

doi: 10.7690/bgzdh.2021.06.016

面向任务携行航材品种确定和消耗预测的特征选择分析

宋传洲¹, 王瑞奇¹, 孙岩¹, 刘战¹, 孟凡娜²

(1. 海军航空大学岸防兵学院, 山东 烟台 264001; 2. 中国人民解放军 91033 部队, 山东 青岛 266000)

摘要: 为进一步提高对携行航材品种确定和消耗数量预测的准确性, 针对非线性多影响因素的品种分类和消耗预测模型中的特征选择问题进行研究。对影响携行航材需求的因素进行分析, 建立三级特征体系, 并针对品种确定和消耗预测提取相应的特征集合; 采用 XGboost、灰色关联度分析 (grey relation analysis, GRA)、决策试验和评价实验法 (decision making trial and evaluation laboratory, DEMATEL) 等方法对各影响特征进行重要性排序和相关性分析; 综合运用定性和定量分析方法筛选特征; 分别建立可用于品种确定和数量预测的精简版特征集合。该研究可为后续提高携行航材品种确定和预测的准确率和运算效率提供参考。

关键词: 面向任务; 携行航材; 特征选择; XGboost; 灰色关联; DEMATEL

中图分类号: TJ02 **文献标志码:** A

Task-oriented Feature Selection and Analysis of Carried Aircraft Spare Parts Variety Determination and Consumption Prediction

Song Chuazhou¹, Wang Ruiqi¹, Sun Yan¹, Liu Zhan¹, Meng Fanna²(1. School of Coast Guard, Navy Aviation University, Yantai 264001, China;
2. No. 91033 Unit of PLA, Qingdao 266000, China)

Abstract: In order to further improve the accuracy of the variety determination and consumption quantity prediction of carried aircraft spare parts, the problem of feature selection in the nonlinear multi-influence factors variety classification and consumption prediction model was studied. The factors affecting the demand for spare parts of carrying aircraft are analyzed, a 3-level feature system is established, and corresponding feature sets are extracted according to the variety determination and consumption prediction. XGboost, grey relation analysis (GRA), decision making trial and evaluation laboratory (DEMATEL) were used to rank the importance degree and analyze the correlation of each influence feature. Synthetically use qualitative and quantitative analysis methods to select features. The feature sets of the simplified edition which can be used for species determination and quantity prediction are established respectively. This study can provide reference for improving the accuracy and computing efficiency of the identification and prediction of spare parts for carrying aircraft.

Keywords: mission oriented; carried aircraft spare parts; feature selection; XGboost; grey correlation; DEMATEL

0 引言

飞行器在执行不同任务时, 会在任务转场、远程异地、驻航母或舰船护航等情况下, 根据任务需要, 携带一定种类和数量的航材进行携行保障^[1]。航材的需求受任务等级、周围环境、工作强度、储存空间和维修能力等因素影响, 且存在很多制约条件对各类因素发生作用, 一定程度上各因素之间也会相互影响, 是典型的非线性多因素影响问题^[2]。以往研究中, 提取特征建立特征指标体系时, 为了提高模型的运算效率及分类或预测的准确率, 往往会对特征进行筛选或合并降维处理, 较常见的有模糊综合评判法^[3]、Topsis 评价法^[4]、灰色关联分析 (GRA)^[5]、AHP 评价法^[6]、灰色局势决策法^[7]、粗

糙集理论^[8-9]、主成分分析法^[10]、逻辑决断法、备件属性决策法、可靠性方法、价值工程法和相似系统法等, 但很多方法只做了单一的定性分析或定量分析, 其局限性导致无法全面分析影响因素的各类特征。笔者通过分析不同任务中影响航材消耗的各种因素来建立特征指标体系, 并从特征体系中分别选出用于携行航材品种确定和数量预测的特征集合; 同时运用机器学习 XGboost^[11]算法与态势分析的灰色关联度分析^[12]2 种定量分析法, 以及定性分析见长的决策试验和评价实验法 (DEMATEL)^[13], 对比 3 种方法的优缺点, 根据实际情况综合分析筛选特征, 建立精简版特征集合, 为后续品种确定和数量预测模型的高效应用做好特征选取准备工作。

收稿日期: 2021-02-22; 修回日期: 2021-03-21

作者简介: 宋传洲 (1989—), 男, 山东人, 硕士, 从事航空器材、仓储勤务、航材预测等技术研究。E-mail: 792868502@qq.com。

1 特征简化的综合设计

1.1 影响因素分析和特征提取

对任务中航材影响因素^[11]进行分析,按照全面

系统、科学简明等原则建立三级特征指标体系,具体如表 1 所示,根据需要从中分别提取用于品种确定和数量预测的特征集合。

表 1 影响携行航材需求的特征

一级特征	二级特征	三级特征
任务因素	任务类型	任务模式、任务需求
	任务等级	危害程度、保障等级、工作强度
	任务强度	飞机起落架次、飞机启动次数、任务飞行时间、历史任务消耗
	任务损伤	受损程度、损伤面积、损伤部位、作战损伤、敌我力量、杀伤概率
内在因素	可靠性水平	失效率(故障率)、可靠度、平均失效(故障)间隔时间、装机寿命、损耗敏感性(耗损敏感度)、故障关联性、环境适应能力
	可更换性	可拆卸能力、可更换单元
	可获得性	通用性、携行空间、可储运性
	可维修性	限时可维修、维修水平、故障可预见
环境因素	自然环境	温度、湿度、盐度、气压、生物作用、气候、地理位置、环境严酷度
	诱发环境	振动、冲击、核辐射、电磁辐射、大气污染
	机场环境	起飞降落的场地铺面、辅助设备等级、周围布局
人为因素	飞行员能力	任务飞行时间、任务飞行次数、操作失误次数、技术等级、心理素质、身体素质
	维修员能力	任务参与次数、维修失误次数、技术等级、心理素质、身体素质
	管理员能力	任务参与次数、管理失误次数、技术等级、心理素质、身体素质
	保障资源	维修参考资料、维修设备、维修工具、维修场地
经济因素	购置费	购买费、周转费
	储运费	储存费、运送费、邮递费、保养费、包装费
	维修费	各类维修费用
其他因素	不明因素	影响航材消耗的偶发未知或尚不清楚原因

1.2 3 种重要性排序方法

综合采用 XGboost、GRA、DEMATEL 3 种方法对特征重要性进行排序。XGboost 算法又称极端梯度提升树,是以 Boosting 的集成学习为基本思想,集成多个不同弱分类器的增强分类器,通过对特征数据持续迭代分割,得到分类节点的增益积累,从而进行重要排序,其特点在于对人的经验依赖较低,计算效率高,可以较为明晰地得到数据间的内在规律。GRA 是灰色理论的重要内容,使用 GRA 方法计算参考变量和比较特征之间的关联度,进而对特征进行分析和筛选,是否携带用于品种确定的参考变量,而航材的历史任务消耗量则是用于消耗预测的参考变量,其特点在于航材消耗的影响因素复杂多变且相互作用,是典型的信息不完全的灰色系统。DEMATEL 又称决策实验室分析法,采用图论和矩阵理论,充分发挥某领域专家的知识 and 经验,最终得到反映特征影响的重要程度的中心度,以及因果方向的原因度。前 2 种方法均为定量分析方法,后 1 种方法是定性分析方法。分别用 3 种方法进行特征重要性排序,再对其平均值的综合排名进行细致深入的分析,既可发挥定性分析的经验优势从实际应用效果来评价和排序,又可通过定量分析的客观评价降低对经验的依赖,找到被忽略的内在规律,

从而对最终的特征集合简化做出更加科学的决策。

1.2.1 XGboost 重要性排序原理

XGboost 算法是集成的梯度提升树模型,近年来在许多领域的应用中都获得良好成效^[14-16],其原理和计算公式如下:

XGboost 模型在处理数据集时,每一层都会选出一个特征分割点,称作叶子节点。这些节点会被累加而获得增益,即分割特征的次数越多,分割后的树增益值就越大,被分割次数越多的特征对整棵树的增益贡献越大,可通过这个积累来说明该特征的重要程度;同理,每次分割特征时的平均增益越大,该特征也越重要,可根据其值进行重要排序。

分割过程中,用 $w(g_i, h_i)$ 表示每个叶子节点的权值, g_i 和 h_i 表示:

$$g_i = \partial_{y^{(t-1)}} l(y_i, \hat{y}^{(t-1)}); \quad (1)$$

$$h_i = \partial_{y^{(t-1)}}^2 l(y_i, \hat{y}^{(t-1)}). \quad (2)$$

式中: $l(y_i, \hat{y}_i)$ 代表训练误差,表示目标值 y_i 和预测值 \hat{y}_i 之间的差距。根据所有叶子节点的权值,每个特征的分割点增益 Gain。该增益可表示为分割后的总权值减去分割前叶子节点的总权值,分割前的权值来自叶子节点左子树总权值与右子树总权值之

和, 对于每一个分割点而言, 其表达式为

$$\text{Gain} = \sum_{\text{left}} w + \sum_{\text{right}} w - \sum_{\text{nosplit}} w. \quad (3)$$

XGBoost 对决策树进行梯度提升算法, 逐层减小前面生成的决策树损失, 并产生新树分支参与模型建构, 实现最终决策的优化, 因此模型随着时间每一次迭代后都会增加一棵树, 对这些树进行组合。

XGBoost 根据特征分裂的次数 FScore、特征平均增益值 AverageGain 取值来进行重要性排序。模型中 API Select From Model 输入自动选定的特征, 选择阈值 thresh 的重要性累加得到了取值后排序。

1.2.2 GRA 重要性排序原理

GRA 原理和计算公式如下:

1) 确定比较特征和参考变量参考变量: 表示为 $x_0 = [x_0(1), x_0(2), \dots, x_0(n)]$, 比较特征表示为 $x_i = [x_i(1), x_i(2), \dots, x_i(n)]$, 其中 i 表示特征表里的特征数, n 为样本数。

2) 计算 x_0 与 x_i 的关联系数, 记作 ξ_{0i} , 其公式如下:

$$\xi_{0i} = \frac{\min |x_0(k) - x_i(k)| + \rho \max |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \max |x_0(k) - x_i(k)|}. \quad (4)$$

式中 ρ 为分辨系数, $\rho \in (0, 1)$, 通常取值 0.5, 用来调整比较公式表示的环境影响程度。

3) 计算灰色关联度并排序, 记作 γ_{0i} , 称作 x_i 对 x_0 的灰色关联度, 其公式如下:

$$\gamma_{0i} = \frac{1}{n} \sum_{k=1}^n \xi_{0i}(k). \quad (5)$$

由式 (5) 可得到关联度序列 $\mathfrak{R} = (\gamma_{01}, \gamma_{02}, \gamma_{03}, \dots, \gamma_{0i})$, 对序列值进行排序, 可直观地对特征集合进行分析和筛选。

1.2.3 DEMATEL 重要性排序原理

DEMATEL^[13]原理如下:

1) 建立矩阵 Y 来表示影响特征量之间的直接影响。

$$Y = (a_{ij})_{n \times n} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}. \quad (6)$$

式中: a_{ij} 表示特征 i 对特征 j 的影响程度, 当 $i=j$ 时, $a_{ij}=0$ 。

2) 对矩阵 Y 进行标准化处理得到标准化影响

矩阵 X 。

$$X = (x_{ij})_{n \times n} = \frac{Y}{\max \sum_{j=1}^n a_{ij}}. \quad (7)$$

式中 $\max \sum_{j=1}^n a_{ij}$ 为矩阵的各行影响特征量之和的最大值。

3) 矩阵变换后得到综合影响矩阵 T 。

$$T = (t_{ij})_{n \times n} = \lim (X + X^2 + \dots + X^k),$$

$$T = (t_{ij})_{n \times n} = X(I - X)^{-1}. \quad (8)$$

式中 I 为单位矩阵。

4) 分别求出单个特征的影响度 r_i 、被影响度 c_i 、中心度 β_i 、原因度 α_i 。

$$r_i = \sum_{j=1}^n t_{ij}, c_i = \sum_{j=1}^n t_{ji}, \quad i = 1, 2, \dots, n;$$

$$\beta_i = r_i + c_i, \alpha_i = r_i - c_i, \quad i = 1, 2, \dots, n. \quad (9)$$

式中 β_i 反映第 i 个特征在特征集合中的重要程度。

1.2.4 特征精简流程

特征精简流程如图 1 所示。

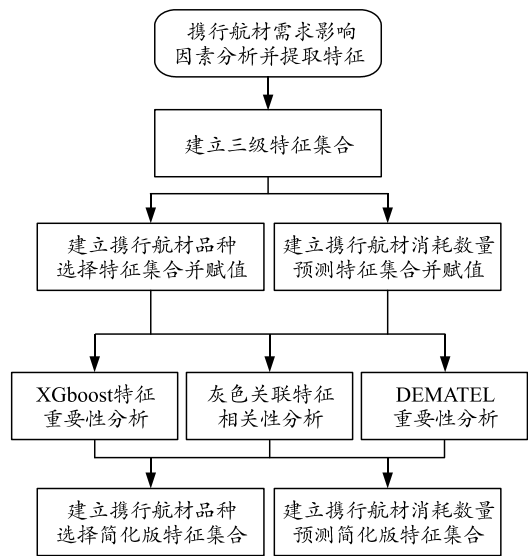


图 1 简化特征集合流程

2 实验分析

2.1 特征提取

通过查询飞行记录、维修记录、文献, 并咨询专家后, 归纳出自身因素、环境因素、任务因素、人员因素、经济因素和其他因素等 6 方面的影响因素。从表 1 的特征体系中分别选出用于携行航材品种确定和数量预测的特征及其取值, 如表 2、表 3 所示。

表 2 用于品种确定的特征取值

代号	特征名称	取值	代号	特征名称	取值
A1	危害程度	1 一级至 7 七级	A12	故障可预见	1 可预见、2 不可预见
A2	任务需求	1 很低至 5 很高	A13	可更换单元	1 可分解、2 不可分解
A3	保障等级	1 低级至 4 最高级	A14	维修水平	1 一等、2 二等、3 三等、4 四等
A4	故障率	1 超极低至 9 超级高	A15	可拆卸能力	1 很难至 5 很容易
A5	限时可维修	1 限时可维修、2 限时不可维修	A16	购置费	1 很低至 5 超高
A6	使用强度	1 很低至 5 很高	A17	储运费	1 低、2 中等、3 高
A7	环境严酷度	1 一级至 6 六级	A18	通用性	1 标准件、2 通用件、3 专用件
A8	飞行员水平	1 合格、2 中级、3 高级、4 特级	A19	携行空间	1 空间允许携带、2 空间不允许携带
A9	任务损伤	1 轻度、2 中度、3 重度、4 严重	A20	可储运性	1 器件可储运、2 器件不可储运
A10	历史消耗	1 低、2 中、3 高、4 很高	D	是否携行	1 携行、0 不携行
A11	装机寿命周期	1 初级期、2 成熟期、3 衰退期			
备注	表中数据均为分类变量，通过采取问卷打分、查询记录、划分区间等形式获得				

表 3 用于数量预测的特征取值

编号	特征	数值类型	取值	编号	特征	数值类型	取值
T1	保障等级	分类变量	1 普通至 5 特级	T9	人员因素	综合量化值	合并飞行员、维修员、管理人员、维修资料等因素
T2	飞机起落架次	次数	任务中飞机起落的总架次	T10	平均温度	摄氏度	查阅资料获得
T3	任务飞行时间	小时	任务中飞行的总时间	T11	异常温度占比	百分比	根据记录计算获得
T4	历史任务消耗数量	个数	任务期间某航材消耗总数量	T12	平均湿度	湿度值	查阅资料获得
T5	可靠性水平	数值	这里主要用故障率来表示。根据元器件性质，拟合指数、正态、威布尔等分布规律计算获得 ^[17] 。	T13	异常湿度占比	百分比	根据记录计算获得
T6	装机寿命周期	分类变量	1 初级期、2 成熟期、3 衰退期	T14	平均盐度	分类变量	查阅资料获得
T7	耗损敏感度	分类变量	1 很低至 5 较高	T15	地理环境因素	量化值	查阅资料 ^[17] 获得
T8	可获得性	分类变量	1 难、2 中等、3 容易	T16	机场环境	分类变量	1 恶劣、2 差、3 普通、4 良好

2.2 数据来源

现有 n 种航材，分别记作 $H\{H_1, H_2, H_3, \dots, H_n\}$ ； m 种用于品种确定的特征，记作 $A\{A_1, A_2, A_3, \dots, A_m\}$ ； c 种用于携行数量预测的特征，记作 $T\{T_1, T_2, T_3, \dots, T_c\}$ ；其中 $m=20$ ， $c=16$ 。笔者选取 $n=400$ 种可能

携带的航材，收集 20 次任务的 8 000 个样本数据，包含样本特征共 160 000 个，据统计，其中有 150 种是 20 次任务均需携行的航材，250 种根据任务不同并不是都携带，从确定携带的航材中选取重要航材 G，对用于其携行数量预测的特征进行分析选择。部分数据如表 4、表 5 所示。

表 4 用于品种确定的部分航材特征数据

编号	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20	D
H01	6	4	4	3	2	4	5	3	1	1	2	1	1	3	3	5	2	3	1	1	1
H02	7	5	4	7	2	4	4	3	2	2	2	1	1	3	3	4	1	3	1	1	1
H03	6	5	4	4	2	4	4	3	2	2	3	1	1	3	3	5	1	3	1	1	1
	...																				
H398	2	3	3	1	2	4	3	4	2	1	1	1	2	4	3	2	3	2	1	2	0
H399	2	3	3	1	2	4	3	4	2	1	2	1	2	4	3	2	2	2	2	2	0
H400	1	2	2	4	1	4	4	4	2	3	2	2	2	4	3	1	1	1	1	1	0

表 5 G 航材部分任务特征数据

任务	T1	T2	T3	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16	T4
1	4	1 000	2 400	0.016	3	3	3	0.05	30	0.26	60	0.30	2	3.0	3	32
2	3	720	1 900	0.023	3	3	2	0.07	26	0.05	50	0.10	2	3.0	3	24
3	4	950	2 100	0.072	1	3	3	0.14	22	0.03	49	0.05	2	3.0	3	28
	...															
118	5	850	1 850	0.035	2	3	3	0.26	25	0.05	60	0.09	2	3.0	1	17
119	4	940	2 010	0.052	1	4	3	0.09	28	0.10	63	0.11	3	3.5	1	21
120	4	110	2 550	0.016	3	4	3	0.14	28	0.13	59	0.12	3	3.5	2	43

2.3 用于品种确定的特征重要性分析

2.3.1 基于 XGboost 的特征重要性排序

基于 XGboost 的特征重要性排序如图 2 所示。

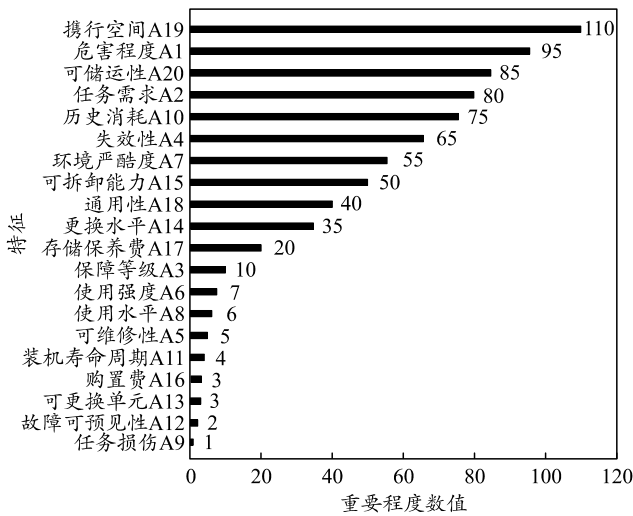


图 2 用于品种确定特征重要性程度对比

表 7 用于品种确定特征之间的影响关系参数

特征	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
<i>r</i>	0.838 1	1.182 2	2.325 6	2.046 5	1.277 9	1.489 8	2.526 8	1.696 1	1.753 9	2.261 3
<i>c</i>	2.360 7	2.070 4	2.088 5	2.519 2	1.661 4	1.473 9	0	0.365 4	2.002 3	2.211 3
<i>n</i>	-1.522 6	-0.888 3	0.237 1	-0.472 7	-0.383 5	0.015 8	2.526 8	1.330 7	-0.248 5	0.049 9
<i>m</i>	3.198 9	3.252 6	4.414 1	4.565 7	2.939 2	2.963 7	2.526 8	2.061 5	3.756 2	4.472 7

特征	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20
<i>r</i>	1.774 8	1.045 0	0.607 3	0.702 2	0.702 3	0	0	0.227 8	0	0
<i>c</i>	0.983 3	1.124 1	1.160 4	0.921 8	0.921 8	0	0	0.272 9	0.159 9	0.159 9
<i>n</i>	0.791 4	-0.079 1	-0.553 1	-0.219 6	-0.219 5	0	0	-0.045 1	-0.159 9	-0.159 9
<i>m</i>	2.758 2	2.169 1	1.767 7	1.624 1	1.624 1	0	0	0.500 7	0.159 9	0.159 9

从上表中心度可以看出，A3、A10、A4、A2、A9、A1 等排在最前，A16、A17、A18、A19、A20 排在最后；从原因度可以看出，A7 环境严酷度值最高，这与任务中环境因素对航材需求发生较大影响实际相符，其次是 A8 飞行员的操作水平对航材消耗影响较大；从影响度和被影响度来看，A4、A3 产生的影响度和被影响度相当，A16、A17 无数值与 A19、A20 影响度无数值说明大部分特征与 A16、A17 没有任何关联，与 A19、A20 关系不大，可看出该方法的局限性。由于对品种确定影响特征间采取单纯的定性分析，并未体现这些特征用来品种确定的最终判定是否携带的目的，但对特征间相互关系的定性分析可为最后的特征筛选提供重要参考。

2.4 用于消耗预测的特征重要性分析

2.4.1 XGboost 消耗预测特征重要性排序

XGboost 消耗预测特征重要性排序如图 3 所示。

2.3.2 基于 GRA 品种确定特征重要性排序

使用 GRA 方法计算特征重要性排序的取值，如表 6 所示。

表 6 品种确定特征关联程度

A2	A1	A19	A4	A10	A20	A7
0.903 2	0.886 1	0.872 4	0.850 7	0.849 7	0.831 3	0.811 9
A6	A4	A8	A5	A14	A15	A9
0.809 0	0.802 3	0.798 9	0.777 1	0.760 8	0.741 0	0.735 7
A12	A3	A11	A17	A13	A16	
0.716 9	0.693 2	0.667 3	0.630 1	0.604 2	0.592 6	

2.3.3 基于 DEMATEL 品种确定特征重要性排序

取值(0,1,2,3)分别对应(无影响、弱影响、中度影响、强影响)表现因素之间的直接影响程度。邀请 15 名航材研究领域专家、20 名勤务保障人员，根据经验打分，判断各特征间的影响关系，最终取分数总平均值四舍五入后取整，得到直接影响矩阵。进一步求得影响度 *r*、被影响度 *c*、原因度 *n*、中心度 *m* 如表 7 所示。

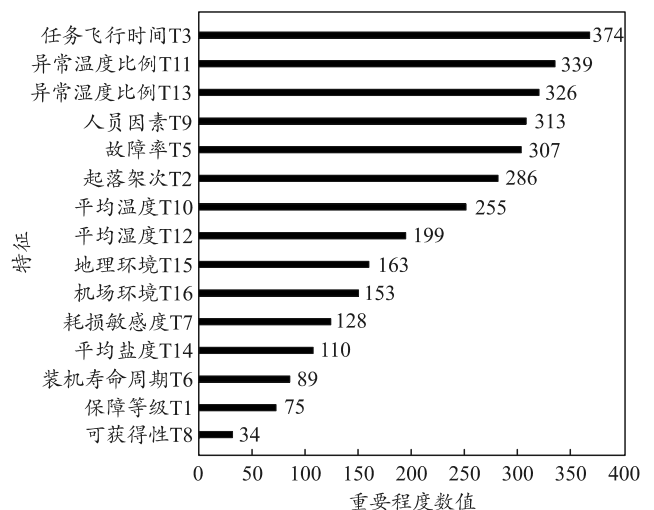


图 3 用于数量预测特征重要性程度对比

2.4.2 GRA 的消耗预测特征重要性排序

影响携行航材消耗需求的参考变量是历史任务消耗量，同品种确定特征排序使用方法一样，求得用于消耗预测特征关联程度如表 8 所示。

表 8 消耗预测特征关联程度

T3	T11	T13	T5	T2	T9	T15	T10
0.919 5	0.895 7	0.889 7	0.851 1	0.826 9	0.791 1	0.787 3	0.771 3
T12	T14	T7	T6	T16	T1	T8	T4
0.769 7	0.761 5	0.751 0	0.731 6	0.698 4	0.636 7	0.591 3	被关联特征

2.4.3 DEMATEL 消耗预测特征重要性排序

同品种确定特征使用方法一样，求得消耗预测直接影响矩阵和特征之间的影响关系参数如表 9 所示。

表 9 影响消耗预测特征之间的影响关系参数

特征	T1	T2	T3	T4	T5	T6	T7	T8
r	0.719 3	1.000 0	1.205 7	0.561 7	0.476 3	0.402 4	0.242 1	0.082 2
c	1.098 2	0.073 5	0.130 0	3.428 4	2.760 7	1.445 0	1.215 0	0.273 8
n	-0.378 9	0.926 6	1.075 7	-2.866 7	-2.284 4	-1.042 6	-0.972 9	-0.191 6
m	1.817 5	1.073 5	1.335 7	3.990 1	3.237 0	1.847 5	1.457 1	0.356 0
特征	T9	T10	T11	T12	T13	T14	T15	T16
r	0.832 3	1.011 4	1.011 4	1.011 4	1.011 4	0.905 5	1.723 3	0.602 3
c	0.396 3	0.245 2	0.245 2	0.245 2	0.245 2	0.179 7	0.706 9	0.110 4
n	0.436 0	0.766 2	0.766 2	0.766 1	0.766 2	0.725 9	1.016 3	0.491 8
m	1.228 6	1.256 6	1.256 6	1.256 6	1.256 6	1.085 3	2.430 2	0.712 7

从上表可以看出，T4 历史消耗的被影响度最高，原因度最低；T3 的原因度最高、T15 的原因度较高；T10、T11、T12、T13 等温度和湿度环境影响特征各类数值相同，因为定性分析时假设温湿度影响程度相同；从中心度可以看出，T5、T15、T6、T1 等定性分析结果较重要；T8、T2、T10、T9 分析结果排在最后。

2.5 综合分析并建立精简版特征列表

2.5.1 品种确定综合分析和特征简化

品种确定特征重要性排序如表 10 所示。

表 10 品种确定特征重要性排序

特征	XGboost	GRA	DEMATEL	平均值
A1	2	2	6	3
A2	4	1	5	3
A3	12	16	3	10
A4	6	4	1	4
A5	15	11	8	11
A6	13	8	7	9
A7	7	7	10	8
A8	14	10	12	7
A9	20	14	4	13
A10	5	6	2	4
A11	16	17	9	14
A12	19	15	11	15
A13	18	19	13	17
A14	10	12	14	12
A15	8	13	14	11
A16	17	20	19	19
A17	11	18	19	16
A18	9	9	18	12
A19	1	3	16	7
A20	3	5	16	8

从上表可知，通过 XGboost 算法和 GRA 方法对数据进行重要性和相关性分析的结果大部分特征排序相差不大，A3 保障等级、A19 携行空间、A20 可储运性前 2 种方法与 DEMATEL 方法的分析排序相差较大；因为 DEMATEL 方法相关性分析是两两比较，难以最终确定是否携带。定性分析时，A3 保障等级是对航材需求影响比较重要的因素，保障

等级低、保障能力差，航材需求量会增多，且会影响飞机任务完成。根据实际数据来看，由于不管执行什么任务保障等级都比较高，取值相对稳定，对最后判定是否携带航材帮助不大，特别是基于数据分析的判定。A9 任务损伤，由于具有对抗性、损耗性很强的任务较为少见，任务损伤定性分析靠前，但其数据对整体判断引起的变化不大，故排在最后；由于 A19 携行空间、A20 可储运性 2 个因素在航材需求分析中与其他因素几乎没有相关性，故定性分析方法将其排在最后，而就数据分析而言，是否携带某航材，是否能够携带和储运是重要的判定特征。综合 3 种方法对其他特征的判定相对一致，故取平均值四舍五入后可综合判定，根据需要去掉 A9、A11、A12、A13、A16、A17，将特征集合简化。

2.5.2 消耗预测综合分析和特征简化

消耗预测特征重要性排序如表 11 所示。

表 11 消耗预测特征重要性排序

特征	XGboost	GRA	DEMATEL	平均值
T1	15	15	5	12
T2	7	6	15	9
T3	2	2	8	4
T4	1	1	1	1
T5	6	5	2	4
T6	14	13	4	10
T7	12	12	7	10
T8	16	16	16	16
T9	5	7	13	8
T10	8	9	9	9
T11	3	3	9	5
T12	9	10	9	9
T13	4	4	9	6
T14	13	11	14	13
T15	10	8	3	7
T16	11	14	6	10

通过上表可知，与品种确定特征重要性排序方法出现的情况类似，XGboost 算法和灰色关联法分析结果相差不大。在 T1、T2、T3、T6、T9 等特征

上,前 2 个方法与 DEMATEL 方法分析结果相差较大;其原因也颇为相似,T1 保障等级、T6 寿命周期定性分析虽然重要,但在定量分析中因随任务改变的数值变化不大,重要性排位靠后。在对 T4 这一特征的判断上 3 种方法完全一致。需要补充的是,实际的航材消耗预测中,T2 飞机起落架次是非常重要的特征因素,特别是对于轮胎、支架、一些传感器等航材消耗影响较大。DEMATEL 方法在定性分析时优势并不明显。3 种方法对其他特征的判定相对一致,故取平均值四舍五入后可综合判定,根据需要去掉 T1、T6、T7、T8、T14、T16,将特征集合简化。

2.5.3 对比分析

对品种确定和消耗预测 2 个特征集合进行对比分析:用于品种确定特征提取 20 个,可简化为 14 个;用于消耗预测特征提取 16 个,除去非重要特征和用来预测的历史任务消耗数量可简化为 9 个;经济因素在品种确定时会有影响,但并不重要,在携行数量预测时经济影响可忽略;是否能顺利完成任务和是否有足够的携带能力是判断携行品种的重要考量,而任务中飞机工作时间与次数和外在诸多环境因素,对航材消耗多少有不同程度的影响。

3 结束语

通过对面向任务的影响携行航材需求的各因素进行分析,建立三级特征体系,提取用于携行航材品种确定和消耗预测的特征集合,提出联合使用定量分析方法 XGboost、GRA 以及定性分析 DEMATEL 方法对特征集合进行特征重要性排序,对排序结果取平均值后结合实际需要筛选特征,分别建立精简版特征集合,可为后续提高携行航材品种确定和预测的准确率以及运算效率提供参考。由于对特征相关性分析不够,后续研究中,若特征间相关性较强,可考虑进一步简化特征集合。

参考文献:

[1] 郭峰,王德心. 航母舰载机航材携行品种及数量确定方法研究[J]. 军事运筹与系统工程, 2015, 29(2): 38-42, 59.
[2] 宋传洲,王瑞奇,刘战. 面向任务携行航材消耗影响因素研究[J]. 航空维修与工程, 2020(8): 68-71.

[3] 季嘉伟,孙国文,罗佳伟. 基于模糊综合评判的电源车战场抢修备件品种研究[J]. 装备制造技术, 2017(7): 214-216.
[4] 金国栋,卢利斌,叶庆. 无人机携行备件品种确定方法[J]. 火力与指挥控制, 2008, 33(10): 144-148.
[5] 张强,刘士通,王开勇. 模糊灰色关联分析在战储维修器材品种优化中的应用[J]. 军事交通学院学报, 2016, 18(11): 63-67.
[6] 陈鹏. L 公司服务备件库存管理优化研究[D]. 秦皇岛:燕山大学, 2019: 10-12.
[7] 周一鸣,王茜,荣鹏辉. 基于灰色局势决策的航空弹药保障装备备件品种确定方法[J]. 舰船电子对抗, 2017, 40(5): 56-60.
[8] 赵建忠,徐廷学,刘勇. 基于粗糙集和 BP 神经网络的导弹备件消耗预测[J]. 兵工自动化, 2012, 31(7): 66-71.
[9] 韩俊骞,何亚群,李继军. 基于混合属性值的航材库存品种确定方法[J]. 数学的实践与认识, 2019, 49(19): 139-144.
[10] 张闯,刘福胜,单志伟,等. 一体化装备保障备件品种确定方法[J]. 装甲兵工程学院学报, 2018, 32(3): 30-34.
[11] CHEN T Q, GUESTRIN C. XGBoost: A Scalable Tree Boosting System[C]//Proceedings of the 22nd ACM SIG-KDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA, USA: Association for Computing Machinery New York, 2016: 785-794.
[12] 高鹏,邢国平,孙德翔. 灰色关联支持向量机在备件库存消耗预测中的应用[J]. 电光与控制, 2012, 19(3): 100-105.
[13] 林原,战仁军,吴虎胜. 基于 DEMATEL-RS 的军事装备采购决策规则获取[J]. 统计与信息论坛, 2020, 35(5): 87-93.
[14] 宋传洲,陈育良,杨宜霖. 基于 XGboost 算法面向任务携行航材品种确定方法[J]. 兵工自动化, 2021, 40(2): 75-80.
[15] LI C, ZHENG X, YANG Z, et al. Predicting Short-Term Electricity Demand by Combining the Advantages of ARMA and XGBoost in Fog Computing Environmen[J]. Wireless Communications and Mobile Computing, 2018(8): 1-18.
[16] GERTZ M, GROBE-BUTENUTH K, JUNGE W, et al. Using The XGboost Algorithm to Classify Neck And Leg Activity Sensor Data Using on-Farm Health Recordings for Locomotor-Associated Diseases[J]. Computers and Electronics in Agriculture, 2020, 173(6): 40-45.
[17] 陈博,徐常凯,任佳成. 基于灰色神经网络的携行航材消耗预测[J]. 指挥信息系统与技术, 2018, 9(5): 86-90.