基于机器学习的内置角钢 CFST 抗压承载力预测模型

叶喆轩

浙大城市学院工程学院，浙江 杭州 310015

DOI:10.61369/ETQM.2026090038

摘　　要： 为提高内置角钢钢管混凝土构件抗压承载力的预测精度，本文建立了包含 133 组试验样本的内置角钢 CFST 抗压承载力数据库。选取钢管直径、壁厚、构件长度、钢材屈服强度、混凝土抗压强度、角钢宽度、角钢厚度及角钢屈服强度等 8 个参数作为输入特征，以轴向承载力为输出目标。分别采用随机森林（Random Forest, RF）、极端梯度提升（eXtreme Gradient Boosting, XGBoost）和轻量级梯度提升机（Light Gradient Boosting Machine, LightGBM）三种集成学习算法建立抗压承载力预测模型，采用贝叶斯优化方法实现超参数自适应寻优，并通过 20 次随机重复试验评估模型稳定性。结果表明：XGBoost 模型综合性能最优，20 次重复试验平均决定系数 R^2 达 0.9284，平均绝对误差 MAE 为 322.61kN，均方根误差 RMSE 为 502.87kN；三种模型的特征重要性分析均表明钢管直径、钢材屈服强度与混凝土抗压强度为关键影响参数，角钢参数呈协同增强效应。研究成果可为内置角钢 CFST 构件的智能化承载力预测提供参考。

关 键 词： 内置角钢 CFST；承载力预测；集成学习；贝叶斯优化；XGBoost

Machine Learning Based Prediction Model for Compressive Bearing Capacity of Built-in angle Steel CFST

Ye Zhexuan

School of Engineering, Hangzhou City University, Hangzhou, Zhejiang 310015

Abstract: To improve the prediction accuracy of compressive bearing capacity for concrete-filled steel tubes reinforced with internal latticed steel angles, this paper establishes a compressive bearing capacity database containing 133 test samples. Eight parameters are selected as input features, including steel tube diameter, wall thickness, member length, steel yield strength, concrete compressive strength, angle steel width, angle steel thickness, and angle steel yield strength, with axial bearing capacity as the output target. Three ensemble learning algorithms, namely Random Forest (RF), eXtreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM), are employed to establish prediction models. Bayesian optimization is adopted for adaptive hyperparameter tuning, and model stability is evaluated through 20 randomized repeated experiments. The results indicate that the XGBoost model achieves the optimal comprehensive performance, with an average coefficient of determination R^2 of 0.9284, mean absolute error (MAE) of 322.61 kN, and root mean square error (RMSE) of 502.87 kN across 20 repeated trials. Feature importance analysis of all three models consistently identifies steel tube diameter, steel yield strength, and concrete compressive strength as the key influencing parameters, with angle steel parameters demonstrating synergistic enhancement effects. The research findings can provide reference for intelligent bearing capacity prediction of concrete-filled steel tubes reinforced with internal latticed steel angles.

Keywords: Concrete-Filled Steel Tubes Reinforced with Internal Latticed Steel Angles; Bearing Capacity Prediction; Ensemble Learning; Bayesian Optimization; XGBoost

引言

钢管混凝土（Concrete-Filled Steel Tube，CFST）结构通过在钢管内部浇筑混凝土，使钢管对核心混凝土产生约束作用，从而显著提高构件的承载力和延性[1,2]。CFST 结构兼具钢结构和混凝土结构的优点，具有承载力高、抗震性能好、施工速度快等优势，已广泛应用于高层建筑、大跨桥梁、输电塔架等重要工程领域[3,4]。为进一步提升 CFST 构件的承载性能，研究者提出了多种内部加劲形式，其中内置角钢 CFST 构件通过在钢管混凝土内部设置角钢骨架，形成外钢管、核心混凝土与内置角钢三者协同工作的复合受力体系，能够有效提高构件的轴向承载力和抗变形能力[5,6]。

作者简介：叶喆轩（2001—），男，硕士研究生，主要从事钢管混凝土结构及机器学习在结构工程中的应用研究。

Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.

内置角钢CFST构件的极限承载力受几何参数（钢管直径、壁厚、构件长度、角钢尺寸等）、材料性能（钢材强度、混凝土强度等）以及构造细节（角钢布置方式、缀条间距等）的多因素耦合影响。传统基于简化假定的解析模型或规范公式难以充分考虑各参数之间的复杂非线性交互作用，导致预测精度与适用范围之间存在矛盾[7,8]。近年来，随着人工智能技术的快速发展，机器学习方法在结构工程领域得到了广泛应用，为构件承载力预测提供了新的技术途径[9]。

集成学习（Ensemble Learning）作为机器学习的重要分支，通过构建多个基学习器并对其预测结果进行组合，能够有效提高模型的精度、稳定性和泛化能力[10]。其中，随机森林（Random Forest，RF）[10]、极端梯度提升（eXtreme Gradient Boosting，XGBoost）[11]和轻量级梯度提升机（Light Gradient Boosting Machine，LightGBM）[12]是三种代表性的集成学习算法，已在多个工程领域的预测问题中展现出优异的性能。陈曦泽等[13]基于XGBoost-SHAP方法建立了钢管混凝土柱轴向承载力预测模型，取得了良好的预测效果。然而，针对内置角钢这一特殊加劲形式的CFST构件，现有机器学习预测模型的研究仍较为匮乏，亟待开展系统性研究。

本文以内置角钢CFST构件为研究对象，系统收集公开发表的轴压试验数据，建立包含133组样本的抗压承载力数据库；选取RF、XGBoost和LightGBM三种集成学习算法建立预测模型，采用贝叶斯优化实现超参数自适应寻优；通过多种评价指标对模型性能进行系统对比分析，并结合SHAP（SHapley Additive exPlanations）可解释性方法揭示关键影响参数的作用规律，以期为内置角钢CFST构件的智能化承载力预测提供方法支撑。

一、方法

（一）数据库建立与特征选取

1. 数据来源与筛选

通过系统查阅国内外公开发表的学术文献，收集已有内置型钢（含角钢）钢管混凝土构件的轴压试验数据。数据来源主要包括王军[3]、Cai等[14]、Xu等[15,16]、侯新宇[17]、Yang等[18]等发表的文献数据，涉及《建筑结构学报》、《工程力学》、Engineering Structures、Thin-Walled Structures、Journal of Constructional Steel Research等国内外权威期刊。从文献中提取的数据包括外钢管几何尺寸（直径D、壁厚t_o、长度L）、角钢几何尺寸（宽度b、厚度t_a）、材料性能参数（钢管钢材屈服强度fy、混凝土立方体抗压强度$f$cu、角钢钢材屈服强度f_{ya}）以及对应的极限承载力N_u。

为保证数据质量，仅选取试验过程完整、破坏模式明确且参数记录齐全的试验结果纳入数据库。考虑到不同学者采用的混凝土强度测量方式存在差异，本文将所有文献中混凝土强度统一转换为150mm×150mm×150mm标准立方体抗压强度f_{cu}。经过筛选预处理后，最终数据库包含133组有效样本。

2. 数据库统计特征

数据库各参数的统计信息如表1所示。由表1可知，数据库样本覆盖了较宽的工程取值区间：钢管直径D=250~400mm，壁厚t_o=1.5~10mm，长度L=500~10500mm，钢材屈服强度fy=206~420MPa，混凝土抗压强度f_{cu}=32.9~70MPa；角钢参数中，宽度b=20~70mm，厚度t_a=2~8mm，屈服强度f_{ya}=227~460MPa。输出目标N_u的均值为5053.63kN，标准差为2339.48kN，说明样本分布具有一定的离散性，能够为回归建模提供丰富的监督信息。

参数	N_u(kN)	D(mm)	t_o(mm)	L(mm)	f_y(MPa)	f_{cu}(MPa)
最小值	967	250	1.5	500	206	32.9
最大值	10750	400	10	10500	420	70
平均值	5054	332	4.53	1974	289	44.1
标准差	2331	47	1.29	1739	37	7.7

表1 抗压承载力数据库参数统计

参数	b(mm)	t_a(mm)	f_{ya}(MPa)	—	—	—
最小值	20	2	227	—	—	—
最大值	70	8	460	—	—	—
平均值	45	4.6	289	—	—	—
标准差	11	1.2	41	—	—	—

表1（续）

皮尔逊相关性分析结果如图1表明：承载力与钢管直径D、壁厚t_o、钢材强度f_y、混凝土强度f_{cu}呈正相关，体现了截面尺度与材料强度对承载力的主导作用；与长度L呈负相关，提示样本中存在稳定性效应。角钢参数b与t_a呈强共线性，说明仅依赖线性相关难以揭示变量的独立贡献，需采用能够刻画非线性交互效应的集成学习模型。

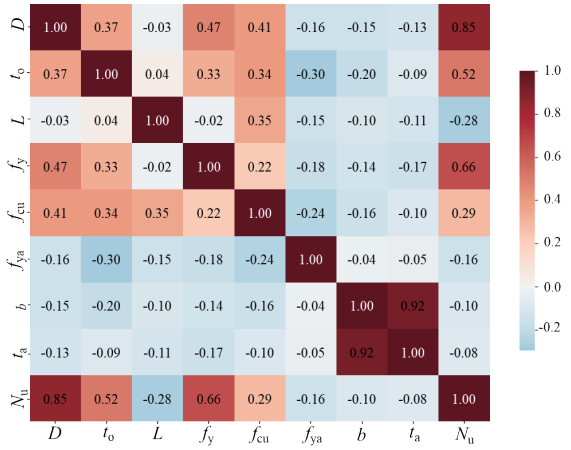


图1 抗压数据库各参数之间的皮尔逊相关系数矩阵

（二）集成学习算法原理

本文选取三种代表性集成学习算法建立内置角钢CFST抗压承载力预测模型，各算法原理简述如下。

（1）随机森林（RF）算法。RF是基于Bagging（Bootstrap Aggregating）框架的并行集成方法[19]。通过对训练样本进行有放回的随机采样（Bootstrap采样）构建多棵决策树，在节点分裂时随机选取部分特征参与最优划分搜索，最终对所有决策树的预测

 Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.

结果取平均值作为输出。RF 通过引入样本随机性和特征随机性，有效降低了不同树之间的相关性，增强了模型的泛化能力和抗过拟合能力。

（2）XGBoost 算法。XGBoost 是一种基于梯度提升框架的串行集成方法[20]。其核心思想是通过多轮迭代逐步构建一系列弱学习器（决策树），每一轮新树拟合前序模型的残差，最终将所有树的预测结果累加作为输出。XGBoost 在目标函数中引入了一阶和二阶梯度信息，并增加了 L1/L2 正则化项以控制模型复杂度，有效抑制过拟合。此外，XGBoost 支持列抽样（column subsampling）和行抽样（row subsampling），进一步提升了模型的泛化性能。

（3）LightGBM 算法。LightGBM 同样基于梯度提升框架，但在树构建策略上进行了创新[12]。不同于 XGBoost 采用的 Level-wise（按层生长）策略，LightGBM 采用 Leaf-wise（按叶子生长）策略，每次选择分裂增益最大的叶子节点进行分裂，在保持预测精度的同时显著提升了训练效率。此外，LightGBM 引入了基于直方图的分裂点查找算法和 Exclusive Feature Bundling（EFB）技术，有效降低了内存消耗和计算复杂度，适合大规模数据集的快速训练。

（三）贝叶斯超参数优化

集成学习模型的预测性能在很大程度上依赖于超参数的设置。本文采用贝叶斯优化（Bayesian Optimization）对三种算法的超参数进行自适应寻优[21,22]。贝叶斯优化通过概率代理模型（通常采用高斯过程）在超参数空间中拟合目标函数的未知映射关系，并利用采集函数（Expected Improvement，EI）在探索未知区域和利用已知最优区域之间取得平衡，从而以较少的评估次数快速逼近全局最优解。相比网格搜索和随机搜索，贝叶斯优化能够在有限评估预算下更高效地找到最优超参数组合。本文基于 Optuna 框架实现贝叶斯优化，以5折交叉验证的 R^2 作为优化目标，各算法经过200次迭代寻优确定最优超参数。贝叶斯优化流程如图2所示。

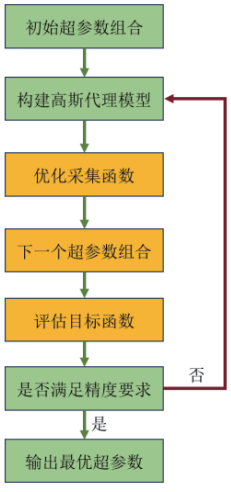


图2 贝叶斯优化流程

（四）模型训练框架

本文采用统一的模型训练框架，以保证不同算法之间的可比性。具体步骤如下：（1）数据清洗与预处理，构建标准化后的输入特征矩阵与输出目标向量；（2）划分训练集与测试集（80%、20%），采用分层随机划分确保训练集和测试集的样本分布一致性；（3）基于训练集进行5折交叉验证，利用贝叶斯优化确定各算法的最优超参数；（4）使用最优超参数在完整训练集上训练最终模型，在独立测试集上评估泛化性能；（5）进行20次不同随机种子的重复划分试验，统计各评价指标的均值与标准差，评估模型对样本划分的敏感程度；（6）基于 SHAP 方法对最优模型进行可解释性分析，揭示各输入特征对预测结果的影响规律和交互效应。

（五）评价指标

为全面客观地评价各模型的预测性能，本文选取以下三个常用指标[23]：

（1）决定系数（Coefficient of Determination，R^2）。R^2 衡量模型对目标变量方差的解释能力，取值范围为 $(-\infty, 1]$，越接近1说明模型拟合效果越好。计算公式为：

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

（2）平均绝对误差（Mean Absolute Error，MAE）。MAE 反映预测值与真实值之间绝对误差的平均水平，对异常值不敏感，能稳健反映模型的典型误差。计算公式为：

$$MAE = \frac{\sum_{i=1}^{N}|\hat{y}_i - y_i|}{N}$$

（3）均方根误差（Root Mean Square Error，RMSE）。RMSE 对较大误差赋予更高权重，能够灵敏地暴露可能导致严重安全风险的预测失准样本。计算公式为：

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

式中：$\hat{y}_i$ 为预测值；yi 为实际值；为预测平均值。

二、结果

（一）最优超参数

通过贝叶斯优化，三种算法获得的最优超参数如表2所示。RF 模型的最优参数组合为：决策树数量 n_estimators=2000，最大深度 max_depth=40，最大特征比例 max_features ≈ 0.727，最小分裂样本数 min_samples_split=2，叶节点最小样本数 min_samples_leaf=1，不采用自助采样 bootstrap=False。该配置下5折交叉验证最优 R^2=0.9513。

XGBoost 模型的最优参数为：最大深度 max_depth=3，学习率 learning_rate ≈ 0.04448，迭代次数 n_estimators=3886，子采样比例 subsample ≈ 0.8021，列采样比例 colsample_bytree= 1.0，正则化参数 γ=5.0，叶节点最小权重 min_child_weight=1.0。对应交叉验证最优 R^2=0.9674。

Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.

LightGBM 模型的最优参数为：学习率 learning_rate=0.2，迭代次数 n_estimators=3490，最大深度 max_depth=2，叶节点数 num_leaves=128，叶节点最小样本数 min_child_samples=3，子采样比例 subsample=0.7，列采样比例 colsample_bytree ≈ 0.9183。对应交叉验证最优 R^2 =0.9496。

超参数	RF	XGBoost	LightGBM
n_estimators	2000	3886	3490
max_depth	40	3	2
learning_rate	—	0.04448	0.2
subsample	—	0.8021	0.7
colsample_bytree	—	1.0	0.9183
max_features	0.727	—	—
min_samples_split	2	—	—
min_samples_leaf	1	—	—
min_child_weight	—	1.0	—
γ	—	5.0	—
num_leaves	—	—	128
min_child_samples	—	—	3
bootstrap	False	—	—

表2 贝叶斯优化后各算法最优超参数

（二）各模型预测结果

1.RF 模型

RF 模型在独立测试集上的评价指标为：R^2=0.852，MAE=513.74kN，RMSE=835.08kN。从实际值与预测值对比分析可知，多数样本点分布在参考线 y = x 附近，总体拟合效果较好，但在高承载力区间（N_u>8000kN）存在少数偏离较大的样本点，导致 RMSE 显著高于 MAE。

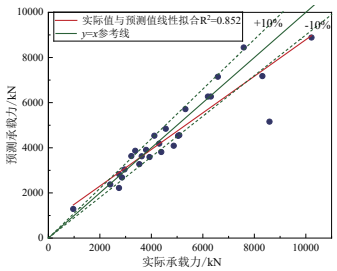


图3 RF 模型实际值与预测值对比

进行20次随机重复划分稳定性检验，统计结果为：R^2 均值0.9110，标准差0.1220；MAE 均值390.24kN，标准差87.95kN；RMSE 均值566.02kN，标准差194.34kN。RF 模型从平均意义上具有较好的总体预测能力，但 R^2 的波动范围较宽，反映出在样本量有限的情况下，不同随机划分会导致较明显的性能波动。

特征重要性分析表明，钢管直径 D 的重要性最高（重要性系数约0.46），其次为混凝土强度 f_{cu}（约0.26）和钢材强度 f_y（约0.16），角钢参数的整体贡献相对较低。该排序与轴压承载力机理一致：D 直接控制截面面积和钢管约束效应，f_{cu} 决定核心混凝土承载力，f_y 影响钢管和角钢的强度贡献。

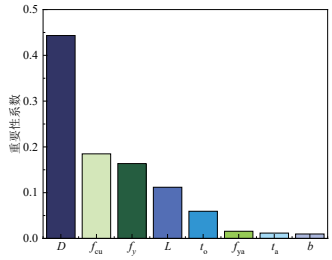


图4 RF 模型特征重要性排序

2.XGBoost 模型

XGBoost 模型在独立测试集上的评价指标为：R^2 =0.9003，MAE=399.35kN，RMSE=659.52kN。与 RF 模型相比，XGBoost 的散点图中样本点更紧密地分布在 y=x 参考线附近，整体离散性更小，特别是在中高承载力区间的预测精度有显著提升。图5为 XGBoost 模型的实际值与预测值对比图。

20次随机重复划分统计结果为：R^2 均值0.9284，标准差0.0941；MAE 均值322.61kN，标准差85.36kN；RMSE 均值502.87kN，标准差193.90kN。与 RF 相比，XGBoost 的平均精度更高且波动幅度更小，说明其整体泛化鲁棒性更强。

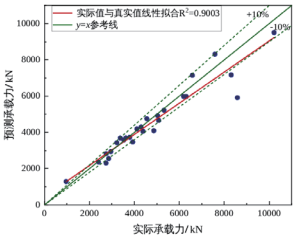


图5 XGBoost 模型实际值与预测值对比

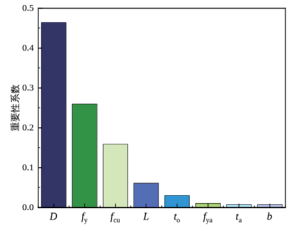


图6 XGBoost 模型特征重要性排序

图6为 XGBoost 模型的特征重要性排序图。结果显示，D 的重要性最高，f_y 次之，f_{cu} 位列第三，三者合计贡献超过80%。构件长度 L 表现为次要影响因子，而角钢参数（f_{ya}、b、t_a）贡献较小，说明在外钢管和混凝土主导承载的机制下，角钢主要起局部增强作用。

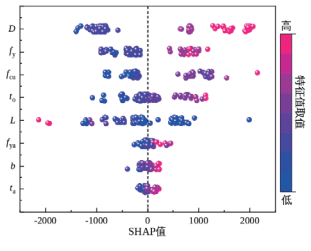


图7 XGBoost 模型 SHAP 可解释性分析

图7为 SHAP 可解释性分析结果。从图中可以观察到：D 的 SHAP 值随直径增大由负转正并显著增大，表明更大的钢管直径稳定地提高预测承载力；f_y 增大整体提升 SHAP 值，且在高 f_{cu} 条件下边际增益更明显，体现钢－混凝土组合构件中材料性能的耦合增强效应；f_{cu} 呈现清晰的正向趋势，低强混凝土区间 SHAP 多为负值，高强区间转为显著正值。角钢参数的 SHAP 值主要分布在零值附近，呈对称分布，表明其对承载力的影响较为有限且与钢管参数存在交互作用。

3.LightGBM 模型

LightGBM 模型在独立测试集上的评价指标为：R^2=0.8751，

 Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.

MAE=446.75kN，RMSE=738.17kN。大部分样本点较接近 y=x 参考线，但在高承载力区仍存在一定偏离，特别是在 N_u>7000kN 的区间，预测偏差相对较大。

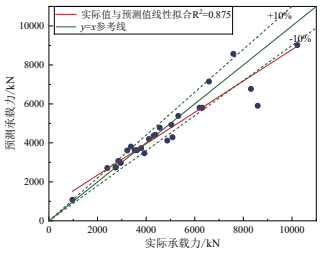


图8 LGBM 模型实际值与预测值对比

20次随机重复划分统计结果为：R^2 均值0.9134，标准差0.0728；MAE 均值393.58kN，标准差82.40kN；RMSE 均值597.69kN，标准差142.56kN。LGBM 相较 RF 模型对随机划分的敏感性更低、稳定性更好（R^2标准差仅为0.0728，三者中最小），但其均值精度略低于 XGBoost。

LightGBM 的特征重要性排序与 XGBoost 基本一致，D 仍为首要控制因素，其次为 f_{cu} 与 f_y。但 LightGBM 对角钢参数的权重分配略高于 XGBoost，说明其在特征选择策略上的差异可能导致对次要因素的不同侧重。

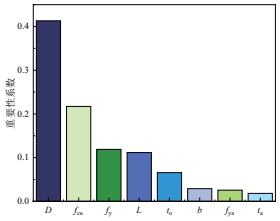


图9 LGBM 模型特征重要性排序

（三）三种模型性能对比

三种集成学习模型20次随机重复试验评价指标对比如表3所示。

评价指标	RF	XGBoost	LightGBM
R^2	0.9110	0.9284	0.9134
MAE/kN	390.24	322.61	393.58
RMSE/kN	566.02	502.87	597.69
R^2标准差	0.1220	0.0941	0.0728
MAE 标准差 /kN	87.95	85.36	82.40
RMSE 标准差 /kN	194.34	193.90	142.56

表3 三种模型20次重复试验评价指标对比

从表3可以得出以下结论：

（1）在拟合优度方面，XGBoost 模型的 R^2 均值最高（0.9284），较 RF（0.9110）提升约1.9%，较 LightGBM（0.9134）提升约1.6%。三种模型的 R^2 均超过0.91，说明集成学习算法整体适用于内置角钢 CFST 抗压承载力预测任务。

（2）在误差指标方面，XGBoost 的 MAE（322.61kN）和 RMSE（502.87kN）均为三者最低，相比 RF 分别降低17.3% 和11.2%，相比 LightGBM 分别降低18.0% 和15.9%。XGBoost 在保持较高拟合优度的同时，对偏差较大的样本具有更强的控制能力。

（3）在稳定性方面，LightGBM 的 R^2 标准差最小（0.0728），表现出最佳的运行稳定性；XGBoost 次之（0.0941）；RF 的波动最大（0.1220）。然而 LightGBM 的均值精度略低于 XGBoost，说明其在追求稳定性的同时牺牲了一部分预测精度。

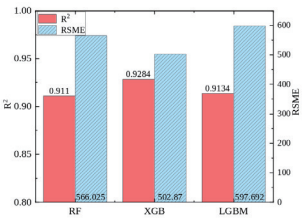


图10 三种模型评价指标对比

图10为三种模型评价指标对比图。从图中可以直观看出，XGBoost 在 R^2、MAE 和 RMSE 三个指标上均占据优势地位。

三、讨论

（一）模型性能差异分析

三种集成学习算法在抗压承载力预测任务中表现出不同的性能特征，其原因可从算法机理和数据特征两个角度进行分析。从算法机理角度看，RF 作为 Bagging 框架的代表，通过并行构建多棵决策树并取平均的方式降低方差，但其对所有树赋予相同权重，难以突出对预测贡献更大的子模型。XGBoost 作为 Boosting 框架的代表，通过串行迭代逐步修正前序模型的错误，对难预测样本赋予更高权重，从而更有效地捕捉复杂非线性关系。本研究中 XGBoost 表现最优，可能原因在于内置角钢 CFST 构件的承载力受多参数耦合影响，不同样本的预测难度差异较大，XGBoost 的梯度提升机制更适应这种数据特征。LightGBM 的 Leaf-wise 策略虽然训练效率更高，但可能导致对训练数据的过拟合倾向，特别是在样本量有限的情况下，其分裂策略的激进性可能限制了泛化性能。

（二）特征重要性机理分析

三种模型的特征重要性排序总体一致：D>fy>fcu>L>to>fya>ta>b，该排序与内置角钢 CFST 构件的轴压承载力机理高度吻合。钢管直径 D 作为首要影响因素，直接决定了截面面积、钢管约束效应和构件抗弯刚度。SHAP 分析显示，当 D>350mm 时，SHAP 值急剧上升至 +1000以上，且与钢材屈服强度 f_y 存在强交互作用，说明大直径构件在高材料强度条件下能够实现承载力的显著提升。

钢材屈服强度 f_y 和混凝土抗压强度 f_{cu} 分别位列第二和第三。值得注意的是，f_y 和 f_{cu} 之间存在显著的协同增强效应：当 f_{cu} 处于高水平时，f_y 的边际增益更明显。这一发现与钢管混凝土的约束增强机理一致，高强度混凝土在钢管约束下能够发挥更高的承载潜力，而高强度钢材则提供了更强的约束能力和材料强度贡献。

角钢参数的整体贡献相对较低，但这并不意味着角钢对承载力无影响。SHAP 分析表明，角钢参数的作用呈现局部修正特征：在钢管参数和混凝土强度确定的前提下，合理的角钢配置能够进一步提升承载力约5%~15%。这一结论与既有试验研究的观测结果相符[5,6]。

（三）模型局限性与改进方向

本文研究存在以下局限性：（1）数据库样本量有限，对于部

Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.

分参数组合的覆盖不够充分，可能影响模型在极端参数条件下的预测可靠性；（2）输入特征未考虑角钢布置方式、缀条间距等构造细节参数，这些参数可能对承载力产生不可忽视的影响；（3）本文仅针对轴压承载力进行预测，尚未拓展至偏压、受弯等复杂受力状态。后续研究可从以下方向改进：（1）持续扩充试验数据库，纳入更多文献数据和数值模拟结果，提升模型的泛化能力；（2）引入角钢构造细节参数，建立更精细化的预测模型；（3）探索物理约束与数据驱动融合建模策略，将经典力学机理嵌入机器学习框架，提升模型在分布外样本上的预测可靠性；（4）将预测模型拓展至偏压、受弯及压弯复合受力状态，构建多工况综合预测体系。

四、结论

本文基于集成学习算法建立了内置角钢 CFST 构件抗压承载力预测模型，通过对 RF、XGBoost 和 LightGBM 三种算法的系统比较和深入分析，得出以下主要结论：

（1）建立了包含133组试验样本的内置角钢 CFST 抗压承载力数据库，输入特征涵盖钢管几何参数、材料强度参数及角钢构造参数，样本覆盖了较宽的工程取值区间。经 Z-score 标准化预处理和80%、20% 随机划分后，能够满足机器学习回归建模的数据需求。

（2）分别采用 RF、XGBoost 和 LightGBM 三种集成学习算法建立了抗压承载力预测模型，通过贝叶斯优化实现了超参数自适应寻优。20次随机重复试验结果表明，XGBoost 模型综合性能最优，平均 R^2 达0.9284，MAE 为322.61kN，RMSE 为502.87kN，在精度和稳定性之间取得了较好平衡，推荐作为内置角钢 CFST 构件抗压承载力预测的首选数据驱动模型。

（3）SHAP 可解释性分析表明，钢管直径 D、钢材屈服强度 f_y 和混凝土强度 f_{cu} 是抗压承载力的关键影响参数，三者之间存在显著的协同增强效应；角钢参数对承载力呈局部修正作用，在钢管主控参数确定的前提下可进一步提升承载力约5%~15%。

（4）三种模型均存在一定的局限性，后续研究可通过扩充数据库规模、引入构造细节参数以及探索物理 - 数据融合建模策略进一步提升模型的预测精度和泛化能力。

致　谢

感谢指导老师浙大城市学院工程学院王军以及各位数据库来源作者在文献查阅和数据整理过程中提供的支持和帮助。

参考文献

[1] 韩林海 . 钢管混凝土结构理论与实践 [M]. 北京：科学出版社，2007.

[2] 韩林海，杨有福 . 现代钢管混凝土结构技术 [M]. 北京：中国建筑工业出版社，2007.

[3] 王军 . 内配格构式角钢圆钢管混凝土力学性能研究及应用 [D]: [博士学位论文]. 杭州：浙江大学，2018.

[4] 陈哲，王军，俞培祥，等 . 内配格构式角钢圆钢管混凝土梁受弯承载力试验及计算 [J]. 钢结构，2019，34(2)：41-45.

[5] 何益斌，张铁成，肖阿林，等 . 极限平衡法分析钢骨 - 钢管混凝土短柱承载力的简化方法 [J]. 铁道科学与工程学报，2006，3(5)：46-49.

[6] 史艳莉，王文达，王亚伟 . 内配型钢方钢管混凝土轴压短柱力学性能及参数分析 [J]. 工程力学，2014，31(1)：201-206.

[7] 王清湘，赵大洲，关萍 . 轴心受压钢骨 - 钢管高强混凝土组合柱力学性能的研究 [J]. 东南大学学报 (自然科学版)，2002，32(5)：710-714.

[8] 陈宝春，韦建刚 . 钢管混凝土拱桥研究现状与展望 [J]. 建筑结构学报，2020，41(8)：1-15.

[9] Tran V L, Thai D K, Kim S E. Application of ANN in Predicting ACC of SCFST Column[J]. Composite Structures, 2019, 228: 111332.

[10] Breiman L. Random Forests[J]. Machine Learning, 2001, 45(1): 5-32.

[11] Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2016: 785-794.

[12] Ke G, Meng Q, Finley T, et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree[C]//Advances in Neural Information Processing Systems. Long Beach: Curran Associates, 2017: 3146-3154.

[13] 陈曦泽，贾俊峰，白玉磊，等 . 基于 XGBoost-SHAP 的钢管混凝土柱轴向承载力预测模型 [J]. 浙江大学学报 (工学版)，2023，57(6)：1061-1070.

[14] Cai J M, Pan J L, Wu Y F. Mechanical Behavior of Steel-Reinforced Concrete Filled Steel Tubular (SRCFST) Columns under Uniaxial Compressive Loading[J]. Thin-Walled Structures, 2015, 97: 1-10.

[15] Xu F, Chen J, Jing W L. Experimental Investigation of Thin-Walled Concrete Filled Steel Tube Columns with Reinforced Lattice Angle[J]. Thin-Walled Structures, 2014, 84: 59-67.

[16] 徐菲 . 薄壁圆钢管混凝土构件与节点力学性能研究 [D]: [博士学位论文]. 杭州：浙江大学，2016.

[17] 侯新宇 . 格构式钢骨 - 钢管混凝土构件轴压性能及承载力计算方法研究 [D]: [硕士学位论文]. 西安：西安建筑科技大学，2020.

[18] Yang Z M, Chen J, Wang F Y, et al. Seismic Performance of Circular Concrete-Filled Steel Tube Columns Reinforced with Inner Latticed Steel Angles[J]. Journal of Constructional Steel Research, 2023, 201: 107735.

[19] Breiman L. Bagging Predictors[J]. Machine Learning, 1996, 24(2): 123-140.

[20] Friedman J H. Greedy Function Approximation: A Gradient Boosting Machine[J]. Annals of Statistics, 2001, 29(5): 1189-1232.

[21] Snoek J, Larochelle H, Adams R P. Practical Bayesian Optimization of Machine Learning Algorithms[C]//Advances in Neural Information Processing Systems 25. Lake Tahoe: Curran Associates, 2012: 2951-2959.

[22] Shahriari B, Swersky K, Wang Z Y, et al. Taking the Human out of the Loop: A Review of Bayesian Optimization[J]. Proceedings of the IEEE, 2016, 104(1): 148-175.

[23] Hyndman R J, Koehler A B. Another Look at Measures of Forecast Accuracy[J]. International Journal of Forecasting, 2006, 22(4): 679-688.

 Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.