

EEMD 分解与多特征结合的地铁一系悬挂故障诊断

许官儒,戴焕云

(西南交通大学 牵引动力国家重点实验室,四川 成都 610031)

摘 要:针对地铁运营过程中各类故障很难及时发现的问题,基于构架垂向振动加速度信号进行分析,将信号时域特征指标与聚合经验模态分解出来的固有模式分量的样本熵组合,构成多维特征向量,使用支持向量机进行故障状态识别。根据仿真实验数据显示,当地铁列车以 76 km/h 速度运行时,故障识别准确率达到并稳定在 94% 以上。

关键词:地铁;聚合经验模态分解;样本熵;时域特征;支持向量机

中图分类号:TP277.3 **文献标志码:**B **文章编号:**1671-5276(2021)05-0191-05

Fault Diagnosis of Metro Primary Suspension Based on EEMD and Multiple Features

XU Guanru, DAI Huanyun

(State Key Laboratory of Traction Power, Southwest Jiaotong University, Chengdu 610031, China)

Abstract: To solve the fault diagnosis difficulty in metro operation, the combination of time-domain characteristic index of the signal and the sample entropy of the intrinsic mode components decomposed by the ensemble empirical mode is made, based on the analysis of the vertical vibration acceleration signal of the frame, thus forming a multi-dimensional feature vector, and the support vector machine is used to identify the fault state. The experimental data show that the accuracy of fault identification and its stability can be obtained at more than 94% when the metro train runs at 76 km/h.

Keywords: metro; ensemble empirical mode decomposition; sample entropy; time domain characteristics; support vector machine

0 引言

地铁列车一系悬挂主要有一系钢弹簧和一系垂向减振器,一系钢弹簧起到连接和缓冲的作用,储存并释放能量,一系垂向减振器可以衰减构架与轮对之间的振动,减小轮轨冲击。在日常检修过程中,多次出现一系钢弹簧断裂^[1]和一系垂向减振器失效的情况^[2]。故障悬挂将会影响车辆动力学性能,甚至影响列车运行安全性,因此监测一系悬挂故障状态很有必要。由于列车振动信号具有较强的非线性与非平稳性特征,传统的时域分析和频域分析方法使用受限,如傅里叶变换通常是建立在平稳信号基础上,因此一些时频分析法引入到信号处理中,如小波分析法^[3],但其依赖于小波基函数的选取,实际应用时效果难以达到预期。秦娜等^[4]提出一种基于聚合经验模态分解(EEMD)并结合样本熵测度对高速列车转向架故障特征进行识别的方法;金子博等^[5]提出城轨车辆二系悬挂故障时,均方根值和方差更适合作为频谱变化的量化指标;刘棋等^[6]将 EEMD 与熵特征结合,准确识别了高速列车蛇形失稳状态。

本文通过选取构架垂向振动信号的多个时域统计指标,与 EEMD 分解后的 IMF 样本熵进行组合,组成高维特征向量输入到支持向量机,提高了故障分类准确率。

1 EEMD 与样本熵计算

1.1 构架振动信号的 EEMD 分解

经验模态分解(EMD)是由黄鹄教授提出的一种对非平稳信号进行自适应分解的信号分析法^[7],但其在使用过程中存在模态混叠的问题。EEMD 是 HUANG 和 WU 等人提出的一种噪声辅助分析法,可抑制 EMD 中出现的模态混叠现象^[8],其原理是利用白噪声频率均匀分布的统计特性使信号在不同尺度上具有连续性。对构架振动信号通过以下步骤进行 EEMD 分解:

1) 在构架振动信号 $s(t)$ 中多次加入高斯白噪声 $n_i(t)$,其标准差取构架振动信号的 0.4 倍,得

$$s_i(t) = s(t) + n_i(t) \quad (1)$$

式中: $s_i(t)$ 表示第 i 次加入高斯白噪声之后的信号。

2) 对 $s_i(t)$ 进行 6 层 EMD 分解,得到 IMF 分量 $c_{ij}(t)$ 和残余量 $r_i(t)$,即

$$s_i(t) = \sum_{j=1}^6 c_{ij}(t) + r_i(t) \quad (2)$$

式中 $c_{ij}(t)$ 是第 i 次加入高斯白噪声后分解得到的第 j 个 IMF 分量。

3) 对 EMD 分解后的每个 IMF 分量总体平均,计算结果即为 EEMD 分解后的 IMF 分量,即

$$c_j(t) = \frac{1}{N} \sum_{i=1}^N c_{ij}(t) \quad (3)$$

原车工况下构架振动垂向加速度信号 EEMD 分解结果如图 1 所示。可以看出,IMF1 是添加的白噪声成分,频率最高,IMF2-IMF6 分量的频率依次降低,并且 EEMD 分解有效抑制了模态混叠现象。

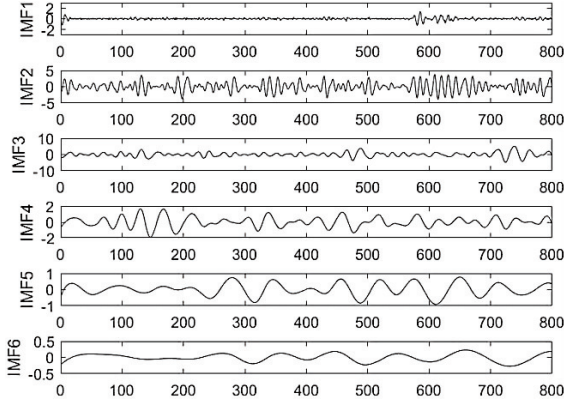


图 1 原车工况 EEMD 分解结果

1.2 样本熵计算方法

样本熵是度量系统复杂度和不规则性的一种指标,适用于复杂机械振动系统的分析。IMF 分量的样本熵计算方法如下^[9]:

1) N 个点组成序列 $x_{(1)}, x_{(2)}, \dots, x_{(n)}$, 选取序列中 m 个连续点构成窗口子序列:

$$x_m(i) = [x(i), x(i+1), \dots, x(i+m-1)], \quad (4)$$

$$i = 1, 2, \dots, N-m+1$$

2) 定义两个窗口子序列 $x_m(i)$ 和 $x_m(j)$ 间距离为 $x_m(i)$ 和 $x_m(j)$ 对应数据点最大差的绝对值, 记作

$$d[x_m(i), x_m(j)] = \max_l |x_m(i+l) - x_m(j+l)|, \quad (5)$$

$$l = 0, 1, 2, \dots, m-1$$

3) 给定阈值 r , 计算每个窗口子序列 $x_m(i)$ 与其他窗口子序列的距离, 找出其中距离小于 r 的个数, 并求其占总个数的比值, 记作

$$B_i^m(r) = \frac{\text{count} \{ d[x_m(i), x_m(j)] < r \}}{N-m}, \quad (6)$$

$$j = 1, 2, \dots, N-m+1 \text{ 且 } i \neq j$$

式中 count 函数表示数量统计。

4) 依次对 $x_m(i)$ 计算 $B_i^m(r)$, 并求其平均值:

$$B_i^m(r) = \frac{\sum_{i=1}^{N-m+1} B_i^m(r)}{N-m+1} \quad (7)$$

5) 窗口子序列长度增加 1, 即 m 增加至 $m+1$, 重复 1) 至 4), 计算得 $B^{m+1}(r)$, 即

$$B^{m+1}(r) = \frac{\sum_{i=1}^{N-m} B_i^{m+1}(r)}{N-m} \quad (8)$$

6) 实际工程应用中, 样本熵计算公式为

$$S_n(m, r, N) = -\ln \frac{B^{m+1}}{B^m} \quad (9)$$

2 地铁列车动力学模型

2.1 SIMPACK 模型建立

故障分析的基础是数据监测, 本文使用 SIMPACK 动力学仿真软件计算各种故障数据。根据某地铁列车的参数建立动力学模型如图 2 所示。

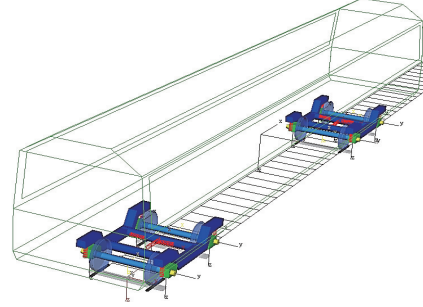


图 2 某型地铁列车 SIMPACK 动力学模型

2.2 一系悬挂故障工况分析

根据对地铁车辆运营过程中出现的故障统计可知, 一系悬挂系统中钢弹簧和垂向减振器是出现故障较多的部件。一系钢弹簧最易出现的故障是弹簧断裂, 导致垂向刚度增加, 在 SIMPACK 仿真时通过改变其垂向刚度值来模拟该故障; 一系垂向减振器容易出现减振器漏油, 致使阻尼值降低, 通过改变垂向阻尼值来模拟。二者均连接轮对轴箱和构架, 对车辆垂向稳定性影响较大, 因此可以选取构架垂向振动作为信号源进行后续分析。

地铁列车运用部门统计, 实际中一系悬挂故障多数情况是单个部件失效, 即单个一系垂向减振器故障或单个一系钢弹簧故障, 本文研究的故障均指单个悬挂元件失效。图 3、图 4 为前转向架单个一系垂向减振器失效、一系钢弹簧失效与正常工况下构架垂向加速度振动信号的对比图(本刊黑白印刷, 相关疑问请咨询作者)。可见当仅一个悬挂元件失效时, 故障特征不明显, 从时域和频域上难以直接判断故障类型, 故传统方法不能准确判断悬挂的故障状态。

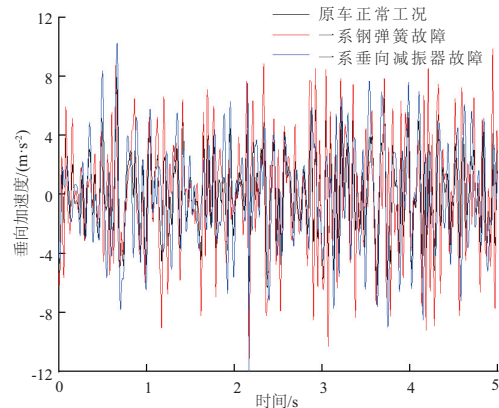


图 3 3 种工况的垂向振动时域对比图

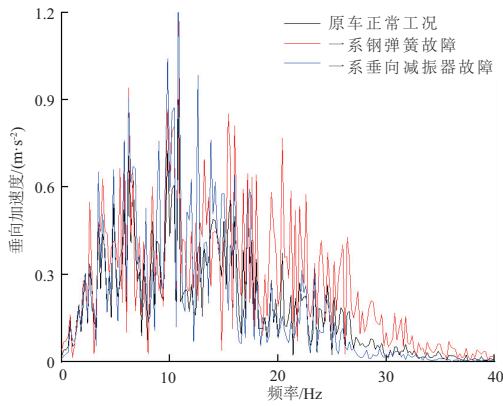


图4 3种工况的垂向振动频域对比图

3 一系悬挂故障诊断仿真实验

3.1 仿真实验方案

SIMPACK 仿真实验所加实测轨道激励为美国五级谱,地铁列车以 76 km/h 运行,在构架端部设置传感器采集垂向振动加速度数据,采样频率为 200 Hz。列车运行工况分为 3 种,即正常工况、一系垂向减振器失效、一系钢弹簧失效,每种工况下各自运行 200 s。选取每 4 s 的运行数据为 1 个样本,每种工况包含 50 个样本,每个样本包含 800 个采样点,3 种工况合计 150 个样本。

3.2 振动信号处理

本文基于 EEMD 分解和多特征的信号处理流程如图 5 所示。首先对构架振动信号滤波,提取区分度最高的信号成分,接着分别计算滤波后信号的时域特征指标和 EEMD 分解后的 IMF 样本熵,选取两类特征中区分度最高的指标,组合成 6 维特征向量,输入支持向量机进行识别。

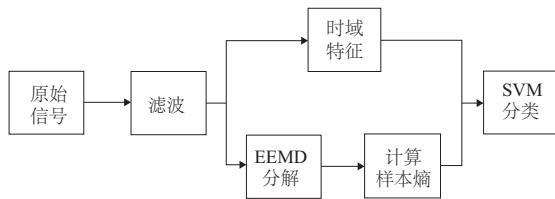


图5 信号处理流程图

1) 信号滤波处理

为了提取 3 种工况区分度最高的信号成分,需要对原始信号进行滤波以去除干扰,提高区分度。根据图 4 频域对比可知,当单个一系垂向减振器出现故障时,在 5~20 Hz 频带内,区分度相对较高,当单个一系钢弹簧出现故障时,与其余工况在 5~30 Hz 频带内的差异较为明显,为全面覆盖 3 种工况信号的主要特点,对时域信号 5~30 Hz 做带通滤波处理。

2) 振动信号 EEMD 分解与样本熵提取

EEMD 分解会根据信号自身的特点自适应地将不同频带的固有振动成分分解至 IMF 中。由于信号 EEMD 分

解后,不可避免会存在一些虚假分量。这些虚假分量会影响后续的特征提取,干扰识别结果,故通过计算各分量与原始信号的相关系数来选择合适的 IMF 分量。3 种工况下 IMF 分量与原信号的相关系数计算结果见表 1。

表1 3种工况 IMF 分量与原始信号相关系数

IMF 分量	正常工况	一系钢弹簧故障	一系垂向减振器故障
IMF1	0.111 2	0.118 4	0.179 4
IMF2	0.797 1	0.801 0	0.780 9
IMF3	0.625 1	0.653 6	0.719 3
IMF4	0.381 7	0.374 7	0.346 5
IMF5	0.119 6	0.177 6	0.152 4
IMF6	0.062 5	0.076 5	0.043 4
IMF7	0.003 9	0.039 4	0.013 7

由表 1 可以看出,信号 EEMD 分解后,IMF1 分量与原始信号的相关系数较低,是因为其含有添加的高频噪声。IMF2、IMF3、IMF4 的相关系数明显比其他分量大,包含了原始信号的大部分有用信息,故选择 IMF2、IMF3、IMF4 这 3 个分量作为特征向量 H1,以表征原信号的特性,其三维空间分布如图 6 所示。

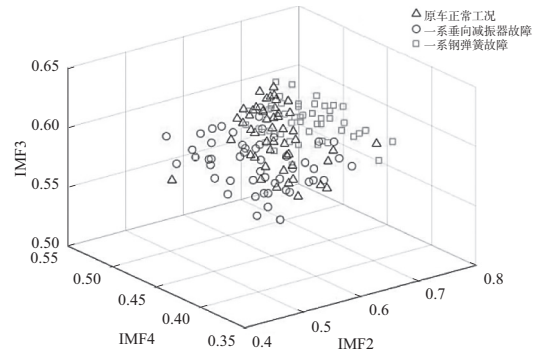
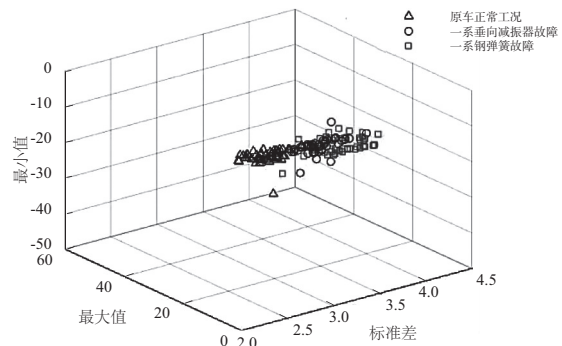


图6 样本熵的空间分布

3) 振动信号的时域统计指标提取

振动信号的时域特征指标主要有标准差、最大值、最小值、均值、峭度、偏度等,地铁列车单个一系悬挂出现故障时,主频变化不够显著,在时域上,幅值变化相对更加明显。3 种工况下 6 种时域指标的三维空间分布如图 7 所示。可见,标准差、最大值、最小值远比均值、峭度和偏度区分度更高、类内聚集性更好。故选取标准差、最大值、最小值作为时域特征向量 H2。



(a) 标准差、最大值、最小值的空间分布

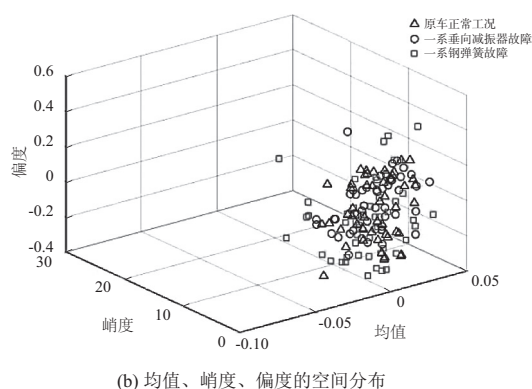


图7 时域指标空间分布

4)将组成特征矩阵2)中提取的3个IMF样本熵特征向量H1与3)中选取的3个时域特征向量H2组合,构成6维特征向量(标准差、最大值、最小值、IMF2样本熵、IMF3样本熵、IMF4样本熵),组成150×6的特征矩阵。

5)分类设计

支持向量机(support vector machine, SVM)是利用统计学与结构风险最小化理论寻求最优分类面发展来的一种工具,学习性能较好,避免了传统方法对样本数量的高要求,可实现小样本和非线性数据的分类。将3)中提取的6维特征向量输入至向量机,在150个样本中,随机选取80个样本进行训练,剩余70个样本作为测试。

3.3 方法对比

为验证本文提出的方法用于地铁列车一系悬挂故障识别的有效性,对比分析4种方案的识别效果如表2所示。

表2 4种特征提取方式的识别准确率

序号	特征提取方法	SVM 识别准确率/%
1	6种时域特征	72.86 (51/70)
2	选取后的3种时域特征	81.43 (57/70)
3	EEMD结合样本熵	82.86 (58/70)
4	EEMD结合样本熵+时域特征	97.14 (68/70)

第1种方案是单独使用时域指标作为特征向量输入到向量机。将6种时域指标(标准差、最大值、最小值、均值、峭度、偏度)组成6维特征向量,构成150×6的特征矩阵输入至向量机中,识别率为72.86%。

第2种是选取标准差、最大值和最小值组成3维特征向量输入至向量机,识别率为81.43%。可见,经过选取之后的时域指标作为特征向量,去除了区分度较低指标的干扰,识别率得到提高。但单独使用时域指标作为特征向量,反映的信息依然不够全面,识别预测情况如图8所示,类别标签1为正常工况,标签2是一系垂向减振器故障,标签3是一系钢弹簧故障。

第3种是单独使用EEMD分解后的IMF分量的样本熵作为特征向量输入,振动信号EEMD分解后,选取前6个固有模态函数的样本熵组成150×6的特征矩阵,其识别率为82.86%。可见EEMD分解在提取非线性、非平稳

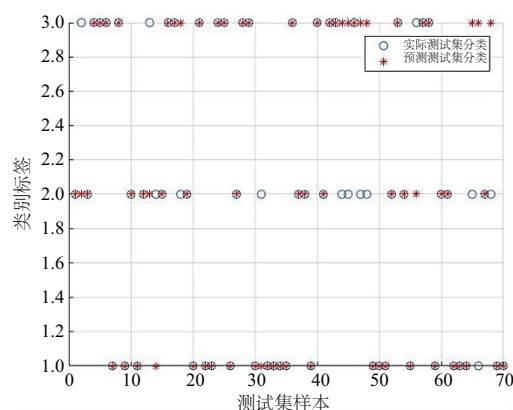


图8 选取3种时域指标识别

信号特征时有较好的效果,但由于地铁列车运行速度较低,故障特征不明显,3种工况样本熵分布没有能够完全区分开,对准确率造成了一定影响,识别预测结果如图9所示。

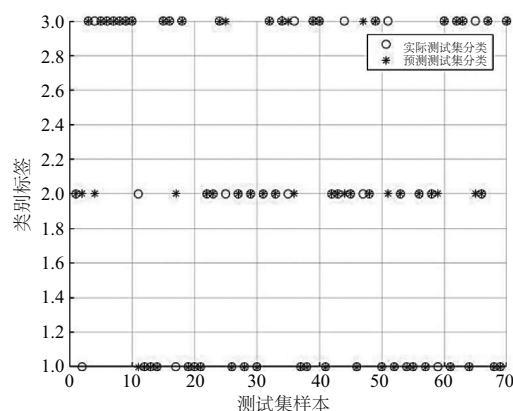


图9 IMF样本熵方法识别结果

第4种是本文提出的IMF分量样本熵与时域特征指标组合的6维特征向量作为输入,分类预测如图10所示,识别准确率为97.14%,与前3种方法相比明显减少了误报和漏报。可见地铁列车单个一系悬挂元件出现故障时,单一类别指标不能完全反映故障信号特点,本文提出的多类指标组合的方法更加全面地囊括了信号的主要特征,能够准确识别一系悬挂故障。

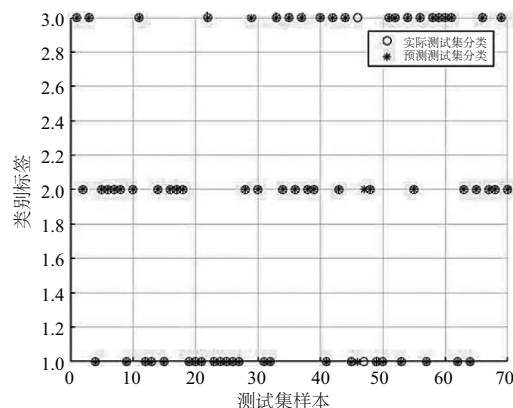


图10 本文方法的分类预测图

4 结语

地铁列车运行速度较低,普遍低于 80 km/h,出现单个一系悬挂部件故障时对振动信号的影响不够显著,且振动信号有较强的非平稳性和非线性特征。本文运用聚合经验模态分解出 IMF 分量样本熵与部分时域特征指标相结合的方法,对地铁列车一系垂向减振器故障、一系钢弹簧故障和正常工况共 3 种工况进行分类识别,与单独使用 EEMD 分解的样本熵作为特征向量和单独使用时域特征作为特征向量相比,信号特征覆盖更加全面,故障诊断准确率得到明显提升,而且减小了 EMD 方法中的模态混叠问题。当列车以 76 km/h 运行时,识别率达到并稳定在 94% 以上。

参考文献:

- [1] 尹太国,李伟,张鹏,等. 地铁车辆一系钢弹簧中高频动态特性分析[J]. 机械,2019,46(6):1-7,41.
- [2] 黄先富,门永林,祝航. 地铁一系垂向减振器失效故障分析[J]. 铁道车辆,2019,57(5):34-36,6.
- [3] TENG W, DING X, ZHANG X L, et al. Multi-fault detection and failure analysis of wind turbine gearbox using complex

wavelet transform[J]. Renewable Energy,2016,93:591-598.

- [4] 秦娜,金炜东,黄进,等. 基于 EEMD 样本熵的高速列车转向架故障特征提取[J]. 西南交通大学学报,2014,49(1):27-32.
- [5] 金子博,方宇,陆远航,等. 基于多体动力学仿真的城市轨道交通车辆二系悬挂故障特征分析[J]. 城市轨道交通研究,2016,19(7):97-101.
- [6] 刘棋,宁静,叶运广,等. 基于 EEMD 能量熵的高速列车蛇行诊断研究[J]. 中国测试,2017,43(5):96-100.
- [7] HUANG N E, SHEN Z, LONG S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proceedings of the Royal Society of London Series A: Mathematical, Physical and Engineering Sciences,1998,454(1971):903-995.
- [8] 卓仁雄,肖金凤. 基于改进的集合经验模态分解的电动机滚动轴承故障诊断研究[J]. 机械制造与自动化,2019,48(1):36-39.
- [9] 楼军伟,胡赤兵,王季,等. 基于 EEMD 样本熵的电机轴承电流信号复杂性评估[J]. 机械制造与自动化,2014,43(5):36-39.

收稿日期:2020-11-05

(上接第 174 页)

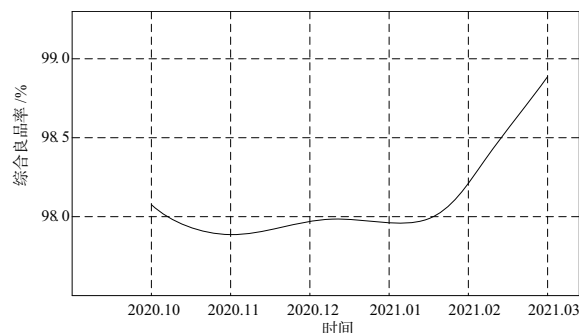


图 4 系统应用后对综合良品率的影响

图 4 中,即便在该系统投入应用的第 1 个月,企业的良品率持续稳定在(98.0±0.05)%左右,2021 年 2 月,企业综合良品率达到 98.24%,2021 年 3 月,企业综合良品率达到 98.75%,保持了持续提升的趋势。此处企业综合良品率是企业良品产量占全部产量的比值,综合了所有生产线和所有车间的良品率数据。

4 结语

综合上述研究发现,企业综合良品率受到企业车间生产线现场管理中的排列熵控制能力影响,但因为当前对排列熵的计算过程,受到当前大型计算机最大算力的制约,无法直接计算。所以该研究通过良品率实际发生值的数据反推产品质量熵(产品质量相关参数的排列熵比值),进而通过产品质量熵的多列神经网络计算企业内的人事、设备、流程管理能力绩效结果。该系统投入应用后,有效

促进了歌尔股份的企业综合良品率,且该提升过程在该系统投入应用后,处于持续上升的过程中,未来预期可将歌尔股份的综合良品率提升到 99% 以上。

参考文献:

- [1] 邢广鑫,周以齐. 基于 LMD 排列熵和 LLE 的滚动轴承故障诊断[J]. 煤矿机械,2021,42(3):164-166.
- [2] MINHAS A S, SINGH S. A new bearing fault diagnosis approach combining sensitive statistical features with improved multiscale permutation entropy method[J]. Knowledge - Based Systems, 2021,218:106883.
- [3] POSE F E, BAUTISTA L, GIANMUSO F, et al. On the permutation entropy Bayesian estimation[J]. Communications in Nonlinear Science and Numerical Simulation,2021,99:105779.
- [4] 王涛,胡定玉,丁亚琦,等. 基于经验模式分解和排列熵的轴承故障特征提取[J]. 噪声与振动控制,2021,41(1):77-81.
- [5] 赵亚辉,胡治钰,赵重阳. 冲压新车初流良品率和 SPH 值快速提升体系的建立[C]//第十六届河南省汽车工程科技学术研讨会论文集. 开封,2019:244-245.
- [6] 贺康. F 半导体公司应用六西格玛理论对测试良品率的改进[D]. 天津:天津大学,2016.
- [7] 段国发. 提高铝药筒阴极电泳涂漆良品率[J]. 表面技术,2015,44(8):126-129.
- [8] 杨新强. 基于 DMAIC 方法论的微控制器产品终测良品率改善[D]. 天津:天津大学,2015.
- [9] 孟丹. A 公司良品率零缺陷管理的改进策略与应用研究[D]. 成都:电子科技大学,2014.
- [10] 李俊英,姚燕燕. 浅谈如何提高隔热盖压制良品率[J]. 科技情报开发与经济,2010,20(26):227-228.
- [11] 罗青. 焊球粘贴生产工艺改进提高良品率[D]. 上海:复旦大学,2010.

收稿日期:2021-03-26