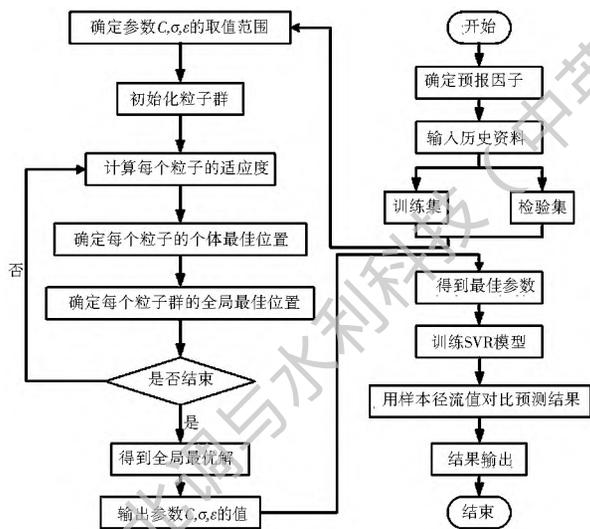


图 1 PSO-SVR 预报模型步骤

Fig. 1 Step diagram of PSO-SVR forecast model

选预报因子数值中的最大值、选定时间段内所选预报因子数值中的最小值, y_i^* 即为该时间上因子的归一化值。遍历所有选择的预报因子, 得到每个预报因子的归一化值。

(3) 选定历史样本, 如某水库或流域近 20 年年径流量数据, 对样本进行分类, 将前 15 年的年径流量数据和步骤(2)中选定的预报因子对应 15 年的归一化值作为训练集, 后 5 年的年径流量数据和步骤(2)中选定的预报因子对应 5 年的归一化值作为检验集。

(4) 采用 PSO 算法确定步骤(3) SVR 模型中 C, σ, ϵ 等各参数值。具体步骤如下。

Step1: 确定 C, σ, ϵ 三个参数的取值范围(粒子群的位置和速度都与三个参数的取值范围有关)。

Step2: 初始化粒子群。即设置粒子群规模、迭代代数、随机位置和速度等。

Step3: 确定适应度评估函数。并由适应度评估函数计算每个粒子的适应度值^[23]。

Step4: 确定每个粒子的个体最佳位置。将每个粒子当前位置的适应度值与其历史最佳位置 P_{best} (即局部最优解)的适应度值作比较, 如果当前位置适应度值大于 P_{best} 的适应度值, 则将其作为当前的最佳位置 P_{best} 。

Step5: 确定整个粒子群的全局最佳位置。将每个粒子当前最佳位置的适应度值与整个群体的当前最佳位置 g_{best} (即全局最优解)的适应度值作比较, 如果当前最好位置适应度值大于 g_{best} 的适应度值, 则将其作为当前的最佳位置 g_{best} 。

Step6: 根据式(12)、式(13)更新粒子速度和位置:

$$v_{i+1} = wv_i + c_1r_1(p_{best} - x_i) + c_2r_2(g_{best} - x_i) \quad (12)$$

$$x_{i+1} = x_i + v_{i+1} \quad (13)$$

式中: i 表示迭代代数; x_i 表示第 i 次迭代时粒子所在位置; v_i 表示第 i 次迭代时粒子的速度; r_1, r_2 为 $(0, 1)$ 之间的两个随机数; c_1, c_2 表示的是增速因子, 它们的取值均大于 0, 一般都取为 2; w 为权重因子, 取值范围是 $(0, 1)$ 。

Step7: 判断算法是否满足结束条件, 不满足则转到 Step3; 满足则输出最优结果, 此时的全局最优解即是三个参数的最优值。

(5) 将 PSO 算法得到的三个参数的最优解输入到 SVR 模型中。

(6) 用样本径流值对比预测结果并输出预测结果。

3 实例应用

3.1 因子选择和参数优化

现有丹江口水库 1981-2016 年的年平均入库流量资料, 其中, 1981-2001 年的数据作为训练集, 2002-2016 年的数据作为检验集。经过相关性分析, 得到百项气候系统指数与年径流值的相关性系数较大的前 20 个因子, 见表 1。

表 1 各预报因子与丹江口年径流值相关性系数大小
Tab. 1 Correlation coefficient between each forecast factor and the annual runoff value of Danjiangkou

序号	因子	相关性系数
1	前 1 年 5 月北半球极涡中心纬向位置指数	-0.458
2	前 1 年 11 月冷空气次数	0.432
3	前 1 年 6 月极地-欧亚遥相关型指数	0.426
4	前 2 年 11 月北大西洋-欧洲区极涡强度指数	0.423
5	前 2 年 2 月北大西洋-欧洲环流 E 型指数	-0.396
6	前 1 年 12 月太平洋区极涡面积指数	0.388
7	前 1 年 5 月极地-欧亚遥相关型指数	-0.388
8	前 2 年 12 月太平洋区极涡面积指数	0.378
9	前 1 年 9 月西风漂流区海温指数	0.372
10	前 2 年 2 月北半球极涡中心经向位置指数	-0.369
11	前 1 年 7 月北大西洋-欧洲环流 W 型指数	-0.365
12	前 1 年 6 月 500hPa 纬向风指数	0.352
13	前 1 年 12 月东亚槽位置指数	-0.351
14	前 1 年 11 月类 ENSO 指数	-0.348
15	前 2 年 1 月东太平洋副高面积指数	0.344
16	前 2 年 1 月东太平洋副高强度指数	0.344
17	前 1 年 6 月太平洋区极涡强度指数	0.343
18	前 2 年 4 月印槽强度指数	0.341
19	前 1 年 9 月太平洋区极涡面积指数	0.338
20	前 1 年北美大西洋副高脊线	0.337

经查阅资料得知, 丹江口水库处于东亚副热带季风区, 其降水主要来源于东南和西南两股暖湿气

流^[24]。大气环流关键区域基本分布在对我国天气有重大影响的北半球极涡、西风带、东亚大槽以及西太平洋副热带高压等区域,东亚大槽是北半球中高纬度对环流西风带形成的低压槽^[25]。而海温对中国天气影响较大的为太平洋、印度洋等海域。并且考虑到大气环流因子影响周期一般不超过一年,再对照表 1 中的因子和相关性系数,选择相关性系数大的,去掉因子间互相关性系数大的其中一个或者对丹江口水库降雨、径流没有物理成因影响的因子。最终选择的预报因子为:前 1 年 5 月北半球极涡中心纬向位置指数(因子 1)、前 1 年 11 月冷空气次数(因子 2)、前 1 年 6 月极地-欧亚遥相关型指数(因子 3)、前 1 年 12 月太平洋区极涡面积指数(因子 6)、前 1 年 9 月西风漂流区海温指数(因子 9)、前 1 年 12 月东亚槽位置指数(因子 13)。

在丹江口水库 36 年的年平均入库流量资料中,将前 21 年的数据作为模型训练集样本,后 15 年的数据作为检验集的样本,并将因子值和径流值归一化。模型输入为归一化后的因子值和实测径流值,模型输出为反归一化后的丹江口水库 36 年年平均径流预测值。经过多次试验,在 PSO 算法的参数设置中,粒子群数 $n=100$,最大迭代次数 $pcount=300$,学习因子 $c_1=2, c_2=2, w=0.8$,同时以确定性系数作为适应度评估函数,寻求确定性系数的最大值。

确定性系数的计算公式为:

$$DC=1-\frac{\sum_{i=1}^n [y_c(i)-y_o(i)]^2}{\sum_{i=1}^n [y_c(i)-\bar{y}_o]^2} \quad (14)$$

式中: DC 为确定性系数; $y_c(i)$ 为预测值; $y_o(i)$ 为实测值; \bar{y}_o 为多年实测值的平均值^[26], n 为样本年数。

经过 PSO 算法迭代后,确定性系数 $DC=0.94$,由此得到丹江口年径流预报 SVR 模型的最佳参数值 $(C, \epsilon, \sigma) = (4.5709, 0.0002, 1.9506)$,然后将三个参数值代入到模型中进行预测,得到训练集和检验集的预测结果。

3.2 结果分析

为了检验本文所述方法的可靠性,本文选取 ANN 模型、SVR 模型、以及 PSO-SVR 模型进行年径流预测对比。其中,ANN 模型采用的是三层神经网络结构,经过多次实验,最终其相关参数设置如下:隐含层节点数 7、输入层节点数 10、输出层节点数 1、学习率 0.77、动量因子 0.95、误差精度 0.05、最大训练次数 8 000;而 SVR 模型采用网格搜索法得到的参数值 $(C, \epsilon, \sigma) = (1.9269, 1.6342, 10.1723)$ 。

以平均相对误差的绝对值以及合格率^[27]来衡量各模型的预报性能,平均相对误差的计算公式为:

$$MRE = \frac{1}{n} \sum_{i=1}^n |y_c(i) - y_o(i)| / y_o(i) \quad (15)$$

式中: MRE 即为平均相对误差;其他参数含义同式(14)。

三种模型的训练集拟合效果和检验集预测效果如图 2 所示。

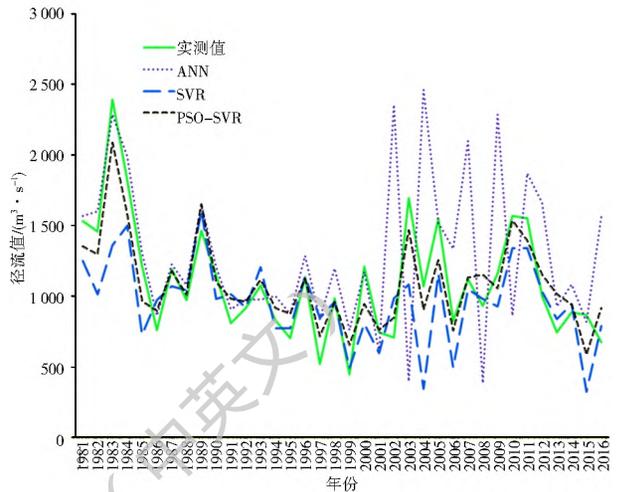


图 2 三种模型预测值与实测值的对比

Fig. 2 Comparison between predictions of the three models and the measured values

三种模型的平均相对误差和合格率如表 2 所示。

表 2 三种模型的平均相对误差和合格率对比

Tab. 2 Comparison of the average relative error and pass rate of the three models

模型	训练集			检验集		
	平均相对误差	合格年数 / 总年数	合格率 (%)	平均相对误差	合格年数 / 总年数	合格率 (%)
ANN	0.1614	16/21	76.19	0.7084	2/15	13.33
SVR	0.1914	13/21	61.9	0.2467	9/15	60
PSO-SVR	0.1365	17/21	80.95	0.1650	12/15	80

由表 2 可以看出,ANN 模型的拟合效果较好,但是检验效果较差,模型的可信度不高;SVR 模型的稳定性能较好,且效果也不错,但其误差仍然较大,难以满足实际预报要求;PSO-SVR 模型整体误差较小,合格率较高,同时加入 PSO 算法实现了模型参数的自动寻优确定,也使得 PSO-SVR 模型的通用性更好,具有一定的应用价值。

4 结论与展望

ANN 模型在进行丹江口年径流预报时,存在拟合效果好、检验效果差,模型稳定性不高等问题,不适用于丹江口水库的年径流预报。在选用 SVR 模型进行预报时,采用网格搜索法进行参数的选择,

存在计算量大、容易陷入局部最优等问题,使得模型参数的选择并非最优值,所以存在拟合、预测误差较大,合格率不高等问题。针对以上模型存在的缺点,本文选择 SVR 模型作为预报模型,并且考虑到参数 C, ε, σ 的选择存在计算量大、选取困难等问题,加之 PSO 算法在参数率定方面的特有优势,将其加入到 SVR 模型中,建立 PSO-SVR 模型,实现了模型参数的自动快速优选。经过多次试验表明,PSO-SVR 模型稳定性高、误差较小,适用于丹江口的年平均入库径流预报。

随着人类活动的加剧,影响径流过程的因素不仅仅局限于太阳活动、大气环流等自然要素,还应包括下垫面改变、水库操作等人工要素,本次研究并未引入人类活动作为预报因子,如何揭示人类活动对径流的影响机理,并在预报中合理的引入下垫面改变、水库操作等预报因子,是作者今后研究的重点。

参考文献(References):

- [1] 沈悦,李阳.南水北调工程水资源生态补偿研究——以丹江口水库为例[J].电网与清洁能源,2016,32(1):119-124.(SHEN Y, LI Y. Study on ecological compensation of water resources in South to North Water Diversion Project- Taking Danjiangkou Reservoir as an example[J]. Power Grid & Clean Energy, 2016, 32(1): 119-124. (in Chinese)) DOI: 10. 3969/j. issn. 1674-3814. 2016. 01. 021.
- [2] 陈刚.水库移民安置补偿政策绩效评价——以湖北丹江口水库为实证[D].华中农业大学,2016.(CHEN G. Performance evaluation of reservoir resettlement compensation policy - Taking Danjiangkou Reservoir as an empirical[D]. Huazhong Agricultural University, 2016. (in Chinese))
- [3] MELOTT A. New fuzzy neural network Markov model and application in mid to long-term runoff forecast[J]. Hydrological Sciences Journal, 2016, 54(8): 150527103244004. DOI: <https://doi.org/10.1080/02626667.2014.986486>.
- [4] 李人姝.基于人工神经网络的荆江河段水文研究[D].广西师范大学,2016.(LI R S. Three level hydrological research of Jingjiang reach based on artificial neural network[D]. Guangxi Normal University, 2016. (in Chinese))
- [5] 赵红标,吴义斌.基于支持向量机的中长期入库径流预报[J].黑龙江大学学报,2009,36(3):55-59.(ZHAO H B, WU Y B. Measurement of runoff from medium and long term based on support vector machine[J]. Journal of Heilongjiang University Engineering, 2009, 36(3): 55-59. (in Chinese)) DOI: 10. 3969/j. issn. 2095-008X. 2009. 03. 001.
- [6] Vapnik V. The Nature of Statistical Learning Theory[M]. New York: Springer, 1995. DOI: 10. 1007/978-1-4757-2440-0.
- [7] Yoo K H, Ju H B, Man G N, et al. Prediction of golden time using SVR for recovering SIS under severe accidents[J]. Annals of Nuclear Energy, 2016, 94: 102-108. DOI: <https://doi.org/10.1016/j.anucene.2016.02.029>.
- [8] 王春燕,夏乐天,孙毓蔓.基于不同核函数的 SVM 用于径流预报的比较[J].人民黄河,2010,32(9):35-36.(WANG C Y, XIA L T, SUN Y M. Comparison of SVM for Runoff Forecast Based on Different Kernel Function[J]. People's Yellow River, 2010, 32(9): 35-36. (in Chinese)) DOI: 10. 3969/j. issn. 1000-1379. 2010. 09. 014.
- [9] 崔东文.人工鱼群算法与支持向量机在枯水期月径流预测中的应用[J].人民珠江,2015,36(3):36-40.(CUI D W. Application of Artificial Fish School Algorithm and Support Vector Machine to Prediction of Monthly Runoff in Dry Season[J]. Penjiang, 2015, 36(3): 36-40. (in Chinese)) DOI: 10. 3969/j. issn. 1001-9235. 2015. 03. 010.
- [10] 郝继升.基于LS-SVR的回归曲线建模[J].江西科学,2007,25(5):563-564.(HAN J S. Study on regression curve modeling based on LS-SVR[J]. Jiangxi Science, 2007, 25(5): 563-564. (in Chinese)) DOI: 10. 3969/j. issn. 1001-3679. 2007. 05. 015.
- [11] 王秋云.对偶原理及其应用[J].吉安师专学报,1990(6):14-18.(WANG Q Y. Polcular Principle and Its Application[J]. Journal of Ji'an Teachers College, 1990(6): 14-18. (in Chinese))
- [12] 曹健,孙世宇,段修生,等.基于KKT条件的SVM增量学习算法[J].火力与指挥控制,2014(7):139-143.(CAO J, SUN S Y, DUAN X S, et al. Advanced SVM incremental learning algorithm based on KKT condition[J]. Fire and Command & Control, 2014(7): 139-143. (in Chinese))
- [13] 董春曦,饶鲜,杨绍全.基于重复训练提高SVM识别率的算法[J].系统工程与电子技术,2003,25(10):1292-1294.(DONG C X, RAO X, YANG S Q. An Algorithm for Improving SVM Recognition Rate Based on Repetitive Training[J]. Systems engineering and electronic technology, 2003, 25(10): 1292-1294. (in Chinese)) DOI: 10. 3321/j. issn: 1001-506X. 2003. 10. 034.
- [14] 黄啸.支持向量机核函数的研究[D].苏州大学,2008.(Huang X. Study on Support Vector Machine Kernel Function[D]. Suhua University, 2008. (in Chinese)) DOI: 10. 7666/d. y1304880.
- [15] 徐莹.基于智能算法的水电站(群)中长期预报调度建模及求解研究[D].大连理工大学,2014.(XU Y, Based on the intelligent algorithm of hydropower station (group) long term forecast scheduling modeling and solving research[D]. Dalian University of Technology, 2014. (in Chinese))
- [16] 李智录,张真真.支持向量机在大坝渗流监测中的应用[J].大坝与安全,2008(1):21-24.(LI Z L, ZHANG Z Z. Application of Support Vector Machine in Dam Seepage Monitoring[J]. Journal of Dam and Safety, 2008(1): 21-24. (in Chinese)) DOI: 10. 3969/j. issn. 1671-1092. 2008. 01. 007.
- [17] 朱永生,张优云.支持向量机分类器中几个问题的研究[J].计算机工程与应用,2003,39(13):36-38.(ZHU Y S, ZHANG Y Y. Study on Several Problems in Support Vector Machine Classifier[J]. Computer Engineering and Applications, 2003, 39(13): 36-38. (in Chinese)) DOI: 10. 3321/j. issn: 1002-8331. 2003. 13. 012.

- [18] 侯伟真, 潘美芹. 高斯核支持向量机最优模型参数选择搜索算法[C]// 中国运筹学会学术交流会, 2006. (HOU W Z, PAN M Q. Gaussian kernel support vector machine optimal model parameter selection search algorithm[C] // China Institute of Operations Academic Exchange, 2006. (in Chinese))
- [19] BAO Y, HAYASHIDA M, AKUTSU T. LBSizeCleave: improved support vector machine (SVM)-based prediction of Dicer cleavage sites using loop/bulge length[J]. *Bmc Bioinformatics*, 2016, 17(1): 487. DOI: 10.1186/s12859-016-1353-6.
- [20] 郝云霄, 闫楚良, 刘克格. 基于支持向量机的机翼载荷模型研究[J]. *科学技术与工程*, 2013, 13(25): 7432-7437. (HAO Y X, YAN C L, LIU K G. Research on wing load model based on support vector machine[J]. *Science Technology and Engineering*, 2013, 13(25): 7432-7437. (in Chinese)) DOI: 10.3969/j.issn.1671-1815.2013.25.028.
- [21] 徐莹, 王嘉阳, 苏华英. 基于遗传算法的支持向量机在径流中长期预报中的应用[J]. *水利与建筑工程学报*, 2014(5): 42-45. (XU Y, WANG J Y, SU H Y. Application of Support Vector Machine Based on Genetic Algorithm in Long term Forecast of Runoff[J]. *Journal of Water Resources and Architectural Engineering*, 2014(5): 42-45. (in Chinese)) DOI: 10.3969/j.issn.1672-1144.2014.05.008.
- [22] 高博, 卢辉斌. 改进型粒子蚁群算法的应用研究[J]. *计算机安全*, 2010(11): 11-13. (GAO B, LU H B. Application of Improved Particle Ant Colony Algorithm[J]. *Computer Security*, 2010(11): 11-13. (in Chinese)) DOI: 10.3969/j.issn.1671-0428.2010.11.004.
- [23] 宫玉琳. 永磁同步电动机伺服系统自适应逆控制策略研究[D]. 长春理工大学, 2013. (GONG Y L. Study on adaptive inverse control strategy of permanent magnet synchronous motor servo system[D]. Changchun University of Science and Technology, 2013. (in Chinese))
- [24] 郭海晋, 金蓉玲. 丹江口水库上游水资源现状及变化趋势分析[J]. *资源科学*, 1997, 24(1): 28-34. (GUO H J, JIN R L. status and dynamics of water upstream of Danjiangkou Reservoir[J]. *Resources Science*, 1997, 24(1): 28-34. (in Chinese))
- [25] 顾思南, 杨修群. 北半球绕极涡的变异及其与我国气候异常的关系[J]. *气象科学*, 2006, 26(2): 135-142. (GU S N, YANG X Q. group of northern circumpolar vortex variation and its relationship with Climate Anomalies in China[J]. *Meteorological Sciences*, 2006, 26(2): 135-142. (in Chinese)) DOI: 10.3969/j.issn.1009-0827.2006.02.003.
- [26] 张作一, 王瑞荣, 王建中, 等. 基于前馈神经网络的潮汐预报[J]. *杭州电子科技大学学报*, 2010, 30(4): 17-21. (ZHANG Z Y, WANG R R, WANG J Z, et al. Tide prediction based on feedforward neural network[J]. *Journal of Hangzhou Dianzi University*, 2010, 30(4): 17-21. (in Chinese)) DOI: 10.3969/j.issn.1001-9146.2010.04.005.
- [27] SL 250-2000, 水文情报预报规范[S]. 北京: 中国水利水电出版社, 2000. (SL 250-2000, Hydrological Information Specification[S]. Beijing: China Water Resources and Hydropower Press, 2000. (in Chinese))

(上接第 37 页)

- [17] 冯增帅, 武桂芝, 李宁宁, 等. 基于季节性河流大沽河河床砂样的侧向非饱和入渗试验研究[J]. *青岛理工大学学报*, 2017, 38(3): 79-83. (FENG Z S, WU G Z, LI N N, et al. Experimental study on lateral infiltration based on the seasonal Daguer riverbed sands[J]. *Journal of Qingdao Technological University*, 2017, 38(3): 79-83. (in Chinese))
- [18] 陈瑞方. 不同计算方法河道水面线差异分析[J]. *水利规划与设计*, 2015(12): 46-48. (CHEN R F. Analysis of river surface water difference by different calculation methods[J]. *Water Resources Planning and Design*, 2015(12): 46-48. (in Chinese)) DOI: 10.3969/j.issn.1672-2469.2015.12.017.
- [19] 贺五洲. 水力学实验[M]. 北京: 清华大学出版社, 2004. (HE W Z. Hydraulic experiment[M]. Beijing: Tsinghua University press, 2004. (in Chinese))
- [20] 汪德耀. 计算水力学理论与应用[M]. 北京: 科学出版社, 2011. (WANG D J. Calculation theory and application of hydraulics[M]. Beijing: Science Press, 2011. (in Chinese))
- [21] 周雪漪. 计算水力学[M]. 北京: 清华大学出版社, 1995. (ZHOU X Y. Computational hydraulics[M]. Tsinghua University Press, 1995. (in Chinese))
- [22] 杨晓佳, 葛永斌. 求解一维扩散方程的一种高精度紧致差分方法[J]. *郑州大学学报(理学版)*, 2016, 48(1): 10-16. (YANG X J, GE Y B. A high resolution compact difference method for solving one dimensional diffusion equations[J]. *Journal of Zhengzhou University (Science Edition)*, 2016, 48(1): 10-16. (in Chinese)) DOI: 10.3969/j.issn.1671-6841.201506013.
- [23] 顾峰峰, 倪汉根. 四点时空偏心隐格式的改进求解[J]. *大连理工大学学报*, 2007, 47(3): 419-423. (GU F F, NI H G. Two improved calculation methods of Preissmann four point linear implicit scheme[J]. *Journal of Dalian University of Technology*, 2007, 47(3): 419-423. (in Chinese)) DOI: 10.3321/j.issn:1000-8608.2007.03.022.
- [24] 宋霞, 姜世强, 盛茂刚. 青岛市大沽河流域洪水库河联合优化调度系统开发总体构思[J]. *山东水利*, 2008(9): 13-15. (SONG X, JIANG S Q, SHEN M G. Overall design of flood control reservoir and river regulation system of Daguer basin in Qingdao city[J]. *Shandong Water Resources*, 2008(9): 13-15. (in Chinese)) DOI: 10.3969/j.issn.1009-6159.2008.09.005.
- [25] 崔素芳. 变化环境下大沽河流域地表水-地下水联合模拟与预测[D]. 济南: 山东师范大学, 2015. (CUI S F. The joint simulation and prediction of surface water and groundwater in the Daguer River basin under the changing environment[D]. Jinan: Shandong Normal University, 2015. (in Chinese))