

引用格式:刘保军,蔡理,罗湘燕,等.基于线性回归的飞机战伤抢修时间估算研究[J].电光与控制,2020,27(2):94-97. LIU B J, CAI L, LUO X Y, et al. Study on ABDR time assessment based on linear regression[J]. Electronics Optics & Control, 2020, 27(2):94-97.

## 基于线性回归的飞机战伤抢修时间估算研究

刘保军<sup>a</sup>, 蔡理<sup>b</sup>, 罗湘燕<sup>a</sup>, 周平<sup>a</sup>

(空军工程大学, a. 航空机务士官学校, 河南 信阳 464000; b. 基础部, 西安 710051)

**摘要:**飞机战伤抢修(ABDR)是损伤飞机“再生”的最佳途径,抢修时间估算是评定飞机战伤等级的一个重要依据,也是实施抢修的一个重要前提和条件。针对抢修时间估算涉及的影响因素多的特点,利用主成分分析法进行降维处理,将较多的抢修时间影响因素转换为较少的主成分,并重新解释新的主成分。建立抢修时间估算的多元线性回归模型,且通过遗传算法对回归参数进行优化。实例表明,利用主成分分析法可将12个影响因素降成6个新的主成分,有效减少了回归模型的输入数量,降低了模型的复杂度,且优化的回归模型估算抢修时间的误差较小,平均相对误差为12.82%。

**关键词:**飞机战伤抢修; 抢修时间; 线性回归; 遗传算法

中图分类号: E92; TJ07 文献标志码: A doi:10.3969/j.issn.1671-637X.2020.02.020

## Study on ABDR Time Assessment Based on Linear Regression

LIU Baojun<sup>a</sup>, CAI Li<sup>b</sup>, LUO Xiangyan<sup>a</sup>, ZHOU Ping<sup>a</sup>

(Air Force Engineering University, a. Aviation Maintenance NCO Academy, Xinyang 464000, China;

b. Department of Basic Science, Xi'an 710051, China)

**Abstract:** Aircraft Battle Damage Repair (ABDR) is the best way for the battle-scarred aircrafts to fly again. The ABDR time assessment is very important for evaluating the battle damage degree of the aircraft, and is also a main precondition for making urgent repair. Considering that the ABDR time assessment is influenced by many factors, the Principal Components Analysis (PCA) method is used for dimension reduction, which transformed the large number of influence factors into uncorrelated new variables, known as principal components, which are linear combinations of the original variables. The achieved principal components are explained from a new insight. A multiple linear regression model is built for assessing ABDR time. The regressive parameters of the model are optimized through Genetic Algorithm (GA). The following results are obtained by using PCA method: 1) The 12 ABDR time assessment influence factors are reduced to 6 new principal components, which decreases the input numbers and complexity of the regression model effectively; and 2) The error of ABDR time evaluated by the optimized regression model is smaller, with an average relative error of 12.82%.

**Key words:** Aircraft Battle Damage Repair (ABDR); ABDR time; linear regression; genetic algorithm

### 0 引言

飞机战伤抢修(Aircraft Battle Damage Repair, ABDR)是保持航空部队持续作战能力最直接、最有效、最经济的途径,是现代战争夺取制空权的重要保障<sup>[1-3]</sup>。抢

修时间是确定战伤飞机是否抢修的一个重要依据<sup>[4]</sup>,因此,能否科学、合理、准确地估算抢修时间,对飞机抢修决策的合理性起着关键作用,严重影响战伤飞机的战斗恢复力。

针对抢修时间估算,人们开展了一些研究。文献[5]基于随机网络仿真的方法对复杂系统的抢修时间进行了研究;文献[6]在抢修流程网络模型的基础上,对航空装备的抢修时间进行了估算;文献[4]基于加权支持向量机,通过训练和参数优化,建立了抢修时间

收稿日期:2019-03-11

修回日期:2019-08-28

基金项目:国家自然科学基金(11975311,11405270)

作者简介:刘保军(1984—),男,山西灵丘人,博士,讲师,研究方向为飞机战伤抢修技术、可靠性分析。

估计模型;文献[7]综合抢修人员数量、备件率、装备损伤程度及运送时间4个指标,对雷达装备的抢修时间进行了估算研究。但这些研究要么以平时修理时间为基础,不能充分反映战时抢修的特点,要么估算模型的通用性不好,亦或抢修时间估算精度不高。同时,由于飞机战伤抢修的时间估算所涉及的影响因素繁多、影响机理复杂,目前的相关研究较少。

通常,为了提高抢修时间估算精度,输入参数必须包括尽可能多的影响因素,如损伤尺寸、损伤位置、战伤等级、操作熟练程度、战伤可达性、备件数、工具设备的适应程度、修理环境、损伤部件结构复杂度等,从而导致抢修时间估算模型的输入节点数量过多,增加模型的复杂度,降低抢修时间估计速度。为有效解决抢修时间估算中存在的问题,本文利用主成分分析(PCA)方法<sup>[8-9]</sup>对原始数据进行处理,去除数据间的相关性,删除其中的部分冗余信息,降低数据的维数,得到主成分分量,然后对这些主成分分量建立多元线性回归模型,并用遗传算法(GA)<sup>[10-11]</sup>优化回归参数。

## 1 主成分分析方法(PCA)

PCA主要用于数据降维<sup>[9]</sup>,其目标是寻找若干个新变量,使它们反映原始数据的主要特征,压缩原有数据矩阵的规模。得到的新变量称为主成分,它们保留了原始数据的大部分信息,且这些变量是互不相关、正交的。PCA算法过程如下。

### 1) 原始数据规范化处理。

设有 $n$ 个样本, $p$ 项指标,可得原始数据矩阵 $\mathbf{X} = (x_{ij})_{n \times p}$ , $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, p$ ,  $x_{ij}$ 表示第 $i$ 个样本的第 $j$ 项指标值。对矩阵 $\mathbf{X}$ 进行规范化处理,得到新的矩阵 $\mathbf{X}^* = (x_{ij}^*)_{n \times p}$ 为

$$x_{ij}^* = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (1)$$

式中: $\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}$ ;  $s_j = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}$ 。

### 2) 计算相关矩阵 $\mathbf{R} = (r_{ij})_{p \times p}$ 。

$$r_{ij} = \frac{\frac{1}{n-1} \sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{s_i^2 s_j^2}} = \frac{1}{n-1} \sum_{k=1}^n x_{ki}^* x_{kj}^* \quad i, j = 1, 2, \dots, p \quad (2)$$

### 3) 计算相关矩阵的特征值和特征向量。

计算相关矩阵 $\mathbf{R}$ 的特征值 $\lambda_i$ ,且 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ ,  $\mathbf{e}_i = (e_{1i}, e_{2i}, \dots, e_{pi})^T$ 为对应于 $\lambda_i$ 特征值的正交单位特征向量。

### 4) 筛选主成分,得到投影矩阵。

第 $i$ 个主成分的方差贡献率为 $\lambda_i/p$ ,前 $m$ 个主成分的累计方差贡献率为

$$M = \frac{1}{p} \sum_{i=1}^m \lambda_i \quad (3)$$

当 $M \geq 85\%$ 时,选择当前 $m$ 个主成分代替原有 $p$ 个变量的综合效果。得到投影矩阵 $\mathbf{W} = (e_1, e_2, \dots, e_m)$ ,利用下式得到 $m$ 个主成分。

$$z_h = \sum_{i=1}^p e_{ih} x_i \quad h = 1, 2, \dots, m \quad (4)$$

式中, $x_i$ 是第 $i$ 个指标。

## 2 抢修时间估算模型

设飞机战伤抢修时间的影响因素个数为 $p$ ,通过实验、仿真等方式得到的样本数为 $n$ ,得到原始数据矩阵 $\mathbf{X} = (x_{ij})_{n \times p}$ ,其中, $x_{ij}$ 表示第 $i$ 个样本的第 $j$ 项因素的数值。抢修时间数据为 $\mathbf{Y} = (y_i)_{n \times 1}$ ,其中, $y_i$ 表示第 $i$ 个样本对应的实际抢修时间。

利用PCA方法,将 $p$ 个影响因素降为 $m$ ( $m < p$ )个主成分 $z_1, z_2, \dots, z_m$ 。建立多元线性回归模型<sup>[12-13]</sup>为

$$f = \omega_0 + \omega_1 z_1 + \omega_2 z_2 + \dots + \omega_m z_m \quad (5)$$

式中: $f$ 为抢修时间估计值; $\omega_0$ 为常数项; $\omega_1, \omega_2, \dots, \omega_m$ 为回归系数。

为了得到回归系数的最优值,利用遗传算法<sup>[11]</sup>对其进行优化。以抢修时间估计值与实际值间的相对误差最小来构建目标函数

$$\min \sum_{i=1}^n |y_i - f|/y_i \quad (6)$$

设置遗传算法的参数(初始种群数、交叉概率、变异概率、遗传代数及终止条件),得到优化的回归参数,代入式(5)即得到抢修时间的估算模型。

## 3 案例分析

为了验证模型的有效性,本文以文献[4]中数据为原始数据来估算抢修时间。抢修时间的影响因素<sup>[4]</sup>包括:抢修操作所需资源复杂性( $x_1$ )、抢修逻辑结构复杂性( $x_2$ )、抢修所需工作量复杂性( $x_3$ )、抢修所需决策知识层次复杂性( $x_4$ )、抢修技术高的人员数( $x_5$ )、抢修技术较高的人员数( $x_6$ )、抢修技术一般的人员数( $x_7$ )、抢修技术较差的人员数( $x_8$ )、工具合适的数量( $x_9$ )、工具基本合适的数量( $x_{10}$ )、工具不合适的数量( $x_{11}$ )、环境恶劣程度( $x_{12}$ )等。

利用式(1)对原始数据进行规范化处理,利用式(2)计算相关矩阵,进而得到相关矩阵的特征值及对应的特征向量,计算方差贡献率及累计贡献率,如表1所示。

表 1 特征值、方差贡献率及累积贡献率

Table 1 Eigenvalue, variance contribution rate and cumulative contribution rate

特征值	方差贡献率/%	累积贡献率/%
3.1957	26.6309	26.6309
2.1170	17.6413	44.2722
1.5721	13.1008	57.3730
1.5427	12.8558	70.2289
1.2610	10.5085	80.7374
0.8684	7.2365	87.9739
0.5634	4.6952	92.6692
0.4279	3.5656	96.2348
0.2082	1.7347	97.9695
0.1731	1.4423	99.4118
0.0587	0.4893	99.9011
0.0119	0.0989	100.0000

由表 1 可知,前 6 项的累积方差贡献率达 87.97%,表明前 6 项主成分基本上已经包含了原始因素所提供的绝大部分信息,用它们来衡量抢修时间是合理的。构造  $z_1, z_2, z_3, z_4, z_5, z_6$  6 个主成分,主成分的特征向量数值大小即反映指标对抢修时间的贡献率大小,如表 2 所示。

表 2 主成分的特征向量

Table 2 Feature vectors of principal components

	$z_1$	$z_2$	$z_3$	$z_4$	$z_5$	$z_6$
$x_1$	0.4737	-0.1886	0.2751	0.0552	-0.1183	0.0005
$x_2$	0.2317	0.2107	0.4464	-0.3785	0.0412	0.1051
$x_3$	0.5367	-0.0321	0.1439	0.0118	0.0219	0.1828
$x_4$	0.4749	0.0912	-0.1184	0.1820	0.0707	0.2304
$x_5$	-0.1678	-0.0292	0.3011	-0.0052	0.6753	0.3080
$x_6$	-0.1566	-0.1059	0.3722	0.4573	-0.4021	0.1305
$x_7$	-0.0206	-0.2935	0.1976	-0.3895	-0.4665	0.1163
$x_8$	0.3113	0.0274	-0.3812	-0.3376	0.0875	-0.2785
$x_9$	-0.0105	-0.4545	0.3742	-0.0252	0.3425	-0.4381
$x_{10}$	-0.0531	-0.4120	-0.2905	-0.0674	0.0912	0.6822
$x_{11}$	0.0369	0.5871	0.1501	0.2362	0.0383	0.0920
$x_{12}$	0.2324	-0.3013	-0.1731	0.5335	0.0936	-0.1835

由表 2 可知:第一主成分与抢修所需工作量复杂性( $x_3$ )、抢修所需决策知识层次复杂性( $x_4$ )和抢修操作所需资源复杂性( $x_1$ )密切相关,解释为抢修任务的复杂性,从工作量、所需资源和决策层面综合衡量任务的复杂程度;第二主成分中贡献最大的是工具不合适的数量( $x_{11}$ ),与工具合适的数量( $x_9$ )、工具基本合适的数量( $x_{10}$ )负相关,解释为抢修工具的整体合适程度;第三主成分中贡献最大的是抢修逻辑结构复杂性( $x_2$ )、工具合适的数量( $x_9$ )和抢修技术较高的人员数( $x_6$ ),解释为人员、工具和任务的协调性因素,反映技术等级较高的人员,利用合适的工具开展复杂抢修逻辑结构工作的协调程度;第四主成分中贡献最大的是环境恶劣程度( $x_{12}$ ),其次是抢修技术较高的人员数

( $x_6$ ),解释为人员工作环境因素;第五主成分中贡献最大的是抢修技术高的人员数( $x_5$ ),其次是工具合适的数量( $x_9$ ),解释为技术等级因素;第六主成分中贡献最大的是工具基本合适的数量( $x_{10}$ )和抢修技术高的人员数( $x_5$ ),解释为工具合适因素。第四、五、六主成分均与人员的技术等级紧密相关,是抢修人员与工具、工作环境相互配合的综合反映。结合表 1 中所列出的各主成分的方差贡献率可知,在开展抢修时间估算中,抢修任务的复杂性( $z_1$ )所包含的信息是最多的,起主要作用,且随着主成分的方差贡献率减小,它们对抢修时间的影响也逐渐减弱。

利用式(4)可得 6 个主成分与原始指标间的关系:

$$\begin{aligned} z_1 &= 0.4737x_1 + 0.2317x_2 + \dots + 0.0369x_{11} + 0.2324x_{12}, \\ z_2 &= -0.1886x_1 + 0.2107x_2 + \dots + 0.5871x_{11} - 0.3013x_{12}, \\ z_3 &= 0.2751x_1 + 0.4464x_2 + \dots + 0.1501x_{11} - 0.1731x_{12}, \\ z_4 &= 0.0552x_1 - 0.3785x_2 + \dots + 0.2362x_{11} + 0.5335x_{12}, \\ z_5 &= -0.1183x_1 + 0.0412x_2 + \dots + 0.0383x_{11} + 0.0936x_{12}, \\ z_6 &= 0.0005x_1 + 0.1051x_2 + \dots + 0.0920x_{11} - 0.1835x_{12}. \end{aligned}$$

以 6 个主成分为变量,利用式(5)建立抢修时间的多元线性回归模型。通过遗传算法求解回归系数,参数设置如下:初始种群数为 40,遗传代数为 1000,交叉概率 0.8,变异概率 0.06。最终,得到的抢修时间估算的回归模型为: $f = -402.65487 + 54.47983z_1 - 21.27044z_2 + 10.00198z_3 - 10.38099z_4 - 0.21041z_5 + 13.36883z_6$ 。

为了检验模型的估计精度,分别计算了用最小二乘法回归和本文的基于 GA 的多元线性回归法得到的抢修时间估计值,结果如图 1 所示。

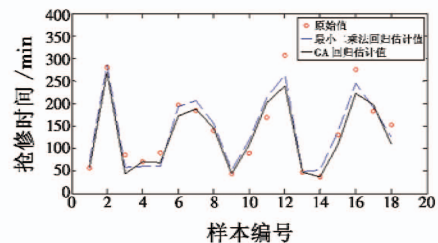


图 1 抢修时间的估计精度比较

Fig. 1 Evaluation accuracy of ABDR time

由结果可看出,虽然 PCA 去掉了部分数据信息,但基于主成分的多元线性回归模型可以较好地估计抢修时间。最小二乘法回归<sup>[12]</sup>和基于 GA 的回归平均相对误差分别为 0.17 和 0.1282,可见,运用 GA 对回归系数进行优化后,回归模型的估计误差更小,精度更高。

## 4 结束语

在估算飞机战伤抢修时间中,考虑的影响因素较多,本文利用主成分分析法对抢修时间影响指标进行降维处理,建立了多元线性回归模型,并用遗传算法对

回归系数进行优化计算,最后通过案例对模型进行了验证。与传统的线性回归模型相比,本文的模型不仅减少了回归模型的输入量,降低了模型的复杂性,且大大降低了误差,提高了模型的精度。然而,由于案例中的数据量较小,PCA 的优势并没有凸显,但随着飞机战伤抢修数据的不断增加及影响因素的不断完善,PCA 和 GA 相结合的优势会越来越明显。同时,由于 PCA 会去除部分信息,给抢修时间的估计造成一定的误差。

### 参考文献

- [1] 王少华,郑毅,吕会强,等. 战场抢修决策的研究现状与展望[J]. 兵器装备工程学报,2017,38(9):130-135.
- [2] 刘晓山,秦宇飞. 飞机战伤抢修保障效能及评估研究[J]. 航空维修与工程,2014(2):58-61.
- [3] 蔡纪伟,贾云献,孙晓,等. 蚁群算法在战损装备抢修任务指派中的应用研究[J]. 数学的实践与认识,2012,42(19):160-165.
- [4] 尤志锋,石全,熊飞. 基于加权支持向量回归的抢修时间估计模型[J]. 现代防御技术,2014,42(4):160-166.
- [5] 石全,李建平,刘祥凯. 复杂系统抢修时间仿真模型研究[J]. 计算机仿真,2002,19(1):93-96.
- [6] 杨新广,陈云翔,刘硕,等. 航空装备抢修时间估算方法[J]. 火力与指挥控制,2012,37(2):96-98.
- [7] 刘飞,杨江平. 雷达装备战场抢修等级量化决策研究[J]. 装备指挥技术学院学报,2009,20(4):116-119.
- [8] SHOKRI S, SADEGHI M T, MARVAST M A, et al. Soft sensor design for hydrodesulfurization process using support vector regression based on WT and PCA [J]. Journal of Central South University of Technology, 2015, 22(2): 511-521.
- [9] 包磊,雷进生,刘琼. 基于主成分分析法的地铁系统安全综合评价[J]. 三峡大学学报(自然科学版),2011,33(4):57-59.
- [10] 李艳军,张建,曹愈远,等. 基于模糊信息粒化优化 SVM 的航空发动机性能趋势预测[J]. 航空动力学报,2017,32(12):3022-3030.
- [11] 崔文斌,叶志锋,彭利方. 基于信息融合遗传算法的航空发动机气路故障诊断[J]. 航空动力学报,2015,30(5):1275-1280.
- [12] 王瑞朝,王远达,刘刚,等. 一种军机备件需求量的多元回归解法[J]. 兵工自动化,2009,28(9):53-54.
- [13] 王卓健,饶学军,沈安慰,等. 部附件送修费用的偏最小二乘回归模型及敏感性分析[J]. 电光与控制,2012,19(10):102-105.
- (上接第 69 页)
- [4] 邹璐. 基于 AADL 的 ARINC653 平台上的大飞机机载软件代码生成技术研究[D]. 北京:北京航空航天大学,2013.
- [5] SAE Aerospace. AS5506/1;SAE architecture analysis and design language (AADL) ANNEX volume1 [S/OL]. [2006-06-05]. <http://www.sae.org/technical/standards/AS5506/1>.
- [6] SAE Aerospace. AS5506/2;SAE architecture analysis and design language (AADL) ANNEX volume2 [S/OL]. [2011-01-17]. <http://www.sae.org/technical/standards/AS5506/2>.
- [7] ARINC. Avionics application software standard interface-Part1;ARINC 65391 [S]. Aeronautical Radio Inc., 2010.
- [8] 孙毅刚,李浩,赵珍,等. 基于 AADL 的综合模块化航空电子通信调度分析与建模[J]. 北京航空航天大学学报,2015,47(4):497-507.
- [9] 周德新,牛亚月,崔海青. IMA 分区间通信的一致性测试模型[J]. 计算机仿真,2018,35(2):43-48.
- [10] 隗立超. 一种基于多处理器的 AADL 模型的可调度性分析方法:107391866A[P]. 2017.
- (上接第 93 页)
- [5] CHUM O, MATAS J. Matching with PROSAC-progressive sample consensus[J]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005:220-226.
- [6] 贾彤. 基于改进 RANSAC 算法的室内视觉定位方法研究[D]. 哈尔滨:哈尔滨工业大学,2017.
- [7] 宋卫艳. RANSAC 算法及其在遥感图像处理中的应用[D]. 北京:华北电力大学,2011.
- [8] 黄于. 基于 FPGA 的 SURF 特征点检测算法的实现[D]. 哈尔滨:哈尔滨工业大学,2016.
- [9] 周德福. 无人机电子稳像的 FPGA 设计与实现[D]. 深圳:深圳大学,2017.
- [10] 张玄. 基于图像的 RANSAC 算法与硬件实现研究[D]. 西安:西安电子科技大学,2018.