

基于改进 SSA 优化预测模型的大曲 发酵湿度监控系统

Monitoring system of Daqu fermentation humidity based on improved
SSA optimization prediction model

廖俊杰 胡光忠 夏秋 朱文昌 宾秋月

LIAO Jun-jie HU Guang-zhong XIA Qiu ZHU Wen-chang BIN Qiu-yue

(四川轻化工大学机械工程学院, 四川 宜宾 644000)

(College of Mechanical Engineering, Sichuan University of Science & Engineering, Yibin, Sichuan 644000, China)

摘要:目的:解决大曲发酵检测困难、发酵状态难以判断和调控等问题。方法:提出 Tent-SSA 优化 BP 神经网络算法大曲发酵湿度预测模型与动态阈值控制算法,实现大曲发酵过程中大曲状态实时判断与调控。结果:该仿真预测模型预测的湿度误差小(0.596%)、鲁棒性强且收敛速度快。结论:基于该模型的大曲监控系统准确可靠。**关键词:**大曲发酵;改进麻雀搜索算法(SSA);BP 神经网络算法;预测;优化算法

Abstract: Objective: Solve the problems of difficulties of Daqu fermentation detection, judge the fermentation state judgement, and controlling. **Methods:** The Tent-SSA optimized BP neural network algorithm Daqu fermentation humidity prediction model and dynamic threshold control algorithm were proposed to realize real-time judgment of Daqu state and Daqu fermentation control during Daqu fermentation process. **Results:** The error predicted by The simulation prediction model for humidity prediction had low error (0.596%), good robust performance and fast convergence speed. **Conclusion:** The Daqu monitoring system based on this model is accurate and reliable.

Keywords: Daqu fermentation; improved sparrow search algorithm (SSA); back propagation; predict; optimization algorithm

大曲是中国传统白酒酿造中特有的关键糖化发酵剂,主要包括菌系、酶系和物系,其中菌系和酶系是乙醇生成和复杂风味化合物产生的必要条件,物系为酿酒提供部分发酵原料和风味形成的前体物质^[1-2]。大曲质量

的好坏决定着白酒的产率和品质,因而素有“曲为酒之骨”之说^[3]。目前,大部分大曲生产车间仍依靠人工监测大曲发酵,人工判断大曲的发酵状态与质量,但人工测量容易带入杂菌影响曲块发酵。大曲的湿度是大曲发酵的重要指标,也是大曲微生物发酵的关键条件之一,因此将大曲湿度作为大曲发酵状态的监测指标具有一定的可行性。

刘国海等^[4]提出了利用 dbiPLS-SPA 建立固体蛋白饲料发酵湿度模型,检测发酵物湿度,但此方法不适用于实时监测发酵。刘慧等^[5]探究了稻谷在不同贮藏温、湿度中的霉菌生长规律,并采用 Logistic 方程拟合了稻谷霉菌在不同贮藏条件下的生长动力学模型。但此方法采集样本比较困难,不适用于实时检测。李清亮等^[6]提出了融合迁移学习的土壤湿度预测时空模型,通过三维卷积层提取土壤湿度滞后时刻的空间特征对网络模型进行训练。但训练过程复杂,且采集样本数量不够,极易影响模型结果。徐佳乐等^[7]提出了双神经网络控制算法与内循环管道结构,仅解决了曲房内温湿度的检测与控制,使得曲房内温度更均匀,但并未对大曲发酵状态进行研究,并不能反映大曲发酵情况。

研究拟建立一种实时预测诊断大曲发酵系统,采用 Tent 改进麻雀搜索算法(SSA)优化的 BP 神经网络预测模型,实时预测大曲发酵湿度。通过预测值与湿度传感器测量值比较,判断大曲当前发酵状态。并采用湿度动态阈值法,通过调控曲房温湿度,改变大曲发酵湿度至正常状态,为保障大曲质量研究提供依据。

1 大曲发酵监控系统

1.1 监控系统架构

系统通过传感器采集大曲发酵过程中的相关参数,传输至上位机数据库,大曲发酵湿度预测模型输出实时

作者简介:廖俊杰,男,四川轻化工大学在读硕士研究生。

通信作者:胡光忠(1972—),男,四川轻化工大学教授,博士。

E-mail:568092170@qq.com

收稿日期:2021-12-12 改回日期:2022-04-15

湿度预测值,系统比较预测值与真实值误差进行大曲发酵诊断,并将诊断结果反馈给工作人员。根据白酒行业标准,若差值超过正常发酵湿度的 10%,则大曲发酵异常,系统通过动态阈值法调用执行设备调控曲房环境温湿度。大曲发酵实时监控结构如图 1 所示。

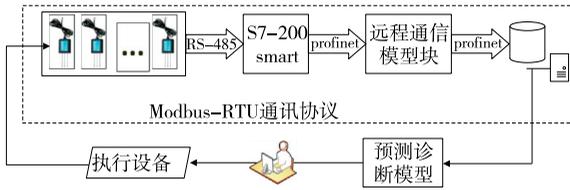


图 1 大曲发酵实时预测诊断结构图

Figure 1 Real-time prediction and diagnostic structure of Daqu fermentation

1.2 大曲湿度仿真预测模型

采用混沌算法(Tent)改进的麻雀搜索算法作优化算法,并利用 Matlab 以 3 种机器学习常用算法(BP/SVM/RF)建立预测仿真模型。通过设置模型最佳参数,比较 3 种仿真预测模型的收敛情况、误差大小、决定系数等,选择其中最适仿真预测模型,并采用改进 SSA 算法优化其预测模型,降低预测模型预测误差至设定阈值范围内。预测模型建立流程图如图 2 所示。

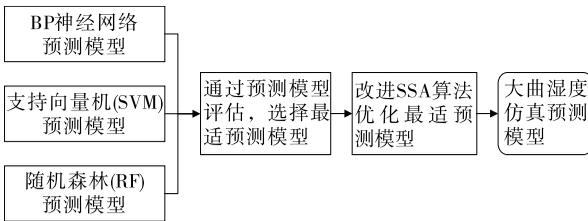


图 2 仿真预测模型建立流程图

Figure 2 Simulation prediction model establishment process

1.2.1 预测模型建立 建立 3 种预测仿真模型(BP/SVM/RF),模型输入设定为发酵时长(Δt)、起始温度(T_f)、起始湿度(H_f)、实时温度(T_t);模型输出设定为大曲实时湿度。利用采集所得的 1 200 组样本进行训练,其中随机将 900 组设定为训练集,其余 300 组为测试集,3 种预测模型分别以最佳参数进行训练。

(1) 输入:

$$X = [x_1 \ x_2 \ x_3 \ x_4] = [\Delta t \ T_f \ H_f \ T_t]$$

(2) 输出:

① BP 神经网络仿真预测模型:

$$y_j = f(s_j) = f\left(\sum_{i=1}^n w_{ji}x_i + b_j\right) = F(W_j X + b) \quad (1)$$

式中:

S_j ——第 j 个神经元输入值;

W ——权值;

b ——偏置。

经样本不断训练确定参数。

② SVM 仿真预测模型:

$$y = f(X) = W\varphi(x) + b \quad (2)$$

式中:

W ——权值向量;

b ——偏置。

利用最小化确定参数。

③ RF 仿真预测模型:

$$f(X) = \frac{1}{b} \sum_{j=1}^b \varphi_j(X) \quad (3)$$

式中:

b ——样本个数。

1.2.2 Tent 改进麻雀搜索优化算法 麻雀搜索算法^[8]

(SSA)是一种近年提出的新型群体优化算法,具有搜索精度高,收敛速度快,稳定性好,鲁棒性强等特点。但 SSA 仍然存在搜索接近全局最优时,容易陷入局部最优。文中利用 Tent 映射初始化种群,引入基于随机变量的混沌映射^[9],使种群分布更加均匀,加快了种群优化效率。根据改进 SSA 算法对基础预测模型进行进一步优化,根据试验法选择最佳模型参数,降低预测误差,仿真预测模型输出大曲湿度与真实发酵的大曲实时湿度基本一致。

(1) 麻雀搜索算法(SSA):SSA 主要模拟麻雀^[10]的觅食与反捕食行为,其发现者位置公式:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j} \exp\left(-\frac{i}{aI_{iter,max}}\right), R2 < ST \\ X_{i,j} + QL, R2 \geq ST \end{cases} \quad (4)$$

式中:

t ——迭代次数;

$X_{i,j}$ ——第 i 个麻雀种群在 j 维度中的位置;

$a \in (0, 1)$;

$I_{iter,max}$ ——最大迭代次数;

Q ——服从正态分布随机数;

L —— $1 \times d$ 且元素均为 1 的矩阵;

$R2 \in (0, 1)$ ——麻雀预警值;

$ST \in (0.5, 1)$ ——麻雀种群安全值。

跟随者更新位置公式^[11]:

$$x_{i,d}^{t+1} = \begin{cases} Q \exp\left(\frac{X_{worst}^t - X_{i,d}^t}{i^2}\right) & , i > N/2, \\ x_{\rho,d}^{t+1} + |x_{i,d}^t - x_{\rho,d}^{t+1}| A^+ L \end{cases} \quad (5)$$

式中:

$x_{i,d}^{t+1}$ —— $t+1$ 次迭代搜索者的最优位置;

x_{worst}^t ——当前全局最差位置;

A ——元素为 1 或 -1 的 d 维矩阵, $A^+ = A^+ (AA^T)^{-1}$ 。

跟随者位置更新的过程主要由 $x_{i,d}^{t+1}$ 和 x_{worst}^t 引导。
种群位置更新公式:

$$X_{i,j}^{t+1} = \begin{cases} X_{best}^t + \beta |X_{i,j}^t - X_{best}^t|, & fi > fg \\ X_{i,j}^t + K \left[\frac{X_{i,j}^t - X_{worst}^t}{(fi - fw) + \varepsilon} \right], & fi = fg \end{cases}, \quad (6)$$

式中:

- X_{best} —— 全局最优位置;
- $\beta \in (-1, 1)$;
- fi —— 麻雀个体适应值;
- fg —— 全局最佳适应值;
- fw —— 全局最差适应值。

(2) Tent 改进 SSA 算法:利用 Tent 混沌算法改进 SSA,避免陷入局部最优,提高全局搜索能力和寻优精度^[12]。

步骤一:将 Tent 变量载波到待求解问题的求解空间。

$$newX_d = \min_d + (\max_d - \min_d)Z_d, \quad (7)$$

式中:

\min_d, \max_d —— d 维变量 $newX_d$ 的最大值与最小值。

步骤二:Tent 扰动公式

$$newX' = (X' + newX) / 2, \quad (8)$$

式中:

- X' —— 需要 Tent 扰动的个体;
- $newX$ —— 产生的 Tent 扰动量;
- $newX'$ —— Tent 扰动后的个体。

1.3 监控系统智能调控

该系统通过 N 个传感器采集实时测量值 y_i 与预测模型对应 N 个实时预测值 y'_i , 比较诊断大曲发酵状态,若系统诊断得出大曲发酵异常,系统通过动态阈值法调用执行设备调控曲房环境,进而调节大曲发酵。控制算法:

(1) $\left(\frac{|y' - y|}{y} \right) \times 100\% < 10\%$ 时,大曲正常发酵,

系统不干涉曲房环境。

(2) $\left(\frac{|y' - y|}{y} \right) \times 100\% > 10\%$ 时,且 $y' > y$,系统输出合适曲房温湿度,通过执行设备提升曲房湿度。

出合适曲房温湿度,通过执行设备提升曲房湿度。

(3) $\left(\frac{|y' - y|}{y} \right) \times 100\% > 10\%$ 时,且 $y' < y$,系统输出合适曲房温湿度,通过执行设备降低曲房湿度。

出合适曲房温湿度,通过执行设备降低曲房湿度。

2 数据采集与样本建立

2.1 样本来源

采用四川宜宾某酒厂发酵大曲作为采集对象,大曲重量约为 4.96 kg,大曲组成为大麦粉约 60%,豌豆粉 40%,制曲时加水量为 48%~50%,成曲水分 37%~

41%,呈中心凸起的长方体,大曲 305 mm×180 mm×85 mm,凸起 110 mm 材料均匀。

2.2 样本建立

通过传感器在曲房中采集的温湿度建立样本。大曲发酵是大曲上的酵母菌、霉菌、细菌等微生物在淀粉、水、氧气、适宜温度等情况下大量增殖产生生物热。大曲上的水分不断被微生物吸收、产生,大曲内湿度在发酵过程中维持一个稳态,因此利用初始温度、初始湿度、发酵时长、实时温度作为模型输入,以实时湿度作为输出建立样本。将所记录样本进行拟合与插值可视化处理,样本数据见表 1。

3 试验分析

3.1 预测模型评估

利用 Matlab 分别将 BP 神经网络算法、支持向量机(SVM)、随机森林(RF)建立仿真预测模型,设定 900 组训练样本,300 组测试样本进行模型训练。训练样本集测试结果的误差平方、平均绝对误差 MAE、均方误差 MSE、均方根误差 RMSE、平均百分比误差 MAPE、相关系数及响应时间见表 2。

将模型响应时间、相关系数、均方根误差 RMSE 作为判定模型的评判标准;仿真预测模型测试如图 3 所示。

根据 3 种模型仿真测试结果与相关系数、RMSE 和相应时间,研究拟采用 BP 预测模型,但 BP 神经网络算法存在极小化、收敛速度慢等问题。因此,采用改进 SSA 优化 BP 算法解决以上问题并减小预测误差,并采用试验法

表 1 样本数据

Table 1 Sample data

样本序号	发酵时长/min	初始湿度/%	初始温度/℃	实时温度/℃	实时湿度/%
1	0	38.6	51.6	38.6	38.6
2	5	38.6	51.6	39.8	39.8
3	10	38.6	51.6	40.6	40.6
4	15	38.6	51.6	41.8	41.8
5	20	38.6	51.6	43.0	43.0
6	25	38.6	51.6	44.2	44.2
⋮	⋮	⋮	⋮	⋮	⋮

表 2 模型评估

Table 2 Model to evaluate

预测模型	相关系数	RMSE	相应时间/s	MAPE/%	MSE	MAE
BP	0.962 52	2.801 5	0.250 906	3.611 8	7.848 5	1.866 3
SVM	0.930 00	3.182 5	0.353 641	4.536 0	10.128 3	2.188 1
RF	0.929 70	3.002 8	2.036 277	3.840 0	9.011 2	3.478 1

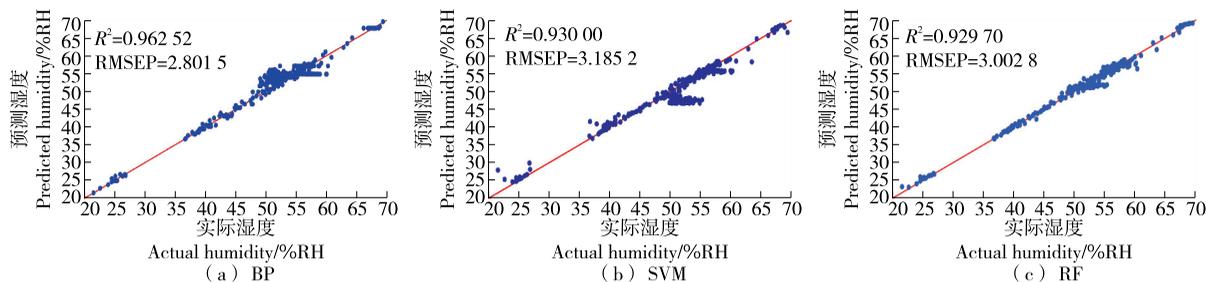


图 3 预测模型测试图

Figure 3 Prediction model test diagram

确定模型结构。

3.2 仿真预测结果

根据改进 SSA 算法优化 BP 神经网络仿真预测模型,确定隐藏层层数与隐藏层节点数。预测模型隐含层设定为 1 层。BP 神经网络算法确定隐含层节点一般依赖经验公式^[13]:

$$L = \sqrt{m+n} + a, \tag{9}$$

式中:

L——隐含层节点数;

m——输入层节点数;

n——输出层节点数;

a——调节常数(1~10)。

由表 3 可知,当隐含层节点数为 9 时,均方差误差最低,因此最佳模型结构确定为 4-9-1。预测模型正反向传播调整权值,将输出值逼近期望值,其模型结构图如图 4 所示。

在正向传播中,输出值:

$$y_j = f(s_j) = f\left(\sum_{i=1}^n w_{ji}x_i + b_j\right) = F(W_j X + b). \tag{10}$$

在反向传播中,样本误差:

$$E_p = \frac{1}{2} \sum_{i=1}^m (y_i - \hat{y}_i)^2 = \frac{1}{2} \sum_{i=1}^m \{y_i - (w x_i + b)\}^2. \tag{11}$$

全局误差:

$$E = \sum_{n=1}^n \sum_{m=1}^m (y_i - \hat{y}_i) = \sum_{m=1}^m E_p. \tag{12}$$

权值调整:

$$\Delta w = -\mu \frac{\partial E}{\partial w_{jk}} = -\mu \frac{\partial}{\partial w_{jk}} \left(\sum_{m=1}^m E_p \right) = \sum_{m=1}^m \left(-\mu \frac{\partial E_p}{\partial w_{jk}} \right), \tag{13}$$

式中:

μ ——学习率;

\hat{y} ——期望值;

表 3 隐含层不同节点误差

Table 3 Hidden layer different node error

节点数	误差	节点数	误差	节点数	误差
4	0.001 839 2	7	0.001 445 7	10	0.001 270 1
5	0.001 380 7	8	0.001 204 4	11	0.001 274 7
6	0.001 406 7	9	0.001 126 9	12	0.001 283 7

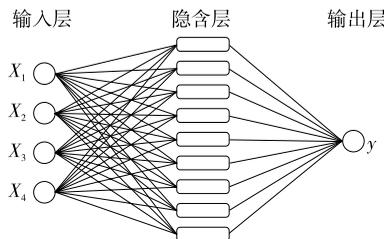


图 4 预测模型结构图

Figure 4 Structure of prediction model

由图 5~图 7 可知,改进 SSA 优化仿真预测模型在迭代 5 次时即获得最小值,均方误差低于 0.625,预测模型稳定且可靠性强。改进 SSA 优化预测模型总体均方误差(MSE)从优化前的 2.511 30 降低至 0.950 17,模型更稳定。

4 结论

大曲发酵湿度预测诊断调控系统采用 Tent-SSA 优化 BP 神经网络预测模型进行大曲发酵湿度实时预测,实现了通过预测值与误差值的比较诊断判断大曲发酵状

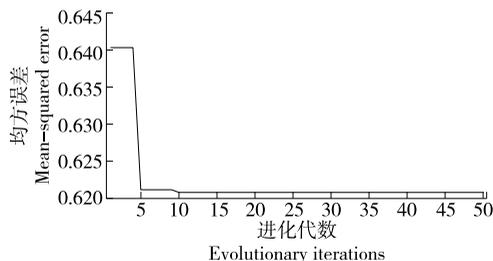


图 5 改进 SSA 优化 BP 模型收敛曲线图

Figure 5 Improved SSA to optimize the convergence curve of BP model

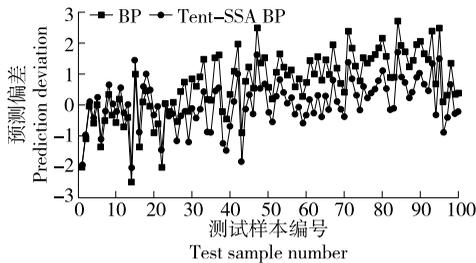


图6 模型预测偏差图

Figure 6 Model prediction bias diagram

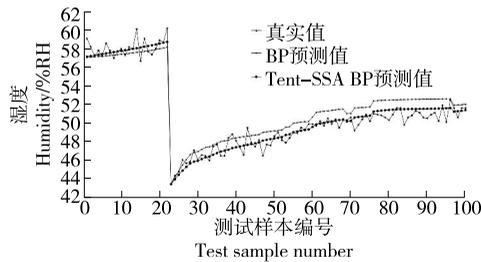


图7 误差对比图

Figure 7 Error contrast diagram

态,同时可利用动态阈值法调用设备调控曲房环境温度。系统具有精准预测诊断大曲发酵的能力,均值百分比误差约 0.596%,且运行稳定,满足大曲发酵监控需求。但模型训练主要采用一个曲房内的的大曲,未考虑普遍采样,且采样时间太短,因此预测模型只适用于大曲阶段性的湿度预测,后续应通过不同曲房采样,同时将采样时间延长至整个大曲发酵周期,使预测模型更稳定,且具有普遍适用性。

参考文献

- [1] 黄和强, 车富红, 陈占秀, 等. 青稞大曲微生物菌群的多样性及其核心菌群的判定[J]. 食品与发酵工业, 2021, 47(23): 305-310. HUANG H Q, CHE F H, CHEN Z X, et al. Determination of microbial diversity and core microbiota in highland barley Daqu[J]. Food and Fermentation Industries, 2021, 47(23): 305-310.
- [2] 申孟林, 张超, 王玉霞. 白酒大曲微生物研究进展[J]. 中国酿造, 2016, 35(5): 1-5. SHEN M L, ZHANG C, WANG Y X. Research progress of microbes in liquor Daqu[J]. China Brewing, 2016, 35(5): 1-5.
- [3] 吴生文, 张志刚, 李旭晖. 大曲微生物在大曲酒生产中的研究开发现状及发展前景[J]. 中国酿造, 2011(5): 8-13. WU S W, ZHANG Z G, LI X H. Research progress on microorganism in Daqu liquor[J]. China Brewing, 2011(5): 8-13.
- [4] 刘国海, 江辉, 梅从立. 基于 dbiPLS-SPA 变量筛选的固态发酵湿度近红外光谱检测[J]. 农业工程学报, 2013, 29(S1): 218-222. LIU G H, JIANG H, MEI C L. Rapid detection of moisture content in solid-state fermentation by near-infrared spectroscopy combined with dbiPLS-SPA[J]. Transactions of the Chinese Society of Agricultural Engineering, 2013, 29(S1): 218-222.
- [5] 刘慧, 周建新, 方勇, 等. 不同储藏温、湿度对稻谷霉菌生长的影响及生长预测模型的建立[J]. 中国粮油学报, 2020, 35(2): 110-115. LIU H, ZHOU J X, FANG Y, et al. Effects of different storage temperature and humidity on the growth of mold in paddy and establishment of growth prediction model[J]. Journal of the Chinese Cereals and Oils Association, 2020, 35(2): 110-115.
- [6] 王学智, 李清亮, 李文辉. 融合迁移学习的土壤湿度预测时空模型[J]. 吉林大学学报(工学版), 2021, 52(3): 675-683. WANG X Z, LI Q L, LI W H. Spatio-temporal model of soil moisture prediction integrated with transfer learning[J]. Journal of Jilin University (Engineering and Technology Edition), 2021, 52(3): 675-683.
- [7] 徐佳乐, 黄丹平, 田建平, 等. 新型曲房内循环温度测控系统设计[J]. 食品与机械, 2020, 36(12): 90-94. XU J L, HUANG D P, TIAN J P, et al. New cycle temperature measurement and control system in kojibank [J]. Food & Machinery, 2020, 36(12): 90-94.
- [8] OUYANG C, ZHU D, QIU Y. Lens learning sparrow search algorithm[J]. Mathematical Problems in Engineering, 2021(2): 1-17.
- [9] CHEN X, HUANG X, ZHU D, et al. Research on chaotic flying sparrow search algorithm[J]. Journal of Physics: Conference Series, 2021, 1 848: 29-31.
- [10] 魏鹏飞, 樊小朝, 史瑞静, 等. 基于改进麻雀搜索算法优化支持向量机的短期光伏发电功率预测[J]. 热力发电, 2021, 50(12): 74-79. WEI P F, FAN X C, SHI R J, et al. Short-term photovoltaic power generation forecast based on improved sparrow search algorithm optimized support vector machine[J]. Thermal Power Generation, 2021, 50(12): 74-79.
- [11] 付华, 刘昊. 多策略融合的改进麻雀搜索算法及其应用[J]. 控制与决策, 2021, 37(1): 87-96. FU H, LIU H. Improved sparrow search algorithm with multi-strategy integration and its application[J]. Control and Decision, 2021, 37(1): 87-96.
- [12] 吕鑫, 慕晓冬, 张钧, 等. 混沌麻雀搜索优化算法[J]. 北京航空航天大学学报, 2021, 47(8): 1 712-1 720. LU X, MU X D, ZHANG J, et al. Chaos sparrow search optimization algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(8): 1 712-1 720.
- [13] 吴少杰, 刘怀举, 张仁华, 等. 基于正交实验和数据驱动的喷丸表面完整性参数预测[J]. 表面技术, 2021, 50(4): 86-95. WU S J, LIU H J, ZHANG R H, et al. Prediction of surface integrity parameters of shot peening based on orthogonal experiment and data-driven[J]. Surface Technology, 2021, 50(4): 86-95.