



中华人民共和国国家标准

GB/T 22394.2—2021/ISO 13379-2:2015

机器状态监测与诊断 数据判读与诊断 技术 第2部分:数据驱动的应用

Condition monitoring and diagnostics of machines—Data interpretation and
diagnostics techniques—Part 2: Data-driven applications

(ISO 13379-2:2015, IDT)

2021-05-21 发布

2021-12-01 实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

前言	I
引言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 实施数据驱动监测的过程	2
4.1 数据驱动监测方法原理	2
4.2 资产关键故障和过程参数选择	2
4.3 数据清洗和重采样	3
4.3.1 总则	3
4.3.2 插值误差	3
4.3.3 数据品质问题	3
4.3.4 数据重采样	3
4.4 模型开发	3
4.4.1 总则	3
4.4.2 模型的定义和相关输入的选择	4
4.4.3 相关工况和数据的选择	4
4.4.4 模型测试准备	4
4.5 模型性能评估	4
4.6 报警设置	5
5 实施数据驱动诊断的过程	5
5.1 总则	5
5.2 自动模式分类方法	5
5.3 简化的自动特征分类方法	6
6 实施数据驱动监测方法的一般建议	7
附录 A (资料性附录) 数据驱动监测应用示例	8
附录 B (资料性附录) 数据驱动诊断应用示例	10
参考文献	11

前 言

GB/T 22394《机器状态监测与诊断 数据判读与诊断技术》分为以下 3 部分：

- 第 1 部分：总则；
- 第 2 部分：数据驱动的应用；
- 第 3 部分：基于知识的应用。

本部分是 GB/T 22394 的第 2 部分。

本部分按照 GB/T 1.1—2009 给出的规则起草。

本部分使用翻译法等同采用 ISO 13379-2:2015《机器状态监测与诊断 数据判读与诊断技术 第 2 部分：数据驱动的应用》。

与本部分中规范性引用的国际文件有一致性对应关系的我国文件如下：

- GB/T 20921—2007 机器状态监测与诊断 词汇(ISO 13372:2004, IDT)。
- GB/T 22394.1—2005 机器状态监测与诊断 数据判读和诊断技术 第 1 部分：总则(ISO 13379-1:2012, IDT)。

本部分由全国机械振动、冲击与状态监测标准化技术委员会(SAC/TC 53)提出并归口。

本部分起草单位：华北电力大学、西安热工研究院有限公司、郑州机械研究所有限公司。

本部分主要起草人：付忠广、张学延、马卫平。

引 言

GB/T 22394 的本部分给出了可用于确定机器相对一系列基线参数状态的一般过程。相对于基线值的变化和与报警条件的对比,用于表示异常状态并产生报警,这样的过程通常称为状态监测。另外,为帮助确定合适的处理措施,识别异常状态原因的过程,通常称为诊断。

机器状态监测与诊断 数据判读与诊断 技术 第2部分:数据驱动的应用

1 范围

GB/T 22394 的本部分给出了实施数据驱动监测与诊断方法的过程,以帮助专业人员,特别是监测中心的专业人员进行分析工作。

虽然现有的工具中已经内嵌了一些步骤,为了更好地使用,仍有必要注意下列步骤:

- 资产、关键故障以及可用过程参数的选择;
- 数据清洗和重采样;
- 模型开发;
- 模型初始化和调整;
- 模型性能评估;
- 诊断过程。

实施这些步骤不需要全面的统计学方法知识,但需要具有首先建立训练模型并将模型应用于监测和诊断过程的能力。

在正常工作状态的机器上实施数据驱动监测模型的训练。故障监测的原理是将观察数据与估计数据进行对比。参数的观察值与期望值之间的不同(称为残差)表示存在异常,该异常可能与设备或仪器有关。

在正常工作状态和故障状态的机器上实施数据驱动诊断模型的训练。诊断方法的原理不是去检测参数的偏差,而是通过将观察到的状况与训练阶段学到的故障进行比较从而识别故障。通常使用的技术是模式识别以及模式分类。

数据可以取自分散式控制系统(DCS)的历史数据,或者取自特定的监测系统。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

ISO 13372 机器状态监测与诊断 词汇(Condition monitoring and diagnostics of machines—Vocabulary)

ISO 13379-1 机器状态监测与诊断 数据判读和诊断技术 第1部分:总则(Condition monitoring and diagnostics of machine—Data interpretation and diagnostics techniques—Part 1: General guidelines)

3 术语和定义

ISO 13372 和 ISO 13379-1 界定的术语和定义适用于本文件。

4 实施数据驱动监测的过程

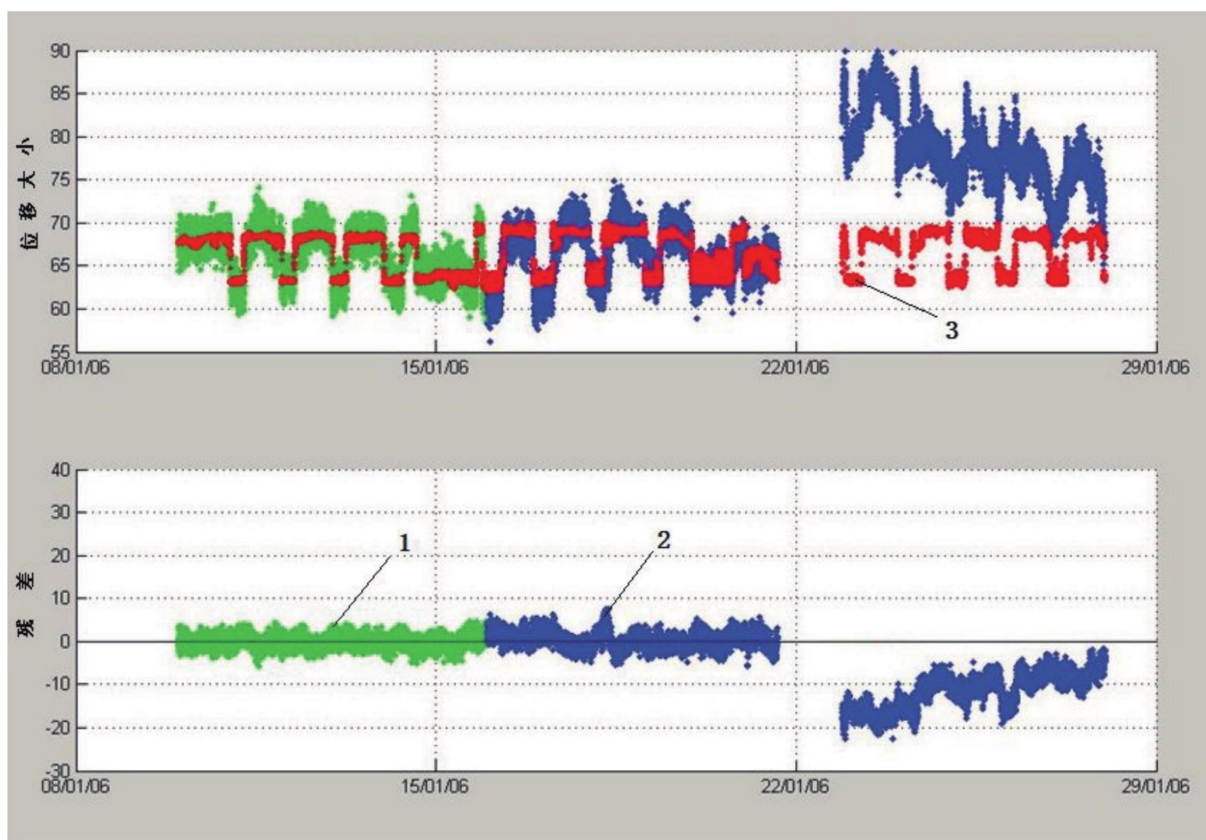
4.1 数据驱动监测方法原理

在状态监测方面传统统计方法是基于单个信号趋势进行观察,现代统计方法则是同时考虑多个设备信号并构建这些信号之间的内在关系。现代统计方法正开始替代传统统计方法。

这些监测方法依照相同的原理监测故障,即将监测数据与估计数据进行对比。

在监测阶段之前,需要在训练阶段建立正常设备状态的模型。这样,可通过系统参数的观察值和期望值之间的偏差进行故障监测。

图 1 所示为在一台燃气轮机上应用的示例,目标是停机后检测异常的轴位移。模型中考虑了几个输入量:有功功率、无功功率和轴承位移。



说明:

- 1——绿色为训练数据;
- 2——蓝色为监测数据;
- 3——红色为预测数据。

图 1 燃气轮机位移大小和残差

数据驱动监测常用方法有自相关核回归法(AAKR)、聚类 and 偏最小二乘法(PLS)、支持向量机法(SVM)和马氏田口法(MT)。

4.2 资产关键故障和过程参数选择

实施数据驱动监测的过程在 ISO 17359 中有精确的描述。包括两个重要的审核:

——设备审核:识别设备及其功能;

——可靠性和关键性审核:绘制可靠性框图,建立设备关键性并完成失效模式、影响及危害度分析。

一旦完成初步研究和识别出关键故障清单,就有必要列出历史数据中或者特定监测系统中的可用过程数据。振动监测系统就是这样的例子。

如果监测关键故障的范围没有完全覆盖,就有可能需要考虑增加传感器或者考虑调整现有传感器的位置。

4.3 数据清洗和重采样

4.3.1 总则

为了构建一个稳健性强的模型,应首先收集信号验证所需要的、涵盖所有系统工况的数据,这些数据是已收集并储存的历史数据。事实上,由于一些诸如插值误差、随机数据误差、缺失数据、重要数据丢失、数据卡顿及其他等异常情况的发生,这些数据可能并不总是表示真实的设备状态。宜一直进行数据检查和纠正数据。

注意:删除数据前应慎重。

4.3.2 插值误差

使用历史数据训练模型时,通常遇到的第一个问题是可用的、处理过的数据与实际数据并不一致,但可以用数据归档程序中压缩的数据代替。通常,历史数据记录软件创建的数据库文件是时间序列数据库文件。然而,并非所有时刻采集的数据都存储。只有数据值的变化超过规定的量时才存储其数据并加上一个时间戳。这种方法只需要很少的存储空间,但却损失了数据的保真。当从历史数据中提取数据时,记录数据之间的数据值或者通过简单的线性插值计算,或者通过第二个数据点的步长计算。数据结果显得是锯齿时间序列,可能严重改变传感器之间的相关性。

总之,用于模型训练的数据宜是实际数据,允差的设定宜尽可能地小或者不设允差。

4.3.3 数据品质问题

几个最常见的数据品质问题是:

——数据缺失;

——噪声或随机数据;

——有缺陷的传感器导致的未更新或标定失效的数据;

——不合理的数据(超出范围)。

大多数这样的数据问题可以直观地识别,或者通过数据清洗实用程序监测到。这些实用程序可剔除坏的数据或者用某种算法得到最可能的数据替代坏数据。常见的是从训练数据集内删除所有的坏数据。大多数软件系统都包括自动数据清洗工具。这些工具软件能很容易地识别极度偏离期望值的数据,但是对期望运行域内的数据偏差通常不敏感。训练数据集内的坏数据点会使模型失效。

4.3.4 数据重采样

数据一旦清洗,可能需要在更低的速率下重采样,该速率由运行模式决定。这样,建议保留所有的时间戳以描述重要运行参数(如:机器停机)的瞬态特征。然而在稳态运行工况,每 10 min 采集一个数据样本(平均或不平均获得)可能就足够了。

4.4 模型开发

4.4.1 总则

模型开发不是简单的。需要完成几个步骤,包括:

- 选择相关特征；
 - 选择相关运行工况和训练数据；
 - 准备模型测试。
- 构建一个数据驱动模型需要：
- 关注于特定类型故障(机械的、电气的、热力的等)的一组参数(传感器)；
 - 已知机器处于良好健康状态期间的数据样本。

4.4.2 模型的定义和相关输入的选择

一旦数据的品质经过了检验,就应定义模型特征。特征可以是原始传感器数值本身,也可以是从传感器值(指数加权移动平均值、平均值、峭度值)导出的数值。一个过程装备有数百个参数需要监测用于关键设备评估。因此,这些参数会被分成一些小的关联组,以关注设备的特定功能(热力的、机械的、冷却的等)。

模型的性能会受到所含特征的强烈影响。多余的特征会诱发虚假报警或屏蔽真实事件,从而降低模型的性能。多余的特征可包含不变的或随机的特征。缺失重要的特征会使一些故障可能监测不到。

构建模型时为了获得精度和稳健性,记住选择功能方面的特征(对特定故障组的监测有用的参数)以及数字量方面的特征。建议在模型中采用相关联的特征,并考虑到独立的参数(如外部过程参数)能说明机器状态正常变化这样的事实。

4.4.3 相关工况和数据的选择

模型应使用涵盖所有工况的数据进行训练。由于工况是由系统结构、传感器值和作业程序界定的,所以不同设备的工况会有显著不同。

一个工况变化的例子是备用泵的定期使用,或者冗余泵的周期使用。为了模型能正确工作,对系统的每个工况都应进行训练。但是对某些不常用工况的过分训练会降低最常用工况下模型的性能。因此,某些设备的工况可能从不列入训练集合。

工况还会由于设备的维修而改变。这种情况下,模型应进行重新训练以考虑新的工况。

最后,工况也会由于一些周期性的现象而改变,如:季节变化。如果一个模型是用夏季收集的数据训练的,在冬季时气温和冷却水温等有了显著的不同,模型可能表现不好。另外,外界环境的异常变化也会影响模型的性能,例如:如果一个模型是用典型夏季数据训练出的,在异常热的夏天冷却水温更高,模型可能不能正确工作。这种情况下,新工况的数据应加到训练数据中。

谨记一些特殊工况如下:

- 宜注意机器维修的时间,因为维修会显著改变机器的性能;
- 有必要考虑维修时间间隔很长的机器,那表明机器性能会显著劣化;
- 构建模型时要注意当前运行工况之前的瞬态影响(例如:之前的功率或速度变化)。

4.4.4 模型测试准备

在识别了所有的典型工况之后,下一步就是选择输入数据并将数据划分为训练数据、检验数据以及测试数据。通常,用户没有优先的分类方法可选,因为其取决于监测工具。

训练和检验数据用于开发和测试生产设备的模型。使用带附加偏差的检验模型组成的测试集,来评估模型的监测性能。

4.5 模型性能评估

监测过程的下一步是选择和优化模型参数,这些参数对于不同的方法可能是不同的(例如:核函数宽度、最大簇直径、特征值的数量和隐层数量等)。

利用训练数据对模型进行训练,利用检验数据进行评估。选取检验数据上性能最好的模型参数。一旦最终的模型参数通过训练数据和检验数据选定后,就可以利用测试数据集评估模型的总体性能。重要的是测试数据不用于调整模型参数,而只是对模型进行总体性能的偏差评估。另外,当传感器有漂移时,模型的性能评估可通过依次给每一个传感器读数一个逐渐加大的偏差,并用模型的输出校正传感器故障的方法来完成。然后,利用有缺陷输入数据的预测值来确定模型的稳健性和溢出效应。稳健性可解释为当传感器由于某种故障导致数值不正确时,数据驱动模型做出正确的传感器预测的能力度量。溢出效应是有缺陷的传感器输入对其他传感器输出的影响。

4.6 报警设置

在一个成功的监测应用中,设置监测模型的预警值和报警值是一项关键的任务。设置模型的预警值是一项平衡模型漏报警和误报警的可能性和敏感度的行为。

设置预警值最常用的方法是选择一个从工程的角度看合理的数(阈值)。该设置值可基于单个参数的残差或者实际样本与训练模型之间的整体差距进行调整,或者基于两者进行调整。

此外,建议对开发的数据驱动模型性能进行量化表示,并根据模型的估计精度对预警设置值进行调整。

报警值定义的最常用方法是使用逻辑滤波规定在给定窗口内发生预警的数目。这类逻辑窗的应用过滤掉了出现大残差值的虚假点,有效地减少了预警的数目。

设定预警值和报警值的决定性要素是选择用哪个监视器的信号给出报警。通常,没必要对模型中所有的信号给出报警。总的来说,决策要基于工程判断,即给定信号对被监测设备健康状态的响应是如何表示的。

5 实施数据驱动诊断的过程

5.1 总则

数据驱动模型能根据模型设置产生一系列的带有相应预警和报警的残差数据流。因此,最终的步骤是利用残差识别观察到的预警和报警处的故障。

各种原因都能在参数残差中引起偏差。首先需要确认观察到的偏差不是因为模型的调整引起的。如果不是模型引起的,其他根本原因可能是仪表或设备故障引起。

目前的数据驱动监测经验通常依赖于人工判读过程。

许多设备劣化的途径是已知的可重复的事件。如果定义了这些途径的失效特征,则异常过程的判读就能自动完成。

诊断机器的状态能采用两个主要的方法,参见 ISO 13379-1 的描述。

——数据驱动的方法;

——基于知识的方法。

很多数据驱动的方法在 ISO 13379-1 有详细解释。下列段落给出两个简单数据驱动方法的应用:

——自动模式分类;

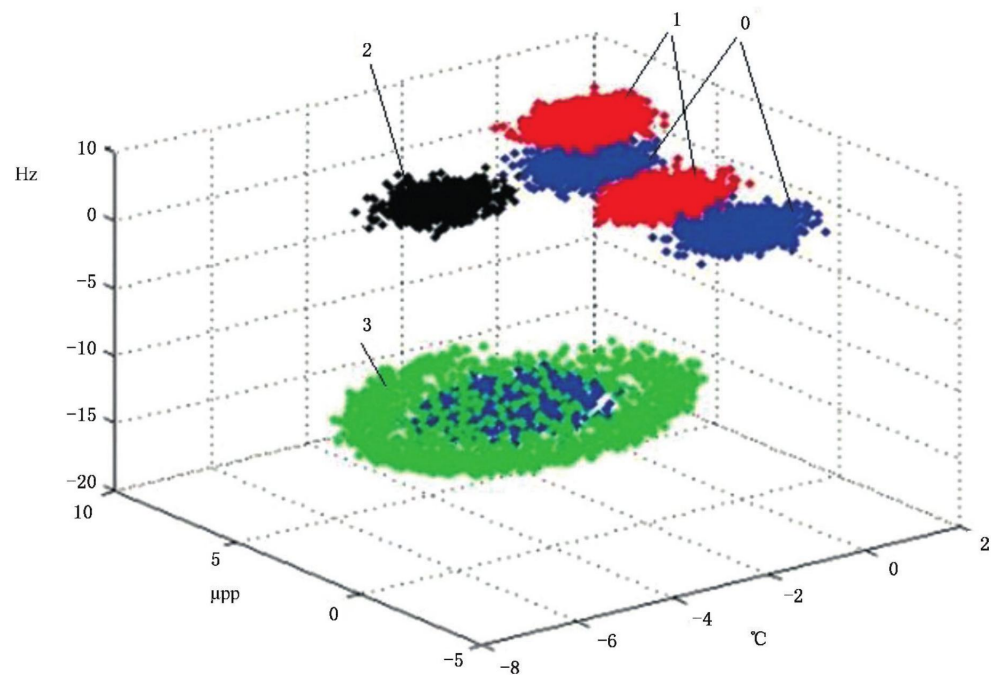
——简化的自动特征分类。

另一个实现自动诊断过程的简单方法是基于所有可用的存储信息(残差、绝对值、预警和报警阈值)编制规则。基于规则的诊断是基于知识诊断的一部分。

5.2 自动模式分类方法

模式分类的目标是根据先验的知识或者根据从模式中提取的统计信息对数据模式进行分类。被分类成的模式通常是一些测量组,在适当多维空间中的定义点群。

分类或描述方案通常基于一组已分类或描述模式的可用性,这组模式叫做训练集(见图 2),产生的学习方法特征是有监督的学习。



说明:
0——蓝色为健康状态;
1——红色为故障状态 1;
2——黑色为故障状态 2;
3——绿色为故障状态 3。

图 2 模式分类示例

模式可能是基于参数的原始值,也可能是基于数据驱动模型监测阶段处理后产生的残差数据。应清晰界定可以接受的区域边界,以避免错误诊断和允许自动诊断。残差数据的应用能使模型不受机器的初始状态的影响,而仅考虑数据流特征因监测到的故障而发生的演变。

对于具有稳定状态特征的故障,模式分类可能容易地实现。对于具有随时间变化特征的故障(磨损、摩擦、裂纹等)分类是可能的,但比较困难。

5.3 简化的自动特征分类方法

简单的自动分类方法可使用带有相应预警和报警的残差数据实现。

部件过程测量的报警列表相当于一组状态指示器,例如:参数 1 高(或正报警);参数 2 和参数 3 无警报;参数 4 低(或负报警)等。

使用这些信息的一种方法是把预警和报警列表转化为数值模式,如[1 0 0 -1]。这表示一种确定部件劣化原因的逻辑方法(见表 1)。

表 1 样本

六个传感器给出的故障特征	状态评估
1 0 0 -1 1 0	部件故障 No.1

表 1 (续)

六个传感器给出的故障特征	状态评估
-1 -1 0 -1 0 0	部件故障 No.2
0 0 0 1 1 1	部件故障 No.3
1 0 0 0 0 0	传感器故障 X1
0 1 0 0 0 0	传感器故障 X2
0 0 1 0 0 0	传感器故障 X3
0 0 0 1 0 0	传感器故障 X4
0 0 0 0 1 0	传感器故障 X5
0 0 0 0 0 1	传感器故障 X6
-1 0 0 0 0 0	传感器故障 X1
0 -1 0 0 0 0	传感器故障 X2
0 0 -1 0 0 0	传感器故障 X3
0 0 0 -1 0 0	传感器故障 X4
0 0 0 0 -1 0	传感器故障 X5
0 0 0 0 0 -1	传感器故障 X6

6 实施数据驱动监测方法的一般建议

为达到优化实施数据驱动监测和诊断方法的目的,宜:

- 列出待识别的严重故障和历史数据库或特定监测系统中的数据;
- 为了监测和诊断的目的,利用不超过 20 个描述符构建模型;
- 为模型中监测数据驱动方法集成相关的描述符;
- 为了监测的目的,清洗相应正常状态的历史数据;
- 为了诊断的目的,清洗相应正常和故障状态的历史数据;
- 为了数据的解释和重新训练,记录训练数据和监测组态;
- 验证不同故障对应的模式之间有清晰的分隔;
- 集成诊断模型中故障描述符的趋势值时,注意模型的稳健性;
- 为了监测的目的,设置合理的残差报警阈值(例如:三级阈值会产生很多报警)。

附 录 A
(资料性附录)
数据驱动监测应用示例

A.1 总则

本附录给出一个数据驱动监测应用的示例。

A.2 联合循环汽轮机的振动特性

运行过程中,这台汽轮机的轴承绝对振动达到异常水平,产生运行报警。

本示例研究的目的,是利用经验模型识别振动变化趋势是由于运行参数的变化引起的,还是由于汽轮机中间轴承的机械问题引起的。

本示例中考虑了几个输入量:有功功率、蒸汽输入流量和背压。输出参数是汽轮机两个中间轴承振动的均方根值,单位:mm/s。

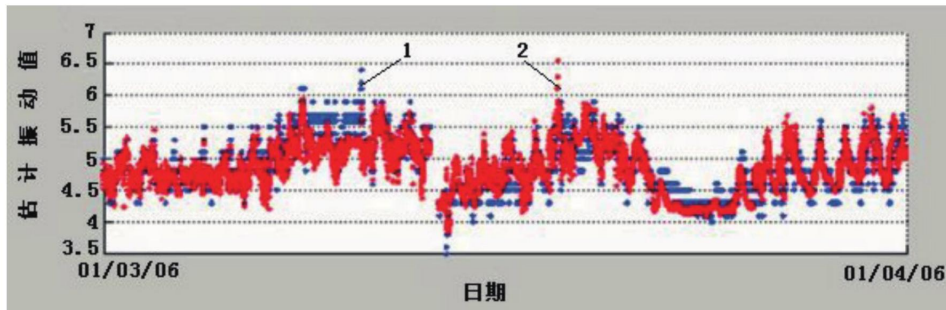
图 A.1 中用红色表示三个阶段(训练阶段、正常行为阶段和异常行为阶段)模型估计的振动值。

这里提出两个理由来解释这种振动特性的变化:

- a) 在上次停机过程中,整体结构刚度出现漂移或者改变,或者两者兼而有之;
- b) 以前从未达到过这么高的背压。

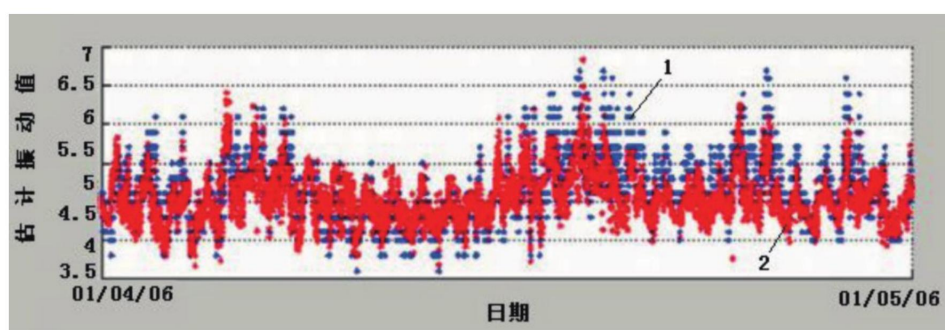
由图 A.1 的分析推断:

- 残差(红线和蓝线之间的差别)的微弱变化趋势,要经过数月的时间才能检测到;
- 变化趋势与运行参数的长期变化无关。

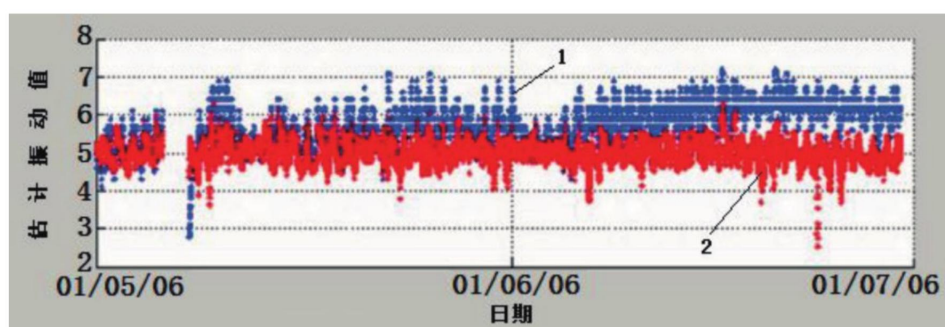


a) 训练阶段

图 A.1 三种汽轮机状态的估计振动值



b) 正常行为阶段



c) 异常行为阶段

说明：

1——蓝色为检测值；

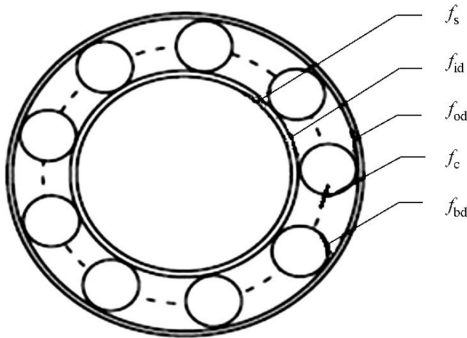
2——红色为预测值。

图 A.1 (续)

附录 B
(资料性附录)
数据驱动诊断应用示例

B.1 使用模式识别技术监测感应电动机轴承故障

感应电动机是许多工业过程的关键部件,常集成到商业应用的设备中。安全性、可靠性、效率和性能是感应电动机应用中最关心的一些问题。对于轴承故障,例如可利用 FFT 分析特征故障谐波幅值来监测。通过人工神经网络(ANNs)的模式识别技术可用于滚动轴承的故障诊断(图 B.1)。



说明:

- f_{bd} ——滚动体故障特征频率;
- f_c ——保持架故障特征频率;
- f_{id} ——内圈故障特征频率;
- f_{od} ——外圈故障特征频率;
- f_s ——轴转速频率。

图 B.1 轴承部件和故障特征频率

多层神经网络包括一个输入层、一个输出层和一个或多个隐层。

输入层连接谱线,输出层表示待诊断轴承的状态(健康,外圈故障/ f_{od} ,内圈故障/ f_{id} ,保持架故障/ f_c 和滚动体故障/ f_{bd})。重要的是记住故障症状与轴转速频率 f_s 有关。

人工神经网络(ANNs)用已知机器状态的实验数据子集训练。

神经网络一旦训练好,就可使用剩余的数据集来评估所推荐的这个方法的效率。

经过训练和评估后,人工神经网络(ANNs)模型就可以利用 FFT 的描述符(这里指频谱特征)进行自动诊断。

参 考 文 献

- [1] GB/T 22393—2015 机器状态监测与诊断 一般指南(ISO 17359:2011,IDT)
 - [2] EPRI Equipment Condition Assessment Modeling Guidelines, Report no. 1010038 December, 2005. Available at: www.epri.com
 - [3] EPRI Application of On-Line Monitoring Techniques to Equipment Condition Assessment-report no. 1010034 December 2005. Available at: www.epri.com
 - [4] HINES J.W., & DAVIS E. Lessons Learned From the U.S. Nuclear Power Plant On-Line Monitoring Programs, Prog. Nucl. Energy. 2005, 46(3-4) pp.176-189
 - [5] On-Line Monitoring for Calibration Extension: An Overview and Introduction to NUREG/CR-6895, “Technical Review of On-Line Monitoring Techniques for Performance Assessment.” Available at: <http://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr6895/overview/>
 - [6] Technical Review of On-Line Monitoring Techniques for Performance Assessment. State-of-the-Art (NUREG/CR-6895, Volume 1). Available at: <http://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr6895/v1/>
 - [7] Technical Review of On-Line Monitoring Techniques for Performance Assessment. Theoretical Issues (NUREG/CR-6895, ORNL/TM-2007/188, Volume 2). Available at: <http://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr6895/v2/>
 - [8] Technical Review of On-Line Monitoring Techniques for Performance Assessment: Limiting Case Studies (NUREG/CR-6895, ORNL/TM-2007/188, Volume 3). Available at: <http://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr6895/v3/>
-

中 华 人 民 共 和 国
国 家 标 准
机器状态监测与诊断 数据判读与诊断
技术 第2部分:数据驱动的应用
GB/T 22394.2—2021/ISO 13379-2:2015

*

中国标准出版社出版发行
北京市朝阳区和平里西街甲2号(100029)
北京市西城区三里河北街16号(100045)

网址:www.spc.org.cn

服务热线:400-168-0010

2021年5月第一版

*

书号:155066·1-67817

版权专有 侵权必究



GB/T 22394.2-2021



码上扫一扫 正版服务到