

# 基于经验知识的机器学习 CFRP/AL 叠层材料制孔质量分层预测

刘大千, 王共冬, 许成阳, 毕晓莹

(沈阳航空航天大学航空宇航学院, 沈阳 110136)

**摘要:** 在碳纤维增强复合材料的钻削制孔加工过程中会产生很多种类的缺陷, 在所有的缺陷中分层对材料的影响最为严重, 因此, 开发一个能够准确预测叠层材料分层的有效模型至关重要。然而材料领域数据具有小样本、纬度高和关系复杂等特点, 利用经验知识增强机器学习建模效果具有必要性和可行性。通过集成经验知识和数据驱动模型的知识引导机器学习(KGML)模型来预测叠层材料分层, 使经验知识被纳入损失函数作为自适应权重, 以便在训练过程中强制执行物理约束。最后, 通过比较未包含知识模型和包含知识模型的预测性能, 可知包含知识模型的  $R^2$  从 0.79 提升到 0.91, 成功地证明了基于经验知识的机器学习的优势, 提供了一种分层预测的通用方法, 以减少研究人员的实验时间和成本。

**关键词:** 经验知识; 机器学习; 制孔质量; 分层预测

**中图分类号:** TB332    **文献标志码:** A    **文章编号:** 1671-1807(2025)13-0022-06

随着制造业的升级和智能制造的推进, 越来越多的企业开始关注机器学习在材料加工中的应用<sup>[1]</sup>。一些航空企业和科研机构正在探索如何利用机器学习优化复合材料的制孔工艺, 以提高飞机结构的强度和轻量化程度。传统的叠层材料制孔工艺往往依赖经验和试错法, 这不仅耗时, 而且容易导致材料浪费; 机器学习可以通过数据驱动的方法, 或者结合制孔经验知识, 快速识别出最佳的制孔参数, 从而提高制孔的效率和质量。

碳纤维增强复合材料(carbon fiber reinforced polymer, CFRP)与铝合金(AL)相结合组成的轻量化复合构件兼具优良承载强度、刚度与延展性, 在航空航天、车辆船舶、医疗等领域有着广泛的应用<sup>[2]</sup>。众所周知, CFRP 制孔过程中会产生许多缺陷, 其中包括微裂纹、纤维断裂、热损伤和分层。在上述缺陷中, 分层是影响 CFRP 使用情况的关键缺陷之一<sup>[3]</sup>。多年来, 人们在分层预测方面做了大量工作。已经提出了多种用于预测分层的解决方案, 如使用仪器的测量方法、建模分析方法和数据分析方法。Xu 和 Davim<sup>[4]</sup>提出了一种识别切削参数的

方法, 该方法结合使用田口技术和方差分析进行无损制孔。Krishnamoorthy 等<sup>[5]</sup>给出了一个用于预测分层的人工神经网络, 其最大误差为 0.81%, 最小误差为 0.03%。Cui 等<sup>[6]</sup>提出了一种基于多传感器数据预测分层的新方法, 提出了极端梯度提升回归树-差分自回归移动平均(XGBoost-ARIMA)模型, 该模型包括回归和用于滚动预测的顺序部分。Zhang 和 Xu<sup>[7]</sup>提出了一种基于高斯过程回归的方法来预测分层因子, 该方法指的是在钻孔过程后使用仪器预测分层, 以获取其形状、大小和位置数据, 这些几何特征可以提供分层损伤的近似估计。Volety 和 Mani<sup>[8]</sup>通过几种机器学习算法, 如随机森林、线性回归和支持向量机(support vector machine, SVM), 再根据已知输入(如进给率、主轴转速)更有效地预测分层因子。

然而, 材料领域数据往往规模较小、维度更高, 并且它们之间的关系错综复杂; 与广泛分布于互联网上的海量信息不同, 科研人员在多年的研究过程中积累了极为丰富的经验。当使用机器学习方法对材料数据进行建模时, 往往会出现“过拟合”现

**收稿日期:** 2025-01-11

**作者简介:** 刘大千(1998—), 男, 四川广元人, 硕士研究生, 研究方向为飞行器数字化制造技术; 通信作者王共冬(1979—), 男, 江苏徐州人, 博士, 教授, 研究方向为复合材料数字化制造; 许成阳(1989—), 男, 黑龙江哈尔滨人, 博士, 副教授, 研究方向为复合材料与金属叠层材料切削加工; 毕晓莹(1995—), 女, 山东威海人, 博士研究生, 研究方向为先进复合材料结构设计与制备。

象,且与领域专家观点不符。因此,如何有效地融合已有的经验与理论,提高模型的可信度与泛化能力,是最大限度地发挥机器学习的优势。比如,将物理代数方程的正则项约束加入损失函数中,或者将知识图谱和神经网络相结合,可以实现利用经验知识对机器学习的提升,这两种方法被广泛应用于物理场重建与分类等领域<sup>[9]</sup>。

但是,目前在材料领域研究中缺乏一种既能精准捕捉到材料特性的定性规律,又能将这些规律高效地整合到机器学习模型中的方法。因此,本文首先采用机器学习方法搭建叠层材料分层预测模型,然后通过知识引导机器学习(knowledge-guided machine learning, KGML)的方法来重构损失函数,将叠层材料制孔经验知识有效地融入算法之中,进而使迭代优化网络权值来提升建模效果。最后,通过对比两种模型预测能力和准确性,证明知识引导的机器学习具有更强的分层预测能力。

## 1 研究方法

### 1.1 数据收集

在CFRP制孔加工中,影响制孔质量的因素很多,主要体现在加工参数、加工方法和加工材料。在加工参数方面有转速、进给率、层合板铺层方式和刀具的类型结构。随着转速的增加,钻削复合材料的轴向力逐渐减小,从而提高了制孔质量;但是如果转速过高,则在切削时容易积累切削瘤,降低刀具的锋利程度,反而造成缺陷;纤维铺层角度不同,层合板的力学性能则不同,切削时纤维的断裂方式不同,所以表面质量有所差异<sup>[10]</sup>。在加工方法上为提高制孔质量可采用导套、支撑以及预制孔,利用导套保证刀具与加工表面的垂直度,容易获得良好的进孔质量;在孔出口处增设支撑,消除了出口轴向力造成的分层,同时减少了毛刺、撕裂等缺陷;预制孔可消除横刃对复合材料的推挤作用,使轴向力减小,出现缺陷的概率也降低。在加工材料方面有增强体材料和刀具材质对制孔质量有较大的影响,增强纤维的不同会使复合材料层合板的力学强度有差异,在相同参数条件下进行钻削制孔时,T800的撕裂长度为T300的1/3左右;不同的刀具材料对制孔质量的影响也非常明显,由于晶粒细化,带涂层的硬质合金刀具的抗磨损性和钻削质量明显优于非涂层的硬质合金刀具。

决定加工质量和加工缺陷产生的重要因素是切削力,切削力可以分解到 $x$ 、 $y$ 、 $z$ 3个方向上,由于

轴向切削力是分层缺陷产生的主要原因,因此,本文将轴向力作为分层预测的结果,其余影响因素作为数据集的输入参数,最终构建数据集,如图1所示。根据以上影响制孔质量因素,从50篇文献共收集320组数据<sup>[11-18]</sup>。

由于制孔过程中有许多特征因素会影响制孔质量且各个特征因素对轴向力的权重不同,为了挖掘出这些特征因素与轴向力的相互影响,引用相关性分析来描述各特征因素与轴向力之间的关联程度。本文的相关性分析是各特征因素与轴向力之间的相关性分析,进而衡量出特征因素之间的相关性程度。因此,为了提高分层预测模型的预测精度、提高算法的运行速度,采用斯皮尔曼相关性(Spearman Rank correlation)分析方法计算各制孔参数之间的相关性,相关系数的绝对值越大,表明两者之间的相关性程度越高,通过剔除相关性较小的特征因素,选择与轴向力相关性较大的特征因素,有利于精简模型。

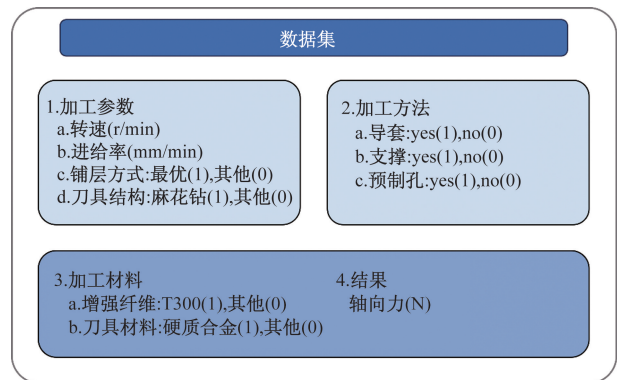


图1 CFRP/AL 叠层材料数据集

### 1.2 预处理

数据预处理的核心目的是为后续的建模和分析提供高质量的数据。首先对数据清洗,包括处理缺失值和处理异常值;其次是数据归一化,将数据缩放到相同的范围(如 $[0, 1]$ )或使数据符合标准正态分布(均值为0,方差为1),最后是特征选择,选择对模型预测最有用的特征,剔除冗余或无关的特征。

Spearman相关性分析结果见表1,从中可以看出轴向力与转速和进给率之间相关性较强,与铺层方式、导套、支撑、预制孔、增强纤维和刀具材质之间相关性较小。

为了加速模型收敛过程,采用Z-score方法对数据进行归一化,将数据转化为均值为0、标准差为1的分布。对于每个样本,计算公式为

表 1 各特征因素的斯皮尔曼相关系数

参数	转速/ ( $r \cdot \min^{-1}$ )	进给率/ ( $\text{mm} \cdot \min^{-1}$ )	最优铺层	导套	支撑	预制孔	增强纤维	刀具材质
相关系数	0.32	0.50	0.12	0.10	0.14	-0.15	0.13	0.12

$$x_i = \frac{x - u}{\sigma} \quad (1)$$

式中： $x$  和  $x_i$  分别为原始数据和标准化数据； $u$  为原始数据的均值； $\sigma$  为标记原始数据的标准差。

### 1.3 模型建立与训练

机器学习模型一般分为 4 类，分别是分类问题、回归问题、聚类问题和降维问题，为实现基于机器学习 CFRP/AL 叠层材料分层预测，本文选择拥有回归问题的 XGBoost 模型。XGBoost 是 boosting 算法的一种实现方式<sup>[17]</sup>，该算法具有精度高、不易过拟合、可扩展、高性能等优点，并且可以分布式处理高维特征值，效率比同类算法高 10 倍以上。该算法可以将分层预测问题转化为机器学习中的监督学习 (super-vised learning) 问题，XGBoost 在构建模型时，每棵决策树之间都有很强的相关性，下一棵决策树都是基于前一棵决策树与轴向力目标的差异进行学习，并进行迭代，直到达到设定的最大迭代次数，从而最终减少分层预测的偏差。因此，最终减少了分层预测的偏差。

XGBoost 通过最小化损失函数和更新树结构来训练模型。损失函数由可微分预测误差和正则化项组成。在更新的第  $t$  步，损失函数可以通过泰勒展开，即

$$L^{(t)} = \sum_{i=1}^n [g_i f_t(x_i) - \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \quad (2)$$

式中： $L^{(t)}$  为第  $t$  个的损失函数值； $g_i$  为损失函数的一阶导数； $h_i$  为损失函数的二阶导数； $f_t(x_i)$  为第  $t$  棵树对第  $i$  个样本的预测值； $\Omega(f_t)$  为正则化项。

### 1.4 知识引导机器学习

知识引导机器学习的流程首先把经验知识转化为规则并将众多规则构建为制孔知识库，然后通过输入制孔参数，根据输入参数的特点从知识库选择对应的规则，选取的规则作为优化模型的约束条件。本文根据实践经验总结出 30 种经验知识，并且通过数学约束得到对应的惩罚项，见表 2。

在机器学习中，模型训练的本质在于损失函数，损失函数的变化会给模型带来不同的预测趋势。一般机器学习的损失函数考虑模型预测值的观

表 2 经验知识

序号	条件	结果	惩罚项
1	转速增加/减少	轴向力减少/增加	$P_1$
2	进给率增加/减少	轴向力增加/减少	$P_2$
3	最优铺层	轴向力减少	$P_3$
...	...	...	...
30	刀具顶角增加/减少 (顶角范围 $90^\circ \sim 180^\circ$ )	轴向力增加/减少	$P_{30}$

测误差，通常是均方误差  $\text{Loss}_{\text{MSE}}$ ，如式(3)所示，忽略经验知识误差  $\text{Loss}_{\text{KGML}}$ ，如式(4)所示。为了解决这个限制，将叠层材料制孔的规则转化为损失函数中的惩罚项纳入误差公式中，从而得到一个改进的损失函数方程。受 Ma 等<sup>[19]</sup>研究的启发，KGML 方法的损失函数增加了自适应知识权重。自适应权重使模型能够辨别各种样本的重要性。例如，物理知识中的预测误差越大，则分配更高的权重。因此，模型将精力集中在拟合这些样本上。

$$\text{Loss}_{\text{MSE}} = \frac{1}{N} \sum_i (y_i - \hat{y}_i)^2 \quad (3)$$

$$\text{Loss}_{\text{KGML}} = \frac{1}{N} \sum_i \lambda_i (y_i - \hat{y}_i)^2 \quad (4)$$

式中： $y_i$  和  $\hat{y}_i$  分别为实际值和模型预测值； $N$  为样本数； $\lambda_i$  为分配给经验知识的权重。

总的损失函数  $\text{Loss}_{\text{Total}}$  是由误差损失函数和经验知识损失函数两部分组成，如式(5)所示，而误差损失函数通过计算 XGBoost 模型预测值与真实值的误差得来，经验知识损失函数通过 KGML 方法来表示。结合多次实验，将经验知识转化为数学约束并嵌入机器学习模型中，在损失函数中加入一个惩罚项，以确保模型的输出符合这一经验规律。

$$\text{Loss}_{\text{Total}} = \text{Loss}_{\text{MSE}} + \text{Loss}_{\text{KGML}} \quad (5)$$

### 1.5 模型预测与评估

将数据集分为训练集和测试集，其中 80% 用于训练，20% 用于测试。为了验证模型预测值的准确性，使用平均绝对误差 (MAE)、均方根误差 (RMSE) 和相关系数 ( $R^2$ ) 作为评估指标。MAE 能准确反映实际预测误差的大小，以及评估真实值和预测值的偏离程度；RMSE 是均方误差的算术平方根，因其以平方根的形式更能突出较大或较小误

差;  $R^2$  用于评估回归模型的解释能力, 表示自变量能够解释因变量方差的比例。MAE、RMSE 和  $R^2$  的计算公式如下:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (6)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (7)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y} - y_i)^2} \quad (8)$$

式中:  $n$  为数据集中的样本数量;  $\bar{y}$  为数据集中实际值的平均值。为了获得更为可靠的评估结果, 通过 MAE、RMSE 和  $R^2$  这些评估指标的平均值, 作为模型的最终性能评估分数。这一分数能够更好地反映模型在整体数据集上的性能, 并消除单次实验验证可能带来的偶然性。

## 2 结果与讨论

为了评估经验知识对机器学习模型性能的影响, 本文建立了包含和不含经验知识的 XGBoost 模型。分别运行包含和不含经验知识的 XGBoost 模型 10 次, 并求取 10 次运行结果的平均数, 基于 MAE、RMSE 和  $R^2$  比较添加经验知识的差异, 对比结果, 见表 3。

从表 3 可以看出, 两种模型都有较好的预测效果, 都能够较准确的预测制孔质量, 决定系数较为接近; 而包含经验知识模型  $R^2$  达到了 0.91, MAE 和 RMSE 都低于不包含经验知识模型。综上所述, 通过知识引导的 XGBoost 模型大大提升了预测制孔质量的精度。

模型的准确度主要体现在对测试集的预测准确性上。模型通过控制树的最大深度来调节树的个数, 不断对模型参数进行调整。运用 Python 和机器学习框架 Pytorch 建立 XGBoost 模型对经过预处理和相关性分析的数据进行学习训练, 得到包含和不含经验知识的分层预测的回归模型, 如图 2 所示。

图 2 展示了模型在测试集上的拟合效果。图中蓝点代表轴向力真实值和预测值的组合, 红色对角线表示预测值和实际值相等。红线周围的点越多, 表示模型拟合越好。从图 2(b) 中可以看出, 包含知识的 XGBoost 模型产生的预测最接近实际值, 几乎所有点都集中在红线附近, 可以看出, 两者的吻合度更好。

在训练完成后, 选取 20% 的数据作为测试集, 数据与经过模型训练得出的预测结果如图 3 所示。

表 3 模型对比结果

参数	不包含知识 XGBoost 模型	包含知识 XGBoost 模型
MAE	8.18	4.28
RMSE	10.39	5.51
$R^2$	0.79	0.91

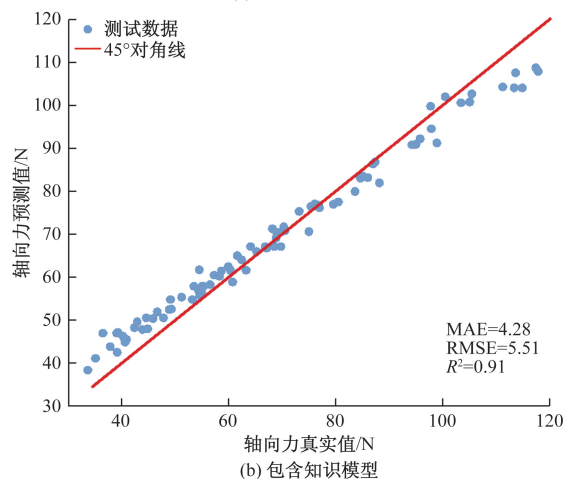
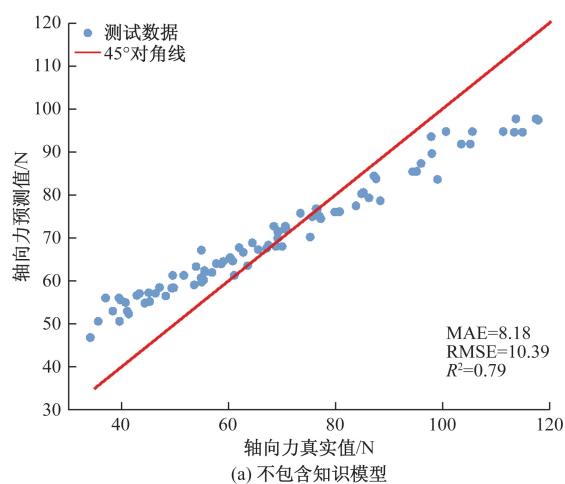


图 2 分层预测回归曲线

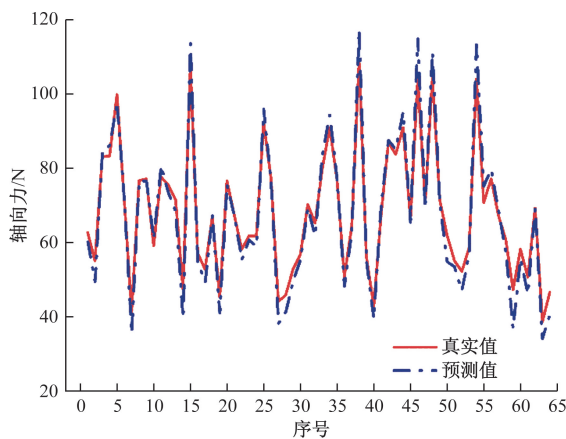


图 3 轴向力预测结果

红色的线为真实值,蓝色的线为 XGBoost 模型预测的值,可以看出两者吻合度较好。结果表明,该模型在预测叠层材料分层方面表现出色, $R^2$ 为 0.91,精度较好。总之,知识引导的 XGBoost 模型预测准确率高、速度快、效率高,不仅能高效地进行数据处理和预测,还节省了实验成本。

### 3 结论

(1)首先将知识的损失函数增加了自适应加权因子嵌入 XGBoost 模型损失函数中,建立了一个集成经验知识增强的机器学习模型。然后通过 CFRP/AL 叠层材料分层预测,并且进行实验验证和机理分析。最终得出基于经验知识引导的机器学习  $R^2$  从 0.79 提升到 0.91,充分证明了该模型的预测能力和准确性。

(2)在样本有限的情况下,证明了经验知识结合的机器学习模型可以有效减少资源浪费,大幅提高研究效率。

未来,为了实现材料领域知识的自动化获取和统一的符号表示,将领域知识嵌入机器学习算法中,对于进一步提高机器学习的通用性、准确性和可解释性具有重要意义。

### 参考文献

- [1] 刘悦,邹欣欣,杨正伟,等. 材料领域知识嵌入的机器学习[J]. 硅酸盐学报, 2022, 50(3): 863-876.
- [2] 白大山,陈五一,陈雪梅. 碳纤维增强复合材料/轻合金叠层结构制孔技术研究进展[J]. 航空制造技术, 2022, 65(9): 82-88.
- [3] 聂鹏,孙国峰,郭勇翼,等. CFRP 制孔分层研究现状与进展[J]. 工具技术, 2023, 57(3): 3-8.
- [4] XU J, DAVIM J P. Investigation of the influence of delamination characteristics on the mechanical properties of CFRP[C]//New Materials, Machinery and Vehicle Engineering: Proceedings of the 2nd International Conference (NMMVE 2023). Amsterdam: IOS Press, 2023, 37: 400.
- [5] KRISHNAMOORTHY A, BOOPATHY S R, PALANIKUMAR K. Delamination prediction in drilling of CFRP composites using artificial neural network[J]. Journal of Engineering Science and Technology, 2011, 6(2): 191-203.
- [6] CUI J, LIU W, ZHANG Y, et al. A novel method for predicting delamination of carbon fiber reinforced plastic (CFRP) based on multi-sensordata[J]. Mechanical Systems and Signal Processing, 2021, 157: 107708.
- [7] ZHANG Y, XU X. Predicting the delamination factor in carbonfiber reinforced plastic composites during drilling through the Gaussian process regression[J]. Journal of Composite Materials, 2021, 55(15): 2061-2068.
- [8] VOLETY R, MANI G. AI-based predictive modeling of delamination factor for carbon fiber-reinforced polymer (CFRP) drillingprocess[J]. Optimization Techniques in Engineering: Advances and Applications, 2023, 1: 139-153.
- [9] CHOI J G, KIM D C, CHUNG M, et al. Multimodal 1D CNN for delamination prediction in CFRP drilling process with industrial robots[J]. Computers & Industrial Engineering, 2024, 190: 110074.
- [10] 王萌,王共冬,李南,等. 加工工艺参数对碳纤维增强复合材料层合板制孔质量的影响[J]. 科学技术与工程, 2020, 20(26): 10683-10689.
- [11] DOMINGUEZ-MONFERRER C, RAMAJO-BALLESTER A, ARMINGOL J M, et al. Spot-checking machine learning algorithms for tool wear monitoring in automatic drilling operations in CFRP/Ti6Al4V/Al stacks in the aircraft industry[J]. Journal of Manufacturing Systems, 2024, 77: 96-111.
- [12] COSTA R, DURO J N S, SOUSA V F C, et al. Drilling of CFRP/Al multi-material stacks using WC-Co CVD diamond coatedtools [J]. Procedia Structural Integrity, 2024, 53: 376-385.
- [13] ZHAO M, WANG F, FU R, et al. Drilling study on CFRP/Al stack with different CFRP thickness using chip-breaking step drill bit[J]. Journal of Manufacturing Processes, 2023, 90: 300-309.
- [14] ALAGAN N T, SAJJA N T, GUSTAFSSON A, et al. Investigation of the quality of Al-CFRP stacks when drilled using innovative approaches[J]. CIRP Journal of Manufacturing Science and Technology, 2023, 43: 260-272.
- [15] JEBARATNAM J M, HASSAN M H. Process enhancement and performance evaluation of single-shot drilling of CFRP/aluminum stacks: areview[J]. The International Journal of Advanced Manufacturing Technology, 2024, 135(7): 3015-3050.
- [16] 李健. 基于力反馈的 CFRP/Al 叠层材料机器人制孔关键技术研究[D]. 哈尔滨: 哈尔滨工业大学, 2023.
- [17] HAMED M, ZHANG C, KHAN A M, et al. Holistic review of drilling on CFRP composites: techniques, FEM, sustainability, challenges, andadvances[J]. The International Journal of Advanced Manufacturing Technology, 2024, 135(5): 2661-2696.
- [18] 索基源,李元奎,崔金龙,等. 基于 XGBoost 算法的船舶油耗预测模型[J]. 中国航海, 2024, 47(2): 153-159.
- [19] MA Y, ZHENG J, LIANG Y, et al. Deeppipe: theory-guided neural network method for predicting burst pressure of corroded pipelines[J]. Process Safety and Environmental Protection, 2022, 162: 595-609.

## Delamination Prediction of Hole-making Quality of CFRP/AL Laminated Materials Based on Empirical Knowledge through Machine Learning

LIU Daqian, WANG Gongdong, XU Chengyang, BI Xiaoxuan

(Aerospace Academy, Shenyang Aerospace University, Shenyang 110136, China)

**Abstract:** Many types of defects are produced during the drilling of holes in carbon fiber reinforced composites, and of all the defects delamination has the most serious effect on the material. Therefore, it is crucial to develop an effective model that can accurately predict delamination in laminated materials. However, materials domain data is characterized by small samples, high latitude and complex relationships, which makes it necessary and feasible to use empirical knowledge to enhance the effectiveness of machine learning modeling. A knowledge-guided machine learning(KGML) model that integrates empirical knowledge and data-driven modeling is used to predict laminated material delamination, the fact that empirical knowledge is incorporated into the loss function as an adaptive weighting in order to enforce physical constraints during the training process. Finally, by comparing the prediction performance of the model without knowledge and the model with knowledge, the  $R^2$  of the model with knowledge was improved from 0.79 to 0.91, which successfully demonstrated the advantages of empirical knowledge-based machine learning, and provide a generalized approach for delamination prediction to reduce the experimentation time and cost for researchers.

**Keywords:** empirical knowledge; machine learning; hole-making quality; delamination prediction