

# 基于 MK-SVM 和时序特征分析的月径流预报模型

雷庆文<sup>1,2,3</sup>, 闫磊<sup>1,2</sup>, 巫晨煜<sup>1,2</sup>, 罗云<sup>4</sup>, 谢笑添<sup>1,2</sup>

(1. 河北工程大学水利水电学院, 河北邯郸 056038; 2. 河北省智慧水利重点实验室, 河北邯郸 056038; 3. 云河(河南)信息科技有限公司, 河南郑州 450003; 4. 云南大学国际河流与生态安全研究院, 云南昆明 650091)

**摘要:** 针对传统径流预报方法预报因子不确定性和预报模型复杂性问题, 基于月径流时序特征重要性分析选择预报因子, 采用混合核函数支持向量机(MK-SVM)模型捕捉径流时序间的非线性关系, 提出动态透镜成像反向学习和 Lévy 飞行等多策略融合的改进灰狼优化算法(IGWO), 并构建了径流预报的 IGWO-MK-SVM 模型。黑河流域莺落峡水文站月径流预报结果表明: IGWO-MK-SVM 模型月径流预报结果的纳什效率系数、均方根误差、Kling-Gupta 效率系数分别为 0.8942、16.9099 m<sup>3</sup>/s 和 0.8639; 与传统 SVM 模型相比, IGWO-MK-SVM 模型在径流预报中的自适应性有所提升, 相较于长短期记忆网络模型和季节性差分自回归移动平均模型, IGWO-MK-SVM 模型能更好地预报月径流的真实变化过程。

**关键词:** 径流预报; 随机森林; 径流预报因子; 混合核函数支持向量机; 改进灰狼优化算法; 黑河流域中图分类号: P338 文献标志码: A 文章编号: 1004-6933(2024)06-0148-07

**Monthly runoff prediction model based on MK-SVM and time series feature analysis**//LEI Qingwen<sup>1,2,3</sup>, YAN Lei<sup>1,2</sup>, WU Chenyu<sup>1,2</sup>, LUO Yun<sup>4</sup>, XIE Xiaotian<sup>1,2</sup>(1. College of Water Conservancy and Hydropower, Hebei University of Engineering, Handan 056038, China; 2. Hebei Key Laboratory of Intelligent Water Conservancy, Handan 056038, China; 3. Yunhe (Henan) Information Technology Co., Ltd., Zhengzhou 450003, China; 4. Institute of International Rivers and Eco-security, Yunnan University, Kunming 650091, China)

**Abstract:** To address the problem of uncertainty of prediction factors and model complexity of traditional runoff prediction methods, prediction factors were selected based on feature importance analysis of monthly runoff time series, and the nonlinear relationship between runoff time series was captured by the mixed kernel function-support vector machine (MK-SVM) model. An improved grey wolf optimizer (IGWO) that integrated multiple strategies, such as dynamic lens imaging reverse learning and Lévy flying strategies, was proposed to enhance the stability of the global parameter optimization of the MK-SVM model, and an IGWO-MK-SVM model for runoff prediction was constructed. The results of monthly runoff prediction at Yingluoxia Hydrological Station in the Heihe River Basin show that the Nash-Sutcliffe efficiency coefficient, root mean squared error, and Kling-Gupta efficiency coefficient of prediction results of the IGWO-MK-SVM model were 0.8942, 16.9099 m<sup>3</sup>/s, and 0.8639, respectively. Compared with the traditional SVM model, the IGWO-MK-SVM model has high adaptability in runoff prediction, and compared with the long short-term memory network model and the seasonal autoregressive integrated moving average model, the IGWO-MK-SVM model can better predict the real change process of monthly runoff.

**Key words:** runoff prediction; random forest; runoff prediction factors; mixed kernel function-support vector machine; improved grey wolf optimizer algorithm; Heihe River Basin

水资源管理过程中的优化配置和调度、水环境保护、防洪抗旱等问题均需要对未来一定时期的径流变化进行预测<sup>[1]</sup>。然而,受全球环境变化和人类活动的影响,径流时序的随机性和非平稳性愈发显著,使得径流预报面临全新挑战<sup>[2-3]</sup>。目前,大量研

究针对数字信号处理、机器学习模型与水文预报中的难点问题,建立了一系列新的预报模型<sup>[4-5]</sup>。基于数据驱动的机器学习模型通常无须依赖径流变化的物理过程,这在一定程度上克服了传统水文模型的不足<sup>[6]</sup>。机器学习径流预报模型通常分为浅层学

基金项目:国家自然科学基金项目(51909053);水利部京津冀水安全保障重点实验室开放研究基金项目(IWHR-KLWS-202305)

作者简介:雷庆文(1994—),男,硕士,主要从事智慧水利和水资源系统分析研究。E-mail:15738519012@163.com

通信作者:闫磊(1990—),男,副教授,博士,主要从事水文分析计算研究。E-mail:yanl@whu.edu.cn

习和深度学习模型。近年来,众多开源神经网络框架的出现,极大地推动了深度学习径流预报模型的发展。熊怡等<sup>[7]</sup>提出利用长短期记忆网络(long short-term memory, LSTM)对径流自适应变分模态分解的每个子序列进行预测,再将子序列的预测结果集成,得到径流的最终预测结果,建立的模型对捕捉多尺度变异的复杂非线性径流序列信息有很好的效果;Xu 等<sup>[8]</sup>针对水文学知识从数据丰富流域转移到数据稀缺流域存在的问题,提出基于 Transformer 的迁移学习框架,通过使用数据丰富流域信息准确预测数据稀缺目标流域的洪水。尽管深度学习模型相较于浅层学习模型的非线性拟合能力更强,但存在数据需求量大、模型结构复杂、参数不确定性强、模型易过拟合的问题<sup>[9]</sup>。

在样本数据少、输入特征有限且明确的情况下,基于智能优化算法的浅层学习模型通常具有更好的径流预测效果<sup>[10]</sup>。Sharifi 等<sup>[11]</sup>在对 Amameh 流域径流预测研究中,发现支持向量机(support vector machine, SVM)模型相较于人工神经网络、局部线性回归、自适应神经模糊推理系统表现更好。SVM 模型的泛化性和准确性非常依赖于核函数,为克服单一核函数的不足,唐奇等<sup>[12]</sup>采用混合核函数支持向量机(mixed kernel function-support vector machine, MK-SVM)模型对武山站和南河川站的月径流进行预测,预测结果明显优于单一核函数的 SVM 模型。基于高性能智能优化算法对 SVM 模型参数调优是提高径流预测精度的另一途径, Samantaray 等<sup>[13]</sup>尝试将樽海鞘群优化算法(salp swarm algorithm, SSA)用于 SVM 模型参数调优,提出了 SSA-SVM 模型,显著提高了传统 SVM 模型的鲁棒性,使得模型具有更好的径流预测精度。但是,浅层学习模型对输入特征信息较为敏感,常需提前确定合适的模型参数和输入变量<sup>[14]</sup>。

本文采用随机森林(random forest, RF)模型分析径流时序特征的重要性,并依据特征重要性顺序选取 MK-SVM 模型的输入变量,提出动态透镜成像反向学习和 Lévy 飞行等多策略融合的改进灰狼优化算法(improved grey wolf optimizer, IGWO)对 MK-SVM 模型参数进行优化,建立了径流预报的 IGWO-MK-SVM 模型。采用黑河流域莺落峡水文站实测径流对模型进行验证,以期在变化环境下探索强健性径流预报模型提供新的思路。

## 1 研究方法

### 1.1 预报因子选择

预报因子的选择是影响数据驱动模型效果的关

键因素之一,但目前没有通用的预报因子选择方法。现有研究中多采用穷举法、偏互信息法、偏自相关函数(partial autocorrelation function, PACF)法等<sup>[5]</sup>,然而,这些方法通常无法描述每个因子对最终结果影响的重要程度。为此,本文采用特征重要性分析(feature importance analysis, FIA)法选择预报因子, FIA 法基于 RF 模型对备选因子集进行分析,并依据因子特征重要性顺序,从中选取使模型预报效果最好的特征作为预报因子<sup>[15]</sup>。

### 1.2 MK-SVM 模型

SVM 模型是一种可有效解决多变量非线性回归问题的机器学习模型,主要通过核函数将输入变量映射到高维特征空间,然后在映射后的特征空间中进行线性回归。不同核函数对 SVM 模型的非线性拟合效果影响很大,局部性核函数的学习能力强,但泛化能力相对较弱,全局性核函数则恰好与之相反。本文基于 Mercer 核理论,结合全局多项式核函数和局部高斯径向基核函数构建 MK-SVM 模型,以提高模型在径流预报中的自适应能力。

$$\kappa_{\text{MK}} = \lambda \kappa_p + (1 - \lambda) \kappa_r \quad (1)$$

式中: $\kappa_{\text{MK}}$  为混合核函数; $\lambda$  为单一核函数在混合核函数中的占比, $\lambda \in [0, 1]$ ; $\kappa_p$ 、 $\kappa_r$  分别为多项式核函数和高斯径向基核函数。

### 1.3 多策略融合的 IGWO 算法

灰狼优化算法(grey wolf optimizer, GWO)依据社会等级将狼群分为  $\alpha$  狼、 $\beta$  狼、 $\delta$  狼和  $\omega$  狼,并通过模拟狼群捕食过程中的跟踪追捕、包围和攻击,实现对目标函数的搜索<sup>[16]</sup>。GWO 算法和其他群体智能优化算法类似,在狼群多样性不足时,存在易陷入局部最优、收敛速度慢和求解精度低的问题。因此,本文引入 Tent-Logistic-Cosine 混沌映射初始化狼群,并提出动态透镜成像反向学习和 Lévy 飞行等多策略融合的 IGWO 算法,以提高对机器学习径流预报模型参数的调优能力。

**a.** Tent-Logistic-Cosine 混沌映射。为克服随机生成策略初始化种群多样性差的问题,本文将 Tent、Logistic 和 Cosine 映射结合,构成 Tent-Logistic-Cosine 混沌系统,可有效克服低维系统的不足,相比高维系统,其复杂度更低,更易实现<sup>[17]</sup>。

**b.** 动态透镜成像反向学习策略。基于群体智能算法优化的 MK-SVM 模型,非常容易陷入局部最优解。为此,本文采用动态透镜成像反向学习策略,使算法在陷入非主导核函数区域时,仍能保持对主导核函数区域的探索,从而促进算法跳出局部最优解。透镜成像反向学习策略基本原理<sup>[18]</sup>为:可行解  $X_\alpha$  ( $\alpha$  狼位置)总存在与之相对应的反向解  $X_\alpha^*$ ,若反

向解  $X_\alpha^*$  的适应度优于可行解  $X_\alpha$ , 则取反向解  $X_\alpha^*$  为当前种群的最优解, 同时更新为  $\alpha$  狼的位置。本文基于透镜成像原理, 采用动态调焦策略(式(2))计算反向学习点的位置(式(3))。

$$1/D_u + 1/D_v = 1/(rf_{\max}) \quad (2)$$

$$X_\alpha^* = \begin{cases} 0.5(B_u + B_l) - D_v & X_\alpha > X_0 \\ 0.5(B_u + B_l) + D_v & X_\alpha \leq X_0 \end{cases} \quad (3)$$

式中:  $D_u$  为物距, 表示可行解  $X_\alpha$  到透镜焦点的距离, 本文选择搜索空间的中心位置  $X_0$  作为焦点;  $r$  为随机数,  $r \in (0, 1]$ ;  $f_{\max}$  为将像控制在约束区间的最大可调焦距;  $D_v$  为像距;  $B_u$ 、 $B_l$  分别为搜索空间的上边界和下边界。

c. Lévy 飞行策略。动态透镜成像反向学习能在一定程度上增强 IGWO 算法跳出局部最优解的能力, 却无法有效地实现狼群从探索阶段到开发阶段的过渡。本文通过引入 Lévy 飞行策略, 结合长距离探索和短距离细致搜索, 从而有效平衡探索和开发的关系<sup>[19]</sup>。

### 1.4 IGWO-MK-SVM 模型

MK-SVM 模型相较于单一核函数的 SVM 模型, 其结构更复杂, 非线性拟合能力更强, 但对模型参数的变化会更为敏感。为此, 采用 IGWO 算法对 MK-SVM 模型的 5 个超参数(多项式核函数参数  $\gamma_p$ 、高斯径向基核函数参数  $\gamma_r$ 、单一核函数在混合核函数中的占比  $\lambda$ 、惩罚因子  $C$ 、不敏感系数  $\varepsilon$ ) 进行优化,

构建了径流预报的 IGWO-MK-SVM 模型, 计算流程如图 1 所示。

## 2 实例验证

为验证所提出的 IGWO-MK-SVM 模型在实际月径流预报中的应用效果, 收集了黑河流域莺落峡水文站 1960—2017 年共 696 月实测径流数据进行分析。莺落峡水文站位于黑河流域上游, 控制集水面积 10018 km<sup>2</sup>, 地表水主要来源于祁连山区的冰雪融化和大气降水, 基流主要由地下水供应, 且大部分水资源在春季和夏季集中。将莺落峡水文站 1960—2007 年月径流数据作为训练集, 2008—2017 年月径流数据作为验证集。

### 2.1 模型输入变量选取

基于径流时序分析的预报模型是通过建立当前预报对象和其滞后值之间的相关关系, 在时间上进行外推, 从而实现对象未来发展趋势的预测。为评估 FIA 法相较于传统 PACF 法在径流预报因子选择中的特点, 分别采用 FIA 法和 PACF 法从备选因子集  $\{x_\tau | \tau = 1, 2, \dots, L\}$  中选择模型的输入变量, 其中  $x_\tau$  为前  $\tau$  月的径流,  $L$  为最大时滞阶数。为选择合适的  $x_\tau$  作为模型输入变量, 分别考虑  $L=12$  月和  $L=24$  月两种情形, 并以此讨论周期性对月径流预报精度的影响。

FIA 法选择径流预报因子时, 采用基于序列模

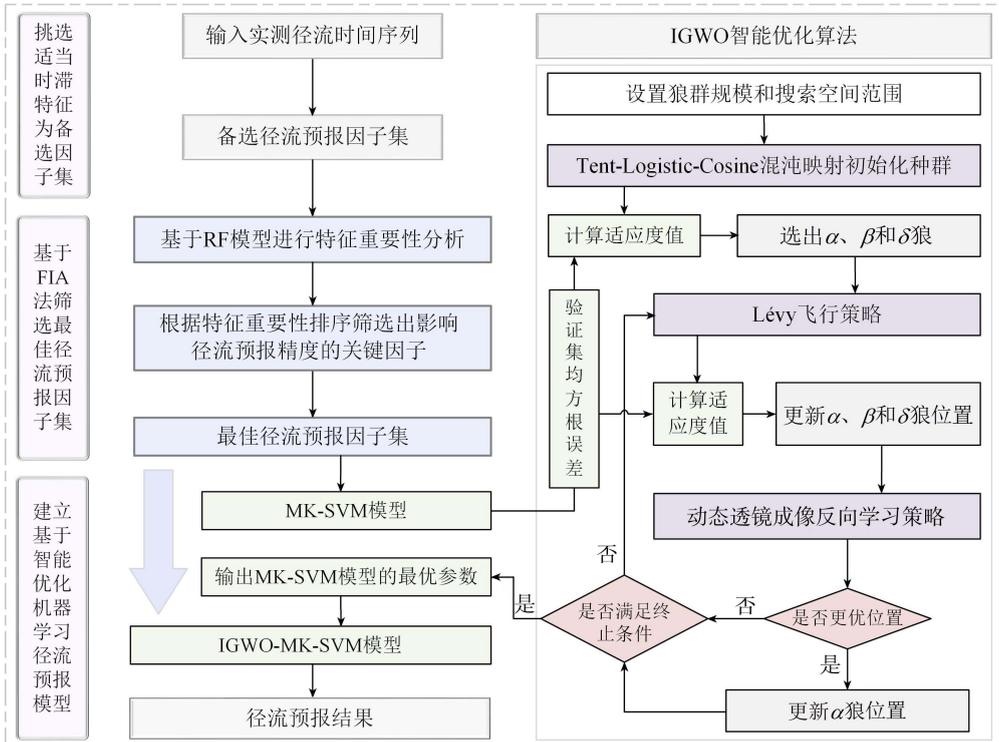


图 1 IGWO-MK-SVM 模型计算流程

Fig. 1 Calculation process of IGWO-MK-SVM model

型的算法配置 (sequential model-based algorithm configuration, SMAC) 对 RF 模型进行调优<sup>[20]</sup>, 搜索最佳输入变量, 优化目标为变量个数最少且交叉验证准确率最高的一组变量集合。PACF 法选择径流预报因子时, 将偏自相关系数通过置信水平为 95% 检验的变量视为有效输入变量<sup>[5]</sup>。由 FIA 法和 PACF 法得到的模型输入变量如表 1 所示, 其中, FIA 法所得模型输入变量依照各因子特征重要性程度排列。

表 1 模型输入变量

Table 1 Model input variables

预报因子选择方法	备选因子集	
	L=12月	L=24月
FIA 法	$x_{12}, x_1, x_9, x_6, x_{11}, x_7$	$x_{12}, x_{24}, x_1, x_{18}, x_9, x_{23}, x_6, x_8$
PACF 法	$x_1 \sim x_8, x_{10} \sim x_{12}$	$x_1 \sim x_8, x_{10} \sim x_{12}, x_{15}, x_{23}, x_{24}$

## 2.2 模型参数优化

MK-SVM 模型相较于单一核函数的 SVM 模型, 自适应能力更强, 但对超参数的选择更为敏感, 不同超参数组合对模型鲁棒性的影响较为明显。针对这些问题, 本文采用 IGWO 算法对 MK-SVM 模型超参数进行优化。为进一步验证 IGWO 算法的全局优化性能, 分别选用粒子群优化算法 (particle swarm optimization, PSO)、SSA、GWO 与 IGWO 算法进行对比。

考虑不同算法的差异性, 各算法参数依据最大化算法性能和相近收敛时间进行配置。为避免随机性差异, 保证评价的客观性, 各算法独立运行 50 次, 对 MK-SVM 模型的超参数搜索空间范围为:  $C \in [2^{-10}, 2^{10}]$ 、 $\gamma_r \in [2^{-10}, 2^{10}]$ 、 $\gamma_p \in [2^{-10}, 2^{10}]$ 、 $\varepsilon \in [0, 1]$ 、 $\lambda \in [0, 1]$ 。采用验证期径流预报结果的均方根误差 (root mean squared error, RMSE) 作为适应度

值, 统计各算法运行 50 次所得 RMSE 的最优值、最劣值、平均值、标准差, 综合评价算法优化效果, 并将与最优 RMSE 偏差较大的解视为局部最优解, 用其出现的频次表征算法跳出局部最优解的能力 (表 2)。

从表 2 可以看出: ①FIA 法相较于传统的 PACF 法所选输入变量更为合理, 其对变量间的非线性关系识别更高效, 且可有效剔除冗余变量对模型的干扰。然而, 由于 FIA 法主要识别的是变量间更为复杂的非线性关系, 通常需要结合更好性能的优化算法才能取得理想效果。②IGWO 算法对 MK-SVM 模型的优化效果最佳, 能最大限度地搜索到保证模型效果最优的超参数, 并且融合多种改进策略, 能有效帮助算法跳出局部最优解和挖掘全局最优解信息, 寻优精度得到较大提高。③最大时滞阶数为 24 月时, 所筛选出的因子能更好地反映径流的实际变化规律, 由于月径流具有较为明显的年周期性, 进行统计分析时, 常需采用季节性差分方式才能将其转化为平稳时间序列。④以往研究多采用高斯径向基核函数  $\kappa_r$  作为 SVM 模型的核函数<sup>[21]</sup>, 而本文所得最优解的  $\lambda = 0.93$ , 说明多项式核函数  $\kappa_p$  为主导性核函数, 可见混合核函数相较于单一核函数更有优势, 由此进一步表明 IGWO-MK-SVM 模型相较于传统 SVM 模型具有更好的自适应性。

## 2.3 结果与分析

为了验证本文提出的 IGWO-MK-SVM 模型对径流预报的有效性, 以 1 月为预见期, 选用 L=24 月的备选因子集, 利用 FIA 法筛选出的预报因子集作为模型输入变量, 并对比了季节性差分自回归移动平均 (seasonal autoregressive integrated moving average,

表 2 各算法对 MK-SVM 模型优化效果评价

Table 2 Evaluation of optimization effects of various algorithms on MK-SVM model

预报因子选择方法	L/月	优化算法	最优值/( $m^3/s$ )	最劣值/( $m^3/s$ )	平均值/( $m^3/s$ )	标准差/( $m^3/s$ )	局部最优频次
FIA 法	12	PSO	17.9012	18.8542	18.2985	0.1611	3
		SSA	18.1677	18.6238	18.3537	0.1113	0
		GWO	18.0575	18.8471	18.3062	0.1672	4
		IGWO	17.9012	18.8468	18.2450	0.1187	1
	24	PSO	16.9100	17.4413	16.9614	0.1470	5
		SSA	16.9127	17.4409	16.9937	0.1270	3
		GWO	16.9383	17.3984	16.9665	0.1070	3
		IGWO	16.9099	16.9112	16.9102	0.0002	0
PACF 法	12	PSO	19.1355	19.4241	19.2022	0.1074	5
		SSA	19.1357	19.4016	19.2101	0.0959	5
		GWO	19.1356	19.3871	19.1915	0.0816	3
		IGWO	19.1355	19.1422	19.1382	0.0028	0
	24	PSO	17.6693	18.3219	17.6947	0.0929	1
		SSA	17.6701	17.9978	17.7024	0.0684	2
		GWO	17.7115	17.9158	17.7166	0.0288	1
		IGWO	17.6693	17.6704	17.6701	0.0003	0

SARIMA) 模型<sup>[22]</sup>、LSTM 模型<sup>[7]</sup>、RF 模型<sup>[14]</sup>、SSA-SVM 模型<sup>[13]</sup>、SVM 模型在验证期的月径流预报效果(表 3)。采用纳什效率系数(Nash-Sutcliffe efficiency coefficient, NSE)、RMSE 和 Kling-Gupta 效率系数(Kling-Gupta efficiency coefficient, KGE)作为径流预报准确性的衡量指标<sup>[23]</sup>。

表 3 验证期不同模型径流预报效果对比

Table 3 Comparison of runoff prediction effect of different models in verification period

模型	NSE	RMSE/(m <sup>3</sup> /s)	KGE
SARIMA	0.8803	17.9923	0.7720
RF	0.8857	17.5764	0.8402
LSTM	0.8565	19.6960	0.8538
SVM	0.5646	34.3075	0.0120
SSA-SVM	0.8863	17.5297	0.8291
IGWO-MK-SVM	0.8942	16.9099	0.8639

由表 3 可见,IGWO-MK-SVM 模型径流预测结果各项指标最优。IGWO-MK-SVM 模型与 SARIMA 模型的建模方法类似,SARIMA 模型通过差分将非平稳径流序列平稳化后,根据序列间的短期自相关性建立统计模型;而 IGWO-MK-SVM 模型则通过 FIA 法对预报因子筛选,利用机器学习模型建立映射关系。从结果来看,IGWO-MK-SVM 模型不仅可以摆脱传统统计模型对径流序列平稳性的要求,而且能捕捉到更准确的径流变化规律。相比深度学习模型强大的非线性拟合能力,浅层机器学习模型尽管存在不足,但在样本数据少和输入特征明确的情况下,IGWO-MK-SVM 模型相较于 LSTM 模型可以获得更好的效果。

从验证期的径流过程拟合效果(图 2)来看,未经调优的 SVM 模型表现最差,仅能捕捉到径流最基本的周期性变化规律,而 SSA-SVM 模型和 IGWO-MK-SVM 模型对枯水期和汛期月径流的预报能力均有较为显著的提高,说明未经智能优化的机器学习模型很难在实际应用中取得好的效果。因此,探索具有良好全局优化性能的智能优化算法对基于机器

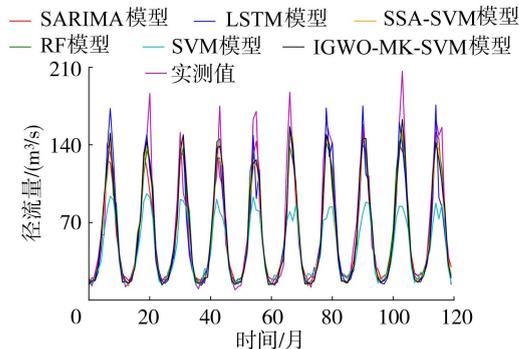


图 2 验证期径流过程拟合效果

Fig. 2 Fitting effects of runoff process in verification period

学习径流预报模型的发展具有重要意义。

从验证期各模型径流预测结果散点图(图 3)来看,SARIMA 模型预测结果的确性系数  $R^2 = 0.901$ ,为所有模型中最优,因为 SARIMA 模型采用季节性差分将非平稳的月径流序列转化为平稳的时间序列,再进行移动平均和自回归<sup>[24]</sup>,使得预测结果与原始径流序列的变化趋势保持了较好的一致性。IGWO-MK-SVM 模型预测结果的  $R^2 = 0.900$ ,也反映出模型具有较好的性能。虽然 SARIMA 模型预测值与实测值的线性趋势较好,但预测值整体偏小,拟合结果偏离  $y=x$  线较为明显。显然,单一评测指标很难全面评价模型的性能,需采用多指标的综合评价体系。

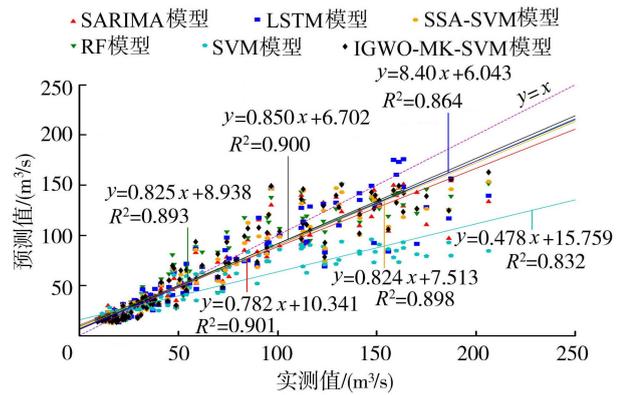


图 3 验证期不同模型径流预测结果散点图

Fig. 3 Scatter diagram of runoff prediction results of different models in verification period

图 4 是验证期各模型径流预测结果的泰勒图,将实测径流序列和模型预测径流序列的相关系数、中心均方根偏差(root mean square deviation, RMSD)和标准差进行整合,能够更直观地表现出预

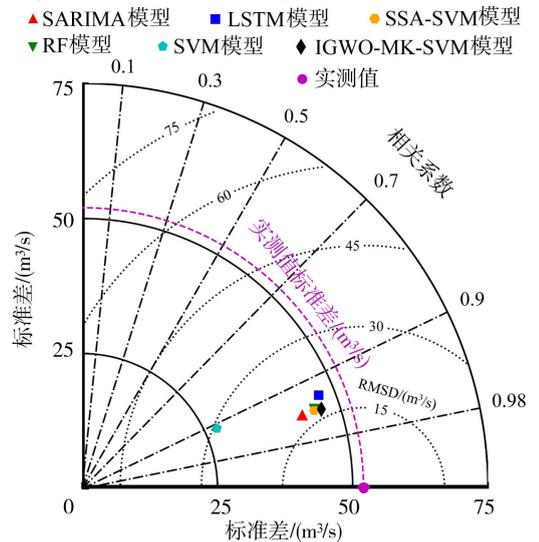


图 4 验证期不同模型径流预测结果泰勒图

Fig. 4 Taylor diagram of runoff prediction results of different models in verification period

测值和实测值间的差异<sup>[25]</sup>。从图 4 可以看出 SSA-SVM 模型和 IGWO-MK-SVM 模型的径流预测结果要更接近实测值。IGWO-MK-SVM 模型相较于 SSA-SVM 模型,采用动态透镜成像反向学习和 Lévy 飞行等多策略融合的 IGWO 算法增强了对 MK-SVM 模型参数的全局搜索能力,且避免了核函数的选择,模型的自适应能力和计算的稳健性均得到提升。

深度学习模型具有良好的学习能力,可用于处理复杂的非线性问题,但也会因为模型过于复杂,将训练数据的细节和噪声视为一般规律,造成过拟合问题,使得模型泛化效果不佳。从以上分析结果来看,相较于深度学习的 LSTM 模型,基于智能优化的 SVM 和 RF 模型效果更好,进一步说明结合智能优化的浅层学习模型在一定条件下,可以达到深度学习模型的效果。

## 2.4 讨论

本文针对径流预报因子选择、智能优化算法性能改进、SVM 模型自适应能力提升 3 方面开展研究。从各模型的预测效果来看,IGWO-MK-SVM 模型对预报精度虽然有一定提高,但对洪水月份的预测结果不够理想,显然仅对原始径流序列进行分析,无法完全捕捉到径流系统多时间尺度的变化规律。月径流序列被视为非线性、非平稳的准周期信号,大量研究通过离散小波变换、模态分解和奇异谱分析将其分解,产生更平稳的子序列作为模型输入,通常可以得到更好的预测效果<sup>[26-27]</sup>。然而,时间序列分解的边界效应问题目前仍然非常棘手<sup>[7]</sup>。可推测,在未来研究中开发基于自适应径流子序列的分解方式,结合 FIA 法选择预报因子,并采用具有良好自适应能力的 IGWO-MK-SVM 模型拟合时序关系,将在一定程度上提高径流预报的准确性。

针对黑河流域莺落峡月径流预报,相较于具有统计意义的 SARIMA 模型、深度学习的 LSTM 模型、浅层学习的 RF 和 SSA-SVM 等模型,本文提出的 IGWO-MK-SVM 模型具有一定优势。但由于莺落峡月径流规律性较强,导致各模型的效果差异并不显著。因此,还需将 IGWO-MK-SVM 模型应用到更多流域的径流预报中,以检验其在不同水文条件下的径流预报效果。

## 3 结论

a. 对比采用 FIA 法和 PACF 法得到的径流预报因子,IGWO-MK-SVM 模型预测结果表明,FIA 法对变量间非线性关系的识别更高效,所选变量更少,可有效剔除冗余变量对模型的干扰。

b. 智能优化算法的性能直接影响机器学习模

型的泛化性,为此,本文提出了基于动态透镜成像反向学习和 Lévy 飞行等多策略融合的 IGWO 算法。各算法对 MK-SVM 模型的优化效果表明,IGWO 算法的全局寻优性能相对较优,IGWO-MK-SVM 模型具有良好的自适应性,对提升径流预报精度具有重要影响。

c. 不同径流时序分析预报模型对黑河流域莺落峡月径流预测结果表明,未经智能优化的 SVM 模型表现最差,其预测结果的 NSE 仅为 0.5646; IGWO-MK-SVM 模型效果有较大幅度的改进,其预测结果的 NSE 提升至 0.8942;相比基于统计理论和深度学习等径流预报模型的精度,IGWO-MK-SVM 模型具有一定的优势。

## 参考文献:

- [1] 张弛,王明君,于冰,等. 松辽流域水资源综合调控研究进展与四大难题探究[J]. 水利学报,2021,52(11): 1379-1388. (ZHANG Chi, WANG Mingjun, YU Bing, et al. Research progress of comprehensive water resources regulation and four major problems in Song-Liao River Basin[J]. Journal of Hydraulic Engineering, 2021, 52(11): 1379-1388. (in Chinese))
- [2] YU Liuyang, ZHAO Xining, GAO Xiaodong, et al. Effect of natural factors and management practices on agricultural water use efficiency under drought; a meta-analysis of global drylands [J]. Journal of Hydrology, 2021, 594: 125977.
- [3] 雷庆文,高培强,李建林. 时序分解和 CNN-LSTM 相融合的月径流预报模型[J]. 长江科学院院报,2023,40(6):49-54. (LEI Qingwen, GAO Peiqiang, LI Jianlin. A monthly runoff forecast model combining time series decomposition and CNN-LSTM[J]. Journal of Changjiang River Scientific Research Institute, 2023, 40(6): 49-54. (in Chinese))
- [4] WANG Wenzhuo, DONG Zengchuan, ZHANG Tianyan, et al. Mixed D-vine copula-based conditional quantile model for stochastic monthly streamflow simulation [J]. Water Science and Engineering, 2024, 17(1): 13-20.
- [5] FENG Zhongkai, NIU Wenjing, TANG Zhengyang, et al. Monthly runoff time series prediction by variational mode decomposition and support vector machine based on quantum-behaved particle swarm optimization [J]. Journal of Hydrology, 2020, 583: 124627.
- [6] ZHAO Xuehua, LYU Hanfang, LYU Shujin, et al. Enhancing robustness of monthly streamflow forecasting model using gated recurrent unit based on improved grey wolf optimizer [J]. Journal of Hydrology, 2021, 601: 126607.
- [7] 熊怡,周建中,孙娜,等. 基于自适应变分模态分解和长

- 短期记忆网络的月径流预报[J]. 水利学报, 2023, 54(2): 172-183. (XIONG Yi, ZHOU Jianzhong, SUN Na, et al. Monthly runoff prediction based on self-adaptive variational mode decomposition and long short-term memory network [J]. Journal of Hydraulic Engineering, 2023, 54(2): 172-183. (in Chinese))
- [ 8 ] XU Yuanhao, LIN Kairong, HU Caihong, et al. Deep transfer learning based on transformer for flood forecasting in data-sparse basins [ J ]. Journal of Hydrology, 2023, 625:129956.
- [ 9 ] YAN Lei, LEI Qingwen, JIANG Cong, et al. Climate-informed monthly runoff prediction model using machine learning and feature importance analysis [ J ]. Frontiers in Environmental Science, 2022, 10: 1049840.
- [ 10 ] NIU Wenjing, FENG Zhongkai, CHENG Chuntian, et al. Forecasting daily runoff by extreme learning machine based on quantum-behaved particle swarm optimization [ J ]. Journal of Hydrologic Engineering, 2018, 23 ( 3 ) : 04018002.
- [ 11 ] SHARIFI A, DINPASHOH Y, MIRABBASI R. Daily runoff prediction using the linear and non-linear models [ J ]. Water Science and Technology, 2017, 76(4) : 793-805.
- [ 12 ] 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. ( TANG Qi, WANG Hongrui, XU Xinyi, et al. Hydrological time series model based on SVM with mixed kernel function and its application [ J ]. Systems Engineering: Theory & Practice, 2014, 34 ( 2 ) : 521-529. ( in Chinese ))
- [ 13 ] SAMANTARAY S, DAS S S, SAHOO A, et al. Monthly runoff prediction at Baitarani River Basin by support vector machine based on salp swarm algorithm [ J ]. Ain Shams Engineering Journal, 2022, 13(5) : 101732.
- [ 14 ] 胡义明, 陈腾, 罗序义, 等. 基于机器学习模型的淮河流域中长期径流预报研究 [ J ]. 地学前缘, 2022, 29(3) : 284-291. ( HU Yiming, CHEN Teng, LUO Xuyi, et al. Medium to long term runoff forecast for the Huai River Basin based on machine learning algorithm [ J ]. Earth Science Frontiers, 2022, 29(3) : 284-291. ( in Chinese ))
- [ 15 ] 姚登举, 杨静, 詹晓娟. 基于随机森林的特征选择算法 [ J ]. 吉林大学学报(工学版), 2014, 44(1) : 137-141. ( YAO Dengju, YANG Jing, ZHAN Xiaojuan. Feature selection algorithm based on random forest [ J ]. Journal of Jilin University ( Engineering and Technology Edition ), 2014, 44(1) : 137-141. ( in Chinese ))
- [ 16 ] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey wolf optimizer [ J ]. Advances in Engineering Software, 2014, 69:46-61.
- [ 17 ] HUA Zhongyun, ZHOU Yicong, HUANG Hejiao. Cosine-transform-based chaotic system for image encryption [ J ]. Information Sciences, 2019, 480:403-419.
- [ 18 ] 喻飞, 李元香, 魏波, 等. 透镜成像反学习策略在粒子群算法中的应用 [ J ]. 电子学报, 2014, 42 ( 2 ) : 230-235. ( YU Fei, LI Yuanxiang, WEI Bo, et al. The application of a novel OBL based on lens imaging principle in PSO [ J ]. Acta Electronica Sinica, 2014, 42 ( 2 ) : 230-235. ( in Chinese ))
- [ 19 ] HEIDARI A A, PAHLAVANI P. An efficient modified grey wolf optimizer with Lévy flight for optimization tasks [ J ]. Applied Soft Computing, 2017, 60: 115-134.
- [ 20 ] KUK E, STOPA J, KUK M, et al. Petroleum reservoir control optimization with the use of the auto-adaptive decision trees [ J ]. Energies, 2021, 14(18) : 5702.
- [ 21 ] HUANG Shengzhi, CHANG Jianxia, HUANG Qiang, et al. Monthly streamflow prediction using modified EMD-based support vector machine [ J ]. Journal of Hydrology, 2014, 511:764-775.
- [ 22 ] VALIPOUR M. Long-term runoff study using SARIMA and ARIMA models in the United States [ J ]. Meteorological Applications, 2015, 22(3) : 592-598.
- [ 23 ] GUPTA H V, KLING H, YILMAZ K K, et al. Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling [ J ]. Journal of Hydrology, 2009, 377(1/2) : 80-91.
- [ 24 ] 练继建, 孙萧仲, 马超, 等. 基于 EEMD-AR 模型的丹江口水库年径流随机模拟与预报 [ J ]. 水利水电科技进展, 2017, 37(5) : 16-21. ( LIAN Jijian, SUN Xiaozhong, MA Chao, et al. Stochastic simulation and prediction of annual runoff in the Danjiangkou Reservoir based on EEMD-AR model [ J ]. Advances in Science and Technology of Water Resources, 2017, 37(5) : 16-21. ( in Chinese ))
- [ 25 ] TAYLOR K E. Summarizing multiple aspects of model performance in a single diagram [ J ]. Journal of Geophysical Research: Atmospheres, 2001, 106 ( D7 ) : 7183-7192.
- [ 26 ] 王秀杰, 乔鸿飞, 曾勇红, 等. 基于 TVF-EMD、GRA 和 LightGBM 的日径流预测组合模型 [ J ]. 水资源保护, 2023, 39(5) : 135-142. ( WANG Xiujie, QIAO Hongfei, ZENG Yonghong, et al. TVF-EMD, GRA, and LightGBM combined daily runoff prediction model [ J ]. Water Resources Protection, 2023, 39 ( 5 ) : 135-142. ( in Chinese ))
- [ 27 ] MENG Erhao, HUANG Shengzhi, HUANG Qiang, et al. A robust method for non-stationary streamflow prediction based on improved EMD-SVM model [ J ]. Journal of Hydrology, 2019, 568:462-478.

(收稿日期:2024-03-06 编辑:施业)