

基于改进凌日搜索算法的风洞天平载荷预测方法

王碧玲^{1,2}, 周 灏³, 沈力华³, 刘博宇^{1,2}, 王 奔³

(1. 中国航空工业空气动力研究院, 沈阳 110034;

2. 高速高雷诺数气动力航空科技重点实验室, 沈阳 110034;

3. 沈阳航空航天大学 机电工程学院, 沈阳 110136)

摘要: 风洞天平是在风洞测试中使用的测力传感器, 在使用之前需要进行校准以测量缩比模型受到的气动载荷; 传统的方法使用预设的多项式函数进行拟合, 忽略了某些变量的存在对测量载荷的负面影响, 导致数据处理结果的失真; 在此, 提出了一种改进的凌日搜索算法 (ITS), 选择对测量载荷更具重要性的特征; 然后, 使用贝叶斯线性回归算法 (BLR) 建立预测模型测量天平载荷, 最后, 在两个天平数据集上测试了该方法, 结果表明 ITS-BLR 方法评估确定了对预测目标具有较高贡献的特征, 进而降低了预测误差, 与最小二乘法得到的综合加载误差相比, 降幅最高达到 60%, 说明提出的方法可以实现对天平载荷的准确预测。

关键词: 风洞天平; 载荷预测; 优化算法; 特征选择; 凌日搜索

A Wind Tunnel Balance Load Prediction Method Based on ITS Algorithm

WANG Biling^{1,2}, ZHOU Hao³, SHEN Lihua³, LIU Boyu^{1,2}, WANG Ben³

(1. AVIC Aerodynamics Research Institute, Shenyang 110034, China;

2. Aviation Key Laboratory of Science and Technology of High Speed and High Reynolds, Shenyang 110034, China;

3. School of Mechatronics Engineering, Shenyang Aerospace University, Shenyang 110136, China)

Abstract: A wind tunnel balance is a force sensor used in wind tunnel testing, which needs to be calibrated to measure the aerodynamic load on the scale model before use. Traditional methods use preset polynomial function for fitting, ignoring the negative influence of some variables on the measured load, which leads to the distortion of data processing results. Based on this, an improved transit search (ITS) algorithm is proposed to select more important features for load measurement. Then, Bayesian linear regression algorithm (BLR) is used to build a prediction model to measure the balance load. Finally, the method is verified on two balance datasets. The results show that the ITS-BLR method evaluates and identifies features with a higher contribution to the prediction target, thereby reducing the prediction error. Compared with the comprehensive loading error that obtained by the least squares method, the error of the ITS-BLR method is reduced up to 60%, which shows that the proposed method can provide an accurate prediction for the balance load.

Keywords: wind tunnel balance; load prediction; optimization algorithm; feature selection; transit search

0 引言

风洞天平是一种单分量或多分量的测力传感器, 广泛应用于风洞试验过程的气动力测量^[1]。一般的六分量天平通过应变片检测结构变形产生的电压信号来测量作用在模型上的法向力、侧向力、轴向力、俯仰力矩、偏航力矩和滚转力矩^[2]。天平校准就是通过校准设备建立天平测量信号与气动载荷的映射关系^[3-4], 从而实现准确的载荷预测, 载荷预测的质量直接决定了风洞试验数据结果的质量。

风洞天平是针对不同的风洞测试要求和环境而专门设计和定制的, 所以需要采用适当的数据处理方法和特征形式, 确保准确预测载荷。传统的天平校准载荷预测使用多

项式拟合校准数据, 这时, 多项式中的变量被视为回归问题中的特征^[5]。Li^[6]指出, 需要提升模型性能时, 获取与预测目标更相关的特征往往能更直接有效地提高模型预测性能。然而, 实际问题中存在一些对目标影响较弱或不相关的特征, 直接加入模型反而可能降低预测性能或引入噪声^[7], 根据天平设计原理^[8], 其他测量元对待测分量会产生一定干扰效应, 为了表征不同分量间的影响需要在数据处理时引入交叉项特征^[9], 其中与预测结果的相关性较弱, 甚至是完全无关的特征会对预测结果产生一定的负面影响。因此, 在校准数据处理时进行一定的特征处理策略是必要的。

收稿日期: 2023-09-21; 修回日期: 2023-12-04。

作者简介: 王碧玲(1981-), 女, 博士, 高级工程师。

通讯作者: 沈力华(1984-), 女, 博士, 讲师。

引用格式: 王碧玲, 周 灏, 沈力华, 等. 基于改进凌日搜索算法的风洞天平载荷预测方法[J]. 计算机测量与控制, 2024, 32(11): 25-33.

特征选择是一种常见的数据降维方法，它利用某种方法删除数据集中冗余的特征，找到最优的特征子集^[10]。已有很多学者对校准数据中的特征选择算法进行了研究，主要分为两类，第一类是使用专门的特征选择算法，N. Ulbrich^[11-13]等使用基于前向选择 (FS, forward selection) 与后向消除 (BE, fackward elimination) 的特征选择方法处理风洞尾撑天平的校准数据，Kammeyers^[14]使用逐步回归的方法处理校准数据，通过搜索有利于校准结果的回归项获取校准误差更低的模型。第二类是借助回归算法本身的能力，如马涛^[15]利用回归树模型进行回归任务，而回归树在回归任务中借助本身的性质来实现特征子集的选择。属于嵌入式方法，但作为回归算法需要较多的数据作为模型的训练输入。汪运鹏^[3]使用卷积神经网络的隐式特征提取可获得较好的模型性能，但可能损失结果的可解释性。特征选择是公认的 NP 难题，尤其是对高维数据^[16]。Zakieh^[17]提出一种改进的正弦余弦非洲秃鹰优化算法来选择检测 DDoS 攻击的有效特征，Marjit^[18]等人提出模拟退火辅助遗传算法从高维微阵列的基因表达谱数据中选择基因特征。相较于传统特征选择算法，启发式算法能够更全面地遍历整个解空间，可以有效解决高维数据的特征选择问题^[19-20]。2022 年，Mirjalili^[21]从宇宙观测中获取设计灵感，提出了凌日搜索算法 (TS, transit search)。该寻优算法具有多种搜索阶段，在全局与局部搜索之间取得了良好平衡，与其他启发式算法相比，在多项测试函数实验中表现出良好的性能优势。然而，根据“无免费午餐定理”，不存在单一算法可完美解决所有问题。需要根据问题的不同设计专门方案以提升效果，启发式算法的性能依赖开发和探索之间的平衡。过度开发和较少探索会导致过早收敛，而较大探索和较少开发可能陷入局部最优值。因此，应该尽量避免搜索策略导致重要特征选择不足和较差的预测性能。本文在凌日搜索算法的基础上，通过引入交叉变异操作来扩展种群多样性，以防止结果陷入局部最优解。另外，借鉴模拟退火和变邻域搜索算法的思想高效率地搜索二进制特征，同时，设置控制参数以平衡全局搜索和局部搜索的能力，以提高收敛精度。优化算法的实现需要使用回归算法计算适应度，目前常用的用于风洞天平校准数据处理的算法有支持向量机、BP 神经网络以及卷积神经网络等，但是由于支持向量机需要进行超参数的寻优^[22]，计算量大，BP 神经网络以及卷积神经网络等神经网络依赖初始值^[23-24]，可能产生较高的误差，进而影响优化算法运行结果。同时，筛选特征后得到的天平校准数据集样本量仍然有限，与最小二乘法 (OLS, ordinary least squares) 相比，贝叶斯线性回归 (BLR, bayesian linear regression) 能够更好地防止过拟合问题^[25]，且不需要交叉验证即可自动计算得到超参值，即使在数据有限或者噪声较大的情况下，也能给出统计学意义上的有效结果，更适用于小样本和高方差的样本数据^[26-28]，在优化算法迭代时可以得到更可靠的系数估计，准确计算每个解的适应度。因此，使用贝叶斯线性回归作为载荷预测模型中的回归算法。

综上所述，为提高风洞天平静态校准数据处理的准确度，本文提出一种基于改进的凌日搜索算法 (ITS, improved transit search) 的风洞天平载荷预测方法，通过特征选择技术与贝叶斯模型的引入，提高了天平测力的抗干扰和容错能力，实现天平载荷的准确预测。

1 改进凌日搜索算法

1.1 凌日搜索算法

凌日搜索是一种模拟天体运行状态的启发式算法，具有星系阶段、恒星阶段、凌日阶段、行星阶段、邻居阶段和开发阶段。

1.1.1 星系阶段

首先随机生成一个 d 维向量 L_r 作为星系中心，然后在其周围选取 $n_s \times S_N$ 个 d 维候选向量作为候选星系。

$$L_R = \begin{cases} L_{\text{Galaxy}} + c_1 L_{\text{Galaxy}} - L_r - \text{Noise} & \text{if } z = 1 \\ L_{\text{Galaxy}} + c_1 L_{\text{Galaxy}} + L_r - \text{Noise} & \text{if } z = 2 \end{cases} \quad (1)$$

对每个向量计算适应度函数。最后从中选择值最小的 n_s 个作为初始集再根据式 (2) 进行更新，从而获得 n_s 个 d 维星体的群体。其中， c_1 、 c_2 是一个 0 到 1 之间的随机数， c_3 是 0 到 1 之间的随机向量。

$$L_S = \begin{cases} L_R + c_3 L_R - c_2 L_r - \text{Noise} & \text{if } z = 1 \\ L_R + c_3 L_R + c_2 L_r - \text{Noise} & \text{if } z = 2 \end{cases} \quad (2)$$

1.1.2 搜索阶段

1) Transit 阶段：对星系阶段的初始集按照式 (2) 计进行更新，计算光度值 L_i 。如果新的光度值小于原始光度值，则确定为凌日发生。

$$L_i = \frac{R_i/n_s}{(d_i)^2} \quad i = 1, \dots, n_s \quad (3)$$

d_i 是随机值代表观测点 L_T ，在优化过程中保持不变，用于计算与第 i 个解的距离。

2) 如果凌日发生，执行行星 (Planet) 阶段：多次计算式 (4) 取平均值确定位置， c_4 、 c_5 是 -1 到 1 之间的随机数。此外， c_6 是值在 -1 到 1 之间的随机向量。

$$\begin{aligned} L_z &= (c_4 L_T + R_L L_S) / 2 \quad i = 1, \dots, n_s \\ L_P &= \begin{cases} L_z + c_5 L_r & \text{if } z = 1 \\ L_z - c_5 L_r & \text{if } z = 2 \\ L_z + c_6 L_r & \text{if } z = 3 \end{cases} \end{aligned} \quad (4)$$

3) 如果凌日没有发生，执行邻域 (Neighbor) 阶段：则根据式 (5) 对周围进行邻域搜索，同样地，也对 L_n 取运行 S_N 次的平均值作为最终位置， c_7 是 0 到 1 之间的随机数， c_8 和 c_9 分别是 -1 和 1 之间的随机数和随机向量。

$$\begin{aligned} L_z &= (c_7 L_s + c_7 L_r) / 2 \\ L_P &= \begin{cases} L_z - c_8 L_r & \text{if } z = 1 \\ L_z + c_8 L_r & \text{if } z = 2 \\ L_z + c_9 L_r & \text{if } z = 3 \end{cases} \end{aligned} \quad (5)$$

4) Exploitation 阶段：按照式 (6) 更新式 (4) (5) 的解。

$$L_E = \begin{cases} c_{11}L_p + c_{10}(c_{12})^p L_r & \text{if } c_k = 1 \\ c_{11}L_p - c_{10}(c_{12})^p L_r & \text{if } c_k = 2 \\ L_p - c_{10}(c_{12})^p L_r & \text{if } c_k = 3 \\ L_p + c_{10}(c_{12})^p L_r & \text{if } c_k = 4 \end{cases} \quad (6)$$

c_{10} 是 0 到 2 之间的随机数, c_{11} 是 0 到 1 之间的随机数, c_{12} 是 0 到 1 之间的随机向量, 每轮迭代后, 根据适应度保存最佳星体。

最后, 从所有迭代最佳星体中选择适应度最优者作为最终星体。

从式 (1) ~ (6) 可以看出, TS 算法有 2 个主要参数, n_s 决定了群体规模, S_N 决定了每个解搜索更新的步长, 算法 1 中总结了详细步骤。

Algorithm 1: Pseudocode of the TS Algorithm

1) 输入:

恒星数 n_s , 信噪比 S_N , 迭代次数 n_i

2) 输出:

最优行星位置 L_B 及其对应的适应度 f_B

3) 初始化: 使用式(1)、式(2)进行星系阶段, 返回特征集合 L_s

4) While (Stopping condition is not met) do

5) 使用式(3)进行凌日阶段, 判断是否凌日

6) for $i = 1 : n_s$,

7) 如果检测到凌日则使用式(4)进行行星阶段

8) 否则使用式(5)进行邻域阶段

9) elseif

10) 返回每颗恒星的 L_p 及对应的适应度 f_p

11) 使用式(6)进行开发阶段

12) end

13) 返回每颗恒星的 L_p 及对应的适应度 f_p

14) end

15) 返回最佳行星位置 L_B 及对应的适应度 f_B

1.2 改进凌日搜索算法

对于风洞天平校准数据, 本文的主要特征包括以下 3 类。

1) 一次项特征: 直接利用校准天平分量直接输出值 $P_1 \sim P_n$ 等作为特征, 表示对待测量分量的线性干扰。

2) 变量的一次交叉项特征: 考虑变量间的线性组合, 如 $P_1 * P_2$ 等交叉项, 能够反映变量之间在校准过程中的交互作用信息。

3) 变量的自身平方项特征: 利用变量的平方项, 如 $P_1 * P_1$, 能够反映变量自身在校准过程中所呈现的非线性特性。

在预测载荷数据时, 多分量天平的数据特征存在一定信息冗余和不相关性, 导致预测结果偏差。可通过降维处理提高预测的准确性。进行载荷预测中的特征选择时以二进制的方式决定特征是否被选取, 0 代表不选择, 1 代表选择。

算法更新后的值为连续值, 因此需要建立连续值和离散值之间的联系, 在每个特征解更新后, 需要根据式 (7) ~ (8) 将连续值转化为离散值。

1	0	1	1	0	...	1	0
F_1	F_2	F_3	F_4	F_5		F_{d-1}	F_d

图 1 二进制表示

$$\text{Sigmoid}(X_i) = \frac{1}{1 + e^{-X_i}} \quad (7)$$

$$X_i = \begin{cases} 0, & \text{if } \text{rand} < \text{igmoid}(X_i) \\ 1, & \text{if } \text{rand} \geq \text{igmoid}(X_i) \end{cases} \quad (8)$$

1.2.1 模拟二进制交叉和 Levy 变异机制

为提高与误差最小化校准特征的搜索效率。采用 Levy 变异和模拟二进制交叉 (SBX, simulated binary crossover) 操作丰富候选特征子集的类型, Levy 变异通过式 (9) 计算:

$$\text{levy} = \frac{u}{|v|^{1/\beta}} \cdot \left| \frac{\Gamma(1+\beta) \cdot \sin\left(\frac{\pi}{2}\beta\right)}{\Gamma\left(\frac{1+\beta}{2}\right) \cdot \beta \cdot 2^{\beta-1/2}} \right|^{1/\beta} \quad (9)$$

u 和 v 是在 (0, 1) 区间内变化的两个随机向量, β 是值为 1.5 的常数。

模拟二进制交叉根据式 (10) 和式 (11) 进行:

$$\beta_1 = \begin{cases} (2u_1)^{1/\eta+1}, & \text{if } u_1 \leq 0.5 \\ \left(\frac{1}{2-2u_1}\right)^{1/\eta+1}, & \text{if } u_1 > 0.5 \end{cases} \quad (10)$$

$$c_{11} = 0.5 \times (1 + \beta_1)p_1 + 0.5 \times (1 - \beta_1)p_2 \quad (11)$$

$$c_{22} = 0.5 \times (1 - \beta_1)p_1 + 0.5 \times (1 + \beta_1)p_2$$

p_1 和 p_2 是交叉父代, c_{11} 和 c_{22} 是生成的子代, 参数, 它控制了新个体与父代的相似性。越小, 新个体与父代的差异越大, β_1 是分布控制扩散系数。

1.2.2 变邻域搜索机制

在风洞天平的载荷预测任务选择特征时, 原始 TS 算法采用连续空间搜索, 通过映射函数将连续解转换为离散解, 这容易陷入局部最优, 针对上述问题, 引入变邻域搜索 (VNS, variable neighbourhood search) 的更新机制^[29]。在此机制中构造了多种邻域结构, 并可以在特征子集间切换, 以二进制形式更新邻域, 以更高效地搜索对降低预测误差有益的特征。表 1 具体描述了 3 种不同的邻域结构的具体搜索行为。

表 1 变邻域搜索的邻域结构

邻域结构	具体内容
邻域结构 1	随机使二元特征向量中的一位偏转 (0→1, 1→0), 同时保持其他值不变;
邻域结构 2	随机选择二元特征向量中的两个位置/元素, 并仅交换这两个元素的值, 同时保持其他值不变;
邻域结构 3	随机选择长度为 L 的子向量生成 2^L 个二进制的格雷码, 并替换原来的子向量, 同时保持其他值不变;

1.2.3 模拟退火机制

在特征子集更新过程中, 可能存在某些子集具有比较重

要的信息，但是适应度较高。模拟退火算法对于初始解不敏感且有概率接受较差解，借鉴其思想^[17]，若新解优于当前解则直接替换，否则以一定概率接受。这样可以提升算法局部搜索的能力，从而在迭代中概率保留有益特征，获得使载荷预测误差最小化的特征组合。接受概率函数定义如下：

$$P_{\text{accept}} = \begin{cases} \exp(-\frac{\Delta E}{T_k}) & \text{if } \Delta E > 0 \\ 1 & \text{if } \Delta E \leq 0 \end{cases} \quad (12)$$

其中： ΔE 为当前解和更新解的适应度差值， T_k 为当前温度。

1.2.4 改进凌日搜索算法流程

图 2 是改进凌日搜索算法流程图，具体步骤如下。

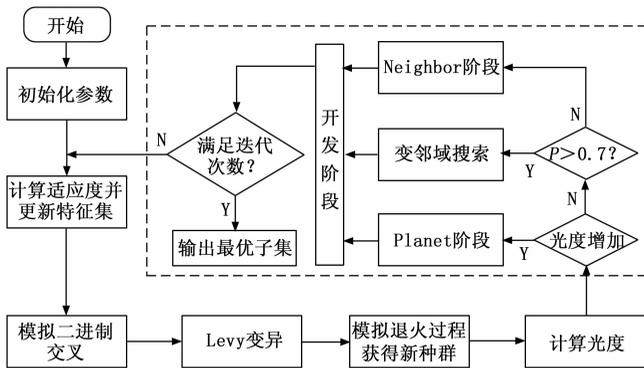


图 2 改进凌日搜索算法流程图

步骤 1：根据式 (1) ~ (2) 初始化获取 n_s 个 d 维星体作为初始特征集，并计算适应度；

步骤 2：按适应度升序排序，从中选取前 k 个特征向量进行模拟二进制交叉，比较后替换升序排序后适应度较高的解；

步骤 3：对更新特征集进行 Levy 变异，并通过模拟退火机制进一步更新；

步骤 4：计算特征集内每一特征子集的光度值；

步骤 5：根据光度值与随机数 P 判断下一步更新方式并更新；

步骤 6：进入开发阶段进一步更新特征集；

步骤 7：保留每轮迭代最优特征，输出满足迭代次数的最终解。

改进的 TS 算法运用了模拟退火、变邻域搜索以及遗传算法中的交叉变异操作来指导风洞天平载荷预测中的特征选择。

为了扩展特征子集的搜索空间，在迭代过程中引入了交叉变异操作。其中，模拟二进制交叉是一种常用的多目标进化算法的交叉算子，它能够充分利用父代个体的信息，并使新个体能够在解空间中进行更大范围的搜索。另外，引入了 Levy 变异，它不仅可以对选择的解进行微调，还能够以一定的突变性跳跃到较远的区域。

模拟退火思想的借鉴，这有助于保留含有关键信息但适应度较高的子集，以上操作可使得特征解集更加多样化，

避免局部最优。原始的 TS 算法中，通过计算光度带来了两种搜索模式，包括探索阶段与邻域搜索阶段，这可以进一步优化搜索结果，但也存在陷入局部最优解的可能。寻找适应度最低的特征解集是一个多目标优化问题，其编码方式是离散的，为此，改进的 TS 算法额外增加了一种搜索方式以协助搜索。变邻域搜索首先定义一组邻域。然后，通过邻域更新初始解，直到适应度不再降低。定义的 3 种邻域如表 1 所示，由于编码特性，邻域更新后特征向量可能不会改变，为了解决这个问题，在对邻域 2 进行更新时设置了一个迭代次数，如果特征向量不改变则继续进行两次交换直至满足迭代次数。邻域 3 更新时根据载荷预测模型的主要特征将二元特征向量分为 3 个搜索区间，然后随机选择区间内长度为 L 的子向量用来生成格雷码并比较适应度。

Algorithm 2: Pseudocode of the VNS Algorithm

- 1) 设置邻域结构集合 $N_k, k = 1, \dots, k_{\max}$
- 2) 选择二元特征向量 x
- 3) $t \leftarrow 0$ // 初始化迭代次数
- 4) while $t < t_{\max}$ do
- 5) $k \leftarrow 1$ // 初始化邻域计数器 k
- 6) repeat
- 7) $x_1 \leftarrow \text{Shake}(x, k)$ // 随机选择 $N_k(x)$
- 8) $x_2 \leftarrow \text{LocalSearch}(x_1)$ // 根据 k 值选择邻域结构作用到 x_1
- 9) if $f(x_2) < f(x_1)$
- 10) $x_1 \leftarrow x_2$, 跳转到第 6 步
- 11) end
- 12) $k \leftarrow k + 1$ 直到 $k = k_{\max}$
- 13) $t \leftarrow t + 1$
- 14) end
- 15) 返回 x

通过以上改进，可以在获得的特征子集上进行细致有效的搜索，并可以跳出局部最优。算法 3 是提出的改进凌日搜索算法的伪代码。

Algorithm 3: Pseudocode of the ITS Algorithm

- 1) 输入: 初始化二元特征向量数量 n_s , 信噪比 S_N
- 2) 输出: 最佳天平特征 L_B 和对应的适应度 f_B
- 3) 初始化: 使用式(1)、式(2)进行星系阶段获取初始化的特征集合 L_s
- 4) While (Stopping condition is not met) do
- 5) 从种群中选择前 k 个特征向量进行模拟二进制交叉操作
- 6) for $i = 1 : n_s$
- 7) 轮盘赌法选出的特征向量采用 Levy 飞行变异
- 8) 将连续的特征向量映射为离散值, 计算适应度
- 9) 采用模拟退火准则评估是否接受变异后新的特征向量进入更新后的特征集
- 10) end
- 11) for $i = 1 \parallel : n_s$
- 12) 如果发生凌日
- 13) 使用式(5)进行 Planet 阶段
- 14) else if $P > 0.7$

- 15) 进行变邻域搜索阶段
- 16) else
- 17) 使用式(6)进行 Neighbor 阶段
- 18) end
- 19) 返回更新后的 L_p 和对应的适应度 f_p
- 20) 使用式(7)进行 Exploitation 阶段
- 21) 返回更新后的 L_p 和对应的适应度 f_p
- 22) end
- 23) 返回最优特征 L_B 和对应的适应度 f_B
- 24) 利用最优特征 L_B 进行载荷预测模型训练和预测分析
- 25) 输出载荷预测值, 进行误差分析和评估预测模型性能

2 基于改进凌日搜索算法的风洞天平载荷预测模型

2.1 数据采集

2.1.1 验证天平

本研究采用两台典型的六分量(包含 3 个力和 3 个力矩)风洞应变天平进行加载验证, 载荷指标如表 2 所示。

表 2 天平自变量信息

天平	力/N 力矩/N×m			28A/28B		
	Y	Z	X	M_z	M_x	M_y
	2 000	300	200	80	20	40

2.1.2 校准系统

校准系统用于被校准的天平的安装、标准载荷的施加、校准过程中天平姿态的检测及调整。如图 3 所示; 本文采用的校准系统属于体轴系天平校准系统, 体轴系校准系统加载时以天平测量端(或与之连接的加载头)为参照坐标系, 天平变形后测量端始终与参照坐标系一致;



图 3 校准系统

在建立载荷预测模型之前, 需要先进行天平的静态校准实验, 天平校准系统模拟了天平在风洞实验时的受力状态, 通过对天平精确地施加静态载荷, 天平元件承受载荷并发生变形, 导致粘贴在天平元件表面上的应变片也同时产生变形导致电阻值发生变化, 测量电路将电阻的变化转换为相应的电压输出量^[30]。实验中, 首先将天平固定安装在校准系统上。然后, 逐步在天平上加加载不同重量的标准载荷, 同时记录下天平传感器输出的对应电压值。通过加加载不同载荷, 收集到输入载荷和输出电压的样本数据 $T = \{F, V\} \in R$ 。V 是电压信号增量数据, F 是载荷数据, 使

用这些数据建立天平载荷和输出电压之间的预测模型。

2.2 模型评价

为了验证预测模型的准确性, 采用综合加载误差 (W_{ci} , combining loading error) 和平均绝对百分比误 (MAPE, mean absolute percent error) 定量评价天平校准模型, 具体描述如式 (14) ~ (15):

$$W_{ci} = \frac{1}{\sqrt{(m-1)P_{\max}}} \sqrt{\sum_{h=1}^m (F_{ih} - P_{ih})^2} \times 100\% \quad (13)$$

$$MAPE = \frac{1}{m} \sum_{h=1}^m \left| \frac{F_{ih} - P_{ih}}{P_{ih}} \right| \times 100\% \quad (14)$$

式中, m 表示各分量载荷的总组数; P_{\max} 为天平各分量设计载荷; P_{ih} 为天平第 i 分量第 h 点综合加载载荷值; F_{ih} 为天平第 i 分量第 h 点综合加载载荷预测值; 综合加载误差和平均绝对百分比误差量化了模型的预测误差, 数值越小表明模型的预测越准确。

2.3 风洞天平载荷预测模型构建

为了去除数据集中的冗余特征, 本文在使用校准系统对风洞天平进行加载采集数据的基础上, 采用改进凌日搜索算法选择与目标相关性较高的特征子集, 然后利用选择出的特征子集通过贝叶斯线性回归建立风洞天平分量载荷的预测模型, 模型的输入变量包括天平各分量电压数据的一次项、交叉项、平方项, 输出变量是对应测量分量载荷。

图 4 展示了载荷预测模型的流程图, 其基本处理流程如下。

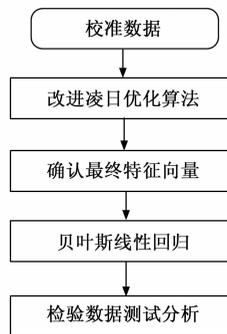


图 4 改进凌日搜索算法的风洞天平载荷预测模型流程图

步骤 1: 数据初始化。对天平进行加载、采集校准数据和检验数据。校准数据用于训练预测模型, 检验数据用于评估预测模型的测试误差。同时, 还需要确定天平分量数 n 与特征数量 D , 并确定 ITS 算法初始参数。

步骤 2: 特征选择。

1) 初始化特征子集。按照图 1 中的编码方法对特征进行编码, 随机生成指定个数的特征子集, 形成初始特征集合。

2) 将特征集输入到贝叶斯线性回归算法中, 根据式 (16) 计算特征子集的适应度并更新特征集, S 是选择特征的数量, $w_1 = 0.99$, $w_2 = 0$, 当 ≤ 0.3 时, $w_2 = 100$ 。

$$fitness = w_1 * err + (1 - w_1) * \frac{S}{D} + w_2 \quad (15)$$

err 是通过十折交叉验证 (10-fold CV, 10-fold cross-

validation) 求解的均方误差 (MSE, mean squared error), MSE 是预测值与实际值之间平方误差的平均值, 可以更好地反映出预测结果与实际结果之间的整体误差, 其计算函数如下:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (16)$$

式中, N 为样本数量, y_i 为真实值, \hat{y}_i 为预测值。

步骤 3: 改进凌日算法搜索出的适应度函数值最小的特征即为最优特征。

步骤 4: 构建载荷预测模型, 将步骤 3 得到的特征输入到贝叶斯线性回归算法中, 在校准数据的基础上获得新的数据集训练载荷预测模型。

步骤 5: 使用检验数据对步骤 5 中确定的预测模型进行测试, 根据设定的评价指标输出结果。

3 模型检验与结果分析

为验证本文所提出 ITS-BLR 载荷预测方法的有效性, 采用 BLR 预测模型、未经改进的 TS-BLR 预测模型以及传统的最小二乘法模型与本文所提出的 ITS-BLR 模型进行对比。根据两台六分量杆式应变天平的加载数据来验证模型, 该天平包含法向力 (Y)、俯仰力矩 (M_z)、轴向力 (X)、滚转力矩 (M_x)、侧向力 (Z) 以及偏航力矩 (M_y) 的测力分量。采用图 3 所示校准系统, 在同一设备和仪器条件下加载 144 组校准数据和 10 组检验数据, 校准数据用于模型训练, 检验数据用于验证模型的性能。

在确定 ITS 算法的结构之后, 需要确定 ITS 算法参数, 首先, 特征子集数量被设置为 30^[31-32], 变邻域搜索次数设置为 10^[33]。通过比较不同参数下交叉验证的适应度确定适宜的交叉概率和变异概率, 变异概率一般在 0.1~0.6 之间, 数值越大, 随机效应越大^[34]。交叉概率一般在 0.5~1 之间^[35]。

固定 ITS 算法的其他参数, 在 28A 天平的 Y 分量上分别取变异概率为 0.1、0.3、0.5 进行仿真, 仿真结果如表 3 所示。

表 3 Y 分量不同变异概率 ITS 算法的适应度对比

变异概率	初始适应度	终止适应度
0.1	8.21×10^{-3}	6.19×10^{-3}
0.3	8.74×10^{-3}	6.06×10^{-3}
0.5	7.92×10^{-3}	6.25×10^{-3}

然后固定变异概率, 分别取交叉概率为 0.6、0.7、0.8、0.9 进行仿真, 仿真结果如表 4 所示。

表 4 Y 分量不同交叉概率 ITS 算法的适应度对比

交叉概率	初始适应度	终止适应度
0.6	8.95×10^{-3}	6.16×10^{-3}
0.7	8.35×10^{-3}	6.08×10^{-3}
0.8	8.74×10^{-3}	6.06×10^{-3}
0.9	9.10×10^{-3}	6.28×10^{-3}

通过对表 3、表 4 中终止适应度的比较, 确定变异概率为 0.3, 交叉概率为 0.8。遗传操作完成后, 需对子代个体进行可行性检验。如果子代个体适应度低于父代个体, 则可行, 将其作为新解加入子代种群; 如果不可行, 则保留父代个体并将其加入子代种群, 在 ITS 算法中, 设置了初始温度为 20, 温度衰减系数 0.99, 在适应度差值较小的情况下, 模拟退火算法会以较大概率接受适应度略差的新解, 当适应度差值较大时, 例如当 $\Delta E = 10$ 时, 根据公式可计算保留的概率约在 $[0.33, 0.61]$ 之间且逐渐减少。意味着即使新解质量严重变差, 算法仍有一定概率接受该劣解, 继续探索解空间, 具体参数如表 5 所示。

表 5 参数设置

参数	值
初始星体规模	30
变异概率 P_m	0.3
交叉概率 P_c	0.8
退火初始温度 T_0	20
温度衰减系数	0.99
变邻域搜索次数	10
迭代次数	80

为了验证改进 ITS 算法的寻优性能, 本文采用标准 TS 和改后的 ITS 同时对特征进行寻优, 其结果如图 5 所示。

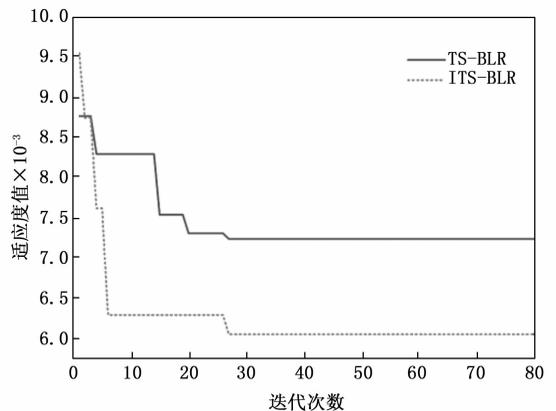


图 5 Y 分量适应度函数变化曲线

根据图 5 的适应度函数变化曲线, 30 次迭代后曲线平稳不变, 说明算法收敛。因此 80 次迭代次数已经足够。可以观察到 TS 算法较容易收敛到局部最优解, 改进 ITS 算法在探索搜索空间时会更加广泛地搜索, 从而搜索到一些适应度值较低的个体, 在 10 次迭代内就达到了低于 TS 算法收敛时的适应度值, 并跳出了局部最优解, 具有更好局部探索和全局寻优能力。图 6 展示了 4 种方法的综合加载误差, ITS-BLR 方法表现最好, 在 M_y 分量上, 综合加载误差下降了 60%。

从表 6 看到, 相比于传统方法, ITS-BLR 方法有效降低了特征的使用数量同时降低了综合加载误差, 尤其在 M_x 、 Z 、 M_y 分量上, 从 TS 的应用来看, TS 算法可以去

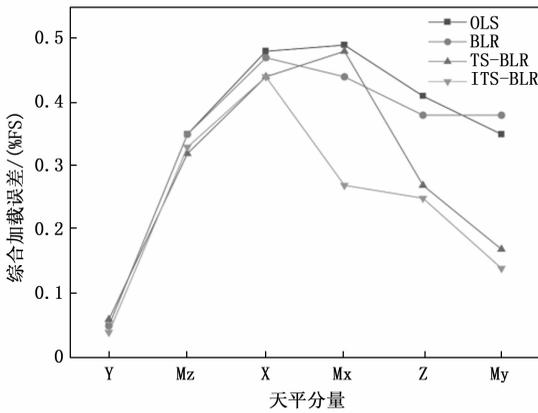


图 6 28A 天平综合加载误差对比

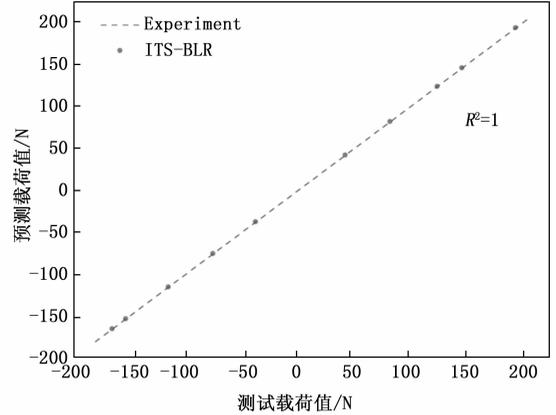


图 7 TS-BLR 模型在 Y 分量上预测值与测试值对比

除较大一部分的特征, 而不会对预测产生负面影响, 但性能略有不足。图 8 中, ITS-BLR 方法的 M_z 分量的 MAPE 略大于 TS-BLR 方法, 增加了 0.05%, 对照表 6, ITS 确定的特征数量相比 TS 减少了一个, 综合加载误差仅增加了 0.01%, 且 ITS-BLR 方法用更少的特征达到了与 TS-BLR 方法接近的预测效果, 证明了 ITS 算法的有效性。表 7 是计算的平均绝对百分比误差, 改进后模型根据 $W_{\text{选}}$ 指标表现较好时, MAPE 也会相应下降, 同样, 如果 $W_{\text{选}}$ 指标变差, MAPE 也会增加, 同一分量下各个模型的 MAPE 和 $W_{\text{选}}$ 指标上的变化趋势存在呈现正相关关系, 且数值都很小, 这说明模型整体上对测试样本的预测效果都比较准确。图 7 为 ITS-BLR 预测模型在 Y 分量的 10 组检验载荷下的预测值, 与测试值对比。从图中可以看出: 预测值紧密分布在对角线上且 $R^2=1$, R^2 在 $[0, 1]$ 之间, 数值越高, 表示预测模型对数据拟合越好。

表 6 28A 天平综合加载误差 % FS

load prediction model	Y	M_z	X	M_x	Z	M_y
OLS	0.05	0.35	0.48	0.49	0.41	0.35
BLR	0.05	0.35	0.47	0.44	0.38	0.38
TS-BLR	0.06 (10)	0.32 (10)	0.44 (20)	0.48 (14)	0.27 (11)	0.17 (12)
ITS-BLR	0.04 (9)	0.33 (9)	0.44 (19)	0.27 (10)	0.25 (9)	0.14 (9)

表 7 28A 天平平均绝对百分比误差 %

load prediction model	Y	M_z	X	M_x	Z	M_y
OLS	0.07	0.61	0.78	0.78	0.92	0.60
BLR	0.07	0.61	0.80	0.76	0.91	0.61
TS-BLR	0.09	0.54	0.76	0.79	0.79	0.41
ITS-BLR	0.07	0.59	0.77	0.58	0.77	0.40

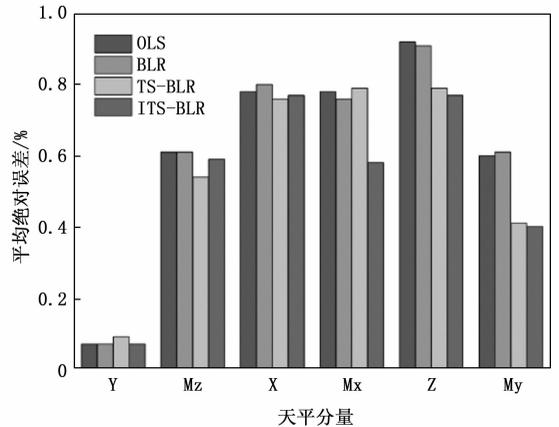


图 8 28A 天平平均绝对百分比误差对比

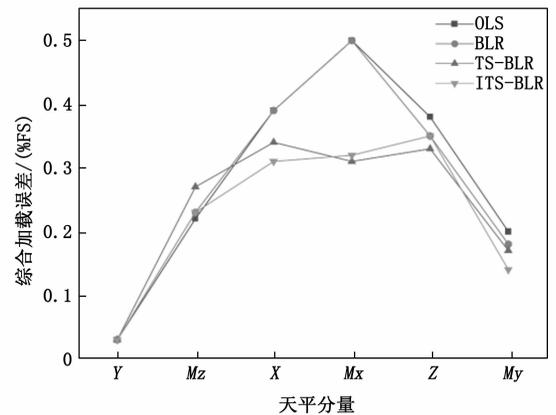


图 9 28B 天平综合加载误差对比

为评估模型泛化能力, 利用另一台六分量天平 28 B 的加载数据集进行验证。该数据集不同于先前的天平数据集, 图 9、10 分别展示了 28 B 天平数据集上的检验数据的综合加载误差和平均绝对百分比误差, X、 M_x 、 M_y 分量上经

过特征选择的综合加载误差相比最小二乘法降低了 21%、36% 和 30%, 平均绝对百分比误差也分别下降了 10%、37% 和 13%, M_z 分量同样使用简化的特征子集, 相较最小二乘法误差上升幅度不大, 但特征数量的减少提高了模型的显著性。对于改进后的 ITS 算法, ITS-BLR 方法在 3 个分量上的综合加载误差低于 TS-BLR 方法, 一个分量上相等, 两个分量略大于但使用特征数量分别减少了 8 个和 2

个。根据表 6~9 的两种误差数据可以看出，直接采用 BLR 算法预测风洞天平的载荷误差和采用 OLS 算法预测的误差结果非常接近，没有明显的差别，表明 BLR 算法可以作为风洞天平载荷预测中一个有效的回归算法。

综合加载误差方面比传统最小二乘法更优；与 TS-BLR 方法相比，ITS-BLR 方法在综合加载误差方面表现出较好的性能，其中有 7 个分量的误差明显低于 TS-BLR 方法。

参考文献：

[1] TROPEA C, YARIN A L, FOSS J F. Springer handbook of experimental fluid mechanics [M]. Berlin: Springer, 2007: 563-616.

[2] JOHNSON T H, PARKER P A, LANDMAN D. Calibration modeling of nonmonolithic wind-tunnel force balances [J]. Journal of aircraft, 2010, 47 (6): 1860-1866.

[3] 汪运鹏, 聂少军, 王 粤, 等. 卷积神经网络在风洞天平静态校准中的应用分析 [J]. 空气动力学学报, 2022, 40 (X): 1-8.

[4] 刘春风, 熊 琳, 刘家骅, 等. 天平校准不确定度的一种评估方法 [J]. 实验流体力学, 2016, 30 (2): 84-90.

[5] GJB 2244A-2011, 风洞应变天平规范 [S]. 2011.

[6] LI G, ZRIMEC J, JI B, et al. Performance of regression models as a function of experiment noise [J]. Bioinformatics and Biology Insights, 2021, 15: 1-10.

[7] GUYON I, ELISSEEFF A. An introduction to variable and feature selection [J]. Journal of Machine Learning Research, 2003, 3: 1157-1182.

[8] MA Y, XIE S, ZHANG X, et al. Hybrid calibration method for six-component force/torque transducers of wind tunnel balance based on support vector machines [J]. Chinese Journal of Aeronautics, 2013, 26 (3): 554-562.

[9] American Institute of Aeronautics and Astronautics. Calibration and use of internal strain-gage balances with application to wind tunnel testing (AIAA R-091A-2020) [J]. 2020.

[10] ALMUALLIM H, DIETTERICH T G. Learning with many irrelevant features [C] //AAAI. 1991, 91: 547-552.

[11] ULBRICH N, BADER J. Analysis of sting balance calibration data using optimized regression models [C] //45th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. 2009: 5372.

[12] ULBRICH N, VOLDEN T. Application of a new calibration analysis process to the MK-III-C balance [C] //44th AIAA Aerospace Sciences Meeting and Exhibit. 2006: 517.

[13] ULBRICH N, VOLDEN T. Strain-gage balance calibration analysis using automatically selected math models [C] //41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. 2005: 4084.

[14] KAMMEYER M E. Balance calibration analysis by sequential regression [C] //AIAA Scitech 2021 Forum. 2021: 0063.

[15] MA T, MIAO L, DU Y, et al. Calibration data prediction of wind tunnel balance based on regression tree [C] //2021 16th International Conference on Intelligent Systems and Knowledge Engineering (ISKE). IEEE, 2021: 362-367.

[16] ZAMAN H R R, GHAREHCHOPOGH F S. An improved particle swarm optimization with backtracking search optimiza-

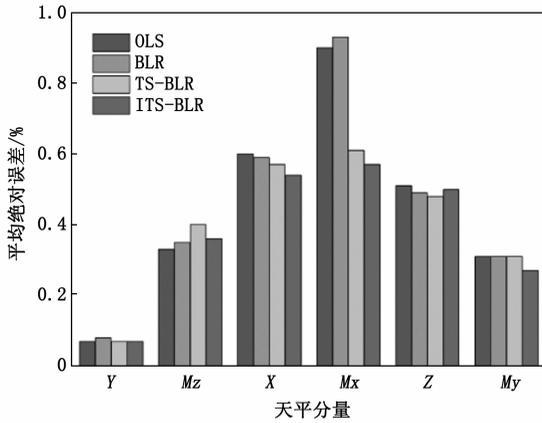


图 10 28 B 天平平均绝对百分比误差对比

表 8 28 B 天平综合加载误差 % FS

load prediction model	Y	Mz	X	Mx	Z	My
OLS	0.03	0.22	0.39	0.50	0.38	0.20
BLR	0.03	0.23	0.39	0.50	0.35	0.18
TS-BLR	0.03 (9)	0.27 (9)	0.34 (19)	0.31 (19)	0.33 (11)	0.17 (16)
ITS-BLR	0.03 (9)	0.23 (9)	0.31 (17)	0.32 (11)	0.35 (9)	0.14 (11)

表 9 28 B 天平平均绝对百分比误差 %

load prediction model	Y	Mz	X	Mx	Z	My
OLS	0.07	0.33	0.60	0.90	0.51	0.31
BLR	0.08	0.35	0.59	0.93	0.49	0.31
TS-BLR	0.07	0.40	0.57	0.61	0.48	0.31
ITS-BLR	0.07	0.36	0.54	0.57	0.50	0.27

4 结束语

本文设计并实践了一种新的风洞天平载荷预测方法，包括特征选择算法和回归算法。使用 2 台六分量天平的检验数据进行测试，可以得到如下结论：

1) 考虑到天平校准数据集较小，且贝叶斯回归不需要交叉验证确定超参数，采用贝叶斯线性回归算法作为预测模型的基础。

2) 相较于传统最小二乘法模型，为了解决存在的冗余特征干扰预测结果的情况，采用凌日搜索算法进行特征选择。但凌日搜索算法在搜索过程中可能陷入局部最优。对其进行改进，最终提出 ITS-BLR 载荷预测模型。

3) 两台天平的 12 个分量中，ITS-BLR 方法有 10 个分量的综合加载误差低于传统最小二乘法，一个分量等于最小二乘法，一个分量略高于最小二乘法，ITS-BLR 方法在

- tion algorithm for solving continuous optimization problems [J]. *Engineering with Computers*, 2022, 38 (Suppl 4): 2797-2831.
- [17] SHARIFIAN Z, BAREKATAIN B, QUINTANA A A, et al. Sin-Cos-bIAVOA: A new feature selection method based on improved African vulture optimization algorithm and a novel transfer function to DDoS attack detection [J]. *Expert Systems with Applications*, 2023, 228: 120404.
- [18] MARJIT S, BHATTACHARYYA T, CHATTERJEE B, et al. Simulated annealing aided genetic algorithm for gene selection from microarray data [J]. *Computers in Biology and Medicine*, 2023, 158: 106854.
- [19] GHAREHAHOPOGH F S. An improved tunicate swarm algorithm with best-random mutation strategy for global optimization problems [J]. *Journal of Bionic Engineering*, 2022, 19 (4): 1177-1202.
- [20] YUSOF N M, MUDA A K, PRATAMA S F, et al. Improving amphetamine-type stimulants drug classification using chaotic-based time-varying binary whale optimization algorithm [J]. *Chemometrics and Intelligent Laboratory Systems*, 2022, 229: 104635.
- [21] MIRRASHID M, NADERPOUR H. Transit search: An optimization algorithm based on exoplanet exploration [J]. *Results in Control and Optimization*, 2022, 7: 100127.
- [22] 李文, 王兴浩, 何云霄, 等. 基于支持向量机的水质浊度补偿研究 [J]. *计算机测量与控制*, 2020, 28 (6): 140-143.
- [23] 颜高洋, 丁贵立, 许志浩, 等. 基于帝王蝶优化算法的 BP 神经网络能源预测模型研究 [J]. *南昌工程学院学报*, 2023, 42 (3): 88-94.
- [24] 车兵辉, 尹欣繁, 彭先敏, 等. 基于 BP 神经网络的天平校准数据处理方法研究 [J]. *计算机测量与控制*, 2020, 28 (10): 165-169.
- [25] BISHOP C M, TIP M E. Bayesian regression and classification [J]. *Nato Science Series sub Series III Computer And Systems Sciences*, 2003, 190: 267-288.
- [26] DU C, SUN L, BAI H, et al. Quantitative detection of azodicarbonamide in wheat flour by near-infrared spectroscopy based on two-step feature selection [J]. *Chemometrics and Intelligent Laboratory Systems*, 2021, 219: 104445.
- [27] CASTILLO I, SCHMIDT-HIEBER J, VAN DER VAART A. Bayesian linear regression with sparse priors [J]. *The Annals of Statistics*, 2015, 43 (5): 1986-2018.
- [28] PEK J, VAN ZANDT T. Frequentist and bayesian approaches to data analysis: evaluation and estimation [J]. *Psychology Learning & Teaching*, 2020, 19 (1): 21-35.
- [29] HANSEN P, MLADENOVI N, TODOSIJEVI R, et al. Variable neighborhood search: basics and variants [J]. *EURO Journal on Computational Optimization*, 2017, 5 (3): 423-454.
- [30] 贺德馨. 风洞天平 [M]. 北京: 国防工业出版社, 2001.
- [31] 问梦飞, 钟建琳, 彭宝营, 等. 基于 WOA-SVR 的电主轴热误差优化建模 [J]. *机床与液压*, 2022, 50 (22): 38-42.
- [32] 徐小平, 朱秋秋, 王峰. 求解圆排列问题的粒子群蚁群优化算法 [J]. *系统仿真学报*, 2017, 29 (2): 248-254.
- [33] 姜天华. 混合灰狼优化算法求解柔性作业车间调度问题 [J]. *控制与决策*, 2018, 33 (3): 503-508.
- [34] 曾强, 沈玲, 潘启东, 等. 基于 NSGA II 的多目标车间设施布局优化方法 [J]. *计算机工程与应用*, 2012, 48 (27): 223-228.
- [35] 尚宇晴, 左钱, 王腾, 等. 改进遗传算法及性能测试 [J]. *机械制造与自动化*, 2020 (1): 60-64.
- ***
(上接第 24 页)
- [12] 李卓, 叶林, 戴斌华, 等. 基于 IDSCNN-AM-LSTM 组合神经网络超短期风电功率预测方法 [J]. *高电压技术*, 2022, 48 (6): 2117-2127.
- [13] 白春垣, 孙有朝. 基于 MSCNN-LSTM 编解码器的飞机辅助动力装置 EGT 预测模型 [J]. *组合机床与自动化加工技术*, 2022, 576 (2): 45-49.
- [14] 王金玉, 金宏哲, 王海生, 等. ISSA 优化 Attention 双向 LSTM 的短期电力负荷预测 [J]. *电力系统及其自动化学报*, 2022, 34 (5): 111-117.
- [15] 钟萍, 徐爱昆, 张艺雯, 等. 一种高效低能耗移动数据采集与无线充电策略 [J]. *软件学报*, 2021, 32 (9): 2867-2886.
- [16] 马礼智, 唐睿, 张睿智, 等. 基于无线能量传输的物联网数据采集系统中资源分配机制的设计 [J]. *信息与控制*, 2023, 52 (2): 220-234.
- [17] 黄光球, 赵羲轩, 陆秋琴. 基于 KPCA-IF-WRF 模型的多源 VOCs 数据清洗方法研究 [J]. *安全与环境学报*, 2022, 22 (6): 3412-3423.
- [18] 韩红桂, 鲁树武, 伍小龙, 等. 基于改进型 SVM 的城市污水处理过程异常数据清洗方法 [J]. *北京工业大学学报*, 2021, 47 (9): 1011-1020.
- [19] 周立, 吴琼, 姚仕明, 等. 江湖系统显式与隐式二维水动力模型比较 [J]. *长江科学院院报*, 2021, 38 (12): 12-18.
- [20] 宋雄, 李大成, 马黎, 等. 二维水动力模型在无资料河段水位流量关系分析中的应用 [J]. *水力发电*, 2022, 48 (4): 17-23.
- [21] 张雨山, 王双银, 臧聪敏, 等. 基于原型观测的输水渠道糙率取值及变化分析 [J]. *水利水电科技进展*, 2021, 41 (1): 24-29.
- [22] 苏磊, 张宝森, 田治宗, 等. 溃口宽度和糙率对黄河下游保护区洪水风险影响分析 [J]. *人民黄河*, 2021, 43 (12): 49-52.
- [23] 周飞飞, 何迎利, 赵华, 等. 基于时间敏感网络的流量调度机制在智能变电站中的研究与实现 [J]. *电力系统保护与控制*, 2021, 49 (19): 107-114.
- [24] 李树平. 管段沿线流量简化前后水头和流行时间差异性分析 [J]. *同济大学学报 (自然科学版)*, 2022, 50 (3): 439-445.