

文章编号: 1671-6612 (2023) 02-175-05

考虑不同状态划分方法的马尔科夫链空调车备件需求预测

王炳忠 毛宇 胡新生 胡海洋
(海军航空大学青岛校区 青岛 266041)

【摘要】 针对空调车备件需求随机性和波动性的特点, 利用灰色马尔科夫组合模型对空调车备件消耗量进行了预测。以状态数量和区间长度为基准, 提出四种马尔科夫链状态划分方式, 得到预测曲线, 并拟合出与实际曲线的相关系数进行比较。结果表明: 第 7, 11, 12, 14 月的灰色预测误差均在 5% 以下, 可考虑在状态划分时加入不变状态; 从状态数量来看, 五状态对应的相关系数整体上高于四状态, 原因是五状态在四状态基础上, 除了能预测备件需求上升和下降的强弱状态, 还可以对上述灰色预测误差较小的月份不作修正, 从而提高预测精度; 从区间长度来看, 五状态下的等值划分和等量划分, 相关系数分别为 0.98 和 0.85, 由于等量划分时的各区间长度较为失衡, 最终使得概率矩阵在第 13 个月状态预测出现偏差。

【关键词】 空调车; 备件; 需求预测; 灰色马尔科夫模型; 状态划分
中图分类号 N941.5 文献标识码 A

Research on Forecast of Air-conditioned Car Spare Parts by Markov Chain Considering Different State Division Methods

Wang Bingzhong Mao Yu Hu Xinsheng Hu Haiyang
(Naval Aviation University Qingdao Campus, Qingdao, 266041)

【Abstract】 Aiming at the randomness and volatility of air-conditioned car spare parts, the gray Marcow combined model is used to predict the consumption of air-conditioning car spare parts. Based on the quantity and interval length, four Malcov chain state division methods are proposed to obtain a predictive curve and compare the correlation coefficients of the actual curve. The results show that: the gray prediction errors in the 7th, 11th, 12th and 14th are below 5%, and you can consider adding an uncomvited state when the state division is divided. From the perspective of the number of states, the correlation coefficient corresponding to the five states is higher than the four states as a whole. The reason is that on the basis of the four states, in addition to predicting the strong or weak state of the demand for spare parts, it can also not modify the month with less gray prediction errors, thereby improving the prediction accuracy. From the perspective of interval length, the equivalent and equivalent division in the five states, the correlation coefficients are 0.98 and 0.85, respectively. Due to the unbalanced length of each interval of equivalent division, the probability matrix predicts the deviation of the state prediction in the 13th month.

【Keywords】 Air-conditioned vehicle; Spare parts; Demand prediction; Gray Markov model; State division

0 引言

在进行飞机地面维护和检查时, 空调车为飞机

设备舱提供冷、热、通风, 以此调节机舱内空气温度和湿度, 确保机载电子设备的正常运行。空调车

备件数量直接影响着保障能力,备件不足会导致空调车战备完好率下降,而备件储备过多既增加管理难度,也浪费经费。因此,准确预测备件需求对于保证空调车保障能力、维持合理备件库存具备重要意义。

备件需求预测方法主要有三大类:基于人工智能、可靠性和历史数据^[1-3]。基于人工智能的预测方法主观性和随机性较强,且需要大量统计样本;基于可靠性主要研究损坏机理,与备件的寿命分布有很大关系;基于历史数据即解析分析法,在预测备件需求时,运用较为广泛的是灰色马尔科夫组合模型。虽然灰色系统理论的 GM(1,1)模型能够挖掘出少量样本数据中的发展趋势,但对于波动明显的的数据预测精度较低,而马尔可夫链预测模型能够很好地弥补这一缺陷,实现优势互补。

文献[4-7]利用灰色马尔科夫模型及其改进模型对雷达备件的月度消耗量、油泵装机后发生故障的时间点、装甲车辆负载需求功率以及 4S 店备件需求进行了研究,预测精度较高,由此可见灰色马尔科夫模型已较为成熟地应用在备件需求、装备故障预测等领域。在马尔科夫模型建立过程中,状态划分对于预测精度是极为重要的,因此本文重点研究马尔科夫链状态划分方式对于预测精度的影响。

1 模型建立^[8]

1.1 GM(1,1)建模

(1) 设 $X^{(0)}$ 为非负原始数据序列, $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)\}$, $X^{(1)}$ 是 $X^{(0)}$ 的一次累加生成序列, $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), x^{(1)}(3), \dots, x^{(1)}(n)\}$, 一次累加方法为:

$$x^{(1)}(t) = \sum_{i=1}^t x^{(0)}(i); t = 1, 2, 3, \dots, n \quad (1)$$

(2) 构造一阶常微分方程逼近累加生成序列,并用最小二乘法求得系统时间响应方程如下:

$$\hat{x}^{(1)}(t+1) = (x^{(1)}(1) - \frac{u}{a})e^{-at} + \frac{u}{a} \quad (2)$$

其中, a 、 u 为待定参数。

(3) 再累减还原,即可得到原始序列的时间响应方程(预测曲线方程):

$$\hat{x}^{(0)}(t+1) = \hat{x}^{(1)}(t+1) - \hat{x}^{(1)}(t) \quad (3)$$

(4) 记 $A = (a \ u)^T$, 则 $A = (B^T B)^{-1} B^T Y$

$$B = \begin{pmatrix} -0.5(x^{(1)}(1) + x^{(1)}(2)) & 1 \\ -0.5(x^{(1)}(2) + x^{(1)}(3)) & 1 \\ \dots & \dots \\ -0.5(x^{(1)}(n-1) + x^{(1)}(n)) & 1 \end{pmatrix}, Y_n = \begin{pmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \dots \\ x^{(0)}(n) \end{pmatrix}$$

1.2 建立灰色—马尔科夫组合模型

(1) 状态划分

根据灰色 GM(1,1) 模型求出预测值 $\hat{x}^{(0)}(t)$, $t = 1, 2, \dots, n$ 。 $\hat{x}^{(0)}(t)$ 曲线较好地反映了原始数据的总体变化趋势,以趋势曲线 $\hat{x}^{(0)}(t)$ 为基准,利用实际值与预测值的比值来划分状态,这样就形成了与趋势曲线平行的若干条形区域,每一个区域构成了一个状态,就将一个符合马尔科夫链特点的非平稳随机序列 $\hat{x}^{(0)}(t)$ 划分成了 n 个状态。

(2) 计算状态转移概率矩阵

状态转移概率为:

$$P_{ij}(m) = \frac{M_{ij}(m)}{M_i}, i = 1, 2, \dots, n \quad (4)$$

式中: M_i 为系统处于状态 i 的原始数据样本数; $M_{ij}(m)$ 为状态 i 经 m 步转移到状态 j 的原始数据样本数。

(3) 计算预测值

若预测系统在某时刻处于 i 状态,则预测值为:

$$\hat{y}^{(t)} = x^{(0)}(t) \times \frac{(A_i + B_i)}{2}, i = 1, 2, \dots, n \quad (5)$$

式中: B_i 和 A_i 分别表示 i 状态的条形区域的上、下限。

2 案例应用

已知某场站某型空调车连续 15 个月的备件消耗数量,如表 1 所示。

表 1 某型空调车备件消耗数量

Table 1 Consumption of spare parts of a certain type of air-conditioned car

月份	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
消耗量	27	33	35	21	27	23	29	31	32	27	30	30	37	29	27

(1) 利用 GM(1,1)来预测备件消耗量

令 $X^{(0)} = \{27, 33, 35, 21, 27, 23, 29, 31, 32, 27, 30, 30, 37, 29, 27\}$, 根据式 (1) 和 (2), 计算得到 GM(1,1)模型的时间响应函数的参数 $a = -0.0059$, $u = 27.8738$, 将每个月份的备件消耗实际值与灰色预测值反映在图 1 中。图 1 缺少了第 1 个月的预测数据, 因为在灰色模型建立过程中, 第 1 个数据作为初始条件求解未知参数, 实际值与预测值相等。由这张图可以看出, 每个月备件需求数量都有波动, 但变化程度不大。预测曲线可以反映出备件需求量呈平缓上升的趋势, 这与实际情况是符合的。为了反映灰色预测误差, 我们将各月实际值与灰色预测值的比值列出, 如表 2 所示。

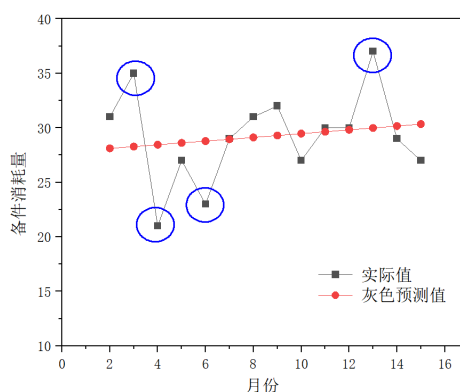


图 1 各月备件消耗实际值与灰色预测值

Fig.1 The actual value and gray forecast value of spare parts in each month

表 2 各月实际值与灰色预测值的比值

Table 2 The ratio of actual values to gray prediction values for each month

月份	2	3	4	5	6	7	8	9	10	11	12	13	14	15
比值	1.103	1.238	0.738	0.944	0.799	1.002	1.065	1.092	0.916	1.012	1.006	1.234	0.961	0.890

结合表 2 中的数据, 可以看出图 1 中所示的第 3, 4, 6, 13 月份的灰色预测误差较大, 其中第 3, 13 月份处于强上升状态, 即预测值要比实际值大许多; 而第 4, 6 月份处于强下降状态, 灰色预测值比实际值小很多。因此在对各月份的灰色预测值进行修正时, 需要对强上升以及强下降月份的状态作出准确预测, 这样才能进行有效修正。而第 7, 11, 12, 14 月份的实际值与预测值比值为 1.002, 1.012, 1.006, 0.961, 这 4 个月份的预测误差均在 5% 以下, 因此可以考虑不进行修正。

(2) 状态划分

在构建马尔科夫模型时, 划分状态的数量以及区间长度对于预测结果有很大影响, 因此本文设置

了四种划分方式, 如表 3 所示。其中四状态分别为: 强下降、弱下降、弱上升、强上升, 而五状态则是多出一种保持不变的状态。划分出来的区间有等值和等量之分, 等值区间即各个区间长度相等, 等量区间是通过保持各个状态数量均衡产生的, 四种划分方式的具体参数设置如表 4 所示。

表 3 划分参数

Table 3 divide parameters

划分方式	状态	区间
方式一	四状态	等值
方式二	四状态	等量
方式三	五状态	等值
方式四	五状态	等量

表 4 四种划分方式具体参数

Table 4 Specific parameters of the four division methods

划分方式	状态 (个数)	划分区间	划分方式	状态 (个数)	划分区间
方式一	1 (2)	[0.700, 0.850)	方式二	1' (3)	[0.700, 0.895)
	2 (7)	[0.850, 1.000)		2' (4)	[0.895, 1.005)
	3 (3)	[1.000, 1.150)		3' (4)	[1.005, 1.100)
	4 (2)	[1.150, 1.300]		4' (3)	[1.100, 1.300]
方式三	1'' (2)	[0.700, 0.820)	方式四	1''' (3)	[0.700, 0.900)
	2'' (2)	[0.820, 0.940)		2''' (3)	[0.900, 1.000)
	3'' (5)	[0.940, 1.060)		3''' (2)	[1.000, 1.010)
	4'' (3)	[1.060, 1.180)		4''' (3)	[1.010, 1.100)
	5'' (2)	[1.180, 1.300]		5''' (3)	[1.100, 1.300]

以划分方式三为例（下同），各月的状态如表 5 所示。

表 5 各月状态

Table 5 The state of each month

月份	2	3	4	5	6	7	8	9	10	11	12	13	14	15
状态	4''	5''	1''	3''	1''	3''	4''	4''	2''	3''	3''	5''	3''	2''

(3) 构造转移概率矩阵

根据式 (3)，得到各步转移概率矩阵为：

$$P_1 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 0 & 1/3 & 0 & 1/3 & 1/3 \\ 1/2 & 0 & 1/2 & 0 & 0 \end{bmatrix}$$

$$P_2 = \begin{bmatrix} 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 1/4 & 1/4 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \end{bmatrix}$$

$$P_3 = \begin{bmatrix} 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1/2 & 1/4 & 1/4 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$P_4 = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1/3 & 1/3 & 1/3 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

(4) 编制预测表

因为是 4 步转移概率矩阵，所以从第 6 个月开始预测。通过第 2, 3, 4, 5 月份对第 6 个月的备件消耗量进行预测，如表 6 所示。

表 6 第 6 个月灰色马尔科夫状态预测

Table 6 The 6th month of gray Marcow status forecast

起始月	5	4	3	2	合计
起始状态	3	1	5	4	
转移步数	1	2	3	4	
状态 1''	1/5	1/2	1	1/3	31/30
状态 2''	1/5	0	0	0	1/5
状态 3''	1/5	0	0	1/3	8/15
状态 4''	1/5	1/2	0	0	7/10
状态 5''	1/5	0	0	1/3	8/15

由表 6 中的合计栏可以看出，状态 1'' 对应的概率最大，所以第 6 个月备件消耗量最有可能处于状态 1''，其灰色预测值为 28.78，由式 (5) 可得，第 6 个月马尔科夫修正值为 21.87，而实际值为 23，误差 4.90%。接着我们对于剩余月份进行灰色马尔科夫状态预测，如表 7 所示。

表 7 第 6-15 月份灰色马尔科夫状态预测

Table 7 Gray Marcow status forecast for 6-15 months

月份	6	7	8	9	10	11	12	13	14	15
实际值	23	29	31	32	27	30	30	37	29	27
实际状态	1''	3''	4''	4''	2''	3''	3''	5''	3''	2''
预测状态	1''	3''	4''	4''	2''	3''	3''	5''	3''	2''
修正值	21.87	28.95	32.61	32.81	25.93	29.64	29.81	37.18	30.16	26.70
误差/%	4.90	0.17	5.21	2.52	3.97	1.21	0.63	0.49	4.01	1.11

(5) 四种划分方式比较

以上我们以方式三为例，对灰色预测值进行了马尔科夫修正，接着我们采用其余三种划分方式进行预测，四种方式的实际值与预测值比较如图 2 所示，四幅图中相关系数的大小表明预测曲线与实际曲线的贴合程度，贴合度从大到小依次为方式三

(0.98)，方式四 (0.85)，方式二 (0.73)，方式一 (0.60)。其中，方式三的贴合程度最高，原因是该方式下的预测状态与实际状态完全一致，因此具有非常高的贴合率。而其他方式下的预测状态与实际状态偏差越大，预测曲线贴合程度越低。

观察图 2 上两幅图可以发现，方式一、二由于

采用了四状态划分, 缺少不变状态, 其预测状态很大程度上偏离实际状态, 概率矩阵失效, 导致预测偏差较大。

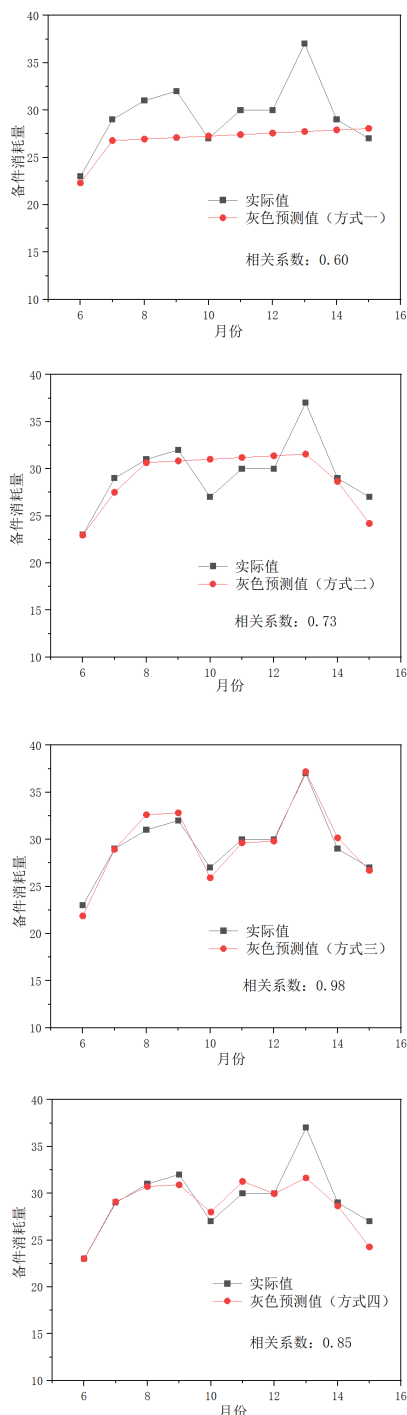


图 2 四种划分方式下实际值与预测值比较

Fig.2 Comparison of actual values and prediction values in the four division methods

而方式四在第 13 个月误差很大, 我们表示出其预测表, 如表 8 所示。可以看出第 13 个月处于

状态 4''' 和 5''' 的概率非常接近, 状态 4''' 的概率略大, 而实际状态为状态 5''' , 这是因为方式四追求各状态数量均衡, 使得区间长短不一, 区间长度分别为 0.2, 0.1, 0.01, 0.09, 0.2。相比之下, 方式三的等值划分, 所得到的转移概率矩阵更加准确。

表 8 方式四第 13 个月灰色马尔科夫状态预测

Table 8 The 13th month of gray Marcov's status

prediction in the fourth way					
起始月	12	11	10	9	合计
起始状态	3'''	4'''	2'''	4'''	
转移步数	1	2	3	4	
状态 1''	0	0	0	1/3	1/3
状态 2''	0	1/3	0	0	1/3
状态 3''	1/2	0	0	1/3	5/6
状态 4''	1/2	1/3	1/2	0	4/3
状态 5''	0	1/3	1/2	1/3	7/6

3 结论

本文通过灰色马尔科夫组合模型对于空调车备件消耗量进行了预测, 并比较了四种马尔科夫链状态划分方法对于预测精度的影响, 结论是:

(1) 在确定马尔科夫状态数量之前, 应对灰色预测值进行误差分析, 得出是否需要加入不变的状态; 如果灰色预测结果中有部分数据与实际误差较小 (5%以下), 则有必要加入不变状态。

(2) 在五状态划分时, 采用等值区间的预测曲线具有更好的贴合度。如果一味追求各状态数量均衡, 会导致各区间长度不一, 最终使得概率矩阵状态预测不准确, 得到的预测结果产生较大偏差。

(3) 第四种划分方式下, 第 13 个月的预测状态处于 4''' 和 5''' 的概率非常接近, 本文采用概率偏大的状态 4''' 进行误差修正, 如果采取加权计算, 会使这一数据的误差有所减小。因此对于数值较小的数据预测, 采用加权法可进一步减小预测误差, 同时后期的研究方向是进一步完善马尔科夫链状态划分方法以及对于特殊数据的处理, 在保证状态转移概率矩阵准确的基础上, 才能确保预测精度。

参考文献:

[1] 龙灏, 向静文, 晋旭博. 备件需求预测方法研究[J]. 工业控制计算机, 2021, 34(5): 68-70.

(下转第 201 页)