

基于故障数据的商用车发动机及底盘系统可靠性评估*

RELIABILITY EVALUATION OF COMMERCIAL VEHICLE ENGINE AND CHASSIS SYSTEM BASED ON FAULT DATA

姜 成**¹ 冯金芝^{1,2} 房启岭³ 赵礼辉^{1,2} 张东东***^{1,2}

(1. 上海理工大学 机械工程学院, 上海 200093)

(2. 机械工业汽车机械零部件强度与可靠性评价重点实验室, 上海 200093)

(3. 福田汽车股份有限公司时代事业部技术研究院, 北京 102206)

JIANG Cheng¹ FENG JinZhi^{1,2} FANG QiLing³ ZHAO LiHui^{1,2} ZHANG DongDong^{1,2}

(1. School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

(2. CMIF Key Lab for Automotive Strength & Reliability Evaluation, Shanghai 200093, China)

(3. Time Division Technical Research Institute, Foton Motor Co., Ltd., Beijing 102206, China)

摘要 发动机及底盘系统可靠性是影响商用车性能的关键因素,准确、全面地评估其可靠性对指导商用车可靠性的改进提升具有重要意义。基于故障大数据,对某商用车制造企业旗下四款商用车发动机及底盘系统可靠性进行评估。收集2021年度四款商用车所有故障数据进行初步统计和分析,随机抽样形成分析样本。根据U检验测试样本与总体数据分布差异性,拟合发动机及底盘系统故障间隔时间概率密度分布函数并推断该分布规律满足威布尔分布。结合传统中位秩法与核密度估计法相互验证计算系统分布参数,采用K-S(Kolmogorov-Smirnov)假设检验验证故障间隔时间样本数据服从威布尔分布。计算可靠度、故障率以及平均故障间隔时间等可靠性评价指标,得出车辆行驶500天后,四款商用车的可靠度分别为0.002、0.003、0.007、 3×10^{-9} ;故障率分别为0.021、0.016、0.015、0.063;平均故障间隔时间分别为146.088天、129.492天、149.755天、70.897天。结合发动机及底盘系统可靠性评估结果与故障内容提出优化方案,为商用车系统可靠性提升工作提供参考。

关键词 可靠性评估 U检验 中位秩法 核密度估计法 威布尔分布 K-S假设检验

中图分类号 U463.3

Abstract The reliability of engine and chassis system is the key factor affecting the performance of commercial vehicles. Accurate and comprehensive evaluation of its reliability is of great significance to guide the improvement of commercial vehicle reliability. Based on big data of faults, a reliability evaluation was conducted for the engine and chassis of four commercial vehicles within a specific manufacturing enterprise. All the fault data of the four commercial vehicles in 2021 were collected for preliminary statistical analysis, and a random sample was drawn to form an analysis sample. Based on the U-test to test the difference between the sample and the overall data distribution, the probability density distribution function of the engine and chassis system fault interval was fitted, and it was inferred that the distribution law satisfied the Weibull distribution. The traditional median rank method and kernel density estimation method were used to mutually verify and calculate the distribution parameters of the system. The Kolmogorov-Smirnov (K-S) hypothesis test was used to verify that the sample data of the fault interval time followed the Weibull distribution. The reliability evaluation indexes such as reliability, failure rate and mean time between failures were calculated. After 500 days of driving, the reliability of the four commercial vehicles were 0.002, 0.003, 0.007 and 3×10^{-9} . The fault rates were 0.021, 0.016, 0.015 and 0.063. The average time between fault were 146.088, 129.492, 149.755 and 70.897 days. Based on the reliability evaluation results and fault contents of engine and chassis system, the optimization scheme was proposed to provide reference for the reliability improvement of the commercial vehicle system.

Key words Reliability evaluation; U-test; Median rank; Kernel density estimation; Weibull distribution; K-S hypothesis test

Corresponding author: ZHANG DongDong, E-mail: dongdongzhang@usst.edu.cn

The project supported by the Shanghai CMIF Key Lab for Automotive Strength & Reliability Evaluation (No.H2022304042).

Manuscript received 20230421, in revised form 20230609.

*20230421收到初稿,20230609收到修改稿。上海机械工业汽车机械零部件强度与可靠性评价重点实验室项目(H2022304042)资助。

**姜 成,男,1998年生,江苏南通人,汉族,上海理工大学机械工程学院硕士研究生,主要研究方向为车辆大数据模型搭建及平台开发。

***张东东(通信作者),男,1986年生,山西晋城人,汉族,上海理工大学机械工程学院副教授,博士,硕士研究生导师,主要研究方向为电驱动系统大数据可靠性评价与设计。

0 引言

由于商用车用途的特殊性,需长时间在复杂恶劣的工况下高强度运行,车辆零部件和系统在承受各种多变苛刻的载荷下,极易遭受损坏和磨损,导致故障或事故的发生。根据国内外商用车相关故障数据统计调查研究,发现发动机及底盘系统是故障率较高的两个系统,对商用车可靠性和安全性有着重要影响^[1-2]。因此,为了保证货物和人员的安全运输,对商用车底盘和发动机进行可靠性评估尤为必要。

传统可靠性评估是通过对系统或设备的运行或试验数据进行统计分析,来评估其在特定工作环境下的可靠性水平,以确定产品的故障率、失效模式及其影响因素等,从而指导产品设计、改进和维护。瞿佳佳等^[3]结合医疗设备故障数据集与 Logistic 分布函数,提出一种医疗设备可靠性分析方法。朱荣^[4]基于初次故障数据,提出了一套客车高压电气系统故障分布规律研究方案。庞欢等^[5]通过汽车质保数据,搭建了一个汽车可靠性评估及决策分析系统,实现汽车故障统计分析。张荧驿等^[6]利用动车组转向架运行数据并结合可靠性模型,对其中部件的维修优化进行建议。刘伟等^[7]在传统可靠性分析的基础上,结合美国陆军器材系统分析活动(Army Materiel Systems Analysis Activity, AMSAA)模型及 Laplace 分析法,验证了某受电弓影响的列车可靠性水平变化趋势。从这些研究中可以发现,传统可靠性评估方法成熟、模型简单、适用范围广,但通常受限于数据收集和分析难度,其总体数据规模不足,导致抽样样本存在偏差而不具备代表性,从而影响评估结果精度,无法全面反映出产品的可靠性水平。

随着数据采集技术及计算能力的提高,基于大数据进行的现代可靠性评估逐渐成为主流趋势。王伟等^[8]利用大数据技术结合故障清单对某轨道列车进行可靠性分析,发现了高频故障的发生原因。赵礼辉等^[9]构造用户运行大数据下的电驱系统的工况特征参数,运用聚类算法完成了对该系统的可靠性分析研究。大数据技术驱动下的现代可靠性评估优势在于数据规模大,且覆盖面广,使抽样样本能更准确地描述总体特征;结合机器学习等数据挖掘算法,深入分析数据潜在的规律和特征,实现更优的数据分布拟合,提高可靠性评估结果的准确性。

针对传统可靠性评估方法的缺陷,本文提出一种基于大数据技术的现代可靠性评估方法。以某商用车制造企业旗下四个品牌的商用车发动机及底盘系统故障大数据为基础进行抽样分析,采用 U 检验法对样本与总体数据分布差异性进行先验,确定最优样本概率分布;对比传统中位秩法和现代核密度估计法两

种统计方法并确定分布参数;结合可靠性评价指标与具体故障内容,为提高和优化商用车发动机及底盘故障系统可靠性水平提供改进建议。

1 商用车发动机及底盘系统故障数据收集与处理

1.1 故障数据来源与类型

本文可靠性评估数据来源于各地区的所有商用车的 T-BOX (Telematics Box) 反馈,相比传统的可靠性实验数据具有两方面的优势。一方面,故障数据包含既全面又有价值的信息,充分涵盖了商用车在实际使用中的故障情况。另一方面,故障数据由大量的用户数据统计整理而得,不需要进行道路试验的数据采集和处理,可以更有效地支持可靠性分析。

以某商用车制造企业旗下四个品牌的商用车发动机及底盘系统为研究对象,收集 2021 年度所有售后故障数据用于可靠性评估,涉及的故障内容包括故障名称、发动机型号、销售日期、故障日期、车辆代号、故障系统。以某辆商用车故障内容明细为例,其具体形式如表 1 所示。

表 1 某商用车故障内容明细示例

Tab. 1 Example of details fault content of a commercial vehicle

品牌	发动机型号	故障名称	销售日期	故障日期	车辆代号	故障系统
Brand	Engine type	Fault name	Sale date	Fault date	Code	Fault system
A	XX,XX	尿素浓度高 High urea concentration	2021-01-27	2021-02-03	XXXX	后处理系统 Post-processing system

1.2 故障数据前处理

首先,对收集的 2021 年度商用车发动机及底盘故障数据进行数据清洗及整理汇总。然后,根据可靠性计算原则,对相同日期下同一辆车发生的相同故障进行合并,避免对同一故障事件进行重复计数而导致故障率和可靠度被高估,确保可靠性评估结果的准确性。最后,删除未知原因及外部原因导致的故障内容。这里的未知原因故障包括待定故障以及未定义故障;外部原因故障包括车辆外接设备故障。

通过车辆的销售日期和故障日期计算故障间隔时间,按照故障间隔时间对故障数据进行排序并统计故障系统数。四个品牌商用车故障车数及系统初步统计结果如表 2 所示。

由表 2 可知,发动机系统包括冷却系统、配气系统、燃油系统、进排气系统、润滑系统、启动系统、配气系统及后处理系统,底盘系统包括制动系统及传动系统。根据表 2 所给数据,将上述四个品牌商用车发动机及底盘系统故障数据汇总,如表 3 所示。

表 2 商用车故障车数及其故障系统数据汇总表

Tab. 2 Data summary of the number of broken down commercial vehicles and their failure systems

故障系统 Fault system	品牌 A Brand A	品牌 B Brand B	品牌 C Brand C	品牌 D Brand D
传动系统 Transmission system	37 243	32	314	14
发动机系统 Engine system	89 552	1 372	631	287
冷却系统 Cooling system	1 339	213	134	5
后处理系统 Post-processing system	170 616	13 213	13 441	5 387
润滑系统 Lubrication system	2 246	121	0	32
燃油系统 Fuel system	13 645	1 020	1 294	241
启动系统 Starting system	473	19	0	5
进排气系统 Intake and exhaust system	97 953	317	1 061	102
电源系统 Electrical power system	2 683	151	97	19
电器系统 Electric system	9 317	358	4 066	73
配气系统 Valve system	1 438	73	122	0
故障车数 Number of fault vehicles	12 654	3 355	8 269	1 948

表 3 商用车发动机及底盘系统故障数据汇总

Tab. 3 Summary of engine and chassis system fault data for commercial vehicles

品牌 Brand	故障总数 Total number of faults	发动机及底盘系统故障数 Number of engine and chassis system faults	发动机及底盘系统故障比例 Ratio of engine and chassis system faults/%
A	428 288	41 4025	96.6
B	16 931	16 342	96.5
C	21 160	16 942	80.1
D	6 165	6 033	97.8

基于前处理后得到的故障数据对商用车发动机及底盘系统故障进行随机抽样,抽取 1 000 条故障数据作为分析样本进行可靠性评估与分析工作。四个品牌商用车发动机及底盘系统对应的故障数据样本如表 4 所示。

表 4 商用车发动机及底盘系统故障数据样本

Tab. 4 Sample of engine and chassis system fault data for commercial vehicles

故障系统 Fault system	品牌 A Brand A	品牌 B Brand B	品牌 C Brand C	品牌 D Brand D
传动系统 Transmission system	83	2	19	1
发动机系统 Engine system	221	84	32	39
冷却系统 Cooling system	4	9	8	1
后处理系统 Post-processing system	422	798	788	886
润滑系统 Lubrication system	6	5	0	9
燃油系统 Fuel system	31	76	86	47
启动系统 Starting system	2	0	0	2
进排气系统 Intake and exhaust system	231	23	60	13
配气系统 Valve system	0	3	7	0
故障总数 Total number of faults	1 000			

2 商用车发动机及底盘系统可靠性模型搭建

传统的可靠性评估流程通常是在假设样本与总体分布一致的情况下直接进行参数估计。然而,在实际应用中,样本与总体数据分布可能存在显著差异,导致参数估计方法会低估或高估系统的可靠性。为避免这个问题,本文采取非参数统计方法中的 Mann-Whitney U 检验方法对样本和总体数据分布差异进行先验,消除抽样偏差因素的影响,确保样本数据信息的全面性和准确性。

已知总体和样本故障数据独立相关且集中可定序,采用 U 检验对样本和总体数据进行比较。该检验是通过两组数据秩和的大小关系来判断两组数据的中位数是否相等,如果两组数据的中位数相等,那么它们的秩和应该相等,同时证明两组数据分布相似。

对于独立相关的两组样本,U 检验经验公式^[10]为

$$U=\min(U+,U-)\tag{1}$$

式中, $U+$ 为正向秩和,即对于每个相同的对比值,选取秩次较小的样本进行排名,并将排名之和相加; $U-$ 为负向秩和,即对于每个相同的对比值,选取秩次较大的样本进行排名,并将排名之和相加。

本文样本与总体数据分布差异 U 检验流程如图 1 所示。

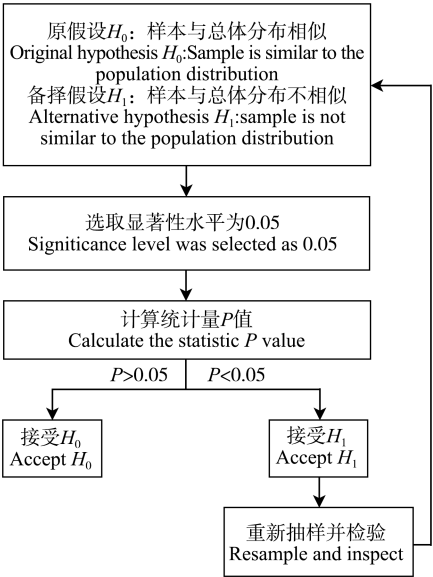


图 1 U 检验计算流程

Fig. 1 Calculation process of U-test

令原假设为样本与总体数据分布相似,选取显著性水平 α 为 0.05,对四个品牌商用车发动机及底盘系统的故障数据样本进行 U 检验并计算 P 值。 P 值一般是用来衡量差异是否足够显著的指标,若 P 值小于显著性水平 α ,则认为两组数据分布存在显著差异性,需重新抽取数据样本并进行检验。四个品牌商用车发

动机及底盘系统的故障样本检验结果如表5所示。

表5 系统故障样本U检验结果
Tab. 5 U-test result of system fault samples

品牌 Brand	A	B	C	D
P值 P value	0.479 1	0.493 2	0.708 2	0.892 7

由表5可知, P 值均大于0.05, 检验结果接受原假设, 四个品牌商用车发动机及底盘系统的故障数据样本和总体数据分布相似。

2.1 故障间隔时间概率密度分布函数

根据四个品牌商用车发动机及底盘系统的故障间隔时间样本数据, 直接拟合概率密度分布函数曲线。对商用车发动机及底盘系统故障间隔时间进行分组, 本文采用改进的Sturges样本分组公式^[11]以固定的组距进行分组, 为

$$K = 1 + 3.3 \lg \left(\frac{n^2}{100} \right) \quad (2)$$

式中, n 为样本总数; K 为分组数目。将 $n=1\,000$ 代入式(2), 得 $K=14.2$, 本次统计可取 $K=15$ 。将故障间隔时间分15组, 得四个品牌商用车发动机及底盘系统的故障间隔时间观测样本如表6所示。

表6 系统故障间隔时间观测样本
Tab. 6 Observation sample of the system fault interval

品牌 Brand	最小故障间隔时间 Minimum time between faults/d	最大故障间隔时间 Maximum time between faults /d	组距 Class interval /d
A	1	484	32.2
B	1	461	30.67
C	1	476	31.67
D	1	224	14.87

概率密度函数观测值计算式为

$$f(t) = \frac{n_i}{n\Delta t_i} \quad (3)$$

式中, n_i 为组间频次; Δt_i 为组距。组间频次 n_i 是指在一个组距内, 系统发生的故障次数。根据式(3)计算并拟合四个品牌商用车发动机及底盘系统的故障间隔时间概率密度函数曲线, 如图2所示。

由图2可知, 故障间隔时间概率密度函数曲线呈单峰型或单调下降型, 与威布尔分布的概率密度函数曲线趋势相似。因此, 本文推测四个品牌商用车发动机及底盘系统的故障间隔时间数据可能符合威布尔分布。但图2中直接拟合的概率密度曲线存在拐点, 均值和方差的估计值可能会发生偏差, 进而对分布参数估计和可靠性评估的精度产生影响。

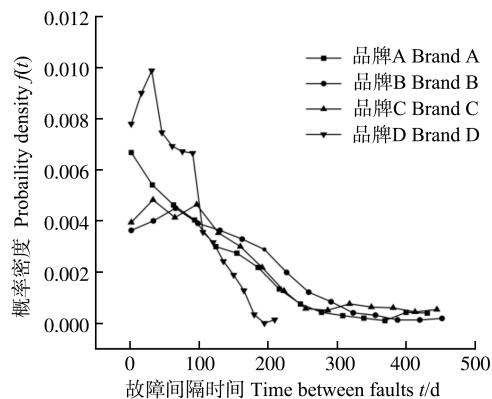


图2 系统故障间隔时间概率密度函数

Fig. 2 Probability density function of the system fault interval

2.2 故障间隔时间分布模型参数估计

假设该四个品牌商用车发动机及底盘系统的故障间隔时间数据符合威布尔分布, 采用最小二乘法^[12]拟合分布参数, 并对累积分布函数进行假设检验, 判断该数据是否满足威布尔分布。

两参数威布尔分布累积分布函数^[13]为

$$F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)^\beta} \quad (4)$$

式中, α 为尺度参数; β 为形状参数。

利用最小二乘法对两参数威布尔分布模型进行参数估计, 需先将威布尔分布模型转化为线性模型, 类似于 $Y = AX + B$ 的形式。由式(4)易得两参数威布尔分布累积函数变式为

$$\ln \frac{1}{1 - F(t)} = \left(\frac{t}{\alpha} \right)^\beta \quad (5)$$

再对两边同时取对数, 得

$$\ln \left[\ln \frac{1}{1 - F(t)} \right] = \beta \ln t - \beta \ln \alpha \quad (6)$$

令 $Y = \ln \left[\ln \frac{1}{1 - F(t)} \right]$, 令 $X = \ln t$, 可得两参数

威布尔分布线性变化模型为

$$Y = \beta X - \beta \ln \alpha \quad (7)$$

结合最小二乘法计算公式与两参数威布尔分布线性模型, 可求得其形状参数与尺度参数值, 计算流程为

$$\beta = \left(n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i \right) / \left[n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \quad (8)$$

$$\beta \ln \alpha = \frac{1}{n} \sum_{i=1}^n y_i - \frac{\beta}{n} \sum_{i=1}^n x_i \quad (9)$$

式中, x_i, y_i 分别为故障间隔时间数据集 t_i 和累积分布函数值数据集 $F(t_i)$ 经过威布尔线性变换后得到的数据集。

由于故障间隔时间数据集 t_i 已知, 累积分布函数集 $F(t_i)$ 未知, 可以选择非参数估计法先行计算累积分布函数估计集 $F'(t)$ 代替累积分布函数集 $F(t_i)$ 求得分

布参数。本文利用中位秩法和核密度估计法两种不同的非参数估计方法进行求解,相互印证分布参数的估计准确性。

2.2.1 中位秩法

中位秩法^[14]是目前工程上较为常用的传统经验分析法的一种改进形式,直接利用数据本身的分布特征来估计分布函数。利用中位秩法在 t_i 时刻求得的累积分布函数估计值 $F'(t_i)$ 为

$$F'(t_i) = \frac{i - 0.3}{n + 0.4} \quad (10)$$

式中, n 为样本总数; i 为样本序号。

将四个品牌商用车发动机及底盘系统的故障间隔时间数据集 t_i 和累积分布函数值数据集 $F'(t_i)$ 代入式(8)~式(10),得相关威布尔分布参数如表7所示。

表 7 中位秩法估计系统分布参数

Tab. 7 System distribution parameters under median rank estimation

品牌 Brand	形状参数 Shape parameters	尺度参数 Scale parameters
A	$\beta = 1.335\ 9$	$\alpha = 156.163\ 3$
B	$\beta = 1.163\ 6$	$\alpha = 121.729\ 7$
C	$\beta = 1.271\ 7$	$\alpha = 148.953\ 4$
D	$\beta = 1.392\ 2$	$\alpha = 73.841\ 1$

利用K-S假设检验方法检验当前样本数据是否符合威布尔分布。通过求经验累积分布函数估计值 $F'(t_i)$ 与假设分布的累积分布函数值 $F(t_i)$ 的最大差值 $D_{\max}$,并与临界 D 值比较,若 $D_{\max} < D$,则接受该显著性水平下的假设分布。

选取显著性水平 α 为0.05,由样本数量及显著性水平可知当前检验的临界 D 值,计算式为

$$D = \frac{1.63}{\sqrt{n}} \quad (11)$$

由式(11)可知,在当前显著性水平下的临界值 $D = 0.051\ 5$ 。

$D_{\max}$ 计算式为

$$D_{\max} = \max_{1 \leq i \leq n} |F(t) - F'(t)| \quad (12)$$

由分布参数及K-S假设检验可得四品牌威布尔累积分布函数 $F(t)$ 及其最大差值 $D_{\max}$,如表8所示。

表 8 系统累积分布函数最大差值

Tab. 8 Maximum difference between cumulative distribution functions of systems

品牌 Brand	最大差值 Maximum difference $D_{\max}$
A	0.042 6
B	0.041 8
C	0.048 2
D	0.039 9

由表8可知,在对应形状参数和尺度参数的条件下,其最大差值 $D_{\max}$ 均小于 D ,即四个品牌商用车发动机及底盘系统的故障间隔时间数据样本服从威布尔分布。

2.2.2 核密度估计法

核密度估计法^[15]是现代可靠性评估所用到的一种分布估计方法,由德国数学家帕罗教授在1962年提出的,也被称为“Parzen”窗口方法。其具体计算式为

$$\hat{f}(t) = \frac{1}{nh} \sum_{j=1}^n K\left(\frac{t - t_j}{h}\right) \quad (13)$$

式中, n 为样本总数; h 为带宽,是控制核密度估计平滑度的参数; K 为核函数,包括高斯内核、三角内核以及多项式内核等,用于对每个样本进行加权计算。核密度估计是通过在每个样本点处放置一个核函数,根据带宽的大小对相邻样本点的位置特性进行平滑处理来降低误差,提高可靠性评估的精度。

基于四个品牌商用车发动机及底盘系统的故障间隔时间样本数据,借助Sklearn机器学习算法库,交叉验证评估不同的核函数与带宽模型的性能,最终选取适用于威布尔分布数据集的高斯内核,同时调整最佳带宽参数并拟合绘制概率密度分布函数曲线。核密度估计与机器学习结合的具体实验流程如图3所示。其中,数据集为故障间隔时间数据样本;参数准备为不同类型核函数组合;网格搜索为指定核函数和带宽的参数范围;交叉验证为评估不同参数组合获得的模型性能。

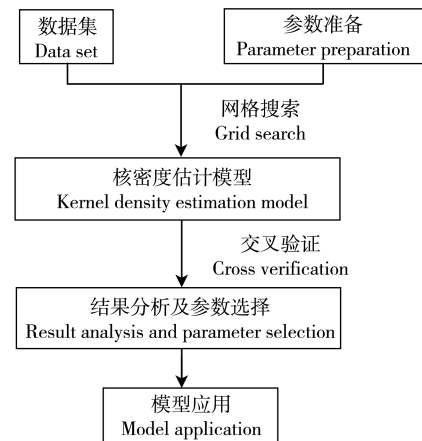


图 3 核密度估计与机器学习结合实验流程

Fig. 3 Combined experimental process of kernel density estimation and machine learning

四个品牌商用车发动机及底盘系统的故障间隔时间概率密度分布函数如图4所示。由图4可以看出,核密度估计拟合的概率密度函数曲线能够捕捉到数据的整体趋势和分布特征,同时消除了拐点的影响,不会出现过多的波动或尖峰,其曲线平滑度远优于图2直接拟合的密度曲线,且趋势符合威布尔分布。

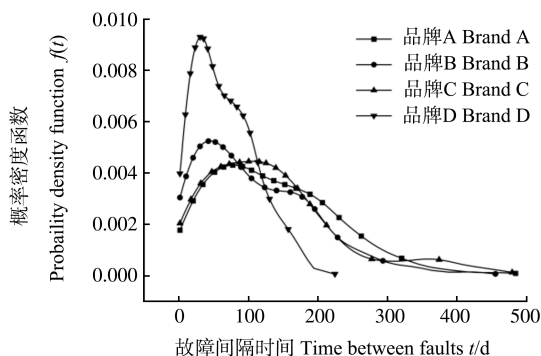


图4 系统核密度估计概率密度函数

Fig. 4 Kernel density estimation probability density function of systems

将四个品牌商用车发动机及底盘系统的故障间隔时间数据代入式(13)并进行积分,得到核密度估计下的累积分布函数估计值 $\hat{F}(t_i)$,代入式(8)~式(10),得到相关威布尔分布参数,如表9所示。

表9 核密度估计系统分布参数

Tab. 9 System distribution parameters under kernel density estimation

品牌 Brand	形状参数 Shape parameters	尺度参数 Scale parameters
A	$\beta = 1.658\ 2$	$\alpha = 163.436\ 6$
B	$\beta = 1.386\ 1$	$\alpha = 141.850\ 3$
C	$\beta = 1.460\ 4$	$\alpha = 165.322\ 2$
D	$\beta = 1.613\ 8$	$\alpha = 79.136\ 9$

由分布参数及K-S假设检验可得核密度估计四品牌的威布尔累积分布函数 $\hat{F}(t)$ 及其最大差值 $D_{\max}$,如表10所示。

表10 系统累积分布函数最大差值

Tab. 10 Maximum difference between cumulative distribution functions of systems

品牌 Brand	最大差值 Maximum difference $D_{\max}$
A	0.033 5
B	0.049 7
C	0.047 6
D	0.036 2

由表10可知,在对应形状参数和尺度参数的条件下,其最大差值 $D_{\max}$ 均小于 D ,即核密度估计下的四个品牌商用车发动机及底盘系统的故障间隔时间数据样本服从威布尔分布。

核密度估计法和传统中位秩法都是用于估计一个未知分布函数的非参数统计方法。通过以上分布计算可以发现,这两种方法都可以得到合理的估计结果。但是,在本文商用车故障大数据基础上进行系统可靠性评估时,核密度估计得到的分布参数结果更适

合,原因如下:

1)更优的精度。核密度估计可以通过选择合适的核函数和带宽参数来适应不同的数据分布和特征,平衡样本数量和估计精度,从而提供更准确的概率密度函数估计。而中位秩法只是基于数据的排序,忽略了大量有价值的信息,同时样本量的增加可能会导致中位数被较多的极端值影响,其结果存在高估或低估的情况。

2)更平滑的概率密度曲线。平滑的概率密度函数可以提高可靠性评估结果的准确性,减少因数据噪声或不足而引起的误差,更好地反映失效概率分布的实际情况,从而使评估结果更加可靠。

3)更高的模型泛化能力。利用Sklearn机器学习库进行核密度估计时,根据数据集的特征,通过交叉验证多次训练和验证模型来评估不同核函数和带宽的模型性能,比较它们在数据集上的表现,并选择具有最佳性能的参数组合来避免模型过拟合或欠拟合,降低模型参数选择误差并提高模型性能和评估精度。

3 商用车发动机及底盘系统可靠性指标分析

为了有效评估商用车发动机及底盘系统的可靠性水平,借助核密度估计法得到的分布参数计算累积分布函数、可靠度、故障率及平均故障间隔时间^[16]4个可靠性评价指标,进一步了解系统的故障状况,帮助确定系统的不足和改进方向,并针对性地采取合适的优化措施来提高其可靠性。

3.1 累积分布函数

累积分布函数(Cumulative Distribution Function, CDF)描述的是在一定时间内,系统出现故障的概率。其函数曲线越陡峭,说明系统在给定时间内发生故障的概率越高。

四个品牌商用车发动机及底盘系统的故障间隔时间累积分布函数曲线对比如图5所示。

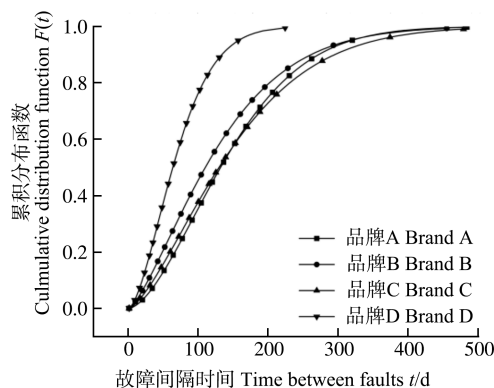


图5 系统累积分布函数曲线对比

Fig. 5 Comparison of curves of system cumulative distribution functions

由图 5 可以看出, 品牌 A、B、C 累积分布函数曲线接近, 品牌 D 曲线的上升斜率明显大于其他 3 条曲线, 趋势集中在 $[0, 200]$ 这个早期故障间隔时间区间内, 说明品牌 D 发动机及底盘系统在早期故障频率集中。

3.2 可靠度

可靠度 (Reliability, R) 是指在一定时间和条件下, 系统或设备在不发生故障的情况下, 能够按照预期性能指标正常工作的概率。可靠度函数形式为

$$R(t) = e^{-\left(\frac{t}{\alpha}\right)^{\beta}} \quad (14)$$

四个品牌商用车发动机及底盘系统的可靠度函数曲线对比如图 6 所示。

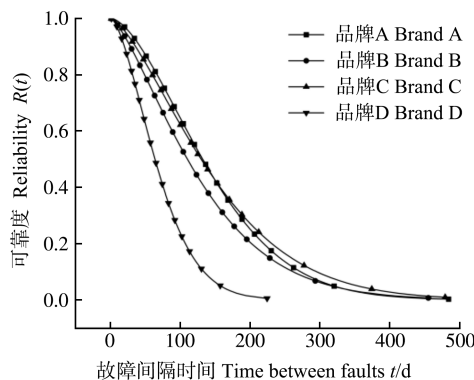


图 6 系统可靠度函数曲线对比

Fig. 6 Comparison of curves of system reliability functions

由图 6 可以看出, 品牌 D 可靠度曲线的下降速度明显大于其他 3 条曲线, 且趋势集中在 $[0, 200]$ 这个早期故障间隔时间区间内, 说明品牌 D 发动机及底盘系统在早期可靠度较差。

根据式 (14) 计算四个品牌商用车行驶 500 天后的发动机及底盘系统可靠度值, 如表 11 所示。

表 11 商用车发动机及底盘系统可靠度

Tab. 11 Reliability of commercial vehicle engine and chassis systems

品牌 Brand	可靠度 Reliability
A	0.001
B	0.003
C	0.007
D	3×10^{-9}

由表 11 可以得出, 品牌 D 发动机及底盘系统可靠度值仅为 3×10^{-9} , 远低于其余 3 个品牌的可靠度值, 说明品牌 D 发动机及底盘系统在行驶 500 天后, 正常运行的概率极低, 性能表现非常不可靠。

3.3 故障率

故障率 (Fault Rate, λ) 是指单位时间内系统或设备出现故障的概率或频率, 描述了系统或设备在一段时间内出现故障的可能性。通常情况下, 故障率在使

用初期较高, 随着时间的推移和维护保养的加强, 故障率逐渐降低, 达到一个较为稳定的水平。故障率 λ 函数形式为

$$\lambda(t) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha}\right)^{\beta-1} \quad (15)$$

四个品牌商用车发动机及底盘系统的故障率函数曲线对比如图 7 所示。

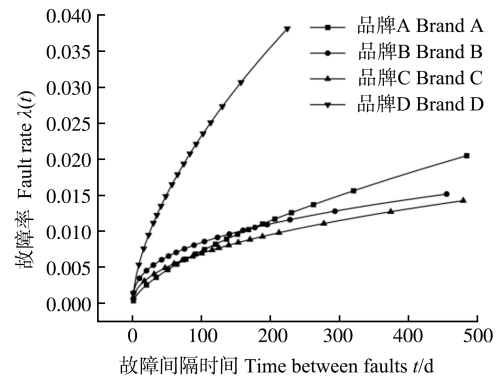


图 7 系统故障率函数曲线对比

Fig. 7 Comparison of curves of system fault rate functions

由图 7 可以看出, 品牌 D 故障率曲线的上升幅度远大于其他 3 条曲线, 且趋势集中在 $[0, 200]$ 这个早期故障间隔时间区间内, 说明品牌 D 发动机及底盘系统在早期故障率较高。另外, 品牌 A 在 200 天之后曲线呈现上升趋势, 可以认为品牌 A 在中后期发动机及底盘系统故障率有增加的趋势。

根据式 (15) 计算四个品牌商用车行驶 500 天后的发动机及底盘系统故障率值, 如表 12 所示。

表 12 商用车发动机及底盘系统故障率

Tab. 12 Fault rate of commercial vehicle engine and chassis systems

品牌 Brand	故障率 Fault rate
A	0.021
B	0.016
C	0.015
D	0.063

由表 12 可以得出, 品牌 D 发动机及底盘系统故障率值为 0.063, 说明品牌 D 发动机及底盘系统在行驶 500 天后, 发生故障的概率最高, 极有可能会出现故障。

3.4 平均故障间隔时间

平均故障间隔时间 (Mean Time Between Faults, MTBF) 是指系统或设备连续正常工作的平均时间, 即从一个故障发生到下一个故障发生的平均时间间隔, MTBF 值越大, 说明系统或设备的可靠性越高, 正常工作的时间越长。一般情况下, T_{MTBF} 计算式为

$$T_{MTBF} = \frac{1}{n} \sum_{i=1}^n t_i \quad (16)$$

式中, n 为样本总数; t_i 为故障间隔时间。对于两参数威布尔分布来说, 其数学期望 E 可以直接看作是 MTBF 的值, 计算式为

$$E(t) = \alpha\tau(1 + \frac{1}{\beta}) \quad (17)$$

式中, τ 为伽马函数。将四个品牌商用车发动机及底盘系统的分布参数代入式 (16), 平均故障间隔时间如表 13 所示。

表 13 系统平均故障间隔时间

Tab. 13 Mean time between faults of the system

品牌 Brand	平均故障间隔时间 Mean time between faults/d
A	146.087 5
B	129.491 6
C	149.755 4
D	70.897 3

结合图 5~图 7 及表 11~表 13 可靠性评价指标曲线及数值, 对四个品牌商用车发动机及底盘系统进行对比分析, 可以得到:

1) 品牌 A、B、C 三款商用车发动机及底盘系统表现较好, 可靠性指标和曲线趋势接近, 寿命较长, 故障

率较低, 可以满足客户的可靠性需求。这三个品牌应当在保持当前可靠性水平的基础上持续优化商用车发动机及底盘系统的可靠性和竞争力, 建立完善的售后服务和维修体系, 进一步提高用户的满意度和信任度。

2) 品牌 D 在故障间隔时间前期, 发动机及底盘系统的可靠性相对较差, 存在较高的故障率和较短的故障间隔时间。特别是当车辆行驶超过 500 天后, 系统可靠度与故障率值极低, 仅为 3×10^{-9} 与 0.063, 基本无法满足商用车正常运行与行驶功能。针对该情况, 需要结合可靠性评价指标及具体故障内容, 采取有关改进措施以提高其商用车发动机及底盘系统的可靠性。对 2021 年度品牌 D 商用车发动机及底盘系统故障发生频率进行排序, 抽取频率最高的前 3 条故障内容, 分析故障原因及提出改进措施。品牌 D 商用车发动机及底盘系统故障分析内容如表 14 所示。

由表 14 可以发现, 故障频率最高的 3 条故障内容都发生于后处理系统。因此, 应加强该系统零部件设计和制造质量的控制, 采取针对性的改进措施提高设备可靠性和耐用性。同时强化故障预警和监测系统, 记录详细故障数据, 定期检查和维修, 及时处理和解决故障。

表 14 品牌 D 商用车发动机及底盘系统故障分析

Tab. 14 Fault analysis of engine and chassis systems for commercial vehicles of brand D

故障名称 Fault name	故障系统 Fault system	故障原因 Fault reason	改进措施 Improvement measure
EGR 阀卡滞故障-大开度 EGR valve stuck fault-wide opening	后处理系统 Post-processing system	EGR 阀门磨损, 积碳, 堵塞 EGR valve wear, carbon accumulation, blockage	定期更换易损零部件, 使用高质量的燃油和润滑油 Replace vulnerable parts regularly, and use high-quality fuels and lubricants
SCR 下游温度传感器动态合理性故障 (偏差超上限) SCR downstream temperature sensor dynamic rationality fault (deviation exceeds the upper limit)	后处理系统 Post-processing system	传感器故障、线路接触不良或损坏 Sensor is faulty, the cable is in poor contact, or damaged	定期更换 SCR 系统零部件, 定期维护下游温度传感器及其相关线路 Regular replacement of SCR system components, regular maintenance of downstream temperature sensors and related lines
尿素消耗量偏差大故障 Urea consumption deviation major fault	后处理系统 Post-processing system	元件损坏、尿素液质量差 Component damage, urea fluid quality is poor	定期检查和更换 SCR 系统元件、清洗管路 Check and replace SCR system components and clean pipes regularly

4 结论

本文提出了一种基于故障大数据的可靠性评估方法, 并将其应用到某商用车制造企业旗下四款商用车发动机及底盘系统的可靠性评估当中。得到以下结论:

1) 提出 U 检验方法确定最优样本数据分布, 保证样本估计总体可靠性水平的准确性。四个品牌商用

车发动机及底盘系统的故障数据样本 U 检验 P 值分别为 0.479 1、0.493 2、0.708 2、0.892 7, 均大于显著性水平 0.05, 证明故障样本数据与总体分布相似。

2) 运用中位秩法及核密度估计法确定威布尔分布参数并对比验证, 发现核密度估计法结合机器学习的交叉验证评估手段下处理的数据分布及参数更适用于商用车故障大数据下的可靠性评估, 同时结合 K-S 假设检验完成分布验证。四个品牌商用车发动机

及底盘系统的故障数据样本K-S假设检验得到的最大差值 $D_{\max}$ 分别为0.033 5、0.049 7、0.047 6、0.036 2,均小于临界值0.515,即四个品牌商用车发动机及底盘系统的故障间隔时间数据样本服从威布尔分布。

3)绘制可靠性评价曲线并计算指标,得出四款商用车在行驶500天后,其可靠度分别为0.002、0.003、0.007、 3×10^{-9} ;故障率分别为0.021、0.016、0.015、0.063;平均故障间隔时间分别为146.088天、129.492天、149.755天、70.897天。根据指标结果与曲线对比,对可靠性最差的品牌D商用车发动机及底盘系统可靠性优化进行讨论分析并提出改进措施,为商用车可靠性改进提供了参考依据。

参考文献(References)

- [1] 中华人民共和国交通部. 中国道路运输发展报告[M]. 北京:人民交通出版社,2021:119-131.
Ministry of Transport of the People's Republic of China. National report on road transport development[M]. Beijing: China Communications Press, 2021: 119-131 (In Chinese).
- [2] JONES M. Commercial vehicle component reliability survey [J]. Commercial Motor, 2019, 144(4188): 16-19.
- [3] 瞿佳佳,金 婷. 基于医疗设备故障数据集的医疗设备可靠性分析[J]. 中国医疗设备, 2022, 37(1): 98-101.
QU JiaJia, JIN Ting. Reliability analysis of medical equipment based on failure data set [J]. China Medical Devices, 2022, 37(1): 98-101 (In Chinese).
- [4] 朱 荣. 基于可靠性的混合动力公交客车高压电气系统故障分布规律研究[J]. 汽车工程, 2021, 50(5): 23-27.
ZHU Rong. Study on fault distribution law of high voltage electrical system of hybrid electric bus based on reliability [J]. Automotive Engineering, 2021, 50(5): 23-27 (In Chinese).
- [5] 庞 欢,石东阳,王道成,等. 基于质保数据的汽车可靠性评估及车企运营决策系统[J]. 汽车安全与节能学报, 2022, 13(2): 250-258.
PANG Huan, SHI DongYang, WANG DaoCheng, et al. Automobile reliability evaluation and automobile-enterprise operational decision system based on warranty date [J]. Journal of Automotive Safety and Energy, 2022, 13(2): 250-258 (In Chinese).
- [6] 张炎驿,傅振亮,田雪艳,等. 列车可修系统的可靠性分析与评估[J]. 城市轨道交通研究, 2018, 21(10): 1-4.
ZHANG YingYi, FU ZhenLiang, TIAN XueYan, et al. Reliability analysis and estimation of train repairable system [J]. Urban Mass Transit, 2018, 21(10): 1-4 (In Chinese).
- [7] 刘 伟,张冬梅,高 磊. 基于运营数据的某型动车组转向架可靠性分析[J]. 铁道车辆, 2021, 59(4): 31-34.
LIU Wei, ZHANG DongMei, GAO Lei. Reliability analysis of bogies on a certain type of multiple units based on operation data [J]. Rolling Stock, 2021, 59(4): 31-34 (In Chinese).
- [8] 王 伟,李永生,段洪亮. 大数据技术在城市轨道交通车辆RAMS验证过程中的应用[J]. 城市轨道交通研究, 2020, 23(2): 1-4.
WANG Wei, LI YongSheng, DUAN HongLiang. Application of big data technology in urban rail transit vehicles RAMS verification [J]. Urban Mass Transit, 2020, 23(2): 1-4 (In Chinese).
- [9] 赵礼辉,王 震,冯金芝,等. 基于用户大数据的电驱动系统可靠性工况特征研究[J]. 汽车工程, 2020, 42(10): 1386-1396.
ZHAO LiHui, WANG Zhen, FENG JinZhi, et al. Research on reliability condition characteristics of electric drive system based on user big data [J]. Automotive Engineering, 2020, 42(10): 1386-1396 (In Chinese).
- [10] 闫 晓. 基于非参数方法的两总体分布相同检验问题[D]. 北京:首都经济贸易大学, 2019: 7-9.
YAN Xiao. Testing problem of two population distributions with same distribution based on nonparametric method [D]. Beijing: Capital University of Economics and Business, 2019: 7-9 (In Chinese).
- [11] 袁非梦,耿化品,王思蕊,等. 基于概率分布特征和代理模型的试验数据扩充方法研究[J]. 机械强度, 2023, 45(4): 871-878.
YUAN FeiMeng, GENG HuaPin, WANG SiRui, et al. Research on test data expansion method based on probability distribution characteristics and proxy model [J]. Journal of Mechanical Strength, 2023, 45(4): 871-878 (In Chinese).
- [12] 曹慧荣. 全最小二乘法在威布尔分布参数估计中的应用[J]. 高师理科学刊, 2023, 43(2): 6-10.
CAO HuiRong. Application of total least squares in parameter estimation of Weibull distribution [J]. Journal of Science of Teachers' College and University, 2023, 43(2): 6-10 (In Chinese).
- [13] 张剑锋,徐 芳,张 燕,等. 三参数威布尔分布的参数估计及置信限计算方法[J]. 机械强度, 2024, 46(5): 1136-1142.
ZHANG JianFeng, XU Fang, ZHANG Yan, et al. Parameter estimation and confidence limit calculation method of three-parameter Weibull distribution [J]. Journal of Mechanical Strength, 2024, 46(5): 1136-1142 (In Chinese).
- [14] 薛光明,宁 鹏,钱明军,等. 基于改进平均秩的最小二乘参数估计方法[J]. 机械强度, 2023, 45(2): 380-385.
XUE GuangMing, NING Peng, QIAN MingJun, et al. Least square parameter estimation method based on improved mean rank [J]. Journal of Mechanical Strength, 2023, 45(2): 380-385 (In Chinese).
- [15] PARZEN E. On estimation of a probability density function and mode [J]. The Annals of Mathematical Statistics, 1962, 33(3): 1065-1076 (In Chinese).
- [16] 刘向阳,韩友国,徐承付,等. 基于威布尔分布的汽车零部件可靠性分析[J]. 汽车零部件, 2022(9): 40-43.
LIU XiangYang, HAN YouGuo, XU ChengFu, et al. Reliability analysis of auto parts based on Weibull distribution [J]. Automotive Parts, 2022(9): 40-43 (In Chinese).