

基于随机森林和梯度提升树的双金属复合管材弯曲壁厚减薄率预测模型研究

朱英霞*, 元琛, 王磊, 陈炜, 许江平

(江苏大学 机械工程学院, 江苏 镇江 212031)

摘要: **目的** 通过机器学习方法构建工艺参数与壁厚减薄率之间的映射关系, 建立弯管壁厚减薄率高精度预测模型, 为缺陷的有效预测提供数据驱动解决方案。**方法** 首先, 以直径、弯曲程度等6项工艺参数为变量设计仿真方案, 并进行实验验证, 建立数据库来源。其次, 分别建立随机森林(Random Forest, RF)与梯度提升回归树(Gradient Boosting Regression Tree, GBRT)2种机器学习模型。最后, 对比分析训练样本量($N=100\sim 808$)对2种机器模型泛化性能的影响规律。**结果** 在中等样本量($N=400$)下, 2种机器模型的预测性能达到最优。其中, GBRT模型测试集决定系数(R^2)高达0.9769, 方差贡献率(VAF)为97.69%, 均方根误差(RMSE)和平均绝对误差(MAE)分别低至0.3470和0.1759, 所有性能指标均显著优于同等条件下的RF模型的性能指标($R^2=0.9229$, VAF=92.36, RMSE=0.6335, MAE=0.3638)。当样本量超过400后, 2种机器模型均出现预测性能饱和甚至衰减的趋势。采用训练区间之外的工艺参数进行GBRT模型预测精度验证, 发现其平均预测误差小于7%。**结论** 构建的GBRT预测模型能够高精度、高效率地映射双金属复合管弯曲工艺参数与壁厚减薄率之间的复杂非线性关系, 验证了机器学习方法在该领域替代或辅助传统有限元分析的可行性。

关键词: 双金属复合管; 弯曲; 壁厚减薄; 机器学习; 预测模型

DOI: 10.3969/j.issn.1674-6457.2026.03.020

中图分类号: TG306 **文献标志码:** A **文章编号:** 1674-6457(2026)03-0184-10

Prediction Model for the Wall Thickness Reduction Rate in Tube Bending Based on the Random Forest and the Gradient Boosting Tree

ZHU Yingxia*, YUAN Chen, WANG Lei, CHEN Wei, XU Jiangping

(School of Mechanical Engineering, Jiangsu University, Jiangsu Zhenjiang 212031, China)

ABSTRACT: The work aims to construct a mapping relationship between process parameters and the wall thickness reduction rate with machine learning methods and establish a high-precision prediction model for the wall thickness reduction rate in tube bending, thereby providing a data-driven solution for effective defect prediction. Firstly, a simulation scheme was designed with

收稿日期: 2025-12-05

Received: 2025-12-05

基金项目: 国家自然科学基金(52575387); 校企合作(HX20241044)

Fund: The National Natural Science Foundation of China(52575387); University-Industry Collaboration(HX20241044)

引文格式: 朱英霞, 元琛, 王磊, 等. 基于随机森林和梯度提升树的双金属复合管材弯曲壁厚减薄率预测模型研究[J]. 精密成形工程, 2026, 18(3): 184-193.

ZHU Yingxia, YUAN Chen, WANG Lei, et al. Prediction Model for the Wall Thickness Reduction Rate in Tube Bending Based on the Random Forest and the Gradient Boosting Tree[J]. Journal of Netshape Forming Engineering, 2026, 18(3): 184-193.

*通信作者(Corresponding author)

six process parameters, such as diameter and bending degree, etc. as variables, and experimental validation was conducted to establish a database. Secondly, two machine learning models—Random Forest (RF) and Gradient Boosting Regression Tree (GBRT)—were developed. Finally, the effect of training sample size (ranging from $N=100$ to 808) on the generalization performance of the two models was compared and analyzed. Under a medium sample size ($N=400$), both machine learning models achieved optimal prediction performance. Among them, the GBRT model exhibited a coefficient of determination (R^2) as high as 0.976 9 on the test set, a variance accounted for (VAF) of 97.69%, and a root mean square error (RMSE) and a mean absolute error (MAE) as low as 0.347 0 and 0.175 9, respectively. All performance metrics were significantly superior to those of the RF model under the same conditions ($R^2=0.922$ 9, VAF=92.36, RMSE=0.633 5, MAE=0.363 8). Beyond a sample size of 400, both models showed a trend of performance saturation and eventual degradation. Using process parameters outside the training range for GBRT model validation revealed an average prediction error of less than 7%. The GBRT prediction model constructed in this study can accurately and efficiently map the complex nonlinear relationship between the bending process parameters of bimetallic composite tubes and the wall thickness reduction rate, demonstrating the feasibility of machine learning methods as alternatives or supplements to traditional finite element analysis in this field.

KEY WORDS: bimetallic composite tube; bending; wall thickness reduction; machine learning; prediction model

双金属复合弯管^[1-5]具有双材料的物理化学性能,能够满足质量轻、耐腐蚀、强度高的工业要求,因此在诸多工业领域中得到广泛应用。然而,当弯曲半径较小时,双金属复合管在弯曲成形后将不可避免地会产生一些成形缺陷,如壁厚过度减薄^[6-8]现象,严重影响了其成形质量和使用性能。因此,必须要对双金属复合管弯曲壁厚减薄率的预测精度展开研究。

有限元仿真是研究管材壁厚减薄的主要方法,可通过整合材料非线性、接触边界与工艺参数,实现高精度量化分析。例如, Xiong 等^[9]利用 COMSOL 仿真平台建立了脉冲涡流检测模型,分析了复合管壁厚减薄的信号特征并建立了映射关系。Hu 等^[10]采用 ABAQUS 仿真平台构建了非切向旋转弯曲模型,揭示了弯曲参数对壁厚分布的影响。Fang 等^[11]则指出,采用随塑性应变变化的弹性模量模型可显著提高预测精度,并将仿真误差从 19.27%降至 7.72%。由此可见,有限元仿真可以为管材弯曲壁厚预测提供可靠的数据来源。但是,有限元方法也存在计算成本高、耗时长、定量分析复杂等突出问题^[12]。因此,寻求一种能够替代有限元仿真的壁厚减薄预测方法非常迫切。

近年来,基于数据驱动的机器学习方法为管材壁厚减薄预测提供了新思路。其中,随机森林(Random Forest, RF)^[13-14]与梯度提升树(Gradient Boosting Tree, GBT)^[15-16]凭借其能够有效处理高维非线性关系、抵抗过拟合以及适用于小样本数据等特点,已成为该领域预测建模的研究热点。

RF 是一种集成学习算法,通过构建多棵决策树并集成其预测结果,在结构化数据回归与分类任务中表现出良好的鲁棒性,已被广泛应用于医疗诊断、金融风控和工业制造等领域^[17-18]。近年来,该算法在管材弯曲缺陷预测方面展现出显著潜力。例如, Amine 等^[19]采用粒子群(PSO)、蚁群(ACO)和鲸鱼优化

(WOA)算法自动优化 RF 超参数,构建了混合预测模型。基于 104 组钢管纯弯试验,发现优化后的模型显著提高了钢管抗弯承载力的预测精度。Wu 等^[20]基于弯管机多轴伺服信号,通过时频特征提取与 Fisher 评分筛选,构建了 RF 分类器以在线识别起皱状态。实验表明,优化后的模型识别准确率达 77%~89%,验证了该方法的实时检测有效性,为替代传统离线测量提供了新途径。上述研究充分表明,RF 在处理高维、小样本工艺参数方面具有优异性能,尤其适用于实验成本较高的金属管材弯曲成形场景。

GBT 通过迭代拟合残差构建高性能预测模型,已在超声波预测与材料科学等领域得到广泛应用^[21-22]。Barton 等^[23]首次将 GBT 应用于英国供水管网失效概率预测,通过整合局部环境特征,建立了 AUC 达 0.89 的高精度模型,并结合风险分布图为管网维护提供可视化支持,体现了 GBT 处理空间异质数据的优势。在工业制造方面,Zhao 等^[24]结合三维点云技术与 GBT,实现了铝钢接头疲劳寿命的非破坏性预测,其预测误差显著低于传统方法,展现了 GBT 在多源异构数据中的卓越泛化能力。研究表明,GBT 能够有效融合结构、环境与工艺等多维度特征,在管材状态预测中具有显著潜力。梯度提升回归树(Gradient Boosting Regression Tree, GBRT)是 GBT 框架专精于连续变量预测的关键分支。相较于通用的 GBT,GBRT 在回归任务中进行了深度优化,使其在需要高精度解析输入变量间复杂交互作用的场景中表现尤为突出。已有研究证实,其在诸如金属表面摩擦系数、磨损率预测等精密回归问题上具有显著优势^[25]。因此,针对本研究中“工艺参数-壁厚减薄率”这一高精度非线性映射的核心需求,GBRT 从理论上即展现出高度的适配性。

在梯度提升算法中,除经典的 GBRT 外,后续涌

现出一系列高性能的改进算法,例如 XGBoost 和 LightGBM^[26]。这些算法通过引入正则化项、基于梯度的单边采样、特征捆绑等创新技术,在计算速度、内存效率及处理特定数据类型(如类别特征)方面表现出显著优势,已成为当前机器学习领域的热点工具。然而,这些先进的模型的核心思想仍源于梯度提升框架。本研究选取其理论基础 GBRT 模型,旨在与 RF 进行一场聚焦于集成学习本源思想的对比,即 Boosting(串行、偏差驱动)与 Bagging(并行、方差驱动)两大根本范式的机理辨析。这一设计能够确保对比的纯粹性与结论的一般性,避免因后续算法的复杂优化而模糊对核心机制作用的评估。同时,明确 GBRT 这一“基础模型”在本任务上的性能基准,也为未来引入更复杂优化算法提供不可或缺的参照系。

综上,本研究选用 RF 和 GBRT 模型,是基于其解决高精度回归问题的理论优势、与本研究目标的高度契合。目前关于双金属复合管弯曲成形中壁厚减薄预测的研究,多依赖于有限元数值仿真方法。尽管 RF 与 GBRT 等机器学习方法具有处理高维、非线性小样本数据的优势,并在超参数优化、多源信号融合等方面取得较大进展,但其在多工艺参数耦合作用下双金属复合管壁厚减薄预测方向上的研究,仍待深入。

因此,本文采用 RF 和 GBRT 2 种机器学习算法,综合考虑金属管直径、壁厚、弯曲半径等工艺参数,构建了金属管弯曲壁厚减薄率的预测模型,探究了样本数量对模型预测性能的影响变化规律,同时对比分析了 2 种模型的性能指标,得出了最佳预测模型。最后,利用样本集之外的数据验证了模型的可靠性。

1 研究方法

本研究针对双金属薄壁复合管弯曲过程中的壁厚减薄预测问题,建立了 RF 和 GBRT 的模型。这 2 种模型都遵循机器学习的通用框架:采用最小化损失函数 $L(\eta_i, \hat{\eta}_i)$ 的方法构建工艺参数向量 $\mathbf{X}$ 和真实壁厚减薄率向量 $\boldsymbol{\eta}$ 之间的映射,最终实现输入工艺参数向量 $\mathbf{X} \in \mathbb{R}^6$, 输出机器学习模型预测的壁厚减薄率 $\hat{\boldsymbol{\eta}} \in \mathbb{R}$, 如式(1)所示。

$$\min \left\{ \frac{1}{N} \sum_{i=1}^N L(\eta_i, \hat{\eta}_i) + \phi \Omega \right\} \quad (1)$$

式中: N 为训练样本数量; η_i 和 $\hat{\eta}_i$ 分别为第 i 个训练样本的真实壁厚减薄率和模型预测壁厚减薄率; Ω 为正则化项(用于控制模型复杂度,防止过拟合); ϕ 为正则化强度系数。

模型训练完成后,本研究在独立的测试集上,对比评估 2 种机器学习模型的预测性能,采用的指标包括决定系数(R^2)、平均绝对误差(MAE)和均方根误差(RMSE)等。整体研究流程如图 1 所示。

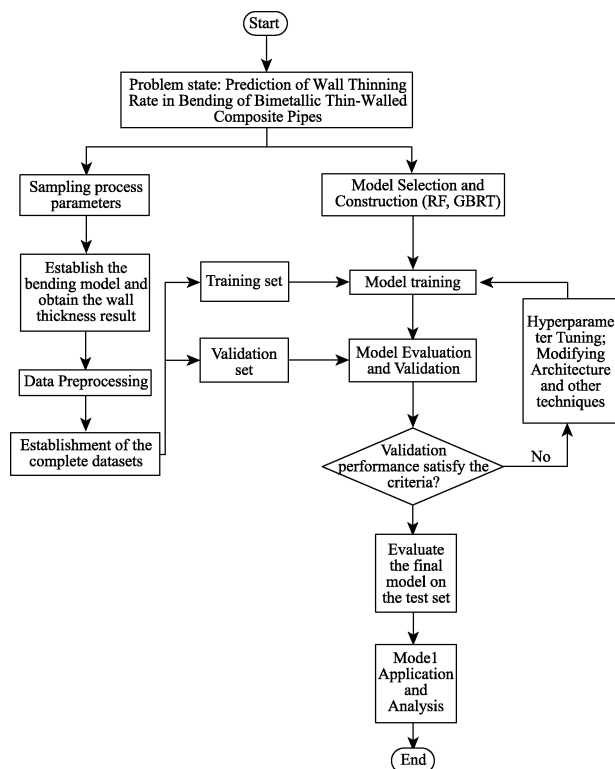


图 1 研究流程图

Fig.1 Flowchart of study

1.1 双金属复合管绕弯成形工艺

复合管绕弯成形原理如图 2 所示。模具包括弯曲模、压块、夹块以及防皱块。图 2a 为复合管弯曲的初始状态,通过夹块与弯曲模将复合管固定,夹块提供夹紧力和绕弯速度,牵引复合管沿弯曲模中心线旋转。压块提供一定压力在管壁上,并向前运动,提供助推作用。防皱块和弯曲模紧密相切,为防止弯曲过程中复合管内侧发生起皱。复合管随弯曲模以一定的角速度旋转,当到达设定弯曲角度时,完成整个弯曲过程,如图 2b 所示。在弯曲过程中,复合管外脊线侧易产生壁厚过度减薄缺陷。复合管绕弯成形以拉弯的方式进行,外侧管材受切向拉应力的作用导致壁厚减薄,由于覆管所受拉应力大于基管,因此一旦覆管减薄到一定程度就会发生破裂,就会直接报废无法使用,这严重影响了使用性能。因此,在分析复合管绕弯成形时,主要考虑覆管外侧壁厚减薄量的变化规律。弯曲过程结束后,夹块和压块从管材上脱开,管材发生回弹。

采用各向同性屈服函数和硬化函数描述铜钛复合管的材料本构关系,如式(2)所示。

$$F = \sqrt{\frac{3}{2}} \mathbf{s} : \mathbf{s} - K \left(\bar{\varepsilon}^p \right)^n = 0 \quad (2)$$

式中: $\mathbf{s}$ 表示偏应力张量; $\bar{\varepsilon}^p$ 为等效塑性应变; K 与 n 为材料硬化常数,其值通过单向拉伸试验获取,见表 1。

表 1 复合管的材料参数
Tab.1 Material parameters of the composite tube

Tubes	Density $\rho/(10^3 \text{ kg} \cdot \text{m}^{-3})$	Elastic modulus E/GPa	Strength factor K/MPa	Hardening index n
T2 base tube	8.86	63.09	416.13	0.172
TA2 covered tube	4.51	85.33	721.92	0.196

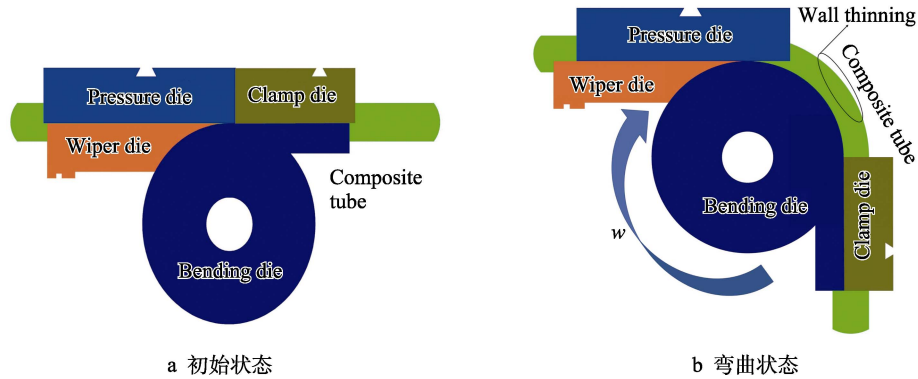


图 2 复合管绕弯示意图
Fig.2 Schematic diagram for rotary-draw bending of the composite tube: a) initial state; b) bending state

1.2 建立数据集 D

1.2.1 建立工艺参数向量 X

数据集的质量对机器学习模型的预测性能影响很大,必须选用对成形缺陷敏感度较高的工艺参数建立工艺参数向量 X 。表 2 为影响双金属弯管壁厚变化的 6 个主要工艺参数 λ 、 ω 、 α 、 β 、 μ 、 δ 的含义及其取值范围。其中,截面中空系数 λ 和弯曲程度 ω 的描述方法基于文献[27],是传统的分析常量 D_2 、 t_0 、 R_0 的函数。角速度 α 、摩擦系数 μ 和模具间隙 δ 的取值受到实际物理约束与工业化生产条件的影响,因此通过前期优化确定其为固定值,以此定义一个稳定的工艺背景。本研究旨在该特定工艺背景下,构建核心几何参数 (λ , ω , β) 与壁厚减薄率之间的精确映射模型。为了实现数据的随机性,采用超拉丁立方采样方法[28],生成了基于表 1 的 808 组样本数据 ($N=808$)。最终建立工艺参数向量 X ,如式 (3) 所示。

$$X = \begin{bmatrix} \lambda_1 & \omega_1 & \beta_1 & \mu_1 & \delta_1 & \alpha_1 \\ \lambda_2 & \omega_2 & \beta_2 & \mu_2 & \delta_2 & \alpha_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_N & \omega_N & \beta_N & \mu_N & \delta_N & \alpha_N \end{bmatrix}_{N=808} \quad (3)$$

1.2.2 基于有限元仿真建立数据来源

充足的训练数据对机器学习模型的性能至关重要。然而,通过实验收集足够的数据往往既昂贵又耗时,特别是在工业场景中。为了避免高成本和材料浪费,本文基于 ABAQUS 有限元仿真建立数据来源,如图 3 所示。其中,基管和覆管网格单元采用 S4R 壳单元,其他模具采用 R3D4 壳单元。有限元模型包括弯曲模型和回弹模型两部分,弯曲模型采用显示迭代算法,回弹模型采用隐式迭代算法。摩擦模型采用经典库仑摩擦模型,基管与覆管、覆管与模具的接触均采用“面-面”接触方式。文献[27]从截面畸变角度验证了该模型的可靠性。本文从壁厚减薄率角度验证该模型的可靠性。图 4 为管材弯曲成形后的壁厚减薄率实验检测 results 和仿真预测结果。可以看到,无论是基管还是覆管,其仿真曲线的变化趋势始终和实验结果一致。对于 $D_220-R_060-t_03$ (覆管外径 $D_2=20 \text{ mm}$,弯曲半径 $R_0=60 \text{ mm}$,厚度 $t_0=3 \text{ mm}$) 薄壁复合管,基、覆管壁厚减薄率的平均预测误差分别为 4.79%、8.09%;对于 $D_234-R_0102-t_03$ ($D_2=34 \text{ mm}$, $R_0=102 \text{ mm}$, $t_0=3 \text{ mm}$) 薄壁复合管,基、覆管壁厚减薄率的平均预测误差分别为 6.36%、7.55%。由此可知,该有限元模型是可靠的。

表 2 几何和工艺参数的符号及取值范围
Tab.2 Symbols and value ranges of geometric and process parameters

Section hollow coefficient λ	Bending degree ω	Angular speed of bending die $\alpha/(\text{rad} \cdot \text{s}^{-1})$	Bending angle $\beta/(^{\circ})$	Friction coefficient between dies and tube μ	Gaps between dies and tube δ/mm	Diameter of covered tube D_2/mm	All thickness t_0/mm	Bending radius R_0/mm
(0, 1.47)	(0.17, 1.5)	0.5	(0, 90)	0.18	0.1	(6, 90)	(2.5, 4)	(30, 200)

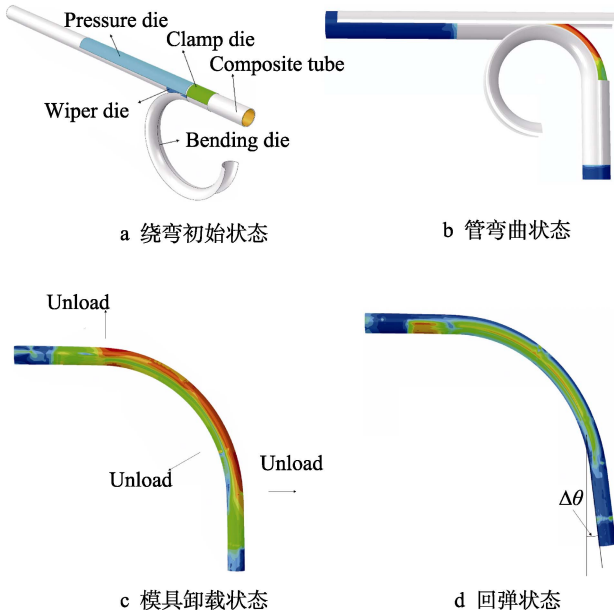


图3 Cu/Ti双金属复合管绕弯-回弹有限元模型

Fig.3 Finite element model for rotary-draw bending-springback of Cu/Ti bimetallic composite tube: a) initial state of rotary-draw bending; b) tube bending state; c) die unloading state; d) springback state

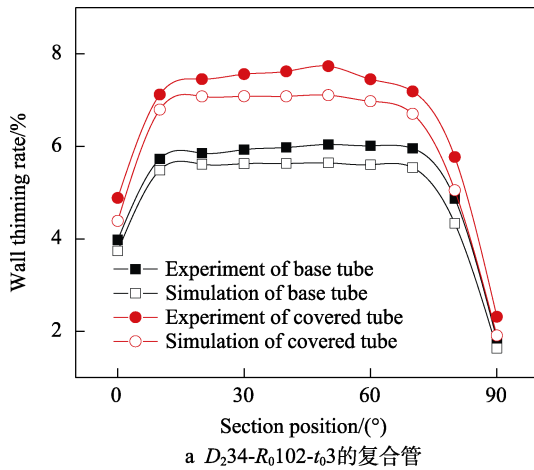
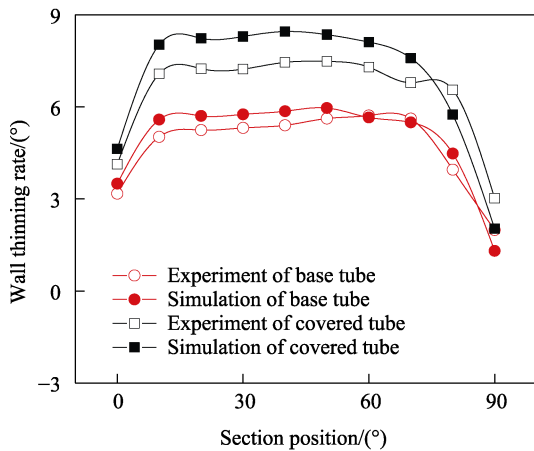
a $D_{234-R0102-t03}$ 的复合管b $D_{220-R060-t03}$ 的复合管

图4 实验和有限元结果之间的壁厚减薄率对比

Fig.4 Comparisons of wall thickness reduction rates between simulative and experimental results: a) composite tubes of $D_{234-R0102-t03}$; b) composite tubes of $D_{220-R060-t03}$

基于上述所建有限元模型, 以及工艺参数向量 $\mathbf{X}$, 仿真输出弯管的壁厚减薄率向量 $\boldsymbol{\eta}$:

$$\boldsymbol{\eta} = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_N \end{bmatrix}_{N=808} \quad (4)$$

由此建立数据集 $\mathbf{D}=[\mathbf{X}, \boldsymbol{\eta}]$ 。其中, 每个样本都包含 6 个特征输入变量和 1 个连续目标变量:

$$\mathbf{D}=[\mathbf{X}, \boldsymbol{\eta}] = \begin{bmatrix} \lambda_1 & \omega_1 & \beta_1 & \mu_1 & \delta_1 & \alpha_1 & \eta_1 \\ \lambda_2 & \omega_2 & \beta_2 & \mu_2 & \delta_2 & \alpha_2 & \eta_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_N & \omega_N & \beta_N & \mu_N & \delta_N & \alpha_N & \eta_N \end{bmatrix}_{N=808} \quad (5)$$

1.3 两种机器学习方法与数学原理

1.3.1 随机森林 (RF) 预测模型原理

RF 模型的构建流程如图 5 所示, RF 中每棵决策树随机选取 2 个特征进行节点分裂, 分裂时需遵循 MSE 最小化分裂准则, 最终 RF 预测值为所有决策树的平均值。其数学原理如下。

1.3.1.1 单颗决策树构建

RF 由 B (取 $B=300$) 棵决策树 T_b ($b=0, 1, \dots, B$) 组成, 对于每棵决策树 T_b , 从原始数据集 $\mathbf{D}$ 中有放回地随机抽取 N 个样本, 形成自助采样集 D_b :

$$D_b = \left\{ (X_i^{(b)}, \hat{\eta}_i^{(b)}) \right\}_{i=1}^N \quad (6)$$

式中: $X_i^{(b)}$ 代表第 b 棵决策树中输入数据的特征。

1.3.1.2 分裂准则

在回归树中, 使用均方误差 (MSE) 作为衡量节点分裂效果的标准。节点 t 的均方误差 ($\delta_{\text{MSE}(t)}$) 定义为节点内所有样本 D_t 的真实值 η_i 与其平均值 $\bar{\eta}$ 的平方误差的平均值:

$$L(\eta_i, \hat{\eta}_i) = \delta_{\text{MSE}(t)} = \frac{1}{|D_t|} \sum_{i=1}^{|D_t|} (\eta_i - \bar{\eta})^2 \quad (7)$$

式中: $|D_t|$ 是 D_t 的样本数。当 $|D_t|$ 达到最小样本分裂值 (取 $\text{min_samples_split}=5$) 或增长 RF 达到最大深度 (取 $\text{max_depth}=20$) 时, 输出节点内样本均值。RF 的最终预测为所有决策树 T_b ($b=0, 1, \dots, B$) 预测值的平均:

$$\hat{\eta} = \frac{1}{B} \sum_{b=1}^B T_b(x) \quad (8)$$

1.3.2 基于梯度提升树 (GBRT) 的预测模型

GBRT 模型的构建流程如图 6 所示。GBRT 通过迭代拟合残差 r_{ib} 来优化模型的预测性能。其基本思想是逐步构建一系列决策树 T_b , 每棵树拟合前一棵树的残差。其数学框架如下。

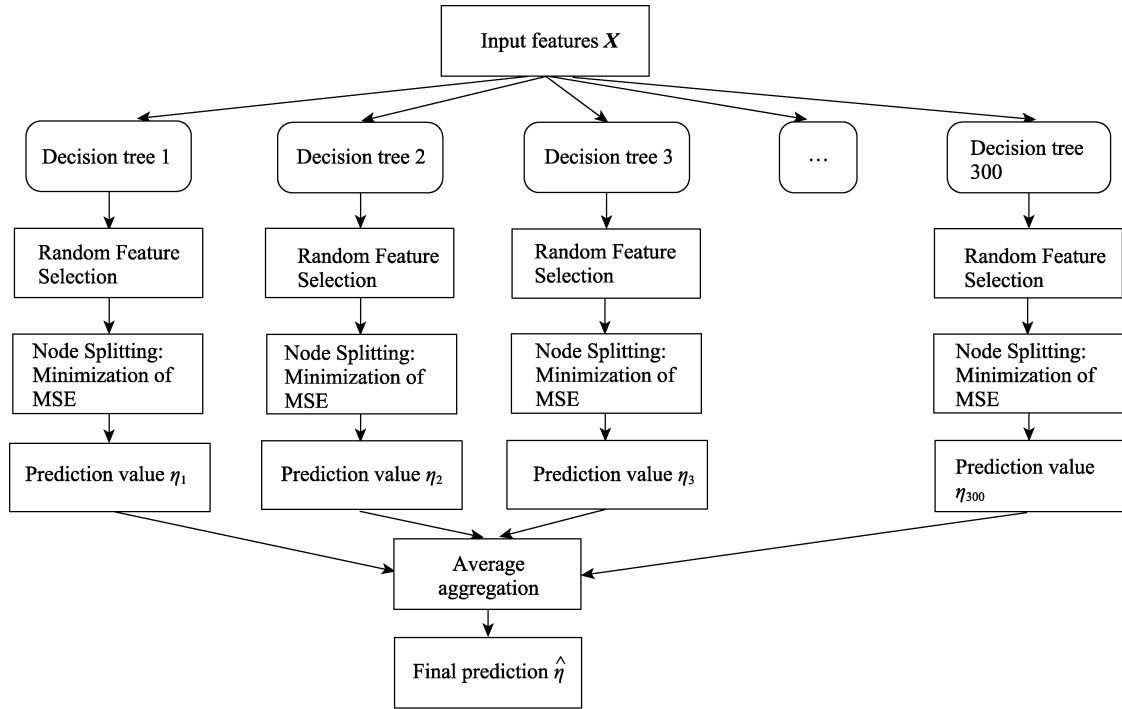


图 5 RF 模型流程图
Fig.5 Flowchart of the RF model

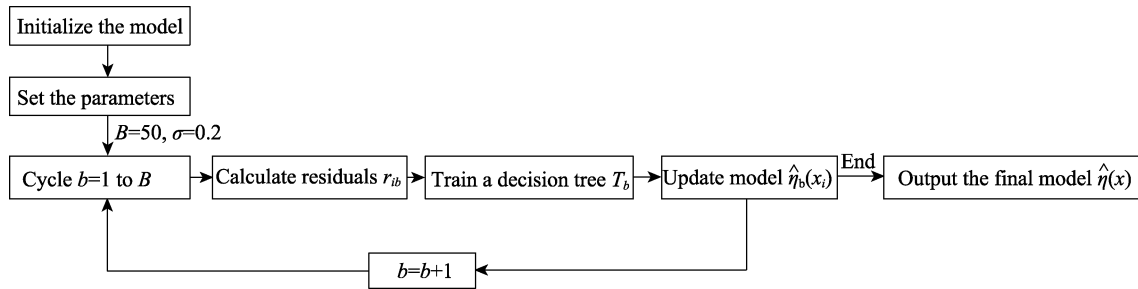


图 6 GBRT 模型流程图
Fig.6 Flowchart of the GBRT model

1.3.2.1 模型形式化

梯度提升回归树通过迭代拟合残差逐步提升模型性能。其模型表示为决策树的加权求和:

$$\hat{\eta}(x) = \sum_{b=1}^B \sigma \cdot T_b(x) \quad (9)$$

式中: b (取 $B=50$) 为决策树数量; σ 为学习率, 经过调试取值 0.2; $T_b(x)$ 代表第 b 棵决策树, 其参数选取通过网格搜索优化。

1.3.2.2 损失函数优化

使用均方误差 (MSE) 作为损失函数:

$$L(\eta_i, \hat{\eta}_i) = \frac{1}{2N} (\eta_i - \hat{\eta}_i(x_i))^2 \quad (10)$$

对于每个样本 i , 计算损失函数关于当前模型 $\hat{\eta}_{b-1}(x_i)$ 的负梯度 (即残差):

$$r_{ib} = - \left[\frac{\partial L(\eta_i, \hat{\eta})}{\partial \hat{\eta}(x_i)} \right] = \eta_i - \hat{\eta}_{b-1}(x_i) \quad (11)$$

在每轮迭代中, 计算当前模型的残差 r_{ib} 并训练

新的决策树 $T_b(x)$ 拟合该残差。通过逐步降低残差, 模型预测精度得到持续提升。

1.4 预测输出 η 的评估

为了评估模型的预测性能, 选用 5 个性能指标, 即平均绝对误差 (MAE)、平均绝对百分比误差 (MAPE)、均方根误差 (RMSE)、决定系数 (R^2), 方差贡献率 (VAF), 表示如下:

$$\delta_{MAE} = \frac{1}{N} \sum_{i=1}^N |\eta_i - \hat{\eta}_i| \quad (12)$$

$$\delta_{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\eta_i - \hat{\eta}_i)^2} \quad (13)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (\eta_i - \hat{\eta}_i)^2}{\sum_{i=1}^N (\eta_i - \bar{\eta}_i)^2} \quad (14)$$

$$\delta_{\text{VAF}} = 100 \cdot \left(1 - \frac{\sum_{i=1}^N (\eta_i - \hat{\eta}_i)^2 / (n-1)}{\sum_{i=1}^N (\eta_i - \bar{\eta}_i)^2 / (n-1)} \right) \quad (15)$$

$$\delta_{\text{MAPE}} = \frac{1}{N} \sum_{i=1}^N \left| \frac{\eta_i - \hat{\eta}_i}{\eta_i} \right| \quad (16)$$

式中: δ_{MAE} 为平均绝对误差; δ_{MAPE} 为平均绝对百分比误差; δ_{RMSE} 为均方根误差; δ_{VAF} 为方差贡献率。 δ_{MAE} 、 δ_{MAPE} 和 δ_{RMSE} 用来衡量预测值和真实值之间的误差大小, R^2 和 δ_{VAF} 表示模型的解释能力。当 δ_{MAE} 、 δ_{MAPE} 和 δ_{RMSE} 越接近 0、 R^2 越接近 1、 δ_{VAF} 越接近 100 时, 模型的性能被认为是优越的。

2 结果与讨论

2.1 样本集数量的有效性分析

90%的数据集用于训练模型, 剩下的 10%用于测试模型的预测性能。研究了样本集数量 N (100~808, 增量为 100) 对 RF、GBRT 2 种机器模型性能的影响。

图 7 为测试集性能指标 R^2 、VAF、MAE 和 RMSE 随 N 的变化趋势, 发现 RF 与 GBRT 模型均表现出早熟特性。当 $N=400$ 、样本容量为中等时, RF 与 GBRT 模型的四项性能指标均已达到最优点 (模型的性能指标最高点), 模型得到了充分学习。当 $N < 400$ 时, 性能指标波动较大。当 $N > 400$ 后, 2 种模型的预测性能反而平稳衰减, 呈现边际收益递减现象, 如当 $N=500$ 时, RF 的 R^2 指标较之 $N=400$ 时降低 5.4%。当 $N=808$ 时, 2 种模型的 MAE 指标达到全样本容量中的最差预测状况。

综上所述, 2 种模型已在样本量达到 $N=400$ 时充分捕捉到了工艺参数与壁厚减薄率之间的核心映射关系。随着样本量的持续增加, 模型性能并未同步提升, 表明其学习能力已趋于饱和。这意味着当前模型架构从新增数据中捕获有效特征的能力已达到上限。因此, 为在保证最佳预测性能的同时兼顾计算效率, 后续分析将聚焦于 $N=400$ 这一最具代表性的条件, 对 2 种模型的性能进行深入对比。

2.2 2 个模型性能指标分析

经过多轮训练, 2 种模型的性能指标如表 3 所示,

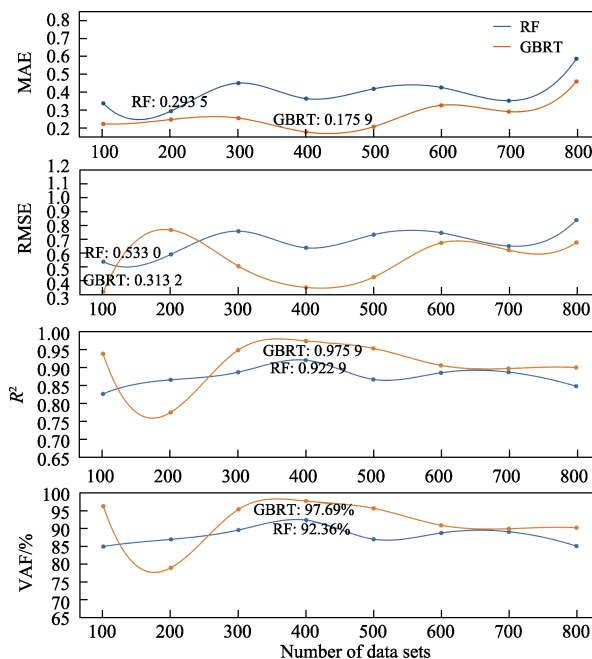


图 7 2 种模型的预测性能在样本集数量上的有效性
Fig.7 Effectiveness of data volume on the prediction performance of two models

将表 3 数据可视化后如图 8 所示。可知, GBRT 在训练集 ($\text{MAE}=0.2520$, $\text{MAPE}=5.7174$, $\text{RMSE}=0.4068$) 和测试集 ($\text{MAE}=0.1759$, $\text{MAPE}=4.8431$, $\text{RMSE}=0.347$) 上都保持最低的 MAE、MAPE 和 RMSE 值, 表明模型在预测目标值上具有更高的准确性。同时, GBRT 在训练集 ($R^2=0.9632$, $\text{VAF}=96.3220$) 和测试集 ($R^2=0.9769$, $\text{VAF}=97.69$) 上始终保持最高的 R^2 和 VAF 值, 表明模型具有最好的数据拟合和解释方差的能力。综合对比, GBRT 模型比 RF 模型的性能指标更优, 在预测准确性和解释方差的能力上更强。

2 种模型的预测值与真实值的对比如图 9 所示, 分析可见, GBRT 的预测值与实际值高度吻合, 进一步验证了模型的预测精度。综上, 数值指标与图形化表征的双重证据, 共同印证了该模型在数据解释与泛化能力方面的可靠性。

2.3 模型的可靠性分析

为验证模型的预测与泛化能力, 本研究使用一组独立于训练样本 D 之外的验证数据 (见表 4) 进行测

表 3 模型预测性能 ($N=400$)
Tab.3 Prediction performance of the model ($N=400$)

Datasets	Model	MAE	RMSE	R^2	VAF	MAPE/%
Train	RF	0.246 4	0.479 3	0.948 9	94.891 1	6.155 4
	GBRT	0.252 0	0.406 8	0.963 2	96.322 0	5.717 4
Test	RF	0.363 8	0.633 5	0.922 9	92.36	8.887 1
	GBRT	0.175 9	0.347	0.976 9	97.69	4.843 1

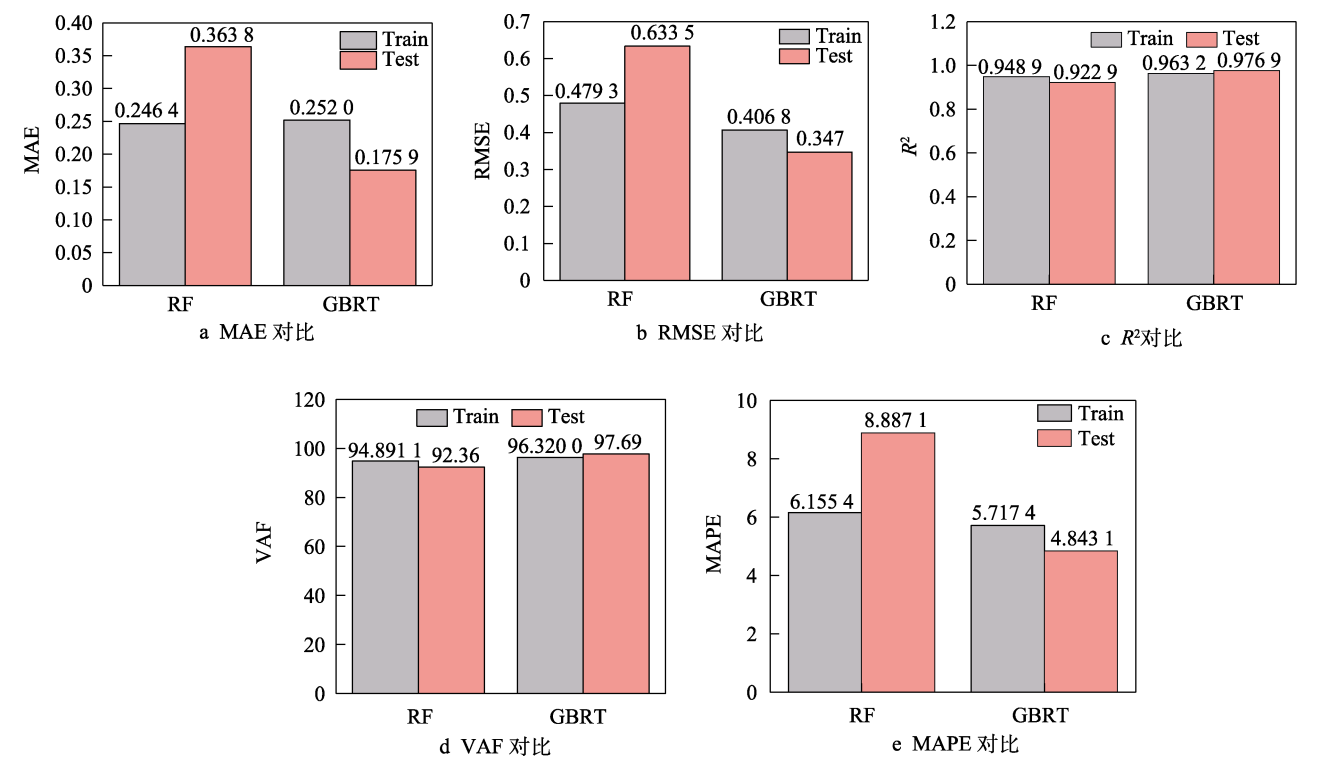


图 8 不同模型的性能评估
Fig.8 Performance evaluations of different models: a) comparison of MAE; b) comparison of RMSE; c) comparison of R^2 ; d) comparison of VAF; e) comparison of MAPE

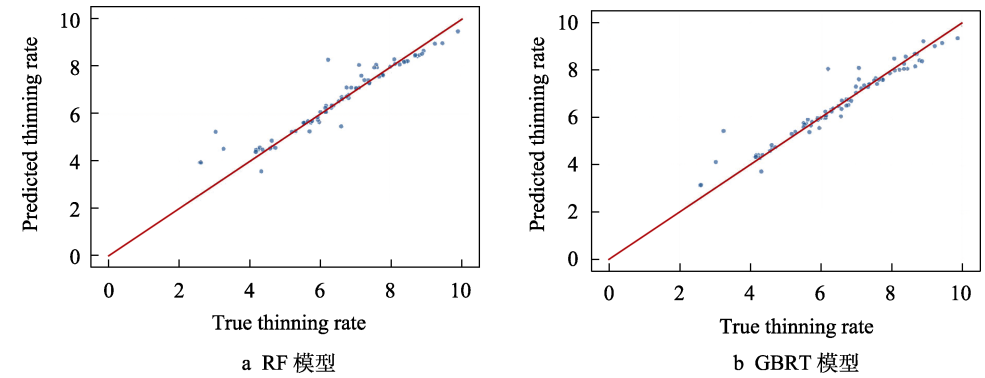


图 9 比较预测值与真实值之间的误差
Fig.9 Comparison of errors between prediction and actual values: a) RF model; b) GBRT model

表 4 几何和工艺参数的符号及取值范围
Tab.4 Symbols and values of geometric and process parameters

Case	λ	ω	β	α	μ	δ
1	0.196	0.225	30	0.5	0.18	0.1
2	0.578	1.5	10	0.5	0.18	0.1
3	0.578	1.32	30	0.5	0.18	0.1
4	0.578	0.56	60	0.5	0.18	0.1
5	0.774	0.45	90	0.5	0.18	0.1
6	0.972	0.45	40	0.5	0.18	0.1
7	1.17	0.375	70	0.5	0.18	0.1

试。结果表明 (图 10), RF 和 GBRT 模型的预测结果均与 FEM 仿真高度吻合, 误差控制在 15%以内。

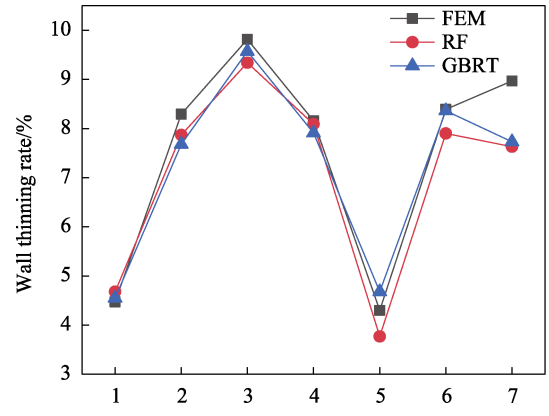


图 10 FEM 与两种模型的预测结果对比
Fig.10 Comparison between prediction results of the two models and the FEM

其中, RF 和 GBRT 模型的最大预测误差分别为 14.9% 和 13.8%, 平均误差分别为 6.9% 和 5.4%。GBRT 模型在精度方面总体优于 RF 模型, 证实其在双金属管弯曲壁厚减薄预测中具有更好的有效性和优越性。

3 结论

本研究首次建立了基于 RF 和 GBRT 2 个机器学习的双金属复合管壁厚减薄预测框架, 通过引入 6 个主要工艺参数, 建立了对应壁厚减薄的非线性映射模型。对比预测精度得到如下结论:

1) RF 与 GBRT 模型在中等样本量 ($N=400$) 下即可达到性能最优, 随着样本量的持续增加, 模型性能并未同步提升, 表明其学习能力已趋于饱和。这意味着当前模型架构从新增数据中捕获有效特征的能力已达到上限。因此, 在保证最佳预测性能的同时兼顾计算效率。在保证数据质量的前提下, 合理控制样本规模对提升建模效率具有重要意义。

2) 在 $N=400$ 的测试集上, GBRT 模型表现出更优的综合预测能力, 其 R^2 、VAF、RMSE 和 MAE 等各项指标均显著优于 RF 模型。具体而言, GBRT 的 R^2 达 0.976 9, RMSE 为 0.347, MAE 为 0.1759, 展现出出色的拟合精度和泛化能力。散点图分析结果进一步表明, GBRT 预测值与真实值吻合度更高, 验证了其在复杂工艺参数映射中的优势。

3) 通过对未参与训练的独立工艺参数点进行预测验证, 2 种模型均表现出良好的泛化能力, 预测误差控制在 15% 以内。其中, GBRT 模型的最大误差和平均误差 (13.8% 和 5.4%) 均低于 RF 模型 (14.9% 和 6.9%), 进一步证实了 GBRT 在壁厚减薄预测中具有更高的准确性和稳健性。

4) 本研究的核心价值在于, 系统地评估与对比了 RF 与 GBRT 模型在双金属复合管绕弯成形预测中的适用性, 明确了 GBRT 模型在中等样本量下的性能优势, 为该工艺的智能化控制提供了有效的工程解决方案。但是本研究仍存在一定局限, 如训练数据主要来源于有限元仿真, 未充分引入实际生产数据; 模型对比仅限于树模型, 未拓展至深度学习等其他架构; 模型可解释性亦有待加强。未来研究可从以下方面展开: 一是引入跨规格、跨材质的迁移学习机制, 提升模型适用性; 二是结合可解释人工智能技术揭示工艺参数与成形缺陷间的物理机制; 三是构建融合物理仿真与机器学习的数字孪生系统, 实现实时预测与闭环控制, 推动该技术在工业中的实际应用。

参考文献:

[1] ZHU Y X, CHEN W, TU W B, et al. Three-Dimensional Finite Element Modeling of Rotary-Draw Bending of

Copper-Titanium Composite Tube[J]. The International Journal of Advanced Manufacturing Technology, 2020, 106(5/6): 2377-2389.

[2] LIU X, LIN Y, FU H, et al. Preparation of the Capillary Copper/Titanium Composite Pipe by Floating-Plug Drawing Processing and Its Microstructure and Properties[J]. Chinese Journal of Engineering, 2017, 39(3): 417-425.

[3] PUGACHEVA N, KRYUCHKOV D, BYKOVA T, et al. Studying the Plastic Deformation of Cu-Ti-C-B Composites in a Favorable Stress State[J]. Materials, 2023, 16(8): 3204.

[4] MOSHIR S K, VAN HOA S, SHADMEHRI F. Structural Analysis of Composite Tubes Using a Meshless Analytical Dimensional Reduction Method[J]. International Journal for Numerical Methods in Engineering, 2021, 122(13): 3191-3217.

[5] JIN K, YUAN Q W, TAO J, et al. Analysis of the Forming Characteristics for Cu/Al Bimetal Tubes Produced by the Spinning Process[J]. The International Journal of Advanced Manufacturing Technology, 2019, 101(1/2/3/4): 147-155.

[6] LI H, SHI K P, YANG H, et al. Significance Analysis of Processing Parameters on Wall Thinning in Tube Bending[J]. Advanced Materials Research, 2012, 622/623: 437-441.

[7] LI H P, LIU Y L, ZHU Y X, et al. Global Sensitivity Analysis and Coupling Effects of Forming Parameters on Wall Thinning and Cross-Sectional Distortion of Rotary Draw Bending of Thin-Walled Rectangular Tube with Small Bending Radius[J]. The International Journal of Advanced Manufacturing Technology, 2014, 74(5): 581-589.

[8] KARTHIKEYAN M, JENARTHANAN M P. Experimental Study of Wall Thinning Behaviour in Boiler Pipes during Bending Process[J]. Australian Journal of Mechanical Engineering, 2020, 18(sup1): S88-S94.

[9] XIONG J Y, LIANG W, DING Y, et al. Quantitative Analysis of Wall Thinning of Bimetallic Clad Steel Tube Based on Pulsed Eddy Current[J]. Process Safety Progress, 2022, 41: 118-128.

[10] HU S H, CHENG C, ABD E A, et al. Forming Characteristics of Thin-Walled Tubes Manufactured by Free Bending Process-Based Nontangential Rotation Bending Die[J]. Thin-Walled Structures, 2024, 194: 111313.

[11] FANG J, OUYANG F, LU S Q, et al. Wall Thinning Behaviors of High Strength 0Cr21Ni6Mn9N Tube in Numerical Control Bending Considering Variation of Elastic Modulus[J]. Advances in Mechanical Engineering, 2021, 13(5): 16878140211021241.

[12] NATH D, ANKIT, NEOG D R, et al. Application of Machine Learning and Deep Learning in Finite Element Analysis: A Comprehensive Review[J]. Archives of Computational Methods in Engineering, 2024, 31(5): 2945-2984.

- [13] VOLKAN GRÜ, BAH M M, EVIK M. Machine Learning for the Prediction of Problems in Steel Tube Bending Process[J]. Engineering Applications of Artificial Intelligence, 2024, 133: 108584.
- [14] YAZDANI A, KNOCHÉ J, ENGEL B, et al. Tube Geometry Prediction in Rotary Draw Bending Process Using Random Forest Regression[J]. Automatisierungstechnik, 2025, 73(4): 223-231.
- [15] AHMAD BHAT S, HUSSAIN I, HUANG N F. Soil Suitability Classification for Crop Selection in Precision Agriculture Using GBRT-Based Hybrid DNN Surrogate Models[J]. Ecological Informatics, 2023, 75: 102109.
- [16] ZHU Z Y, ZHENG Y H, WANG X Y, et al. Forecasting China's Precious Metal Futures Volatility: GBRT Models and Time-Model Dimension Combination of Tree SHAP[J]. International Review of Financial Analysis, 2025, 104: 104249.
- [17] HE C. Enterprise Financial Risk Warning Based on Random Forest Algorithm[J]. Procedia Computer Science, 2025, 261: 1229-1237.
- [18] TIRUMANADHAM N S K M K, PRIYADARSHINI V, PRAVEEN S P, et al. Optimizing Lung Cancer Prediction Models: A Hybrid Methodology Using GWO and Random Forest[M]. Cham: Springer Nature Switzerland, 2025: 59-77.
- [19] SEGHER M E A B, PLEVRIS V, SOLORZANO G. Random Forest-Based Algorithms for Accurate Evaluation of Ultimate Bending Capacity of Steel Tubes[J]. Structures, 2022, 44: 261-273.
- [20] WU T Y, HUNG Y C. Tube Wrinkle State Estimation in Bending Machine Using Servo Signal Analysis and Random Forests[J]. The International Journal of Advanced Manufacturing Technology, 2025, 136(7): 3329-3344.
- [21] 岳峰丽, 郭威, 刘明华, 等. 基于机器学习算法的超声波预测 TP2 铜材晶粒度模型优化研究[J]. 精密成形工程, 2025, 17(9): 185-194.
- [22] YUE F L, GUO W, LIU M H, et al. Optimization of Ultrasonic Prediction of TP2 Copper Grain Size Model Based on Machine Learning Algorithm[J]. Journal of Netshape Forming Engineering, 2025, 17(2): 185-194.
- [22] MALASHIN I, TYNCHENKO V, GANTIMUROV A, et al. Boosting-Based Machine Learning Applications in Polymer Science: A Review[J]. Polymers, 2025, 17(4): 499.
- [23] BARTON N A, HALLETT S H, JUDE S R, et al. Predicting the Risk of Pipe Failure Using Gradient Boosted Decision Trees and Weighted Risk Analysis[J]. NPJ Clean Water, 2022, 5(1): 22.
- [24] ZHAO Y J, LAI M, WU Y Q, et al. Fatigue Life Prediction of Aluminum-Steel Magnetic Pulse Crimped Joints Based on Point Cloud Measurement and Gradient Boosting Regression Trees[J]. International Journal of Fatigue, 2025: 109020.
- [25] WANG G Q, RUAN Y L, WANG H X, et al. Tribological Performance Study and Prediction of Copper Coated by MoS₂ Based on GBRT Method[J]. Tribology International, 2023, 179: 108149.
- [26] LIANG W Z, LUO S Z, ZHAO G Y, et al. Predicting Hard Rock Pillar Stability Using GBDT, XGBoost, and LightGBM Algorithms[J]. Mathematics, 2020, 8(5): 765.
- [27] ZHU Y X, WAN M M, WANG Y, et al. Size Effect Mechanism of Cross-Section Deformation and Section Hollow Coefficient-Bending Degree of the Thin-Walled Composite Bending Tube[J]. Materials & Design, 2021, 212: 110274.
- [28] MCKAY M D, BECKMAN R J, CONOVER W J. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code[J]. Technometrics, 2000, 42(1): 55-61.