

基于改进樽海鞘群寻优 SVM 的土壤含水量预测算法

赵小强, 杨帆, 晏珠峰

(西安邮电大学, 陕西 西安 710121)

摘要: 针对传统土壤含水量预测算法存在的精度和效率较低等问题, 采用支持向量机 (SVM, support vector machine) 建立预测模型, 提出一种改进樽海鞘群算法 (SSA, salp swarm algorithm) 优化 SVM 的土壤含水量预测算法。首先, 引入反向学习和混沌优化对标准樽海鞘群算法进行改进, 解决算法易陷入局部最优解和收敛速度慢的问题; 其次, 利用改进的樽海鞘群算法对影响 SVM 性能的参数进行优化, 构建对应的预测模型; 最后, 将所提模型与粒子群优化 SVM 预测模型、鲸鱼算法优化 SVM 预测模型进行对比。实验结果表明, 所提模型的均方误差和决定系数分别为 0.42 和 0.901, 与其他两种模型相比性能更优, 验证了所提算法的有效性。

关键词: 土壤含水量预测; 支持向量机; 樽海鞘群算法; 反向学习; 混沌优化

中图分类号: TP181

文献标识码: A

doi: 10.11959/j.issn.2096-3750.2021.00192

Prediction method of soil water content based on SVM optimized by improved salp swarm algorithm

ZHAO Xiaoqiang, YANG Fan, YAN Zhufeng

Xi'an University of Posts and Telecommunications, Xi'an 710121, China

Abstract: Aiming at the problems of low accuracy and low efficiency of traditional soil water content prediction methods, support vector machine (SVM) was used to establish a prediction model, and the soil water content prediction method based on SVM optimized was proposed by the improved salp swarm algorithm. Firstly, the opposition-based learning and chaotic optimization were introduced to improve the standard salp swarm algorithm to solve the problem that the algorithm was easy to fall into the local optimal solution and its convergence speed was slow. Secondly, the improved salp swarm algorithm was used to optimize the parameters that affect the performance of SVM and the corresponding prediction model was built. Finally, the proposed model was compared with the particle swarm optimization SVM and the whale algorithm optimized SVM prediction model. The experimental results show that the mean square error and decision coefficient of the proposed model are 0.42 and 0.901, which are better than the other two models which verified the effectiveness of the proposed method.

Key words: soil water content prediction, support vector machine, salp swarm algorithm, opposition-based learning, chaotic optimization

收稿日期: 2020-04-29; 修回日期: 2021-01-15

通信作者: 杨帆, 2513199608@qq.com

基金项目: 国家自然科学基金委员会—雅砻江联合基金资助项目 (No.U1965102); 陕西省创新人才推进计划—科技创新团队 (No.2019TD-28); 西安市科技计划项目 (No.201806117YF05NC13-2); 陕西省教育厅产业化培育项目 (No.18JF029); 陕西省国际合作计划项目 (No.2018KW-025)

Foundation Items: The Yalong River Joint Funds of the National Natural Science Foundation of China (No.U1965102), The Shaanxi Innovative Talent Promotion Plan-Science and Technology Innovation Team (No.2019TD-28), The Science and Technology Projects of Xi'an (No.201806117YF05NC13-2), The Industrialization Cultivation Project of Shaanxi Provincial Department of Education (No.18JF029), The Shaanxi International Cooperation Project (No.2018KW-025)

1 引言

随着我国现代化建设的不断发展以及人口的快速增长,水资源短缺的问题日益凸显。农业作为我国的基础产业,其用水需求量巨大。近年来,物联网技术的广泛应用为农业的发展注入动力。农业物联网高度集成和综合运用多种新一代信息技术解决农田信息采集、面向不同应用的智慧决策及管理难题,逐步实现农业生产精细化和智慧化。土壤含水量是陆地与大气能量交换的重要因素,在农业领域具有重要的作用^[1],提高土壤含水量预测水平是实现节水灌溉、提高水资源利用率与作物产能、促进农业可持续发展的关键环节^[2]。同时,科学有效的土壤含水量预测也会反作用于农田物联网监测系统,促使系统的决策功能更完善。因此,不断提升土壤含水量预测水平是研究学者关注的课题之一。

目前,常用的土壤含水量预测算法包括经验法、土壤水量平衡法、时间序列法、神经网络法以及支持向量机等。Lewin 等^[3]提出一种基于土壤含水量与贮水量变化关系的经验预测模型。王玮^[4]选取作物需水量、有效降水量以及作物地下水利用量等参数作为输入数据,利用土壤水量平衡模型对土壤的含水量进行动态预测并验证了模型的精度。白冬妹等^[5]利用土壤含水量时间序列建立自回归预测模型,实验结果表明,此模型能有效预测土壤含水量。安小宇等^[6]对标准樽海鞘群算法进行改进,并利用改进后的算法优化 BP (back propagation) 神经网络,实验结果表明,所提方法提高了预测精度。Prakash 等^[7]基于 SVM 模型并利用 3 个不同的数据集对未来 1 天、2 天、7 天的土壤含水量分别进行预测,并成功验证了模型的有效性。经验法、土壤水量平衡法和时间序列法的应用前提都是其对应的物理量关系确定,因此应用范围受限;BP 神经网络则存在易陷入局部极值点和收敛速度慢等问题;虽然 SVM 应用广泛,但其性能受参数取值的影响。为了弥补 SVM 预测的不足,本文提出利用群智能算法对其参数进行优化。

本文提出一种基于改进樽海鞘群算法优化 SVM 的土壤含水量预测算法,针对标准樽海鞘群算法的不足,利用两种方法进行改进,并将改进后的算法用于 SVM 参数的优化。以陕西省渭南市的物联网示范基地的监测数据为对象进行实验,并使用不同预测模型进行对比,验证了所提方法的有效

性,以期后续可以指导农田物联网系统进行更科学合理地节水灌溉。

2 樽海鞘群算法

2.1 标准樽海鞘群算法

樽海鞘是一种透明桶状体、与水母高度相似的海洋生物,种群个体首尾相连构成一条链,依次跟随移动。樽海鞘群形成的链由领导者和追随者两部分组成,其中领导者位于链首,它的环境判断力是种群中最好的,除了领导者之外的个体均为追随者。Mirjalili 等^[8]受樽海鞘群体行为的启发,于 2017 年首次提出樽海鞘群算法。

假设樽海鞘种群个体数为 N ,求解问题的维度为 D ,则樽海鞘在 D 维空间的位置可写为 $X_n = [(X_{n1}, X_{n2}, \dots, X_{nD})]^T, n = 1, 2, \dots, N$,在捕食空间 $N \times D$ 中存在食物 $F = [F_1, F_2, \dots, F_D]^T$ 。搜索区域的上限表示为 $\mathbf{ub} = [\mathbf{ub}_1, \mathbf{ub}_2, \dots, \mathbf{ub}_D]$,下限可表示为 $\mathbf{lb} = [\mathbf{lb}_1, \mathbf{lb}_2, \dots, \mathbf{lb}_D]$ 。式(1)描述了种群初始化过程为

$$X_{N \times D} = \text{rand}(N, D) \times (\mathbf{ub} - \mathbf{lb}) + \mathbf{lb} \quad (1)$$

在樽海鞘种群中,领导者与追随者的每一维状态分别可以表示为 X_d^l 和 X_d^m ,其中 $d = 1, 2, \dots, D$ 为维数, $m = 2, 3, \dots, N$ 为追随者序号。领导者负责在搜索空间中寻找食物,引导樽海鞘群进行移动,更新计算式如式(2)所示。

$$X_d^l = \begin{cases} F_d + c_1((\mathbf{ub}_d - \mathbf{lb}_d)c_2 + \mathbf{lb}_d), & c_3 \geq 0.5 \\ F_d - c_1((\mathbf{ub}_d - \mathbf{lb}_d)c_2 + \mathbf{lb}_d), & c_3 < 0.5 \end{cases} \quad (2)$$

其中, c_2 、 c_3 均为 $0 \sim 1$ 的随机数, c_1 被称作收敛因子,其表达式为 $c_1 = 2e^{-(4l/l_{\max})^2}$,其中, l 是当前的迭代轮数, l_{\max} 是对应最大迭代次数,计算式的选取根据 c_3 的取值而定。

追随者位置更新受自身的初始位置、运动速度和运动时的加速度影响,利用牛顿运动定律可以求得更新计算式为

$$X_d^{m*} = \frac{1}{2}(X_d^{m-1} + X_d^m) \quad (3)$$

其中, X_d^m 表示第 m 个追随者第 d 维的位置, X_d^{m*} 对应其更新后的位置。

2.2 改进樽海鞘群算法

2.2.1 反向学习

反向学习 (OBL, opposition-based learning) 被

Tizhoosh^[9]首次引入并应用于各种优化问题。反向学习的思想是在算法进行每一轮迭代时,同时计算当前解与其反向解,通过对比选取当前解集合和反向解集合中更能促进算法进化的解,以降低算法在迭代寻优时的盲目性。

Wang 等^[10]提出了广义反向学习 (GOBL, generalized opposition-based learning) 策略,该策略引入权重因子控制反向解的取值区间,本文采用此策略求得反向解。假设 $\mathbf{P}=(x_1, x_2, \dots, x_D)$ 是一个 D 维坐标系的点,并且 $x_i \in [a_i, b_i], i=1, 2, \dots, D$, 权重因子 k 是 $0 \sim 1$ 的随机数。则其反向点 $\mathbf{OP}=(ox_1, ox_2, \dots, ox_D)$ 利用式(4)求解, ox_i 是组成 \mathbf{OP} 的元素,通过求解 ox_i 即可得到 \mathbf{OP} 。

$$ox_i = k \cdot (a_i + b_i) - x_i \quad (4)$$

2.2.2 混沌优化

混沌优化方法利用混沌变量替换随机变量,与随机搜索相比,它可以用更快的速度执行搜索工作^[11]。近年来,混沌优化由于其动态行为在优化领域得到广泛认可,其不重复遍历搜索空间的特性有助于优化算法更动态和全局地进行探索。在混沌优化中,可以选择 $[0,1]$ 中的任一数字作为初始值,也可根据具体混沌映射的范围选取初始值。本文选用 Sinusoidal 映射来产生相应的混沌变量,其表达式为

$$c_{k+1} = \mu c_k^2 \sin(\pi c_k) \quad (5)$$

其中, c_k 为随机数, k 为迭代的次数, μ 为混沌系数, $c_k \in [0,1]$ 。当 $\mu=2.3$ 时,系统会成为混沌系统。

2.2.3 基于反向学习的混沌樽海鞘群算法

为了解决原始 SSA 容易陷入局部最优解和收敛速度较慢的问题,在改进后的基于反向学习的混沌樽海鞘群算法 (OCSSA, opposition-based chaotic salp swarm algorithm) 中引入了混沌优化和反向学习两种方法。混沌优化具有不重复遍历搜索区域的特点,避免算法陷入局部最优,反向学习的加入可以使算法的解接近全局最优解的概率更大。OCSSA 的改进包括:对种群初始化阶段种群的位置进行改进;对食物位置更新进行优化处理。

1) 改进的种群初始化

SSA 种群初始化的方法是随机生成樽海鞘群位置,导致算法性能不稳定。若算法初始化种群位置接近全局最优解的位置,则算法的收敛速度以及其获得全局最优解的能力都将得到提升,然而随机生成的初始解的位置大多不理想。基于此,OCSSA

引入了混沌优化,此措施可以在初始化种群时提供更可靠的初始种群位置,从而确保算法的收敛速度不会出现太大的波动,并且能够在一定程度上提高算法的性能。

使用随机方法生成 N 个 D 维樽海鞘组成的种群的方法如式(6)所示。

$$x_{ij} = l_{ij} + k \cdot (u_{ij} - l_{ij}) \quad (6)$$

利用混沌优化方法生成种群的操作可以用式(7)说明。

$$x_{ij} = l_{ij} + ch_{ij} \cdot (u_{ij} - l_{ij}) \quad (7)$$

其中, $x_{ij}=(x_{i1}, \dots, x_{ij}, \dots, x_{iD})$ 表示种群的个体, $i=1, 2, \dots, N, j=1, 2, \dots, D$ 。 $l_{ij}=(l_{i1}, \dots, l_{ij}, \dots, l_{iD})$ 和 $u_{ij}=(u_{i1}, \dots, u_{ij}, \dots, u_{iD})$ 分别代表樽海鞘个体 x_{ij} 的上、下限, ch_{ij} 是使用 Sinusoidal 混沌映射产生的混沌变量, k 为 $0 \sim 1$ 的随机数。利用混沌优化方法生成种群 X_{mic} 并随机初始化生成种群 X_{mir} 后,从 X_{mic} 和 X_{mir} 两个种群中选取更好的个体形成新种群 X_{inew} , 并利用该种群进行反向学习。利用式(4)求解种群 X_{inew} 的反向种群 X_{oi} , 选取 X_{inew} 和 X_{oi} 两个种群中更好的个体并形成最终初始化种群 X_i , 利用该种群开始算法的迭代寻优。

2) 改进的食物位置更新

由式(2)表明,种群中领导者的位置根据食物位置的变化而改变,所以对整个算法来说,食物位置至关重要。但在优化问题中全局最优是未知的,因此在算法中将当前得到的最优位置作为食物位置,并在搜索迭代时利用最优位置不断更新食物位置。食物位置即种群的最佳位置,即使樽海鞘种群位置在搜索过程中变差,也不会影响食物位置,因此对食物位置的处理显得至关重要。

为了优化算法搜索过程中的食物位置,OCSSA 加入了混沌优化方法与反向学习方法。如果食物位置一直不变或者只存在微小的变化,算法就存在陷入局部最优解的可能,OCSSA 采用混沌优化方法来解决领导者陷入局部最优的问题。在算法寻优阶段,通过设置一个阈值控制混沌优化的搜索次数,设置一个食物历史位置记录每一轮寻优食物的位置,并利用式(7)更改食物的位置。当找到更好的食物位置或达到搜索次数上限时,OCSSA 将结束混沌优化并开始执行反向学习阶段的任务。在得到混沌优化后的食物位置后,利用式(4)求解反向的食物

位置,在两个食物位置中选取更优的一者作为某一轮寻优的最终食物位置并代入下一轮迭代继续寻优。

2.2.4 OCSSA 性能验证

为验证 OCSSA 的性能,选取 23 个测试函数^[8]中的 4 个函数 (Sphere 函数、Rastrigin 函数、Griewank 函数、Penalize2 函数)对算法进行检验,其中 Sphere 函数为单峰函数,其余 3 个函数为多峰函数,测试函数的数学表达式及属性如表 1 所示。另外,选取粒子群优化 (PSO, particle swarm optimization)算法^[12]、鲸鱼优化算法 (WOA, whale optimization algorithm)^[13]、SSA 与改进算法进行对比。本文中所有算法的种群数为 30,最大迭代次数为 500。每个基准测试函数独立运行 50 次,最大程度地降低结果的统计误差。实验中的对比算法 WOA

和 SSA 的相关参数值采用两种原算法的参数值,PSO 算法的参数学习因子取值为 1.6,惯性因子取值为 0.8,测试函数的运行结果如表 2 所示,其中加粗的数值为最优值。

从表 2 可以看出,无论是应用于单峰函数还是多峰函数,OCSSA 的收敛精度在 4 种算法中都是最好的,证明了改进算法的全局寻优能力较强。算法收敛速度对比如图 1 所示,从图 1 可得,改进后的 OCSSA 的收敛速度明显快于其他 3 种对比算法的收敛速度。

3 基于 SVM 模型的土壤含水量预测

3.1 SVM

SVM 建立在统计学习理论的基础上,通过使结

表 1 测试函数的数学表达式及属性

函数名称	计算式	范围	最优值
Sphere	$F_1(x) = \sum_{i=1}^n x_i^2$	[-100,100]	0
Rastrigin	$F_9(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10]$	[-5.12,5.12]	0
Griewank	$F_{11}(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	[-600,600]	0
Penalize2	$F_{13}(x) = 0.1 \left\{ \sin^2(3\pi x_1) + \sum_{i=1}^n (x_i - 1)^2 [1 + \sin^2(3\pi x_i + 1)] + (x_n - 1)^2 [1 + \sin^2(2\pi x_n)] \right\} + \sum_{i=1}^n u(x_i, 5, 100, 4)$	[-50,50]	0

表 2 测试函数的运行结果

函数名称	统计值	OCSSA	PSO	SSA	WOA
Sphere	最优值	0	4.13×10^{-6}	4.15×10^{-9}	3.68×10^{-88}
	平均值	0	2.81×10^{-5}	5.69×10^{-9}	9.25×10^{-78}
	标准差	0	2.88×10^{-5}	1.74×10^{-9}	1.81×10^{-77}
Rastrigin	最优值	0	2.60×10^1	2.98×10^1	0
	平均值	0	3.23×10^1	4.76×10^1	5.68×10^{-15}
	标准差	0	5.90	9.96	1.04×10^{-14}
Griewank	最优值	0	3.94×10^{-2}	4.89×10^{-7}	0
	平均值	0	1.83×10^{-1}	6.41×10^{-3}	0
	标准差	0	1.18×10^{-1}	6.33×10^{-3}	0
Penalize2	最优值	6.00×10^{-12}	2.25×10^{-7}	1.10×10^{-2}	4.47×10^{-2}
	平均值	9.00×10^{-12}	2.10×10^{-6}	1.43×10^{-2}	1.81×10^{-1}
	标准差	1.69×10^{-12}	1.09×10^{-6}	6.62×10^{-3}	1.02×10^{-1}

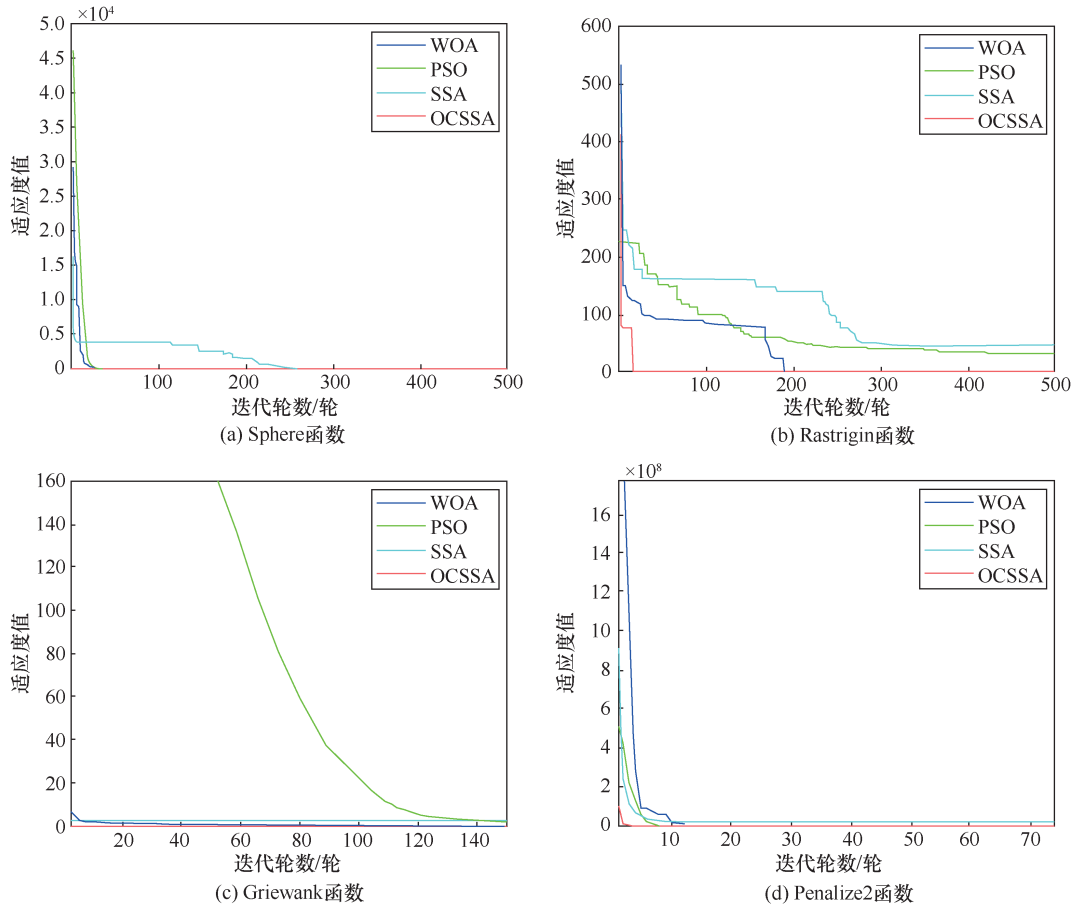


图 1 算法收敛速度对比

构化风险最小提升泛化能力，达到经验风险与置信风险最小化的目标。即使在统计样本量较少的情况下，SVM 仍然可以得到理想的统计规律，本文采用支持向量回归（SVR, support vector regression）算法进行预测。

若给定训练样本为 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, $y_i \in R$, 期望通过学习得到一个形如 $f(x) = w^T x + b$ 的模型，使 $f(x)$ 与 y 尽可能接近， w 和 b 是待确定的模型参数。假设 SVR 能容忍的 $f(x)$ 与 y 的偏差最多为 ε , 即只有在 $|f(x) - y| > \varepsilon$ 时才会计算损失。引入松弛变量 ξ_i 和 $\hat{\xi}_i$, SVR 问题可形式化为

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i + \hat{\xi}_i) \quad (8)$$

$$\text{s.t.} \begin{cases} f(x_i) - y_i \leq \varepsilon + \xi_i \\ y_i - f(x_i) \leq \varepsilon + \hat{\xi}_i \\ \xi_i \geq 0, \hat{\xi}_i \geq 0, i = 1, 2, \dots, m \end{cases}$$

SVR 问题的对偶问题为

$$\max_{\alpha, \hat{\alpha}} \sum_{i=1}^m y_i (\hat{\alpha}_i - \alpha_i) - \varepsilon (\hat{\alpha}_i + \alpha_i) - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m (\hat{\alpha}_i - \alpha_i) (\hat{\alpha}_j - \alpha_j) x_i^T x_j \quad (9)$$

$$\text{s.t.} \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) = 0, 0 \leq \alpha_i, \hat{\alpha}_i \leq C$$

其中， C 为惩罚因子， α_i 和 $\hat{\alpha}_i$ 是拉格朗日乘子。假设核函数为 $K(x, x_i)$, 则最终决策函数为

$$f(x) = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) K(x, x_i) + b \quad (10)$$

其中， $b = y_i + \varepsilon - \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) K(x, x_i)$ 。

径向基核函数结构简单，参数少且应用范围广，因此选取径向基核函数作为本文的核函数。其表达式是 $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$, γ 为核函数参数。由上述推导过程可以看出，惩罚因子 C 和核函数参数 γ 是影响 SVM 性能的重要因素，需要对

两个参数进行优化。

3.2 基于 OCSSA-SVM 的土壤含水量预测

3.2.1 实验数据来源及处理

本文的实验数据来源于陕西省渭南市葡萄园区的官方数据，即园区大棚内监测获取的与土壤环境相关的数据，选取渭南市葡萄园区大棚内 2019 年 9 月 8 日至 24 日的监测数据，时间间隔为 1 h，共计 408 个样本点作为样本数据集。输入向量对土壤含水量预测模型的性能有较大影响，本文考虑选取土壤含水量、空气温度、空气湿度、土壤温度、光照强度等构成输入向量。部分原始数据信息如表 3 所示。

采用 [0,1] 标准化对数据进行预处理，对于异常值利用 3σ 准则对其进行剔除。然后，将预处理后的样本数据集分为训练集和测试集两部分，随机选取 300 个样本数据作为训练集，选取 100 个样本数据作为测试集，对土壤含水量预测模型进行训练和测试。

3.2.2 OCSSA-SVM 算法

由第 3.1 节可知，参数 γ 与 C 对 SVM 预测准确性有至关重要的作用。本文利用 OCSSA 对模型中的参数 C 和 γ 进行联合寻优，OCSSA 的 SVM 参数寻优步骤如下。

- 1) 确定 OCSSA 的相关参数，如种群数量 N 、 C 和 γ 取值的上、下限等。
- 2) 将原始数据分为训练集合测试集，并对数据进行预处理。
- 3) 设置适应度函数。在寻优过程中，选用均方误差 (MSE, mean squared error) 作为寻优算法的适应度函数，其计算式为 $Fitness = \frac{1}{n} \sum_{i=1}^n (Q_i - \hat{Q}_i)^2$ ，其中 n 为数据总数， Q_i 表示真实的数据值， \hat{Q}_i 代表模型预测值。
- 4) 初始化种群 X_{ini} ($ini = 1, 2, \dots, N$)，其中每个

樽海鞘个体由 (C, γ) 组成，利用种群 X_{ini} 进行混沌优化，形成新的种群 X_{inew} ($i = 1, 2, \dots, N$)。

5) 计算种群反向位置 X_{oi} ($i = 1, 2, \dots, N$)，计算两个种群 X_{inew} 和 X_{oi} 中的个体适应度值，并选择两者中适应度值更小的作为樽海鞘个体位置，形成最终初始化种群 X_i 进行算法的迭代寻优。

6) 计算每个樽海鞘个体的适应度值并对其按适应度值大小排序，适应度值排在首位的为食物的位置 X^* 。

7) 计算种群个体的适应度值，根据适应度值大小选出领导者和追随者，并更新其位置。计算更新位置后的适应度值并与当前食物位置的适应度值做比较，选出更优的位置作为食物的新位置。

8) 利用混沌优化和反向学习生成食物的新位置 X_{new}^* ，计算 X^* 和 X_{new}^* 的适应度值，选取更小适应度值的食物位置作为最终食物的位置。

9) 判断是否满足终止条件 (即是否达到最大迭代次数)，若满足终止条件，则结束；否则跳转至上一步，继续进行算法的迭代寻优过程。

10) 输出最优参数 (食物位置) $X^* = (C^*, \gamma^*)$ 。

3.2.3 实验结果对比与分析

在本次实验中，构建了 OCSSA 优化 SVM (OCSSA-SVM)、PSO 优化 SVM (PSO-SVM)、WOA 优化 SVM (WOA-SVM) 及 SSA 优化 SVM (SSA-SVM) 这 4 种土壤含水量预测模型。SVM 的核函数选取径向基函数，参数 C 和 γ 的取值范围均为 $[2^{-8}, 2^8]$ 。OCSSA、PSO、WOA 和 SSA 这 4 种算法的种群大小为 30，最大迭代次数为 300，PSO 中学习因子值为 1.6，惯性因子值为 0.8，OCSSA 最大混沌优化次数为 10 次。选取 MSE 和决定系数评价模型预测性能的优劣，MSE 反映真实值与预测值之间的差，决定系数代表模型拟合程度的好坏，用 R^2 表示，其值越接近 1，模型准确度就越高，预测性能越好；反之则表示模型性能较差。

表 3 部分原始数据信息

时间	空气温度/°C	空气湿度	土壤温度/°C	土壤含水量	光照强度/lx
2019/9/9 12:00	40.4	47.6%	32.1	27.6%	23.9
2019/9/10 14:00	41.0	44.5%	34.3	26.4%	28.0
2019/9/11 16:00	41.4	38.9%	34.4	24.9%	31.8
2019/9/12 18:00	37.4	47.2%	32.2	24.4%	1.1
...
2019/9/23 13:00	40.4	17.3%	31.1	22.8%	23.3
2019/9/24 16:00	39.1	16.7%	31.7	22.0%	28.4

首先，随机选取 300 个数据作为训练集，利用 OCSSA 对 SVM 模型参数进行优化，训练 SVM 模型。然后，将剩余数据中的 100 个数据作为测试集，用于验证优化后预测模型的性能。由于训练集与测试集的选取具有随机性，因此，在相同环境下进行 5 次重复实验，并记录参数寻优后的模型参数 C 、 γ 以及评价模型准确度的 MSE 与 R^2 的值，OCSSA-SVM 模型预测结果如表 4 所示。

表 4 OCSSA-SVM 模型预测结果

运行次数	模型参数 C	模型参数 γ	MSE	R^2
1	14.843 0	7.356 1	0.42	0.903
2	20.999 9	7.047 6	0.39	0.909
3	17.957 1	6.583 7	0.44	0.897
4	9.728 3	5.233 8	0.44	0.897
5	17.326 9	6.878 5	0.43	0.901

由表 4 可以看出，OCSSA-SVM 模型预测土壤含水量的性能较好，测试集决定系数稳定在 0.9 左右，模型的泛化能力较强，此模型可用于土壤含水量的预测。另外，可以看出每次由 OCSSA 优化后的模型参数值变化较大，SVM 模型的参数值保留至小数点后 4 位，取值范围广，OCSSA 的引入可以在更大范围内搜索模型参数的全局最优解。

利用 PSO-SVM、WOA-SVM、SSA-SVM 以及 OCSSA-SVM 这 4 种方法对土壤含水量进行预测，验证 OCSSA-SVM 模型预测的有效性。使用 PSO、WOA、SSA 和 OCSSA 算法分别对 SVM 参数进行寻优，对 5 次重复实验的参数结果取平均值，4 种模型在不同参数下的 SVM 预测结果如表 5 所示。

表 5 4 种模型在不同参数下的 SVM 预测结果

模型名称	MSE	R^2
OCSSA-SVM	0.42	0.901
SSA-SVM	0.59	0.864
PSO-SVM	0.67	0.845
WOA-SVM	0.62	0.859

通过表 5 可知，PSO-SVM 模型是 4 种模型中预测性能最差的，得到的 MSE 和 R^2 分别为 0.67 和 0.845；WOA-SVM 模型和 SSA-SVM 模型的预测性能在 4 种模型中居中，WOA-SVM 模型的 MSE 和 R^2 分别是 0.62 和 0.859，SSA-SVM 模型的 MSE 和 R^2 分别是 0.59 和 0.864；本文提出的 OCSSA-SVM 模型是 4 种模型中预测性能最好的，OCSSA-SVM 模型的 MSE 和 R^2 分别为 0.42 和 0.901。对比 4 个

模型可以看出，OCSSA-SVM 模型在 MSE 和 R^2 两个指标上都取得了最优值。

4 种预测模型在测试集上的预测结果如图 2 所示，由图 2 可以看出，OCSSA-SVM 模型的预测值最接近监测值，OCSSA-SVM 模型的拟合度在 4 种方法中最高。另外，4 种模型的预测误差如图 3 所示，其中，OCSSA-SVM 方法出现较大误差的数据点最少，验证了所提方法的有效性。

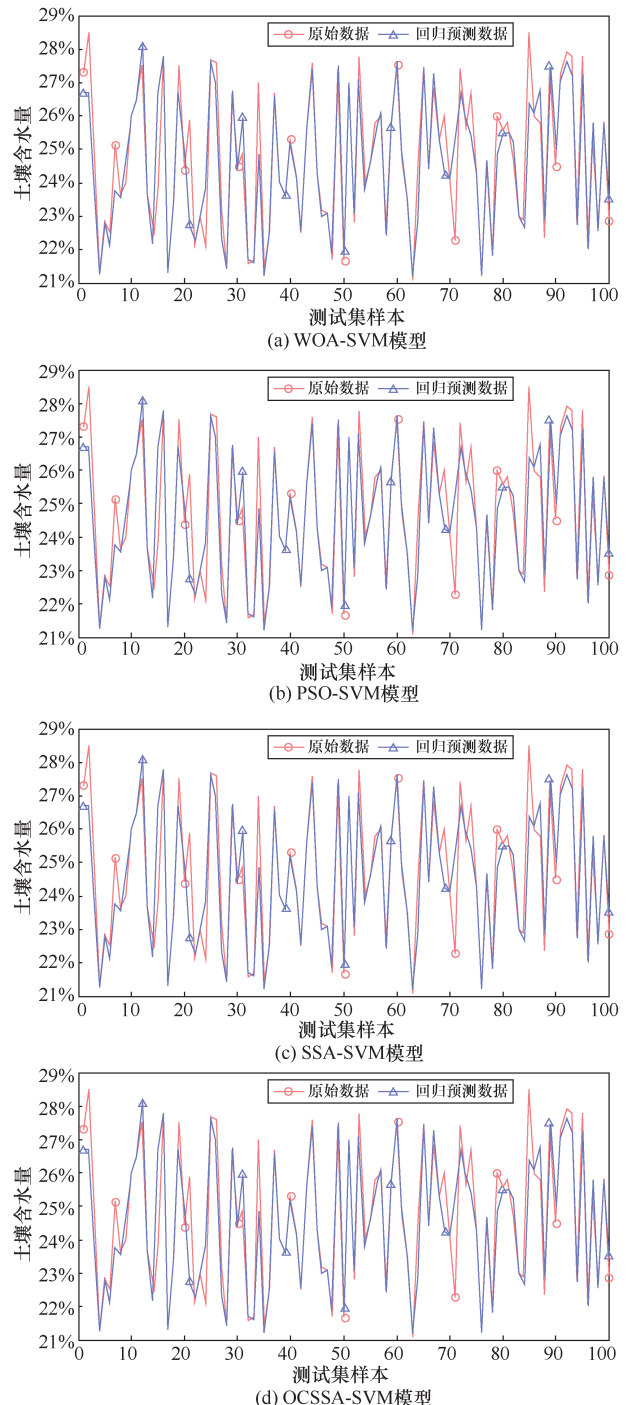


图 2 4 种预测模型在测试集上的预测结果

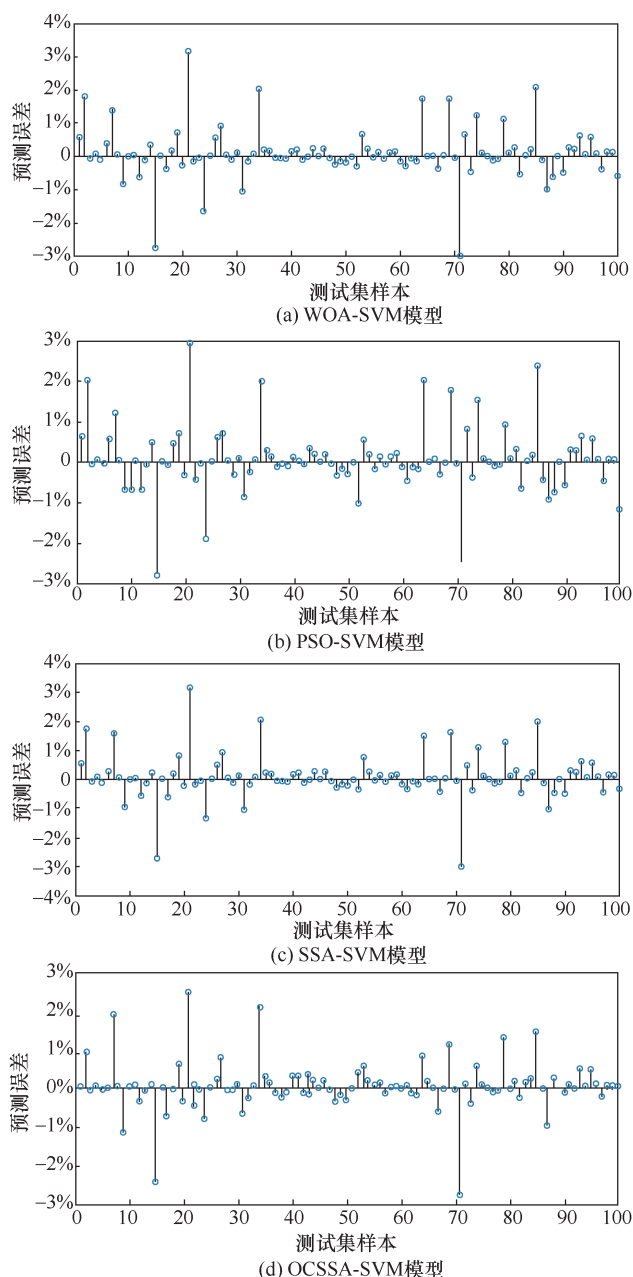


图 3 4 种模型的预测误差

4 结束语

针对传统土壤含水量预测精度低的问题，本文提出一种改进樽海鞘群算法优化 SVM 的土壤含水量预测算法，主要工作内容和成果如下。

1) 由于标准樽海鞘群算法存在收敛速度慢、易陷入局部最优解等问题，提出基于反向学习的混沌樽海鞘群算法。反向学习使得算法能更快地接近全局最优解，加快收敛速度。混沌优化的加入能有效地使算法避免陷入局部最优，可以提升算法所得解的质量。

2) 采用改进后的 OCSSA 对 SVM 模型中的惩

罚因子 C 与径向基核函数中的参数 γ 进行寻优，将 OCSSA-SVM 算法应用于土壤含水量预测中，构造基于 OCSSA-SVM 的土壤含水量预测模型。

3) 采用渭南市葡萄园区大棚的监测数据作为原始实验样本数据，结合 LIBSVM 工具箱对 OCSSA-SVM 土壤含水量预测模型进行训练和预测，并与 PSO-SVM、WOA-SVM、SSA-SVM 预测模型做对比。实验结果表明，基于 OCSSA-SVM 算法的预测模型相比其他 3 种预测模型性能更优。

参考文献:

- [1] QIU J X, GAO Q Z, WANG S, et al. Comparison of temporal trends from multiple soil moisture data sets and precipitation: the implication of irrigation on regional soil moisture trend[J]. International Journal of Applied Earth Observation and Geoinformation, 2016, 48(5): 17-27.
- [2] 王露娇. 基于神经网络的茶园土壤墒情预测方法研究[D]. 合肥: 安徽农业大学, 2018.
WANG L J. Study on the prediction method of soil moisture in tea garden based on neural network[D]. Hefei: Anhui Agricultural University, 2018.
- [3] LEWIN J. Simple soil water simulation model for assessing the irrigation requirements of wheat[J]. Israel Journal of Agricultural Research, 1972, 22(4): 201-213.
- [4] 王玮. 黄河干流灌区土壤墒情动态预测与用水需求模型研究[D]. 西安: 西安理工大学, 2019.
WANG W. Study on dynamic prediction of soil moisture and water demand in the main reaches irrigation area of yellow river[D]. Xi'an: Xi'an University of Technology, 2019.
- [5] 白冬妹, 郭满才, 郭忠升. 时间序列自回归模型在土壤水分预测中的应用研究[J]. 中国水土保持, 2014(2): 42-45.
BAI D M, GUO M C, GUO Z S. Application of time series autoregressive model in soil moisture prediction[J]. Chinese Society of Soil and Water Conservation, 2014(2): 42-45.
- [6] 安小宇, 鲁奎豪, 崔光照. 基于改进樽海鞘优化 BP 神经网络的土壤墒情预测[J]. 中国农机化学报, 2019, 40(11):124-130.
AN X Y, LU K H, CUI G Z. Prediction of soil moisture based on BP neural network optimized by adaptive salp swarm algorithm[J]. Journal of Chinese Agricultural Mechanization, 2019, 40(11): 124-130.
- [7] PRAKASH S, SHARMA A, SAHU S S. Soil moisture prediction using machine learning[C]//Proceedings of Second International Conference on Inventive Communication and Computational Technologies (ICICCT). [S.l.: s.n.], 2018: 1-6.
- [8] MIRJALILI S, GANDOMI A, MIRJALILI S Z, et al. Salp swarm algorithm: a bio-inspired optimizer for engineering design problems[J]. Advances in Engineering Software, 2017, 114(12): 163-191.
- [9] TIZHOOSH H R. Opposition-based learning: a new scheme for machine intelligence[C]//Proceedings of IEEE International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06). Piscataway: IEEE Press, 2005: 695-701.
- [10] WANG H, WU Z, RAHNAMAYAN S, et al. Enhancing particle swarm optimization using generalized opposition-based learning[J]. Information Sciences, 2011, 181(20): 4699-4714.
- [11] KOHLI M, ARORA S. Chaotic grey wolf optimization algorithm for

constrained optimization problems[J]. Journal of Computing and Information Science in Engineering, 2018, 5(4): 458-472.

[12] BONYADI M, MICHALEWICZ Z. Particle swarm optimization for single objective continuous space problems: a review[J]. Evolutionary Computation, 2017, 25(1): 1-54.

[13] MIRJALILI S, LEWIS A. The whale optimization algorithm[J]. Advances in Engineering Software, 2016, 95(5): 51-67.

[作者简介]



赵小强（1977- ），男，博士，西安邮电大学教授，主要研究方向为物联网技术及应用。



杨帆（1994- ），男，西安邮电大学硕士生，主要研究方向为物联网技术及应用。



晏珠峰（1995- ），男，西安邮电大学硕士生，主要研究方向为物联网技术及应用。