

基于 QAR 数据的航空公司飞行安全韧性评价方法

何鹏^{1,2}, 孙瑞山³

(1. 新疆大学交通运输工程学院, 乌鲁木齐 830017; 2. 新疆大学新疆交通基础设施绿色建养与智慧交通管控重点实验室, 乌鲁木齐 830017; 3. 中国民航大学民航安全科学研究所, 天津 300300)

摘要: 为客观量化航空公司飞行安全韧性, 提出一种基于 QAR 数据的飞行安全韧性评价方法 (QAR-FSRAM)。首先, 通过理论分析界定飞行安全韧性的概念, 并描述其飞行绩效数据特征; 然后, 基于 QAR 数据, 采用短时间序列表达挖掘技术 (STEM) 构建 QAR-FSRAM; 最后, 以 A 航空公司在进近阶段的飞行数据作为研究对象进行实例评价。结果表明: 飞行安全韧性的绩效数据特征分为趋势和偏离两个维度; QAR-FSRAM 包含 4 个标准化步骤, 分别为构建航班韧性绩效表征变量、量化表征变量趋势特征、计算表征变量偏离特征, 和分析航空公司航班绩效结构; A 航空公司在进近阶段的飞行安全韧性不高, 韧性值仅为 0.51, 将韧性评价结果与飞行品质监控 (FOQA) 超限检测结果进行对比, 验证了 QAR-FSRAM 的有效性。

关键词: 航空公司; 飞行安全; 趋势特征; 偏离特征; 韧性评价; 快速存取记录器 (QAR)

中图分类号: U8

文献标志码: A

文章编号:

An airline flight safety resilience assessment method based on QAR data

HE Peng^{1,2} SUN Ruishan^{3*}

(1.School of Traffic and Transportation Engineering, Xinjiang University, Urumqi 830017, China; 2.Xinjiang Key Laboratory of Green Construction and Smart Traffic Control of Transportation Infrastructure, Xinjiang University, Urumqi 830017, China; 3.Research Institute of Civil Aviation Safety, Civil Aviation University of China, Tianjin 300300, China)

Abstract: In order to objectively quantify the flight safety resilience of airlines, a QAR data based flight safety resilience assessment method (QAR-FSRAM) is proposed. Firstly, the concept of flight safety resilience is defined through theoretical analysis, and its characteristic manifestations in flight performance data are described. Subsequently, based on QAR data, the QAR-FSRAM is constructed using the Short Time-series Expression Miner (STEM) technique. Finally, an empirical evaluation is conducted using the flight data of Airline A during the approach phase as a case study. The results show that the performance data characteristics reflecting flight safety resilience are characterized by two dimensions: trend and deviation. Four standardized steps are involved in the QAR-FSRAM: construction of flight resilience performance characterization variables, quantification of their trend characteristics, calculation of their deviation characteristics, and analysis of the airline's overall flight performance structure. It is revealed by the findings that relatively low flight safety resilience is exhibited by Airline A during the approach phase, with a calculated resilience index of only 0.51. The effectiveness of the proposed QAR-FSRAM method is validated through a comparison between the resilience assessment results and Flight Operations Quality Assurance (FOQA) exceedance detection.

Keywords: airlines; flight safety; trend characteristic; deviation characteristic; resilience assessment; Quick Access Recorder (QAR)

基金项目: 新疆维吾尔自治区人才发展基金“天池英才”引进计划青年博士项目; 中国民航局安全能力建设项目 (No. ASSA2023/19)。

收稿日期: XXXX-XX-XX; 修订日期: XXXX-XX-XX

通信作者: 何鹏, 男, 讲师, E-mail: hepeng@xju.edu.cn

0 引言

安全是决定航空业可持续发展的关键因素。安全韧性作为韧性概念的分支,是安全领域新概念^[1,2],也是进一步优化系统安全绩效的新方法。Hollnagel等^[3]认为韧性是系统在变化和干扰之前、期间或之后用来调整其功能的内在能力,以期在预期和非预期条件下维持所需性能。这种维持所需性能的能力,通常会在系统的实际运行绩效数据中有所体现。黄浪等^[4]将安全韧性定义为系统在一定时空内面对风险的冲击扰动时,维持、恢复和优化系统安全状态的能力。随着韧性理论的不断演进,国际民航业开始意识到,管理非预期风险不仅要依赖事故、征候、超限事件等负面样本,更应从“安全-II”视角出发,深入分析和挖掘全运行航班的安全规律和正向安全贡献,并以评价、提升系统的安全韧性^[5,6],实现更有效的航空事故预防。

近年来,国际航空业在安全韧性方面取得了显著进展。例如,ICAO^[7]提出循证训练(Evidence Based Training, EBT)的理念,并以此强化飞行机组的知识、态度、技能,使其在预期与非预期情形下具有韧性。美国航空(American Airlines, AA)^[8]在2020年提出了机组安全韧性的评价方法,通过观察和收集飞行运行中的机组韧性行为(Flight Crew Resilient Behavior, FRCB),对机组的安全韧性进行了描述性分析。欧洲航行安全组织^[9]则提出了系统韧性潜能管理的概念,并通过韧性分析矩阵(Resilience Analysis Grid, RAG)分析方法对空管机构的安全韧性进行了半定量评价。国内方面,王兴隆^[10]等通过构建贝叶斯网络模型,从宏观、中观、微观等3个层次自下而上对空中交通运输体系的韧性进行评估。郭九霞^[11]和Li^[12]等学者基于网络视角,分别量化了航空运输网络的韧性及脆弱性,以此分析新冠疫情、自然灾害等冲击对宏观航空运输系统的影响。

尽管安全韧性概念已在航空业得到广泛关注,但如何将其与航空公司日常运行产生的大规模数据相结合,以实现客观、量化的评价,是当前面临的关键挑战。航空公司积累的海量快速存取记录器(Quick Access Recorder, QAR)数据为此提供了宝贵的数据基础,QAR数据以高频度、高精度记录了飞行过程中的运行参数,蕴含了反映航班运行状态、动态稳定性及对环境扰动响应能力的丰富信息。然而,当前QAR数据的应用主流仍集中在飞行品质监控(Flight Operations Quality Assurance, FOQA)框架下,其主要目标是识别超限事件,这在本质上属于传统安全管理(Safety-I)的范畴,侧重于对失效和偏差的分析^[13,14]。近年来,数据挖掘

与机器学习技术的发展推动了部分研究利用QAR数据进行更深层次的探索^[15],例如通过聚类算法识别典型的飞行运行模式^[16],或评估飞行操纵品质^[17]。上述工作为从数据层面理解航班运行的复杂性与适应性提供了新的视角,其分析结果与韧性概念具有潜在相关性。但是,将此类数据分析方法与安全韧性理论(特别是强调系统适应与成功的Safety-II视角)进行系统性整合,构建能够量化评估航空公司飞行安全韧性水平的综合性评价方法,目前仍是研究的薄弱环节,相关的公开研究成果较为有限。

鉴于此,本文聚焦于飞行安全韧性在航班运行绩效层面的体现,拟提出一种基于QAR数据的航空公司飞行安全韧性评价方法。该方法的核心在于通过分析航班运行核心参数的时间序列模式(特别是趋势的稳定性和偏离程度),提取航班QAR特征参数构建韧性绩效表征变量,采用短时间序列表达挖掘方法(Short Time-series Expression Miner, STEM)量化航班簇趋势特征、偏离特征,实现韧性绩效模式划分,进而量化评价飞行安全韧性。本文特别选取了运行风险较高的进近阶段作为研究切入点,并以A航空公司在该阶段的运行数据为例展开案例评价,将其评价结果与该阶段的FOQA超限检测结果进行对比验证,以期为航空公司利用自有运行数据评估和提升飞行安全韧性提供理论方法支撑。

1 飞行安全韧性

1.1 飞行安全韧性概念界定

在航空领域,飞行机组被认为是系统安全韧性的主要来源^[18]。面对飞行运行中预期与非预期的危险,飞行机组为保障运行安全所作出的调整、适应等行为可以帮助飞行运行维持可接受的安全水平。AA^[8]于2020年提出FRCB的概念,并通过定义FRCB间接对机组飞行安全韧性和航空公司飞行安全韧性进行界定。FRCB被认为是“doing something above and beyond SOP”,既包含了飞行机组在飞行运行过程中高于规章标准的行为,也包含了在规章覆盖不到或者规章没有明确指出的情况下,飞行机组为保证运行安全所采取的行动,这些行为都反映了飞行机组在预期与非预期情形下对运行安全的正面贡献,使得飞行安全绩效尽可能地不低于基线绩效。若不考虑航空公司在航空器硬件、管理因素等方面的差异,航空公司飞行安全韧性可以认为是既定统计时段内航空公司所有飞行机组飞行安全韧性的总体水平。

综上所述,本研究结合航空运输特殊行业属性,借鉴AA对于安全韧性的概念界定,将“航空

公司飞行安全韧性”定义为：航空公司以机组飞行安全韧性为基础，面对飞行运行中的风险扰动而保持安全绩效不低于基线绩效的能力。这一定义强调了航空公司整体飞行安全韧性对一线运行人员（机组）韧性的依赖，并将关注点聚焦于系统在扰动下维持关键安全绩效（不低于特定基线）的实际表现。需要注意的是，本研究并不试图为这一定义本身构建单一的数学表达式来直接度量韧性，而是主张通过分析大规模 QAR 数据中能够反映系统飞行安全绩效稳定性与演变规律的关键统计特征，来定量地评估航空公司的飞行安全韧性水平。这些关键数据特性的具体定义、识别与量化方法，将在后文进行详细阐述。

1.2 飞行安全韧性的绩效数据特征

韧性概念并被广泛应用于不同学科领域^[19-21]，但本质上，其韧性概念都可以采用一个相似的图形化表示来阐述，即图 1 所示的韧性曲线，该曲线最初由 Bruneau 等^[22]提出，后来被称为“韧性三角”。韧性曲线以动态视角描绘了系统受到冲击扰动后的绩效特征： t_1 时刻，系统受到扰动，绩效下降； t_2 时刻，系统绩效达到最低值； t_2-t_3 时刻，系统绩效逐渐恢复并达到基线水平。

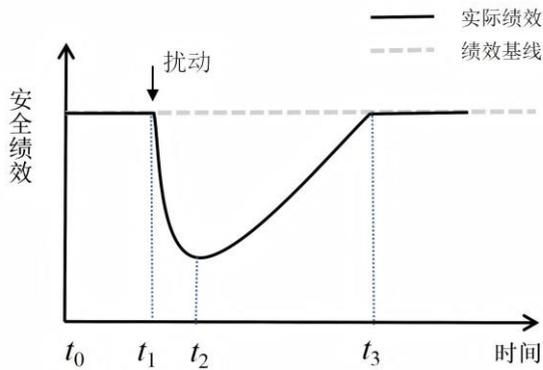


图 1 韧性曲线

Fig.1 Resilience curve

然而，图 1 中的韧性曲线是概念性的，在飞行安全韧性研究中，航班安全绩效由多个 QAR 飞行过程参数决定，因此韧性曲线具有更广泛的轨迹特征。尽管轨迹形态多样，但其动态过程的核心特征可以从两个关键维度进行概括和量化：一是趋势特征，二是偏离特征。

趋势特征描述了在特定运行情境下，大量航班样本在某一或某组关键飞行参数上表现出的整体性、方向性的统计趋势。它反映了航空公司在该情境下航班安全绩效的平均水平随时间（或其他条件）的演变方向或集中倾向。通过分析此特征，可以评估系统维持安全绩效的宏观表现。例如，若某参数的安全裕度随时间推移呈现持续增大或稳

定的趋势，表明系统在该方面较好地维持了安全绩效水平，体现了韧性；反之，持续下降的趋势则可能预示韧性不足，绩效有跌破基线的风险。

偏离特征衡量了在特定运行情境下，个体航班在某一或某组关键飞行参数上的实际表现值与其所属群体趋势特征之间的离散程度或波动幅度。它量化了个体运行表现围绕群体中心趋势的变异性与一致性，反映了系统在面对各类（包括未显性记录的）扰动时，保持运行稳定性和一致性的能力。较低的偏离特征意味着航班运行更为稳定、可预测，系统能有效吸收扰动，将绩效维持在预期范围内（有助于保持在基线之上）；较高的偏离特征则表明运行一致性差，系统对扰动的吸收能力可能不足或不稳定，增加了绩效偏离预期的风险。

通过对这两类特征的综合分析，可以更全面地评价航空公司在不同运行情境下的飞行安全韧性水平，即在扰动下保持安全绩效不低于基线的能力。

2 飞行安全韧性评估方法

本研究通过建立航班韧性绩效表征变量、量化韧性绩效表征标量趋势特征、偏离特征，分析航空公司飞行绩效结构，构建基于 QAR 数据的飞行安全韧性评价方法。

2.1 建立韧性绩效表征变量

航班韧性表征变量 $R = \{R_1, R_2, \dots, R_n\}$ 被定义为综合安全绩效 $I = \{I_1, I_2, \dots, I_n\}$ 相较安全绩效基线 $B = \{B_1, B_2, \dots, B_n\}$ 的偏离。将飞行过程分为 n 个时段 T_1 至 T_n ，则航班在时段 T_i 的韧性表征变量 R_i 可由下式计算：

$$R_i = I_i - B_i \quad (1)$$

式中，安全绩效基线 B_i 是一个固定标准，由大范围的飞行数据样本综合决定。

航班在时段 T_i 的综合安全绩效 I_i 如下：

$$I_i = \sum_j \omega_{ij} m_{ij} \quad (2)$$

式中， m_{ij} 为 QAR 特征参数 x_{ij} 在时段 T_i 的安全裕度， ω_{ij} 为权重。

m_{ij} 可由下式计算：

$$m_{ij} = \min \begin{cases} \frac{U(x_{ij}) - \bar{x}_{ij}}{U(x_{ij}) - O(x_{ij})} \\ \frac{\bar{x}_{ij} - L(x_{ij})}{O(x_{ij}) - L(x_{ij})} \end{cases} \quad (3)$$

式中， \bar{x}_{ij} 为 x_{ij} 的均值； $U(x_{ij})$ 和 $L(x_{ij})$ 分别为 x_{ij} 在

FOQA 中的监控上限和下限； $o(x_{ij})$ 为 x_{ij} 的最佳值或者推荐值。

为后续分析方便，在 \mathbf{R} 中增加虚拟参考点 R_0 作为序列起点，即 $\mathbf{R}=\{R_0, R_2, \dots, R_n\}$ ，并令 R_0 为 0，假设运行的初始状态处于安全绩效基线。

2.2 量化韧性绩效表征变量趋势特征

采用 STEM 聚类技术对航班韧性绩效表征变量的趋势特征进行分析。STEM 聚类由卡耐基梅隆大学 Jason Ernst 等^[23]开发，技术优势在于良好的短时间序列数据处理能力，能够快速实现大量短时间序列数据的动态模式捕捉；其次，STEM 基于一组预定义的、具有明确趋势含义的趋势画像（Temporal Profiles）进行聚类，可以直接识别并量化如稳定、上升、下降等与韧性表现相关的趋势特征，相比 K-Means 等传统聚类方法在趋势识别和解释性上更具优势；此外，STEM 内置了基于置换检验（Permutation Test）的统计显著性评估方法，可以检验趋势画像及其聚类航班数量的显著性，为趋势特征分析提供统计支持，增强结论的客观性。

趋势特征分析步骤如下：

（1）确定参考趋势画像

根据韧性表征变量的长度 n ，设定时间序列相邻点的变化幅度 c 以及参考趋势画像个数 z ，从而确定参考趋势画像集合 $\mathbf{S}=(s_1, s_2, \dots, s_z)$ ，每个参考趋势画像都由 1 个自然数序列表示。

（2）航班样本 STEM 聚类

筛选在安全绩效基线 \mathbf{B} 附近的航班。记总航班样本为 \mathbf{F} ，满足筛选条件的航班样本记为 \mathbf{F}_b ， \mathbf{F}_b 中所有的航班的运行模式定义为“基线模式”。

通过 STEM 聚类将剩余航班的韧性表征序列与参考趋势画像按照最小距离原则进行聚类。参考趋势画像 s_k 对应的航班簇记为集合 \mathbf{F}_k ，采用置换检验检查每种参考趋势画像的显著性。

（3）分析航班簇的趋势特征

对于显著趋势画像 s_k ，其表达式为自然数序列 $(V_0, V_1, V_2, \dots, V_n)$ ，计算趋势特征指标 TI ：

$$TI = \frac{1}{n} \sum_{i=0}^{i=n-1} (V_{i+1} - V_i) \quad (4)$$

根据 TI 的符号将航班簇划分为 2 类： $TI \geq 0$ ，则对应航班簇是正向趋势主导。反之，则是负向趋势主导。一般而言，正向趋势主导说明飞行机组对运行安全做出了正向贡献，飞行安全韧性较高。

2.3 计算航班簇的实际偏离（global deviation）

本节通过定义累积代数偏离（Cumulative Algebraic Deviation, CAD）和累积绝对偏离（Cumulative Absolute Deviation, CSD）指标，对

航班簇相对于安全绩效基线的偏离特征进行量化。这两个指标从不同维度刻画了航班簇韧性表征变量的偏离情况，其中，CAD 关注净偏离方向与大小，CSD 则关注总偏离幅度或波动性。

（1）累积代数偏离

累积代数偏离 CAD 被定义为韧性表征变量在各个时段取值的代数和。对于航班 f ，其韧性表征变量序列 $\mathbf{R}^f = (R_0^f, R_1^f, R_2^f, \dots, R_n^f)$ ，则航班 f 的累积代数偏离 CAD^f 如下：

$$CAD^f = \sum_i R_i^f \quad (5)$$

对于航班簇 \mathbf{F}_k ，其累积代数偏离 CAD^k 可用下式计算：

$$CAD^k = \frac{1}{N(\mathbf{F}_k)} \sum_{f \in \mathbf{F}_k} CAD^f \quad (6)$$

CAD^k 反映了航班簇 \mathbf{F}_k 相对于基线绩效的平均净偏离方向与幅度。由于计算中正负偏离会相互抵消， CAD^k 主要衡量整体偏离态势。 CAD^k 大于等于 0，说明航班簇 \mathbf{F}_k 中的航班趋近于更高的安全水平，是“正向偏离主导”；反之，则是“负向偏离主导”。 CAD^k 越大，意味着更高的安全韧性。

（2）累积绝对偏离

为全面衡量航班绩效偏离基线的总幅度（即波动程度或不稳定性），定义累积绝对偏离 CSD，即航班韧性表征变量在各个时段的偏离量绝对值之和。

航班 f 的累积代数偏离 CSD^f 可用下式计算：

$$CSD^f = \sum_i |R_i^f| \quad (7)$$

航班簇 \mathbf{F}_k 的累积代数偏离 CSD^k 如下：

$$CSD^k = \frac{1}{N(\mathbf{F}_k)} \sum_{f \in \mathbf{F}_k} CSD^f \quad (8)$$

CSD^k 值越大，说明 \mathbf{F}_k 的综合安全绩效相对于基线的累积波动程度越大，即飞行过程中的不稳定性或偏离幅度更大，安全管理需要对 CSD 值高的航班或航班簇投入更多的关注。

2.4 分析航班韧性绩效模式

飞行安全韧性作为航空公司的一种内部能力，会直接地影响到外显的飞行安全绩效，进而使得航班样本展现出不同的绩效模式。通过划分航班韧性绩效模式并分析总航班样本的绩效结构，可以实现航空公司飞行安全韧性评价。

（1）绩效模式划分

整体上，可以将全运行航班样本的聚类簇划分为如下三个模式：

模式 I—基线模式：航班簇 F_b 的综合安全绩效 I 在安全绩效基线 B 附近，即航班韧性表征变量 R 对应基线趋势画像 $s_b = (0, 0, \dots, 0)$ 。

模式 II—特征模式：航班簇 F_k 的综合安全绩效 I 较安全绩效基线 B 一定程度偏离，同时又通过置换检验，呈现显著变化趋势。模式 II 的航班簇对应显著参考趋势画像 $s_k, 0 \leq k \leq z$ 。

模式 III—混乱模式：航班簇 F_c 较基线绩效具有一定程度偏离，但其趋势画像不显著。该模式认为是由随机现象导致，大概率对应异常安全状态的航班。

进一步，根据趋势特征 (TI 符号) 以及偏离特

征 (CAD 符号)，将模式 II 中的航班簇划分为 4 个子模式：

- a) 模式 II-1：正向趋势与正向偏离主导，记为 (+, +)
- b) 模式 II-2：负向趋势与正向偏离主导，记为 (-, +)
- c) 模式 II-3：负向趋势与负向偏离主导，记为 (-, -)
- d) 模式 II-4：正向趋势与负向偏离主导，记为 (+, -)

综上所述，航空公司的飞行运行样本，都可以划分到如图 2 所示的类别中。



图 2 航空公司飞行安全韧性绩效模式划分

Fig. 2 Modes of airline flight safety resilience performance

(2) 绩效结构分析

值得注意的是，不同于“安全—不安全”“违规—不违规”这些二元模式划分，图 2 将系统飞行安全绩效划分为更精细的模式。本研究分别提出了 2 个反映系统绩效结构的三元组，来衡量航空公司的飞行安全韧性。

第 1 个三元组反映模式 I、模式 II，和模式 III 航班数量占总航班数量的比重，记为总体绩效结构 $G = (G_1, G_2, G_3)$ ， G_1 、 G_2 和 G_3 分别被称为基线模式运行率、特征模式运行率，以及混乱模式运行率，可由式 (9) 计算：

$$G_i = \frac{N_i}{N} \quad (9)$$

式中， $i \in \{1, 2, 3\}$ ， N 为总航班数， N_i 为模式 i 对应的航班数量。

如果航空公司的 G_1 很大， G_3 很小，说明该公司的飞行安全绩效越稳定，安全韧性越高。特征模式中里面包含了不同趋势特征与偏离特征的模式，因此不能通过 G_2 的大小直接推断航空公司的安全绩效高低。第 2 个三元组为航班样本的特征绩效结构 $C = (C^+, C, C^-)$ ，其中 C^+ 、 C 和 C^- 分别表示“正向趋势与正面偏离主导的航班占比”，

“正向趋势或正面偏离主导的航班占比”，以及“负向趋势与负面偏离主导的航班占比”，可通过式 (10) 至 (12) 计算：

$$C^+ = \frac{N_{21}}{N_2} \quad (10)$$

$$C = \frac{N_{22} + N_{24}}{N_2} \quad (11)$$

$$C^- = \frac{N_{23}}{N_2} \quad (12)$$

式中， N_{21} 、 N_{22} 、 N_{23} 、 N_{24} 分别为模式 II-1 至 II-4 对应的航班数量。

在特征模式中，如果 C^+ 很大，则说明该公司的运行安全绩效如果发生偏离，多会朝着正向的趋势、正面结果发展，是高安全韧性的体现。反之， C^- 很大则表明安全绩效的发展趋势和偏离结果是负面的，是典型的安全韧性不足。 C 则表示大多数的偏离只能具备一种正面的发展动态。

(3) 飞行安全韧性评价

基于上述分析，为方便多个航司进行飞行安全韧性的横纵向比较，综合 G 与 C ，计算飞行安全韧

性指标 (Flight Safety Resilience Index, FSRI) :

$$FSRI = \omega_1 G_1 + G_2 (\omega_2 C^+ + \omega_3 C - \omega_4 C^-) - \omega_5 G_3 \quad (13)$$

式中, ω_i 为韧性绩效管理乘数, 可以由飞行安全管理者的管理偏好决定。

FSRI 越大, 说明安全韧性越高。每个运行主体, 可以通过对 FSRI 的持续评估、分析与韧性绩效改进, 实现飞行安全韧性绩效的管理。

3 案例研究

本研究以中国境内的 A 航空公司为例, 收集了该公司 2021 年度第一季度 875 个航班的 QAR 数据。3.1-3.4 节展示了该航空公司飞行韧性评价的过程及其结果, 3.5 节将韧性评价结果与 FOQA 超限检测结果进行对比, 验证安全韧性评价的有效性。

表 1 最终进近阶段的特征飞行参数

Tab.1 Characteristic flight parameters for the final approach phase

| 高度 | 2000—1000 英尺 | 1000—500 英尺 | 500—200 英尺 | 200—50 英尺 | 50—0 英尺 |
|----|--------------|--------------------|--------------------|--------------------|-------------------|
| 时段 | T_1 | T_2 | T_3 | T_4 | T_5 |
| 参数 | 垂直下降率 | 垂直下降率 空速 滚转角 | 垂直下降率 空速 滚转角 | 垂直下降率 空速 滚转角 | 垂直下降率 / 滚转角 |
| | 滚转角 | 下滑道偏离 航向道偏离 | 下滑道偏离 航向道偏离 | 下滑道偏离 航向道偏离 | 下滑道偏离 航向道偏离 |

3.2 趋势分析结果

将总航班样本综合安全绩效 I 的均值作为安全绩效基线 B , 并以 $B \pm 0.1$ 为筛选条带, 从 875 个航班样本中筛选出 454 个航班, 将其划分为“模式 I-基线模式”。图 3 为基线模式航班簇 F_b 的韧性绩效表征变量 R 的变化情况。设定 ± 0.1 阈值的主要目的是界定一个围绕安全绩效基线的窄带区域, 该区域内的波动被视为操作上可接受的微小偏离或系统固有噪声。为在实际数据中有效地区分出稳定维

3.1 航班进近阶段韧性绩效表征变量

在飞行过程中, 最终进近阶段是影响飞行安全最为关键的阶段, 因此以该阶段为例, 构建韧性绩效表征变量。

在进近过程中, 不稳定进近会诱发飞行失控 (LOC-I), 可控飞行撞地 (CFIT) 和跑道偏出 (RE) 等 3 类事故, CFIT 事故可以理解为偏离下滑道、航向道规定范围的结果; RE 大多是由于过大的空速和垂直下降率使飞机携带过多能量造成的; LOC-I 事故一般由不安全的飞行姿态或者空速小于最低阈值导致。基于上述分析, 提取空速、垂直下降率、滚转角、下滑道偏离和航向道偏离度等 5 个参数作为进近过程的特征飞行参数, 如表 1 所示。结合 FOQA 监控标准, 依据飞行高度将最终进近过程划分出 5 个阶段, 添加虚拟参考点 R_0 后, 构造长度为 6 的韧性绩效表征变量。

持在基线水平附近的航班 (模式 I) 和发生了统计或操作意义上显著偏离的航班, 从而提升对后者的分析效率和针对性, 阈值的具体取值也可以根据数据特性和分析需求进行调整。

通过 STEM 软件, 设置 $n=6, c=2, z=50$, 构建 50 个参考趋势画像, 如图 4 所示, 每个方框中的折线图代表一种趋势画像, 方框左上角的数字为趋势画像的编号。

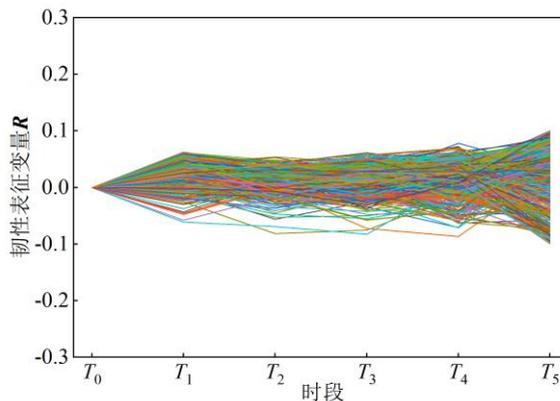


图 3 基线模式的航班簇 F_b

Fig. 3 Flight clusters F_b for the baseline mode

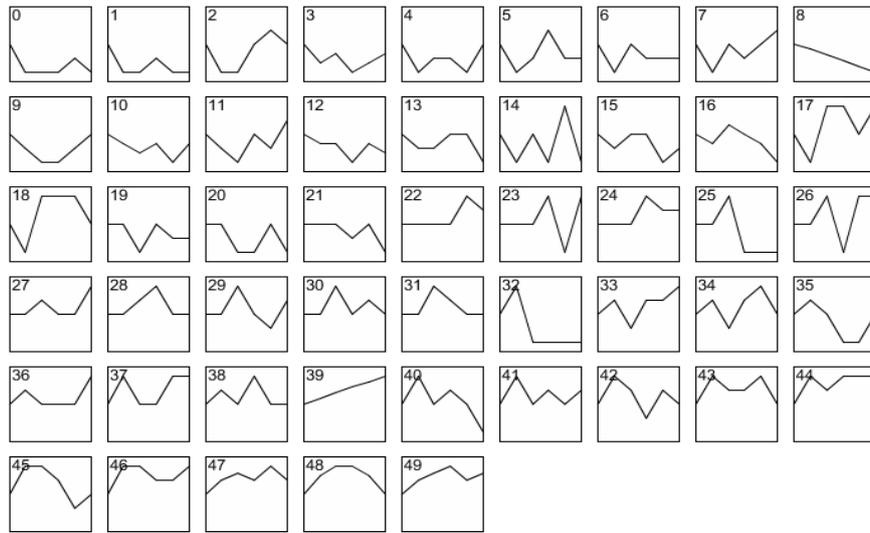
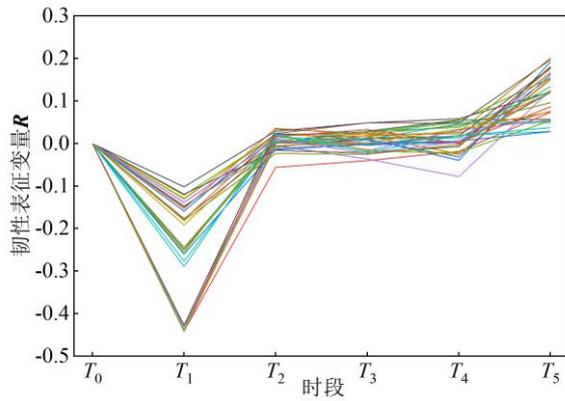


图 4 参考趋势画像集合 S

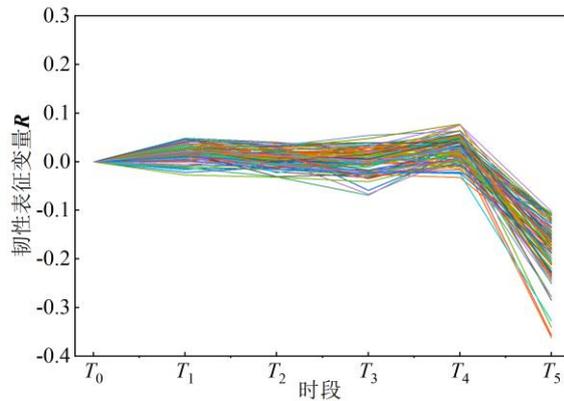
Fig. 4 The set of reference temporal profiles S

将剩余的 421 个航班与图 4 中的 50 个参考趋势模型进行匹配, 根据最小距离原则将航班聚类到不同的趋势画像, 利用 STEM 内置的置换检验方法, 识别出 5 个显著的趋势画像 ($p < 0.05$), 分别为: $s_7 (0, -2, 0, -1, 0, 1)$, $s_{21} (0, 0, 0, -1, 0, -2)$, $s_{27} (0, 0, 1, 0, 0, 2)$, $s_{36} (0, 1, 0, 0, 0, 2)$ 和 $s_{40} (0, 2, 0, 1, 0, -2)$ 。对应的航班簇分别记为 F_7 , F_{21} , F_{27} , F_{36} 和 F_{40} , 上述

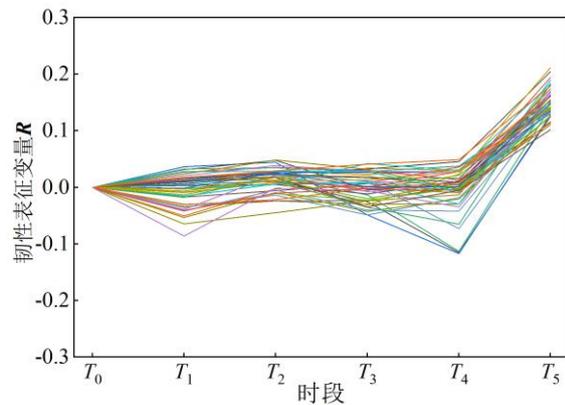
航班簇一共包含 355 个航班, 属于特征模式。如图 5 所示, F_7 中的韧性表征变量呈现波动上升的趋势; F_{27} 和 F_{36} 中, 韧性表征变量在最后一个时段 (50—0 英尺) 出现了明显的上升趋势; F_{21} 中, 韧性表征变量在 50—0 英尺则出现了明显的下降趋势; F_{40} 的趋势与 F_7 几乎相反, 韧性表征变量呈现波动下降的趋势。



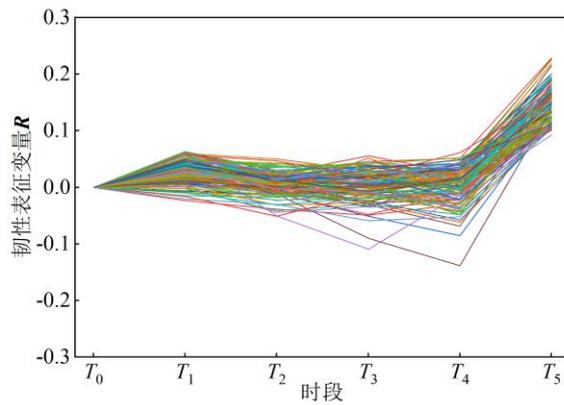
(a) F_7



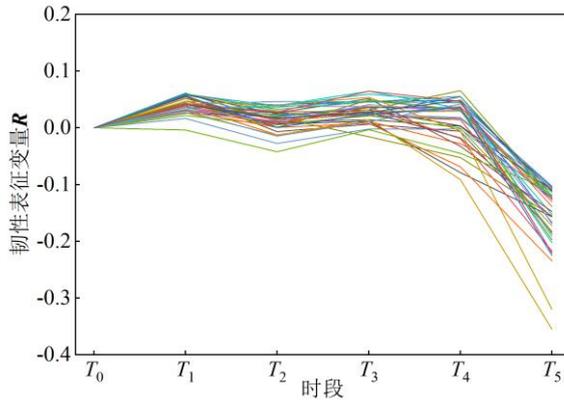
(b) F_{21}



(c) F_{27}



(d) F_{36}



(e) F₄₀

图 5 特征模式对应的航班簇

Fig. 5 Flight clusters for the characteristic mode

除去基线模式和特征模式中的航班，剩余 86 个航班的韧性表征变量由于其趋势画像不显著，被划分入“混乱模式”，混乱模式的航班簇 F_c 如图 6 所示，这些序列的变化趋势较为复杂，并且某些航班的韧性表征变量取值非常低。

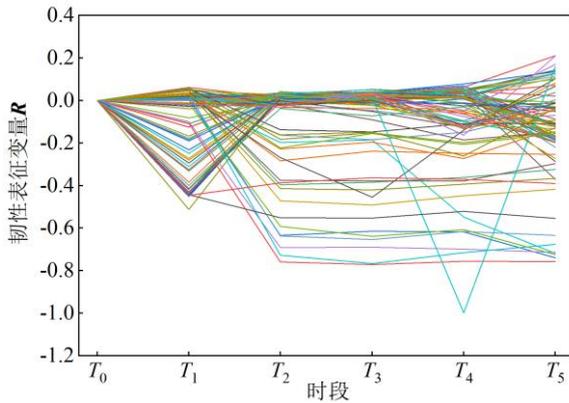


图 6 混乱模式的航班簇

Fig. 6 Flight clusters F_c for the chaotic mode

通过计算趋势模式的 *TI*，可以将所有航班簇为两类。正向趋势主导的航班簇为：F_b (*TI*=0)，F₇ (*TI*=0.2)，F₂₇ (*TI*=0.4)，F₃₆ (*TI*=0.4)；负向

表 2 不同航班簇的偏离

Tab. 2 Deviation of different flight clusters

| 航班簇 | F _b | F ₇ | F ₂₁ | F ₂₇ | F ₃₆ | F ₄₀ | F _c |
|------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|----------------|
| CAD | 0.073 | -0.121 | -0.127 | 0.151 | 0.191 | -0.064 | -0.542 |
| CSD | 0.144 | 0.437 | 0.272 | 0.239 | 0.239 | 0.285 | 0.7 |
| 航班数量 | 454 | 33 | 89 | 46 | 128 | 39 | 86 |

3.4 航空公司安全韧性评价及管理启示

(1) 安全韧性评价结果

结合 3.1-3.3 节的分析，将航班样本划分到不同的绩效模式并计算总体绩效结构 *G* 与特征绩效结构 *C*，如表 3 所示。

总体绩效结构 *G* 表明，该航空公司超过一半

趋势主导的航班簇为：F₂₁ (*TI*=-0.4)，F₄₀ (*TI*=-0.4)。混乱模式没有固定的趋势画像，因此趋势指数无法计算。

3.3 实际偏离测量结果

表 2 为不同航班簇的 *CAD* 与 *CSD*。

根据 *CAD* 的符号，可以判断 F_b，F₂₇ 和 F₃₆ 都是正向偏离主导，这些簇中的航班整体上具有更高的安全绩效；F₇，F₂₁，F₄₀ 和 F_c，都是负向偏离主导，这些航班簇的安全绩效水平要低于安全绩效基线。

从 *CSD* 的大小来看，显而易见，F_b 的 *CSD* 值最小，因为基线模式航班的综合安全绩效都在一个较窄的条带内；F_c 的 *CSD* 值最大，为 0.7，接近 F_b 的 5 倍，反映出混乱模式航班的绩效波动程度非常大，加之 F_c 的 *CAD* 值为负，推断这类航班可能多为传统飞行安全管理中关注的超限运行。

(*G*₁=51.9%) 航班都能保持在安全绩效基线附近，这构成了其运行韧性的基础；接近十分之一 (*G*₃=9.8%) 的运行由于伴随强随机性或异常趋势，被归入混乱模式，这部分航班代表了运行中难以预测和控制的不稳定因素，直接削弱了整体韧性；其余近四成 (*G*₂=38.3%) 的航班呈现出显著的趋势特征，属于特征模式。

该航空公司特征模式（模式 II）内部的绩效结构 $C=(51.9\%, 9.9\%, 38.2\%)$ 进一步揭示了运行趋势的多样性：代表正向韧性表现（正向趋势或正面偏离主导）的航班合计占 61.8%，其中，正向趋势与正面偏离同时主导（模式 II-1）的航班占特征模式航班量的 51.9%，仅有一个正向维度主导（模式 II-2, II-4）的航班占比为 9.9%。然而，负向趋势与负面偏离同时主导（模式 II-3）的航班占比达到 38.2%，这部分航班代表了系统性地偏离理想状态，是另一个显著拉低韧性的因素。

综合 G 与 C ，根据公式(15)并取权重为 1，计算得到该航司的飞行安全韧性指数 $FSRI$ 为 0.51。这个数值位于 0 到 1 标度的中间区域，直观地反映了该航空公司安全韧性的现状：既有超过半数航班稳定运行的积极面（ G_1 主导），但也显著受到混乱模式（ G_3 ）和特征模式中负向运行（ C^- ）的削弱。具体而言， $FSRI=0.51$ 意味着该航司在维持常态运行稳定性方面表现尚可，但在有效吸收干扰、防止负向趋势蔓延以及管理极端混乱运行方面存在明显不足，整体韧性水平有较大的提升空间。显然，由于 G_3 与 C^- 的值都相对较大，导致 $FSRI$ 值未能达到更高水平。

(2) 管理启示

该 $FSRI$ 得分及绩效结构分布为航空公司的安全管理提供了具体的改进方向。

一是，重点关注与干预低韧性模式。模式 III 和模式 II-3 是拉低整体韧性的主要绩效模式，应作为安全管理的优先关注点。建议安全管理部门深入分析这两类模式航班的具体运行特征、发生环境（如天气、机场条件）、机组构成、偏离参数等，

结合飞行数据、机组报告、签派记录等多源信息，识别导致混乱或系统性负向偏离的根本原因，制定并实施精准的干预措施。

二是，巩固并推广高韧性模式经验。模式 I 代表了稳定的常态运行，模式 II-1 代表了积极应对并展现良好趋势的运行。应分析这些航班的成功经验和保护性因素，并通过案例学习、经验分享会、修订培训材料等方式，将这些成功的经验在整个飞行队伍中推广，以提升整体运行的稳定性和适应性。

三是，利用韧性评价结果优化安全管理体系。将 QAR-FSRAM 的评价结果（包括绩效模式的识别、 G 、 C 等结构参数、 $FSRI$ 值）整合到航空公司的安全管理体系中。例如，将 $FSRI$ 作为一项新的安全绩效指标进行趋势监控；将识别出的模式 III 和 II-3 航班作为风险管理的重点输入，用于评估风险水平和制定缓解措施；将对低韧性模式的干预效果和高韧性模式经验的推广情况纳入安全保障环节的监控范围，评估管理措施的有效性。通过这种方式，助推韧性评估从静态的评价工具转变为驱动安全管理体系持续改进的动态引擎。

考虑到 $FSRI$ 分值与评价等级之间的映射关系可能需要根据多个航空公司 $FSRI$ 分布确定，因此本研究未对航空公司飞行安全韧性建立评价等级。此外，韧性评估的重点，并不仅仅是得出 $FSRI$ 数值，更在于通过这个过程，帮助航空公司的安全管理者精准识别本公司的绩效结构，并视情结合舱音记录、运行日志、机组报告、航后总结等安全信息，深入探索和了解不同运行模式航班的产生原因，进而提出针对性的对策，使得飞行运行朝着更加稳定、更加具有韧性的方向发展。

表 3 A 航空公司的飞行绩效结构

Tab. 3 Flight performance structure of Airline A

| 绩效模式 | 模式I | | | 模式II | | | 模式III |
|-------|--------------|--------------|----------|--------------|----------|-----------|-------------|
| 航班簇 | F_b | F_{27} | F_{36} | F_{21} | F_{40} | F_7 | F_c |
| TI | 0 | 0.4 | 0.4 | -0.4 | -0.4 | 0.2 | / |
| CAD | 0.073 | 0.151 | 0.191 | -0.127 | -0.064 | -0.121 | -0.542 |
| 航班数量 | 454 | 46 | 128 | 89 | 39 | 33 | 86 |
| 子模式 | / | 模式II-1 | | 模式II-3 | | 模式II-4 | / |
| C | / | $C^+=51.9\%$ | | $C^-=38.2\%$ | | $C=9.9\%$ | / |
| G | $G_1=51.9\%$ | | | $G_2=38.3\%$ | | | $G_3=9.8\%$ |

3.5 安全韧性评价结果的验证

为客观评估本研究提出的基于 QAR 的飞行安全韧性评价模型（QAR-FSRAM）的实际效用，特别是其识别真实飞行运行风险的能力，采用飞行运行品质监控（FOQA）对安全韧性评价结果进行验证。FOQA 是行业广泛认可的飞行安全监控手段，其核心在于中国民航局相关规定（如 AC-121/135-FS-2012-45R1）及航空公司运行标准，

监测航班飞行参数是否超过预设的监控标准阈值。FOQA 超限事件（Exceedance）是对飞行操作偏离既定标准或安全边界的客观记录，因此，FOQA 超限事件的发生频率可视为衡量实际飞行操作中风险水平的关键指标。

本研究提出的 QAR-FSRAM 旨在分析量化航空公司整体飞行运行的韧性水平（体现为 $FSRI$ 值），并识别不同的运行绩效模式。理论上，较低的安全

韧性水平应对应着更高的实际运行风险暴露。根据 2.4 节，模式 III（混乱模式，对应 G_3 指标）和模式 II-3（负向趋势与负向偏离主导，对应 C^- 指标）是拉低 $FSRI$ 得分的关键因素，代表了较低的飞行安全韧性状态。因此，本研究的验证逻辑设定为：检验由 QAR-FSRAM 识别出的、导致 $FSRI$ 得分偏低低韧性模式（模式 III 和模式 II-3），是否确实与更高的 FOQA 超限事件发生风险显著相关联。

为执行验证，本研究首先参照 FOQA 监控标准，对该航空公司研究时段内的 875 个航班样本进行了超限检测。如表 4 所示，共检测到 116 次飞行参数超限，分布在 5 个关键飞行阶段，其中 T_5 时段（50 英尺—0 英尺）的下滑道偏离最为突出。这 116

次超限共涉及 52 个独立航班，表明部分航班存在多次超限情况。超限航班占样本总数的约 6.0%（52/875），这初步印证了该航空公司在研究时段内整体飞行安全绩效存在改进空间，与 3.4 节中 $FSRI=0.51$ 所反映的“韧性不高”的评价提供了初步的印证。

验证的关键在于考察超限事件在不同韧性绩效模式间的分布差异。如图 7 所示，在 52 个超限航班中，51 个航班被识别在低韧性的模式中，其中 24 个航班属于模式 III，27 个航班属于模式 II-3，仅有 1 个超限航班属于韧性最高的模式 I（基线模式）。

表 4 超限航班检测结果

Tab. 4 Exceedance flight detection results

| 超限次数 | T_1 | T_2 | T_3 | T_4 | T_5 |
|--------|-------|-------|-------|-------|-------|
| 垂直下降超限 | 0 | 0 | 0 | 0 | 0 |
| 空速超限 | 0 | 0 | 0 | 1 | 2 |
| 下滑道超限 | 0 | 7 | 7 | 10 | 41 |
| 航向道超限 | 0 | 12 | 12 | 12 | 12 |
| 超限航班 | 0 | 13 | 13 | 16 | 48 |

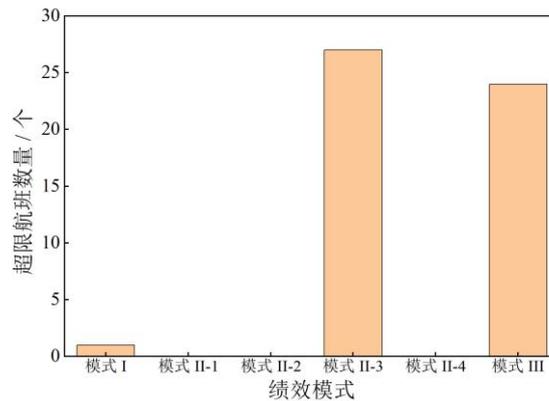


图 7 超限航班在不同绩效模式中的分布

Fig. 7 Distribution of exceedance flights in different performance modes

进一步，本研究分析了这 52 个超限航班在不同运行模式中的分布，如图 7 所示。为了更精确地量化风险关联，本研究计算了不同模式下的航班超限率，即该模式下发生超限的航班数占该模式总航班数的比重。结果表明，模式 III 和模式 II-3 的航班超限率分别 27.9% 和 21.1%，远高于模式 I（基线模式）的超限率 0.2%，以及模式 II-1（正向趋势与正面偏离）和模式 II-4（单一正向维度主导）的超限率 0.0%。超限率的量级差异，清晰的揭示了低韧性模式（II-3, III）与显著升高的 FOQA 超限风险之间的强关联性，以及其他模式（I, II-1, II-4）的低风险特性。这表明：

(1) QAR-FSRAM 方法能够有效区分具有显

著不同实际运行风险水平的航班运行模式。

(2) 被模型识别为拉低整体韧性得分的关键低韧性模式（模式 III 和 II-3），是 FOQA 超限事件的“重灾区”，其风险水平远超其他运行模式。

因此，FOQA 的验证结果有力地支持了本研究构建的 QAR-FSRAM 方法的有效性和实际应用价值，证明了该模型不仅能评估整体韧性水平，更能识别出需要优先关注和干预的高风险运行模式。

4 结论

(1) 本研究构建了一种基于 QAR 数据的飞行安全韧性评价方法（QAR-FSRAM）。该方法通过建立韧性绩效表征变量，量化表征变量趋势特征和

偏离特征，从而揭示航班样本绩效结构并实现航空公司飞行安全韧性评价。

(2) 以 A 航空公司进近阶段 QAR 数据为例，案例研究结果表明，该航空公司飞行安全韧性值 (FSRI) 为 0.51，存在明显进步空间，混乱模式 (模式 III) 和负向特征模式 (模式 II-3) 航班占比较高是削弱该航空公司韧性的主要因素。进一步，通过与 FOQA 超限事件数据的对比分析，验证了 QAR-FSRAM 识别出的低韧性绩效模式与实际运行风险之间存在相关性，证实了该方法在识别潜在运行风险方面的有效性。

(3) 基于韧性评价结果和航班韧性绩效模式分析，QAR-FSRAM 可为航空公司的安全管理提供

决策支持：指引管理层优先关注和干预特定的高风险运行模式，并将韧性评估指标 (如 FSRI、绩效模式分布) 作为前瞻性的安全绩效指标纳入安全管理体系，以促进安全管理的精细化和前瞻性。

(4) 本研究的案例分析基于单一航空公司特定时段和飞行阶段的数据，结论的普适性尚需更大范围的数据验证。未来的研究应致力于扩展数据样本 (涵盖不同航空公司、更长时间周期、不同飞行阶段及运行环境)，以检验 QAR-FSRAM 方法的广泛适用性，并探索建立可供比较的韧性基准。

参 考 文 献

- [1] 黄浪, 吴超, 杨冕, 等. 韧性理论在安全科学领域中的应用[J]. 中国安全科学学报, 2017,27(03): 1-6.
Huang Lang, Wu Chao, Yang Mian, et al. Application of resilience theory in the field of safety science[J]. China Safety Science Journal, 2017, 27(3): 1-6.
- [2] 罗通元. 安全韧性学基本概念和理论体系探讨[J]. 安全与环境学报, 2022,22(01): 280-291.
Luo Tongyuan. Basic concepts and theoretical system of safety resilience[J]. Journal of Safety and Environment, 2022, 22(1): 280-291.
- [3] HOLLNAGEL E, PARIÈS J, WOODS D, et al. Resilience Engineering in Practice: A Guidebook: Resilience Engineering in Practice: A Guidebook[M]. England: Ashgate, 2010.
- [4] 黄浪, 吴超, 王秉. 系统安全韧性的塑造与评估建模[J]. 中国安全生产科学技术, 2016,12(12): 15-21.
Huang Lang, Wu Chao, Wang Bing. Modeling and evaluation of system safety resilience[J]. Journal of Safety and Technology in Production, 2016, 12(12): 15-21.
- [5] HOLLNAGEL E. Safety-II in Practice: Developing the Resilience Potentials[M]. Routledge, 2017.
- [6] Flight Safety Foundation. Learning From All Operations: Expanding the Field of Vision to Improve Aviation Safety[EB/OL]. [(2021-07-29)[2024-06-10]. <https://flightsafety.org/wp-content/uploads/2021/07/Learning-from-All-Operations-FINAL.pdf>.
- [7] ICAO. Doc 9995: Manual of evidence-based training[EB/OL]. [(2013-12-30)[2022-10-15]. <https://skybrary.aero/sites/default/files/bookshelf/3177.pdf>.
- [8] American Airlines. Trailblazers into Safety-II: American Airlines' learning and improvement team, a white paper outlining AA's beginnings of a Safety-II journey[EB/OL]. [(2020-01-06)[2023-10-15]]. <https://skybrary.aero/bookshelf/trailblazers-safety-ii-american-airlines%E2%80%99-learning-and-improvement-team>.
- [9] EUROCONTROL. The Systemic Potentials Management - building a basis for resilient performance:a white paper[EB/OL]. (2021-09-30)[2023-10-15].<https://skybrary.aero/bookshelf/systemic-potentials-management-building-basis-resilient-performance>.
- [10] 王兴隆, 赵俊妮, 贺敏. 基于贝叶斯网络的空中交通运输系统韧性评价[J]. 南京航空航天大学学报, 2022, 54(6):1121-1130.
Wang Xinglong, Zhao Junni, He Min. Resilience evaluation of air transportation system based on Bayesian network[J]. Journal of Nanjing University of Aeronautics and Astronautics, 2022, 54(6): 1121-1130.
- [11] 郭九霞. 新一代民航运输系统安全韧性理论与方法研究[D]. 成都: 电子科技大学, 2022.
Guo Jiuxia. Research on the theory and methodology of safety resilience in the next generation civil aviation transportation system[D]. Chengdu: University of Electronic Science and Technology of China, 2022.
- [12] LI H, LIU X Y, ZHANG Y F, HU X B. Comparison on vulnerability of European and Chinese air transport networks under spatial hazards[J]. Transactions of Nanjing University of Aeronautics & Astronautics,2020,37(2):300-310
- [13] Barry D J. Estimating runway veer-off risk using a Bayesian network with flight data[J]. Transportation Research Part C: Emerging Technologies, 2021, 128: 1-23. (103280)
- [14] WANG L, WU C, SUN R. An analysis of flight quick access recorder (QAR) data and its applications in preventing landing

-
- incidents[J]. *Reliability Engineering & System Safety*, 2014, 127: 86-96.
- [15] ZHANG H. A review of the research of quick access recorder data[J]. *Academic Journal of Science and Technology*, 2023, 5(1):95—101.
- [16] LI X, SHANG J, ZHENG L, et al. CurveCluster+: Curve clustering for hard landing pattern recognition and risk evaluation based on flight data[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 8(23): 12811-12821.
- [17] LI C, SUN R, PAN X. Takeoff runway overrun risk assessment in aviation safety based on human pilot behavioral characteristics from real flight data[J]. *Safety Science*, 2023, 158: 105992.
- [18] ICAO. Doc 10151: Manual on human performance (HP) for regulators[EB/OL]. <https://www.icao.int/safety/OPS/OPS-Section/Documents/Advance-unedited.Doc.10151.alltext.en.pdf>.
- [19] 李瑞奇, 黄弘, 周睿. 基于韧性曲线的城市安全韧性建模[J]. *清华大学学报(自然科学版)*, 2020,60(01): 1-8.
Li Ruiqi, Huang Hong, Zhou Rui. Modeling of Urban Safety Resilience Based on Resilience Curve[J]. *Journal of Tsinghua University (Science and Technology)*, 2020, 60(1): 1-8.
- [20] 魏强, 刘加奇, 王景春, 等. 基于理想模糊物元的隧道施工安全韧性评估[J]. *中国安全科学学报*, 2021,31(08): 62-68.
Wei Qiang, Liu Jiaqi, Wang Jingchun, et al. Evaluation of Safety Resilience of Tunnel Construction Based on Ideal Fuzzy Matter Element[J]. *China Safety Science Journal*, 2021, 31(8): 62-68.
- [21] 陈超, 谭心鑫, 王迪, 等. 油库抗震韧性动态评估模型[J]. *中国安全科学学报*, 2024,34(04): 87-92.
Chen Chao, Tan Xinxin, Wang Di, et al. Dynamic Evaluation Model of Seismic Resilience of Oil Depots[J]. *China Safety Science Journal*, 2024, 34(4): 87-92.
- [22] BRUNEAU M, CHANG S E, EGUCHI R T, et al. A framework to quantitatively assess and enhance the seismic resilience of communities[J]. *Earthquake spectra*, 2003,19(4): 733-752.
- [23] ERNST J, NAU G J, BAR-JOSEPH Z. Clustering short time series gene expression data[J]. *Bioinformatics*, 2005,21(Suppl 1): 159-168.
-