



引用格式: 卫宗敏. 民机驾驶舱人机交互脑力负荷预测模型[J]. 科学技术与工程, 2021, 21(13): 5270-5274.

Wei Zongmin. A model for prediction of pilot's mental workload of human machine interface in civil flight deck[J]. Science Technology and Engineering, 2021, 21(13): 5270-5274.

医药、卫生

民机驾驶舱人机交互脑力负荷预测模型

卫宗敏

(中国民航管理干部学院航空安全管理系, 北京 100102)

摘 要 针对民机驾驶舱人机交互中飞行员脑力负荷问题,设计不同难度的人机交互飞行实验任务,分析飞行作业过程中飞行员脑力负荷在评估指标上的变化。在此基础上,提出一种新的改进的多元线形回归方法,探索人机交互中飞行员脑力负荷变化预测模型的构建方法。结果显示:飞行作业过程中脑力负荷在反应时间、正确率、NASA-TLX、SDNN 4 个评估指标上变化显著;在 RRI Count、Max RRI、Minimum RRI、Mean RRI、Max/Min 5 个评估指标上变化不显著。改进的多元线形回归模型可对不同飞行难度下个体脑力负荷水平进行预测和等级划分,平均预测准确率为 87.5%。提出的预测模型与实测数据吻合性较好,能够较准确地反映民机驾驶舱人机交互中飞行员脑力负荷变化特性,可为未来民机驾驶舱人机交互中工效学评价与优化设计提供依据。

关键词 民机驾驶舱; 脑力负荷; 人机交互; 预测; 数学模型; 多元线形回归

中图法分类号 R857.14 V7; **文献标志码** A

A Model for Prediction of Pilot's Mental Workload of Human Machine Interface in Civil Flight Deck

WEI Zong-min

(Development of Aviation Safety Management, Civil Aviation Management Institute of China, Beijing 100102, China)

[Abstract] Aiming at the pilot's mental task, human machine interface flight task was implemented at different levels, mental workload changes in different evaluation indicators during the flight was analyzed, a new improved multiple linear regression model was proposed, and a construction method of pilots' mental workload change prediction model in human-machine interaction was discovered. The results suggest that the mental workload of pilots during flight operation changed significantly in the evaluation indicators of response time, accuracy rate, NASA-TLX and the SDNN, while the evaluation indicators of RRI Count, Max RRI, Minimum RRI, Mean RRI, and Max/Min do not change significantly. By the improved multiple linear regression model, the individual mental workload with different levels of flight difficulty can be predicted and ranked, with average prediction accuracy of 87.5%. Accordingly, the proposed prediction model fit well with the measured data, and can accurately reflect the characteristics of flight deck human-machine interaction mental workload change, which can provide a basis for ergonomic evaluation and optimization design of future aircraft flight deck human-machine interaction.

[Keywords] flight deck; mental workload; human machine interaction; prediction; mathematical model; multiple linear regression

飞行特情下的脑力负荷问题是导致飞行机组人为差错的重要因素^[1]。自 20 世纪 80 年代起,研究人员开始持续关注这一问题。目前,飞机驾驶舱人机交互系统中的脑力负荷问题已上升为航空、航天、航海、核电等大型复杂系统人机工效领域热点研究课题^[2]。国外针对相关航空事故分析的相关研究指出,飞行员脑力负荷过高导致 60% 以上的事

故^[3-4]。国内一份针对中国民航事故原因进行分析统计的研究结果也表明,约 20% 的飞行事故与脑力负荷密切相关^[5]。

近年来,相关研究人员已经意识到,要解决飞行员脑力负荷问题,不仅仅要考虑加强对飞行员的培训和飞行特情训练,更应该在飞机驾驶舱人机交互的设计阶段就应考虑脑力任务设计问题,并就该

收稿日期: 2020-10-30; 修订日期: 2020-12-21

基金项目: 国家自然科学基金(61972021)

作者简介: 卫宗敏(1984—),男,汉族,山西运城人,博士,讲师。研究方向:航空事故分析与适航技术、人机交互、数学建模。E-mail: weizongmin8411@163.com。

问题开展了系列研究^[6]。但是这些研究绝大多数都着眼于探讨不同类型飞行任务下的不同的评价指标与脑力负荷之间的关系,对飞行员的脑力负荷实判别预测模型的构建关注较少。考虑到在人们所获取的信息中,80%是通过视觉通道获取的^[7],因而有必要在飞机驾驶舱显示系统人机交互设计的早期阶段,开拓测评方法,开展该系统设计与脑力负荷相关研究。

针对飞机驾驶舱人机交互界面脑力负荷的测量和评价,中外相关研究人员开展了一系列研究,采用的方法包括主观评价法、绩效测评法(包括主任务测评法和次任务测评法)、生理测评法以及多指标综合测评方法。大量的研究表明,不同测评方法分别适用于不同的飞行任务,试图以一种测量方法或指标来完成该项测评是难以实现的。同时,脑力负荷的复杂特性也决定了对其作综合评估的必要性^[8]。

在综合评估方法建模方面,目前在这一领域较常用的建模技术主要有回归分析和神经网络建模等^[9-11]。例如,Mazaeva 等^[9]采用自组织映射(self-organizing map,SOM)网络方法,对飞行模拟试验中所采集的多项脑电指标进行分析和建模,构建了预测脑力负荷水平的 SOM 网络模型。Ryu 等^[10]分别采用因素分析法和回归分析方法将多项生理指标综合对脑力负荷进行综合评价。李金波等^[11]采用反向传播(back propagation,BP)神经网络和 Elman 神经网络两种方法,构建了脑力负荷预测模型,两种模型的预测准确率分别在 61.40%~80.49%、70.07%~89.65%。

然而该建模方法在理论基础方面仍存在较大争议^[11]。而多元线性回归分析方法,能够有效用于研究不同影响因素对于多分类变量的效应,并且能够有效保留所选评价指标,防止信息丢失,预测结果稳定性较好,但是采用传统的多元线性回归方法无法对类别进行直接判定,因而采用该方法对脑力负荷类别进行数学建模,目前还未引起飞行领域的广泛重视。

现尝试改进传统的多元线形回归分析方法,并用该方法对飞行任务相关的脑力负荷进行理论建模。首先以飞行任务人机交互实验为基础,通过飞行模拟器测评人机交互飞行实验过程中被试的飞行任务绩效(反应时和正确率)、主观评价(NASA-TLX 量表)和生理变化(心率和心率变异性)来间接测量脑力负荷,探索在不同难度飞行任务过程中脑力负荷在这三类评估指标上的变化;然后筛选出对脑力负荷敏感的评估指标,基于改进的多元回归分

析方法构建脑力负荷预测模型,以达到对人机交互过程中的脑力负荷等级变化进行预测的目的。

1 实验测量方法

1.1 实验设计

实验采用被试内设计,每个被试共进行 4 次飞行模拟实验,包括对照组、低、中、高四种脑力负荷飞行任务。实验中仪表信息会随机呈现异常,不同难度飞行任务条件下需监视的仪表信息如表 1 所示。

表 1 不同难度飞行任务条件下需监视的仪表信息
Table 1 Flight instrumentation information need to be monitored under different difficulty tasks

对照组	低负荷组	中负荷组	高负荷组
—	空速信息	空速信息	空速信息
—	俯仰角信息	俯仰角信息	俯仰角信息
—	气压高度信息	气压高度信息	气压高度信息
—	—	航向角信息	航向角信息
—	—	滚转角信息	滚转角信息
—	—	方向舵状态	方向舵状态
—	—	—	副翼位置信息
—	—	—	起落架状态
—	—	—	发动机状态

驾驶舱显示系统中所显示的飞行信息包括空速、俯仰角、气压高度、滚转角、航向角、方向舵状态、副翼位置、起落架状态、发动机状态等,异常范围的设置与文献[6]的系列研究相同。实验中,要求被试对显示系统上呈现的异常信息迅速准确地做出反应。

1.2 被试

被试者为北京航空航天大学 16 名在地面飞行模拟器上受过培训的飞行学员(21~28 岁),身体健康,无睡眠障碍,右利手,视力或矫正视力正常。实验前对被试者进行模拟飞行培训,使被试者了解实验过程、熟悉实验操作方法及实验要求。

1.3 实验仪器和测试工具

实验在飞行模拟器上进行,心电测量中采集每个被试每 5 min 内的心率、全部窦性心搏相邻心搏间的距离(R-R 间期)时间序列以及该时间段内心电图等数据。心率数值显示范围为 20~300 次/min,心率检测精度为 ± 2 次/min,采样频率为 0.05~150 Hz,波形记录速度为 25 mm/s。

实验中主观评定采用通用 NASA-TLX 量表^[12],为便于被试操作员更准确完成量表,本研究以[0,100]范围内的分值等效替代 NASA-TLX 中的刻

度线,分值越高负荷越大,从6个维度综合评定被试人员脑力负荷。

1.4 实验程序

针对每个被试,逐个开展实验。实验前,每个被试均完成一次无任何异常信息呈现的实验任务,该任务测量数据即为实验参照数据。被试每完成一次飞行任务后,即开展被试 NASA-TLX 量表。实验预设低、中、高3种复杂飞行任务,这3种飞行任务顺序在被试之间做了交叉平衡。完成实验任务的同时,计算机和心电测量设备同步相关测量数据。

1.5 实验数据记录

- 实验中记录如下数据:
- (1)飞行任务正确探测率。利用 E-prime 编制的程序自动记录每次实验中被试对异常飞行信息正确响应的次数占所有异常信息出现次数的比例就作为该实验任务条件下的飞行任务正确探测率。
 - (2)反应时间。基于 E-prime 编程,由系统自动记录自飞行异常信息出现至被试操作的反应时间,以它作为该实验任务条件下的飞行任务反应时。
 - (3)心电图 (electrocardiogram, ECG) 数据。平均心率 (mean HR)、R-R 间期数量 (RRI Count)、平均 R-R 间期 (Mean RRI)、最大 R-R 间期 (Maximum RRI)、最小 R-R 间期 (Minimum RRI)、最大 R-R 间期与最小 R-R 间期的比值 (Max:Min RRI ratio)、R-R 间期的标准差 (SDNN)。
 - (4)主观评定数据。每个负荷下的实验任务结束后用 NASA-TLX 量表度量负荷分值。

2 实验测量结果

2.1 任务难度对脑力负荷的影响

计算各个脑力负荷评估指标在不同难度的飞行任务条件下的均值,并分别对各个评估指标在不同难度的飞行任务条件下均值的差异进行方差分析,结果如图1所示。

从图1可以看出,随着飞行任务难度的增加,对飞行任务主观评价分值增大;对异常信息的反应时增加,对异常信息的正确探测率降低;Mean HR、RRI Count、Max:Min RRI ratio 呈现增加的趋势,而 Maximum RRI、Minimum RRI、SDNN 呈现降低的趋势。方差分析结果显示,在不同难度的飞行任务条件之间,NASA-TLX 主观评价分值、反应时、正确率、Maximum RRI、SDNN 均表现出极显著的差异 ($P < 0.01$)。而 Mean HR、RRI Count、Max:Min RRI ratio 和 Minimum RRI 变化的差异不显著 ($P > 0.05$)。由此可见,对飞行任务难度变化敏感的评估指标为 NASA-TLX 分值、反应时、正确率、Maximum RRI、SDNN 等。

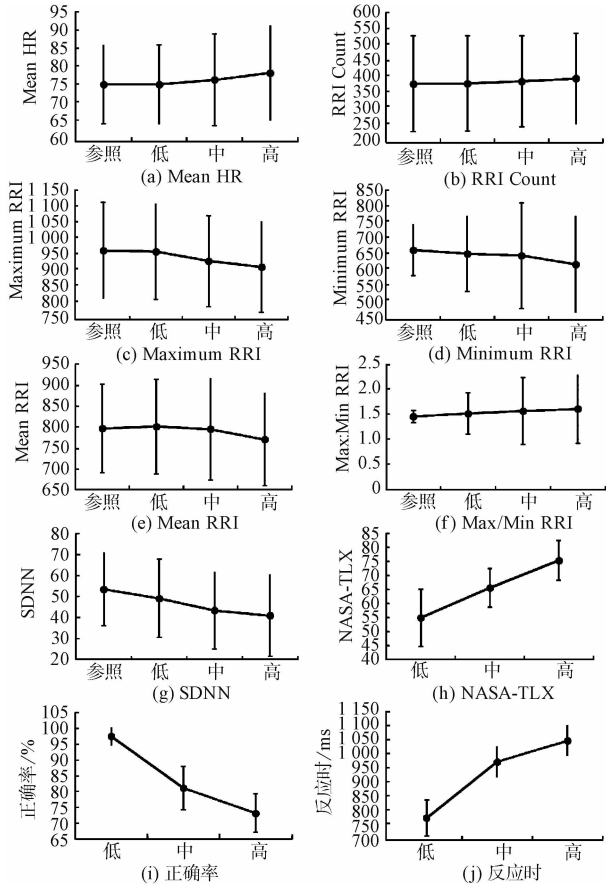


图1 不同难度飞行任务条件下均值的差异方差分析
Fig.1 Variance analysis of the differences of the mean values under different difficulty tasks

2.2 脑力负荷变化预测模型

2.2.1 模型的建立及使用说明

基于实验测量的分析结果,使用改进的多元回归分析方法,构建飞机驾驶舱人机交互脑力负荷预测模型,并用于驾驶舱人机交互中脑力负荷等级的判定。

为了保证方程中各变量的显著性,采用逐步筛选法建立回归分析模型,即 NASA-TLX 分值、反应时、正确率和 SDNN 按照逐步筛选策略进入回归模型。为便于对类别的预测,在传统多元回归分析方法的基础上做了相应的改进,所构建的回归模型如式1所示:

$$\begin{cases} y = \begin{cases} [x] + 1, & x - [x] \geq 0.5 \\ [x], & x - [x] < 0.5 \end{cases} \\ x = 0.230 + 2.223x_1 - 2.650x_2 + 0.024x_3 + 0.008x_4 \end{cases} \quad (1)$$

式(1)中: y 为预测结果; x_1 为心率变异性指标 SDNN 测量值; x_2 为 NASA-TLX 计算值; x_3 为反应时测量值; x_4 为正确率测量值。根据 x_1 、 x_2 、 x_3 和 x_4 的数值可以计算出相应 x 的值,进而求出 y 的值。若 y 的值为1,则认为属低脑力负荷水平;若 y 的值为2,则认为属中脑力负荷水平;若 y 的值为3,则认为属高脑力负荷水平。

为属高脑力负荷水平。

2.2.2 模型判别准确性的检验

模型的调整的 R^2 值为 0.827 ,较接近 1,拟合效果较好。模型预测准确率的检验,基于所构建的多元回归模型,对不同类型的认知负荷分类进行预测,其结果如表 2 所示。

表 2 脑力负荷水平等级预测结果

Table 2 Mental workload level prediction results					
实际类别	预测结果	实际类别	预测结果	实际类别	预测结果
1	1	2	2	3	3
1	1	2	2	3	3
1	1	2	2	3	3
1	1	2	2	3	3
1	1	2	2	3	3
1	1	2	2	3	3
1	1	2	2	3	3
1	1	2	2	3	3
1	1	2	2	3	2
1	1	2	2	3	2
1	1	2	3	3	3
1	2	2	3	3	3
1	1	2	2	3	3
1	1	2	2	3	2
1	1	2	2	3	3
1	1	2	2	3	3

从表 2 的预测结果可知,采用改进的多元线形回归方法建模,模型总体预测正确率为 87.50%。48 组测试样本脑力负荷中,实测结果与预测结果吻合的有 42 组,有 6 组样本实测结果与预测结果相差一级,实测结果与预测结果无相差两级以上的情况。由此可见,利用改进的多元线形回归模型对个体脑力负荷水平进行分类具有较高的准确度。

3 讨论

3.1 三类评估指标对脑力负荷变化的敏感性

目前,中外学者对飞行脑力负荷的研究大都采用模拟器试验飞行方法,即通过设置飞行任务难度或梯度,研究各个测评指标对于脑力负荷状态的敏感性^[13-15]。研究通过改变飞行任务中所需监视的仪表异常信息的数量、呈现时间和间隔时间来设置并探讨脑力负荷与三类评估指标的数学关系。NASA-TLX 量表分值呈阶梯形增加趋势,与前期研究结果一致^[16-18]。该结果说明,从本实验所选拔的被试者个人主观角度,实验任务设计与预期吻合。被试者的绩效呈阶梯形下降趋势,与前期其他飞行任务测试结果一致^[16-18]。该结果支持了 Sweller 的

认知负荷理论^[19] 和 Lavie 等^[20] 的注意负荷理论。

3.2 模型预测结果的准确性

模型整体评估结果如表 3 所示。从总体分类预测准确来看,采用改进的多元线性回归预测模型可对不同飞行难度下个体脑力负荷水平进行预测和等级划分。该模型的分类预测准确率在 81.25%~93.75%,平均预测准确率为 87.50%。

表 3 模型对脑力负荷等级的预测结果
Table 3 Results of predication accuracy rate

负荷等级	预测结果/%			
	低	中	高	预测准确率
低	93.75	6.25	0	93.75
中	0	87.50	12.50	87.50
高	0	18.75	81.25	81.25

3.3 研究的意义

通过设置动态飞行过程中所需监视的显示系统异常试验任务,获取飞行作业中脑力负荷评估指标,采用改进的多元线形回归方法建立起驾驶舱人机交互脑力负荷预测模型。

使用该模型可以对飞行机组脑力负荷水平是否在合理范围内进行实时分析或对可能出现的超脑力负荷现象进行有效预测,以利于在驾驶舱人机交互系统设计的早期阶段能够随时调整所设计的脑力任务,从而有效降低预期可能出现的脑力负荷超高或超低现象,保证航空安全。

同时,提供了如何构建飞机驾驶舱人机交互系统作业过程中脑力负荷变化预测的方法,即通过指标测量并基于改进传统回归分析方法,构建飞机驾驶舱人机交互脑力负荷预测模型,这一方法对于构建各类人机交互系统中的脑力负荷的预测模型可能具有普遍意义。

4 结论

研究获得以下结论。

(1)动态飞行模拟试验中,飞行作业正确率、反应时、心率变异性指标 SDNN、NASA-TLX 量表分值这 4 个指标对飞行任务相关的脑力负荷的变化有显著的敏感性。

(2)改进的多元线性回归预测模型克服了传统方法的不足,可对不同难度飞行任务人机交互中脑力负荷水平进行预测和等级划分。

(3)研究构建了一个简洁易用的飞机驾驶舱人机交互脑力负荷预测数学模型,该模型的分类预测准确率在 81.25%~93.75%,平均预测准确率为 87.50%。

参 考 文 献

- [1] Dismukes R K, Kochan J A, Goldsmith T E. Flight crew errors in challenging and stressful situations[J]. *Aviation Psychology and Applied Human Factors*, 2018, 8(1): 35-46.
- [2] Young M S, Brookhuis K A, Wickens C D, et al. State of science: mental workload in ergonomics[J]. *Ergonomics*, 2015, 58(1): 1-17.
- [3] 卫宗敏, 郝红勋, 徐其志, 等. 飞行员脑力负荷测量指标和评价方法研究进展[J]. *科学技术与工程*, 2019, 19(24): 1-8.
Wei Zongmin, Hao Hongxun, Xu Qizhi, et al. Review of research on pilots' mental workload evaluation indices and measurement methods[J]. *Science Technology and Engineering*, 2019, 19(24): 1-8.
- [4] 姚永杰. 军机飞行员空中脑力负荷的心理生理评定[D]. 西安: 第四军医大学, 2007.
Yao Yongjie. Mental workload assessment in military pilots using psychophysiological measures in flight[D]. Xi'an: Fourth Military Medical University, 2007.
- [5] 徐胜平, 李华明. 81 起民航等级飞行事故的机组资源管理分析[J]. *中华航空航天医学杂志*, 2003, 9(14): 174-175.
Xu Shengping, Li Huaming. Analysis of crew resource management of 81 civil aviation flight accidents[J]. *Chinese Journal of Aerospace Medicine*, 2003, 9(14): 174-175.
- [6] 卫宗敏, 完颜笑如, 庄达民. 飞机座舱显示界面脑力负荷测量与评价[J]. *北京航空航天大学学报*, 2014, 40(1): 86-91.
Wei Zongmin, Wanyan Xiaoru, Zhuang Damin. Measurement and evaluation of mental workload for aircraft cockpit display interface[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2014, 40(1): 86-91.
- [7] Pulat B M. Fundamentals of industrial ergonomics[M]. Englewood Cliffs, NJ: Prentice Hall, 1992.
- [8] 康卫勇, 袁修干, 柳忠起, 等. 飞机座舱视觉显示界面脑力负荷综合评价方法[J]. *航天医学与医学工程*, 2008, 21(2): 103-107.
Kang Weiyong, Yuan Xiugan, Liu Zhongqi, et al. Synthetic evaluation method of mental workload on visual display interface in airplane cockpit[J]. *Space Medicine & Medical Engineering*, 2008, 21(2): 103-107.
- [9] Mazaeva N, Ntuen C, Lebby G. Self-organizing map(SOM) model for mental workload classification[C]//IFSA World Congress and 20th NAFIPS International Conference. New York: IEEE, 2001: 1822-1825.
- [10] Ryu K, Myung R. Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic[J]. *International Journal of Industrial Ergonomics*, 2005, 35(11): 991-1009.
- [11] 李金波, 许百华, 田学红. 人机交互中认知负荷变化预测模型的构建[J]. *心理学报*, 2010, 42(5): 559-568.
Li Jinbo, Xu Baihua. Synthetic assessment of cognitive load in human-machine interaction process[J]. *Acta Psychological Sinica* 2010, 42(5): 559-568.
- [12] Hart S G, Staveland L E. Development of NASA-TLX (task load index): results of empirical and theoretical research[M]. *Advances in Psychology*. Amsterdam: North Holland Press, 1988.
- [13] Grassmann M, Vlemingx E, Von L A, et al. Individual differences in cardiorespiratory measures of mental workload: an investigation of negative affectivity and cognitive avoidant coping in pilot candidates[J]. *Applied Ergonomics*, 2017, 59: 274-282.
- [14] Mansikka H, Virtanen K, Harris D. Comparison of NASA-TLX scale, modified Cooper-Harper scale and mean inter-beat interval as measures of pilot mental workload during simulated flight tasks[J]. *Ergonomics*, 2019, 62(2): 246-254.
- [15] Orlandi L, Brooks B. Measuring mental workload and physiological reactions in marine pilots: building bridges towards redlines of performance[J]. *Applied Ergonomics*, 2018, 69: 74-92.
- [16] 卫宗敏. 飞机驾驶舱显示界面脑力负荷实验测量与数学建模[D]. 北京: 北京航空航天大学, 2014.
Wei Zongmin. The experimental measurement and mathematical modeling of mental workload of aircraft cockpit display interface[D]. Beijing: Beihang University, 2014.
- [17] Wei Z, Zhuang D, Wanyan X, et al. A model for discrimination and prediction of mental workload of aircraft cockpit display interface[J]. *Chinese Journal of Aeronautics*, 2014, 27(5): 1070-1077.
- [18] 陆 崑, 卫宗敏, 庄达民, 等. 飞机驾驶舱显示界面脑力负荷判别预测生理模型[J]. *北京航空航天大学学报*, 2016, 42(4): 685-693.
Lu Kun, Wei Zongmin, Zhuang Damin, et al. Integrated physiological model for mental workload assessment and prediction of aircraft flight deck display interface[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2016, 42(4): 685-693.
- [19] Sweller J. Cognitive load during problem solving: effects on learning[J]. *Cognitive Science*, 1988, 12(2): 257-285.
- [20] Lavie N, Hirst A, De Fockert J W, et al. Load theory of selective attention and cognitive control[J]. *Journal of Experimental Psychology: General*, 2004, 133(3): 339-354.