

军用飞机预测与健康管理系统性能度量体系研究

苗学问¹, 杨 云¹, 雷 迅¹, 张 卫²

(1. 空军装备研究院 装备总体论证研究所, 北京 100076; 2. 空军飞行训练试验基地 理训中心, 河北 沧州 061022)

摘要: 基于健康退化曲线对军用飞机故障预测与健康管理 (PHM) 技术的内涵、基本功能和能力需求进行探讨, 在此基础上, 以科学评价 PHM (prognostics and health management) 系统的诊断和预测能力为目标, 从能力需求出发提出 PHM 系统性能度量方法体系 (包括诊断性能度量、预测性能度量以及综合度量), 并对各个度量方法的定义和应用进行详细阐述。为 PHM 系统算法设计、改进及系统能力验证奠定基础, 具有一定的工程应用价值。

关键词: 故障预测与健康管理; 能力需求; 性能度量

中图分类号: TP277 **文献标识码:** A **文章编号:** 1000-8829(2011)12-0106-05

Performance Metrics of Aerocraft PHM System

MIAO Xue-wen¹, YANG Yun¹, LEI Xun¹, ZHANG Wei²

(1. Air Force Equipment Academy, Beijing 100076, China;

2. The Flight Test and Training Center of Air Force, Cangzhou 061022, China)

Abstract: The basic functionalities, essential function and the performance requirements of prognostics and health management (PHM) technology are firstly discussed by using equipment failure progression curve. Aiming at how to properly evaluate the performance of PHM system, a set of diagnosis & prognosis algorithm performance metrics approaches for the PHM system on the performance requirements is described. The definitions and applications of these metrics are explained in detail, including diagnosis performance metrics, prognosis performance metrics and integrated metrics. It is concluded that performance metrics works will significantly support the process of algorithm design, improvement and verification and validation (V&V) of the PHM system.

Key words: PHM; performance requirements; performance metrics

随着军用飞机的复杂性、信息化和综合化程度大幅度提高, 装备保障由从前的机械修复为主, 逐步转变为通过信息的获取、处理和传输做出保障决策为主, 这推动着机内测试技术不断拓展, 并最终发展为与飞机各系统高度耦合的独立“功能系统”——故障诊断、预测与健康管理系统 (PHM) 系统。PHM 系统利用分布于全机的先进传感器 (含虚拟传感器) 的信息, 借助各种数学模型自底向上综合, 实现故障诊断、隔离与预测, 并给出保障决策支持。PHM 系统是实现飞机基于状态的维修和自主式保障, 进而提高装备保障性水平和降低寿命周期费用的关键因素之一。分析表明, 采用

PHM 系统的 F-35 联合攻击机与 F-16 飞机相比, 出动架次率增加 25%, 维护人员减少 20~40%, 保障设备减少 50%, 维修保障费用降低约 30%^[1-2]。由于 PHM 系统的诸多优势, 美英等军事强国越来越重视其研发与应用, 采用 PHM 系统将成为先进军用飞机的重要标志之一^[3-5]。

发展 PHM 系统面临的难题之一是如何对其功能和性能进行评价。国外进行了一系列的探索性研究并取得了一定的成果。如美国空军大学开发了自主保障系统的仿真工具 (ALSIm) 并应用于 JSF 的 PHM 的诊断算法演示验证。IMPAC 公司开发的性能评价工具 (MET), 用于对基于振动特征的健康监控方法进行评价。美军方为有效开展 F-35 飞机 PHM 系统的评价, 采用 OSA-CBM 标准搭建的基于网络的评估系统, 来评估和确认 PHM 系统在各个层次上的性能; G. Vachtsevanos^[6] 对 PHM 系统的性能度量进行了探讨。但从公开发表的文献看, 当前尚未建立 PHM 系统性能度量

收稿日期: 2011-04-18

基金项目: 中国博士后科学基金资助项目 (20100481500)

作者简介: 苗学问 (1977—), 男, 河南周口人, 博士, 主要研究方向为装备“四性”论证; 杨云 (1969—), 男, 甘肃人, 高级工程师, 主要研究方向为装备“四性”论证。

体系,更缺乏被广泛认可的评价程序与方法。在国内,相关研究基本是空白。

笔者从 PHM 系统的技术内涵和能力需求出发,提出 PHM 系统的度量参数体系及参数计算方法,作为 PHM 系统验证、PHM 建模方法分析与改进的研究基础。

1 PHM 技术内涵及能力需求分析

1.1 技术内涵

由于损伤发展的随机性、载荷的不确定性以及失效模式的多样性等原因,部件/系统使用寿命的分散度很大,如图 1 所示。但是部件/系统的健康水平(部件或系统及其子系统的整体状态)将经历状态良好、初步损伤(早期的潜在故障)、故障发展和最终失效的历程,这个过程被称为健康退化过程^[7-8]。对于不修复系统,其健康退化过程示意图如图 1 所示。

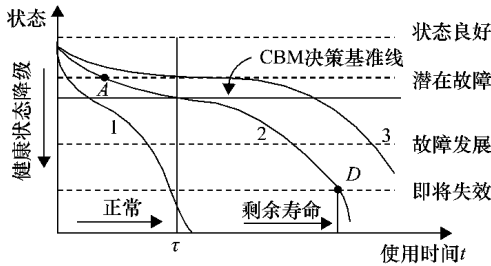


图 1 健康退化曲线

如图 1 所示,由于寿命的分散性,采用定时维修,即基于使用寿命的统计学估计确定维修间隔 τ ,既不能有效避免因极端恶劣的或者不可预知的工况所导致的部件突然失效(如图 1 中的部件“1”)而造成的飞行事故,也不能有效避免因替换下有很长剩余使用寿命的部件(如图 1 中的部件“3”)而造成的浪费,以及增加例行检查和更换而造成的人为故障。因此,基于状态的维修(CBM)可优化装备保障,并确保飞行安全。实现 CBM 的技术基础就是进行军用飞机的故障预测与健康管理。

PHM 系统的技术内涵即是利用先进传感器的集成,并借助各种智能模型来预测、诊断、监控和管理飞机的状态,并给出最优的保障策略。如图 2 所示,PHM 系统强调对剩余使用寿命(或正常运行时间)的预测,目标是在早期的潜在故障阶段,通过对故障征候的分析,预判实际故障的发生时间(剩余使用寿命)。这样可以根据诊断/预测信息、可用资源和使用需求对维修活动做出适当决策,获得系统效益最大化,即实现健康管理。此外,具备了增强的诊断和可靠的预测能力,可以对飞机系统的相关设计特性进行优化权衡,以实现费用约束下各种目标的综合平衡。也就是说 PHM 系统与飞机系统设计、维修保障策略和保障管理等紧密

相连,因而也成为了先进战斗机提高可靠性、维修性、测试性、保障性和安全性,以及降低寿命周期费用的关键系统。

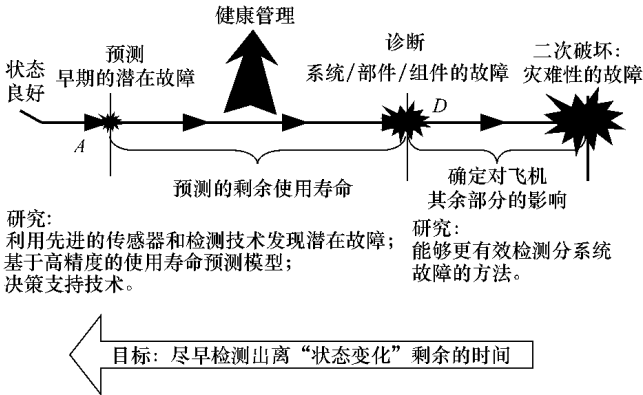


图 2 故障预测与诊断过程

1.2 能力需求

如上分析,PHM 的引入是为了在准确时间(right time)对准确部件(right components)采取准确维修活动(right maintenance actions),并为保障计划的制定、实施和调整提供可靠依据,相应的 PHM 系统应具备早期故障检测与隔离能力、故障演变进程的监测能力,以及故障的报告能力。具体能力需求如下:① 故障检测能力;② 故障隔离能力;③ 针对选定部件进行故障预测的能力;④ 系统/部件状态的监视能力;⑤ 有寿件剩余使用寿命的预计能力;⑥ 故障选择性报告能力;⑦ 对飞机各系统/设备的配置管理能力;⑧ 可靠性高、重量轻、数据处理与存储迅速等。

针对 PHM 系统的能力需求,可得出军用飞机对 PHM 系统的性能需求,进而可以得出性能度量体系,如表 1 所示。

表 1 性能度量指标体系表

顶层目标	能力分解	性能度量指标
实现飞机的健康管理	故障诊断与隔离	诊断性能度量
	故障预测或使用寿命预计	预测性能度量
	其他要求	其他(综合)度量

2 诊断和预测算法性能度量

2.1 诊断性能度量

2.1.1 能力需求分解

PHM 系统应具备较强的故障检测、隔离能力和较高的鲁棒性。诊断能力可分为两个层次,一是对故障征兆的判断,即要求对微弱故障信息有足够高的分辨灵敏度;二是对故障能有效检测和精确定位。可分解为:

① 故障的检测与隔离能力强。一是对故障能有效检测和精确定位;二是误报和漏报应尽可能少;三是

对故障征兆的判断准确;

- ② 检测隔离的速度应尽量快,即延迟时间短;
- ③ 对背景噪声和工况变化应足够稳定;
- ④ 置信度高。

2.1.2 度量的定义和应用

根据 PHM 能力需求分解的结果,提出性能度量参数:

① 成功检测率(POD):检测到的故障占有故障的比例,它与漏报率的和始终为 1。

$$\text{POD} = P(D_1/F_1) = \frac{a}{a+c} \quad (1)$$

② 误报率(POFA):故障报警在所有的非故障事件中所占的比例

$$\text{POFA} = P(D_1/F_0) = \frac{b}{b+d} \quad (2)$$

③ 准确度(accuracy):应用所有的可用数据进行分析,进而衡量算法正确区分故障事件与非故障事件的有效性

$$\text{准确度} = P(D_1/F_1 \& D_0/F_0) = \frac{a+d}{a+b+c+d} \quad (3)$$

上式中参数依据表 2 的决策矩阵得到。其中, $C(s)$ 是对某一指定严重度的检测置信水平。

表 2 决策矩阵

输出	故障(F_1)	没有故障(F_0)	总计
肯定(D_1)	a	b	$a+b$
检测到故障	正确报警数	误报个数	报警总数
否定(D_0)	c	d	$c+d$
检测无故障	漏报个数	正确否定个数	未报警总数
	$a+c$	$b+d$	$a+b+c+d$
	故障总数	无故障事件总数	检测事件总数

④ 稳定性(stability):用以衡量某故障严重度在峰一峰值变化过程中相应置信水平的变化范围。公式表示为

$$\text{稳定性} = 1 - \int_0^1 [C_H(s) - C_L(s)] ds \quad (4)$$

式中, $C_H(s)$ 和 $C_L(s)$ 分别为故障变化过程中,检测置信水平的极大与极小值。

⑤ 工况敏感度(duty sensitivity):用以衡量在不同的工况状态下,PHM 输出差异。公式表示如下

$$\text{工况敏感度} = 1 - \int_0^1 [C_1(s) - C_2(s)]^2 ds \quad (5)$$

式中, $C_1(s)$ 、 $C_2(s)$ 分别为两种不同的负载或者不同工况状态下,PHM 算法输出的检测置信水平。

⑥ 噪声敏感度(noise sensitivity):用以衡量一个 PHM 输出的鲁棒性或抗噪能力。公式表示如下

$$\text{噪声敏感度} = \{1 - \int_0^1 [C_1(s) - C_2(s)]^2 ds\} (\Delta \text{False Positive}) \quad (6)$$

式中, $C_1(s)$ 为在噪声 1 下算法的检测置信水平;

$C_2(s)$ 为在噪声 2 下算法的检测置信水平; s 为故障严重程度; $\Delta \text{False Positive}$ 为两种噪声下误报率差值的绝对值。

⑦ 总体置信度(overall confidence);用以表示对 PHM 系统诊断结果的检测置信水平,通过下式得到

$$\text{总体置信水平} = \int_0^1 C(s) ds \quad (7)$$

式中, $C(s)$ 为对某一指定故障严重度的检测置信水平。故障严重程度表示“故障严重的程度”,属于区间 $[0, 1]$, “0”代表“健康状态”, “1”代表“功能故障”。

此外,诊断算法的性能可以用误报和漏报之类的度量来进行表征。但如果检测的阈值发生改变,或者需要对其进行权衡折衷,此类度量的标准可能会发生改变。可使用接受者操作特性曲线(ROC curve)度量,详见文献 [6]。

所提出的上述度量指标,是针对 PHM 系统由传统的机内测试(BIT)即电子产品领域扩展至机械产品,由部件级拓展为系统级这一特点所提出的。在工程实践中,针对 BIT 设计,其能力度量仍然可采用故障检测率、故障隔离率等指标衡量。

2.2 预测性能度量

2.2.1 能力需求分解

PHM 系统应具备针对选定部件进行故障预测的能力,针对此能力需求,可分解为:

- ① 应覆盖可行的所有对象;
- ② 预测应准确。一是预测的准确度高;二是预测曲线簇“聚集”;
- ③ 预测及时性好;
- ④ 置信度高;
- ⑤ 对背景噪声和工况变化应足够稳定;
- ⑥ 对损失演变敏感。

2.2.2 度量的定义和应用

根据 PHM 能力需求分解的结果,提出性能度量参数:

(1) 预测覆盖率(PCR)。

PCR 描述 PHM 系统预测能力的覆盖程度,定义为被测部件在规定的时间内,用规定的方法完成预测,其中所有可预测故障模式(即不考虑不可预测的故障模式)中预测结果能够达到预测要求的故障模式所占百分比。

$$\text{PCR} = \frac{\text{大于等于预测要求的故障模式数}}{\text{所有可预测的故障模式数}} \times 100\% \quad (8)$$

上述计算方法不考虑不可预测的故障模式,原因是若把其纳入到预测覆盖率指标中,不同部件或系统没有可比性,操作也很困难。

(2) 准确度和精度。

一个预测算法应该输出两个基本结果:预测失效

时间和置信区间。基于统计方法,对预测算法进行大量的测试,最后利用对大量实验结果的统计来测量预测算法的性能。由此得到两个性能指标:准确度和精度。准确度反应预测值和真实值的接近程度,精度反应预测值的聚集程度。

① 准确度(accuracy):是失效时间的一个估计点与真实失效时间接近程度的度量。假定在第 i 次试验中,失效时间的真实值和预测值分别是 $t_{af}(i)$ 和 $t_{pf}(i)$,预测算法在具体预测时刻 t_p 的准确度被定义为

$$\text{Accuracy}(t_p) = \frac{1}{N} \sum_{i=1}^N e^{-D_i/D_0} \quad (9)$$

式中, $D_i = |t_{pf}(i) - t_{af}(i)|$ 为真实失效时间与预测失效时间之间的距离; D_0 为正规化因子, N 为试验的次数。这里应用指数函数给出一个平滑单调递减曲线: e^{-D_i/D_0} 随着 D_i 的增大而减小,并且当 $D_i=0$ 时它为1,当 D_i 趋于无穷时它趋于0。即当预测值与真实值相同时,准确度是最高的,当预测值偏离真实值时准确度降低。

② 精度(precision):是衡量剩余寿命所覆盖区间的狭窄程度指标,是由基于大量试验所产生的预测值的方差定义的,如果预测值聚集在真实值周围,那么相关预测算法的精度就很高,如果预测值分散在真实值周围,那么精度就会很低。图3显示了预测值的分布与精度的关系。

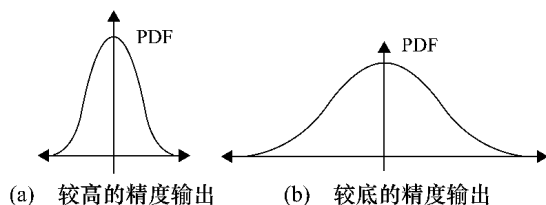


图3 预测值概率密度函数

若定义 t_p 时刻的预测误差 $E_i = t_{pf}(i) - t_{af}(i)$,得出预测误差的平均值和方差为

$$\bar{E} = \frac{1}{N} \sum_{i=1}^N E_i \quad (10)$$

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (E_i - \bar{E})^2 \quad (11)$$

那么, t_p 时刻的预测精度的具体计算公式为

$$\text{Precision}(t_p) = \left(\frac{1}{N} \sum_{i=1}^N e^{-R_i/R_0} \right) \times e^{-\sigma^2/\sigma_0} \quad (12)$$

式中, σ_0^2 和 R_0 分别为取决于实际情况的正规化因子; R_i 为每次试验的预测置信区间。同样,这里也应用一个指数函数来定义预测误差的方差、预测置信区间和精度之间的关系。精度在0~1范围内取值,用1表示精度最高,0表示精度最低。较窄的置信区间和较紧簇的预测值将得到一个更高的精度值。

(3) 预测及时性。

准确度和精度没有考虑预测的及时性,不能解释失效预测时间(TTF)的及时性影响,即不能考虑真正的失效点与TTF概率密度函数(PDF, probability density function)的覆盖区域在时间轴上的相关位置。如图3所示,失效发生在预测时域之内是最理想的(曲线2);应至少避免失效在预测时域之前发生(曲线3)。

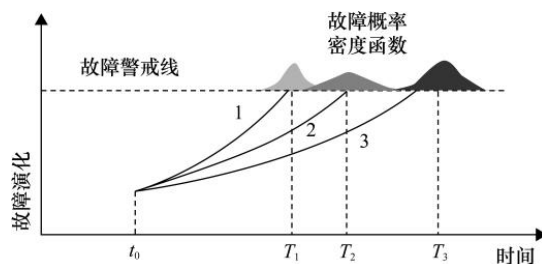


图4 预测及时性示意图

(4) 预测置信度。

假定某故障预测算法在相应不确定性范围内,预测到的故障演化过程如图5所示。在 t_0 时刻,一种故障被检测到并被隔离出来,根据预测规则得到平均失效时间 T_{fm} ,最早失效时间 T_{fe} ,最晚失效时间 T_{fl} ;警戒线指出部件在其中任一点失效时所对应的故障级别。

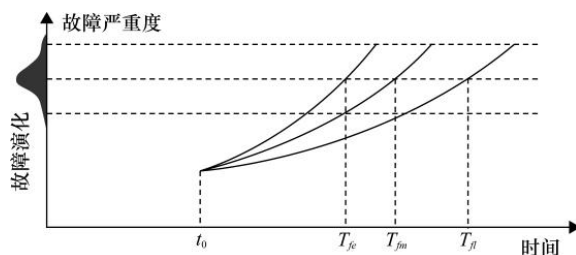


图5 不确定性范围和置信界限

进一步假定,失效点的数据可以是一种分布或几种分布的叠加,如图6所示。

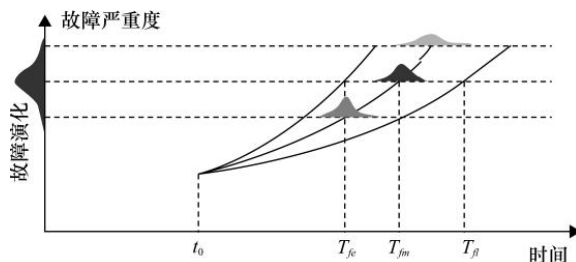


图6 置信界限的概率密度函数

假定在 T_{fm} 时刻(与警戒线交汇)的分布如图7所示。

假设预定的任务要求在时刻 T 前考虑产品的可用性。概率密度曲线从时刻 T 到无穷远的积分可得到其置信水平。该置信水平就是产品在任务完成前不会失效的程度。

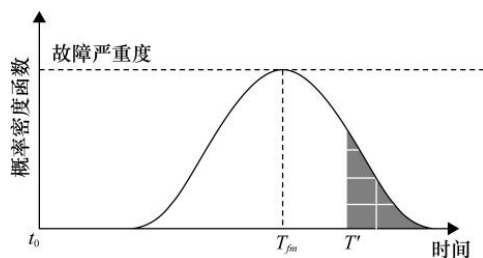


图7 警戒线处的分布

另一方面,给定一个置信水平(如95%),想要找到时刻 T'' ,在这个时刻下产品还能工作,没有完全失效。对概率密度曲线从时刻 T'' 到无穷时刻进行积分,并让积分值等于给定的置信水平(本例中为95%),计算出 T'' 。这样就能得到从开始时刻 t_0 到 T'' 的时间长度,亦即在期望置信水平下产品可以工作的时间。如图8所示。

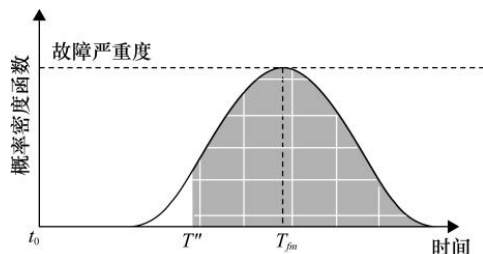


图8 某置信水平下的失效时间计算

(5) 相似度(similarity)。

相似度用以比较从开始时刻到将来某特定时刻的多个预测时间序列和常规序列的拟合程度。可表示为

$$\text{Similarity}(x, y) = \sum_{i=1}^n \left(1 - \frac{|x_i - y_i|}{\max_i - \min_i} \right) \quad (13)$$

式中, x_i 和 y_i 为两个不同的时间序列中的第 i 个元素; \max_i 和 \min_i 为所有第 i 个元素中的最大值和最小值,如图9所示。

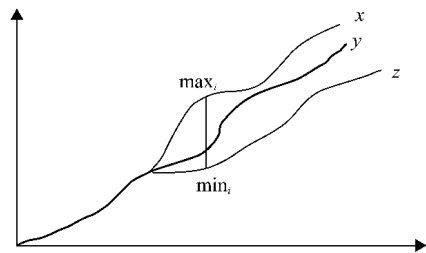


图9 相似度计算

(6) 灵敏度(SN)。

灵敏度表征故障预测算法对系统输入变化或外部干扰的敏感程度。它的定义如下

$$\text{SN} = \frac{\sum_{i=1}^N \frac{\Delta_i^{\text{out}}}{\Delta_i^{\text{in}}}}{N} \quad (14)$$

式中, Δ^{out} 为两个连续输出之间的差值; Δ^{in} 为两个连续

输入之间的差值。

2.3 其他/综合度量

作为一个与全机各系统高度耦合的功能系统,为了更好地发挥其能力,真正地提高装备的保障效能,要求其可靠性高、重量轻、数据处理与存储迅速等。因此,应提出相应的综合度量参数,包括数据下载时间、数据存储容量、数据处理精度、平均故障间隔飞行小时、重量,以及引入PHM系统的费用-效益权衡指标——总体价值。

$$\text{总体价值} = \sum_{\text{Failure Modes}} \text{技术价值}_i - A - O - (1 - P_c) \delta \quad (15)$$

式中, A 为获取和实施费用; O 为整个寿命周期运行与维修费用; P_c 为计算机资源需求概率; δ 为标准计算机系统费用。

上述度量的分值越高,代表的评估结果越理想。根据各个度量的定义,可以指导相应算法的开发,例如,若算法中计算机资源评分 P_c 较高,则应更注重计算机资源在此算法中的利用。

3 结束语

笔者从PHM系统的技术内涵和能力需求出发,提出PHM系统性能度量参数体系,为开展PHM系统的要求论证、诊断预测算法设计与优化,以及PHM系统的能力评价奠定技术基础,具有一定的工程实用价值。

参考文献:

- [1] Hess A, Fila L. The joint strike fighter (JSF) PHM concept: potential impact on aging aircraft problems [A]. Aerospace Conference Proceedings [C]. 2002:3021-3026.
- [2] Malley M E. A methodology for simulating the joint strike fighter's (JSF) prognostics and health management system [D]. Air Force Institute of Technology, 2001.
- [3] Bolkcom C. F-35 joint strike fighter (JSF) program: background, Foreign Affairs, Status, and Issues [R]. Defense, and Trade Division, 2007.
- [4] Smith G, Schroeder J B, Navarro S, etc. Development of a prognostics and health management capability for the joint strike fighter [A]. IEEE Autotestcon Proceedings [C]. 1997.
- [5] 苗学问, 杨云. 预测与健康管理系统的应用与发展 [J]. 装备论证, 2010, (1): 25-29.
- [6] Vachtsevanos G, Lewis F L, Roemer M, et al. Intelligent fault diagnosis and prognosis for engineering systems [M]. Hoboken: John Wiley & Sons, 2006.
- [7] 梁旭, 李行善, 张磊, 等. 支持视情维修的故障预测技术研究 [J]. 测控技术, 2007, 26(6): 5-8.
- [8] Hong J, Miao X W, Han L, et al. Prognostics model for predicting aero-engine bearing grade-life [A]. Proceedings of ASME Turbo Expo 2009: Power for Land, Sea and Air [C]. 2009-06.

□