风洞设备PHM模型规范化管理研究

咨询报告

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | | | |
|  | | 编 写 |  |
| 校 对 |  |
| 审 核 |  |
| 批 准 |  |
|  | | | | |

目 录

[**1 前言** 5](#_Toc83732391)

[**2 PHM模型定义规范** 7](#_Toc83732392)

[2.1模型驱动的PHM模型 8](#_Toc83732393)

[2.1.1故障诊断模型 10](#_Toc83732394)

[2.1.2故障预测模型 12](#_Toc83732395)

[2.1.3健康状态评估模型 15](#_Toc83732396)

[2.2数据驱动的PHM模型 17](#_Toc83732397)

[2.2.1故障诊断模型 19](#_Toc83732398)

[2.2.2故障预测模型 25](#_Toc83732399)

[2.2.3健康状态评估模型 31](#_Toc83732400)

[2.3模型推荐 40](#_Toc83732401)

[**3 模型存储运维规范** 45](#_Toc83732402)

[3.1 模型接口与参数规范 45](#_Toc83732403)

[3.2 模型编码规范 46](#_Toc83732404)

[3.2.1 范围 46](#_Toc83732405)

[3.2.2 引用文件 47](#_Toc83732406)

[3.2.3 术语和定义 47](#_Toc83732407)

[3.2.4 编码要求 48](#_Toc83732408)

[3.2.5 PHM模型码编码规范 48](#_Toc83732409)

[3.3 模型存储规范 50](#_Toc83732410)

[3.3.1 模型存储流程 50](#_Toc83732411)

[3.3.2 模型数据备份与恢复 51](#_Toc83732412)

[3.4 模型调用规范 53](#_Toc83732413)

[3.4.1 API规范 53](#_Toc83732414)

[3.4.2 可视化控件建模规范 55](#_Toc83732415)

[3.5 模型运行规范 56](#_Toc83732416)

[3.6 模型性能评价规范 58](#_Toc83732417)

[3.6.1 分类型模型性能评价规范 59](#_Toc83732418)

[3.6.2回归型模型性能评价规范 63](#_Toc83732419)

[3.7 模型更新规范 64](#_Toc83732420)

[3.7.1 PHM模型更新流程 64](#_Toc83732421)

[3.7.2更新日志规范 72](#_Toc83732422)

[3.8 模型简介规范 74](#_Toc83732423)

[**4 平台PHM模型库技术要求** 76](#_Toc83732424)

[4.1 PHM模型库设计原则 76](#_Toc83732425)

[4.2 PHM模型库开发技术要求 76](#_Toc83732426)

[4.3 PHM模型库设计说明 76](#_Toc83732427)

[**参考文献** 84](#_Toc83732428)

[**附件1 风洞PHM模型管理软件技术规格书** 86](#_Toc83732429)

[F1 项目背景 86](#_Toc83732430)

[F2 建设目标 86](#_Toc83732431)

[F3 设计原则 86](#_Toc83732432)

[F3.1规范性 86](#_Toc83732433)

[F3.2稳定性 87](#_Toc83732434)

[F3.3易用性 87](#_Toc83732435)

[F3.4易维护性 87](#_Toc83732436)

[F3.5可扩展性 88](#_Toc83732437)

[F4 研制内容 88](#_Toc83732438)

[F5总体研制要求 88](#_Toc83732439)

[F5.1硬件部署要求 88](#_Toc83732440)

[F5.2系统总体设计方案 89](#_Toc83732441)

[F5.3.1 用户创建 89](#_Toc83732442)

[F5.3.2 用户管理 90](#_Toc83732443)

[F5.3.3 权限管理 90](#_Toc83732444)

[F5.3.4 PHM模型定义 90](#_Toc83732445)

[F5.3.5 PHM模型编码 91](#_Toc83732446)

[F5.3.6 PHM模型存储 91](#_Toc83732447)

[F5.3.7 PHM模型调用 92](#_Toc83732448)

[F5.3.8 PHM模型运行 92](#_Toc83732449)

[F5.3.9 PHM模型评估 92](#_Toc83732450)

[F5.3.10 PHM模型更新 92](#_Toc83732451)

1 前言

风洞作为空气动力装备的重要试验设施和设备，是发展高新飞行技术和先进飞行器的重要战略装备和基础。随着航空航天行业的飞速发展，风洞设备的规模和数量不断增加，风洞试验任务量也呈现出迅猛增长的趋势，所模拟的试验工况也变得日渐严苛和恶劣，导致风洞设备的可靠性和安全性面临极大的考验。为了保障风洞设备在运行过程中保持良好的性能状态，防止故障导致的意外事件发生，降低风洞设备的运维成本，就必须对风洞设备开展预测与健康管理（Prognostics and Health Management, PHM）工作。

为满足风洞设备的可靠性和安全性需求，提高试验装备运维保障工作效率，基地积极推动PHM技术的应用，先后在4米×3米风洞、2.4米跨声速风洞、Φ1米高超声速风洞等开展了单体风洞或单体关键部段的健康管理系统建设，并开发了对应的健康管理软件系统。与此同时，基地中众多在用与在建风洞都申请或即将开展各自的设备健康管理系统建设，开发工程量和成本日益增加。然而，基地目前采取的是“一座风洞、一个平台、一套软件”的独立分散建设运行模式，导致每个风洞系统拥有一套独立的健康管理系统，不同风洞设备健康管理系统之间PHM模型的接口与参数、存储管理、访问管理、运行管理、更新管理和性能评价等存在很大差异，使得PHM模型只能在特定风洞设备健康管理系统中正常有效地运行，导致系统中的模型复用性差、健康管理系统重复建设、系统开发成本重复投入和研发深度浮于表层的问题严重，模型难以在系统间共享使用而形成孤岛。

因此，针对PHM模型管理方面遇到的问题，基地拟开展PHM模型规范化管理研究。针对PHM业务建立及应用的各类模型特点，诸如物理模型、故障诊断/预测模型、状态评估模型等，开展模型管理研究，主要包括模型存储管理、模型更新管理、模型访问管理、模型评价、模型运行管理、模型接口和参数等的规范化要求等，同时提出平台PHM模型库技术要求，以指导模型库建设。

本咨询报告主要分为5章，主要对风洞设备PHM模型规范化研究的内容进行阐述，各章节主要内容如下：

第一章为前言，主要介绍基地现阶段风洞设备PHM模型管理的作用和意义。

第二章为需求分析，主要对基地现阶段所面临的PHM模型规范化管理问题进行剖析，进而对基地所提出的任务和需求进行分析。

第三章为PHM模型定义规范，通过对PHM模型进行分类的方式，以模型分类结果为主线，对不同类型的模型的输入输出参数、构建过程、模型特点、适用范围等进行规范。

第四章为模型存储使用运维规范，从模型的编码、建模、存储、调用、运行、性能评估、更新和简介的角度，开展PHM模型接口、模型存储管理、模型访问管理、模型运行管理、模型评价、模型更新管理研究。

第五章为平台PHM模型库技术要求，针对一体化平台模型集成的要求，从模型库设计原则、开发技术要求和设计说明三个方面，对PHM模型库设计和开发的技术要求进行规范。

2 PHM模型定义规范

模型定义规范是风洞设备PHM模型规范化管理研究的基础性环节，为后续PHM模型存储运维规范提供依据。本部分针对风洞设备健康管理系统中模型复用性差、重复建设问题严重的现状，通过制定模型的定义规范化模板的方式，对PHM模型的类别、输入输出参数、构建过程、适用范围等进行统一。

PHM模型种类繁多，导致模型调用、运行、存储等管理工作容易出现混淆。因此，在PHM模型定义规范制定之前，需了解各类PHM模型的特点，以分类的方式更加清晰地对模型进行归类管理。根据应用方法、算法原理和模型构建过程，PHM模型分为模型驱动和数据驱动两类。而PHM模型在风洞设备的实际应用中，根据PHM任务类型的不同，可分为故障诊断、故障预测和健康状态评估三种任务。不同的PHM任务类型，有着不同的功能边界和适用范围。下面将简要介绍三种PHM任务的基本概念。

故障诊断是对系统运行异常情况做出分析判断，并为维修决策提供信息支撑。要对系统进行故障诊断，首先必须对其进行检测，在发生系统故障时，对故障类型、故障部位及原因进行诊断，最终给出解决方案，实现故障恢复。

故障预测是对尚未发生故障的系统所出现的早期故障或退化征兆进行检测和识别，进而预测系统未来一段时间内将要发生的故障类型和剩余寿命。

健康状态评估是使用当前监测数据及历史运行数据，利用各种评估模型和算法对设备系统或子系统当前健康状态（指系统性能相较于期望的正常性能状态退化或偏离的程度）做出评估。

三种任务的主要区别和特点如表2-1所示。

表2-1 PHM任务的主要特点

|  |  |  |  |
| --- | --- | --- | --- |
| PHM任务 | 功能 | 适用范围 | 本质任务 |
| 故障诊断 | 故障诊断 | ①设备已发生故障 | 分类任务 |
| 故障预测 | 故障类型、位置预测 | ①设备未发生故障，且出现早期故障征兆  ②仅适用于退化型设备  ③具有故障样本数据 | 分类任务 |
| 剩余寿命预测 | 回归任务 |
| 健康状态评估 | 健康状态等级划分 | ①仅适用于退化型设备  ②具有故障和退化样本数据 | 回归或分类任务 |

开展以上三种任务所使用的模型不尽相同，模型的基本信息包括输入、输出参数等和需要的风洞设备监测数据也有差异。综上所述，本项目首先依据PHM模型的原理以及任务类型两个维度对PHM模型进行分类，划分得到模型驱动和数据驱动的故障预测、故障诊断和健康状态评估模型。通过对各种PHM模型的类别、输入输出参数、训练方法以及适用范围等进行总结，对模型的定义和创建过程进行规范，并基于风洞设备特点推荐合适的模型，以便后续PHM模型存储运维工作的开展，为构建一体化PHM模型库、提高模型复用性打下坚实的基础。

2.1模型驱动的PHM模型

模型驱动的PHM模型一般要求对象系统的数学模型是己知的，其基本思想是通过建立能够描述风洞设备的物理特性和失效模式的数学模型，从而实现对风洞设备的故障诊断、预测和健康状态评估。模型驱动的方法进一步可以分为基于物理模型的方法和基于滤波器的方法。

基于物理的PHM模型研究的是风洞设备内部机理，揭示的是其输入输出以及某些内部参数之间的关系。它使用数学的方法对风洞设备的物理失效机理进行表述，从物理机理出发来深入研究和理解部件中典型故障模式的特性。基于物理的PHM模型，其基本构建流程如图2-1所示。

基于滤波器的PHM模型先建立风洞设备的测量模型，通过对含有噪声的观测信号的处理，利用目标的动态信息，设法去掉噪声的影响，得到一个估计值用于一步或多步预测。基于滤波器的PHM模型，其基本构建流程如图2-2所示。

除上述模型外，近年来PHM领域中基于数字孪生特征的模型逐渐崭露头角。数字孪生是以数字化方式创建物理实体的虚拟模型，借助数据模拟物理实体在现实环境中的行为，通过虚实交互反馈、数据融合分析、决策迭代优化等手段，为物理实体增加或扩展新的能力。数据孪生技术驱动下的PHM系统将各类传感器采集的数据实时感知并传输到虚拟模型中。虚拟模型输入真实物理设备的数据，进行同步工作的仿真模拟，通过与历史数据库全方位比对及算法模拟进行故障预测等工作。

模型驱动的方法主要有以下特点：

（1）能够深入设备本质研究其性质，实现实时故障预测、诊断和健康状态评估，精度比较高。

（2）需要提前了解设备的结构以及退化机理，需要大量的专业知识，尤其是对专家系统过于依赖，对其故障进行物理的建模依赖较为专业的系统知识。

（3）针对复杂的设备，建立和抽象物理模型很难实现。

在工程实践中，多数设备均具有复杂结构，且内部退化机理不明确，内部结构的关联难以完全掌握，构建有效反映设备故障特性的数学模型往往无法实现。随着系统复杂度的增加，以及内部关联度指数倍的提升，模型驱动的方法受到极大的限制。同时，针对不同的任务，调研发现可供研究的模型驱动的PHM模型数量有限。因此，本部分基于PHM任务类型对相关模型作简要定义规范研究，其中PHM任务主要包括故障诊断、故障预测和健康状态评估。



图2-1 基于物理模型的PHM模型基本构建流程



图2-2 基于滤波器的PHM模型基本构建流程

2.1.1故障诊断模型

由于前文说明的局限性，在模型驱动的故障诊断模型中，本部分仅针对未知输入观测器和粒子滤波两类模型，总结各自的适用范围、需求数据特点、原理、模型特点、功能边界以及输入输出，如表2-2所示。

表2-2 模型驱动的故障诊断模型

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 模型类型 | 适用范围 | 需求数据特点 | 原理 | 模型特点 | 功能边界 | 输入输出 |
| 未知输入观测器 | ①含未知干扰  ②非线性系统 | 含干扰、建模误差等不确定因素 | 利用设备可监测的输入输出信号，设计状态观测器，通过产生对故障有指示作用的残差信号，实现故障诊断 | ①实现方式简单灵活  ②响应速度快 | 无法处理线性系统 | 输入：观测器状态向量以及设备系统的输入和输出向量  输出：残差信号，通过与阈值比较进行诊断 |
| 粒子滤波 | ①非线性系统 | 大样本 | 在状态空间抽取一组带权值的粒子集，使用这组粒子集来近似逼近状态的后验概率分布函数 | ①对系统的过程噪声和测量噪声没有任何限制  ②采用蒙特卡洛思想 | 无法处理线性系统 | 输入：系统状态向量的经验分布（处于各个状态的先验概率）  输出：状态估计值，进而判断是否故障状态 |

具体地，针对未知输入观测器和粒子滤波两类模型，详细的模型介绍和构建流程规范如下：

（1）未知输入观测器

①模型简介：通过将设备模型中的干扰、建模误差等不确定因素作为未知输入进行处理，利用干扰解耦方法设计观测器，生成对于未知输入等不确定项不敏感却对故障量非常敏感的残差信号，实现故障诊断。

②模型构建：如图2-3，首先确定未知输入观测器状态方程，然后计算观测器参数，得到状态估计值以及残差信号，从而判断设备是否故障。



图2-3 未知输入观测器构建过程

a.确定观测器状态方程：利用设备系统的输入和输出向量作为驱动，通过设计系数矩阵，使得未知输入观测器给出状态向量的一个线性组合的估计。

b.计算观测器参数：基于线性矩阵不等式等计算观测器参数矩阵；

c.状态估计：通过状态方程产生状态估计值；

d.残差评估：选择合适的阈值，当残差信号超过此阈值时则诊断系统出现故障，否则认为系统处于正常运行状态。

（2）粒子滤波

①模型简介：粒子滤波是一种优化算法，通过寻找一组在状态空间中传播的随机样本来近似地表示概率密度函数，用样本均值代替积分运算，进而获得系统状态的最小方差估计。

②模型构建：如图2-4，首先初始化粒子，然后重复进行粒子权值更新、重采样以及状态估计三个步骤。

a.粒子初始化：从先验分布中抽取*N*个粒子，各粒子权值初始化为1/*N*；

b.更新粒子权值：基于重要性分布函数抽取*N*个粒子，计算权值并归一化；

c.重采样：根据有效粒子数判断是否重采样而得到新的粒子集；

d.状态估计：利用当前粒子集权重加权求和得到状态估计值；

e.迭代更新：判断是否结束，是则退出，得到状态估计值；否则转到步骤b。



图2-4 粒子滤波构建过程

2.1.2故障预测模型

在模型驱动的故障预测模型中，本部分仅针对卡尔曼滤波和粒子滤波两类模型，总结各自的适用范围、需求数据特点、原理、模型特点、功能边界以及输入输出，如表2-3所示。

具体地，针对卡尔曼滤波和粒子滤波两类模型，详细的模型介绍和构建流程规范如下：

（1）卡尔曼滤波

①模型简介：卡尔曼滤波通过采用递推线性最小方差估计方法，利用设备可观测的输出估计误差去修复不可观测的状态估计误差，从而消除数据流中的噪声干扰，提高状态量的估计精度。

②模型构建：如图2-5，根据建立的设备状态方程和量测方程，基于故障特征参数的历史数据序列依次进行状态变量的估计与预测，然后计算剩余寿命。

a.建立状态方程和量测方程：以故障特征参数真实值作为状态变量，以采集到的故障特征参数值作为测量值，离散化设备的性能退化模型，得到系统的状态方程和测量方程；

b.获取故障特征参数的历史数据序列：由仿真或实验得到故障特征参数随时间的变化值；

c.估计当前时刻状态变量：根据已知的当前时刻的故障特征参数，运用卡尔曼滤波递推算法来估计当前时刻的状态变量；

d.预测状态变量：把从步骤c获取的状态变量初始值估计值代入到迭代方程中，从而计算出未来时刻的状态变量；

表2-3 模型驱动的故障预测模型

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 模型类型 | 适用范围 | 需求数据特点 | 原理 | 模型特点 | 功能边界 | 输入输出 |
| 卡尔曼滤波 | ①线性系统  ②噪声服从高斯分布 | 时序数据 | 过滤设备的噪声影响，利用过程噪声、测量噪声来推算设备的故障状态 | ①采用状态空间概念描述对象结构  ②通过递推计算，具有自适应特性 | 无法处理非线性系统 | 输入：故障特征参数时间序列  输出：设备剩余寿命 |
| 粒子滤波 | ①中长期预测  ②非线性系统 | 大样本 | 在状态空间抽取一组带权值的粒子集，使用这组粒子集来近似逼近状态的后验概率分布函数 | ①对系统的过程噪声和测量噪声没有任何限制  ②采用蒙特卡洛思想 | 无法处理线性系统 | 输入：退化特征时间序列观测值  输出：设备故障时间分布 |



图2-5 卡尔曼滤波构建过程

e.剩余寿命计算：当状态变量预测值快要达到失效阈值时，发出故障警报，根据设定的失效阈值预测当前时刻设备的剩余寿命。

（2）粒子滤波

①模型简介：粒子滤波是一种优化算法，通过寻找一组在状态空间中传播的随机样本来近似地表示概率密度函数，用样本均值代替积分运算，进而获得系统状态的最小方差估计。

②模型构建：如图2-6，根据已有的观测值，通过对既定的预测模型参数进行迭代估计，然后将各参数的估计值代入预测模型，外推预测时间序列的演化情况。



图2-6 粒子滤波构建过程

a.确定预测模型：根据对历史数据拟合误差的比较选择，或者可以直接利用已有的经验模型；

b.粒子样本初始化：由先验概率随机产生粒子群；

c.模型参数估计：利用状态转移方程估计下一时刻粒子状态，根据该时刻的新观测值，计算粒子样本与观测值的似然概率密度函数，并以此更新粒子权值且归一化，则可得当前时刻未知参数的最小均方估计；未达到观测阈值前，根据上一步中粒子权值大小进行重采样，得到新的粒子样本集，重复该估计步骤；

d.预测与结果统计：根据所有粒子在指定时刻的状态，统计得到该时刻的系统状态概率密度分布，根据所有粒子的失效时间，统计得到系统失效时间的概率密度分布。

2.1.3健康状态评估模型

相对故障诊断和预测而言，关于模型驱动的健康状态评估的研究相对更少，本部分仅针对卡尔曼滤波与扩展卡尔曼滤波，总结各自的适用范围、需求数据特点、原理、模型特点、功能边界以及输入输出，如表3-4所示。

具体地，卡尔曼滤波与扩展卡尔曼滤波的模型介绍和构建流程规范如下：

（1）卡尔曼滤波

①模型简介：详见2.1.2。

②模型构建：如图2-7，依次按照4个递推步骤进行更新，可由当前时刻递推得到下一时刻状态向量及其误差协方差、卡尔曼增益的更新值，实现不同时刻状态向量的估计。



图2-7 卡尔曼滤波构建过程

a.状态估计的更新：根据当前时刻的状态向量后验估计值递推下一时刻状态向量先验估计值；

b.状态估计误差协方差的更新：根据当前时刻状态向量后验估计误差的协方差递推下一时刻状态向量先验估计误差的协方差；

c.卡尔曼增益的观测更新：基于上述状态估计及其误差协方差的时间更新，对卡尔曼增益进行更新；

d.状态估计值的观测更新：根据下一时刻观测值、卡尔曼增益、状态向量先验估计值以及控制向量，得到下一时刻状态向量的后验估计值，也即状态向量的最优估计值；

表2-4 模型驱动的健康状态评估模型

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 模型类型 | 适用范围 | 需求数据特点 | 原理 | 模型特点 | 功能边界 | 输入输出 |
| 卡尔曼滤波 | ①线性系统  ②噪声服从高斯分布 | 时序数据 | 过滤设备的噪声影响，利用过程噪声、测量噪声来推算设备的故障状态 | ①采用状态空间概念描述对象结构  ②通过递推计算，具有自适应特性 | 无法处理非线性系统 | 输入：状态观测值以及状态先验分布  输出：状态估计值 |
| 扩展卡尔曼滤波 | ①非线性系统 | 时序数据 | 将非线性问题转换为能用卡尔曼滤波处理的线性问题 | ①易于实现  ②相对卡尔曼滤波较为复杂 | 要严格定义状态方程 | 输入：状态观测值以及状态先验分布  输出：状态估计值 |

（2）扩展卡尔曼滤波

①模型简介：卡尔曼滤波主要用于估计线性时变系统，而当被估计的系统模型是非线性的时候，需对非线性状态空间模型进行线性化处理，然后再采用卡尔曼滤波算法进行状态量的估计，这种改进后的模型被称作扩展卡尔曼滤波。

②模型构建：如图2-8，将非线性的状态空间模型进行线性化，然后采用基本卡尔曼滤波进行实现。



图2-8 卡尔曼滤波构建过程

a.非线性函数线性化：求取状态方程和测量方程的偏导，得到相应的泰勒级数展开式，将非线性映射函数进行线性化，获得对应于卡尔曼滤波迭代计算中所对应的系数矩阵；

b.构建卡尔曼滤波：按照图2-7，实现状态变量的最优估算。

2.2数据驱动的PHM模型

数据驱动的PHM模型核心是对可利用的历史数据或实时监测数据进行特征提取，并转化和表征为能够反映风洞设备性能变化的信息，以采集的数据为研究基础，采用先进的数据处理与分析软件和方法对数据进行深入挖掘，找出风洞设备监测数据中隐含的健康状态或退化特征信息。数据驱动的方法进一步可以分为基于统计分析的方法和基于人工智能的方法。

基于统计分析的PHM模型由于其在获取准确全面数据资源的前提下，通过对风洞设备健康状态参数的监测数据进行统计分析，描述数据输出关系和相关参数以用于风洞设备的故障诊断、预测与健康状态评估。基于统计分析的PHM模型，其基本构建流程如图2-9所示。基于人工智能的PHM模型首先是选取若干历史数据作为训练样本，然后用某种机器学习算法（如支持向量机、卷积神经网络、循环神经网络、长短期记忆网络等）对已采集的历史数据进行训练，使其满足一定的精度后再基于当前监测数据进行风洞设备的故障预测、诊断和健康状态评估。基于人工智能的PHM模型，其构建流程如图2-10所示。

如前所述，关于两类数据驱动的PHM模型，不同的PHM任务可以使用不同的PHM模型，相同的模型可以应用于不同的任务中但其构建和应用过程有差异。因此，本部分基于PHM任务类型开展相关模型定义规范研究，其中PHM任务主要包括故障诊断、故障预测和健康状态评估。



图2-9 基于统计分析的PHM模型基本构建流程



图2-10 基于人工智能的PHM模型基本构建流程

2.2.1故障诊断模型

数据驱动的故障诊断模型按照不同的原理可以继续划分为统计分析模型和人工智能模型，这两类模型的适用范围、需求数据特点、原理、模型特点、功能边界以及输入输出存在一定的差异性，为方便这些模型的存储运维，对其进行总结如表2-5所示。

表2-5 数据驱动的故障诊断模型

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 模型类型 | | 代表模型 | 适用范围 | 需求数据特点 | 原理 | 模型特点 | 功能边界 | 输入输出 |
| 统计分析模型 | 信号  分析 | 小波变换 | 全系统适用 | 无特殊要求 | 利用各种信号分析技术提取信号时域和频域的特征 | ①简单易行，便于掌握  ②通常与其他数据驱动的方法相结合  ③只适用于信号与设备故障相关的场合 | —— | 输入：关键时域信号  输出：频域信号 |
| 经验模态分解 | 时频域数据 | —— | 输入：时频域特征信号或振动信号  输出：内涵模态分量 |
| 其他  统计  模型 | 高斯过程回归 | ①恒定工况  ②退化型设备 | 无标签  非小样本数据 | 从过去故障历史数据的统计特性角度进行故障诊断 | ①参数设置比神经网络简单  ②需要合适的核函数 | 无法处理小样本数据 | 输入：设备性能退化数据，包括时间与对应的性能退化量  输出：预测点的预测分布估计（均值、方差） |
| 隐马尔可夫模型 | ①恒定工况  ②设备运行状态的驻留时间近似服从指数分布更好  ③状态间的转移概率矩阵不随当前状态大幅变化 | 低噪声 | ①适用于动态建模能力要求更高的机械故障诊断领域  ②适用于故障种类较少且区分明显的简单情况 | 无法处理含噪声数据 | 输入：状态观测序列（等间距时刻对应的设备状态观测值）  输出：最大似然率对应的状态 |
| 人工智能模型 | | 支持向量机 | ①小样本 | 时序、非线性、大样本数据，需平滑降噪处理 | 通过对己知故障数据的训练，建立输入输出之间的映射关系进行诊断 | ①避免复杂的数学推导和专家知识  ②适合非线性复杂系统  ③依赖于数据的完整性和真实性  ④通常需要大量数据，网络训练时间较长 | 不宜处理大样本数据 | 输入：带标签的训练数据集（如各故障类型下的设备振动信号）或设备运行特征数据  输出：设备故障诊断结果，故障或正常 |
| BP神经网络  循环神经网络  深度置信网络  卷积神经网络 | ①恒定工况或复杂工况  ②大样本 | 无法处理小样本数据 |

具体地，针对几种典型的故障诊断模型，详细的模型介绍和构建流程规范如下：

（1）信号分析

一般而言，信号分析通常与其他数据驱动的方式相结合，因为通过传感器所采集到的信号会掺杂很多无用的信号，故很难直接通过信号对设备的故障原因进行判断。因此需要对设备的信号进行进一步的处理，如与人工智能或与统计分析相结合，以此来发现设备更加明显的特征，使对设备的诊断更加智能化和准确化，本部分介绍两种常用方法。

①小波变换

方法简介：小波变换在某种程度上也是一种窗函数的分析方法，相较于傅里叶变换的窗函数形状是固定的，小波变换中窗函数的面积是固定的，但是窗函数的形状却可以根据信号而做出改变，从而具有了更高的分辨率。在信号的低频部分能够有更高的频率分辨率，而在高频部分有更高的时间分辨率，时频分析中，传统的傅里叶变换使用的地方都能够用小波变换进行取代，可见小波变换相对于傅里叶变换具有很好的优越性。对于短时傅里叶变换在时间分辨率和频率分辨率方面的问题上，小波变换在一定的条件下能够很好地进行兼顾，让变换的结果在这两者之间保持很好的平衡。所以小波变换在特征提取和异常检测等方面都有了很好的应用。

方法原理：小波变换是傅里叶变换基础上的改进。傅里叶变换的把无限长的三角函数作为基函数：



这个基函数会伸缩、会平移（其实本质并非平移，而是两个正交基的分解）。缩得窄，对应高频；伸得宽，对应低频。然后，该基函数不断与信号做相乘。某一个尺度（宽窄）下乘出来的结果，可以理解成信号所包含的当前尺度对应频率成分有多少。于是，基函数会在某些尺度下，与信号相乘得到一个很大的值，因为此时二者有一种重合关系，那么就可以得到信号包含该频率的成分的多少。

如上述，小波做的改变就在于，将无限长的三角函数基换成了有限长的会衰减的小波基：



从公式可以看出，不同于傅里叶变换，变量只有频率，小波变换有两个变量：尺度（scale）和平移量（translation）。尺度控制小波函数的伸缩，平移量控制小波函数的平移。尺度对应于频率（反比），平移量对应于时间。

当伸缩、平移到这么一种重合情况时，也会相乘得到一个大的值。这时候和傅里叶变换不同的是，这不仅可以知道信号有这样频率的成分，而且知道它在时域上存在的具体位置。而当在每个尺度下都平移着和信号乘过一遍后，可以得到信号在每个位置都包含哪些频率成分。

②经验模态分解

方法简介：经验模态分解（EMD）是一种新型自适应信号时频处理方法，特别适用于非线性非平稳信号的分析处理。该方法是依据数据自身的时间尺度特征来进行信号分解，无须预先设定任何基函数。这一点与建立在先验性的谐波基函数和小波基函数上的傅里叶分解与小波分解方法具有本质性的差别。正是由于这样的特点，EMD方法在理论上可以应用于任何类型的信号的分解，因而在处理非平稳及非线性数据上，具有非常明显的优势,适合于分析非线性、非平稳信号序列，具有很高的信噪比。

方法原理：经验模态分解的特点在于自适应的基函数（基波），使得它可以处理任意信号，内涵模态分量（IMF）是原始信号被EMD分解之后得到的各层信号分量。其具体原理步骤如下所示：

a.根据原始信号上下极值点，分别画出上、下包络线；

b.求上、下包络线的均值，画出均值包络线；

c.原始信号减均值包络线，得到中间信号；

d.判断该中间信号是否满足IMF的两个条件，如果满足，该信号就是一个IMF分量；如果不是，以该信号为基础，重新做1~4的分析。IMF分量的获取通常需要若干次的迭代。

用上述方法得到第一个IMF后，用原始信号减IMF1，作为新的原始信号，再通过1~4的分析，可以得到IMF2，以此类推，完成EMD分解。

（2）高斯过程回归

①模型简介：高斯过程（GPR）是一种基于定量知识的数据驱动方法，对于无法建立精确系统模型的非动态系统能够保持良好的计算精度，可保证故障诊断的准确性。高斯过程回归分析作为一种常用的统计学习算法，且其参数设置问题

相比神经网络简单，这使得高斯过程回归能更好适用于工业现场。

②模型构建：如图2-11，首先确定核函数与先验模型，然后优化超参数以测试模型。



图2-11 高斯过程回归构建过程

a.确定核函数与先验模型：选取合适的核函数，并设置超参数的初始值，从而概率分布的形式确定GPR的先验模型；

b.优化超参数：通过训练学习样本，将先验模型转换为后验模型，同时对核函数的超参数进行优化，获得最优超参数；

c.模型测试：将测试样本应用到优化后的GPR后验模型中。

（3）隐马尔可夫模型

①模型简介：隐马尔可夫模型（HMM）由一个隐藏的不能观测的随机过程（马尔可夫链）和一个实际可观测的随机序列过程构成，记为，为初始状态概率分布，为状态转移概率矩阵，为观测概率矩阵。

②模型构建：基于多组设备状态观测序列，使用参数估计算法（如Baum-Welch算法）估计出多个，基于Viterbi算法计算各下的似然率。

（4）卷积神经网络

①模型简介：卷积神经网络（CNN）属于多级神经网络，从作用上，将其划分为滤波级和分类级。滤波级将输入特征信号进行特征提取，包括卷积层、激活层以及池化层共3层基本单元或者其他一些优化单元。分类级主要是对特征进行分类，一般由全连接层组成。CNN模型的网络参数是通过两个层级共同训练得到。一般CNN结构主要由输入层、卷积层、池化层、全连接层和输出层组成。

②模型构建：如图3-12，首先进行特征提取与降采样，然后更新参数以测试模型。



图2-12 一维卷积神经网络构建过程

a.特征提取：输入信号经过卷积操作后利用ELU函数激活成一组特征图；

b.降采样：利用最大池化方法进行降采样，在全连接层将提取的深层特征展平并连接；

c.更新参数：使用Adam优化器反复迭代以更新网络参数；

d.模型测试：测试数据代入训练完成的模型。

（5）BP神经网络

①模型简介：BP神经网络，是一种按照误差逆向传播算法训练的多层前馈神经网络，是应用最广泛的神经网络。其能够自适应、自主学习，根据预设的参数更新规则，不断地调整神经网络中的参数，以达到最符合期望的输出。

②模型构建：如图2-13，首先进行网络初始化与样本数据计算，然后训练模型并测试。



图2-13 BP神经网络构建过程

a.网络初始化：对BP神经网络的一些初始值进行赋值，如初始权重、学习率、目标误差、迭代次数等；

b.样本数据计算：取一组样本数据，进行正向输入，计算各层神经元输入与输出，计算误差，并将误差反向传播，直到所有学习数据用尽；

c.模型训练：计算网络平均误差，检测误差是否达到目标允许误差范围，若达到，结束训练；未达到允许误差范围，则整个流程进行迭代，直到达到迭代上限，结束训练；

d.模型测试：输入新的设备运行数据，利用训练好的BP神经网络对设备进行故障诊断并输出结果。

（6）循环神经网络

①模型简介：循环神经网络（RNN）是一类以序列数据为输入，在序列的演进方向进行递归且所有节点（循环单元）按链式连接的递归神经网络。循环神经网络具有记忆性、参数共享并且图灵完备（Turing Completeness），因此在对序列的非线性特征进行学习时具有一定优势。

②模型构建：如图2-14，首先通过输入数据更新网络参数，检测模型精度后开展测试。



图2-14 循环神经网络构建过程

a.数据输入：以时间为收集轨道，按照一定的时间步长收集出连续的设备运行特征数据，作为算法的直接输入；

b.更新网络参数：小批量数据向前传播，误差反向传播并更新网络参数，直到单个周期所有数据迭代完成，迭代该过程直到达到总迭代次数，保存训练好的网络模型；

c.检测精度：将训练样本添加噪声，输出诊断结果用以测试网络训练精度；

d.模型测试：利用训练好的RNN网络模型对新的设备数据进行分析，输出设备故障诊断结果。

2.2.2故障预测模型

数据驱动的故障预测模型按照不同的原理可以继续划分为统计分析模型和人工智能模型，这两类模型的适用范围、需求数据特点、原理、模型特点、功能边界以及输入输出存在一定的差异性，为方便这些模型的存储运维，对其进行总结如表2-6所示。

具体地，针对几种典型的故障预测模型，详细的模型介绍和构建流程规范如下：

（1）自回归模型

①模型简介：自回归模型是一种时序模型，不仅可以揭示动态数据的规律，预测其未来值，而且还能够从多方面研究系统的有关特性。

②模型构建：如图2-15，通过模型识别与定阶确定自回归模型，然后基于参数估计的结果进行模型评估、检验与预测。

表2-6 数据驱动的故障预测模型

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 模型类型 | | 代表模型 | 适用范围 | 需求数据特点 | 原理 | 模型特点 | 功能边界 | 输入输出 |
| 统计分析模型 | 时间  序列 | 自回归模型 | ①恒定工况  ②短时间序列下的故障预测  ③退化型设备 | 平稳性时间序列数据，对于非平稳的时间序列，需要进行差分处理 | 将预测对象的历史数据按照一定的时间间隔进行排列，构成随时间变化的统计序列 | ①一种回归模型  ②简单易行，便于掌握，计算速度快 | 不便解决高度离散程度数据的预测问题 | 输入：随时间变化的观测信号值  输出：预测时刻的观测信号值 |
| 可靠性模型 | 比例风险模型 | ①恒定工况  ②设备具有硬故障失效模式 | 故障集合、测  点集合以及  其逻辑关系 | 从过去故障历史数据的统计特性角度进行故障预测 | ①体现了同类设备的共性属性和服役设备的个体差异  ②需要大量样本可靠性试验历史数据估计参数  ③需要失效率差异与协变量差异的比例，即比例性假设 | 适用于不随工况变化明显的数据 | 输入：设备故障数据  输出：设备剩余寿命 |
| 随机  过程 | 维纳过程  伽玛过程  逆高斯过程 | ①恒定工况（变工况需要引入协变量）  ②退化型设备 | 退化量变化规律与一种随机过程相符 | 利用随机过程相关知识，建立随机过程模型来描述退化轨迹以得到概率框架下的剩余寿命概率分布 | ①适合描述当前设备失效产生机制和运行环境的时变不确定性  ②便于统计分析 | 无法处理小样本数据 | 输入：设备性能退化数据，包括时间与对应的性能退化量  输出：设备剩余寿命 |
| 人工智能模型 | | 支持向量机 | ①小样本 | 时序、非线性、大样本数据，需平滑降噪处理 | 通过对己知故障数据的训练，建立输入输出之间的映射关系进行预测 | ①避免复杂的数学推导和专家知识  ②可以进行多参数、多步预测，动态自适应能力强  ③适合非线性复杂系统  ④依赖于数据的完整性和真实性  ⑤通常需要大量数据，网络训练时间较长 | 不宜处理大样本数据 | 输入：带标签的训练数据集（如各故障类型下的设备振动信号）或设备运行特征数据  输出：设备故障预测结果 |
| BP神经网络  循环神经网络  深度置信网络  卷积神经网络 | ①恒定工况或复杂工况  ②大样本 | 无法处理小样本数据 |



图2-15 自回归模型构建过程

a.模型识别：对已处理的平稳时间序列，通过其自相关函数和偏相1.关函数的统计性来初步判断该序列是否适合自回归模型。设样本的自相关函数为，可被定义为如下：



设样本的偏相关函数列为，若时间序列的自相关函数被负指数函数控制收敛到零，且偏相关函数在指定步数内截尾，则可以判断该平稳时间序列是可以被自回归的；

b.模型定阶：对于平稳随机序列的模型拟合来说，模型阶数估计是一个重要问题，阶数选择不同，会严重影响到模型拟合结果的准确性。针对上述提到的自回归分析方法的偏相关系数的截尾性，可以利用偏相关函数何时截尾作为阶数的初步参考值，再结合相应的方法确定出模型的精确阶数。在实际的应用中，常用准则函数定阶法来确定模型的阶数，该放方法综合考虑了对模型的拟合精度和模型参数的个数，在保证拟合结果的基础上，极大缩短了计算时间。常用BCI准则函数被定义如下：



其中，为模型的阶数，为时间序列的长度，为采用不同阶数时的模型残差方差的极大似然估计，的最小值点即为确定的阶数；

c.参数估计：对于P阶自回归模型，已知观测样本，先计算样本的自协方差函数估计值，求出回归系数的估计值回代到模型中,再求样本自协方差函数，随后解得所有参数值和；

d.模型评估与检验：在参数估计之后，通过检查模型的假设是否满足所给条件来检验模型是否恰当，如不恰当则需返回前一步骤，重新选定模型。在估计完成模型的参数后，可以得到模型的残差序列，即为白噪声序列。假设残差序列为白噪声序列，设为的自相关函数，当样本个数扩充的时候，是互相不关联的，且近似于正态分布，即：



因此，如果则认为在0.05的显著性水平下，是独立的。

e.模型预测：用序列在时刻对时的观察值进行预测。该预测以为原点，向步长为预测，并以的条件期望作为的预测值，即：



（2）比例风险模型

①模型简介：比例风险（PH）模型由两部分组成：第一部分为基准风险率函数，用来描述设备的运行时间或使用寿命对硬故障风险率函数的影响；第二部分为一个恒为正值的连接函数，用来将监测得到的退化信息融入到针对硬故障风险率函数的评估中，二者之间一般通过乘性的关系结合起来。

②模型构建：如图2-16，首先构建威布尔比例风险模型，然后基于参数估计的结果进行寿命预测。



图2-16 比例风险模型构建过程

a.构建威布尔比例风险模型：综合考虑设备的役龄时间与当前时刻的状态信息，通过历史数据与退化信息建立比例风险模型如下：



其中，表示设备的工作时间，表示失效率，表示在时刻的协变量所构成的列向量，为协变量所对应的回归参数所构成的行向量，为只和时间有关的基底失效率，这里取基低函数为最常用的威布尔分布；

b.参数估计：需要应用可靠性理论对模型中的未知参数进行估计，如基于可靠性的参数估计方法主要有矩估计法和极大似然估计法等。极大似然估计在处理截尾数据等不完全数据方面具有很大的优势。因此，采用极大似然估计方法对比例风险模型进行参数估计，极大似然估计函数如下:



式中，为失效轴承样本个数，为尚未失效的轴承样本截尾个数。将可靠度函数和失效概率密度函数代入上式，则上述似然函数公式可化简为:



式中，为轴承失效时间或截尾时间，其对数函数如下：



将个样本数据代入上式，分别对参数求偏导，并令偏导数为零，采用牛顿迭代法、单纯形等方法进行求解，从而建立威布尔比例风险模型；

c.寿命预测：采用基于差分的ARMA趋势预测方法对设备的失效率曲线进行趋势预测，并由可靠度的变化曲线设定失效阈值，得到设备的剩余寿命；

（3）随机过程

①模型简介：随机过程是一系列随机变量在时间轴上的组合，各个时刻的值均是随机变量。

②模型构建：如图2-17，首先进行模型选择，然后基于参数估计的结果进行寿命预测。

a.模型选择：选择合适的模型（如维纳过程、伽玛过程和逆高斯过程），描述退化数据变化规律；

b.参数估计：基于样本数据使用极大似然方法或贝叶斯方法估计模型参数；

c.寿命预测：结合设备的故障阈值以及当前退化信息估计该设备的剩余寿命分布。



图2-17 随机过程构建过程

（4）卷积神经网络

①模型简介：详见2.2.1。

②模型构建：如图2-18，首先进行数据集划分与参数初始化，然后开展模型训练与测试。



图2-18 卷积神经网络构建过程

a.数据集划分：先对数据集划分为训练集，验证集和测试集；

b.参数初始化：初始化模型普通参数和超参数；

c.模型训练：以交叉熵损失函数作为模型的损失函数，训练模型，直至模型收敛；

d.模型测试：通过验证集选出最优模型，最后使用测试集评估选出来的模型。

由于人工智能模型种类繁多，但本质上均按图2-4进行模型的构建，因此此处不再逐一列出。

2.2.3健康状态评估模型

数据驱动的健康状态评估模型按照不同的原理可以继续划分为统计分析模型、人工智能模型，另外，有一类以层次分析法为代表的模型是综合底层设备的性能指标来评价顶层系统的健康状态。这三类模型的适用范围、需求数据特点、原理、模型特点、功能边界以及输入输出存在一定的差异性，为方便后续模型的存储运维，本部分对其进行总结，如表2-7所示。

具体地，针对几种典型的健康状态评估模型，详细的模型介绍和构建流程规范如下：

（1）隐马尔可夫模型

①模型简介：隐马尔可夫模型（HMM）由一个隐藏的不能观测的随机过程（马尔可夫链）和一个实际可观测的随机序列过程构成，记为，为初始状态概率分布，为状态转移概率矩阵，为观测概率矩阵。

②模型构建：基于多组设备状态观测序列，使用参数估计算法（如Baum-Welch算法）估计出一个。

（2）高斯混合模型

①模型简介：高斯混合模型采用多个高斯概率密度来描述参数空间中数据的空间分布特性。它将每一个状态的数据用一个高斯分布来近似，通过增加高斯模型成员的数目可以以任意精度逼近任何的连续分布。

②模型构建：如图2-19，首先基于工况划分结果构建健康评估基准模型，然后建立工况识别模型以开展在线健康评估。



图2-19 高斯混合模型构建过程

a.工况划分：选择工况参数组成工况特征向量，状态特征向量数据归一化，利用K-means聚类算法进行聚类，得到工况空间划分结果；

b.健康评估基准模型构建：选择无故障时设备中反映健康状态的参数，并根据工况聚类情况将其划分到相应工况下，组成该工况下的训练样本集，EM算法估计每个工况下的基准模型GMM参数；

c.工况识别模型构建：利用步骤a工况聚类结果给工况特征向量打标签，即一个工况向量对应一个聚类标签，利用带有标签的工况数据训练Softmax分类模型，得到工况识别模型；

d.在线健康评估：将当前时刻工况特征向量输入训练好的Softmax模型中，得到当前时刻工况属于所划分每一种工况的概率，计算当前时刻状态特征向量与每种工况基准模型GMM的马氏距离，利用GMM公式计算当前时刻健康指标。

表2-7 数据驱动的健康状态评估模型

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 模型类型 | 代表模型 | 适用范围 | 需求数据特点 | 原理 | 模型特点 | 功能边界 | 输入输出 |
| 统计分析模型 | 隐马尔可夫模型 | ①恒定工况  ②设备运行状态的驻留时间近似服从指数分布  ③状态间的转移概率矩阵不随当前状态大幅变化 | 低噪声 | 对设备健康状态参数的监测数据进行统计分析，通过对比与健康状态数据的偏移距离、分布差异等，评估各类状态指标的健康指数 | ①一个双重随机过程  ②需要大量数据建模 | 无法处理噪声数据 | 输入：状态观测序列（等间距时刻对应的设备状态观测值）  输出：设备的状态概率分布，即处于各个健康状态的概率 |
| 高斯混合模型 | ①多工况  ②大部分情况下应该使用的分布数未知，一般选取3～5个 | 无标签、非小样本数据 | ①一种无监督的聚类算法  ②抗干扰和去噪能力较突出  ③适用于低维度情况 | 无法处理小样本数据 | 输入：当前时刻设备状态参数组成的特征向量  输出：当前时刻设备健康指标 |
| 人工智能模型 | 支持向量机 | ①小样本 | 时序、非线性、大样本数据，需平滑降噪处理 | 通过对己知故障数据的训练，建立输入输出之间的映射关系对设备状态进行评估 | ①避免复杂的数学推导和专业知识  ②适合非线性复杂系统  ③依赖于数据的完整性和真实性  ④通常需要大量数据，网络训练时间较长 | 不宜处理大样本数据 | 输入：带标签的训练信号数据集或设备运行特征数据  输出：设备健康状态评估结果 |
| BP神经网络  循环神经网络  深度置信网络  卷积神经网络 | ①恒定工况或复杂工况  ②大样本 | 无法处理小样本数据 |
| 综合评价模型 | 层次分析法 | ①多层次系统  ②存在多个影响健康状态的因素  ③最底层的指标可以被量化，并能够通过一定的方法测量 | 底层指标量化数据 | 按照问题组成因素间的相互关联影响以及隶属关系将因素按不同的层次聚集组合，从而最终使问题归结为最低层相对于最高层(目标)的相对重要权值的确定 | ①具有高度的逻辑性、系统性、简洁性和实用性  ②可以分析一些难以量化的问题 | 无法用于结构不确定系统 | 输入：设备底层性能指标的监测数据或健康状态  输出：设备健康状态等级 |

（3）支持向量机

①模型简介：支持向量机是一类按监督学习方式对数据进行二元分类的广义线性分类器，基本想法是求解能够正确划分训练数据集并且几何间隔最大的最优分类超平面，该样本学习求解的最大边距超平面是其决策边界。需要注意的是，数据样本集合，包含训练样本与测试样本，二者的格式相同，其具体样本格式如下。

针对某一设备，设具有个状态评估指标，种故障模式，种正常模式，则对状态评估指标的某一样本为对应一种故障模式或者正常模式。训练样本格式为：



整个样本集合为：



②模型构建：如图2-20，首先收集样本集，然后选取核函数与分类算法进行模型训练，以开展设备健康状态评估。



图2-20 支持向量机构建过程

a.收集样本集：按要求收集样本集用于支持向量机模型训练与测试；

b.选取核函数与分类算法：核函数的选取对测试样本评估的准确度有很大影响，选取核函数要视样本而定。一般情况下，高斯径向基核函数能达到要求，是首选的核函数。分类算法一般选取一对多支持向量机分类算法或M-ary分类算法；

c.模型训练：将收集到的样本集，按比例分为训练样本和测试样本，用训练样本训练出设备健康状态评估的支持向量机模型，并用测试机检测模型训练精度；

d.设备健康状态评估：将评估数据代入训练好的支持向量机模型中，得到设备健康状态的评估结果。

（4）自组织映射神经网络

①模型简介：自组织映射神经网络是一种人工竞争型神经网络，它具有模拟生物神经系统特征映射的功能。该网络的基本结构由输入层和竞争层两层网络构成，主要是采用无监督方式进行自组织学习。与其它神经网络相比较，它可以提取输入数据的模式，经过训练对新的输入数据进行自动地模式识别。自组织映射神经网络训练完成之后，竞争层的神经元会形成一个二维网络，这样的二维网络代表了相邻神经元之间的连接关系，依靠神经元之间互相竞争逐步优化网络。

②模型构建：如图2-21，首先初始化网络权重以及确定优胜节点，然后迭代更新权重以开展设备健康状态评估。

a.权重初始化：与其他神经网络相同，需要给自组织映射神经网络结构中的权重初始化为很小的随机数；

b.确定优胜节点：取一个输入样本，遍历竞争层中每一个节点：计算与节点之间的相似度(通常使用欧式距离)选取距离最小的节点作为优胜节点(winner node)，有时也叫BMU(best matching unit)；



图2-21 自组织映射神经网络构建过程

c.迭代更新权重：根据邻域半径确定优胜邻域将包含的节点，并通过neighborhood function计算它们各自更新的幅度（基本思想是：越靠近优胜节点，更新幅度越大；越远离优胜节点，更新幅度越小），更新优胜邻域内节点的权重；完成一轮迭代（迭代次数+1），返回第二步，直到满足设定的迭代次数；

d.设备健康状态评估：将评估数据代入训练好的模型中，得到设备健康状态评估结果。

（5）层次分析法

①模型简介：层次分析法是一种解决多目标的复杂问题的定性与定量相结合的决策分析方法。该方法将定量分析与定性分析结合起来，用决策者的经验判断各衡量目标之间能否实现的标准之间的相对重要程度，并合理地给出每个决策方案的每个标准的权数，利用权数求出各方案的优劣次序，比较有效地应用于那些难以用定量方法解决的课题。

②模型构建：如图2-22，首先建立设备的层次模型，由专家构造判断矩阵，并进行一致性检验，然后依次进行层次单排序和层次总排序，确定底层性能指标隶属于各个状态的隶属度，进一步计算出设备的健康状态。

a.建立层次模型：首先，研究设备的基本结构和运行原理，建立设备结构指标体系。其次，筛选出能准确反应设备当前运行状态的各项性能指标，提取出关键的可监测的性能指标作为层次结构的底层指标，摒除无用低效的性能指标，建立设备故障指标体系。综合考虑设备结构指标体系和故障指标体系，建立设备层次结构；

b.建立判断矩阵：对每一层中的各个指标进行两两比较，通过给定标度值来确定它们的相对重要性，例如对于上一层的指标，它对下一层指标有支配关系，针对，两个指标和哪个更重要，需要对“重要性”进行量化，建立两两比较判断矩阵，其中，表示对于上一层指标，指标相对于的重要程度。为了量化决策判断，形成数值判断矩阵，常用1~9标度方法对同一层次中各指标重要程度进行打分，即标度范围为1~9，标度越大说明前者指标与后者指标相对越重要，若要表示前者指标比后者指标相对次要，则用1~9的倒数表示，1表示两指标相同重要。再通过对同层次中各指标重要性标度进行两两比较，构造权重判断矩阵。判断矩阵1~9标度及其含义如表2-8所示；



图2-22 层次分析法构建过程

表2-8 “1~9”标度及含义

|  |  |  |
| --- | --- | --- |
| 标度 | 定义 | 说明（对于而言） |
| 1 | 同样重要 | 和同样重要 |
| 3 | 稍微重要 | 比稍微重要 |
| 5 | 明显重要 | 比明显重要 |
| 7 | 强烈重要 | 比..强烈重要 |
| 9 | 极端重要 | 比极端重要 |
| 2，4，6，8 | 相邻判断的中值 | 相邻判断的中间值 |
| 倒数 | 满足正互反性 | 和比较得，则 |

c.矩阵的一致性检验：用来度量判断矩阵是否满足一致性，即：



式中，为判断矩阵阶数，为判断矩阵最大特征值。越大，表明偏离一致性程度越大，越小，表明矩阵更满足对于一致性的要求。由于判断矩阵阶数不同对的要求也不同，引入一致性指标，的值如表2-9所示：

表2-9 一致性指标*RI*表

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 矩阵阶数 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|  | 0.58 | 0.90 | 1.12 | 1.24 | 1.32 | 1.41 | 1.45 |

当判断矩阵的阶数小于等于2时，不需进行一致性检验，定义一致性指标和平均随机一致性指标之比为随机一致性比率：



值越小，表明矩阵一致性越好，通常认为当时，一致性基本满足要求。否则，进行优化判断矩阵，再次进行一致性检验判断直到判断矩阵满足一致性要求；

d.层次单排序及总排序：层次单排序问题可归结为求判断矩阵最大特征根及其特征向量的问题，即求得判断矩阵的最大特征根对应的特征向量，对进行归一化，即得该层因素对上层某指标的权重向量，这就是层次单排序，权重越大，对应指标对上层指标健康状态等级影响程度越大；层次总排序是进行最终排序，计算底层性能指标关于总目标的权重；

e.确定隶属度矩阵：利用模糊综合评价法，确定底层性能指标的隶属度矩阵。为各层级指标设置健康等级，如，以描述指标的健康程度；

f.综合评估：计算组合权重矩阵**W**与隶属度矩阵的乘积，得到设备综合评价向量，将其与设备健康状态评估的隶属度相比较，便可得到对应的健康状态等级。

2.3模型推荐

风洞设备种类繁多，各设备运行时，其一些部件甚至其本身可进行不同形式的运转，致使这些设备在性能降低、故障信号等方面表现出一定的差异性，导致对设备开展PHM任务时需要选择性的使用PHM模型，例如当设备的历史监测数据缺乏时，不宜选择需要大量数据建模的神经网络等模型。因此，本部分基于不同类型的风洞设备，包括机械设备、机电设备和电子设备，针对典型的PHM任务，依据设备故障机理、监测数据的特点或工况的变化与否等因素，推荐合适的PHM模型。

机械设备是指部件或本身进行机械运动的设备，经过大量的试验可以证明，机械设备的故障是有规律的，即故障率曲线，大部分机械设备的故障率是时间的函数。

机电设备最初是指既有机械系统又有电子系统的设备的统称，例如设备运动功能的实现主要是由机械部分完成，而自动化控制等则是由电路板、继电器和传感器等电子元器件组成的电子系统来实现，后来发展为集机械、微电子、计算机技术、信息控制技术等为一体的全新技术混合系统，故其系统组成与故障机理一般非常复杂。

电子设备是指由集成电路、晶体管、电子管等电子元器件组成，应用电子技术（包括）软件发挥作用的设备，包括电子计算机以及由电子计算机控制的机器人、数控或程控系统等。通常，电子设备退化状态无明显的外在表现。

基于上述各类型设备特点，为了方便风洞设备PHM任务的开展，我们首先给出如表2-10所示的模型推荐建议，然后以几种风洞设备为代表，总结它们各自的信息，并提供模型推荐示例，如表2-11所示。

表2-10 模型推荐

|  |  |  |  |
| --- | --- | --- | --- |
| 设备类型 | 任务 | 适用情况 | 推荐模型 |
| 机械设备 | 故障  诊断 | ①含未知干扰  ②非线性系统 | 未知输入观测器 |
| ①大样本  ②非线性系统 | 粒子滤波 |
| ①设备性能逐渐降低，且有指标能够反映  ②性能指标的监测数据需是恒定工况下的 | 高斯过程回归 |
| ①设备各运行状态区分明显  ②设备在各个运行状态停留时间服从指数分布  ③某一时刻设备真实状态不可观测，而由观测状态依概率确定 | 隐马尔可夫模型 |
| ①设备组成及故障机理复杂  ②监测数据多  ③恒定工况或复杂工况 | 神经网络 |
| ①监测数据较少 | 支持向量机 |
| 故障  预测 | ①线性系统  ②噪声服从高斯分布 | 卡尔曼滤波 |
| ①大样本  ②非线性系统  ③中长期预测 | 粒子滤波 |
| ①设备性能逐渐降低，且有指标能够反映  ②性能指标的监测数据是恒定工况下的  ③预测未来较短时期内的故障 | 自回归模型 |
| ①设备性能逐渐降低，且有指标能够反映  ②性能指标的监测数据是恒定工况下的 | 随机过程 |
| ①设备性能逐渐降低，且有指标能够反映  ②性能指标的监测数据是多工况下的 | 带协变量的  随机过程 |
| ①设备组成及故障机理复杂  ②监测数据多  ③恒定工况或复杂工况 | 神经网络 |
| ①监测数据较少 | 支持向量机 |
| 健康状态评估 | ①线性系统  ②噪声服从高斯分布 | 卡尔曼滤波 |
| ①非线性系统 | 扩展卡尔曼滤波 |
| ①设备各运行状态区分明显  ②设备在各个运行状态停留时间服从指数分布  ③某一时刻设备真实状态不可观测，而由观测状态依概率确定 | 隐马尔可夫模型 |
| ①监测数据是多工况下的 | 高斯混合模型 |
| ①设备组成及故障机理复杂  ②监测数据多  ③恒定工况或复杂工况 | 神经网络 |
| ①监测数据较少 | 支持向量机 |
| ①多层次系统  ②存在多个影响健康状态的因素  ③最底层的指标可以被量化，并能测量 | 层次分析法 |
| 机电设备 | 故障  诊断 | ①含未知干扰  ②非线性系统 | 未知输入观测器 |
| ①大样本  ②非线性系统 | 粒子滤波 |
| ①设备组成及故障机理复杂  ②监测数据多  ③恒定工况或复杂工况 | 神经网络 |
| ①监测数据较少 | 支持向量机 |
| 故障  预测 | ①线性系统  ②噪声服从高斯分布 | 卡尔曼滤波 |
| ①大样本  ②非线性系统  ③中长期预测 | 粒子滤波 |
| ①设备组成及故障机理复杂  ②监测数据多  ③恒定工况或复杂工况 | 神经网络 |
| ①监测数据较少 | 支持向量机 |
| 健康状态评估 | ①线性系统  ②噪声服从高斯分布 | 卡尔曼滤波 |
| ①非线性系统 | 扩展卡尔曼滤波 |
| ①监测数据是多工况下的 | 高斯混合模型 |
| ①设备组成及故障机理复杂  ②监测数据多  ③恒定工况或复杂工况 | 神经网络 |
| ①监测数据较少 | 支持向量机 |
| ①多层次系统  ②存在多个影响健康状态的因素  ③最底层的指标可以被量化，并能测量 | 层次分析法 |
| 电子设备 | 故障  诊断 | ①含未知干扰  ②非线性系统 | 未知输入观测器 |
| ①大样本  ②非线性系统 | 粒子滤波 |
| ①设备性能逐渐降低，且有指标能够反映  ②性能指标的监测数据需是恒定工况下的 | 高斯过程回归 |
| ①设备各运行状态区分明显  ②设备在各个运行状态停留时间服从指数分布  ③某一时刻设备真实状态不可观测，而由观测状态依概率确定 | 隐马尔可夫模型 |
| ①设备组成及故障机理复杂  ②监测数据多  ③恒定工况或复杂工况 | 神经网络 |
| ①监测数据较少 | 支持向量机 |
| 故障  预测 | ①线性系统  ②噪声服从高斯分布 | 卡尔曼滤波 |
| ①大样本  ②非线性系统  ③中长期预测 | 粒子滤波 |
| ①设备性能逐渐降低，且有指标能够反映  ②性能指标的监测数据是恒定工况下的  ③预测未来较短时期内的故障 | 自回归模型 |
| ①设备是硬故障，即在某一瞬间功能突然完全丧失 | 比例风险模型 |
| ①设备组成及故障机理复杂  ②监测数据多  ③恒定工况或复杂工况 | 神经网络 |
| ①监测数据较少 | 支持向量机 |
| 健康状态评估 | ①线性系统  ②噪声服从高斯分布 | 卡尔曼滤波 |
| ①非线性系统 | 扩展卡尔曼滤波 |
| ①监测数据是多工况下的 | 高斯混合模型 |
| ①设备组成及故障机理复杂  ②监测数据多  ③恒定工况或复杂工况 | 神经网络 |
| ①监测数据较少 | 支持向量机 |
| ①多层次系统  ②存在多个影响健康状态的因素  ③最底层的指标可以被量化，并能测量 | 层次分析法 |

表2-11 模型推荐示例

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 设备 | 所属类型 | 设备信息 | | | | 模型推荐 | |
| 主要故障模式 | 故障演变特性 | 主要监测参数 | 特点 | 任务 | 模型 |
| 挤压供气装置 | 机械设备 | 磨损  活塞阻滞  拉缸 | 慢速演变 | 振动 | ①退化型  ②无层级结构  ③恒定工况 | 故障诊断 | 高斯过程回归 |
| 故障预测 | 随机过程 |
| 健康状态评估 | 隐马尔可夫模型 |
| 调节阀 | 机械设备 | 磨损  裂纹  泄露 | 慢速演变 | 振动 | ①退化型  ②无层级结构  ③复杂工况  ④大样本 | 故障诊断 | 神经网络 |
| 故障预测 | 带协变量的随机过程 |
| 健康状态评估 | 神经网络 |
| 冷却水系统 | 机电设备 | 泄露  磨损 | 慢速演变 | 振动、噪声 | ①退化型  ②有层级结构  ③含未知干扰 | 故障诊断 | 未知输入观测器 |
| 故障预测 | 卡尔曼滤波 |
| 健康状态评估 | 层次分析法 |
| 快速阀 | 机电设备 | 泄露  结构损伤 | 慢速演变 | 振动 | ①退化型  ②无层级结构  ③复杂工况  ④大样本 | 故障诊断 | 神经网络 |
| 故障预测 | 神经网络 |
| 健康状态评估 | 高斯混合模型 |
| 点火系统 | 电子设备 | 功率不足  点火频率低  点火头烧蚀 | 慢速演变 | 电压、电流 | ①退化型  ②有层级结构  ③恒定工况 | 故障诊断 | 高斯过程回归 |
| 故障预测 | 自回归模型 |
| 健康状态评估 | 层次分析法 |
| 加热器 | 电子设备 | 爆闪  爆燃  伸缩变形 | 突变 | 振动、位移 | ①硬故障，即在某一瞬间功能突然完全丧失  ②无层级结构  ③大样本 | 故障诊断 | 神经网络 |
| 故障预测 | 比例风险模型 |
| 健康状态评估 | 神经网络 |

3 模型存储运维规范

随着预测与健康管理技术的高速发展，PHM模型呈现出复杂多样、数量庞大的特点。然而，由于现阶段基地采用“一座风洞、一个平台、一套软件”的分散建设、独立运行模式，导致每个风洞系统拥有一套独立的健康管理系统，系统中的模型复用性差、重复建设问题严重，模型难以在系统间共享而形成孤岛。鉴于上述特点，应该对模型的存储使用运维进行规范，促进各个风洞模型的一体化。如果缺乏相关规范，将导致模型问题凸显，影响后续正常使用。因此，为了保障风洞一体化平台模型的综合质量，需要对模型的接口与参数规范、编码规范、调用运行、性能评估、数据存储、数据更新、模型简介等方面做出相应的规范进行约束，具体流程如图3-1所示。

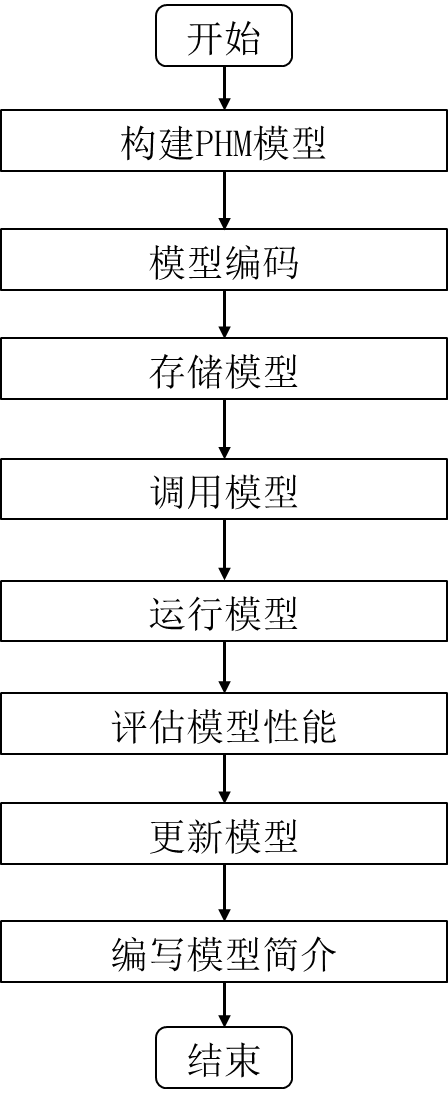


图3-1 模型存储运维规范流程

3.1 模型接口与参数规范

PHM的模型可根据其中应用方法或算法的原理和构建过程，可分为模型驱动和数据驱动两类，由于不同种类的模型原理和任务的不同，模型的输入输出参数和模型接口存在很大差异，导致模型在不同系统之间的共享困难。因此，需要对不同类型的模型的接口、参数进行规范。本报告对不同类型的PHM模型的原理、输入输出、模型接口等进行了梳理，详细信息如表3-1所示。

表3-1 模型的接口参数要求

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 模型类型 | | 原理 | 代表模型 | 输入输出参数 | 模型接口 |
| 模型驱动模型 | 基于物理模型的PHM模型 | 物理失效模型 | ①Paris-Erdogan模型  ②状态空间模型  ③裂纹增长力学定律  ④非线性损伤定律  ⑤随机物理失效模型 | 输入：物理失效模型、模型参数  输出：设备性能退化曲线 | 函数、输入参数 |
| 基于滤波器的模型 | 滤波器 | ①卡尔曼滤波  ②粒子滤波 | 输入：状态和测量模型、状态模型参数初始分布  输出：设备性能退化曲线 | 状态方程、观测方程、噪声等级 |
| 数据驱动模型 | 基于统计分析的模型 | 统计分析模型 | ①自回归移动平均模型  ②比例风险模型  ③随机过程模型  ④高斯混合模型 | 输入：故障样本数据、退化样本数据  输出：可靠性曲线 | 函数、输入数据 |
| 基于人工智能的模型 | 机器学习算法 | ①支持向量机  ②BP神经网络  ③循环神经网络  ④深度置信网络  ⑤卷积神经网络 | 输入：多源状态监测数据  输出：故障诊断、故障预测、健康状态评估结果 | 函数、类、API或控件、输入数据 |
| 综合评价方法 | | 系统级算法 | 层次分析法 | 输入：设备底层性能指标的监测数据或健康状态  输出：设备健康状态等级 | 图形化接口、输入参数 |

3.2 模型编码规范

由于风洞设备类型繁多、数量庞大，针对不同的可维修单元所构建的PHM模型也各式各样。因此，为了更加清晰地对不同的PHM模型进行区分和调用，本部分结合气动中心装备管理编码规范，对通用型和风洞设备相关模型的PHM模型的编码结构提出具体的规范和要求。下面将介绍PHM模型编码的范围、引用文件、术语定义、编码要求和具体的模型码编码规范。

3.2.1 范围

本规范从通用型模型和风洞设备相关模型两个角度出发，一方面针对通用型模型编码，依据模型分类对PHM模型的编码进行规范。另一方面，基于气动中心装备管理编码规范，以装备的最小维修单元或元件为对象，规范了以装备的最小维修单元或元件为对象的PHM模型管理编码要求。

本规范适用于气动中心装备的PHM模型的管理以及相关信息系统的开发。

3.2.2 引用文件

下列文件中的条款通过引用而成为本规范的条款。

《中国人民解放军军语》

《中国人民解放军装备管理条例》

《全军信息服务基础平台白皮书》

GB/T 2260 中华人民共和国行政区划分代码

GB/T 10113 分类编码通用术语

GB/T 24463.3-2009 交互式电子技术手册 第3部分：公共源数据库要求

GJB 4017-2000 风洞试验代码

GJB 7000-2010 军用物资和装备分类

GJB 6600.2-2009 装备交互式电子技术手册 第2部分：数据模块编码和信息控制编码

Q/QD 9-2021中国空气动力研究与发展中心标准 气动中心装备管理编码规范

3.2.3 术语和定义

（1）PHM模型

本规范所谓“PHM模型”是指用于风洞设备预测与健康管理（Prognosis and Health Management，PHM）任务或工作模型的统称。

（2）编码结构

编码结构指编码符号排列的逻辑顺序。

（3）PHM模型原理对照表

模型原理对照表用于描述编码模型的原理（或构建流程）和码段字符串值之间一一对应的关系。

（4）PHM模型任务与功能对照表

模型任务与功能对照表用于描述编码模型的任务（或功能）和码段字符串值之间一一对应的关系。

（5）PHM模型结构码

PHM模型结构码是用于表示一个特定PHM模型的从属关系、相关关系及其对装备的最小维修单元或元件从属关系的编码。

3.2.4 编码要求

（1）通用要求

①具有唯一性；

②采取分段式编码。

（2）字符要求

①数字字符“0”至“9”；

②大写英文字母“A”至“Z”，但不包括“I”和“O”（易与数字“1”和“0”混淆）。

3.2.5 PHM模型码编码规范

（1）编码依据

在整个PHM模型中，一部分模型具有通用性和普适性，而另一部分模型与风洞设备直接相关。因此，本编码规范针对通用型PHM模型和风洞设备相关的PHM模型，分别设计对应的模型编码规范。其中，Q/QD 9-2021对气动中心装备管理编码规则进行了定义，风洞设备相关的PHM模型码编码规范便可基于Q/QD 9-2021中的装备台账码编码规范进行编写。

（2）通用型PHM模型编码设计要求

通用型PHM模型码由“模型原理类别码+模型任务与功能类别码+模型版本号编码”组成。通用型的PHM模型码结构如图3-2所示。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ××××× | ××× | - | ×××××××× |
| 模型原理类别码 |  |  |  |  |
| 模型任务与功能类别码 |  |  |  |  |
| 模型版本号编码 |  |  |  |  |

图4-2 通用型的PHM模型码结构

①“模型原理类别码”由5位大写英文字母组成。具体地，第一位大写字母表示模型按原理划分后为“模型”或“数据”驱动。模型驱动类型为“M”，数据驱动类型为“D”；中间一位大写字母表示模型算法原理的类型，如基于物理模型的类型为“P”，基于滤波器的类型为“F”，基于统计模型的类型为“S”，基于人工智能的类型为“A”。最后三位大写字母表示模型的具体算法，如基于粒子滤波的模型为“PFB”，基于数字孪生的模型为“DTM”，基于卷积神经网络的模型为“CNN”。例如，基于深度卷积神经网络的模型被表示为“DACNN”类。

②“模型任务与功能类别码”由3位大写英文字母组成。具体地，第一位大写字母表示模型按任务与功能划分后为由“分类”或“回归”任务。分类任务类型为“C”，回归任务类型为“R”；后两位大写字母表述具体的任务与功能。如故障诊断类型为“FD”，故障预测类型为“FP”，健康状态评估类型为“HE”，剩余寿命预测类型为“RP”等。例如，故障诊断模型可被表示为“CFD”类。

③“模型版本号编码”由10位数字组成，包括8位数字的模型创建日期和2位数字的模型创建顺序号。具体地，模型创建日期由8位数字组成，如2021年9月16日所创建模型的创建时间表示为“20210916”；模型创建顺序号由2位数字组成，如某天所创建的第二个相同或相似模型的创建顺序号表示为“02”。例如，在2021年9月16日所创建的第二个相同或相似模型的版本号可以表示为“2021091602”。

举例：在2021年9月16日构建的第二个基于深度卷积神经网络的故障诊断模型可被表示为DACNNCFD-2021091602。

（3）风洞设备相关的PHM模型码编码设计要求

风洞设备相关的PHM模型码由“管理机构码+系统类别码+系统序号+装备类别码+设备序号+模型原理类别码+模型任务与功能类别码+模型版本号编码”组成。风洞设备相关的PHM模型码结构如图3-3所示。其中，装备台账码由“管理机构码+系统类别码+系统序号+装备类别码+设备序号”组成。

Q/QD 9-2021已经对装备台账码进行了规范，具体编码规则详见气动中心装备管理编码规范第8章。

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 510×× | - | ××× | ××× | - | ×××× | ××× | - | ××××× | ××× | - | ×××××××× |
| 管理机构码（前5位） |  |  |  |  |  |  |  |  |  |  |  |  |
| 所属系统类别码 |  |  |  |  |  |  |  |  |  |  |  |  |
| 所属系统序号 |  |  |  |  |  |  |  |  |  |  |  |  |
| 装备类别码 |  |  |  |  |  |  |  |  |  |  |  |  |
| 设备序号 |  |  |  |  |  |  |  |  |  |  |  |  |
| 模型原理类别码 |  |  |  |  |  |  |  |  |  |  |  |  |
| 模型任务与功能类别码 |  |  |  |  |  |  |  |  |  |  |  |  |
| 模型版本号编码 |  |  |  |  |  |  |  |  |  |  |  |  |

图3-3 风洞设备相关的PHM模型码结构

①“模型原理类别码”由5位大写英文字母组成。具体地，第一位大写字母表示模型按原理划分后为“模型”或“数据”驱动。模型驱动类型为“M”，数据驱动类型为“D”；中间一位大写字母表示模型算法原理的类型，如基于物理模型的类型为“P”，基于滤波器的类型为“F”，基于统计模型的类型为“S”，基于人工智能的类型为“A”。最后三位大写字母表示模型的具体算法，如基于卷积神经网络的模型为“CNN”，基于粒子滤波的模型为“PFB”，基于数字孪生的模型为“DTM”。例如，基于深度卷积神经网络的模型被表示为“DACNN”类。

②“模型任务与功能类别码”由3位大写英文字母组成。具体地，第一位大写字母表示模型按任务与功能划分后为由“分类”或“回归”任务。分类任务类型为“C”，回归任务类型为“R”；后两位大写字母表述具体的任务与功能。如故障诊断类型为“FD”，故障预测类型为“FP”，健康状态评估类型为“HE”，剩余寿命预测类型为“RP”等。例如，故障诊断模型可被表示为“CFD”类。

③“模型版本号编码”由10位数字组成，包括8位数字的模型创建日期和2位数字的模型创建顺序号。具体地，模型创建日期由8位数字组成，如2021年9月16日所创建模型的创建时间表示为“20210916”；模型创建顺序号由2位数字组成，如某天所创建的第二个相同或相似模型的创建顺序号表示为“02”。例如，在2021年9月16日所创建的第二个相同或相似模型的版本号可以表示为“2021091602”。

举例：低速空气动力研究所4米×3米低速风洞风扇，其装备台账码为51013-AFL12Z-0104001，而用于该设备上的在2021年9月16日构建的第二个基于深度卷积神经网络的故障诊断模型可被表示为DACNNCFD-2021091602。于是该PHM模型的编码为：51013-AFL12Z-0104001-DACNNCFD-2021091602。

3.3 模型存储规范

在构建PHM模型的时候需要存储大量的模型数据，为了保证模型整体的一致性和完整性，在存储模型数据的时候应该遵循统一规范的存储流程和一致的命名方式。在存储模型数据之后还应该注意到会有意外事件的发生，应对数据进行备份处理以便意外事件突发时能迅速恢复数据。因此，本小节将从模型储存流程和模型数据备份与恢复两个方面进行说明。

3.3.1 模型存储流程

对模型的存储进行规范就是要对模型存储过程中的各个步骤进行规范。模型的存储可以分为设计模型存储表、设置表的编码类型、设计表中字段、规范字段命名方式等步骤。

其中，设计模型存储表就是要对模型中应包含有哪些数据表、数据表的名称等进行设计并说明，构建出模型数据的整体框架。设置表的编码类型是指为了让计算机能够处理表中存储的数据，要对数据进行编码并校验，为了能够支持中文，可以使用utf-8编码并使用utf-8-cgi校验。设计表中字段指对数据表中所包含的多个字段及其类型进行设计并说明，如表中应包含有id、time等字段用来作为表中数据的唯一标识和时间标志，id字段应设计为int类型等。规范字段命名指的是应该对表中各个字段的命名方式进行约束，要规定统一风格的命名方式保证全局数据名称的一致性。其整体的流程图如下图3-4所示：

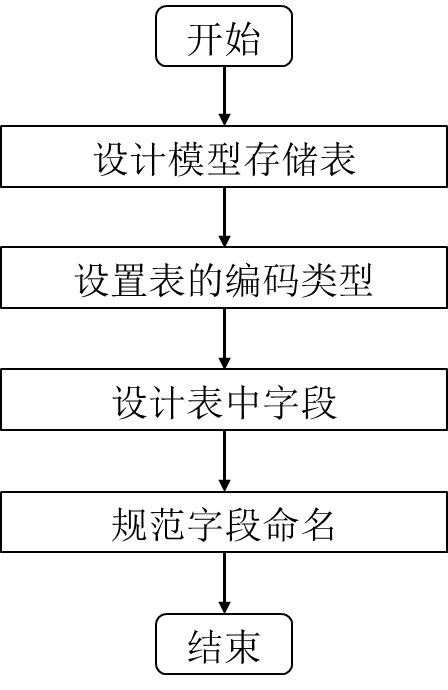


图3-4 模型存储流程

3.3.2 模型数据备份与恢复

在使用模型库存储数据文件时，可能会出现以下场景导致数据丢失或损坏：

①人为操作失误造成某些数据被误操作；

②软件BUG造成部分数据或全部数据丢失；

③硬件故障造成数据库部分数据或全部数据丢失；

④全漏洞被入侵数据恶意破坏。

还可能会由于特殊应用场景下基于时间点的数据恢复、开发测试环境模型库搭建、相同模型库的新环境搭建、模型库或者数据迁移等情况需要对数据进行迁移处理，因此应该对模型库本身及其数据进行备份，以便各种事件发生时能够迅速恢复数据。模型库备份的内容应该包括模型库中的数据、模型库结构和模型库定义文件。应按照模型库结构（字段）备份模型库数据。数据的备份方式可以有整体备份、拆分备份、增量与差异备份、在线备份和离线备份以及自动备份几种方式。

整体备份是指当模型库的容量小于备份介质的容量时，无需对模型库做任何处理，直接对模型库整体进行复制备份。拆分备份是指当模型库的容量大于备份介质的容量时，应将模型库拆分成容量小于备份介质容量的若干个子模型库，然后将各子模型库分别备份到备份介质上，保留原有的访问控制策略，并保证原模型库的完整性。增量与差异备份是指对模型库中新增的数据进行增量备份，对被修改的数据进行差异备份。在线备份指将模型库中的数据、模型库结构和模型库定义文件备份到在线存储介质上。离线备份指是将模型库中的数据、模型库结构和模型库定义文件备份到离线存储介质上。自动备份通过软件的控制方式将模型库中的数据、模型库结构和模型库定义文件进行备份。

在风洞一体化平台数据的备份时可以采取增量与差异备份，这种方法不会产生重复的备份数据，备份数据需要时间较短且相较于其它方法更能节省磁盘空间。

在对模型中存储的数据表进行单独备份时，可以对该表进行锁定，允许其他线程继续查询该表但是不能对表中数据进行修改。锁定表之后可以使用SQL语句对数据进行备份和恢复，备份完成之后可以解除对该表的锁定。可以使用SELECT INTO OUTFILE语句备份数据，并用LOAD DATA INFILE语句恢复数据。但是这种方法只能导出数据的内容，而不包括表的结构，如果表的结构文件损坏，必须要先恢复原来的表的结构。对表中数据进行备份的步骤及SQL语句大概如下：

①首先锁定数据表，避免在备份过程中表被更新，使用SQL语句“LOCK TABLES READ tbl\_name”；

②然后导出数据，使用SQL语句“SELECT \* INTO OUTFILE ‘tbl\_name.bak’ FROM tbl\_name”；

③备份完成后解锁表，使用SQL语句“UNLOCK TABLES”。

使用备份数据还原数据表时的步骤及SQL语句如下：

①首先为表增加锁定，使用SQL语句“LOCK TABLES tbl\_name WRITE”；

②然后恢复数据，使用SQL语句“LOAD DATA INFILE ‘tbl\_name.bak’ ->REPLACE INTO TABLE tbl\_name”；

③最后对表进行解锁。

数据表备份及恢复的完成流程如下图3-5所示：

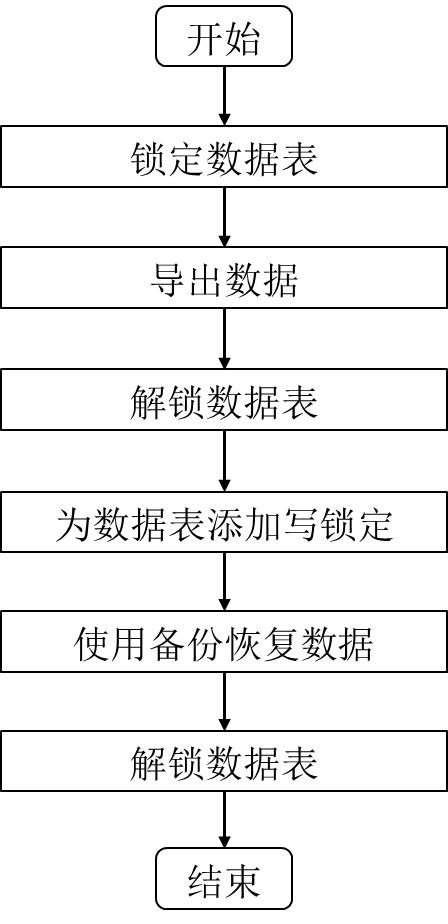


图3-5 数据表备份及恢复流程

3.4 模型调用规范

模型的源代码是模型调用、运行、训练、测试等一系列工作的准则和源泉，为了在后期的运行、训练、测试过程中顺利地实现模型的功能，需要对PHM模型的调用进行规范。而在PHM模型调用过程中，分为应用程序接口（Application Programming Interface，API）和可视化控件。根据业务场景和使用复杂度的不同，可使用API解决较为简单和已经集成完备的模型问题，使用可视化控件解决复杂调用业务场景。下面将分别对API和可视化的调用进行规范。

3.4.1 API规范

API是一些预先定义的接口（如函数、HTTP接口）或指软件系统不同组成部分衔接的约定。可用来提供应用程序与开发人员基于某软件或硬件得以访问的一组例程，而又无需访问源码，或理解内部工作机制的细节。API的一个主要功能是提供通用功能集。API同时也是一种中间件，为不同模块提供数据共享。在模型调用设计的实践中，编程接口的设计首先要使系统模型的职责得到合理划分。良好的接口设计可以降低系统各部分的相互依赖，提高组成单元的内聚性，降低组成单元间的耦合程度，从而提高系统的可维护性和可扩展性，常见API执行流程如图3-6所示。



图3-6 API执行流程图

绝大部分的服务需要编写API接口，对外提供服务。在开发之前需要确定一种API风格。API风格也可以理解为API类型，目前业界常用的API风格有三种：REST、RPC和GraphQL。

表3-2 API风格特点

|  |  |  |  |
| --- | --- | --- | --- |
| 名称 | 优点 | 缺点 | 适用范围 |
| REST | 系统耦合性低、调用者无需了解内部实现、增强了语义表现力、API可随时间改进 | 缺少约束、缺少简单统一的规范、可能调用繁琐 | 关注对象与资源、有多种不同的客户端或需要良好的可发现性和文档 |
| RPC | 面向开发者简单、且易于理解、轻量级、高性能 | 耦合性高、调用者调用者必须足够了解系统 | 面向动作或者指令、仅需要简单的交互、消息密集型或对系统性能有较高 |
| GraphQL | 网络开销低、适合图状数据结构的业务场景 | 语法较为复杂、技术细节待验证、版本控制待优化 | 具有图状的特点并且希望对于高延迟场景可以有更好的优化 |

REST是一种设计网络应用时的构架风格，这种风格约束了网络应用中各个模块之间的交互。REST描述了以抽象的“资源”为中心，各个模块之间通过转移沟通资源的“表征状态”来实现应用的功能。在REST里，“资源”是一切可命名事物的抽象。REST的优点是对于系统本身耦合性低，调用者不再需要了解接口内部处理和实现细节，重复使用了一些http协议中的已定义好的部分状态动词来增强语义表现力，API可以随着时间而不断演进。同时REST也缺少约束，缺少简单、统一的规范；有时调用API会比较繁琐。

RPC就是从一台机器（客户端）上通过参数传递的方式调用另一台机器（服务器）上的一个函数或方法（可以统称为服务）并得到返回的结果。RPC有着他的优点与缺点。优点是面向开发者简单并且易于理解，并且是一种轻量级的数据载体同时具备高性能。缺点是对于系统本身耦合性高，调用者需要知道所调用接口的函数名、参数格式、参数顺序、参数名称等等。因此，调用者必须足够了解系统，从能够知道如何正确的调用这些接口，但是对于接口调用者往往不需要了解过多系统内部实现细节。

GraphQL是API的查询语言，它显示了服务器提供的不同类型的数据，然后客户端可以准确地选择它想要的内容。GraphQL的优点是网络开销低，可以在单一请求中获取REST中使用多条请求获取的资源，强类型Schema可以根据规范形成文档、IDE、错误提示等生态工具，特别适合图状数据结构的业务场景（比如好友、流程、组织架构等系统）。但同时GraphQL也具有着一些缺点。它本身的语法相比较REST和RPC均复杂一些，对于API的版本控制当前没有完善解决方案，仍然是新鲜事物，很多技术细节仍然处于待验证状态。

当API关注于对象与资源，有多种不同的客户端或需要良好的可发现性和文档这种情景下使用REST风格可能会更好。如果API具有面向动作或者指令、仅需要简单的交互、消息密集型或对系统性能有较高要求的特点时建议使用RPC。当API的数据类型是具有图状的特点并且希望对于高延迟场景可以有更好的优化时，这种场景无疑GraphQL是最好的选择。

针对项目特点可采用REST或GraphQL的API风格.

而一个优秀的API应该遵守以下规范：

（1）简单：即使它可能会有所帮助，也要努力去避免不必要的复杂性。好的API应当表现出简单性；

（2）提供有用的抽象：API应该隐藏它对用户的详细信息，同时使其对用户有用。如果用户需要深入代码或执行来理解，则表明该API抽象程度不够；

（3）可发现的：通过相应的计划和设计，包括文档和示例，以及自我描述访问点让API具备可发现性；

（4）一致性：用相同的方式命名相同的东西，保持一个共同的风格。从而让API可预测；

（5）遵循最不惊奇的原则：一个系统应该以与该组件的用户可能期望的行为一致的方式运行。用户不应该对它的行为感到惊讶。

针对系统中可能出现的模型接口进行声明

表3-3 故障诊断模型

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 接口类型 | 数据源 | 名称 | 数据格式 | API风格 |
| 输入 | 数据库 | 状态向量 | 数值 | REST |
| 数据库 | 特征参数时间序列 | 数值 | REST |
| 传感器 | 时域信号 | 处理为图片 | GraphQL |
| 传感器 | 时频域信号 | 处理为图片 | GraphQL |
| 数据库 | 标签数据集 | 数值 | REST |
| 输出 | 数据库 | 故障状态评估结果 | 数值/字符 | REST |

表3-4 故障预测模型

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 接口类型 | 数据源 | 名称 | 数据格式 | API风格 |
| 输入 | 数据库 | 状态向量 | 数值 | REST |
| 数据库 | 特征参数时间序列 | 数值 | REST |
| 传感器 | 故障数据 | 字符 | REST |
| 传感器 | 时频域信号 | 处理为图片 | GraphQL |
| 数据库 | 标签数据集 | 数值 | REST |
| 输出 | 数据库 | 剩余寿命 | 数值 | REST |
| 数据库 | 故障预测结果 | 字符 | REST |

表3-5 健康状态评估模型

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 接口类型 | 数据源 | 名称 | 数据格式 | API风格 |
| 输入 | 数据库 | 状态观测序列 | 数值 | REST |
| 数据库 | 特征向量 | 数值 | REST |
| 传感器 | 状态先验分布 | 数值 | REST |
| 数据库 | 标签数据集 | 数值 | REST |
| 输出 | 数据库 | 健康指标 | 字符/数值 | REST |
| 数据库 | 健康状态评估结果 | 字符 | REST |
| 数据库 | 健康状态等级 | 数值/字符 | REST |

3.4.2 可视化控件建模规范

可视化控件建模是利用标准的图形语言对模型的结构进行全面的分析与设计，将系统中的信息直观地显示出来的过程。在此过程中我们应用面向对象的方法，其主要思想是把现实世界中需要解决的业务流程、控制逻辑等问题映射到图形语言中去，利用图形语言中的对象代表现实世界中的实体，并利用对象之间的交互描述现实世界中实体之间的动态关系。常见可视化控件执行流程如图3-7所示



图3-7 可视化流程框架

（1）建模的必要性

建模是帮助用户更好地理解算法、更容易地构建模型的最佳途径。

①使用户更好地理解模型：人们对复杂问题的理解能力是有限的，随着业务环境的不断改变，模型复杂性不断增加，这为用户的使用带来了独特的挑战。通过建模，将整个大的模型按照功能的不同划分为一些细小的模块等等，简化了模型的复杂性。

②使用户着眼于模型的组件如何相互作用：模型的最终表现形式是可运行的代码，而最终的软件代码是非常复杂的，包含了太多的细节信息，建模使用户只关注于算法逻辑，不需要考虑代码的具体实现。

③通过使用一个共同的图形语言，改进跨团队的沟通：使用容易理解的共同的图形语言会改进团队内和团队间的沟通，提高效率和质量。

（2）建模技术

可视化控件建模技术是指在部件视图中，通过拖放模型库中的组件并设置组件参数，然后建立组件之间的连接线来创建模型的技术。可视化建模操作简单，易于理解，可以快速高效地创建模型。用户不需要熟悉具体的语法规范，只需专注于选择合适的组件来实现模型的功能，降低了使用要求。

常用的可视化控件包含jsPlumb和D3.js等。

jsPlumb是一个强大的JavaScript连线库，提供HTML元素的拖放、连线等功能，可绘制不同类型、样式的连线，适用于开发Web页面的图表、建模工具等。

D3js是一个可基于数据来操作文档的JavaScript库。可以帮助使用HTML，CSS，SVG以及Canvas来展示数据。D3遵循现有Web标准，可以不需要其他任何框架独立运行在现代浏览器中，它结合强大的可视化控件来驱动DOM操作。

可视化控件建模的主要功能应当包括但不限于图标的创建、模型库的创建、拖放式建模、组件不重复命名、组件属性的设置、组件的几何变换、连接线的正交化和文本关联修改技术等。

针对业务场景简单，模型复杂度不高时，可采用jsPlumb进行绘制，针对模型较为复杂时可采用D3js进行模型搭建。

（3）建模规范

①简易性：要方便用户使用，便于用户更快构建模型，减少用户选择发生错误的可能性；

②一致性：模型的构建结构必须清晰且一致，设计风格必须与内容相一致；

③易用性：有完善的使用文档或提示，方便用户快速上手，减少用户学习成本；

④复用性：模型元素应当利用共性构建，避免模型元素的冗余。

3.5 模型运行规范

通常情况下，随着模型复杂度的提升，训练误差会不断减小，而测试误差会先减小后增大。控制模型的复杂度可以帮助找到合适的模型来拟合数据。复杂度高的模型通常偏差低而方差高；复杂度小的模型通常偏差高而方差低。控制复杂度本质上是帮助实现偏差-方差的权衡。其主要思想是通过训练数据来估计期望的测试误差，进而得到最适合该模型的训练测试方法来实现偏差-方差的权衡。表3-6展示了常见的模型训练测试方法。

（1）试错法

通过不断试验和消除误差，是一种探索具有黑箱性质的系统的方法。模型的输入经过调整参数后的系统会得到不同的输出，通过对输出结果与历史经验数据的对比，可以得到效果最好的参数，进一步可以得到最贴切的模型结构。

（2）拟合优度检验

拟合优度检验是用卡方统计量进行统计显著性检验的重要内容之一。它是依据总体分布状况，计算分类变量中各类别的期望频数，与分布的观察频数进行对比，判断期望频数与观察频数是否有显著差异，从而达到从分类变量进行分析的目的。

表3-6 常见的模型训练测试方法

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 模型类型 | | 原理 | 训练测试方法 | 特点 | 适用范围 |
| 模型驱动模型 | 基于物理模型的PHM模型 | 物理失效模型 | 试错法 | 暴力尝试 | 物理模型完善且参数丰富 |
| 基于滤波器的模型 | 滤波器 | 暂无 | 暂无 | 暂无 |
| 数据驱动模型 | 基于统计分析的模型 | 统计分析模型 | 拟合优度检验 | 需知总体分布 | 已知总体分布的任何情况 |
| AIC信息准则 | 引入惩罚项  降低过拟合 | 普遍适用 |
| BIC信息准则 | 防止模型精度过高造成的模型复杂度过高。 | 倾向于选择参数少的简单模型 |
| 基于人工智能的模型 | 机器学习算法 | 留出法 | 划分互斥集合 | 任何情况 |
| 留一法 | 计算量大 | 数据缺乏 |
| N折交叉验证法 | 避免固定划分数据 | 数据丰富 |

（3）AIC信息准则

衡量统计模型拟合优良性的一种标准。它建立在熵的概念基础上，可以权衡所估计模型的复杂度和此模型拟合数据的优良性。在一般的情况下，AIC可以表示为：。其中是参数的数量，是似然函数。假设条件是模型的误差服从独立正态分布。

（4）BIC信息准则

又称贝叶斯信息准则，与AIC准则类似，用于模型的选择。，其中为模型参数个数，为样本数量，为似然函数。训练模型时，增加参数数量，也就是增加模型复杂度，会增大似然函数，但是也会导致过拟合现象。惩罚项在样本数量较多的情况下可有效防止模型精度过高造成模型复杂度过高的问题，避免维度灾难现象。

（5）留出法

留出法直接将数据集划分为两个互斥的集合，其中一个集合作为训练集，另一个作为测试集，即。在上训练出模型后，用来评估其测试误差，作为对泛化误差的估计。

（6）留一法

每次从个数为的样本集中取出1个样本作为验证集，剩下的个样本作为训练集，重复进行次，依次取遍所有个数据作为验证集，最后将平均的个数据的结果作为泛化误差的估计。一般在数据缺乏时使用。

（7）折交叉验证

*N*折交叉验证也称循环估计，是一种统计学上将数据样本切割成较小子集的实用方法。将训练集分份，轮流将其中份作为训练数据，1份作为测试数据，进行试验。每次试验都会得出相应的正确率。次结果的正确率的平均值作为对算法精度的估计。具体步骤如下：

①将数据集分为训练集和测试集，将测试集放在一边。

②将训练集分为份。

③每次使用份中的1份作为验证集，其他全部作为训练集。

④通过次训练后，得到个不同的模型。

⑤评估个模型的效果，从中挑选效果最好的超参数

⑥使用最优的超参数，然后将份数据全部作为训练集重新训练模型，得到最终模型。

3.6 模型性能评价规范

预测与健康管理模型本质是分类或回归问题，为了使模型能够发挥其实际作用，必须在模型的调用和运行后选择合适的性能指标来评估模型输出结果的误差。目前广泛使用的多种PHM模型评价方法分别基于不同的评价指标，对于不同种类的模型其评价效果也大不相同，因此这里需要对PHM模型的评价方法进行规范。本节内容包含了PHM模型性能评价指标在不同场景中使用特定指标的适用性、每个指标的优缺点，以及在选择用于评估不同模型性能的指标时应考虑的各种因素。其中不仅包含传统的通用型指标，还介绍了面向PHM的指标。本节内容可以帮助操作人员为他们的PHM模型选择合适的性能评价指标。

3.6.1模型性能评价规范

完整的性能综合评价可以分为如下几个阶段：确立评价指标、确立标准数据集、收集数据、评估结果，如图3-8所示。



图3-8 模型性能评价规范流程

由于各种数据集的复杂性和随机性，同一模型在不同数据集上的性能表现可能参差不齐，因此需要确立标准数据集，标准数据集是检验算法性能好坏的评价基准之一，构建标准数据集有以下要求：

（1）数据集规模适中且类别分布平衡：一般工况下，数据集规模十分有限，且存在严重的分类非平衡问题。对于分类问题，这将直接影响到模型的性能。在构建标准数据集时应尽可能使数据集中不同类别数据所占的比例

（2）数据偏差小：由于不同的采集条件和数据标注的主观性，导致数据偏差大的问题在数据集中十分常见。数据集的数据偏差现象较弱时，能够客观真实地反映现实工况信息。

（3）标注信息一致：一个标准的数据集不仅要包含大量的标注而且确保要标注的准确性，准确的标注能够提升算法的效率。PHM作为一个数据驱动型任务，大量的标注数据对于模型捕捉装备特征参数变化而言十分重要。

3.6.2分类型模型性能评价指标

本方案将从通用指标和面向PHM的指标两方面介绍分类型模型性能评价规范。通用指标大多来自统计研究领域，是研究人员广泛使用的指标。PHM追求尽可能避免意外故障，并延长其使用寿命。因此，通用方法在工程应用中具有局限性，如通用指标没有考虑错误分类带来的成本影响等。针对PHM领域的特点，已有部分面向PHM的性能评价指标得到广泛应用。

（1）通用方法

对于故障预测，一种有效的方法是开发分类模型来预测故障。这种建模技术是利用分类算法开发分类器，这种分类器能够将系统运行状态分为“正常（负）”或“异常（正）”。由于预测的特殊性，对分类模型的评估更具挑战性。对于分类型任务，模型性能的评价指标包括通用方法和PHM领域方法。其中，通用方法又包括评分方法和图形方法。表3-7列出了分类型模型基于评分方法的性能指标及适用范围，表3-8为分类型模型基于图形方法的性能指标及适用范围。

（2）面向PHM的方法

通用方法，无论是评分方法还是图示方法，在分类器评估中都发挥了重要作用。但也有一些局限性。由于固有的缺陷，现有的通用方法可能无法验证分类器对于一些现实问题的性能。通用评估方法的不足可以总结如下。

首先，通用方法要求在评估中进行独立同分布（independent and identically distributed, i.i.d.）取样。实际上，来自现实问题的许多数据可能不满足实例独立性的要求。例如，来自预测应用的时间序列中的实例不是独立的，它们依赖于时间。基于i.i.d.的随机抽样可以将相关数据分成不同的组，如训练和测试数据集。事实上，与时间序列相关联的实例应该保持在同一个组中。

其次，通用度量本身也有一些不足。评分标准没有考虑错误分类的成本或评估的错误率。这是一个严重的问题，因为在不同的问题中，一些错误可能比其他错误花费更多。例如，在风洞设备健康状态管理应用中，设备的假阴性故障可能比假阳性故障要贵得多。最后，从通用指标中解释评估结果可能很难提取有意义的信息，甚至会产生误导。例如，ROC曲线下的面积（Area Under Curve, AUC）指标是一个介于0和1之间的标准化值。理论上，AUC值越高，分类器的性能越好。因此，AUC值为0.8的分类器应该优于AUC值为0.75的分类器。然而，这种解释对于最终用户来说可能没有意义或无用。从使用价值的角度来看，解释可能完全不同。

表3-7 分类型模型通用模型评价的性能指标

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 评价指标 | 公式 | 功能描述 | 使用前提或数据需求 | 取值 |
| 正确率或准确率（ACR） |  | 预测正确的样本数在样本总数中的占比 | 各类别样本的数量基本一致，不适合用于非平衡数据集。没有考虑错误分类的代价，不适合用于代价敏感问题 | ,完美模型值为1 |
| 错误率（MCR） |  | 预测错误的样本数在样本总数中的占比 |
| 误检率或虚警率（FPR） |  | 预测为正的负类样本在所有负类样本中的占比 |
| 漏检率（FNR） |  | 预测为负的正类样本在所有正类样本中的占比 |
| 精确率或查准率（Precision） |  | 预测为正的样本中真正的正样本占比 | 适用于全部数据集 |
| 召回率或查全率（TPR） |  | 样本中的正类被预测正确的比率 | 适用于全部数据集，但两者难以同时取得理想值，根据不同的需求，可以有偏向地关注这两个评价指标 |
| 几何平均（GM） |  | 倾向于最大化正召回率和负召回率，同时保持它们相对平衡。 | 能更客观评价分类器在非平衡问题中的表现，可用于非平衡数据集 |

续表4-3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 评价指标 | 公式 | 功能描述 | 使用前提或数据需求 | 功能边界 |
| 特异度（Specificity） |  | 表示的是所有负例中被分对的比例，衡量了分类器对负例的识别能力 | 普遍适用 | ,完美模型值为1 |
| F度量（FM） |  | F是P和R的加权调和平均，*β*通常取1 | P和R有时候会出现的矛盾的情况，此时需要综合考虑 |

注：True Positive（TP）表示把正类样本预测为正类；True Negative（TN）表示把负类样本预测为负类；False Positive（FP）表示把负类样本预测为正类；False Negative（FN）表示把正类样本预测为负类；*β*为超参数。非平衡数据集中各类别样本的数量明显不一致。

表3-8分类型模型通用模型评价的性能指标（图形方法）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 评价指标 | 公式 | 功能描述 | 使用前提或数据需求 | 取值 |
| ROC曲线 | / | 纵坐标为真阳性率（灵敏度）、横坐标为假阳性率 | 适用于二分类问题，能够在测试数据集正负样本分布的平衡性或分类阈值发生变化时，保持评估能力 | 完美模型取值1 |
| ROC曲线下的面积（AUC） |  | 接受者操作特征曲线（ROC）与坐标轴围蔽起来的面积 | ROC曲线一定程度上可以反映分类器的分类效果，但不够直观。AUC适用于非平衡数据分类，未考虑类分布或错误成本 | 值越大越好 |

注：其中和分别是正和负示例的数量，是测试数据集中所有示例的排名列表中第i个正示例的排名。ROC曲线绘制方法：1.分类模型输出样本类型，根据分类结果可以得到一个对应的混淆矩阵，然后计算FPR和TPR的值，得到一个(FPR,TPR)点坐标。通过不断地更改阈值，会得到不同的分类结果，也就是多个(FPR,TPR)点坐标。2.将FPR作为横轴，TPR作为纵轴绘制平面直角坐标系。然后将步骤1得到的点坐标(FPR,TPR)在平面直角坐标系中标出并且连接起来，就得到了曲线。

在实际应用中部署后，AUC值为0.8的分类器可能不会比AUC值为0.75的分类器更优。

因此，针对上述通用方法的局限，建议使用以面向PHM的模型性能评价指标作为补充。面向PHM的模型评估指标需要考虑的特性包括故障预测能力和剩余使用寿命估计。从剩余使用寿命的角度来看，期望预测模型能够在故障之前的目标时间窗口内生成警报。过早预测故障的模型会导致非最佳的组件使用。面向PHM的模型性能评价指标分为两方面：基于评分的方法和基于成本的方法，基于评分的方法普遍适用，而基于成本的方法是作为前者的补充，当用户关心计算结果所引发的成本问题时，可以选择基于成本的方法。

①基于评分的方法

基于评分的方法能够解决预测应用的特殊性和有效地评估预测模型的性能。基于评分首先定义奖励函数来预测正确的实例结果。预测正面实例的奖励基于警报生成和实际故障之间的故障时间，即确定的目标窗口。时间目标窗口被形成为。参数和是根据预测应用的要求确定的。当模型在部件故障之前预测目标窗口中的故障时，获得最大增益。在这个目标窗口之外，预测失败会导致负面的奖励阈值。因为这样的预测相当于误导性的建议。因此，假阳性预测(没有失败时的失败预测)与真阳性预测(有失败时的失败预测)的1.0奖励相比，会受到-1.5奖励的惩罚失败)。为了评估模型覆盖率，必须考虑不同故障情况下的警报分布。



其中，是测试数据集中正预测的数量；*NrDetected*是在目标时间间隔内至少包含一个警报的失败次数；*NrofCase*是在目标时间间隔内至少包含一个警报的失败次数；Sign是的符号，当Sign<0且brDetected=0时，Score设置为0；是使用上述奖励函数为每个警报计算。

使用测试数据集运行所有模型，并使用上式计算它们各自的分数。得分最高的模型被认为是应用程序的最佳模型。

②基于成本的方法

尽管基于评分的方法在评估预测分类器时考虑了RUL预测和问题检测覆盖率，但计算出的分数并没有告知最终用户预测模型的预期成本节约。成本可能会随着时间和部署环境的变化而变化。因此，提出了两种不同的成本节约估算指标：一种是准确的成本信息，另一种是不确定或遗漏的成本信息。当获得准确的成本信息时，可以使用节省成本的指标来估计最终用户的业务价值。通过使用该度量，需要四种成本信息：①错误警报的成本（没有组件更换的检查）；②早期更换的按比例的成本；③修复故障组件的成本以及；④未检测到的故障的成本（即，在没有任何来自预测模型的预先预测的情况下操作期间的功能故障）。前三个成本通常很容易获得，而最后一个成本很难精确估算。这是因为操作过程中的故障可能会导致各种其他成本，这些成本本身很难估计。以下是成本节约估算的细节。

为了估计一个模型的成本节约，通过下式计算没有模型的运行成本和有预测模型的成本之间的差异。





其中，*a*是提前更换的比例成本；*b*是虚假警报的成本；*c*是未检测到故障的成本(运行期间故障的直接成本)；*d*是更换组件的成本(故障后或警报后)；*N*是未检测到的故障数量；*M*是检测到的故障数量；；*F*是错误警报的数量。成本参数*a*、*b*、*c*和*d*由最终用户提供，、*F*、*M*和*N*在给定模型应用于测试数据集后计算。

3.6.3回归型模型性能评价指标

与分类问题一致，回归问题也需要良好的性能度量。以往更多的关注点是开发好的算法和模型。而不是评估性能，一个好的预测模型不仅应该改进RUL估计，还应该确保合理的预测范围和预测的置信水平。对此，有一些面向PHM领域的模型性能评价指标。本小节将从通用指标和面向PHM的指标两方面开展描述。

（1）通用方法

对于回归型模型，模型测试值与真实值之间的差距作为模型性能的评价指标。通用度量标准构成了大多数PHM性能度量的基础。鉴于用于确定PHM算法性能的基本方法依赖于将预测的RUL值与真实值进行比较，基于统计的度量是目前采用的最常见的度量。研究表明，均方误差(MSE)、均方根误差(RMSE)、平均绝对误差(MAE)和平均绝对百分比误差(MAPE)是测量人工智能算法性能最广泛使用的指标。表3-9列出了上述通用性能指标及适用范围。

（2）面向PHM的方法

回归问题的多数通用指标都是来自预测领域中用来衡量预测误差的指标，这些指标主要是统计性的，尽管基于统计的度量标准很受欢迎，并且仍在广泛使用，但一些研究人员已经为PHM模型开发了定制的性能度量，这些指标更适应RUL预测估计值随时间的推移而改善的程度问题，包括直接衡量算法性能的指标和基于成本效益的指标。在应用中，通用性能指标仍然很受欢迎，并且这些定制算法还没有被广泛接受和应用。因此，本方法有助于用户在流行的通用性能指标和特定于PHM的指标之间做出选择。

①性能度量

性能度量是衡量模型效果的评判标准，性能度量反映了任务需求，在对比不同模型的能力时，使用不同的性能度量往往会导致不同的评判结果。表3-10将更详细地描述四个主要指标及适用范围。

②基于成本的度量

上述指标都是为了直接衡量RUL预测的质量。然而，做出正确预测的实际好处是最大限度地减少与计划外干预相关的成本。成本效益指标衡量在业务中采用PHM的预期效益，如节约生命周期成本或降低风险。表3-11是基于成本考虑的主要指标。

表3-12为模型性能评价指标总表，汇总了本章所提的全部PHM模型性能评价指标。

3.6.4评价指标适用性分析

评价指标根据适用程度大致可分为完全适用和部分适用。

完全适用性指标：可完全覆盖对象PHM的性能评价要求，且指标测试难度低、计算方法明确、测试费用不高且测试周期较短如在电子系统PHM评价过程中，检测率、漏检率、检测门限值等评价指标的符合性均可定量化描述，指标计算和获取有特定的理论计算依据，指标可通过软硬件故障注入试验等方式快速获取，此类标准定义为完全适用类标准

部分适用类标准：指标可部分覆盖评价该对象PHM系统的功能性能，且有部分指标虽可作为评价该系统的依据，但验证评价过程中指标获取难度较大，成本较高，部分指标评价结果准确度较低，且部分指标无法定量描述，评价客观度较低，部分指标只能参考类似系统或经验进行定性评价，如在液压、传动、结构等系统或部件中，针对监测率、漏检率等指标，由于系统的退化类故障模式较多，指标验证过程中故障注入试验时间长，且由于试验样本数量的局限性，评价结果的置信度较低大多数情况下只能依靠同类型号产品或这理论计算值定性评价，指标的适用性有一定的局限性。

表3-9 回归任务PHM模型通用性能评价指标

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 模型性能评价指标 | 公式 | 意义 | 使用前提或数据需求 | 取值 |
| 均方误差  （MSE） |  | 评估预测值与真实值之间的误差均方和 | MSE可以评价数据的变化程度，MSE的值越小，说明预测模型描述实验数据具有更好的精确度 | ,完美模型值为0 |
| 均方根误差（RMSE） |  | 评估预测值与真实值之间的误差均方根 | 均方根误差是均方误差的算术平方根，对一组测量中的特大或特小误差反映非常敏感，能够很好地反映出测量的精密度 |
| 平均绝对误差（MAE） |  | 评估预测值与真实值之间的误差绝对值之和 | MAE是一种线性分数，所有个体差异在平均值上的权重都相等，避免偏差相互抵消的问题 |
| 平均绝对百分比误差（MAPE） |  | 对同一预测范围内多个被测单元预测的绝对百分比误差进行平均，降低了个别离群点带来的绝对误差的影响 | MAPE为0%表示完美模型，MAPE大于100%则表示劣质模型。当真实值有数据等于0时，存在分母0除问题，该公式不可用 |
| 对称平均绝对百分比误差（sMAPE） |  | 对同一预测范围内多个被测单元预测的绝对百分比误差进行平均。 | 百分比是根据预测的平均值和实际情况计算的。这可以防止百分比误差对于接近真相的情况来说太大了 |

注：为回归序列长度，和分别为样本的真实值序列和预测值序列，为样本的预测误差，为样本的预测误差平均值。

表3-10 回归任务面向PHM模型的性能评价指标

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 指标 | 公式 | 描述 | 使用前提或数据需求 | 取值范围 |
|  |  | 通过确定在指定时间的预测是否在规定的限度内，对预测质量进行量化 | 在特定时间λ的预测精度是否在实际RUL的α\*100%内。需要RUL真值。使用灵活，可以实现可视化 |  |
| Prognostic Horizon |  | 预测结果首次满足指定的预测准确度要求的时间点与设备失效时间点之间的时间差值 | 这个度量指标表明预测的估计是否在实际EoL周围的指定范围内。指标性能可以根据真实EoL附近的允许误差界(α)来指定。易于计算和理解 | ,完美模型值为0 |
| 收敛指标 |  | 它衡量任何指标随时间而改进的方式，例如，当预测朝着EoL前进时，它朝着实际RUL收敛的速度有多快 | 它是使用根据预测结果计算的其他性能度量来计算的。需要RUL真值，难以处理大范围预测 | ,完美模型值为 |
| 相对精度 |  | RUL预测相对于真实RUL的误差的瞬时测量 | 用于比较多个模型，需要RUL真值 | ,完美模型值为1 |

注：PH指标中，EOL：End-of-Life；i是预测满足α界时的第一个时间指数。收敛指标中，EOP：End-of-Prediction；i是第i个被测单元；M(i)是非负预测误差精度或精度度量。指标、RA指标中，l是被测单元的索引合；为λ时刻系统的真实RUL；为λ时刻系统的预测RUL。

表3-11 基于成本的指标

|  |  |  |  |
| --- | --- | --- | --- |
| 指标 | 公式 | 描述 | 适用前提或数据需求 |
| 生命周期成本 |  | 计算PHM系统下的采购、运营和维护的总成本，并与没有PHM决策系统时的成本进行比较。为了证明采用PHM的合理性，PHM的成本应该更低 | 适用于全部数据集 |
| 平均故障间隔时间与平均故障间隔时间之比 |  | 它被定义为平均故障间隔时间(通过可靠性方法估计)与平均设备更换间隔时间(在PHM实施后)之比。这一指标表明了预测的有效性。较低的比率表征了PHM决策系统有较高的效率 |
| 投资汇报率 |  | 年平均利润占PHM实施初始投资的百分比 |
| 技术价值 | / | 技术价值衡量关键故障模式正确预测的好处，以及错误预测和相关资源需求的成本。另一方面，总价值着眼于PHM系统能够有效覆盖的所有故障模式的好处，减去与PHM实施相关的所有成本。 |

注：为平均故障间隔时间；为平均设备更换间隔时间。

表3-12 模型性能评价指标总表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 问题类型 | 方法 | 指标 | 使用前提或数据需求 | 取值范围 |
| 分类问题 | 通用方法 | 正确率 | 不适合用于非平衡数据集 |  |
| 错误率 |
| 虚警率 |
| 漏检率 |
| 查准率 | 普遍适用 |
| 召回率 |
| 几何平均 |
| 特异度 |
| F度量 |
| ROC曲线 | 适用于二分类问题 |  |
| ROC曲线下面积 | 普遍适用 | 越大越好 |
| PHM定制 | 基于评分 | 适用于各类PHM模型 | 无限制 |
| 基于成本 | 无限制 |
| 回归问题 | 通用方法 | 均方误差 | 普遍适用 |  |
| 均方根误差 |
| 平均绝对百分比误差 | 当真实值有数据等于0时，存在分母0除问题，该指标不可用 |
| 对称平均绝对百分比误差 | 普遍适用 |
| PHM定制  （基于模型性能） |  | 需要RUL真值 |  |
| PH指标 | 普遍适用 |
| 收敛指标 | 需要RUL真值，难以处理大范围预测 |  |
| 相对精度 | 用于横向比较多个模型，需要RUL真值 |  |
| PHM定制  （基于成本） | 生命周期成本 | 需在计划财务部门支持下使用 | 无限制 |
| 平均故障间隔时间与平均故障间隔时间之比 |
| 投资汇报率 |
| 技术价值与总价值 |

表3-13 评价指标适用性分析结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 类别 | 指标名称 | 适用模型 | | | |
| 故障诊断 | 健康状态评估 | 故障预测 | 剩余寿命预测 |
| 模型性能评价 | 正确率 | √ |  | √ |  |
| 错误率 | √ |  | √ |  |
| 虚警率 | √ |  | √ |  |
| 漏检率 | √ |  | √ |  |
| 查准率 | √ |  | √ |  |
| 召回率 | √ |  | √ |  |
| 几何平均 | √ |  | √ |  |
| 特异度 | √ |  | √ |  |
| F度量 | √ |  | √ |  |
| ROC曲线 | √ |  | √ |  |
| ROC曲线下面积 | √ |  | √ |  |
| 基于评分 | √ |  | √ |  |
| 基于成本 | √ |  | √ |  |
| 均方误差 |  | √ |  |  |
| 均方根误差 |  | √ |  |  |
| 平均绝对百分比误差 |  | √ |  |  |
| 对称平均绝对百分比误差 |  | √ |  |  |
|  |  |  |  |  |
| PH指标 |  | √ |  |  |
| 收敛指标 |  | √ |  |  |
| 相对精度 |  | √ |  |  |
| 生命周期成本 |  | √ |  | √ |
| 平均故障间隔时间与平均设备更换间隔时间之比 |  | √ | √ | √ |
| 投资汇报率 |  | √ | √ | √ |
| 技术价值与总价值 |  | √ |  | √ |

3.7 模型更新规范

PHM模型更新的目的是及时、准确、科学地对模型进行必要的修改与完善，模型更新是实现PHM模型高效运维的基础。PHM模型更新需要经过模型备份、参数更新等一系列过程，为了保证模型更新过程中模型数据的一致性需要对模型更新过程进行规范。本节将介绍PHM模型更新流程和更新日志规范。

3.7.1 PHM模型更新流程

PHM更新流程规定了规范操作的基本要求，旨在指导操作人员科学进行模型的更新维护工作，确保模型更新准确、有效。整个模型更新可分为模型备份、参数更新和撰写更新日志，具体流程如图3-9所示，下面将按照此流程展开描述。

首先，进行模型备份。PHM模型应按本章存储规范中的数据备份模块进行容灾备份，容灾备份功能是增强数据持久保存的工具。操作人员应根据计算存储和带宽约束等资源约束并确定备份机制来满足所有的容灾备份需求。由于平台要求高可靠性，建议采用冗余备份机制。冗余备份机制指重复地在不同节点上备份，备用服务节点可以在故障发生时及时替换故障服务节点，实现服务的快速恢复。冗余备份是应对单节点故障的有效方法。



图3-9 模型更新流程图

其次，对模型进行参数更新。一个PHM模型通常具有一个称为主变量（primary\_variable）的模型参数，若干称为辅助性变量（secondary\_variable）的参数以及和主变量相对应的更新参数（update\_variable）。模型的核心函数（KernelFunction）在执行过程中，会产生相应的更新数据，因此需要描述各个变量的更新逻辑，即如何根据执行过程中产生的更新数据去更新相应的变量。主变量和各个辅助变量之间的键具有不同的映射关系，因此需要将这种映射关系记录在变量中。这样，运行时系统会根据主变量的键以及先前提供的各种键的映射函数，自动推导出所需的各种辅助性的变量的键，然后根据核心函数的逻辑进行运算，产生更新数据。当进行模型参数的更新操作时，运行时系统通过调用更新函数（Update Function），将更新数据推送到模型参数的维护用户，当模型参数的维护用户收到更新数据时，利用先前维护的更新逻辑去更新相应的变量。

用户应在管理系统中开发一个名为Variable的API，用于描述各种变量，它是张量的进一步封装。Variable对象内部应含有一个映射函数的向量（KeyProjection），这个向量保存了编程人员编写的主变量和各个辅助性变量之间的键的映射关系。主变量和其对应的更新变量的键的映射关系蕴含的是一对一的关系，也即主变量的键和主变量的更新变量的键是相同的。KeyProjection描述一类相同类型的函数，该类型的函数的参数是两个键向量的引用，而返回值为空。这两个键向量，一个表示源键向量（也即主变量的键向量），另一个表示目标键向量（某个辅助性变量）。针对变量的更新函数，它描述的是如何根据主变量的更新变量更新主变量的值。更新函数是这样一类的函数：函数的参数是键向量的引用，而返回值为空。参数更新规范流程图如图3-10所示。



图3-10 模型参数更新规范

下面举例说明模型参数的更新过程。假设PHM平台应用的某个阶段需以下模型参数变量V1、UpdateV1、V2和V3。其中V1是主变量而V2和V3是辅助性变量，UpdateV1是针对V1的更新变量。主变量V1被分布式存储在参数维护节点，UpdateV1、V2、V3都是局部模型变量。

该阶段的核心函数、主变量的更新函数、更新变量和其他辅助变量的键映射函数如图3-11所示。从图3-11中可以看到，更新UpdateV1的新值需要辅助参数模型V2和V3，且根据UpdateV1的更新函数UpdateV1\_update可以看出，UpdateV1的键和V2以及V3的键的关系是：V2和V3的键是UpdateV1的键的2倍和3倍的倍数关系；更新变量的更新逻辑蕴含在该阶段内的kernel函数中，而主变量V1的更新逻辑是对于V1的一个确定的键key，V1[key]等于UpdateV1[key]。此外，图3-11中的注释部分表明更新逻辑也可以是V1[key]等于其自身加上UpdateV1[key]。

最后，撰写更新日志。撰写更新日志的目的是，当模型工作不正常时及早发现，发生严重问题是可以根据日志快速定位到之前问题原因。下面将给出PHM模型管理日志规范。

|  |
| --- |
| KernelFunction kernel = [&](vector<Key> & range){  Foreach(update\_id:range){  UpdateV1[update\_id] = 10\*V2[2\* update\_id] + 20\*V3[3\* update\_id];  }  }  UpdateFunction V1\_update = [&](vector<Key> & range){  Foreach(key:range){  V1[key] = V1\_update[key];  //V1[key] += V1\_update[key];  }  }  KeyProjcetion UpdateV1\_v2 = [&](vector<Key>& src,vector<Key>&target){  Foreach(update\_k:src){  target.push\_back(2\* update\_k);  }  }    KeyProjcetion UpdateV1\_v3 = [&](vector<Key>& src,vector<Key>&target){  Foreach(update\_k:src){  target.push\_back(3\* update\_k);  }  } |

图3-11 更新函数示例图

3.7.2更新日志规范

对于每一条日志应含有的信息包括日期、时间、日志级别、模型代码、日志内容、版本号、操作人员、更新目的、更新结果等信息，其相应内容需要规范格式。

①日期和时间：应采用统一格式，以准确描述日志产生的日期。日期格式示例：2018-05-22；时间格式示例：15:35:53。

②日志级别：通过定义每一条日志信息的级别，能够更加细致地控制日志的生成过程。六级日志级别示例：TRACE、DEBUG、INFO、WARN、ERROR和FATAL。日志级别对应描述如表3-14所示

表3-14 日志级别级描述

|  |  |
| --- | --- |
| 日志级别（由高到底） | 描述 |
| FATAL | 表示需要立即被处理的模型级错误。当该错误发生时，表示服务已经出现了某种程度的不可用，模型管理员需要立即介入。 |
| ERROR | 该级别的错误也需要马上被处理，但是紧急程度要低于FATAL级别。当ERROR错误发生时，已经影响了用户的正常访问。 |
| WARN | 该日志表示模型可能出现问题，也可能没有，如模型的计算超时等。这种情况对于那些目前还不是错误，然而不及时处理也会变为错误的情况 |
| INFO | 该种日志记录模型的正常运行状态，例如某个子模型的初始化，某个请求的成功执行等等。 |
| DEBUG和TRACE | 这两种日志具体的规范应该由用户自己定义。该级别日志的主要作用是对模型每一步的运行状态进行精确的记录。通过该种日志，可以查看某一个操作每一步的执行过程，可以准确定位是何种操作，何种参数，何种顺序导致了某种错误的发生。 |

③模型代码：详见本方案第三部分——模型库技术要求。

④日志内容：日志内容应包含模型名、更新内容、更新位置、更新原因等。

⑤版本号：版本号是更新版本的标识号。版本号能使用户了解所使用的模型是否为最新的版本。版本号示例：以3段版本号为例，分为X.Y.Z三段，分别代表主版本号、次版本号和补丁版本号。

⑥操作人员：日志系统应自动记录操作人员信息。具体地，当操作人员完成一次更新其ID或计算机用户ID反映在用户标识字段中。

⑦更新目的：更新目的从表3-15中的6个方面体现。

表3-15更新目的选项表

|  |  |
| --- | --- |
| 目的 | 描述 |
| 新增（Features） | 新增功能 |
| 修复（Fixed） | 修复bug |
| 变更（Changed） | 对于某些已存在功能所发生的逻辑变化 |
| 优化（Refactored） | 性能或结构上的优化，并未带来功能的逻辑变化 |
| 即将删除（Deprecated） | 不建议使用或在以后的版本中即将删除的功能 |
| 删除（Removed） | 已作废的功能 |

⑧更新结果：更新结果应指出当前版本和上一个版本之间的差异和效果。

更新日志示例：

|  |
| --- |
| 日期：<date:yyyy-mm-dd>  时间：<time:xx:xx:xx>  日志级别：<log-level>  模型代码：<model-id>  日志内容：<log-message>  用户标识：<user-id>  更新目的：<update-aim>  更新结果：<results-message>  版本号：<x.y.z> |

3.8 模型简介规范

模型简介是指使用一小段简洁精炼的文字对模型的功能、主体、输入和输出等进行描述，能够方便读者更快速地理解模型，因此在编写模型简介时，应当具备以下特点：

（1）简要：模型简介的篇幅不宜过长，应当抓住模型的关键点进行阐述。

（2）完备：模型简介的内容编写必须包含模型的属性及功能简介。

（3）清晰：编写的模型简介应当正确描述模型，清晰无二义性。

例：基于卷积神经网络的低速风洞风扇故障诊断模型，其模型简介可写为“基于卷积神经网络的低速风洞风扇故障诊断模型，其主体为卷积神经网络，输入为1024×1024的RGB图像，输出为故障诊断结果（如“0”为“正常”，“1”为“风扇叶片故障”等），模型采用API调用格式为：FAN\_RNN\_FD\_001(include\_top=True, weights='imagenet', input\_tensor=None, input\_shape=None, pooling=None, classes=1000)，模型可用于恒定工况下风扇的故障诊断，现阶段在基础验证数据集上故障诊断准确率为100%。

其中，模型API调用格式的解释如下：

该模型在Theano和TensorFlow后端均可使用,并接受channels\_first和channels\_last两种输入维度顺序。

模型的默认输入尺寸时224x224

①参数：

nclude\_top：是否保留顶层的全连接网络。

weights：None代表随机初始化，即不加载预训练权重。'imagenet'代表加载预训练权重。

input\_tensor：可填入Keras tensor作为模型的图像输出tensor。

input\_shape：可选，仅当include\_top=False有效，应为长为3的tuple，指明输入图片的shape，图片的宽高必须大于197，如(200,200,3)。

pooling：当include\_top=False时，该参数指定了池化方式。None代表不池化，最后一个卷积层的输出为4D张量。‘avg’代表全局平均池化，‘max’代表全局最大值池化。

classes：可选，图片分类的类别数，仅当include\_top=True并且不加载预训练权重时可用。

②返回值：用户定义模型对象。

4 平台PHM模型库技术要求

为了将数量庞大且类型各异的PHM模型统一到一体化平台中，就必须构建PHM模型库，对各种各样的模型进行集成，构成一体化平台的“工具库”，与数据库和知识库相互独立并行构建，分别为PHM工作的开展提供工具、数据原料和规则。因此，本部分针对将要开发PHM模型库，从PHM模型库设计原则、开发技术原则和设计说明三方面，提出了相应的技术要求和规范。

4.1 PHM模型库设计原则

①用户友好原则：PHM模型库是为应用人员而不是为机器设计的应用程序接口（Application Programming Interface，API）。它把用户体验放在首要和中心位置。PHM模型库遵循减少认知困难的最佳实践：提供一致且简单的API，将常见用例所需的用户操作数量降至最低，并且在用户错误时提供清晰和可操作的反馈。

②模块化原则：PHM模型库被理解为由独立的、完全可配置的模块构成的序列或可视化控件。这些模块可以以尽可能少的限制组装在一起。比如，不同的模块都是可以结合起来构建新模型的模块。

③易扩展性原则：新的模块是很容易添加的（作为新的类和函数）。由于能够轻松地创建可以提高表现力的新模块，PHM模型库更加适合高级研究。

④基于编程语言或可视化控件实现原则：PHM模型库没有特定格式的单独配置文件。模型定义在程序或可视化控件中，模型的代码紧凑易于调试且易于扩展。

4.2 PHM模型库开发技术要求

①PHM模型库中的模型需要给出明确的程序或可视化控件源代码存储地址，以方便后续的检索和调用；

②PHM模型库中模型需要明确调用接口、调用格式、调用参数的数据类型和范围；

③PHM模型库中模型需要给出测试案例或程序，明确模型的运行方式、适用范围和功能边界等。

4.3 PHM模型库设计说明

（1）PHM模型库技术指标要求

PHM模型库作为一种存储模型各项属性的数据库，应具备以下技术指标要求：

①具备支持跨平台异类数据库的良好数据互联，实现对文件数据和文件系统的高效访问；

支持当前市场流行的的应用拓扑架构，如C/S、B/S等；

②具有保密性和安全性设计，实施正确有效的权限管理，保证只有经过认证且符合权限的用户可以合法地读写数据库；

③数据库基于主流通用的关系型数据库进行开发。

④PHM模型数据库除了需要合理存储相关数据，还需保证在多人同时访问数据库时正常工作，将数据库的响应时间控制在较短的时间内，数据库还应具有一定的可扩展性以存储目前没涉及到的数据。

（2）功能设计

数据库功能设计包括PHM模型库模块设计和系统综合管理模块设计。其中数据库模块包括PHM模型库维护管理子模块、数据导入导出子模块和数据备份与恢复3个子模块。

PHM模型库维护管理子模块用于定期或不定期存储、更新模型各项属性等信息，便于用户对数据进行快速查阅和使用。

数据导入导出子模块用于实现良好的数据互联，快速响应数据读取和输出功能。由于部分监测内容下的数据量较大，可能使用文件形式存储，因此数据库需要有导入和导出规范化文件的功能。

数据备份与恢复子模块用于保证数据库的安全性，防止因系统崩溃或其他意外关闭情况导致数据丢失和不完整等情况发生。

PHM模型库中，针对特定的模型的部分存储属性和信息如表4-1所示。

表4-1 PHM模型的部分存储属性和信息表

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| PHM  模型名称 | 模型编码 | 模型分类 | 模型功能 | 存储地址 | 存储类型 | 调用格式 | 输入类型 | 输出类型 | 模型参数 | 模型中间运算结果 | 适用范围 | 特点 | 模型性能 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |

其中，模型名称严格遵循3.3.1节中模型命名规范进行填写。模型编码遵循3.1节模型编码规范进行填写。模型分类遵循3.1.5节中模型原理类别的说明和分类依据进行填写，模型功能遵循模型3.1.5节中任务与功能类别的说明和分类依据进行填写。存储地址用于存储模型源文件或可执行文件的计算机硬盘地址，也可用于存储模型的动态链接库等。存储类型需要明确填写模型存储的形式，如模型源文件、可执行文件、动态链接库等。调用格式填写模型调用的方式（如：API、可视化控件等）和对用的调用接口。输入输出类型需要明确模型的输入输出表或字段的内容和形式。模型中间运算结果通过表的形式被暂存到数据库中，以方便后续对模型的调整。适用范围和特点通过文本的方式被存储到数据库中。模型性能存储模型性能评估后的数值量化结果，具体可遵循3.6节模型性能评价规范进行填写。模型参数的填写和存储需要遵循数据库中数据表的要求和规范，其中部分模型的数据库表设计如下。

a.维纳过程

维纳过程包含的参数有漂移参数和扩散参数，均为浮点类型数值，因此可以在存储时指定其字段类型为float。其参数表名设置为wiener\_parameter，在对字段进行命名时，使用“wiener\_”+参数英文名的形式进行命名，如wiener\_drift表示维纳过程中的漂移参数。

其二维表结构可以设计成如下表4-2所示：

表4-2 wiener\_parameter二维表结构

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 类型 | 是否为空 | 注释 |
| wiener\_drift | float | 否 | 维纳过程模型中的漂移参数 |
| wiener\_diffusion | float | 否 | 维纳过程模型中的扩散参数 |

b.伽马过程

伽马过程包含的参数有形状参数和尺度参数，均为浮点类型数值，因此可以在存储时指定其字段类型为float。其参数表名设置为gamma\_parameter，在对字段进行命名时，使用“gamma\_”+参数英文名的形式进行命名，如gamma\_shape表示维纳过程中的形状参数。

其二维表结构可以设计成如下表4-3所示：

表4-3 gamma\_parameter二维表结构

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 类型 | 是否为空 | 注释 |
| gamma\_shape | float | 否 | 伽马过程模型中的形状参数 |
| gamma\_scale | float | 否 | 伽马过程模型中的尺度参数 |

c.逆高斯过程

逆高斯过程包含的参数有形状参数和尺度参数，均为浮点类型数值，因此可以在存储时指定其字段类型为float。其参数表名设置为inverse\_gaussian\_parameter，在对字段进行命名时，使用“inverse\_gaussian\_”+参数英文名的形式进行命名，如inverse\_gaussian \_shape表示逆高斯过程中的形状参数。

其二维表结构可以设计成如下表4-4所示：

表4-4 inverse\_gaussian\_parameter二维表结构

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 类型 | 是否为空 | 注释 |
| inverse\_gaussian \_shape | float | 否 | 逆高斯过程模型中的形状参数 |
| inverse\_gaussian \_scale | float | 否 | 逆高斯过程模型中的尺度参数 |

d.神经网络模型

神经网络模型包含的参数表中保存的格式为“模型名字模型参数路径”，因此可以将多个模型参数共同保存在同一张表中。在表中设置两个字段，第一个字段命名为“model\_name”，类型设置为varchar，第二个字段命名为“parameter\_path”，类型同样设置为varchar，表的名字设置为neural\_network\_parameter。

其二维表结构可以设计成如下表4-5所示：

表4-5 neural\_network\_parameter二维表结构

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 类型 | 是否为空 | 注释 |
| model\_name | varchar | 否 | 神经网络模型名称 |
| model\_path | varchar | 否 | 神经网络模型存放地址 |
| model\_parameter | float | 否 | 神经网络模型参数 |
| parameter\_path | varchar | 否 | 神经网络模型参数存放路径 |

此外，系统的综合管理模块包括用户注册信息及权限管理、数据资源管理以及日志管理3个子模块，用于对系统访问用户登录信息、访问权限以及PHM模型的信息进行增加、删除和修改，并对系统的操作日志进行管理以便对非法操作行为进行审计追踪。

①用户和权限管理子模块

本子模块用于对用户登录信息及访问权限进行管理，包含的功能如下：

a.用户创建：系统管理员可创建新的用户，录入用户信息，为用户分配权限类别以及创建初始密码，访问用户注册信息如表4-6所示；

表4-6 访问用户注册信息表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 用户名 | 登录密码 | 个人姓名 | 所属部门 | 权限等级 | 人员岗位 | 用户状态 |
|  |  |  |  |  |  |  |

b.用户管理：系统管理员可以可折叠的树状结构浏览、删除和修改用户和用户信息；系统管理人员可查看和修改用户的登录密码；系统管理员可对用户执行启用和停用的操作；

c.部门管理：系统管理员可创建、修改和删除部门和部门信息；

d.权限管理：系统管理员可浏览和管理系统所有的可分配权限，并修改权限的名称和分组。本软件设置3种访问权限，不同权限的用户只能对特定的软件功能模块进行操作，如表4-7所示；

表4-7 访问用户权限设置表

|  |  |
| --- | --- |
| 权限类别 | 权限具体描述 |
| 超级管理员  （角色1） | 可以使用系统软件的所有功能进行系统维护管理，能查看所有数据内容，可进行所有软件模块操作。 |
| 管理人员  （角色2） | 可以查看本系统所有数据分析结果信息，但不能进行数据修改。 |
| 负责人（角色3） | 可以录入、修改和查询所负责PHM模型的所有信息，对本人负责的设备进行健康管理评估和故障诊断分析。 |

e.用户类别管理：超级管理员可创建、修改和删除各个权限类别，为每个用户类别分配访问权限；每个用户类别的用户只能访问到有权查看的模块、页面、功能点和数据内容；初始用户类别如下表所示，超级管理员可根据需求增加或减少每个类别的权限（系统管理员类别除外），或增加更多的用户类别。

②数据资源管理子模块：

本子模块用于对PHM模型的信息进行数据库管理和可视化显示，具体包含功能如下：

a.参数一览：能以可折叠的树状结构分类显示各类PHM模型信息概要；

b.参数修改：能以可折叠的树状结构添加、修改和删除关键的有关信息。

③日志管理子模块：

本子模块对系统各类日志信息进行管理，以便系统管理员通过查询操作日志详细掌握有关访问用户对系统软件功能模块的操作行为、IP地址和访问时间等信息，对访问用户的数据篡改、恶意操作等非法行为进行责任追踪。系统涉及的日志包括用户（含系统管理员和一般访问用户）操作日志、系统异常日志和账户安全日志三种，具体功能如下：

a.用户操作日志：记录用户对系统各个软件子模块的操作，包括用户的真实信息、注册用户名、操作模块、操作内容、IP地址、日志信息类别（分为调试、报告、警告、错误、严重警告）和操作时间，操作日志信息如表4-8所示；

表4-8 操作日志信息表

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 登录人  姓名 | 登录人注  册用户名 | 软件操作  功能模块 | 操作  类型 | 操作  内容 | 访问计算机  IP地址 | 日志信息  类别 | 访问时间 |
|  |  |  |  |  |  |  |  |

b.系统异常日志：记录系统运行过程中产生的预警和报错信息，如故障报错、意外中断、存储空间不足等；

c.账户安全日志：对用户的异常行为如使用错误密码登录等进行记录，对用户不合理的数据修改及潜在的恶意操作等非法行为进行识别和记录；

d.日志管理：系统管理员能够导出、删除和清空已经存储的日志信息，同时可以通过关键词、用户、时间范围等方式查询日志信息。

（3）详细设计方案

根据以上提出的PHM模型库设计与研发要求，本节制定了更为详细的设计方案，包括数据库选型、数据库设计、技术指标、数据库性能等。

①数据库选型

PHM模型库的选型主要考虑可靠性与安全性，可靠性高的数据库需要保证数据的一致性，目前主流数据库可分为关系型数据库与非关系型数据库，非关系型数据库注重最终一致性；关系型数据库注重数据整个生命周期的强一致性。

针对系统数据存储的设计中，数据存取效率是考虑的重点。本系统将覆盖对各类PHM模型的存储，存储对象为PHM模型的各项属性和信息以及最新记录的模型运行输出数据等，包括文档类数据存储和模型类数据存储。总的来说，存在以下三个方面的特点：

a.数据种类繁多：本次系统将覆盖文档类数据来源于换料检修相关技术总结报告、现场试验、工程实施数据记录文件、工艺说明书、设备说明书等，经过整理和提取后形成易于存储和提取的结构化数据，以及相关设备工艺的图片等信息数据。模型类数据主要是由三维软件设计生成的三维数字化模型文件等；

b.数据存取效率要求高：大量数据在采集完成后，需要以最快的速度完成结构化并进行分析处理，并提供运维人员查看。

c.历史表具备1年历史数据的在线存储能力，支持对历史数据的在线检索功能。

针对以上数据存储方面的特点，由于PHM模型库系统对数据库的高可靠性要求，项目组选择了有事务机制且能确保数据强一致性的Oracle、DB2、SQL/Access和MySQL等关系型数据库。

通过使用关系型数据库，PHM模型库系统可以对用户行为进行约束，并能有效满足涉密数据管理要求。

②数据库设计

数据库的设计开发分为六个阶段：需求分析阶段、概要设计阶段、逻辑设计阶段、物理设计阶段、实施阶段和运行维护阶段。

a.需求分析阶段的重点是分析数据与处理过程，在模型库系统项目中，就是结构化、规范化存储涉及到的所有数据，了解数据项、数据结构、数据流、处理过程，最终形成数据字典，作为后续设计开发数据库表的基础；

b.概要设计阶段需要将需求分析阶段的需求转化为抽象模型，在模型库系统项目中，就是将各个PHM模型对象抽象为诸多实体以及实体之间的关系，绘制实体关系图，形成关系数据库的基本雏形；

c.逻辑设计阶段需要将概要设计阶段的实体关系转化为逻辑结构，将实体转化为表，关系转化为外键，并进行一定的优化；

d.物理设计阶段就是为数据库设计物理存储结构，即合适的数据库管理系统，定义PHM模型的数据库、表、字段的命名规范，选择合适的字段数据类型，并进行一定的反范式化处理，提高查询和维护的效率；

e.实施阶段指在服务器中创建数据库，将数据存入数据库中，进行试运行，包括功能测试及性能测试，即测试数据库的增删查改等功能，以及数据库的响应时间，同时响应的最大请求数量等；

f.运行维护阶段指数据库投入使用后的管理工作，包括监控数据库的运行状况，进行数据一致性检查，磁盘空间检查，数据库备份与恢复。数据库备份与恢复是其中最关键的步骤，PHM模型数据库系统中有专门的数据库管理模块进行数据库的日常维护工作。

③技术指标

a.数据库系统：数据存储总容量不低于2TB，支持存储容量平滑扩展；

b.数据库响应：数据库查询响应速度在1000ms以内；

c.数据接口：采用具有签名机制的JSON API设计，保证只有经认证的符合权限的用户可以合法地读写数据库数据；

d.存储管理：但系统存储容量低于10%或数据库存储异常，及时向系统管理人员发出预警；

e.分布式架构：使用应用服务器和数据备份服务器完成数据库分布式架构。

④数据库性能

a.数据库系统依照国内外软件可靠性相关标准规范进行研制与开发，采用的算法与报表满足相关国家军用标准。

b.数据库查询响应时间控制在1000毫秒以内，专家系统数据库设计通过以下措施保证数据操作的性能：软件具有灵活的功能扩展性，便于后续结合实际需求进行软件二次开发；

c.系统具有良好的容错性，对有Bug的程序写入的错误数据有足够的适应能力，对人为操作具有较好的容错性；

d.系统软件具备不少于20个用户终端的并发访问能力；

e.系统平均无故障时间（MTBF）≥1000小时，平均恢复时间（MTTR）≤30分钟，可7\*24小时持续运行；

f.系统运行期间，CPU平均占用率≤60%，内存资源总占用率≤60%，用户界面响应迅速。

g.系统支持主流的关系型数据库；

h.系统数据库初步设计为可支撑5000套以上设备的可靠性数据以及设备资源信息存储，数据存储总容量不小于2TB，支持存储容量平滑扩展；

i.支持EXCEL表格、WORD或PDF格式的可靠性分析评估结果数据导出；

j.存储数据安全，当系统存储容量低于10%或系统存储设备损坏，应及时提醒系统管理员。

参考文献

1. 范彬. 面向飞行器关键部件健康管理的故障预测方法研究[D]. 国防科学技术大学, 2015.
2. 李旻. 复杂电子系统综合诊断与故障预测关键技术研究[D]. 电子科技大学, 2014.
3. 姜正. 水轮发电机组轴系故障趋势预测与评估[D]. 华中科技大学, 2012.
4. 肖飞. 基于时间序列分析和智能算法的故障预测方法研究[D]. 北京化工大学, 2014.
5. 王岩. 基于威布尔比例风险模型的滚动轴承剩余寿命预测[D]. 哈尔滨理工大学, 2020.
6. 潘东辉. 基于退化数据的产品可靠性建模与剩余寿命预测方法研究[D]. 华中科技大学, 2014.
7. 袁容. 基于性能退化分析的可靠性方法研究[D]. 电子科技大学, 2015.
8. 马雁春. 基于数据挖掘的航空PHM中预测方法的研究[D]. 南京航空航天大学, 2010.
9. 尚小税. 基于大数据的公交车辆状态细分与故障预测研究[D]. 北京交通大学, 2018.
10. 苏航. 基于BP神经网络的地铁信号设备故障预测[D].华南理工大学,2013.
11. 张炳森. 轨道交通自动检票机设备故障预测技术研究[D]. 东南大学, 2018.
12. 梁天辰. 基于多深度置信网络融合的航空电子设备故障预测[J]. 电讯技术, 2021, 61(02): 248-253.
13. 计炜梁. 基于深度学习的无线网络日志故障预测算法研究[D]. 中国科学技术大学, 2018.
14. 李旭东. 基于深度学习的旋转部件故障诊断研究[D]. 中国科学院大学(中国科学院国家空间科学中心), 2021.
15. 祁丽洁. 基于高斯过程回归模型的贝叶斯滤波故障诊断方法研究[D]. 北京交通大学, 2017.
16. 张晔. 某高压电源板的早期故障检测与性能退化预测[D]. 哈尔滨工业大学, 2019.
17. 董晗拓. 基于隐马尔可夫模型的结构损伤识别研究[D]. 哈尔滨工业大学, 2019.
18. 夏丽莎. 基于隐马尔可夫模型的故障诊断及相关算法研究[D]. 华中科技大学, 2014.
19. 陈扬文. 支持向量机在航空发动机气路故障诊断中的应用[D]. 上海交通大学, 2014.
20. 李永亮. 基于机器学习的故障预测与健康管理（PHM）方法研究[D]. 电子科技大学, 2017.
21. 吴渊. 基于BP神经网络的车载设备故障诊断与预测研究[D]. 北京交通大学, 2016.
22. 李健. 基于深度学习的变循环发动机气路故障诊断[D]. 上海交通大学, 2019.
23. 代杰杰. 基于深度学习的变压器状态评估技术研究[D]. 上海交通大学, 2018.
24. 孟祥峰. 基于深度学习的滚动轴承的故障诊断及预测[D]. 电子科技大学, 2020.
25. 廖广纯. 数据驱动的设备健康评估与维修决策[D]. 华中科技大学, 2016.
26. 崔诗好. 基于劣化隐马尔可夫模型的设备状态评估研究[D]. 重庆大学, 2018.
27. 张静. 基于数据的风电机组发电机健康评估与故障预测研究[D]. 上海交通大学, 2019.
28. 王玉静. 滚动轴承振动信号特征提取与状态评估方法研究[D]. 哈尔滨工业大学, 2015.
29. 常田. 城市供水管网管线健康状态评估方法研究及其应用[D]. 清华大学, 2016.
30. 李惠柯. 基于数据的风电设备健康评估研究[D]. 上海交通大学, 2017.
31. 赵帅. 基于数据驱动的数控机床关键功能部件健康评估[D]. 上海交通大学, 2018.
32. 罗嗣棂. 基于长短时记忆神经网络的风电机组健康状态评估与预测研究[D]. 华北电力大学(北京), 2019.
33. 罗阳青. 基于模糊层次分析法的服役石拱桥健康评估研究[D]. 湖南大学, 2011.

附件1 风洞PHM模型管理软件技术规格书

F1 项目背景

为满足风洞设备的可靠性和安全性需求，提高试验装备运维保障工作效率，基地积极推动PHM技术的应用，先后在4米×3米风洞、2.4米跨声速风洞、Φ1米高超声速风洞等开展了单体风洞或单体关键部段的健康管理系统建设，并开发了对应的健康管理软件系统。与此同时，基地中众多在用与在建风洞都申请或即将开展各自的设备健康管理系统建设，开发工程量和成本日益增加。然而，基地目前采取的是“一座风洞、一个平台、一套软件”的独立分散建设运行模式，导致每个风洞系统拥有一套独立的健康管理系统，系统中的模型复用性差、重复建设问题严重，数据、模型和知识难以在系统间共享而形成孤岛。因此，针对PHM模型管理方面遇到的问题，基地拟开展PHM模型规范化管理研究。

F2 建设目标

针对PHM业务建立及应用的各类模型特点，诸如物理模型、故障诊断/预测模型、状态评估模型等，开展模型管理研究，主要包括模型存储管理、模型更新管理、模型访问管理、模型评价、模型运行管理、模型接口和参数等的规范化要求等，同时开发PHM模型管理的软件系统，促进风洞群一体化平台的建设、管理与维护。

F3 设计原则

系统以技术规范、运行稳定、易使用、易维护、可扩展作为基本原则进行设计，规划系统的整体架构和各个功能模块。

F3.1规范性

（1）在系统设计上使用先进且在市场上具有充分实践的方案与架构，减少在开发和使用阶段出现未知问题的可能性，使维护人员和用户在培训之外，还能自行通过互联网查阅到丰富可靠的参考资料；

（2）符合国家计算机软件质量保证计划规范和相应使用机构的技术规范；符合国内外PHM软件系统的相关标准规范，采用的算法与报表满足相关国家标准。软件开发参考的相关国标如下：

——GJB/Z 102-1997：软件可靠性和安全性设计准则；

——GJB 437-86：军用软件开发规范；

——GJB 439-88：军用软件质量保证规范；

——GJB 1267-91：军用软件维护；

——GJ/T 15532-2008：计算机软件测试规范。

F3.2稳定性

（1）容错能力：对有Bug的程序写入的错误数据有足够的适应能力，对人为操作具有一定的容错性；

（2）并发能力：软件具备不少于20个用户终端的并发访问能力；

（3）稳定运行能力：系统平均无故障时间（MTBF）≥1000小时，平均恢复时间（MTTR）≤30分钟，可7×24小时持续运行；

（4）资源调度能力：系统运行期间，CPU 平均占用率≤60%，内存资源总占用率≤60%，用户界面响应时间≤1000ms。

F3.3易用性

（1）交互友好：提供现代化且友好的图形界面，实现良好的可视性，提供符合用户直觉的交互方式，尽可能减少用户的学习和适应成本，使各个功能的操作和表现保持一致；

（2）反馈机制：对用户的操作以及当前的状态给予适当的反馈；在用户进行步骤较多的操作时，向用户展示当前的进度；在用户完成一项操作时，向用户反馈操作是否成功；但后台需要进行长时间的响应时，通过进度条及文字提示等方式向用户反馈处理进度；

（3）启发性机制：为用户提供充分的操作提示，帮助用户在执行关键动作时进行确认，对于可能造成数据丢失或损坏的高风险操作，给予醒目的风险提醒；

（4）合法性检查：在表单操作中对用户的各项输入进行合法性检查，对于非法输入当即进行警告和提示，确保所有输入合法之后，才为用户开放提交权限；

（5）导航系统：帮助用户快捷地寻找目标功能，在各个页面中帮助用户快速定位当前位置，在关系紧密的功能或流程之间提供快捷的跳转通道；

（6）数据接口：对于已有的各业务子系统数据，可根据各自的接口协议进行知识的归档衔接，如采用的是TFP文件传输协议提供数据，维稳系统建立起与各业务系统的FTP协议通道，通过FTP协议发送请求获取所需的知识数据，同时还会考虑到各系统归档数据版本的同步更新问题。

F3.4易维护性

（1）文档一致性：软件设计文档与软件代码一致，设计更改与甲方协商，签署设计变更确认单；

（2）代码规范性：软件代码命名规范，可读性强，注释应为功能性的，有效注释的行数不得少于源程序总行数的20%；

（3）程序员手册：提供描述总体架构、模块调用关系、各模块输入输出及异常、接口使用说明的程序员手册；

（4）用户手册：提供便于查询、理解，可操作性强的用户使用手册；

（5）上线培训：提供软件使用维护培训服务，培训人数不少于3人。

F3.5可扩展性

设备具有良好的扩展性，根据功能需求的变化、数据的扩增、用户量的生长可以进行快捷的二次开发以及灵活的扩容与升级。

F4 研制内容

根据项目建设目标和设计原则，PHM模型管理系统研制内容主要有：

（1）对基地现阶段所面临的PHM模型规范化管理问题进行剖析，进而对基地所提出的任务和需求进行分析；

（2）通过对PHM模型进行分类的方式，以模型分类结果为主线，对不同类型的模型的输入输出参数、构建过程、模型特点、适用范围等进行规范；

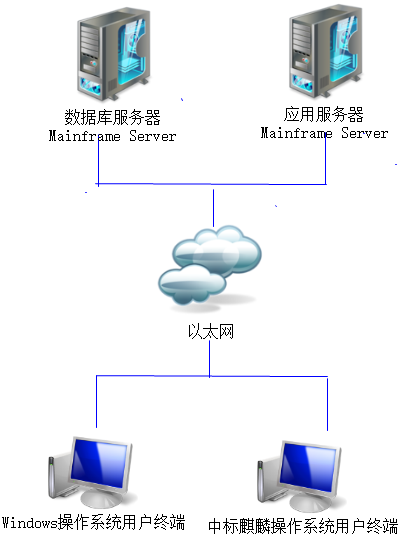
（3）从模型的编码、建模、存储、调用、运行、性能评估、更新和简介的角度，开展PHM模型接口、模型存储管理、模型访问管理、模型运行管理、模型评价、模型更新管理研究；

（4）针对一体化平台模型集成的要求，从模型库设计原则、开发技术要求、软件研制要求等方面开发一套PHM模型管理软件系统。

F5总体研制要求

F5.1硬件部署要求

项目研制内容主要硬件应包括1台应用服务器、1台数据库服务器、2台用户访问终端（1台安装Windows操作系统，1台安装类Unix操作系统，两种类型用户终端均可访问应用服务器的系统软件并完成相应可靠性分析评估任务），4台显示器以及1台网络交换机，部署方式如图F1-1所示。其中，网络架构接口拟采取类型包括远程过程调用、标准查询语言、文件传输、信息交付；其标准包括 ANSI 标准 SQL API；拟使用的网络协议为TCP/IP协议及TCP/IP五层架构（物理层、数据链路层、网络层、传输层、应用层）中涉及的协议；拟采用超五类网络布线系统的百兆以太网实施网络通信；拟采用防火墙、访问控制等安全措施；具体的网络架构中的接口、协议、网线等的选择根据实际的情况可进行调整，以满足用户访问体验的实时性要求。



图F1-1 硬件部署示意图

F5.2系统总体设计方案

开发的PHM模型管理系统主要功能如下：用户综合管理（包括用户创建、用管理和权限管理）、PHM模型定义、PHM模型编码、PHM模型存储、PHM模型调用、PHM模型运行、PHM模型评估和PHM模型的更新，详见图F1-2。



图F1-2 软件系统主要组成

F5.3.1 用户创建

系统管理员可创建新的用户，录入用户信息，为用户分配权限类别以及创建初始密码，访问用户注册信息如表F1-1所示；

表F1-1 访问用户注册信息表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 用户名 | 登录密码 | 个人姓名 | 所属部门 | 权限等级 | 设备岗位 | 用户状态 |
|  |  |  |  |  |  |  |

在用户创建的过程中，对输入的用户名、密码、姓名等进行合法性检测，避免将非法字段输入到数据库中。

F5.3.2 用户管理

超级管理员可以可折叠的树状结构浏览、删除和修改用户和用户信息；超级管理员可查看和修改用户的登录密码；超级管理员可对用户执行启用和停用的操作。此外，超级管理员可创建、修改和删除部门和部门信息。

F5.3.3 权限管理

系统管理员可浏览和管理系统所有的可分配权限，并修改权限的名称和分组。本软件设置3种访问权限，不同权限的用户只能对特定的软件功能模块进行操作，如表F1-2所示；

表F1-2 访问用户权限设置表

|  |  |
| --- | --- |
| 权限类别 | 权限具体描述 |
| 超级管理员  （角色1） | 可以使用系统软件的所有功能进行系统维护管理，能查看所有数据内容，可进行所有软件模块操作。 |
| 管理人员  （角色2） | 可以查看本系统所有数据分析结果信息，但不能进行数据修改。 |
| 负责人（角色3） | 可以录入、修改和查询所负责PHM模型的所有信息，对本人负责的设备进行健康管理评估和故障诊断分析。 |

F5.3.4 PHM模型定义

在各类PHM模型被导入系统时，PHM模型管理软件具备PHM模型定义规范功能，主要内容包括：

（1）PHM模型涵盖内容：PHM模型定义规范模块实现FD设备故障诊断模型、故障预测模型及健康状态评估模型的集中管理，相应的模型需有故障诊断、故障预测或健康状态评估的标签或编码；

（2）PHM模型的输入输出：不同的PHM模型需要基于各自的特性，匹配相符的输入数据，需要基于任务需求，输出合适的结果数据；

（3）PHM模型的适用范围：对于各类PHM模型，用户可以根据模型特点编辑模型的适用范围或功能边界；

（4）PHM模型的推荐：软件针对某种设备存在的故障问题或状态监测数据，提供模型推荐样表帮助用户选择模型。

F5.3.5 PHM模型编码

模型编码的规范用于对一体化平台中的PHM模型的编码结构提出具体的规范和要求，具体功能要求如下：

（1）模型自动编码功能：PHM模型被构建后，系统能自动为该模型编码，模型码与模型一一对应，可用于模型的存储、调用和检索等;

（2）区分通用模型和设备相关模型的功能：PHM模型可分为通用型模型和风洞设备相关的模型，通用型模型可用对象较广，风洞设备相关的模型以装备的最小维修单元或元件为对象。因此，系统对模型的编码，能够将通用模型和设备相关模型进行区分;

（3）遵循编码规范的功能：模型编码应该严格遵循4.2节模型编码规范的总体要求。

F5.3.6 PHM模型存储

对模型的存储进行规范就是要对模型存储过程中的各个步骤进行规范，具体功能如下：

（1）为适应不同内容的数据格式以及元数据需要，模型库中字段的类型应包括字符串、数值、日期、时间、文本、二进制等；

（2）在字符的编码和校验过程中，为了能够支持中文，应该使用utf-8编码，使用utf-8-cgi校验；

（3）对模型库本身以及其内部的表和字段进行规范的命名，模型的命名使用四级命名结构，每级之间使用“-”隔开；第一级名称统一使用PHM，表示模型属于PHM模型一类；第二级命名可以使用CORE和AUXILIARY，用以表示模型属于核心算法模型还是辅助模型；第三级命名可以使用模型本身英文名称或是英文名称缩写；第四级命名可以使用如“001”所示的三位数字，用以表示模型的版本；

（4）在使用模型库存储数据文件时，可能会由于各种原因导致数据丢失或损坏，还可能会由于一些原因需要对数据进行迁移处理，因此应该对模型库本身及其数据进行备份，以便各种事件发生时能够迅速恢复数据。

F5.3.7 PHM模型调用

根据已有模型，对模型通过拖拽生成或调用接口的方式对模型进行调用。具体功能如下：

（1）调用方式选取：根据用户习惯，用户选择通过API调用或通过可视化控件调用；

（2）API调用：系统需提供的API接口，用户通过输入指定参数，可调用系统内置函数执行相关操作；

（3）可视化控件调用：系统提供可视化控件，用户通过拖拽连接控件生成所需模型，系统需对模型进行保存，并基于用户输入的关键参数运行模型。

F5.3.8 PHM模型运行

模型运行的具体功能如下：

（1）列表展示：系统提供多种内置方法，根据模型调用类型为用户提供对应训练测试方法列表；

（2）单选：用户单选列表选项，系统则针对所选选项展示详细运行数据及运行结果；

（3）多选：用户多选列表选项：系统针对多选内容提供运行比较。

F5.3.9 PHM模型评估

PHM模型性能评价指在不同场景中使用特定指标的适用性、每个指标的优缺点，以及在选择用于评估不同模型性能的指标时应考虑的各种因素。因此，系统需提供多元化评价指标函数，用户选择指定函数生成对应结果，并可以通过图形化或列表方式展示对比。

F5.3.10 PHM模型更新

PHM模型更新的目的是及时、准确、科学的对模型进行必要的修改与完善，具体功能如下：

（1）模型备份：系统为用户提供模型备份功能，支持用户后台备份，选择备份及更新备份；

（2）参数更新：系统提供导入功能和修改功能，支持用户通过导入或输入的方式对参数进行更新；

（3）编写更新日志：需提供基于更新的更新日志，方便用户查看更新记录与更新人员。