

DOI: 10.13382/j.jemi.B2205746

基于 OOB-GWO-SVR 的风电机组齿轮箱故障预警*

刘杰 曹静 赵昕

(沈阳工业大学机械工程学院 沈阳 110870)

摘要:针对风电机组齿轮箱超温出现的故障问题,提出了基于改进参数优化机器学习算法的风电机组齿轮箱故障预警模型。首先,通过随机森林袋外估计确定特征变量,并采用滑动平滑滤波对输入变量进行滤波处理。其次,构建灰狼算法优化支持向量回归模型,根据最优模型输出的偏差值确定状态识别指标。最后,通过时移滑动窗口设置阈值范围,当状态识别指标超出阈值范围之外时立即报警。实验结果表明,该模型能提前 87 min 对风电机组齿轮箱温度异常发出故障预警,并且预警效果优于距离相关系数-GWO-SVR 模型、Pearson-GWO-SVR 模型和 OOB-SVR 模型。

关键词:风电机组;齿轮箱;故障预警;特征提取;智能算法;支持向量回归

中图分类号: TM315 **文献标识码:** A **国家标准学科分类代码:** 480.6040

Wind turbine gearbox fault warning based on OOB-GWO-SVR

Liu Jie Cao Jing Zhao Xin

(School of Mechanical Engineering, Shenyang University of Technology, Shenyang 110870, China)

Abstract: Aiming at the fault problem of wind turbine gearbox overtemperature, a fault early warning model of wind turbine gearbox based on improved parameter optimization machine learning algorithm is proposed. Firstly, the characteristic variables are determined by random out-of-the-pocket estimation, and the input variables are filtered by sliding smoothing filtering. Secondly, the gray wolf algorithm optimization support vector regression model is constructed, and the state identification index is determined according to the residual value of the optimal model output. Finally, the threshold range is set by the time-lapse sliding window, and the alarm is immediately issued when the status identification indicator exceeds the threshold range. Experimental results show that the model can issue a fault warning for the temperature abnormality of the wind turbine gearbox 87 minutes in advance, and the early warning effect is better than that of the distance correlation coefficient-GWO-SVR model, Pearson-GWO-SVR model and OOB-SVR model.

Keywords: wind turbines; gearbox; fault warning; feature extraction; intelligent algorithms; SVR

0 引言

风能是一种重要的绿色能源,已被世界各国广泛利用。随着风电装机容量的不断增加,风电场的运行和维护成本也随之上升,风电机组关键部件中最易损坏是齿轮箱,一旦齿轮箱损坏,就会出现维修成本高、维修过程复杂等问题^[1]。因此,在风机齿轮箱失效之前,对其进行有效的识别和早期的故障预警,可降低经济损失,延长风机的使用寿命。

目前,已有相关文献基于 BP 神经网络^[2]、极限学习机^[3]、贝叶斯方法^[4]、支持向量回归^[5]等对风电机组齿轮箱进行评估。但是,BP 神经网络参数多、学习速度慢;极限学习机的非线性能力差并且预测结果不稳定;贝叶斯方法需要有大量样本支撑,不能满足实际需求;支持向量回归(support vector regression, SVR)由于结构风险最小化原理具有较好的泛化能力,使得它在预测问题中发挥着重要的作用。SVR 的性能受到核函数和惩罚因子的影响,所以在实施过程中对二者进行优化是关键环节。传统的 SVR 参数优化方法包括交叉验证、梯度下降、网格

收稿日期: 2022-08-12 Received Date: 2022-08-12

* 基金项目:辽宁省教育厅(LQGD2020016)项目资助

搜索等。乔海升等^[6]采用交叉验证方法优化 SVR, 获得预测效果和泛化能力较好; 莫赞等^[7]将梯度下降法和量子遗传算法相结合解决支持向量机参数进行寻优问题; 文献[8-9]使用网格搜索法优化后的 SVR 算法提高预测精度, 同时大幅减少了参数寻优的时间。但是交叉验证方法存在运算工作量大, 难以完成多个参数的选择; 梯度下降法对初始值敏感, 并不能保证找到全局最优解; 网格搜索法目标函数过多或者取值的范围较大时, 也需要大量的计算时间。

部分学者也使用群智能算法对 SVR 的重要参数进行优化, 如粒子群优化算法、灰狼算法、鲸鱼算法等。文献[10-13]提出群智能算法对支持向量回归机的参数优化, 有效提高学习参数选择效率和预测精确度。上述文献在不同程度上改善了 SVR 的预测精度, 但仍然存在搜索效率低、收敛速度差、全局优化能力差、容易陷入局部最优的问题。灰狼算法^[14-15]相较于其他寻优算法具有参数更少、收敛性速度快的特点和通用性强。

此外, 文献[16]利用皮尔逊相关系数方法选取与目标变量相关性高的特征变量作为输入变量, 但该方法受异常数据的影响较大, 当数据量比较小时, 相关系数波动较大。文献[17-18]采用随机森林算法对参数特征进行降维, 保证了有效的特征选取。随机森林算法在特征选择的方面具有数据选择随机性和运行时间短的优势。

因此, 本文提出一种随机森林袋外估计与灰狼算法优化支持向量回归相结合的风电机组齿轮箱故障预警方法。运用随机森林袋外估计提取特征变量, 并采用滑动平滑滤波对输入变量进行滤波处理。构建灰狼算法优化支持向量回归模型, 基于时移滑动窗口确定阈值的范围实现风电机组齿轮箱故障预警。

1 算法原理

1.1 随机森林袋外估计

在选取输入变量特征时, 并非状态参数特征量越多, 准确度就越高, 过多的特征量会提高模型的复杂性, 造成训练效果较差。由于随机森林可以处理高维数据, 并且计算复杂度低, 使其应用于特征提取方面具有较好的降维效果。

随机森林(random forest, RF)算法是在 2001 年由美国贝尔实验室提出, 属于传统决策树的一种集成学习算法^[19]。一颗决策树的采样过程中, 采样方式为有放回的随机抽样(Bootstrap)。其中, 有部分数据(约 36.8%)未进入决策树的采样中, 把这部分未采样的数据称为袋外数据(out of bag, OOB)。用未赋予随机数的 OOB 数据在训练好的基学习器上运行, 可得到袋外误差; 然后把 OOB 数据样本中的每个特征赋予一个随机数, 得到是加入随

机数的袋外误差。最后通过两种的袋外样本的误差, 建立一个无偏估计评估变量重要性计算公式为:

$$VI = \frac{\frac{1}{n} \sum_{i=1}^n (errOOB'_i - errOOB_i)}{\sqrt{\frac{1}{n} \sum_{i=1}^n (errOOB_i - errOOB_1)^2}} \quad (1)$$

其中, n 为决策数的个数, $errOOB_i$ 是袋外误差, $errOOB'_i$ 是为加入随机数的袋外误差, $errOOB_1$ 是平均袋外样本误差。

1.2 支持向量机回归

支持向量回归是机器学习求解回归问题的一种有效方法^[20]。

设训练集 $T = \{(x_1, y_1), \dots, (x_l, y_l)\} \in (X^n \times Y)^l$, 其中 $x_i \in R^n, y_i \in Y = R, i = 1, \dots, l$ 。

$$\min_{w, b} \frac{1}{2} \|w\|^2 + c \sum_{i=1}^l l_\varepsilon(f(x_i) - y_i) \quad (2)$$

其中, c 为惩罚因子, l_ε 是不敏感损失函数: 当 $y_i - f(x_i) \leq \varepsilon$ 时 l_ε 取值为 0; 当 $y_i - f(x_i) \geq \varepsilon$ 时 l_ε 取值为 $|y_i - f(x_i)| - \varepsilon$ 。

引入松弛变量 ξ_i^* 和 ξ_i , 不敏感系数 $\varepsilon > 0$, 则原始最优化问题可转化为凸二次规划问题。

$$\begin{aligned} \min_{w, b, \xi, \xi^*} & \frac{1}{2} \|w\|^2 + c \cdot \sum_{i=1}^l (\xi_i + \xi_i^*) \\ \text{s. t.} & \begin{cases} f(x_i) - y_i \leq \varepsilon + \xi_i \\ y_i - f(x_i) \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, i = 1, \dots, l \end{cases} \end{aligned} \quad (3)$$

为解决线性不可分问题, 通常方法是寻找核函数。把低维空间中线性不可分的样本集变成线性可分的, 在这个过程中需要核函数, 它能够使得一个低维的样本集向高维做映射。

引入拉格朗日乘子 $\alpha_i \geq 0$ 、核函数 $K(x_i, x_j)$ 和惩罚因子 $c > 0$, 则原始最优化问题的对偶问题如下:

$$\begin{cases} \min_{\alpha_i} & \frac{1}{2} \sum_{i=1}^N y_i y_j \alpha_i \alpha_j K(x_i, x_j) - \sum_{i=1}^N \alpha_i \\ \text{s. t.} & \sum_{i=1}^N \alpha_i y_i = 0, 0 < \alpha_i < c, i = 1, 2, \dots, N \end{cases} \quad (4)$$

支持向量回归算法中核函数选取高斯核函数为:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (5)$$

其中, σ 为核参数, 因此, 为使得风电机组齿轮箱油温预测模型更加精确, 必须要选择合适的参数。

1.3 灰狼算法

灰狼算法(grey wolf optimizer, GWO)是一种优化搜索方法, 其灵感来自于灰狼的捕食行为。灰狼捕食猎物的过程分狼群划分、围剿阶段、捕猎阶段和攻击阶段 4 个

阶段。在构建 GWO 模型的过程中,按照适应度值由高到低的顺序,将灰狼分为 4 个等级: α 、 β 、 δ 、 ω 。

首先,狼群找到目标猎物,然后包围目标猎物:

$$X(t+1) = X_p(t) - A \times |C \times X_p(t) - X(t)| \quad (6)$$

式中: t 为当前迭代; A 和 C 为系数向量; X_p 为目标猎物的位置向量; X 为当前灰狼的位置向量。

$$\begin{aligned} A &= 2 \times a \times r_1 - a \\ C &= 2 \times r_2 \end{aligned} \quad (7)$$

式中: r_1 和 r_2 为随机向量,在 $[0, 1]$ 范围内随机取值; a 为收敛因子,随着迭代次数从 2 线性减少到 0 的距离控制系数。

$$a = 2 - 2t/t_{\max} \quad (8)$$

式中: i 为当前迭代次数; t_{\max} 为最大迭代次数。

其次,由 α 、 β 和 δ 狼领导,灰狼 ω 自动位置更新并进行猎捕。

$$D_i(t) = |C \times X_i(t) - X(t)| \quad (9)$$

$$X_i(t+1) = X_i(t) - A \times D_i(t) \quad (10)$$

$$X(t+1) = \sum X_i(t) / 3 \quad (11)$$

式中: i 为依次表示 α 、 β 和 δ ; $X(t+1)$ 为位置更新后的潜在最优解; D_i 为剩余的个体狼与 α 狼、 β 狼和 δ 狼之间的距离。

最后,灰狼群体的攻击行为主要通过 A 来实现。当 $|A| > 1$ 时,表现灰狼在整个区域是分散的寻找猎物,希望找到更合适的猎物。当 $|A| < 1$ 时,表现灰狼会集中精力在某一区域寻找猎物,狼群向猎物发起攻击。

1.4 灰狼算法优化支持向量回归

SVR 算法中惩罚因子 c 和核参数 σ 的数值选取能直接决定 SVR 算法的预测精度。因此,本文采用 GWO 算法对支持向量机回归参数进行优化^[21-22]。

为了验证 GWO 算法的优势,本文选取了蝴蝶优化算法(butterfly optimization algorithm, BOA)和粒子群算法(particle swarm optimization, PSO)进行比较。通过单峰、多峰和定维多峰 3 种基准测试函数对 3 种算法进行测试,算法的输入参数如表 1 所示。

表 1 输入参数

Table 1 Input parameters

函数	维度	搜索范围	种群数量	迭代次数
$F_1(x)$	30	$[-10, 10]$	30	500
$F_2(x)$	30	$[-32, 32]$	30	500
$F_3(x)$	2	$[-65, 65]$	30	500

单峰基准测试函数为:

$$F_1(x) = \sum_{i=1}^n |x_i| + \prod_{i=1}^n |x_i| \quad (12)$$

多峰基准测试函数为:

$$F_2(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10] \quad (13)$$

定维多峰基准测试函数为:

$$F_3(x) = \left(\frac{1}{500} + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=2}^2 (x_i - a_{ij})^6} \right)^{-1} \quad (14)$$

式中: x 为基准函数变量, a_{ij} 为参数变量。

由表 2 可知,GWO 算法对 3 种验证函数平均适应度值均达到最低,其中对 $F_1(x)$ 优化结果较 PSO 算法提高 15 个数量级; $F_2(x)$ 优化结果较 PSO 算法提高 23 个数量级; $F_3(x)$ 优化结果较 BOA 算法提高 183 个数量级。

表 2 各算法平均适应度

Table 2 The average fitness of different algorithms

	BOA	GWO	PSO
$F_1(x)$	2.19×10^{-8}	1.01×10^{-16}	1.44×10^{-1}
$F_2(x)$	7×10^{-11}	1.34×10^{-28}	3.97×10^{-5}
$F_3(x)$	1.96×10^{-14}	5.38×10^{-197}	1.53×10^{-38}

适应度标准差值越小,证明其稳定性越高。分析表 3 可见,在多数情况下 GWO 算法的标准差适应度值最低,所以 GWO 优化稳定度最高,其中 GWO 算法对 $F_1(x)$ 的优化结果较 PSO 算法提高 16 个数量级;对 $F_2(x)$ 的优化结果较 PSO 算法提高 22 个数量级。

表 3 各算法标准差适应度

Table 3 The standard fitness of different algorithms

	BOA	GWO	PSO
$F_1(x)$	6.84×10^{-9}	9.95×10^{-17}	3.18×10^{-1}
$F_2(x)$	6.90×10^{-12}	1.85×10^{-28}	1.32×10^{-4}
$F_3(x)$	1.39×10^{-13}	0	6.36×10^{-38}

综合图 1、2 和 3 的 PSO、GWO 和 BOA 算法的寻优曲线及测试运行 50 次的标准差适应度和平均适应度结果可知,GWO 算法相比 BOA 算法和 PSO 算法,寻优能力最好且最稳定,因此,采用 GWO 算法对支持向量回归模型进行参数优化。

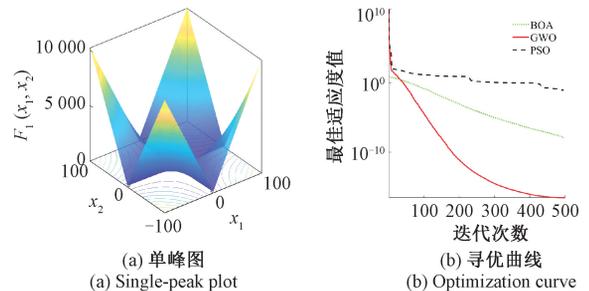


图 1 单峰基准测试函数

Fig. 1 Single-peak benchmark function

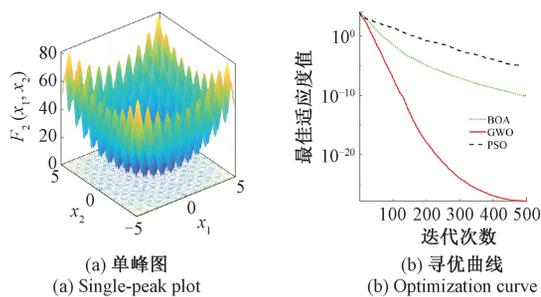


图 2 多峰基准测试函数

Fig. 2 Multimodal benchmark function

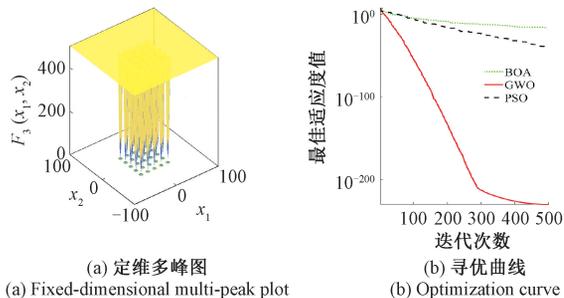


图 3 定维多峰基准测试函数

Fig. 3 Fixed-dimensional multimodal benchmark function

本文以齿轮箱油温为研究对象,利用随机森林袋外估计与灰狼算法优化支持向量回归模型对齿轮箱故障进行预警,研究内容如图 4 所示。

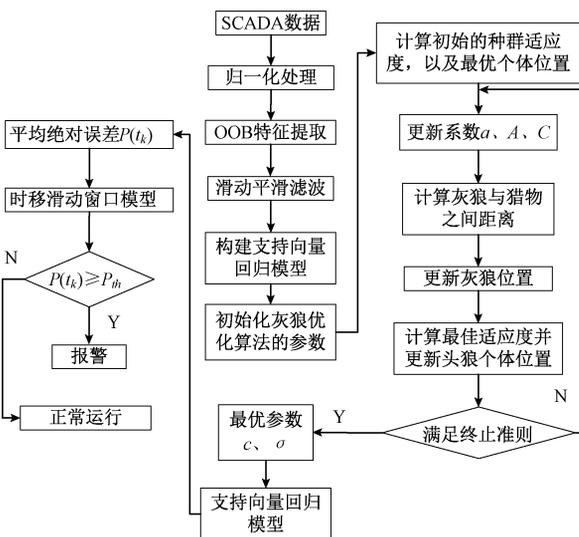


图 4 风电机组齿轮箱故障预警流程

Fig. 4 Wind turbine gearbox fault early warning process

1) 运用随机森林袋外估计算法选择与齿轮箱油温相关性高的特征变量;采用滑动平滑滤波对输入变量进行滤波处理。

2) 设置灰狼算法的种群数量、优化参数的个数、参数取值范围和最大迭代次数。

3) 初始化 SVR 的参数,计算适应度值,使 SVR 算法每次迭代的均方根误差 (mean square error, MSE) 最小。

4) 根据灰狼个体的适应度值确定等级,然后更新其位置信息;

5) 根据 α, β, δ 的位置信息更新 ω 的位置;

6) 更新 a, A, C , 确定新的 X_α, X_β 和 X_δ ;

7) 当迭代次数满足要求时停止迭代,获得最优参数 c 和 σ , 否则跳至 6);

8) 得到最优参数,用于模型预测;

9) 基于时移滑动窗口模型对 GWO-SVR 模型的齿轮箱油温偏差值进行分析,确定预警阈值。

2 实例验证

2.1 数据描述

某风场于 2017 年 7 月 1 日 14:00 发生故障,选取故障前 24 h 的数据采集与监视控制系统 (supervisory control and data acquisition, SCADA) 监测数据。数据采样间隔为 1 min,共采样 1 440 组数据,风电机组状态参数为 20,如表 4 所示。

表 4 风电机组状态参数

Table 4 Wind turbine status parameters

编号	状态参数	单位	编号	状态参数	单位
0	齿轮箱油温	℃	10	发电机 v1 绕组温度	℃
1	齿轮箱前轴温度	℃	11	发电机 v2 绕组温度	℃
2	齿轮箱后轴温度	℃	12	发电机 w1 绕组温度	℃
3	齿轮箱进口油温	℃	13	发电机 w2 绕组温度	℃
4	变频器温度	℃	14	发电机转速	r/min
5	发电机前轴承温度	℃	15	机舱 X 方向振动	$m \cdot s^{-2}$
6	发电机后轴承温度	℃	16	机舱 Y 方向振动	$m \cdot s^{-2}$
7	发电机冷却水入口温度	℃	17	机舱前端温度	℃
8	发电机 u1 绕组温度	℃	18	塔基控制柜温度	℃
9	发电机 u2 绕组温度	℃	19	偏航方位角	(°)

2.2 模型参数选择

由于风电机组各个状态参数的单位不统一,将其进行归一化处理:

$$x'_{ij} = \frac{x_{ij} - \min\{x_{ij}\}}{\max\{x_{ij}\} - \min\{x_{ij}\}} \quad (15)$$

其中, x_{ij} 是第 $j(j=1, 2, \dots, m)$ 个状态参数在第 $i(i=1, 2, \dots, n)$ 时刻的初始值; x'_{ij} 是数据归一化后的值。

结合 1.1 节随机森林袋外估计进行特征选择的方法。对上述的状态参数进行筛选,选取能反映齿轮箱油温变化的状态参数作为输入变量。

随机森林袋外估计算法主要两个参数为叶子节点数和决策树的数量。对这两个参数进行寻优,达到最佳的

状态特征参数的提取。采用均方差作为随机森林袋外估计寻优的评价指标。如图 5 和 6 所示。

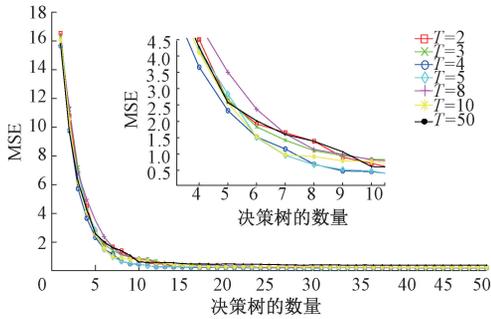


图 5 叶子节点数寻优

Fig. 5 The number of leaf nodes is optimal

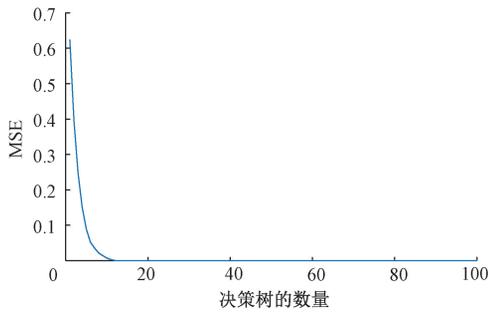


图 6 决策树的数量寻优

Fig. 6 The number of decision trees is optimal

通过图 5 可得叶子节点数的选择对状态特征提取影响不大,但是叶子节点数为 5 收敛速度最快,因此叶子节点数为 5。由图 6 可知,当决策树数量 $CART < 10$ 时, MSE 下降速度很快,但决策树数量 $CART > 20$ 时, MSE 基本稳定,因此决策树选为 20 棵。

通过上述对随机森林袋外估计算法中两个主要参数进行寻优选择,得出最优参数是叶子节点数为 5,决策树的数量为 20,作为参数输入可得到状态特征参数的重要性排序图,如图 7 所示。

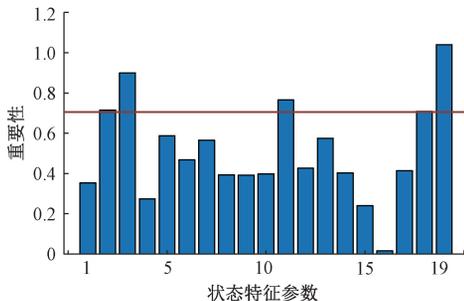


图 7 状态特征参数的重要性

Fig. 7 The importance of the state characteristic parameter

选取前 4 个的状态特征参数为齿轮箱后轴温度、齿轮箱进口油温、发电机 v2 绕组温度和偏航方位角,构建特征集。

2.3 滑动平滑滤波

本文使用具有相同窗口大小的移动平均值生成具有相同平滑度的多个新时间序列。

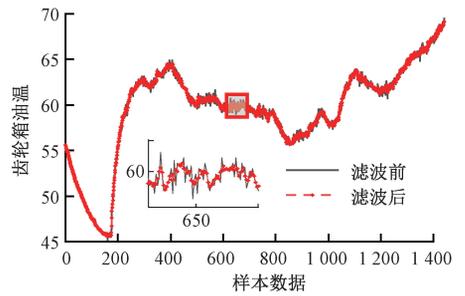
对于窗口大小为 l 的滤波器,对第 i 个数据到第 $i=l-1$ 的齿轮箱油温数据进行滤波,滤波处理如式 (16) 所示,原时间序列如式 (17) 所示,那么滤波处理后的数据如式 (18) 所示:

$$\bar{t}_{i+l-1} = \frac{t_i + t_{i+1} + \dots + t_{i+l-1}}{l} \tag{16}$$

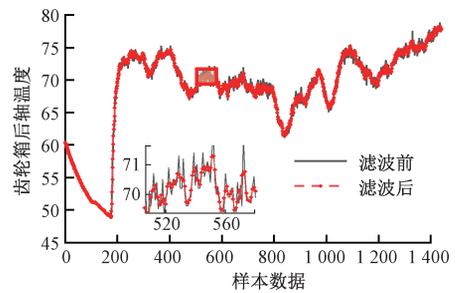
$$T = \{t_1, t_2, \dots, t_i, t_{i+1}, \dots, t_n\} \tag{17}$$

$$T' = \{t_1, t_2, \dots, \bar{t}_{i+l-1}, \dots, t_n\} \tag{18}$$

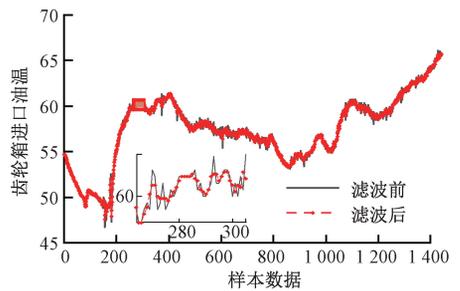
通过滑动平滑滤波,对所有特征参数进行滤波处理,滤波器的窗口大小为 3,对 1 440 组数据进行滤波处理,滤波处理如图 8 所示。



(a) 齿轮箱油温
(a) Gearbox oil temperature



(b) 齿轮箱后轴温度
(b) Gearbox rear axle temperature



(c) 齿轮箱进口油温
(c) Gearbox inlet oil temperature

基于随机森林袋外估计的状态特征参数选择,最终

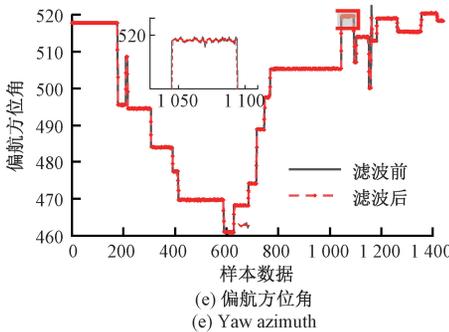
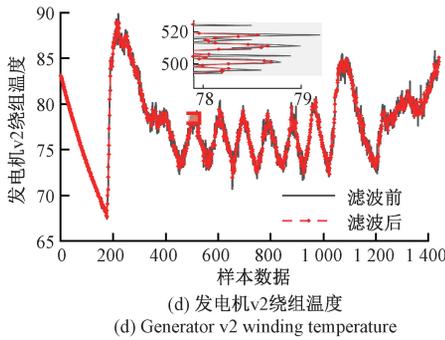


图 8 特征参数

Fig. 8 Feature parameters

2.4 GWO-SVR 模型建立

通过对比分析风电机组齿轮箱油温预测值曲线与真实值曲线故障前的变化趋势,得出风电机组齿轮箱油温异常状态。当风电机组齿轮箱正常工作时,模型对齿轮箱油温数据的预测值与真实值应基本重合;在齿轮箱临近发生故障时,预测的齿轮箱油温曲线与真实齿轮箱油温曲线会在某一时间点发生明显的偏离趋势。该偏离趋势表示为齿轮箱油温预测值与实际值的残差波动,当偏差波动范围超出一定范围,则对风电机组齿轮箱的工作状况做出故障警示。

设模型输入向量为 $X = \{x_1, x_2, x_3, x_4\}$, $Y = \{y_1\}$ 为输出变量,齿轮箱后轴温度、齿轮箱进口油温、发电机 v2 绕组温度、偏航方位角和齿轮箱油温依次作为模型的输入向量变量与输出向量。

将总样本的 80% (1 152 组数据) 作为训练集建立 GWO-SVR 模型。狼群算法中参数设置为种群大小为 5, 迭代次数为 200, 搜索空间为 $[0, 20]$ 。通过灰狼算法对 SVR 的惩罚因子 c 和核函数参数 σ 进行寻优, 得到最优的预测模型。

如图 9(a) 为 SVR 模型中两个重要参数对预测结果的影响图, 将预测值和实际值之间的差异定义为 $loss$ 。 $loss$ 越小, 差异越小, 相对来说模型预测的效果就越好。如图 9(b) 为模型适应度值图, 当迭代次数达到 100 时, 适应度值趋于平稳, GWO-SVR 模型训练集精度达到

99.93%。通过图 9 可得出 SVR 算法中惩罚因子 c 为 19.87 和核函数参数 σ 为 20。

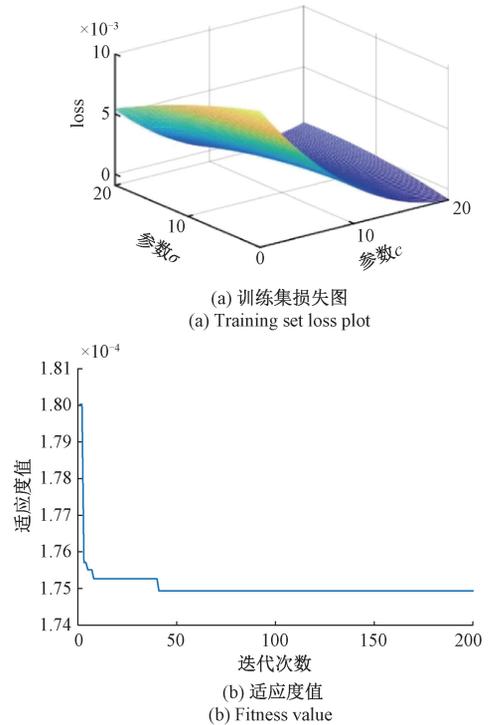


图 9 模型训练结果

Fig. 9 Model training results

经过模型的训练和参数调优, 可得到最优的 GWO-SVR 预测模型, 选取剩余 20% (288 组) 数据输入到惩罚因子为 19.87 和核函数参数为 20 的 GWO-SVR 模型中进行验证, 由图 10 可知 GWO-SVR 模型输出的齿轮箱油温预测值与真实值的偏差图。

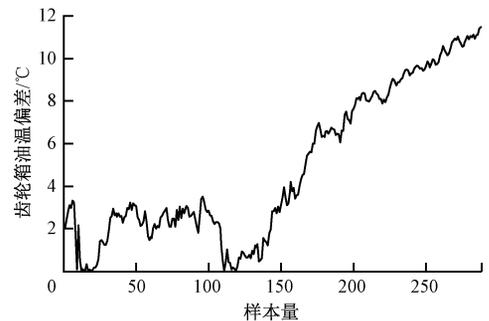


图 10 模型验证结果

Fig. 10 Validation set predictions

2.5 时移滑动窗口预警

建立滑动窗口模型, 从而实现风电机组齿轮箱油温异常状态预警^[23], 其原理如图 11 所示。设滑动窗口宽度为 h , 时间增量为 Ω , 第 k 时刻的窗口内输入 $t_{k-h} \sim t_k$ 之间的数据, 即 S_k 。

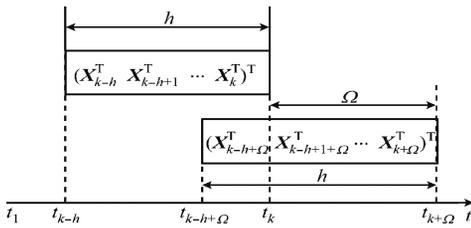


图 11 时移滑动窗口模型

Fig. 11 Slide the time window model

采用平均绝对误差 $P(t_k)$ 作为状态识别指标,其表达式如下:

$$P(t_k) = \frac{1}{h} \sum_{i=t_{k-h}}^{t_k} |E'_i - E_i| \quad (19)$$

式中: E'_i 为预测模型输出预测值; E_i 为样本的实际值; t 为当前时刻。

基于统计学中的区间估计理论,无需选取置信下限,只需选取置信上限 P_{th} , 区间 $[0, P_{th}]$ 为状态指标 P 置信度为 $1-\alpha$ 的置信区间。置信上限 P_{th} 采用式(19)计算:

$$P_{th} = \bar{P} + \frac{S_{std}}{\sqrt{k}} t_{\alpha}(k-1) + 3S_{std} \quad (20)$$

式中: \bar{P} 为状态识别指标的均值; S_{std} 为状态识别指标的标准差; k 为滑动时间窗口的个数; t_{α} 为随着置信度 $1-\alpha$ 取值不同而进行改变的小概率值。

为验证所提出预警方法的有效性,将 GWO-SVR 模型输出的齿轮箱油温偏差作为时移滑动窗口模型的样本量,通过对齿轮箱油温偏差数据进行时移滑动确定出状态识别指标。

风电机组齿轮箱在正常状态工作时,平均绝对误差曲线应该低于阈值上限。当齿轮箱油温偏差数据的平均绝对误差均超过阈值上限,立即发出故障报警。选取滑动时间窗口宽度为 5; 时间增量为 1; 置信度为 0.000 5。通过式(20)可知齿轮箱油温的预警阈值为 7.677 3℃。

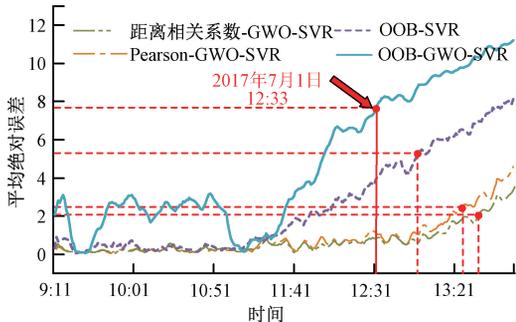


图 12 基于时移滑动窗口的模型对比

Fig. 12 Model comparison chart based on time-shifted sliding windows

将距离相关系数-GWO-SVR 模型、Pearson-GWO-

SVR 模型、OOB-SVR 模型与本文 OOB-GWO-SVR 模型的预警时间进行对比。如图 12 可知,距离相关系数-GWO-SVR 模型的预警时间为 2017 年 7 月 1 日 13:38,距风电机组齿轮箱发生故障大约早 23 min; Pearson-GWO-SVR 模型的预警时间为 13:27,距风电机组齿轮箱发生故障也大约早 33 min; OOB-SVR 模型的预警时间为 12:59,距风电机组齿轮箱发生故障大约早 61 min; 本文使用的 OOB-GWO-SVR 模型的预警时间为 12:33,距风电机组齿轮箱发生故障大约早 87 min。

3 结论

本文提出一种基于 OOB-GWO-SVR 的风电机组齿轮箱故障预警的方法,实现了风电机组齿轮箱超温故障报警并得出如下结论:

1) 随机森林袋外估计算法不需要单独划分交叉验证集,直接使用袋外数据验证特征变量的重要性,并建立 VI 作为评估特征变量重要性指标,其中对算法的叶子节点数与决策树数量进行寻优处理,可提取到最佳的特征向量。

2) 通过将 GWO 算法与 BOA 算法和 PSO 算法进行比较分析,得出 GWO 算法的稳定性与收敛速度最优,因此,采用 GWO 算法对 SVR 模型中惩罚因子和核函数宽度系数进行优化处理,有效地提高了预测模型的准确率。

3) 本文模型与 OOB-SVR, Pearson-GWO-SVR 和距离相关系数-GWO-SVR 相比较能够提前发现风电机组齿轮箱早期故障,可以有效地降低风电机组齿轮箱出现停机带来的损失。

参考文献

[1] FU J, CHU J C, GUO P, et al. Condition monitoring of wind turbine gearbox bearing based on deep learning model[J]. IEEE Access, 2019, 7: 7078-75087.

[2] 董开松,李涛涛,尹浩霖. 风力发电机组故障分析与智能诊断[J]. 高压电器, 2016, 52(10): 176-181.

[3] 李辉,邓奇. 基于改进 LMD 方法的风电机组齿轮箱故障诊断研究[J]. 自动化仪表, 2021, 42(3): 60-65.

[4] LI H, DENG Q. Research on gearbox fault diagnosis of wind turbine based on improved LMD method [J]. Automation Instrumentation, 2021, 42(3): 60-65.

[5] SONG ZH, ZHANG Z, JIANG Y, et al. Wind turbine health state monitoring based on a Bayesian data-driven approach[J]. Renewable Energy, 2018, 125.

[6] 王新. 基于 SCADA 数据的风电机组轴承故障预警[D]. 重庆:重庆交通大学, 2020.

- WANG X. Early warning of bearing failure of wind turbines based on SCADA data [D]. Chongqing : Chongqing Jiaotong University, 2020.
- [6] 乔海升,赵文升. 基于支持向量回归和聚类算法的主蒸汽压力优化[J]. 广东电力,2022,35(7): 72-78.
- QIAO H SH, ZHAO W SH. Principal vapor pressure optimization based on support vector regression and clustering algorithms [J]. Guangdong Electric Power, 2022, 35(7): 72-78.
- [7] 莫赞,刘希良,谢海涛. 基于混合梯度下降算法的支持向量机参数优化[J]. 网络新媒体技术,2014,3(2): 54-58.
- MO Z, LIU X L, XIE H T. Support vector machine parameter optimization based on hybrid gradient descent algorithm[J]. Network New Media Technology, 2014, 3(2): 54-58.
- [8] 杜聪,邵建华,杨薇,等. 网格搜索法优化的支持向量机室内可见光定位[J]. 激光杂志,2021,42(3): 104-109.
- DU C, SHAO J H, YANG W, et al. The grid search method optimizes the indoor visible light localization of the vector machine[J]. Laser Magazine, 2021, 42(3): 104-109.
- [9] 李森娟,张萍,岳大为,等. 基于支持向量机的风电机组故障预测[J]. 计算机仿真,2022,39(5): 84-88,180.
- LI S J, ZHANG P, YUE D W, et al. Failure prediction of wind turbines based on support vector machines[J]. Computer Simulation, 2022, 39(5): 84-88,180.
- [10] 高峰,邓星星,刘强,等. 大型风电机组电动变桨系统变桨角度故障诊断[J]. 太阳能学报,2020,41(5): 98-106.
- GAO F, DENG X X, LIU Q, et al. Pitch angle fault diagnosis of electric pitch system for large wind turbines[J]. Journal of Solar Energy, 2020, 41(5): 98-106.
- [11] 余畅文,潘万宝,刘练,等. 基于斑点鬣狗算法优化支持向量机的短期风电功率预测[J]. 电工技术,2022(15): 4-6.
- YU CH W, PAN W B, LIU L, et al. Short-term wind power prediction of support vector machine is optimized based on spotted hyena algorithm[J]. Electrotechnical Technology, 2022(15): 4-6.
- [12] 王维高,魏云冰,滕旭东,等. 基于麻雀搜索优化支持向量机的短期风机发电功率预测[J]. 智能计算机与应用,2022,12(1): 119-123.
- WANG W G, WEI Y B, TENG X D, et al. Short-term wind turbine power prediction based on sparrow search optimization support vector machine [J]. Intelligent Computers and Applications, 2022, 12(1): 119-123.
- [13] 梁迪,郭启航,姜廷霖. 布谷鸟搜索算法优化支持向量机的停车位预测[J]. 沈阳大学学报(自然科学版), 2021,33(3): 234-239,266.
- LIANG D, GUO Q H, JIANG T L. The cuckoo search algorithm optimizes the parking space prediction of the support vector machine [J]. Journal of Shenyang University(Natural Science Edition), 2021, 33(3): 234-239,266.
- [14] 胡璇,李春,叶柯华. 灰狼算法优化支持向量机在风力机齿轮箱故障诊断中的应用[J]. 机械强度,2021,43(5): 1026-1034.
- HU X, LI CH, YE K H. Gray wolf algorithm optimizes the application of support vector machine in wind turbine gearbox fault diagnosis[J]. Mechanical Strength, 2021, 43(5): 1026-1034.
- [15] 熊军华,师刘俊,康义. 基于灰狼算法优化支持向量机的变压器故障诊断[J]. 信息技术与信息化,2020(11): 144-147.
- XIONG J H, SHI L J, KANG Y. Optimize transformer fault diagnosis of support vector machine based on gray wolf algorithm [J]. Information Technology and Informatization, 2020(11): 144-147.
- [16] ZI L, LIU X L. Wind power forecasting of an offshore wind turbine based on high-frequency SCADA data and deep learning neural network [J]. Energy, 2020, 201: 117693.
- [17] 刘杰,付雪娇,蒋树旗,等. 基于 RF-LSTM 网络的风电机组状态参数预测[J/OL]. 控制工程:1-6[2022-10-20].
- LIU J, FU X J, JIANG SH Q, et al. Prediction of wind turbine status parameters based on RF-LSTM network[J/OL]. Control Engineering: 1-6[2022-10-20].
- [18] 常梦容,王海瑞,肖杨. mRMR 特征筛选和随机森林的故障诊断方法研究[J]. 电子测量与仪器学报,2022,36(3):175-183.
- CHANG M R, WANG H R, XIAO Y. Research on mRMR feature screening and fault diagnosis method of random forest[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(3): 175-183.
- [19] 张航,史兆培,束垠,等. 基于 OOB-BO-LightGBM 的风电机组故障诊断方法[J/OL]. 中国测试:1-7[2022-07-16].
- ZHANG H, SHI ZH P, SHU Y, et al. Fault diagnosis method for wind turbine based on OOB-BO-LightGBM[J/OL]. Tested in China: 1-7[2022-07-16].
- [20] ZHANG L W, ZHOU Z D, XU Y F, et al. Classification

- of imagined speech EEG signals with DWT and SVM[J]. Instrumentation, 2022, 9(2): 56-63.
- [21] YANG Z S, WANG Y H, KONG C Z. Remaining useful life prediction of lithium-ion batteries based on a mixture of ensemble empirical mode decomposition and GWO-SVR model[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-11.
- [22] 胡璇,李春,叶柯华. 灰狼算法优化支持向量机在风力机齿轮箱故障诊断中的应用[J]. 机械强度, 2021, 43(5): 1026-1034.
- XU X, LI CH, YE K H. The gray wolf algorithm optimizes the application of support vector machine in the fault diagnosis of wind turbine gearbox[J]. Mechanical Strength, 2021, 43(5): 1026-1034.
- [23] 马良玉,孙佳明,於世磊,等. 基于 DBSCAN 和 SDAE 的风电机组异常工况预警研究[J]. 动力工程学报, 2021, 41(9): 786-793, 808.

MA L Y, SUN J M, WU SH L, et al. Research on early warning of abnormal working conditions of wind turbines based on DBSCAN and SDAE[J]. Journal of Power Engineering, 2021, 41(9): 786-793, 808.

作者简介



刘杰(通信作者), 2002 年于辽宁科技大学获得学士学位, 分别于 2005 年和 2008 年于东北大学获得硕士学位和博士学位, 现为沈阳工业大学副教授, 主要研究方向为设备智能运维及健康管理。

E-mail: starliujie@126.com

Liu Jie (Corresponding author) received her B. Sc. degree from University of Science and Technology Liaoning in 2002 and M. Sc. and Ph. D. degrees from Northeastern University in 2005 and 2008. Now she is an associate professor in Shenyang University of Technology. Her main research interests include intelligent operation and health management of equipment.