文章编号:1674-8190(2026)04-191-11

基于LSTM-XGBoost的民航高原航线运行负荷评估模型研究

张先浩[1],王晨旭[2],冯江垚[1]
(1.中国民用航空飞行学院 民航飞行技术与飞行安全科研基地,广汉 618307)
(2.中国民用航空飞行学院 航空电子电气学院,成都 641419)

摘 要: 民航高原航线特殊工况易引发运行负荷评估难、单一模型精度不足等问题。为此,提出一种面向高原航线场景的长短期记忆网络(LSTM)—极端梯度提升(XGBoost)融合评估模型。依据《中国民用航空规章第121部》及飞行品质监控(FOQA)咨询通告,从飞机性能、气象环境、飞行操作3个维度选取22项指标,经灰色关联度分析与熵权法筛选出15项关键指标;以某航A320机型执飞的6条高原航线8 640组快速存取记录器(QAR)实测数据为样本,经预处理后由LSTM提取时序特征、XGBoost融合静动态特征建模;采用准确率、均方根误差(RMSE)、平均绝对误差(MAE)对模型性能进行综合评价,并结合典型复杂工况开展鲁棒性验证。结果表明:该模型在测试集上评估准确率为96.30%,RMSE和MAE分别为0.042和0.031,在典型复杂工况下准确率仍保持在92%以上,整体性能优于各对比模型。

关键词: 高原航线;运行负荷;LSTM-XGBoost;QAR数据;熵权法;灰色关联度分析
中图分类号: V35; TP183 **文献标识码:** A
DOI: 10.16615/j.cnki.1674-8190.2026.04.19

Research on operational load assessment model for civil aviation high-plateau routes based on LSTM-XGBoost

ZHANG Xianhao[1], WANG Chenxu[2], FENG Jiangyao[1]
(1. Civil Aviation Flight Technology and Flight Safety Research Base, Civil Aviation Flight University of China, Guanghan 618307, China)
(2. Institute of Electronic and Electrical Engineering, Civil Aviation Flight University of China, Chengdu 641419, China)

Abstract: This paper presents a hybrid LSTM-XGBoost model for operational load assessment on civil aviation high-plateau routes, addressing the limitations of single-model methods. Following CAAC Part 121 and FOQA guidelines, 22 initial indicators (aircraft performance, meteorology, flight operations) are reduced to 15 key ones via grey relational analysis and entropy weighting. Using 8 640 QAR data records from A320 flights on six high-plateau routes, LSTM extracts temporal features and XGBoost fuses static and dynamic features for modeling. On the test set, the model attains 96.30% accuracy, with RMSE of 0.042 and MAE of 0.031. Under complex conditions, accuracy stays above 92%. The hybrid approach outperforms benchmarks, offering a reliable tool for quantitative load evaluation and safety decisions on high-altitude operations.

Key words: plateau air route; operational load; LSTM-XGBoost; QAR data; entropy weight method; grey relational analysis

收稿日期: 2026-02-04; **修回日期:** 2026-07-03
基金项目: 中央高校基本科研业务费;中国民用航空局飞行技术与飞行安全重点实验室项目(FZ2025ZX07)
通信作者: 王晨旭(2006—),男,本科生。E-mail: 1252888738@qq.com
引用格式: 张先浩,王晨旭,冯江垚. 基于LSTM-XGBoost的民航高原航线运行负荷评估模型研究[J]. 航空工程进展, 2026, 17(4): 191-201.
ZHANG Xianhao, WANG Chenxu, FENG Jiangyao. Research on operational load assessment model for civil aviation high-plateau routes based on LSTM-XGBoost[J]. Advances in Aeronautical Science and Engineering, 2026, 17(4): 191-201. (in Chinese)

0 引 言

高原航线通常指途经海拔1 524～2 438 m区域、受高原地形与气候影响显著的航线[1]。受空气密度降低、复杂地形及气流扰动等因素影响，高原航线运行过程中更易出现载荷波动与操纵负荷增加等问题[2]。研究表明，高原区域高强度颠簸事件发生率明显高于平原地区，易增加机组操作压力和飞机结构疲劳风险[3]。因此，建立面向高原航线的运行负荷评估方法，对于提升运行安全监测与风险预警能力具有重要意义。

现有运行负荷评估方法多采用单一机器学习模型开展建模分析。典型模型包含支持向量机(Support Vector Machine，SVM)、反向传播神经网络(Backpropagation Neural Network，BPNN)等，该类方法可依托海量飞行运行数据挖掘负荷变化内在规律，但仅依靠单一模型难以兼顾长时序依赖挖掘、静动态特征耦合关联建模，同时在湍流、结冰等典型复杂工况下鲁棒性存在明显短板。Lee等[4]基于机器学习算法完成了民航安全事件关键影响参数识别；Li X Y等[5]依托快速存取记录器(Quick Access Recorder，QAR)飞行数据搭建民用飞机跑道冲出风险评估模型；Rey等[6]构建面向全量飞行数据的驱动型安全风险识别框架。上述研究均采用单一机器学习模型完成航空安全相关评估建模工作。

从航空运行数据建模的现有研究脉络来看，相关工作主要分为3类：第一类聚焦航空结构载荷预测，例如，陈农田等[7]基于QAR数据对高高原机场进近着陆风险开展建模分析；Yan Q等[8]采用主成分分析与偏最小二乘回归方法实现飞行载荷定量计算。第二类围绕飞机系统状态与运行参数预测展开，例如，Sun C X等[9]构建多深度学习融合架构实现飞机载荷预测；杨军利等[10]基于长短期记忆网络(Long Short-Term Memory，LSTM)—极端梯度提升(Extreme Gradient Boosting，XGBoost)开展民用飞机燃油流量预测；Ahmed[11]利用类似混合模型实现涡扇发动机剩余寿命预测。第三类关注飞行风险与飞行员脑力负荷评估，例如，卫宗敏[12]构建多维脑力负荷评估体系，对人为负荷及典型飞行风险进行量化分析。总体来看，上述研究成果为航空运行数据建模打下了重要基础，但面向高原航线运行负荷的系统性评估体系仍较为欠缺。

具体而言，现有研究主要存在3方面不足：1) 在指标体系方面，尚未形成面向高原航线运行特性的统一评价框架，缺乏对高原气象扰动、复杂地形及飞机性能变化的系统融合；2) 在建模方法方面，多数研究侧重单一时序特征或静态特征建模，缺乏静动态特征协同表达机制；3) 在验证层面，对典型复杂工况的鲁棒性分析仍不充分。

针对上述问题，本文基于民航高原航线QAR实测数据，提出一种面向高原航线运行负荷评估的LSTM-XGBoost融合评估方法。首先，结合运行机理与规章要求构建高原航线运行负荷指标体系，并利用灰色关联度分析和熵权法筛选关键指标；其次，采用LSTM提取飞行过程中的时序动态信息，并通过XGBoost完成静动态特征融合评估；最后，借助对比实验、消融实验和典型复杂工况测试，验证模型在所选样本中的评估精度和复杂工况稳定性。

1 高原航线运行负荷理论基础与指标体系构建

1.1 高原航线运行负荷定义与构成

运行负荷是指飞机在航线的全运行周期内，由多种因素共同作用形成的，作用于飞机结构、动力系统和飞行机组的综合动态负荷状态[1]。结合高原航线的运行特性，本文将运行负荷划分为结构负荷与操作负荷两类。

结构负荷主要由气动力变化、湍流扰动及发动机性能调节等因素引起，其核心表征参数为机翼负荷系数(n)。由航空动力学理论，负荷系数n定义为机翼升力L与飞机重力G之比：

$$n=\frac{L}{G}=\frac{L}{mg} \tag{1}$$

操作负荷主要反映机组在复杂工况下的操控压力，常通过正常加速度n_z(沿机体z轴，单位为g)进行表征。在稳态飞行条件下，n_z可视为n的直接

测量结果;在非稳态工况下,n_z能够更敏感地反映瞬态负荷变化特征。因此,本文同时采用n与n_z分别表征飞机结构负荷与机组操作负荷。

1.2　高原航线运行负荷影响因素分析

高原航线运行负荷主要受三重因素共同影响。1)飞机性能方面,高海拔空气密度降低导致发动机推力衰减,为维持升力需适度提高飞行速度,进而增加机翼负荷与燃料消耗。发动机排气温度(EGT)、振动值等参数直接反映发动机运行负荷。2)气象环境方面,高原地区的山地湍流、相对湿度和温度变化、侧风等会导致飞机姿态波动、轻度结冰、操作难度增加,是高原航线运行负荷升高的重要外部因素。3)飞行操作方面,高原机场周边地形复杂,进近和着陆阶段航迹调整空间有限,610~305 m进近阶段下降率过大易引发重着陆风险[13]。航向变化量、接地垂直加速度和操作频次能够反映机组操纵压力及运行风险水平。

1.3　高原航线运行负荷指标体系构建

1.3.1　指标选取原则

遵循针对性、客观性、可量化原则:针对性聚焦高原航线核心特性,剔除高高原相关指标;客观性基于QAR实测数据与CCAR规章,避免主观评价[14];可量化确保指标能通过机载设备直接采集或计算获得。

1.3.2　初始指标体系

为构建适配高原航线运行负荷评估的指标框架,本文综合行业规范、已有研究与飞行品质监控(FOQA)运行监测要求开展指标筛选。中国民用航空局相关运行规范对高原飞行场景下发动机性能监控及运行裕度控制提出了明确要求[15],已有研究亦证实大气环境扰动会显著改变飞机全阶段受载状态[16]。在此基础上,结合FOQA体系对飞行操作行为的全过程监测与量化分析要求,从飞机性能、气象环境与飞行操作3个维度构建初始指标体系,筛选得到22项初始指标,如表1所示,表中1 ft=0.304 8 m。

表1　高原航线运行负荷初始指标体系

Table 1　Initial index system of operational load for plateau routes

一级指标	二级指标	单位
飞机性能指标	发动机排气温度(EGT)	℃
	滑油温度	℃
	发动机振动值	mm/s
	真空速度	km/h
	机翼负荷系数	无
	正常加速度	g
	燃油流量	kg/h
	飞机载重系数	无
气象环境指标	海拔高度	m
	环境温度	℃
	风速	m/s
	风向角	(°)
	能见度	km
	湍流强度	级
	相对湿度	%
飞行操作指标	航向变化量	(°)
	610~305 m进近下降率	ft/min
	153~15 m进近下降率	ft/min
	接地垂直加速度	g
	航迹偏差	(°)
	操作频次	次/分钟
	复飞次数	次

1.3.3　关键指标筛选

采用灰色关联度分析与熵权法组合筛选[17],保留关联度≥0.7且权重≥0.03的15项关键指标,确保指标体系适配高原航线特性。

1)灰色关联度分析

设运行负荷的目标参考序列为x_0,各评价指标的比较序列为x_i,对各序列进行无量纲化处理后,计算参考序列与比较序列在第k个样本处的绝对差值:$\Delta_i(k)=|x_0(k)-x_i(k)|$。

灰色关联度系数定义为

$$\xi_i(k)=\frac{\min\limits_{i,k}\Delta_i(k)+\rho\max\limits_{i,k}\Delta_i(k)}{\Delta_i(k)+\rho\max\limits_{i,k}\Delta_i(k)} \tag{2}$$

式中:ρ为分辨系数,取值范围为$0<\rho<1$,本文取$\rho=0.5$。

第i个指标的灰色关联度为

$$r_i=\frac{1}{n}\sum_{k=1}^{n}\xi_i(k) \tag{3}$$

2）熵权法

为进一步评估指标的信息量与区分能力，引入熵权法计算各指标权重[18-19]。

对已通过灰色关联度初筛的指标进行归一化处理，得到第i个指标在第k个样本处的标准化值：

$$p_{ik}=\frac{x_{ik}}{\sum_{k=1}^{n}x_{ik}} \tag{4}$$

第i个指标的信息熵为

$$e_i=-\frac{1}{\ln n}\sum_{k=1}^{n}p_{ik}\ln p_{ik} \tag{5}$$

其对应的熵权为

$$w_i=\frac{1-e_i}{\sum_{i=1}^{m}(1-e_i)} \tag{6}$$

式中：m为通过灰色关联度筛选后的指标数量。

筛选结果如表2所示，最终保留15项关键指标，涵盖上述3个维度，关联度均≥0.72，权重均≥0.032，可全面表征高原运行负荷。

表2 高原航线运行负荷关键指标筛选结果

Table 2 Screening results of key indicators for operational loads on high-altitude routes

一级指标	二级指标	灰色关联度	熵权
飞机性能指标	EGT	0.89	0.087
	发动机振动值	0.82	0.063
	真空速度	0.78	0.051
	机翼负荷系数	0.92	0.095
	正常加速度	0.91	0.092
气象环境指标	海拔高度	0.85	0.072
	风速	0.76	0.048
	风向角	0.73	0.039
	湍流强度	0.88	0.081
	相对湿度	0.72	0.032
飞行操作指标	航向变化量	0.86	0.075
	610～305 m进近下降率	0.87	0.078
	153～15 m进近下降率	0.84	0.069
	接地垂直加速度	0.90	0.089
	操作频次	0.77	0.045

2 基于LSTM-XGBoost的高原航线运行负荷评估模型设计

2.1 模型整体架构

LSTM-XGBoost组合模型采用两阶段架构：第一阶段通过LSTM网络提取运行负荷的时序特征，处理随时间动态变化的高原航线核心指标；第二阶段将LSTM提取的时序特征与静态特征融合，输入XGBoost模型实现负荷等级与量化值的评估，全程适配高原航线特性。

2.2 LSTM网络时序特征提取模块

2.2.1 LSTM网络结构

LSTM适合提取高原航线运行负荷的时序依赖关系[20]。其中，发动机排气温度、发动机振动值、真空速度、机翼负荷系数、正常加速度、湍流强度、航向变化量、610～305 m下降率、153～15 m下降率及接地垂直加速度共10项动态特征作为LSTM输入变量，用于时序特征提取。提取得到的时序特征向量与海拔高度、风速、风向角、相对湿度及操作频次5项非时序特征在融合阶段共同输入XGBoost完成运行负荷评估。具体如下：

1）门控机制：

$$f_t=\sigma(W_f\cdot[h_{t-1},x_t]+b_f) \tag{7}$$

$$i_t=\sigma(W_i\cdot[h_{t-1},x_t]+b_i) \tag{8}$$

$$\tilde{C}_t=\tanh(W_C\cdot[h_{t-1},x_t]+b_C) \tag{9}$$

$$o_t=\sigma(W_o\cdot[h_{t-1},x_t]+b_o) \tag{10}$$

$$h_t=o_t\odot\tanh(C_t) \tag{11}$$

2）细胞状态更新：

$$C_t=f_t\odot C_{t-1}+i_t\odot\tilde{C}_t \tag{12}$$

式中：$\odot$为哈达玛积。

2.2.2 时序特征处理

采用滑动窗口法构建时序样本，并基于TensorFlow框架搭建LSTM模型。模型参数通过网格搜索优化，参数设置适配高原航线QAR数据特性，最终参数设置如表3所示。

表3 LSTM模型参数设置

Table 3 LSTM model parameter settings

参数	数值	参数	数值
时间步长	50	优化器	Adam
输入维度	10	学习率	0.001
隐藏层数	2	批处理大小	32
隐藏层神经元数	64	损失函数	MSE
输出特征维度	32	迭代次数	100
Dropout率	0.2		

2.3 XGBoost特征融合评估模块

2.3.1 XGBoost算法原理

XGBoost是一种基于梯度提升树（Gradient

Boosting Decision Tree, GBDT)的集成学习算法[21],目标函数包括损失函数与正则化项:

$$\mathrm{Obj}(\theta)=\sum_{i=1}^{n} l(y_i,\hat{y}_i^{(t)})+\sum_{k=1}^{t}\Omega(f_k) \tag{13}$$

目标函数经二次近似展开后,叶节点最优权重为

$$w_j^*=-\frac{\sum_{i\in S_j} g_i}{\sum_{i\in S_j} h_i+\lambda} \tag{14}$$

其中,变量定义参见文献[21]。

2.3.2　XGBoost模型参数设置

通过网格搜索优化XGBoost参数,结合高原航线数据特性调整参数取值,基于Python的XGBoost库构建模型,同样采用网格搜索法优化参数。最优参数设置如表4所示。

表4　XGBoost模型参数设置

Table 4　XGBoost model parameter settings

参数	数值
学习率	0.01
最大树深度	6
叶子节点最小权重	1
树的数量	200
列抽样比例	0.8
正则化系数λ	0.1
正则化系数γ	0.01

2.4　模型训练流程

1)数据预处理:对高原航线QAR数据进行去噪、归一化、缺失值填充,剔除异常数据。

2)数据划分:按7∶2∶1比例划分为训练集、验证集、测试集,数据均来自典型高原航线;为避免样本泄露带来的评估偏差,本文在数据划分过程中以"航班"为基本单元进行样本划分,确保同一航班的QAR数据不会同时出现在训练集与测试集中;同时,在模型训练过程中引入早停机制与L2正则化策略,防止模型过拟合。

3)LSTM训练:用训练集训练LSTM网络,提取高原航线运行负荷时序特征向量。

4)特征融合:将时序特征向量与静态特征拼接,形成融合特征集。

5)XGBoost训练:用融合特征训练XGBoost模型,通过验证集优化参数。

6)模型评估:采用测试集评价模型性能。模型架构如图1所示。

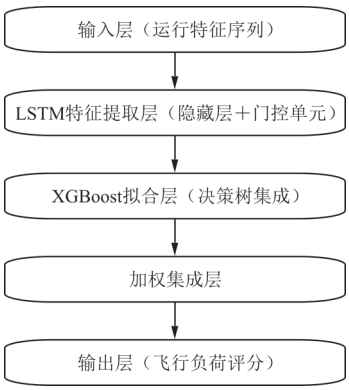

图1　模型架构

Fig. 1　Model architecture

3　实验验证与结果分析

3.1　实验数据来源与预处理

3.1.1　数据来源

实验数据来源于国内某航司A320机型执飞的6条高原航线QAR数据及配套气象数据。各航线目的地机场海拔为1 559～2 242 m,满足本文高原航线样本筛选范围,航线信息如表5所示。原始QAR数据采样频率为16 Hz,经航段截取、时间同步和滑动窗口切分后形成9 200个样本窗;剔除缺失率过高及明显异常样本后,保留8 640个有效样本窗。数据来源于某航司FOQA项目[22],已完成脱敏处理,仅用于学术研究。

表5　实验高原航线信息

Table 5　Information of the experimental high-altitude flight route

航线编号	起降机场	目的地机场海拔/m	样本数量/组
H1	成都双流—丽江三义	2 242	1 820
H2	昆明长水—保山云瑞	1 665	1 560
H3	西安咸阳—西宁曹家堡	2 184	1 480
H4	重庆江北—昆明长水	2 103	1 350
H5	兰州中川—嘉峪关酒泉	1 559	1 230
H6	贵阳龙洞堡—大理凤仪	2 155	1 200

3.1.2　数据预处理

1)去噪处理:由于QAR数据与气象数据在采集过程中易受干扰而混入一定程度的高频噪声,选取Daubechies4(db4)小波,对原始数据序列进行3层离散小波分解[23-24]。

针对高频分量,采用软阈值函数进行去噪处

理[23]，函数表达式为

$$w'_{j,k}=\begin{cases}\text{sign}(w_{j,k})(|w_{j,k}|-\lambda) & (|w_{j,k}|\geqslant\lambda)\\ 0 & (|w_{j,k}|<\lambda)\end{cases} \tag{15}$$

式中：$w_{j,k}$、$w'_{j,k}$分别为阈值处理前后的小波系数；λ为阈值参数。

阈值采用基于噪声标准差估计的通用阈值法确定，其计算公式为

$$\lambda=\sigma\sqrt{2\ln n} \tag{16}$$

$$\sigma=\frac{\text{median}(|cD_i|)}{0.674\,5} \tag{17}$$

式中：n为样本长度；σ为噪声标准差；median($|cD_i|$)为第i层高频系数绝对值中位数。

高频系数阈值处理后，与低频系数$cA3$重构得到去噪序列。

2）缺失值填充：采用线性插值补填缺失值，缺失率超5%的样本直接剔除。

3）归一化处理：为消除不同指标量纲差异对模型训练的影响，本文对所有连续型指标进行最小—最大归一化处理（Min-Max Normalization）[25]，将数据映射至[0,1]区间。归一化参数仅由训练集计算，并用于验证集和测试集，以避免测试信息泄露。其计算公式为

$$x'_i=\frac{x_i-x_{i,\min}}{x_{i,\max}-x_{i,\min}} \tag{18}$$

式中：x_i、x'_i分别为归一化前后的指标值；$x_{i,\min}$与$x_{i,\max}$分别为第i个指标在训练集中的最小值与最大值。

3.1.3 运行负荷等级划分

本文基于FOQA规则和熵权法构建代理负荷标签，用于表征运行参数组合下的相对运行负荷水平，而非实际安全事件风险。参考FOQA相关咨询通告及行业标准，结合高原航线运行特点，选取机翼载荷系数、正常加速度等关键参数进行综合评价，并依据评分结果将运行负荷划分为4个等级，如表6所示。综合评分范围为0～1，分值越高表示运行负荷等级越高。

表6 高原航线运行负荷等级划分标准

Table 6 Classification criteria for operational load levels on high-altitude routes

负荷等级	评分范围	核心指标阈值	工况描述
0级（低负荷）	0～0.25	$n\leqslant1.2$，$n_z\leqslant1.1g$，湍流强度≤1级	高原平稳飞行，操作难度低，无明显气流干扰
1级（中低负荷）	0.25～0.50	$1.2<n\leqslant1.5$，$1.1g<n_z\leqslant1.3g$，湍流强度=2级	高原轻度扰动，常规操作可应对，轻度颠簸
2级（中高负荷）	0.50～0.75	$1.5<n\leqslant1.8$，$1.3g<n_z\leqslant1.5g$，湍流强度=3级	高原中度湍流，需精准操作，航迹微调频繁
3级（高负荷）	0.75～1.00	$n>1.8$，$n_z>1.5g$，湍流强度≥4级	高原典型复杂工况，中强湍流/侧风超限，操作风险高

3.2 实验环境与对比模型

3.2.1 实验环境

实验在Python3.10环境下完成，主要使用TensorFlow2.10、XGBoost1.7.5和Scikit-learn1.2.2。

3.2.2 对比模型

为验证所提模型在高原航线的有效性，选取LSTM、XGBoost、SVM、粒子群优化反向传播神经网络（Particle Swarm Optimization-Back Propagation Neural Network，PSO-BP）、LSTM-DNN 5种常用模型作为对比，所有模型采用相同数据集和评价指标进行测试，以保证结果的可对比性。

3.3 评价指标

为全面评估模型性能，首先构建运行负荷连续评分模型（回归任务），再依据表6中的评分区间将预测结果映射为对应负荷等级（分类任务）。因此，连续负荷预测采用均方根误差（RMSE）、平均绝对误差（MAE）及决定系数（R^2）进行评估，其计算公式如下：

$$e_{\text{RMSE}}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \tag{19}$$

$$e_{\text{MAE}}=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i| \tag{20}$$

$$R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \tag{21}$$

式中：y_i为真实负荷值；$\hat{y}_i$为模型预测值；$\bar{y}$为真实值均值；n为样本数量。

在负荷等级判别层面，采用准确率（Accuracy）来评价模型对不同运行负荷等级的识别能力，其计算公式为

$$a_{ACC}=\frac{N_{correct}}{N_{total}} \tag{22}$$

式中：$N_{correct}$为负荷等级预测正确的样本数量；N_{total}为样本总数。

为量化模型对不同负荷等级的分类能力，本文采用精确率(Precision)、召回率(Recall)、F1值及宏F1值(Macro-F1)作为指标，定义如下：

$$p_{PRE}=\frac{T_{P_i}}{T_{P_i}+F_{P_i}} \tag{23}$$

$$r_{Rec,i}=\frac{T_{P_i}}{T_{P_i}+F_{N_i}} \tag{24}$$

$$f_{F1,i}=\frac{2\times p_{PRE,i}\times r_{Rec,i}}{p_{PRE,i}+r_{Rec,i}} \tag{25}$$

$$f_{MF1}=\frac{1}{C}\sum_{i=1}^{C}f_{F1,i} \tag{26}$$

式中：T_{P_i}、F_{P_i}、F_{N_i}分别为第i类真正例、假正例和假反例数；C为负荷等级类别数，$C=4$。Macro-F1对各等级赋予相同权重，反映整体识别能力。

准确率、F1值越高、e_{RMSE}与e_{MAE}越小、R^2越接近1，模型在高原航线的适配性能越优，同时也表明模型对高原航线运行负荷等级的综合判别能力越强。

3.4　实验结果与分析

3.4.1　样本分布

以“航班”为基本单元，按7∶2∶1比例划分训练集、验证集、测试集，且各子集内负荷等级分布与原始数据集一致，具体样本分布如表7所示，可以看出：训练集、验证集、测试集共计8 640个总样本，其中0级2 160个、1级2 585个、2级2 426个、3级1 469个，各等级占比分别为25.0%、29.9%、28.1%、17.0%。

表7　高原航线运行负荷样本分布表

Table 7　Sample distribution of operational load on high-altitude routes

数据集	总样本数	0级(低负荷)	1级(中低负荷)	2级(中高负荷)	3级(高负荷)	各等级占比/%
训练集	6 048	1 512	1 814	1 703	1 019	25.0/29.9/28.2/16.9
验证集	1 728	432	514	481	301	25.0/29.7/27.8/17.4
测试集	864	216	257	242	149	25.0/29.7/28.0/17.2

3.4.2　模型整体性能对比

各模型在高原航线测试集上的性能指标统计结果如表8所示，准确率、e_{RMSE}、e_{MAE}与R^2对比结果分别如图2～图5所示。

表8　各模型在高原航线的性能对比结果

Table 8　Performance comparison results of each model on high-altitude routes

模型	准确率/%	e_{RMSE}	e_{MAE}	R^2
LSTM-XGBoost	96.30	0.042	0.031	0.958
LSTM	90.21	0.087	0.069	0.872
XGBoost	88.47	0.095	0.076	0.851
SVM	85.14	0.112	0.089	0.813
PSO-BP	88.23	0.092	0.073	0.864
LSTM-DNN	94.18	0.058	0.045	0.923

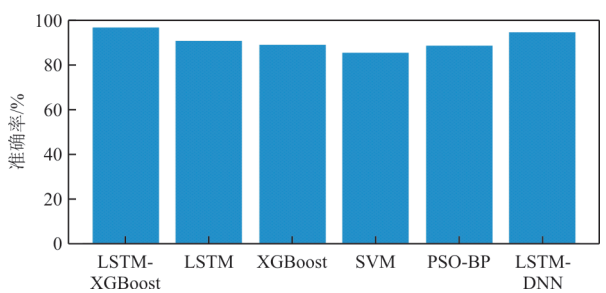

图2　不同模型准确率对比

Fig. 2　Accuracy comparison of different models

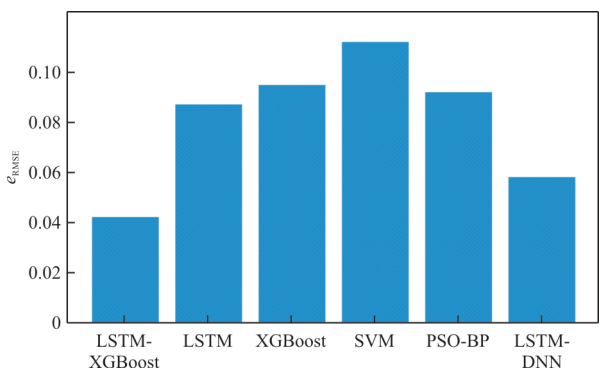

图3　不同模型均方根误差对比

Fig. 3　RMSE comparison of different models

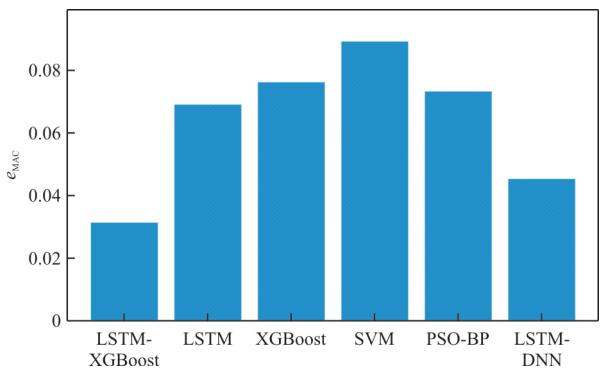

图4　不同模型平均绝对误差对比

Fig. 4　MAE comparison of different models

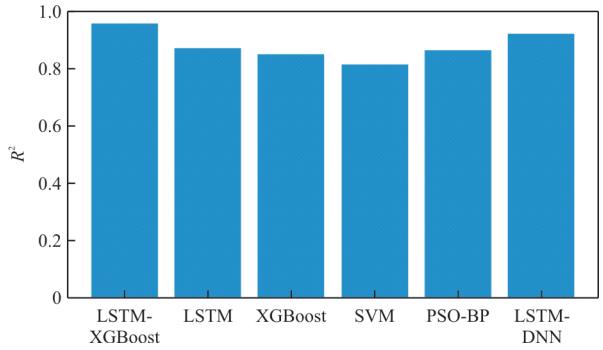

图 5 不同模型决定系数对比

Fig. 5 R^2 comparison of different models

从表 8 和图 2～图 5 可以看出：LSTM-XGBoost 模型在各评价指标上均优于对比模型。与 LSTM、XGBoost、SVM、PSO-BP 和 LSTM-DNN 相比，其准确率分别提高 6.09、7.83、11.16、8.07 和 2.12 个百分点，说明时序特征提取与静动态特征融合能够有效提升高原航线运行负荷评估精度。

3.4.3 不同运行负荷等级分类性能分析

采用混淆矩阵量化各等级分类细节，结果如图 6 和表 9、图 7 和表 10 所示。

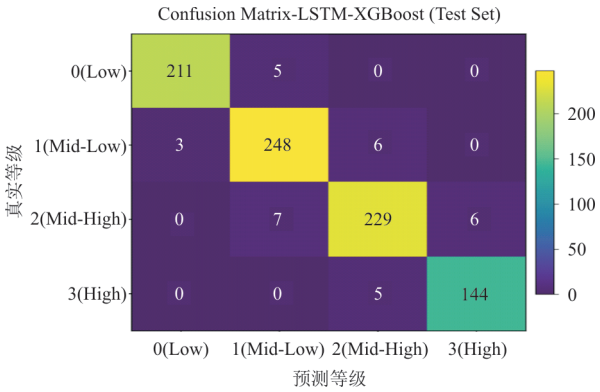

图 6 LSTM-XGBoost 混合模型测试集的混淆矩阵

Fig. 6 Confusion matrix of the LSTM-XGBoost hybrid model on the test set

表 9 LSTM-XGBoost 模型混淆矩阵(测试集)

Table 9 Confusion matrix of LSTM-XGBoost model (test set)

真实等级/预测等级	0级(低负荷)	1级(中低负荷)	2级(中高负荷)	3级(高负荷)	召回率/%	精确率/%
0级(低负荷)	211	5	0	0	97.69	98.60
1级(中低负荷)	3	248	6	0	96.50	95.38
2级(中高负荷)	0	7	229	6	94.63	95.42
3级(高负荷)	0	0	5	144	96.64	96.00

从表 9 可以看出：LSTM-XGBoost 对 0 级和 3 级样本识别效果较好；各等级误判均集中在相邻等级之间，未出现跨两级及以上误判；其中，2 级样本召回率最低，为 94.63%，说明其处于常规运行与高负荷状态之间的过渡区间，是主要边界混淆场景。总体上，各等级精确率和召回率分别不低于 95.38% 和 94.63%，表明模型具有较好的分等级识别能力。

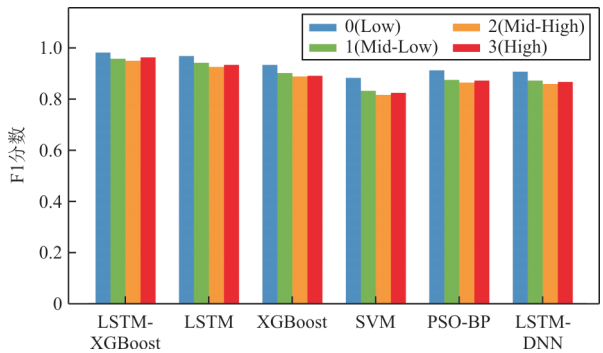

图 7 不同模型在测试集上的分负荷等级 F1 分数对比

Fig. 7 Comparison of per-load-level F1 scores for different models on the test set

表 10 各模型分等级 F1 值对比

Table 10 F1-score comparison of each model by load level

模型	0级(低负荷)	1级(中低负荷)	2级(中高负荷)	3级(高负荷)	平均F1值
LSTM-XGBoost	0.981	0.959	0.950	0.963	0.963
LSTM-DNN	0.968	0.942	0.927	0.935	0.943
LSTM	0.935	0.901	0.887	0.892	0.904
PSO-BP	0.912	0.876	0.864	0.871	0.876
XGBoost	0.908	0.873	0.859	0.868	0.877
SVM	0.884	0.832	0.815	0.823	0.838

从表 10 可以看出：LSTM-XGBoost 平均 F1 值为 0.963，高于其他对比模型；相较 LSTM-DNN，平均 F1 值提高 0.020，在 2 级样本上的 F1 值提高 0.023。说明融合模型不仅提升了整体准确率，也改善了边界负荷等级的识别稳定性。

3.4.4 典型复杂工况下的稳定性测试

典型复杂工况样本筛选自测试集，按主导工况归类后，包括中强湍流 92 组、侧风超限 78 组、轻度结冰 65 组和重着陆风险 51 组，共 286 组，占测试

集的33.1%,各模型测试结果如表11所示。

表11　各模型在典型复杂工况下的性能表现
Table 11　Performance of each model under typical complex conditions

模型	准确率/%	e_{RMSE}	e_{MAE}	R^2	性能准确率下降幅度(百分点)
LSTM-XGBoost	92.31	0.057	0.043	0.932	3.99
LSTM-DNN	88.81	0.079	0.061	0.895	5.37
LSTM	83.92	0.108	0.085	0.821	6.29
PSO-BP	81.47	0.115	0.092	0.803	6.76
XGBoost	80.77	0.121	0.098	0.791	7.70
SVM	77.62	0.143	0.116	0.745	7.52

从表11可以看出:LSTM-XGBoost模型在典型复杂工况下准确率为92.31%,较整体测试集下降3.99个百分点,且e_{RMSE}、e_{MAE}和R^2整体优于对比模型。相比之下,XGBoost模型准确率下降7.70个百分点,SVM和PSO-BP在复杂工况下准确率分别为77.62%和81.47%,表明单一模型对高扰动样本的适应性相对不足。

3.4.5　关键指标贡献分析

为分析各关键指标对运行负荷评估结果的影响,对训练后模型的指标贡献得分进行归一化排序,各指标贡献得分总和约为1,偏差主要由三位小数保留造成。关键指标贡献排序如表12所示。

表12　关键指标贡献排序
Table 12　Ranking of key indicator contributions

一级指标	二级指标	贡献得分	排名	指标贡献说明
飞机性能指标	机翼负荷系数	0.126	1	结构负荷主要表征参数
飞机性能指标	正常加速度	0.117	2	操作负荷直接表征参数
飞行操作指标	接地垂直加速度	0.103	3	着陆阶段负荷关键参数
气象环境指标	湍流强度	0.096	4	高原航线气流扰动核心影响因素
飞机性能指标	发动机排气温度(EGT)	0.085	5	发动机运行状态表征参数
飞行操作指标	610~305 m进近下降率	0.077	6	进近阶段操作精度核心指标
气象环境指标	海拔高度	0.071	7	影响空气密度与发动机性能的基础参数
飞行操作指标	航向变化量	0.064	8	航迹控制难度表征,适配高原狭窄航路
气象环境指标	风速	0.047	9	侧风影响及着陆扰动相关参数
飞行操作指标	153~15 m进近下降率	0.041	10	着陆前关键阶段操作负荷指标
飞机性能指标	发动机振动值	0.038	11	发动机稳定性表征,关联机械负荷
飞机性能指标	真空速度	0.036	12	高原升力补偿核心参数
飞行操作指标	操作频次	0.034	13	反映机组操控强度
气象环境指标	风向角	0.033	14	侧风影响辅助表征
气象环境指标	相对湿度	0.031	15	轻度结冰风险辅助指标

从表12可以看出:机翼负荷系数、正常加速度、接地垂直加速度、湍流强度和EGT位列前五,累计贡献得分为0.527,说明结构受载、操纵响应、着陆冲击、气流扰动和发动机运行状态是高原航线运行负荷的主要影响因素。海拔高度、进近下降率和航向变化量等指标也具有一定贡献,表明运行负荷受飞机性能、气象环境和飞行操作共同影响。表12为模型评估过程中的指标贡献排序,不等同于表2中的熵权筛选权重。

3.4.6　消融实验

为分析LSTM时序特征提取模块和XGBoost特征融合模块对模型性能的影响,在保持数据划分、预处理方法和主要训练参数一致的条件下进行消融实验,结果如表13所示。

表13　消融实验结果
Table 13　Ablation experiment results

模型配置	准确率/%	e_{RMSE}	e_{MAE}	R^2
LSTM+XGBoost	96.30	0.042	0.031	0.958
仅XGBoost	88.47	0.095	0.076	0.851
仅LSTM	90.21	0.087	0.069	0.872

从表13可以看出:移除LSTM模块后,模型准确率较完整模型下降7.83个百分点,e_{RMSE}升至0.095,说明时序特征提取对运行负荷评估具有重要作用;移除XGBoost模块后,模型准确率较完整模型下降6.09个百分点,e_{MAE}升至0.069,说明特

征融合模块有助于提升模型对复杂特征的处理能力。

综上，LSTM时序特征提取与XGBoost特征融合对模型性能提升均具有贡献，组合结构能够提高高原航线运行负荷评估的准确性和稳定性。

4 结 论

1）结合飞机性能、气象环境和飞行操作特征，构建了高原航线运行负荷指标体系，并通过灰色关联度分析与熵权法筛选得到15项关键指标。其中，机翼负荷系数、正常加速度、接地垂直加速度、湍流强度和发动机排气温度对运行负荷评估贡献较大，可作为高原航线运行负荷的重要表征参数。

2）所提出的LSTM-XGBoost模型在测试集上的准确率达到96.30%，e_{RMSE} 和 e_{MAE} 分别为0.042和0.031，整体性能优于LSTM、XGBoost、SVM、PSO-BP及LSTM-DNN等对比模型，表明时序特征提取与静动态特征融合能够有效提升高原航线运行负荷评估精度。

3）在中强湍流、侧风超限、低温高湿及重着陆风险等典型复杂工况下，模型准确率仍保持在92.31%，误判主要集中于相邻负荷等级，表明该模型具有较好的复杂工况适应能力和稳定性。

4）本文研究基于单一机型高原航线QAR数据开展验证，后续可进一步扩展至多机型、多区域运行数据，并结合实际运行风险事件开展研究，以提升模型的泛化能力和工程应用价值。

参考文献

[1] 中国民用航空局飞行标准司. 高原机场运行：AC-121-FS-2015-21R1[S]. 北京：中国民用航空局，2015.
Flight Standards Department, Civil Aviation Administration of China. High-altitude airport operations: AC-121-FS-2015-21R1[S]. Beijing: Civil Aviation Administration of China, 2015. (in Chinese)

[2] 余江. 高原/复杂地形机场和航线运行的飞机性能分析[M]. 2版. 成都：西南交通大学出版社，2015.
Yu Jiang. Aircraft performance analysis for operations in high-altitude/complex terrain airports and routes[M]. 2nd ed. Chengdu: Southwest Jiaotong University Press, 2015. (in Chinese)

[3] 张林意，孙华波，王纯，等. 基于QAR飞行大数据的空中颠簸风险时空分布模式探索与分析[J]. 武汉大学学报(信息科学版)，2024，49(3)：482-490.
Zhang Linyi, Sun Huabo, Wang Chun, et al. Spatiotemporal pattern of air turbulence risks with QAR flight big data[J]. Geomatics and Information Science of Wuhan University, 2024, 49(3): 482-490. (in Chinese)

[4] Lee H, Madar S, Sairam S, et al. Critical parameter identification for safety events in commercial aviation using machine learning[J]. Aerospace, 2020, 7(6): 73.

[5] Li X Y, Zhang L, Shang J X, et al. A runway overrun risk assessment model for civil aircraft based on quick access recorder data[J]. Applied Sciences, 2023, 13(17): 9828.

[6] Rey M, Aloise D, Soumis F, et al. A data-driven model for safety risk identification from flight data analysis[J]. Transportation Engineering, 2021, 5: 100087.

[7] 陈农田，满永政，李俊辉. 基于QAR数据的民机高高原进近着陆风险评估方法[J]. 北京航空航天大学学报，2024，50(1)：77-85.
Chen Nongtian, Man Yongzheng, Li Junhui. Risk assessment method for civil aircraft approach and landing at high plateau based on QAR data[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(1): 77-85. (in Chinese)

[8] Yan Q, Yang C, Wan Z Q. A comparative regression analysis between principal component and partial least squares methods for flight load calculation[J]. Applied Sciences, 2023, 13(14): 8428.

[9] Sun C X, Li H Y, Dui H N, et al. A multi-model architecture based on deep learning for aircraft load prediction[J]. Communications Engineering, 2023, 2: 47.

[10] 杨军利，李立坤，钱宇，等. 基于LSTM-XGBoost的国产民用飞机燃油流量预测[J]. 航空工程进展，2025，16(5)：207-214.
Yang Junli, Li Likun, Qian Yu, et al. Fuel flow prediction of domestic civil aircraft based on LSTM-XGBoost[J]. Advances in Aeronautical Science and Engineering, 2025, 16(5): 207-214. (in Chinese)

[11] Ahmed A. Hybrid and deep learning architectures for predictive maintenance: evaluating LSTM and attention-based LSTM-XGBoost on turbofan engine RUL[C]// MATEC Web of Conferences. Les Ulis: EDP Sciences, 2025: 07008.

[12] 卫宗敏. 面向复杂飞行任务的脑力负荷多维综合评估模型[J]. 北京航空航天大学学报，2020(7)：1287-1295.
Wei Zongmin. A multi-dimensional comprehensive evaluation model of mental workload for complex flight missions[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020(7): 1287-1295. (in Chinese)

[13] 马林波. 从一起重落地事件试析特殊机场运行风险[J]. 民航新型智库研究，2021(11)：5276.
Ma Linbo. Analysis of special airport operating risks based on a heavy aircraft landing incident[J]. Civil Aviation New Think Tank Research, 2021(11): 5276. (in Chinese)

[14] 中华人民共和国交通运输部. 大型飞机公共航空运输承运

人运行合格审定规则[S]. 北京:中华人民共和国交通运输部, 2024.
Ministry of Transport of the People′s Republic of China. Rules for the operational approval of public air transport carriers of large aircraft[S]. Beijing: Ministry of Transport of the People′s Republic of China, 2024. (in Chinese)

[15] 中国民用航空局飞行标准司. 高高原机场运行特殊程序:OSB-2020-05[S]. 北京:中国民用航空局, 2020.
Flight Standards Department, Civil Aviation Administration of China. Special procedures for operations at high-altitude airports: OSB-2020-05[S]. Beijing: Civil Aviation Administration of China, 2020. (in Chinese)

[16] Wu Z, Cao Y, Ismail M. Gust loads on aircraft[J]. The Aeronautical Journal, 2019, 123: 1216-1274.

[17] 刘思峰. 灰色系统理论及其应用[M]. 5版. 北京:科学出版社, 2010.
Liu Sifeng. Grey system theory and its applications[M]. 5th ed. Beijing: Science Press, 2010. (in Chinese)

[18] 李航. 统计学习方法[M]. 2版. 北京:清华大学出版社, 2019.
Li Hang. Statistical learning methods[M]. 2nd ed. Beijing: Tsinghua University Press, 2019. (in Chinese)

[19] 王斌会. 多元统计分析及R语言建模[M]. 5版. 北京:高等教育出版社, 2020.
Wang Binhui. Multivariate statistical analysis and modeling with R language[M]. 5th ed. Beijing: Higher Education Press, 2020. (in Chinese)

[20] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.

[21] 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.

[22] 中国民用航空局飞行标准司. 飞行数据分析方案实施与管理要求:AC-121-FS-45R2[S]. 北京:中国民用航空局, 2024.
Flight Standards Department, Civil Aviation Administration of China. Requirements for the implementation and management of flight data analysis program: AC-121-FS-45R2[S]. Beijing: Civil Aviation Administration of China, 2024. (in Chinese)

[23] Ouyang C Y, Cai L M, Liu B, et al. An improved wavelet threshold denoising approach for surface electromyography signal[J]. EURASIP Journal on Advances in Signal Processing, 2023(1): 108-117.

[24] Donoho D L, Johnstone I M. Ideal spatial adaptation by wavelet shrinkage[J]. Biometrika, 1994, 81(3): 425-455.

[25] 周志华. 机器学习[M]. 北京:清华大学出版社, 2016.
Zhou Zhihua. Machine Learning[M]. Beijing: Tsinghua University Press, 2016. (in Chinese)

(编辑:马文静)

(上接第180页)

Deng Yongheng, Sun Hong, Zhang Peiwen. Recognition method of rectangular traffic pattern base on SD card data[J]. Journal of Civil Aviation Flight University of China, 2020, 31(2): 5-9. (in Chinese)

[16] 孙宏, 张兆阳, 兰舰, 等. 基于胜任力与CatBoost算法的商照训练评估方法[J]. 航空工程进展, 2025, 16(5): 183-189.
Sun Hong, Zhang Zhaoyang, Lan Jian, et al. Commercial pilot license stage flight training evaluation method based on competency and CatBoost algorithm[J]. Advances in Aeronautical Science and Engineering, 2025, 16(5): 183-189. (in Chinese)

[17] 刘伟, 刘玉妍, 杨泳. 人工和数据相结合的飞行技术量化评价方法研究[J]. 航空工程进展, 2025, 16(5): 171-182.
Liu Wei, Liu Yuyan, Yang Yong. Research on quantitative evaluation method of flight technology combining artificial and data[J]. Advances in Aeronautical Science and Engineering, 2025, 16(5): 171-182. (in Chinese)

[18] Guevarra M, Das S, Wayllace C, et al. Augmenting flight training with AI to efficiently train pilots[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(13): 16437-16439.

[19] Sun H, Zhou X, Zhang P W, et al. Competency-based assessment of pilots′ manual flight performance during instrument flight training[J]. Cognition, Technology & Work, 2023, 25(4): 345-356.

[20] Yang X, Ren J, Li J C, et al. Data-driven long-landing event detection and interpretability analysis in civil aviation[J]. IEEE Access, 2022, 10: 64257-64269.

[21] 何健, 钱宇, 唐盛香. 基于LSTM的民用飞机重着陆风险预测评估研究[J]. 航空工程进展, 2025, 16(5): 51-57.
He Jian, Qian Yu, Tang Shengxiang. Research on heavy landing risk prediction and evaluation of civil aircraft based on LSTM[J]. Advances in Aeronautical Science and Engineering, 2025, 16(5): 51-57. (in Chinese)

(编辑:丛艳娟)