

# 摩擦学学报

TRIBOLOGY



## 群智能搜索在基础油性能预测模型中的优化效能

夏延秋, 王裕兴, 冯欣, 蔡美荣

### Optimization Efficiency of Swarm Intelligence Search in Base Oil Performance Prediction Model

XIA Yanqiu, WANG Yuxing, FENG Xin, CAI Meirong

在线阅读 View online: <https://doi.org/10.16078/j.tribology.2021304>

您可能感兴趣的其他文章

Articles you may be interested in



关注微信公众号, 获得更多资讯信息

DOI: 10.16078/j.tribology.2021304

# 群智能搜索在基础油性能预测模型中的优化效能

夏延秋<sup>1</sup>, 王裕兴<sup>1</sup>, 冯欣<sup>1\*</sup>, 蔡美荣<sup>2</sup>

(1. 华北电力大学 能源动力与机械工程学院, 北京 102206;

2. 中国科学院兰州化学物理研究所 固体润滑国家重点实验室, 甘肃 兰州 730000)

**摘要:** 润滑油基础油组成成分的变化对其性能有重要影响, 本文中针对复合基础油体系中KN4010、PAO40和PriEco 3000三种成分含量的不同配比, 对润滑油运动黏度、黏度指数和旋转氧弹性能指标变化的影响, 建立基于最小二乘支持向量机(LSSVM)基础性能预测模型, 并进行对比分析及筛选; 选用经典的粒子群优化算法(PSO)、蜻蜓算法(DA)和鲸鱼优化算法(WOA)等仿生群智能搜索策略构建混合模型, 对优选的预测模型进行参数优化. 测试结果表明: 机器学习技术对油液性能具有良好的预测能力, 并且LSSVM基础模型可以在小样本条件下得到相对较好的预测结果, 而WOA-LSSVM能够显著降低模型的预测误差; 并且通过测试和留一交叉验证法分析, WOA-LSSVM的预测结果明显优于其余模型, 具有良好的预测精度和泛化能力.

**关键词:** 基础油; 含量配比; 机器学习; 性能预测; 群智能优化搜索

中图分类号: TH117.2

文献标志码: A

文章编号: 1004-0595(2023)04-0429-10

## Optimization Efficiency of Swarm Intelligence Search in Base Oil Performance Prediction Model

XIA Yanqiu<sup>1</sup>, WANG Yuxing<sup>1</sup>, FENG Xin<sup>1\*</sup>, CAI Meirong<sup>2</sup>

(1. School of Energy Power and Mechanical Engineering, North China Electric Power University, Beijing 102206, China

2. State Key Laboratory of Solid Lubrication, Lanzhou Institute of Chemical Physics, Chinese Academy of Sciences, Gansu Lanzhou 730000, China)

**Abstract:** Changes in the composition of lubricant base oils have important effects on their performance. While traditional lubricant development needs to be carried out by repeatedly testing and profiling a large number of oil samples for performance, machine learning analysis and processing techniques for data can not only improve efficiency, save costs and eliminate human factor interference, but also make it possible to develop corresponding oil products in a precise and adjustable manner. This paper investigated the effects of different ratios of three component contents of mineral oil (KN4010), polyalpha-olefin synthetic oil (PAO40) and polyol ester (PriEco 3000) in the composite base oil system on the changes of lubricant kinematic viscosity, viscosity index and rotational oxygen bomb performance index. Firstly, a least squares support vector machine (LSSVM) based base performance prediction model was developed and compared with four algorithms, namely, random forest (RF), back propagation neural network (BPNN), extreme learning machine (ELM) and multiple linear regression (MLR), to examine the predictive ability of the base model on three physicochemical performance indicators, namely, kinematic viscosity, viscosity index and rotational oxygen bomb,

Received 21 December 2021, revised 28 February 2022, accepted 28 February 2022, available online 10 March 2022.

\*Corresponding author. E-mail: fengxinemail@126.com, Tel: +86-10-61772251.

This project was supported by China's National Key Research and Development Program (2018YFB0703802) and the Open Project of the State Key Laboratory of Solid Lubrication (LSL-1814).

中国国家重点研究与发展计划项目(2018YFB0703802)和固体润滑国家重点实验室开放课题(LSL-1814)资助.

for model selection. Then the classical Particle Swarm Optimization (PSO) algorithm and the emerging Dragonfly Algorithm (DA) and Whale Optimization Algorithm (WOA) in recent years are used to construct the hybrid prediction model, and the parameters of the preferred LSSVM base prediction model are optimized, i.e., the kernel function width ( $\sigma^2$ ) and the regularization parameter ( $\gamma$ ) are selected optimally. And examine the effects of different parameter optimization methods on the convergence speed, stability and prediction accuracy of the model under different physical and chemical properties with examples. Finally, the classical, base and hybrid models were evaluated and analyzed by comparing the prediction results of the models on the three performance metrics and the leave-one-out cross-validation (LOOCV) test, and their prediction ability in the case of limited data samples was examined, especially the expression of model applicability for sample generalization performance. The results showed that machine learning techniques had good predictive capability for oil performance, and the LSSVM base model could obtain relatively good prediction results under small sample conditions, while WOA-LSSVM could significantly reduce the prediction error of the model, with a mean absolute percentage error of 0.95%, a root mean square error of 1.16 and a coefficient of determination  $R^2$  of 0.998 9 for kinematic viscosity, for example. And WOA-LSSVM enabled the model to search for better results, and its convergence accuracy and stability were better than the two optimization models, PSO-LSSVM and DA-LSSVM, and it was less likely to fall into local optimal solutions and converges faster. Through model testing and analysis by the leave-one-out cross-validation method, the prediction results of WOA-LSSVM were significantly better than the rest of the models, with excellent performance in all aspects, good prediction accuracy and generalization ability, and suitable for studying the problem of lubricant physicochemical performance prediction in small sample situations. The results of the above study provided an effective way for future information studies on the correlation between the composition components and performance indexes of lubricant products.

**Key words:** lubricating oil; content ratio; machine learning; performance prediction; swarm intelligence optimization search

润滑油是应用广泛的重要石化产品,主要由基础油和添加剂组成<sup>[1]</sup>.其中基础油是润滑油的主要成分,对润滑油的黏温性能、高低温性能和挥发性等基础性能起重要作用,添加剂用于改善基础油的原有性能或赋予新的性能<sup>[2]</sup>,通过调整添加剂的种类和用量可以配制满足不同需求的润滑油.人工研发新油品是一项繁琐的工作,耗时费力,且严重依赖人员经验及多种昂贵测试设备.伴随着信息科学的进步,具有强大数据分析和数据处理能力的机器学习技术已延伸至摩擦润滑研究领域,从而减少试验的次数和降低试验研究成本<sup>[3-5]</sup>.在组分与性能关系方面,Altay等<sup>[6]</sup>使用线性回归(LR)、支持向量机(SVM)和高斯过程回归(GPR)机器学习算法建立了精准的铁合金涂层磨损量预测模型;Wang等<sup>[7]</sup>利用小波神经网络(WNN)和遗传算法(GA)模型,研究三种润滑油组分对三类性能指标的影响,预测结果比传统的经验公式计算值更接近试验值;Yu等<sup>[8]</sup>通过反向传播神经网络(BPNN)和遗传算法(GA)构建润滑油抗磨性能预测模型,在寻找三种磷化合物添加剂用量最佳组合方面做了有益尝试.

但是,机器学习方法作回归预测,可能会出现预测训练速度慢、陷入局部极值或过拟合等现象.润滑油组分种类和含量与性能指标的关联关系是1个复杂的系统,且经常要尽量减少油样性能试验样本的需求,所以不仅要求构建的模型具有强大的预测能力<sup>[9]</sup>,

而且还要求在小样本情况下的润滑油性能预测具有良好的预测精度和泛化能力.支持向量机(SVM)在解决小样本方面具有特有的优势,而最小二乘支持向量机(LSSVM)对SVM进行了改进,对数据处理有较好的准确度,降低了求解难度,提高了求解速度<sup>[10-11]</sup>.LSSVM是否适合解决小样本润滑油性能预测问题,一方面需要对LSSVM基础模型的泛化能力和预测精度进行验证,留一交叉验证法(LOOCV, Leave-one-out cross validation)被认为适用于小样本数据条件下的模型性能检验<sup>[12-13]</sup>,其可以用来解决模型验证问题;另一方面需要筛选LSSVM回归预测模型中严重影响预测结果的参数初始值的赋值策略<sup>[14]</sup>,群智能(SI, Swarm intelligence)优化技术通过模拟生物的群体行为,抽象出数学模型,其利用群体高智能的启发式搜索,替代某些参数的人为设定或随机生成,可以有效避免出现收敛速度慢和陷入局部最优解的问题.

本文中基于LSSVM构建润滑油理化性能预测基础模型,通过与随机森林(RF)<sup>[15]</sup>、反向传播神经网络(BPNN)、极限学习机(ELM)<sup>[16]</sup>和多元线性回归(MLR)四种模型的对比分析,考核基础模型对小样本实例的预测精度和泛化能力;采用经典的粒子群优化算法(PSO)<sup>[17]</sup>、蜻蜓算法(DA)<sup>[18]</sup>和鲸鱼优化算法(WOA)<sup>[19]</sup>三种群智能搜索优化方法,分别对核函数宽度( $\sigma^2$ )和正则化参数( $\gamma$ )进行寻优选取,并构建LSSVM混合模

型, 结合实例考察不同的参数优化方法对模型的收敛速度、稳定性及预测精度的影响。

## 1 润滑油样本

### 1.1 样本制备及参数选择

选择润滑油复合基础油体系为测试对象, 该油品由克拉玛依KN4010(矿物油)、美孚PAO40(聚 $\alpha$ -烯烃合成油)和纳克PriEco 3000(多元醇酯)复合而成, 物理特性列于表1中。其中, 矿物油价格低廉, 来源可靠, 作为降低润滑油成本的必要成分; 聚 $\alpha$ -烯烃是配制高性能润滑油较为理想的合成基础油, 具有良好的热稳定性和低温性能; 而合成酯类基础油的黏温性能和环境友好性比较好, 但水解稳定性较差<sup>[20]</sup>。将三者复合使用, 可以优势互补, 避免各自存在的不足。

后续研究中, 输入数据为复合基础油体系油品的KN4010、PAO40和PriEco 3000三种组分的质量分数, 输出数据则选择对应的运动黏度(40 °C)、黏度指数和旋转氧弹三种性能指标。

### 1.2 试验样本集划分

为了观察和测试不同成分分配比对润滑油性能的影响, 设计了30组试验样本, 具体各成分质量分数( $\omega$ )的设计方案列于表2中。

在本次研究过程中, 为使模型达到较好的预测效果, 小样本测试训练集大约取整个数据集的2/3~4/5。据此, 将上述30组样本数据进行划分, 训练集样本为样本设计方案中的前21组(样本编号1~21), 测试集样本为设计方案中后9组(样本编号22~30)。

## 2 构建模型

### 2.1 最小二乘支持向量机

LSSVM是以传统SVM算法为基础的一种改进算法, 将等式约束代替不等式约束, 在构造损失函数时应用最小二乘准则, 实现将复杂二次规划问题转化为线性方程组问题的求解, 从而降低计算的复杂程度, 加快收敛速度, 并且在非线性系统中效果会更好。基本原理如下:

给定训练样本集 $\{(x_i, y_i) | i = 1, 2, \dots, n\}$ , 其中,  $x_i \in R^d$ 为 $d$ 维训练样本输入,  $y_i \in R$ 为训练样本输出,  $n$ 为训练集样本个数。

用如下高维特征空间的线性函数对样本进行拟合:

$$f(x) = \omega^T \varphi(x) + b \tag{1}$$

式中:  $\omega$ 为权系数向量,  $b$ 为阈值,  $T$ 为矩阵转置符号, 输入数据通过非线性映射函数 $\varphi(x)$ 被映射到高维特征空间。

根据结构风险最小准则, LSSVM回归可以表示为约束优化问题, 最优 $\omega$ 和 $b$ 可经下述函数最小化得到:

$$\min_{\omega, b, \zeta} J(\omega, \zeta) = \frac{1}{2} \omega^T \omega + \frac{1}{2} \gamma \sum_{i=1}^n \zeta_i^2 \tag{2}$$

约束条件为

$$y_i = \omega^T \cdot \varphi(x_i) + b + \zeta_i \tag{3}$$

式中:  $\gamma$ 为正则化参数,  $J$ 为目标函数,  $\zeta$ 是拟合误差,  $i$ 为样本数。通过调节 $\gamma$ 可以平衡 $\zeta_i$ 在目标函数中的权重, 避免最小寻优过程中出现过拟合问题。

为解决LSSVM的优化问题, 构造如下拉格朗日函数:

$$L(\omega, b, \zeta, \alpha) = \frac{1}{2} \omega^T \omega + \frac{1}{2} \gamma \sum_{i=1}^n \zeta_i^2 - \sum_{i=1}^n \alpha_i [\omega^T \varphi(x_i) + b + \zeta_i - y_i] \tag{4}$$

式中:  $\alpha_i$ 为拉格朗日乘子, 由KKT条件(Karysh-Khun-Tucker)计算得到:

$$\begin{cases} \frac{\partial L}{\partial \omega} = 0 \rightarrow \omega - \sum_{i=1}^n \alpha_i \varphi(x_i) = 0 \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^n \alpha_i = 0 \\ \frac{\partial L}{\partial \zeta_i} = 0 \rightarrow \alpha_i = \gamma \zeta_i \\ \frac{\partial L}{\partial \alpha_i} = 0 \rightarrow \omega^T \varphi(x_i) + b + \zeta_i - y_i = 0 \end{cases} \tag{5}$$

表 1 基础油物理特性

Table 1 Physical properties of base oil

Sample	Kinematic viscosity (100 °C)/(mm <sup>2</sup> /s)	Kinematic viscosity (40 °C)/(mm <sup>2</sup> /s)	Viscosity index	Pour point/°C	Flash point/°C
KN4010	10.54	127	51	-15	220
PAO40	39.0	396	147	-36	281
PriEco 3000	4.5	20.0	140	-51	250
Experimental methods	ASTM D445	ASTM D445	ASTM D2270	ASTM D97	ASTM D92

表2 样本设计方案

Table 2 Sample design solutions

Number	w (KN4010)/%	w (PAO40)/%	w (PriEco 3000)/%	Number	w (KN4010)/%	w (PAO40)/%	w (PriEco 3000)/%
1	55	25	20	16	30	40	30
2	50	20	30	17	40	40	20
3	40	45	15	18	40	35	25
4	45	30	25	19	50	25	25
5	45	25	30	20	55	20	25
6	35	35	30	21	35	30	35
7	60	20	20	22	45	20	35
8	35	45	20	23	30	50	20
9	30	45	25	24	50	30	20
10	55	30	15	25	40	30	30
11	45	40	15	26	50	15	35
12	45	35	20	27	35	40	25
13	60	25	15	28	30	35	35
14	50	35	15	29	35	50	15
15	30	55	15	30	40	25	35

消去变量 $\omega$ ,  $\zeta$ , 可得矩阵方程:

$$\begin{bmatrix} 0 & \mathbf{e}_{1 \times n} \\ \mathbf{e}_{n \times 1} & R + \frac{1}{\gamma} \mathbf{E} \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (6)$$

式中:

$$R = \{K(x_i, x_j) | i, j = 1, 2, \dots, n\};$$

$$y = [y_1, y_2, \dots, y_n]^T;$$

$$\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]^T;$$

$\mathbf{e}_{1 \times n}$  是  $1 \times n$  的单位行向量,  $\mathbf{e}_{n \times 1}$  是  $n \times 1$  的单位列向量,  $\mathbf{K}$  为核矩阵,  $\mathbf{E}$  是  $n \times n$  的单位矩阵.

由于模型的特征数较小并且与训练样本数的比值大小适中, 所以本次优化问题的求解可选择RBF(径向基)核函数作为核函数, 表达式如下:

$$\mathbf{K}(x, x_j) = \exp\left(-\frac{\|x - x_j\|^2}{\sigma^2}\right) \quad (7)$$

式中:  $\sigma$  为核函数宽度系数, 如果  $\sigma$  较大, 易把所有样本点归为同一类; 反之则会出现过拟合问题.

由以上各式求得  $b$  和  $a$ , 最终由非线性方程变换得到线性模型, 那么对于新的样本  $x$ , 则有LSSVM模型的输出  $y$  为

$$y = \sum_{i=1}^n \alpha_i K(x, x_j) + b \quad (8)$$

## 2.2 留一交叉验证法

留一交叉验证法(LOOCV)是一种特殊的交叉验证方法, 用于衡量模型在数据集上的泛化能力. 整个过程虽然计算繁琐, 但样本利用率高, 尤其适合在小样本情况下进行模型选择, 避免出现过拟合现象. 操

作步骤为在原有的训练集中只保留1个样本用做测试, 而将其他样本归入训练集; 以此类推, 对集合中的其他样本顺次进行迭代, 直至遍历训练集中的每一个样本, 如图1所示. 假如原训练集有  $n$  个样本, 那么模型需要训练  $n$  次、测试  $n$  次.

## 2.3 群智能搜索算法

群智能(SI)优化算法来源于对昆虫、兽群、鸟群和鱼群等生物群体行为的模仿, 从中抽象出具有群体高智能的启发式搜索数学模型, 用于解决各类优化问题, 可以有效地避免收敛速度慢和陷入局部最优解情况的发生.

对群智能优化算法的研究一直非常活跃, 其中粒子群优化算法(PSO)是比较经典的一种方法, 最早由Eberhart和Kennedy于1995年提出, 源于对鸟群捕食行为的研究, 根据鸟自身及其同伴的飞行经验来调整自己的飞行速度和位置最终实现更新与优化. 近年来, 新兴的仿生群智能算法不断涌现, 具有代表性的有: Seyedali等源于自然界中蜻蜓群体静态的觅食行为和动态的迁徙行为, 于2016年提出了蜻蜓算法(DA, Dragonfly algorithm); Mirjalili等模拟鲸鱼群包围猎物、气泡攻击猎物和搜索猎物3个阶段的捕食行为, 于2016年提出鲸鱼优化算法(WOA, Whale optimization algorithm). 以上这些群智能方法对问题是否适用, 尚需要多种类型的试验来验证.

## 2.4 基础模型

对于回归模型, 除LSSVM之外, 目前应用较多的算法还有随机森林(RF)、反向传播神经网络(BPNN)、极限学习机(ELM)及多元线性回归(MRL)等. 与LSSVM

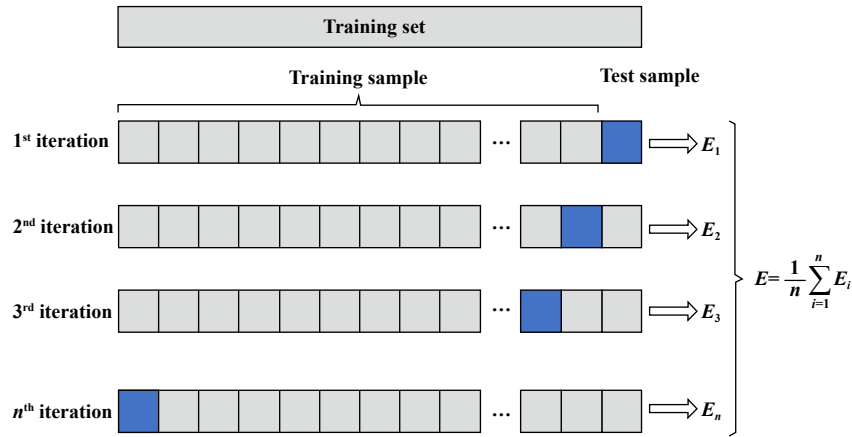


Fig. 1 Leave-one-out cross validation (LOOCV)

图 1 留一交叉验证法(LOOCV)

不同, RF通过组合多个决策树算法对样本训练得出最后预测结果, 对异常值和噪声具有较好的容忍度, 在测试中根据样本数量及特征数大小设置参数生长树的数目为500以及在每一个分裂节点处样本预测器个数为1, 其他为默认值; BPNN通过反向传播训练来不断调整网络的权值和阈值, 使预测输出不断逼近期望输出, 在训练中根据输入与输出维数设置隐含层神经元个数为5, 又设置适宜的最大训练次数为1 000和学习率为0.01; ELM则随机产生且保持不变的输入层与隐含层间的连接权值、隐含层神经元的阈值, 通过设置隐含层神经元个数训练模型, 其测试中隐含层神经元个数为5; 多元线性回归(MRL)是利用数理统计中回归分析来研究1个应变量依赖多个自变量的变化关系, 来确定其定量关系的一种统计分析方法, 在模型实践中不需要额外设置参数. 显然, 这些方法各具优势, 通过润滑油样本实例, 构建以上四种方法的润滑油性能预测模型, 经对比分析, 评判LSSVM模型在解决小样本润滑油性能预测问题的预测精度和泛化能力.

### 2.5 混合模型

对LSSVM模型预测性能影响较大的参数是核函数宽度( $\sigma^2$ )和正则化参数( $\gamma$ ): 其中 $\sigma^2$ 是核函数自带的1个参数, 影响着数据映射到新特征空间的分布,  $\sigma^2$ 越大, 支持向量越少,  $\sigma^2$ 值越小, 支持向量越多, 而支持向量的个数影响训练与预测的速度;  $\gamma$ 是对误差的宽容度, 其过大或过小, 都会使模型泛化能力变差. 所以不合理的初始值设置, 将导致计算时间长、效率差以及预测精度下降. 为此, 分别采用粒子群优化算法(PSO)、蜻蜓算法(DA)和鲸鱼优化算法(WOA)三种群智能搜索策略对参数寻优, 构建LSSVM混合模型, 总体步骤如下:

Step1: 对样本集进行归一化, 分别设定五种群智

能优化算法运行过程参数的初值, 其中种群数(SearchAgents\_no)为20, 最大迭代次数(Max\_iter)为100, 参数寻优区间为0.01~1 000;

Step2: 初始化种群各个体位置, 其位置代表参数 $\sigma^2$ 和 $\gamma$ , 将其代入LSSVM模型计算预测值与真实值的均方根误差(RMSE)并定义为算法的适应度函数;

Step3: 计算种群中每个个体的适应度(fitness), 选择出当前最优适应度值个体, 将该个体位置设置为当前的最优位置;

Step4: 经过不断迭代更新搜索位置, 返回Step3, 重新计算适应度(fitness)并进行比较, 找出并更新至最优位置, 重复进行, 直至满足结束条件;

Step5: 将最后寻优得到的 $\sigma^2$ 和 $\gamma$ 参数再次代入到LSSVM基础模型中, 完成对LSSVM预测模型的训练过程.

混合模型工作流程示意图如图2所示.

### 2.6 评价标准

本文中采用平均绝对百分误差(MAPE, Mean absolute percentage error)、均方根误差(RMSE, Root mean square error)和决定系数( $R^2$ , Coefficient of determination)作为模型综合性能评价标准, 计算公式如下:

$$MAPE = 100 \times \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (9)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (10)$$

$$R^2 = \frac{\left( n \sum_{i=1}^n y_i \hat{y}_i - \sum_{i=1}^n y_i \sum_{i=1}^n \hat{y}_i \right)^2}{\left[ n \sum_{i=1}^n \hat{y}_i^2 - \left( \sum_{i=1}^n \hat{y}_i \right)^2 \right] \left[ n \sum_{i=1}^n y_i^2 - \left( \sum_{i=1}^n y_i \right)^2 \right]} \quad (11)$$

式中:  $n$  是样本总量,  $\hat{y}_i$  为测试样本的预测值,  $y_i$  为真实值. 决定系数  $R^2$  介于 0 到 1 之间, 且越接近 1, 模型拟合度越高, 性能越好.

### 3 结果与讨论

#### 3.1 样本数据

按照试验方案配置 30 组试验样本, 加热至 60 °C 左右, 恒温搅拌 1 h 即可制备出所需试样. 并遵循相关行业制定的标准, 测定每组样本的运动黏度(40 °C, 100 °C)

和旋转氧弹等指标, 并依据石油产品黏度指数计算法计算黏度指数. 为了直观展示所设计的 30 组试验样本, 绘制三种性能指标测试结果的空间分布图(图 3), 图中以颜色表明输入数据与输出数据的对应关系.

#### 3.2 模型选取

将样本划分完成的训练集和测试集数据代入模型中, 分别建立基于 RF、BPNN、ELM、MLR 和 LSSVM 的理化性能预测模型. 经过反复多次测试, LSSVM 的表现相对突出, 而其余几种算法模型的测试结果都出

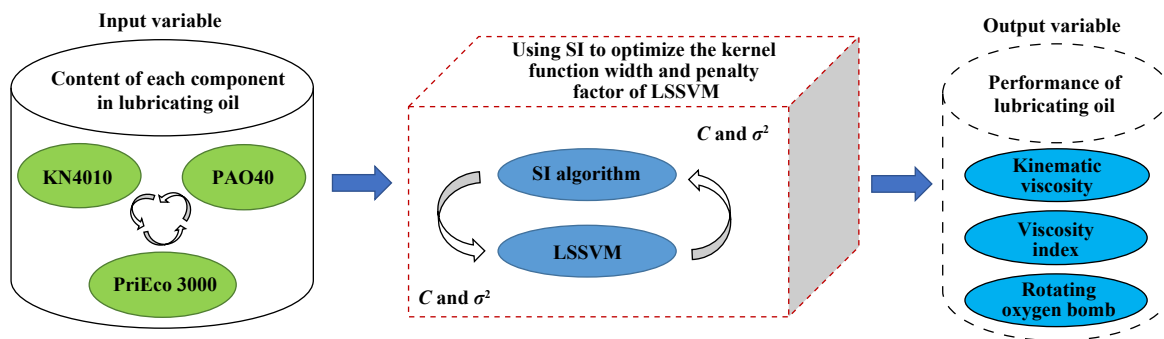


Fig. 2 Flow chart of hybrid model for lubricant performance prediction

图 2 润滑油性能预测混合模型工作流程图

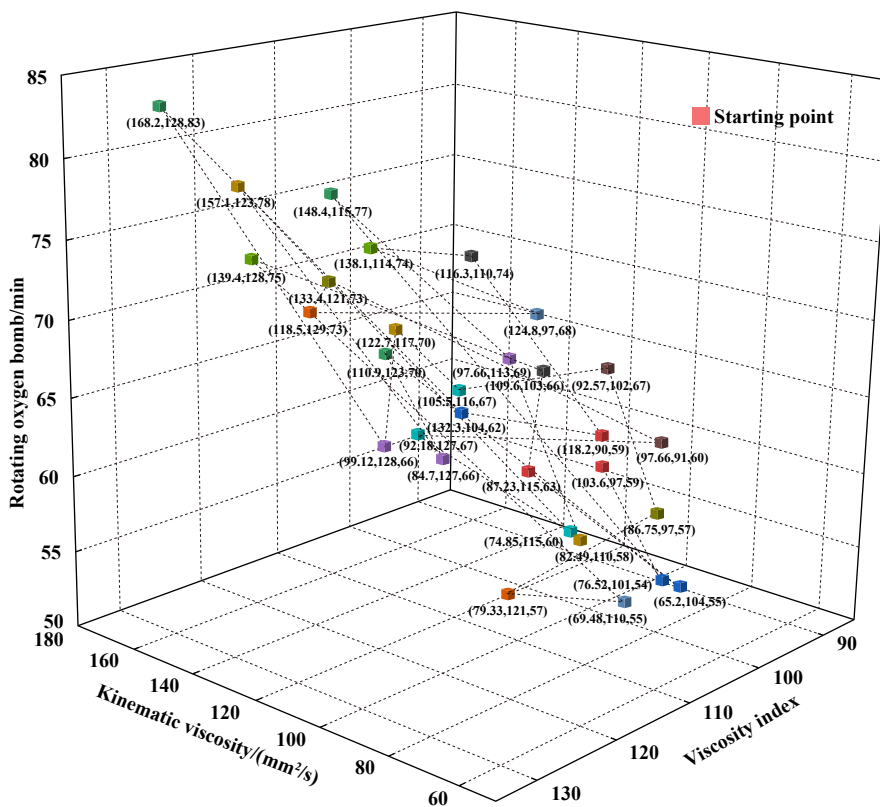


Fig. 3 Spatial distribution of test results

图 3 测试结果空间分布图

现较大的误差, 无法满足本次研究需要. 分别对几种算法模型的3个理化指标模型的MAPE、RMSE和 $R^2$

进行计算, 其中纵坐标为各指标百分误差值(Relative percentage error, RPE), 结果如图4所示.

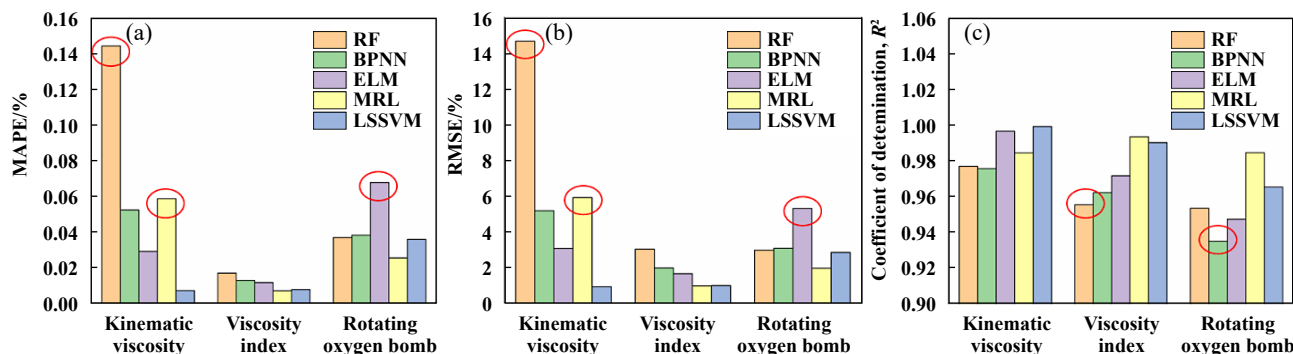


Fig. 4 Evaluation of basic model predictions: (a) MAPE; (b) RMSE; (c)  $R^2$

图 4 基础模型预测评价: (a) MAPE; (b) RMSE; (c)  $R^2$

由图4可以看出, RF在运动黏度(40 °C)和黏度指数的预测上出现较大误差; BPNN和ELM则是在旋转氧弹上表现较差; MLR虽然在黏度指数和旋转氧弹两项指标表现优异, 但是对于运动黏度(40 °C)则是表现异常. 而LSSVM对于3个理化性能的预测都表现出良好的预测性能, 平均相对百分误差都在6%以内, 决定系数均在0.96以上. 显然, 其余四种模型在指标数据集上出现不同程度的误差, 而LSSVM适合润滑油基础理化性能的预测模型.

在LSSVM的基础模型中, 两个关键参数的产生

是随机选取的, 具有不确定性, 参数确定方法会严重影响模型训练时间和预测精度. 本文中采用交叉验证 (fold=10) 的方式来调整参数, 而其参数确定仍是简单粗略的. 从预测效果来看, 还有不少改进空间, 需要更加强化的参数寻优方法来解决.

### 3.3 模型优化

按照2.5节中的步骤, 分别构建PSO-LSSVM、DA-LSSVM和WOA-LSSVM三种群智能搜索算法与LSSVM的混合预测模型, 考察这三种混合模型在训练中的迭代进化过程如图5所示.

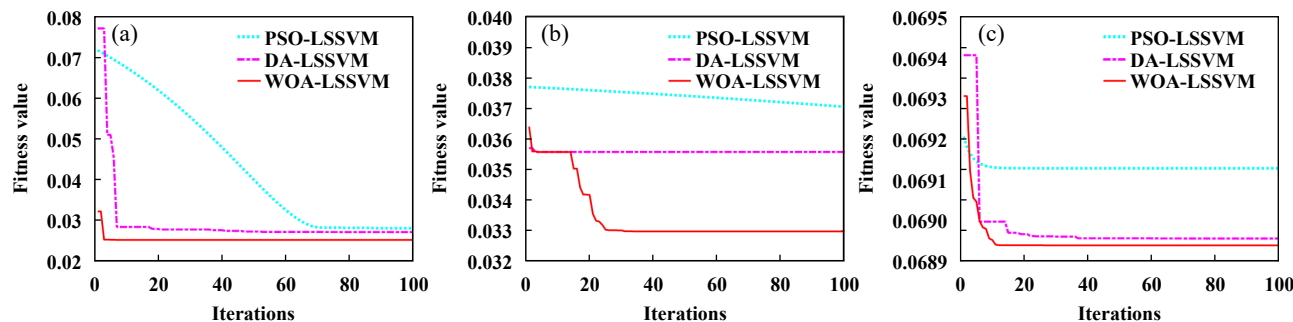


Fig. 5 Iterative curves of three hybrid models: (a) kinematic viscosity; (b) viscosity index; (c) rotating oxygen bomb

图 5 三种混合模型迭代曲线: (a)运动黏度; (b)黏度指数; (c)旋转氧弹

经过反复测试, 对比三种混合模型的收敛速度、稳定性以及收敛精度可知: PSO-LSSVM各项指标都是最差的; DA-LSSVM的迭代速度较快, 收敛稳定性也较差, 容易使模型陷入局部最优解; WOA-LSSVM能够使模型搜索到更优的结果, 其收敛精度和稳定性均优于其余的模型, 并且收敛速度更快. 因此, 采用鲸鱼优化算法(WOA)对模型参数进行优化, 可以使模型

的预测效果更好.

以其中黏度指数为例, LSSVM经过三种群智能算法寻优之后的 $\gamma_{best}$ 和 $\sigma^2_{best}$ 列于表3中.

### 3.4 模型评价

通过LSSVM、WOA-LSSVM以及文献[7]中GAWNN分别建立模型, 将全部划分好的训练集和测试集代入三种模型中进行预测, 其在3个性能指标测试



表3 三种混合模型最佳参数

Table 3 Optimal parameters of three hybrid models for kinematic viscosity

Model	$\gamma_{\text{best}}$	$\sigma^2_{\text{best}}$
PSO	351.45	450.38
DA	450.38	1000
WOA	11.37	8.74

结果的相对百分误差如图6所示。

可以看出, 经过参数优化后的模型WOA-LSSVM, 在3个理化性能数据上都有优异的表现。其中, 旋转氧弹和黏度指数上都有较大提升; 在运动黏度(40 °C)上虽然与基础模型表现相近, 但是两者预测精度已经达到较好的程度。计算三种模型的MAPE、RMSE和 $R^2$ , 其中纵坐标为各指标误差值, 其对比图如图7所示。

WOA-LSSVM的三项误差指标与其他方法相比明显较优, 数据的拟合效果好, 测试集的预测精度更高。由此可见, WOA-LSSVM各项预测结果均明显优于GA-WNN和LSSVM, 而对于基础模型LSSVM表现不佳的黏度指数和旋转氧弹两个数据集上都有较大提升。

采用已经划分好的数据集代入模型取得的预测结果, 对于未知样本无法较好的表现出适用性。为了

充分表现模型对于新样本数据的预测能力, 使用LOOCV法不断重新划分数据集来训练数据, 以此考察模型的泛化能力。将30组样本分别代入GA-WNN、LSSVM和WOA-LSSVM模型中, 其在3个性能指标数据集上测试结果的相对百分误差如图8所示。

综合来看, WOA-LSSVM模型通过LOOCV法验证的相对百分误差结果优于其他两种模型, 而基础模型LSSVM又优于GA-WNN。同时计算三种模型的MAPE、RMSE和 $R^2$ , 其中纵坐标为各指标误差值, 结果如图9所示。显然, WOA-LSSVM与LSSVM和GA-WNN相比, 三项评价指标都有较大提升, 因此WOA-LSSVM具有更好的预测精度和泛化能力。

### 3.5 预测结果

上述经对比分析可知, 群智能搜索优化算法对模型的预测性能有较大提升, 其中WOA-LSSVM模型具有更好的迭代速度和预测精度。用30组样本(21组为训练集, 其余9组为测试集)代入模型进行预测, 将WOA迭代寻优后的参数 $\gamma_{\text{best}}$ ,  $\sigma^2_{\text{best}}$ 代入LSSVM中进行预测, 结果列于表4中。

WOA-LSSVM模型的预测结果与试验值的对比结果如图10所示。

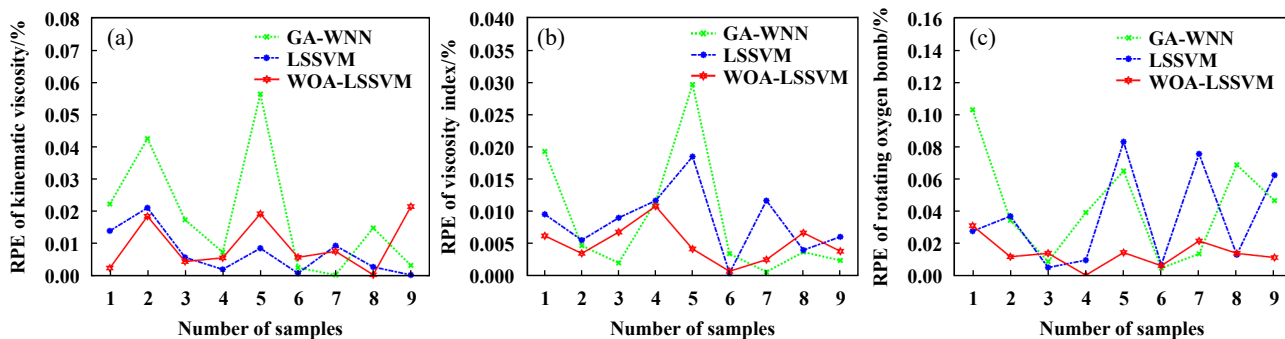
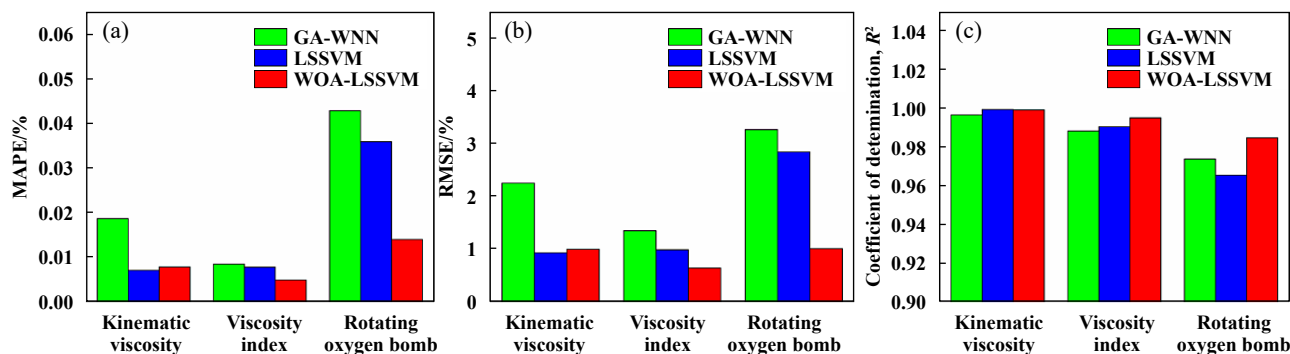


Fig. 6 Prediction results of model: (a) kinematic viscosity; (b) viscosity index; (c) rotating oxygen bomb

图6 模型预测结果: (a)运动黏度; (b)黏度指数; (c)旋转氧弹

Fig. 7 Evaluation of model predictions: (a) MAPE; (b) RMSE; (c)  $R^2$ 图7 模型预测评价: (a) MAPE; (b) RMSE; (c)  $R^2$

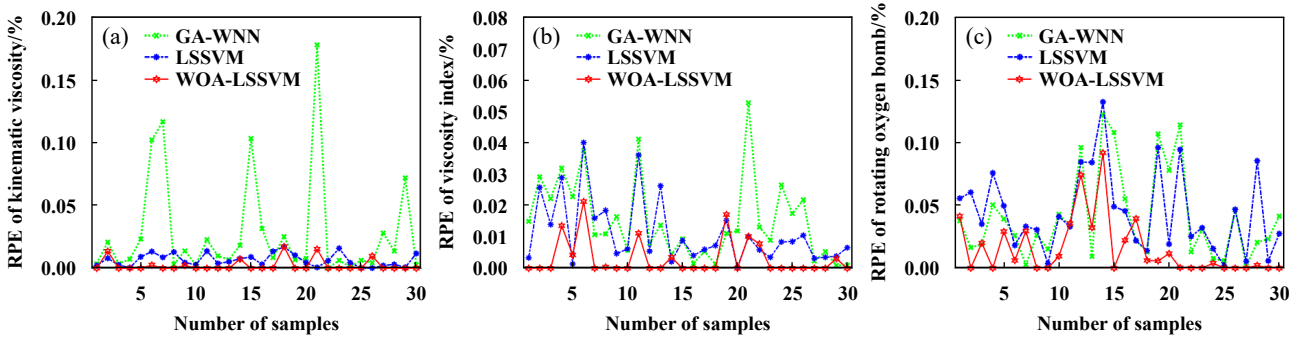


Fig. 8 Prediction results of model by LOOCV: (a) kinematic viscosity; (b) viscosity index; (c) rotating oxygen bomb

图 8 模型LOOCV法预测结果: (a)运动黏度; (b)黏度指数; (c)旋转氧弹

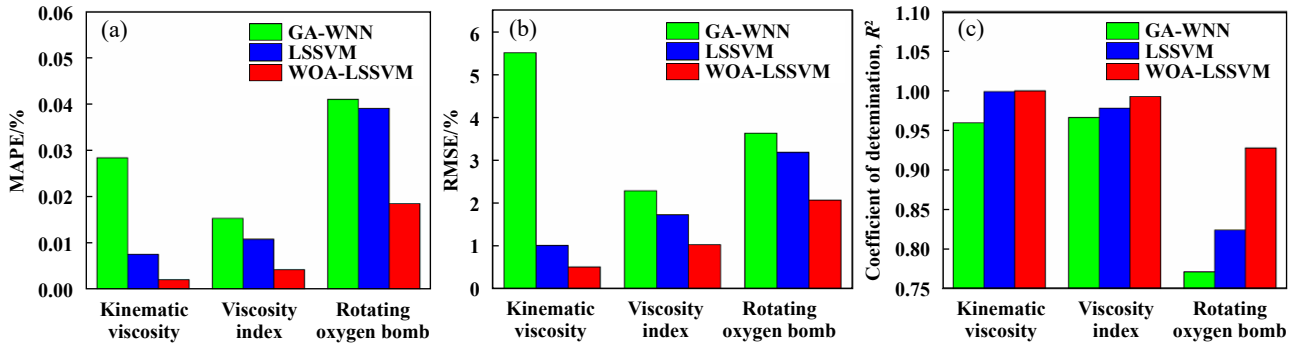


Fig. 9 Evaluation of model by LOOCV: (a) MAPE; (b) RMSE; (c)  $R^2$

图 9 模型LOOCV法评价: (a) MAPE; (b) RMSE; (c)  $R^2$

表 4 WOA-LSSVM预测结果

Table 4 Prediction results of WOA-LSSVM

WOA model	Kinematic viscosity/(mm <sup>2</sup> /s)	Viscosity index	Rotating oxygen bomb/min
1	69.31	109.43	56.67
2	141.98	128.20	75.87
3	110.08	104.06	65.05
4	86.75	115.72	63.06
5	66.45	103.38	53.99
6	111.54	121.99	69.48
7	84.06	127.56	64.76
8	157.05	122.38	76.86
9	73.24	115.48	59.37

将数据代入模型WOA-LSSVM测试后,可以看出数据的拟合结果良好,各项计算误差指标都处于较小范围,在性能指标预测上表现优异。

## 4 结论

a. 通过对多种回归模型的测试分析,发现LSSVM相较于其余模型具有良好的预测能力,适合于研究小样本润滑油性能预测问题。

b. 群智能搜索算法对LSSVM基础模型的正则化参数( $\gamma$ )和核函数宽度( $\sigma^2$ )进行参数寻优,可以显著降低预测误差。

c. 混合模型WOA-LSSVM收敛速度更快,能够搜

计算WOA-LSSVM模型的MAPE、RMSE和 $R^2$ 列于表5中。

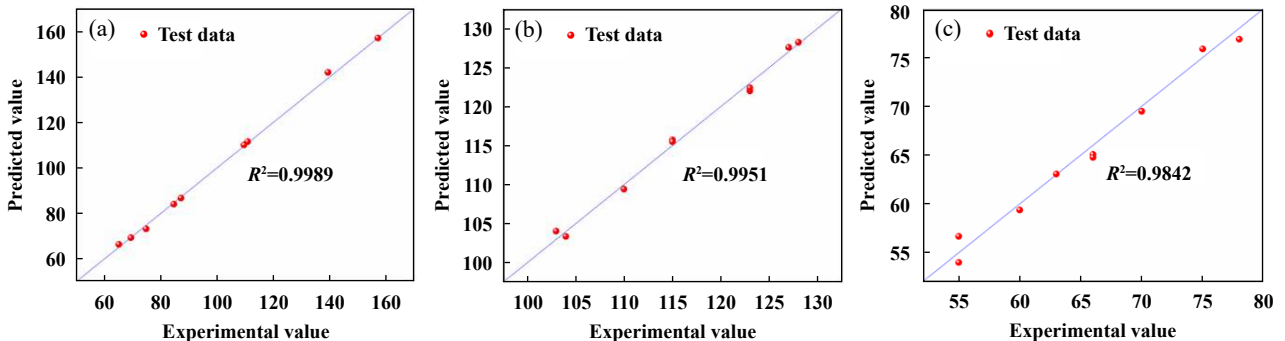


Fig. 10 Comparison of model prediction results: (a) kinematic viscosity (40 °C); (b) viscosity index; (c) rotating oxygen bomb

图 10 模型预测结果对比: (a)运动黏度(40 °C); (b)黏度指数; (c)旋转氧弹

表5 WOA-LSSVM模型预测评价

Table 5 Evaluation of WOA-LSSVM model predictions

Model	Kinematic viscosity (40 °C)	Viscosity index	Rotating oxygen bomb
MAPE	0.95%	0.49%	1.40%
RMSE	1.16	0.64	1.01
$R^2$	0.998 9	0.995 1	0.984 2

索到更优的结果,在各项性能上的预测精度和泛化能力最优。

### 参考文献

- [1] Guo Shoujing, Bai Tianzhong, Liang Xuemei, et al. Research on blending of different base oils and application in diesel engine oil[J]. *Petroleum Refinery Engineering*, 2021, 51(9): 53–56 (in Chinese) [郭守敬, 白天忠, 梁雪美, 等. 不同基础油之间调配及在柴油机油中应用研究[J]. *炼油技术与工程*, 2021, 51(9): 53–56]. doi: 10.3969/j.issn.1002-106X.2021.09.015.
- [2] Ge Xiangyu, Xia Yanqiu, Feng Xin, et al. Electrical conductivities and tribological properties of lithium salts conductive grease[J]. *Journal of Mechanical Engineering*, 2015, 51(15): 61–66 (in Chinese) [葛翔宇, 夏延秋, 冯欣, 等. 锂盐型电力复合脂的导电性和摩擦学性能[J]. *机械工程学报*, 2015, 51(15): 61–66]. doi: 10.3901/JME.2015.15.061.
- [3] Xia Yanqiu, Wang Chen, Feng Xin. GA-BPSO Hybrid Optimization of Middle Infrared Spectrum Feature Band Selection of Lubricating Oil Additive Type Identification Technology[J]. *Tribology*, 2022, 42(1): 142–152 (in Chinese) [夏延秋, 王宸, 冯欣. GA-BPSO混合优化中红外光谱特征波段筛选的润滑油添加剂种类识别技术[J]. *摩擦学学报*, 2022, 42(1): 142–152]. doi: 10.16078/j.tribology.2020164.
- [4] Ulas M, Altay O, Gurgenc T, et al. A new approach for prediction of the wear loss of PTA surface coatings using artificial neural network and basic, kernel-based, and weighted extreme learning machine[J]. *Friction*, 2020, 8(6): 1102–1116. doi: 10.1007/s40544-017-0340-0.
- [5] Li Xiaohu, Feng Xin, Xia Yanqiu. IR spectra of grease optimization based on cuckoo search[J]. *Spectroscopy and Spectral Analysis*, 2017, 37(12): 3703–3708 (in Chinese) [李晓鹤, 冯欣, 夏延秋. 布谷鸟搜索的润滑脂特征红外光谱波段优选技术[J]. *光谱学与光谱分析*, 2017, 37(12): 3703–3708].
- [6] Altay O, Gurgenc T, Ulas M, et al. Prediction of wear loss quantities of Ferro-alloy coating using different machine learning algorithms[J]. *Friction*, 2020, 8(1): 107–114. doi: 10.1007/s40544-018-0249-z.
- [7] Wang Guomin, Wu Yuanyuan, Jiang Haifu, et al. Physical and chemical indexes of synthetic base oils based on a wavelet neural network and genetic algorithm[J]. *Industrial Lubrication and Tribology*, 2020, 72(1): 116–121. doi: 10.1108/ilt-03-2019-0101.
- [8] Yu Tong, Yin Peng, Zhang Wei, et al. A compounding-model comprising back propagation neural network and genetic algorithm for performance prediction of bio-based lubricant blending with functional additives[J]. *Industrial Lubrication and Tribology*, 2021, 73(2): 246–252. doi: 10.1108/ILT-05-2020-0165.
- [9] Parikh H H, Gohil P P. Experimental investigation and prediction of wear behavior of cotton fiber polyester composites[J]. *Friction*, 2017, 5(2): 183–193. doi: 10.1007/s40544-017-0145-y.
- [10] Zhou Liang. Toward prediction of kinematic viscosity of biodiesel using a robust approach[J]. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 2018, 40(23): 2895–2902. doi: 10.1080/15567036.2018.1513099.
- [11] Sepchr M, Baghban M, Ghanbari A, et al. Modeling dynamic viscosity of n-alkanes using LSSVM technique[J]. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 2018, 40(16): 1966–1973. doi: 10.1080/15567036.2018.1486906.
- [12] Huang Pinghua, Yang Zhongyuan, Wang Xinyi, et al. Research on Piper-PCA-Bayes-LOOCV discrimination model of water inrush source in mines[J]. *Arabian Journal of Geosciences*, 2019, 12(11): 1–14. doi: 10.1007/s12517-019-4500-3.
- [13] Miri M, Ghassoun Y, Dovlatbadi A, et al. Estimate annual and seasonal PM<sub>1</sub>, PM<sub>2.5</sub> and PM<sub>10</sub> concentrations using land use regression model[J]. *Ecotoxicology and Environmental Safety*, 2019, 174: 137–145. doi: 10.1016/j.ecoenv.2019.02.070.
- [14] Zhao Haoran, Guo Sen, Zhao Huiru. Energy-related CO<sub>2</sub> emissions forecasting using an improved LSSVM model optimized by whale optimization algorithm[J]. *Energies*, 2017, 10(7): 874. doi: 10.3390/en10070874.
- [15] Gholizadeh M, Jamei M, Ahmadianfar I, et al. Prediction of nanofluids viscosity using random forest (RF) approach[J]. *Chemometrics and Intelligent Laboratory Systems*, 2020, 201: 104010. doi: 10.1016/j.chemolab.2020.104010.
- [16] Xia Yanqiu, Xu Dayi, Feng Xin, et al. Identification and content prediction of lubricating oil additives based on extreme learning machine[J]. *Tribology*, 2020, 40(1): 97–106 (in Chinese) [夏延秋, 徐大祎, 冯欣, 等. 基于极限学习机和优化算法的润滑油添加剂种类识别与含量预测[J]. *摩擦学学报*, 2020, 40(1): 97–106]. doi: 10.16078/j.tribology.2019107.
- [17] Ahmadi H, Ahmadi H, Baghban A. Modeling vaporization enthalpy of pure hydrocarbons and petroleum fractions using LSSVM approach[J]. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 2020, 42(5): 569–576. doi: 10.1080/15567036.2019.1587103.
- [18] Mirjalili S. Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems[J]. *Neural Computing and Applications*, 2016, 27(4): 1053–1073. doi: 10.1007/s00521-015-1920-1.
- [19] Mirjalili S, Lewis A. The whale optimization algorithm[J]. *Advances in Engineering Software*, 2016, 95: 51–67. doi: 10.1016/j.advengsoft.2016.01.008.
- [20] Zhang Xin ai, Zhao Yuzhen, Ma Kai, et al. Friction behavior and wear protection ability of selected base lubricants[J]. *Friction*, 2016, 4(1): 72–83. doi: 10.1007/s40544-016-0106-x.