

基于混合核 SHTS - SVM 的年径流预测

周有荣¹, 崔东文²

(1. 临沧润汀水资源科技服务有限公司, 云南 临沧 677000; 2. 云南省文山州水务局, 云南 文山 663000)

摘要: 为提高年径流预测精度, 利用同热传递搜索(SHTS)算法对混合核支持向量机(SVM)关键参数和混合权重系数进行优化, 提出混合核 SHTS - SVM 年径流预测模型。通过 6 个不同维度的标准测试函数对 SHTS 算法进行仿真验证, 并与当前寻优效果较好的教学优化(TLBO)算法、灰狼优化(GWO)算法进行对比验证。利用两个年径流预测算例对混合核 SHTS - SVM 模型进行实例验证, 并与多项式核 SHTS - SVM、高斯核 SHTS - SVM 及 SHTS - BP 模型预测结果进行对比。结果表明: SHTS 算法寻优精度优于 TLBO、GWO 优化算法, 具有较好的极值寻优能力和稳健性能。混合核 SHTS - SVM 模型综合了多项式全局核函数和高斯局部核函数二者优点, 在预测精度、泛化能力等方面均优于对比模型, 具有较好的实际应用价值。

关键词: 径流预测; 同热传递搜索算法; 混合核函数; 支持向量机; 参数优化; 仿真验证

中图分类号: P338

文献标识码: A

文章编号: 1672-643X(2019)03-0066-07

Annual runoff prediction based on mixed kernel simultaneous heat transfer search - support vector machine

ZHOU Yourong¹, CUI Dongwen²

(1. Lincang Runtong Water Resources Technology Services Limited, Lincang 677000, China;

2. Wenshan Water Bureau of Yunnan Province, Wenshan 663000, China)

Abstract: In order to improve the accuracy of annual runoff prediction, the simultaneous heat transfer search (SHTS) algorithm was used to optimize the key parameters and the mixed weight coefficients of hybrid nuclear support vector machine (SVM). A mixed kernel SHTS - SVM annual runoff prediction model was proposed. The SHTS algorithm was verified by six standard test functions at different dimensions, and verified with the teaching optimization (TLBO) algorithm and gray wolf optimization (GWO) algorithm. Two of the annual runoff prediction examples were used to verify the mixed kernel SHTS - SVM model and compared with the prediction results of polynomial kernel SHTS - SVM, Gaussian kernel SHTS - SVM and SHTS - BP models. The results showed that the optimization accuracy of SHTS algorithm is better than the TLBO and GWO optimization algorithms, and it has better extreme value searching ability and robust performance. The mixed kernel SHTS - SVM model combines the advantages of the polynomial global kernel function and the Gaussian local kernel function. It is superior to the comparison model in terms of prediction accuracy, generalization ability and so on, and has good practical application value.

Key words: runoff forecasting; simultaneous heat transfer search(SHTS) algorithm; mixed kernels; support vector machine(SVM); parameter optimization; simulation verification

1 研究背景

提高径流预测精度一直是水文预测预报中的热

点和难点。由于河川径流预测的复杂性和不确定性, 传统数理统计等方法难以获得理想的预测效果^[1]。BP^[2-3]、Elman^[4]、RBF^[5-6]、GRNN^[7]等神经

收稿日期: 2018-06-25; 修回日期: 2018-10-12

作者简介: 周有荣(1971-), 男, 云南巍山人, 学士, 高级工程师, 主要从事水资源管理保护、水电开发、水环境评价等研究工作。

通讯作者: 崔东文(1978-), 男, 云南玉溪人, 学士, 教授级高级工程师, 主要从事水资源管理保护及智能算法在水文水资源系统中的应用等研究工作。

网络因其较好的自学习能力和函数逼近能力,在水文预测预报中得到广泛应用。但 BP 等神经网络基于大样本、经验风险最小以及易陷入局部极值等不足限制了它在水文中的应用^[8]。研究表明,径流的形成和变化过程异常复杂,仅依靠单一的常规或非正规方法进行径流预测,往往难以达到理想的预测效果和精度,选择合理的模型与方法,建立有效的预测模型是提高水文预测预报精度的关键。

支持向量机(Support Vector Machines, SVM)基于结构风险最小化原则,能有效避免维数灾、过拟合等问题,具有高容错性、智能化和自学习等优点,已成功应用于水文预测预报^[9-11]。但在实际应用中, SVM 存在两方面的困难和不足:(1) SVM 核函数等关键因子选取困难。常规试错法、经验法等选取 SVM 核函数等参数的方法已不能满足 SVM 应用需求。目前,除遗传算法^[12](GA)等传统智能算法用于优化 SVM 参数外,人工鱼群算法^[13](AFSA)、果蝇优化算法^[14](FOA)、布谷鸟搜寻(CS)算法^[15]、灰狼优化(GWO)算法^[10]、文化算法^[10](CA)、SCE - UA 算法^[10]、花授粉算法^[10](FPA)、混合蛙跳算法^[16](SFLA)、入侵杂草优化(IWO)算法^[16]、帝国竞争算法^[16](ICA)、生物地理学优化(BBO)算法^[16]等群体智能优化算法被尝试用于 SVM 关键参数的选取,并获得一定的应用效果,但存在智能算法仿真对比验证的不足。(2)单一核函数制约 SVM 性能问题。核函数的选取是提高 SVM 预测精度的关键,每种核函数都有其优势和不足。高斯核函数属典型的局部性核函数,其局部学习能力强,但泛化能力弱;多项式核函数属典型的全局性核函数,其泛化能力强,但局部学习能力弱^[17]。目前, SVM 在水文预测预报中普遍采用高斯核函数,通过构建混合核函数, SVM 的预测精度尚有进一步提升的空间^[18]。鉴于核函数的选取和核函数参数优化对于改善 SVM 性能的关键性作用,为进一步提高年径流预测精度,本文利用一种新型智能算法——同热传递搜索(Simultaneous Heat Transfer Search, SHTS)算法优化混合核 SVM 关键参数和混合权重系数,提出混合核 SHTS - SVM 年径流预测模型。通过 6 个不同维度的标准测试函数对 SHTS 算法进行仿真验证,再与当前寻优效果较好的教学优化(TLBO)算法、GWO 算法进行对比验证,并通过两个年径流预测算例对混合核 SHTS - SVM 模型进行实例验证,并与多项式核 SHTS - SVM、高斯核 SHTS - SVM 及 SHTS - BP 模型预测结果进行对比,旨在验证混合核 SHTS - SVM 模型用于年径流预测的可行性和有效性。

2 同热传递搜索算法及仿真验证

2.1 同热传递搜索算法

热量传递搜索(Heat Transfer Search, HTS)算法是 Patel 等^[19]于 2015 年提出的种群迭代随机搜索算法,该算法灵感来源于热量传递原理,通过传导、对流和辐射 3 种热量传递方式与周围环境系统相互作用来实现热平衡,其每次迭代包括传导、对流、辐射 3 个阶段之一,即在 0 和 1 之间生成一个随机数 R ,如果 R 小于 1/3,则算法执行传导算子;如果 R 大于 1/3 且小于 2/3,则执行对流算子;如果 R 大于 2/3,则执行辐射算子, HTS 算法通过不断迭代直至获得问题最优解,算法描述参见文献^[19]。

SHTS 算法是一种基于 HTS 改进的变体算法,该算法具有较低的计算复杂度和全局极值寻优能力,在解决高维、复杂优化问题时具有竞争性。在 SHTS 中,每次迭代中将种群随机分成 3 组,每个组分别为热量传递搜索算法的 3 个传热阶段之一,所有进行任何传热模式的个体都将产生一个潜在的新个体,该新个体通过适应度值评判选择接受或丢弃,且在每次迭代过程中传导、对流和辐射是并行执行的,从而大大减少了计算时间^[20-21]。

参考文献^[20-21], SHTS 算法简述如下:

(1) 种群划分。随机将种群 N 分成 3 组,并将其分配到传导、对流和辐射 3 种传热模式。令 X_D 为传导阶段的个体集合, X_V 为对流阶段的个体集合, X_R 为辐射阶段的个体集合,令 $N_s = N/3$ 。

(2) 传导阶段。利用式 $X_D^N = X_D - r_1 X_D$ 生成新个体 X_D^N , r_1 为 $[0, 1]$ 之间随机数矩阵,大小为 $N_s \times D$ 。

(3) 对流阶段。利用式 $c = r_2 X_{best} R$ 确定最佳 c 值。式中: r_2 为 $[0, 1]$ 之间的随机矩阵,其大小为 $1 \times D$; R 为大小为 $1 \times D$ 的 $[0, 1]$ 之间的随机矩阵,值为 1 或 2。并生成一个矩阵 C ,使得 C 的大小为 $N_s \times D$; 利用式 $X_V^N = X_V + C$ 生成新个体 X_V^N 。

(4) 辐射阶段。从辐射个体集合 X_R 中随机将 X_R 分成两个种群大小相等的组 $X_{R'}$ 和 $X_{R''}$ 。利用公式(1)、(2)生成新的个体 $X_{R'}^N$ 和 $X_{R''}^N$, 利用 X_D^N 、 X_V^N 、 $X_{R'}^N$ 和 $X_{R''}^N$ 生成种群 X_{new} , 并利用公式(3)、(4)将新个体约束在搜索区域。

$$X_{R'}^N = X_{R'} - r_3 (X_{R'} - X_{R'}^P) \quad (1)$$

$$X_{R''}^N = X_{R''} - r_3 (X_{R''} - X_{R''}^P) \quad (2)$$

$$X_{new} = \max(X_{new}, L) \quad (3)$$

$$X_{new} = \min(X_{new}, U) \quad (4)$$

式中: $X_R^p, X_{R'}^p$ 分别为新个体 X_R 和 $X_{R'}$ 的合作伙伴; r_3 为 $[0, 1]$ 之间的随机矩阵, 其大小为 $N/6 \times D$; U, L 分别为搜索空间的上、下限值。

2.2 仿真实验

为验证 SHTS 算法寻优能力, 对 6 个标准测试函数在 5 维、30 维和 100 维条件下进行仿真实验 (见表 1), 并与当前寻优效果较好的 TLBO 算法和 GWO 算法的仿真结果对比, 见表 2。

为确保验证的公平性, 本文所有验证均在 Inter (R) Core (TM) i7 - 6500U CPU @ 2.5GHz 4GB 内存的机器上实现, 软件运行环境为 Matlab2010b。实验参数设置: SHTS 算法、TLBO 算法和 GWO 算法最大迭代次数 $G_{\max} = 200$, 种群规模 $N = 100$, 其他参数采用各种算法默认值。为避免寻优效果的偶然性, 并证明算法的稳健性, 采取 3 种算法均对测试函数寻优 20 次, 利用 20 次寻优结果的平均值、标准差对 3 种算法的优化性能进行评估。

SHTS 算法在实验中的性能分析如下:

(1) 对于单峰函数 Sphere、Sumsquares、Quadric, 无论是低维还是高维, SHTS 算法寻优精度均优于 TLBO 算法、GWO 算法 20 个数量级以上, 表明 SHTS 算法具

有较快的收敛速度和寻优精度。对于多峰函数 Griewank、Rastrigin, 无论是低维还是高维, SHTS 算法 20 次寻优均获得理论最优值 0, 寻优精度优于 TLBO 算法、GWO 算法 (除 TLBO 对函数 Griewank 30 维、100 维寻优外); 对于 Ackley 函数, SHTS 算法在 5 维、30 维和 100 维实验条件下, 其寻优结果相同, 20 次寻优结果均为 8.88×10^{-16} , 标准差均为 1.97×10^{-31} , 寻优精度同样优于 TLBO 算法、GWO 算法, 表明 SHTS 算法具有较好的全局极值寻优能力和跳出局部极值能力。

(2) 对于单峰函数, 随着维度的增加, 3 种算法的寻优精度均有所下降, 尤其是 TLBO 算法、GWO 算法, 当维度达到 100 维时, 其寻优精度低于同维度的 SHTS 算法 44 个数量级以上; 对于多峰函数, 无论是低维还是高维, SHTS 算法寻优结果和标准差均相同, 表明该算法的稳定性较好。对于 TLBO 和 GWO 算法, 除 Griewank 函数外, 其寻优精度随着维度的增加下降明显。

可见, 不论是单峰函数还是多峰函数, 抑或低维或高维函数, SHTS 算法的寻优精度、稳定性均优于 TLBO、GWO 算法, 表明 SHTS 算法具有较好的极值寻优能力和稳健性能。

表 1 基准函数

| 函数名 | 表达式 | 属性 | 维数 | 取值范围 | 理论最优解 |
|------------|---|----|----------|-----------------|--------------|
| Sphere | $f_1(x) = \sum_{i=1}^D x_i^2$ | 单峰 | 5/30/100 | $[-100, 100]$ | $\min f = 0$ |
| Sumsquares | $f_2(x) = \sum_{i=1}^D ix_i^2$ | 单峰 | 5/30/100 | $[-100, 100]$ | $\min f = 0$ |
| Quadric | $f_3(x) = \sum_{i=1}^D (\sum_{k=1}^i x_k)^2$ | 单峰 | 5/30/100 | $[-10, 10]$ | $\min f = 0$ |
| Griewank | $f_4(x) = \frac{1}{4000} \sum_{i=1}^D x_i^2 - \prod_{i=1}^D \cos(\frac{x_i}{\sqrt{i}}) + 1$ | 多峰 | 5/30/100 | $[-600, 600]$ | $\min f = 0$ |
| Rastrigin | $f_5(x) = \sum_{i=1}^D [x_i^2 - 10 \cos(2\pi x_i) + 10]$ | 多峰 | 5/30/100 | $[-5.12, 5.12]$ | $\min f = 0$ |
| Ackley | $f_6(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^D x_i^2}\right) - \exp\left(\frac{1}{D} \sum_{i=1}^D \cos 2\pi x_i\right) + 20 + e$ | 多峰 | 5/30/100 | $[-32, 32]$ | $\min f = 0$ |

3 混合核 SHTS - SVM 预测模型

3.1 混合核 SVM

SVM 是基于核函数原理将低维空间回归问题映射到高维特征空间进行求解, 算法原理见文献 [9 - 12]。

为进一步提高 SVM 预测精度, 综合多项式全局核函数和高斯局部核函数二者优势, 弥补二者在应用上的不足 [18, 22 - 24], 本文采用这两个函数的混合, 基于 libsvm 工具箱构造一种满足 Mercer 条件的混

合函数, 表达式为:

$$K_{\min} = \rho K_{poly} + (1 - \rho) K_{RBF} \quad (5)$$

式中: $K_{poly} = [g(x \cdot x_i) + 1]^3$ 三次多项式核函数; $K_{RBF} = \exp(-g \|x - x_i\|^2)$, $g > 0$, 表示 RBF 核函数; ρ 为权重系数。

研究表明, 惩罚因子 C 、核函数参数 g 和不敏感系数 ε 的合理选取决定着混合核 SVM 性能 [10]。 C 取值过小则易导致网络欠拟合, 训练样本误差大; 取值过大则网络过拟合, 导致网络泛化能力差; g 取值小

则拟合误差小,但过小的 g 值会导致模型过拟合; ε 值用于控制模型的预测能力, ε 值小易导致模型欠拟合, ε 值大则易导致模型过拟合; ρ 值决定单一核函数在混合核函数中所占的比重^[18]。

表 2 函数优化对比结果

| 函数名 | 算法 | 维度 | 平均值 | 标准差 | 函数名 | 算法 | 维度 | 平均值 | 标准差 |
|------------|------|-----|------------------------|------------------------|-----------|------|-----|------------------------|------------------------|
| Sphere | SHTS | 5 | 1.02×10^{-94} | 1.79×10^{-94} | Griewank | SHTS | 5 | 0 | 0 |
| | | 30 | 2.12×10^{-73} | 4.42×10^{-73} | | | 30 | 0 | 0 |
| | | 100 | 1.20×10^{-68} | 2.88×10^{-68} | | | 100 | 0 | 0 |
| | TLBO | 5 | 1.35×10^{-51} | 1.43×10^{-51} | | TLBO | 5 | 3.83×10^{-2} | 1.29×10^{-2} |
| | | 30 | 1.91×10^{-32} | 1.11×10^{-32} | | | 30 | 0 | 0 |
| | | 100 | 5.07×10^{-27} | 2.98×10^{-27} | | | 100 | 0 | 0 |
| | GWO | 5 | 7.03×10^{-61} | 1.72×10^{-60} | | GWO | 5 | 3.67×10^{-2} | 2.62×10^{-2} |
| | | 30 | 2.40×10^{-14} | 1.97×10^{-14} | | | 30 | 4.68×10^{-3} | 7.45×10^{-3} |
| | | 100 | 7.65×10^{-5} | 3.67×10^{-5} | | | 100 | 4.38×10^{-5} | 2.18×10^{-5} |
| Sumsquares | SHTS | 5 | 8.73×10^{-91} | 2.13×10^{-90} | Rastrigin | SHTS | 5 | 0 | 0 |
| | | 30 | 1.71×10^{-73} | 3.93×10^{-73} | | | 30 | 0 | 0 |
| | | 100 | 2.24×10^{-67} | 4.69×10^{-67} | | | 100 | 0 | 0 |
| | TLBO | 5 | 7.13×10^{-49} | 1.12×10^{-48} | | TLBO | 5 | 6.29×10^{-5} | 1.47×10^4 |
| | | 30 | 1.21×10^{-30} | 4.56×10^{-31} | | | 30 | 2.88×10^{-2} | 4.24×10^{-2} |
| | | 100 | 8.34×10^{-26} | 3.98×10^{-26} | | | 100 | 0 | 0 |
| | GWO | 5 | 1.77×10^{-31} | 4.13×10^{-31} | | GWO | 5 | 6.35×10^{-1} | 1.3696 |
| | | 30 | 5.17×10^{-13} | 6.37×10^{-13} | | | 30 | 8.8749 | 3.8846 |
| | | 100 | 2.22×10^{-2} | 6.79×10^{-3} | | | 100 | 5.07×10 | 1.79×10 |
| Quadric | SHTS | 5 | 5.93×10^{-65} | 1.12×10^{-64} | Ackley | SHTS | 5 | 8.88×10^{-16} | 1.97×10^{-31} |
| | | 30 | 5.77×10^{-50} | 1.41×10^{-49} | | | 30 | 8.88×10^{-16} | 1.97×10^{-31} |
| | | 100 | 4.26×10^{-43} | 1.02×10^{-42} | | | 100 | 8.88×10^{-16} | 1.97×10^{-31} |
| | TLBO | 5 | 5.33×10^{-32} | 5.14×10^{-32} | | TLBO | 5 | 1.69×10^{-15} | 1.42×10^{-15} |
| | | 30 | 4.10×10^{-3} | 8.50×10^{-3} | | | 30 | 6.86×10^{-5} | 1.63×10^{-15} |
| | | 100 | 3.67×10 | 4.98×10 | | | 100 | 1.37×10^{-2} | 3.34×10^{-2} |
| | GWO | 5 | 7.92×10^{-34} | 1.33×10^{-33} | | GWO | 5 | 3.96×10^{-15} | 1.07×10^{-15} |
| | | 30 | 2.06×10^{-2} | 2.22×10^{-2} | | | 30 | 4.21×10^{-8} | 2.80×10^{-8} |
| | | 100 | 2.04×10^3 | 1.31×10^3 | | | 100 | 1.02×10^{-3} | 2.06×10^{-4} |

3.2 混合核 SHTS - SVM 预测实现步骤

SHTS 算法优化混合核 SVM 关键参数有惩罚因子 C 、核函数参数 g 、不敏感系数 ε 和权重系数 ρ , 其预测实现步骤可归纳如下(多项式核 SHTS - SVM、高斯核 SHTS - SVM 和 SHTS - SBP 模型预测实现步骤可参考实现):

Step1 合理划分各算例训练样本和检验样本。设定惩罚因子 C 等 4 个待优化参数搜寻范围。

Step2 确定适应度函数。本文选用检验样本的平均相对误差绝对值之和作为适应度函数。该适应

度函数描述如下:

$$\begin{cases} \min f(C, g, \varepsilon, \rho) = \left| \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \right| \\ \text{s. t } C \in [C_{\min}, C_{\max}], g \in [g_{\min}, g_{\max}], \\ \varepsilon \in [\varepsilon_{\min}, \varepsilon_{\max}], \rho \in [0, 1] \end{cases} \quad (6)$$

式中: y_i 为第 i 个样本实测值; \hat{y}_i 为第 i 个样本模拟值。

Step3 设置 SHTS 算法种群大小 N 、最大迭代次数 G_{\max} 和变量维度及变量上、下限值;设置当前迭代次数 $t = 0$ 。随机将种群 N 分成 3 组,并将其分配到传

导、对流和辐射3种传热模式。

Step4 基于公式(6)计算种群个体适应度值,并依据适应度值确定种群中最佳适应度值 $f(X)$ 和对应最佳个体 X 。

Step5 利用SHTS算法中传导、对流和辐射3个阶段生成新个体 X_{new} 及适应度值 $f(X_{new})$ 。

Step6 比较 $f(X)$ 和 $f(X_{new})$ 。若 $f(X_{new})$ 优于 $f(X)$,则 X_{new} 代替 X , X_{new} 进入下一次迭代;否则舍去 X_{new} ,保留 X 。

Step7 令 $t = t + 1$,判断算法是否达到最大迭代次数 G_{max} 。若达到,则输出最佳个体 X_{best} 及所对应的个体适应度值 $f(X_{best})$;否则转至Step5,直到满足算法终止条件,算法结束。

Step8 输出最佳个体 X_{best} 及其所对应的个体适应度值 $f(X_{best})$,即待优化问题的最优解及最优适应度值。

Step9 利用SHTS算法优化获得的惩罚因子 C 、核函数参数 g 、不敏感系数 ε 和权重系数 ρ 代入混合核SHTS-SVM模型进行预测。

4 算例

两个算例的参数设置均相同,即SHTS算法最大迭代次数 $T = 200$,种群规模 $N = 100$ 。待优化参数搜索空间设置为:惩罚因子 $C \in [2^{-10}, 2^{10}]$ 、核函数参数 $g \in [2^{-10}, 2^{10}]$ 、不敏感系数 $\varepsilon \in [2^{-10}, 2^{10}]$ 、交叉验证参数 $V = 5$,变量维度 D 为4维,权重系数 ρ 搜索空间为 $[0, 1]$ 。

(1)数据来源。算例1和算例2的数据来源于云南省西洋街水文站和革雷水文站。西洋街水文站设立于1959年1月,位于云南省广南县西洋江干流上,控制径流面积2473 km²。西洋江属珠江流域西江水系,发源于广南县者兔乡那腊村九龙山西麓,于富宁县洞巴出境进入广西田林县与驮娘江汇合,为滇桂省际河流。革雷水文站设立于1970年4月,位于云南省丘北县清水江干流上,控制径流面积3186 km²。清水江发源于文山州砚山县者腊乡老毛山北麓,汇入南盘江,为云南、广西界河。本文利用算例1中1962-2005年、算例2中1971-2005年的实测资料进行预测分析。两个算例年径流与1-10月月均流量的相关性见表3。

从表3可以看出,算例1中,年径流与各月均流量均呈正相关关系,相关性并不十分显著。算例2中,除1月份外,年径流与各月均流量均呈正相关关系,相关性不显著。利用标准化法将算例1和算例

2各径流序列数据处理在 $[0.1, 0.9]$ 之间。

(2)预测评价。对于算例1,选取5-10月月径流作为年径流预测的影响因子,以前30a实测资料为训练样本,后14a资料为检验样本;对于算例2,选取5-9月月径流作为年径流预测的影响因子,以前25a实测资料为训练样本,后10a资料为检验样本。分别基于MatlabR2011b软件环境创建6输入1输出和5输入1输出的年径流预测模型,选取平均相对误差绝对值MRE和最大相对误差绝对值MaxRE两个评价指标对混合核SHTS-SVM等4种模型的拟合、预测结果进行评价,见表4和图1、2。

(3)优化结果。利用SHTS算法对两个算例所有参数进行寻优计算,可以确定算例1混合核SVM惩罚因子 $C = 2^{8.6549}$ 、核函数参数 $g = 2^{-3.4951}$ 、不敏感系数 $\varepsilon = 2^{-3.2893}$ 和权重系数 $\rho = 0.2518$;算例2混合核SVM惩罚因子 $C = 2^{9.9799}$ 、核函数参数 $g = 2^{-3.8242}$ 、不敏感系数 $\varepsilon = 2^{-9.8875}$ 和权重系数 $\rho = 0.7455$ 。

依据表3~4及图1~2可以得出以下结论:

(1)对于算例1,无论是训练样本还是检验样本,混合核SHTS-SVM模型拟合、预测的MRE和MaxRE均优于多项式核SHTS-SVM等3种模型,其检验样本预测精度分别比多项式核SHTS-SVM、高斯核SHTS-SVM和SHTS-BP提高了57.0%、35.5%和24.2%;对于算例2,除训练样本的maxRE略低于多项式核SHTS-SVM模型外,其余训练样本的MRE、检验样本的MRE和maxRE均优于多项式核SHTS-SVM等3种模型,其检验样本预测精度分别比多项式核SHTS-SVM、高斯核SHTS-SVM和SHTS-BP提高了37.7%、26.2%和59.7%。两个算例验证结果表明,混合核SHTS-SVM模型能有效综合多项式全局核函数和高斯局部核函数二者优势,弥补二者在应用上的不足,从而提高混合核SHTS-SVM模型的预测精度和泛化能力;同时验证了本文提出的混合核SHTS-SVM模型用于年径流预测的可行性和有效性。

(2)从表3来看,算例1中年径流与月径流相关性要优于算例2,但预测效果不如算例2,其原因在于算例1中多项式全局核函数和高斯局部核函数二者互补性不如算例2。算例2中虽然年径流与月径流相关性并不显著,其最大相关系数仅为0.546,但多项式全局核函数和高斯局部核函数能很好弥补二者之间的不足,从而提升了混合核SHTS-SVM模型的预测精度和泛化能力。

表 3 两个算例年径流与 1 - 10 月月均流量的相关系数

| 算例 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|------|--------|-------|-------|--------|---------|---------|---------|---------|---------|---------|
| 算例 1 | 0.056 | 0.185 | 0.088 | 0.374* | 0.576** | 0.681** | 0.593** | 0.727** | 0.713** | 0.684** |
| 算例 2 | -0.086 | 0.052 | 0.292 | 0.390* | 0.431** | 0.435** | 0.404** | 0.546** | 0.469** | 0.291 |

表 4 年径流预测结果及其比较

| 算例 | 样本 | 评价指标 | 混合核 | 多项式核 | 高斯核 | SHTS - BP |
|------|-----------------|------------|------------|------------|------------|-----------|
| | | | SHTS - SVM | SHTS - SVM | SHTS - SVM | SHTS - BP |
| 算例 1 | 训练样本 | MRE / % | 4.37 | 7.86 | 7.30 | 5.13 |
| | (1962 - 1991 年) | max RE / % | 12.64 | 16.33 | 18.08 | 14.20 |
| | 检验样本 | MRE / % | 3.29 | 7.66 | 5.10 | 4.34 |
| | (1992 - 2005 年) | max RE / % | 8.09 | 19.89 | 9.32 | 20.85 |
| 算例 2 | 训练样本 | MRE / % | 2.34 | 3.93 | 7.00 | 3.98 |
| | (1971 - 1995 年) | max RE / % | 15.01 | 14.03 | 21.41 | 15.94 |
| | 检验样本 | MRE / % | 2.87 | 4.61 | 3.89 | 7.13 |
| | (1995 - 2005 年) | max RE / % | 10.76 | 14.82 | 12.89 | 19.08 |

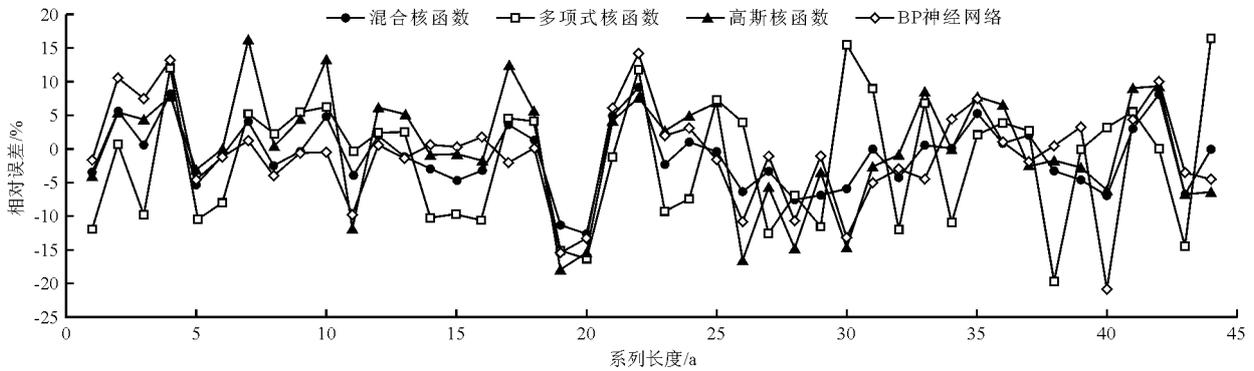


图 1 1962 - 2005 年算例 1 年径流拟合 - 预测相对误差效果图

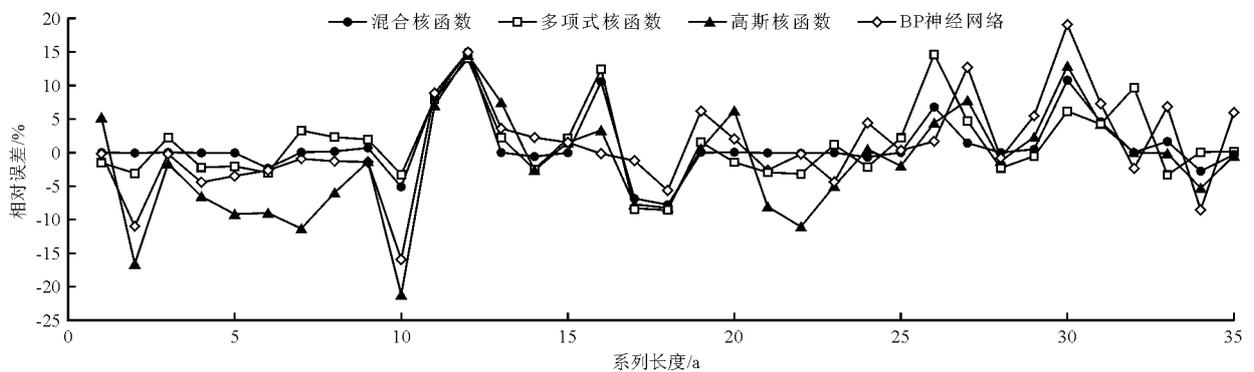


图 2 1971 - 2005 年算例 2 年径流拟合 - 预测相对误差效果图

(3) 从 SHTS 算法对两个算例权重系数的优化结果来看,对于算例 1,权重系数 $\rho = 0.2518$,表明高斯核函数占主导地位;对于算例 2,权重系数 $\rho = 0.7455$,表明多项式核函数占主导地位。

(4) 从图 1、2 及表 4 来看,算例 1 中 4 种模型的拟合、检验精度相差不大,4 种模型均未出现“欠拟

合”与“过拟合”特征。对于算例 2,混合核 SHTS - SVM 模型、多项式核 SHTS - SVM 的拟合、检验精度相差不大,具有较好的预测效果和泛化能力;但高斯核 SHTS - SVM 模型表现出“欠拟合”特征,而 SHTS - BP 模型呈“过拟合”状态。

5 结 论

(1)介绍一种新型智能算法——同热传递搜索(SHTS)算法,通过6个不同维度的典型测试函数对SHTS算法进行仿真验证,并与当前寻优效果较好的TLBO算法、GWO算法作对比,结果表明SHTS算法具有较好的全局极值寻优能力和稳健性能,是一种全新高效的全局优化算法。

(2)构造线性混合核SVM,利用SHTS算法同时优化混合核SVM关键参数和混合权重系数,提出混合核SHTS-SVM年径流预测模型。利用两个算例对混合核SHTS-SVM模型进行实例验证,并与多项式核SHTS-SVM、高斯核SHTS-SVM和SHTS-BP模型的预测结果进行对比,结果表明混合核SHTS-SVM模型的预测精度优于多项式核SHTS-SVM等3种模型,表明利用SHTS算法能有效优化混合核SVM关键参数和混合权重系数;且混合核SHTS-SVM模型能有效综合多项式全局核函数和高斯局部核函数二者优势,弥补二者在应用上的不足,从而提高混合核SHTS-SVM模型的预测精度和泛化能力。

(3)两个算例验证结果表明,混合核SHTS-SVM模型用于水文预测预报是合理可行的,模型具有较好的预测精度和泛化能力,是提高预测精度的有效方法。

参考文献:

- [1] 崔东文. 基于教学优化算法的多元超越回归模型及其在径流预报中的应用[J]. 华北水利水电大学学报(自然科学版), 2016, 37(6): 61-66+71.
- [2] 崔东文. 多隐层BP神经网络模型在径流预测中的应用[J]. 水文, 2013, 33(1): 68-73.
- [3] 郭淳, 李祚泳, 党媛. 基于免疫进化算法的BP网络模型在径流预测中的应用[J]. 水资源保护, 2009, 25(5): 1-4.
- [4] 崔东文. 改进Elman神经网络在径流预测中的应用[J]. 水利水电工程学报, 2013(2): 71-77.
- [5] 黄剑竹. 改进RBF-Adaboost模型及其在年径流预测中的应用[J]. 人民珠江, 2015, 36(1): 32-36.
- [6] 任磊, 岳春芳, 何训江. RBF神经网络模型在金沟河流域径流预测中的应用[J]. 水资源与水工程学报, 2011, 22(1): 94-97.
- [7] 覃光华, 宋克超, 周泽江, 等. 基于WA-GRNN模型的年径流预测[J]. 四川大学学报(工程科学版), 2013, 45(6): 39-46.
- [8] 张拥华, 曾凡仔. 基于混合核支持向量机的金融时间序列分析[J]. 计算机工程与应用, 2008, 44(19): 220-222.
- [9] 崔东文, 金波. 基于改进的回归支持向量机模型及其在年径流预测中的应用[J]. 水力发电学报, 2015, 34(2): 7-14.
- [10] 崔东文. 几种智能算法与支持向量机融合模型在中长期月径流预测中的应用[J]. 华北水利水电大学学报(自然科学版), 2016, 37(5): 51-57.
- [11] 李月玉, 李磊. 免疫粒子群算法与支持向量机在枯水期月径流预测中的应用[J]. 水资源与水工程学报, 2015, 26(3): 124-128+135.
- [12] 代兴兰. 遗传算法与最小二乘支持向量机在年径流预测中的应用[J]. 水资源与水工程学报, 2014, 25(6): 231-235.
- [13] 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.
- [14] 崔东文, 金波. 果蝇优化算法与支持向量机在年径流预测中的应用[J]. 人民珠江, 2015, 36(2): 29-32.
- [15] 赖锦辉, 梁松. 一种新的基于GCS-SVM的网络流量预测模型[J]. 计算机工程与应用, 2013, 49(21): 75-78.
- [16] 崔东文, 郑斌. 几种智能优化算法与支持向量机相融合的月径流预测模型及应用[J]. 人民珠江, 2016, 37(3): 18-25.
- [17] 周金明, 王传玉, 何帮强. 基于混合核函数FOA-LSSVM的预测模型[J]. 计算机工程与应用, 2015, 51(4): 133-137.
- [18] 唐奇, 王红瑞, 许新宜, 等. 基于混合核函数SVM水文时序模型及其应用[J]. 系统工程理论与实践, 2014, 34(2): 521-529.
- [19] PATEL V K, SAVSANI V J. Heat transfer search (HTS): a novel optimization algorithm[J]. Information Sciences, 2015, 324: 217-246.
- [20] MAHARANA D, KOTTECHA P. Simultaneous heat transfer search for computationally expensive numerical optimization[C]// Evolutionary Computation. IEEE, 2016.
- [21] MAHARANA D, KOTTECHA P. Simultaneous Heat Transfer Search for single objective real-parameter numerical optimization problem[C]//2016 IEEE Region 10 Conference (TENCON). IEEE, 2016: 2138-2141.
- [22] 刘伟, 王建平, 刘长虹, 等. 基于粒子群寻优的支持向量机番茄红素含量预测[J]. 农业机械学报, 2012, 43(4): 143-147+155.
- [23] 周建友, 张凯威. 改进布谷鸟算法优化混合核LSSVM的卷烟销售量预测[J]. 计算机工程与应用, 2015, 51(19): 250-254.
- [24] 王新华, 桂卫华, 王雅琳, 等. 混合核函数支持向量机的磨矿粒度预测模型[J]. 计算机工程与应用, 2010, 46(12): 207-209+214.
- [25] 耿俊豹, 孙林凯, 陈是学. 支持向量机的混合核函数参数优选方法[J]. 计算机应用, 2013, 33(5): 1321-1323+1356.