本章目录

故障诊断与健康预测

第七章 健康管理与安全评估

7.1 健康管理与安全评估基础
7.2 统一故障建模框架
7.3 虚实安全强化与故障数据方法
7.4 故障危害与安全评估
7.5 故障诊断与健康预测
7.6 安全决策与整机改进
7.7 故障注入接口与算法开发框架
7.8 例程实验与虚实验证
7.9 习题与课堂讨论

本节 4 个教学单元

7.5.1 残差与基线
教材对应：第7.1.2.1、7.1.5.4节

7.5.2 融合诊断
教材对应：第7.1.5.4节

7.5.3 评价与跨工况
教材对应：第7.1.5.4节

7.5.4 健康预测
教材对应：第7.1.2.1节（电池退化预测）

7.5.1 残差与基线：健康管理中的判断

检测、诊断和预测回答不同的问题

故障预测与健康管理（PHM）：感知当前状态，预测退化，并支持任务与维护决策。

判断	输入和依据	典型输出
异常检测	观测与正常范围或模型比较	当前窗口是否异常
故障诊断	多通道特征、残差与已知故障模式	部件、故障类别及类别分数
健康预测	历史健康指标、退化规律 及未来使用条件	未来健康趋势、剩余可用寿命 （RUL）

健康指标可连续变化；功能是否够用，还取决于当前任务的要求。

7.5.1 残差与基线：工况与异常

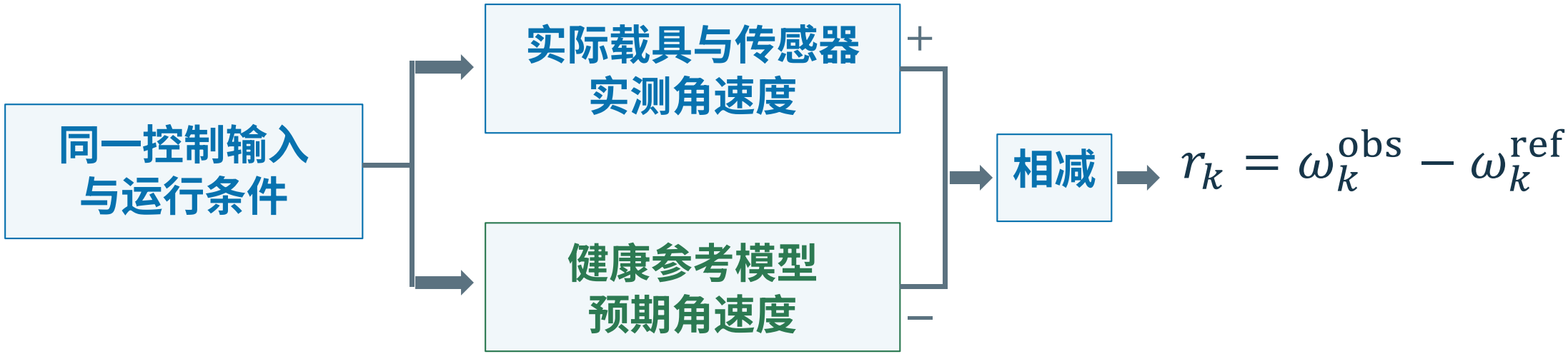
正常机动也会让传感器出现大波动

情形	观测为什么变化	有用的对照
正常加速或转弯	控制输入改变， 健康载具的响应随之改变	同指令下的健康模型响应
传感器异常	单通道偏置或噪声改变， 可能破坏跨通道一致性	冗余观测与跨通道关系
电机效能下降	执行能力下降， 控制器可能增大输出补偿	执行器量、运动响应 与健康残差

健康参考随工况变化，异常证据才有可比较的基准。

7.5.1 残差与基线：残差的方向与单位

残差比较实测与同工况健康响应



ω_k^{obs} 为实测；ω_k^{ref} 为健康参考；本例均为机体系滚转角速度。

两路对齐：时间、坐标系与单位一致；**残差还含：**模型偏差、噪声与时间错位。

$$r_k = 0.70 - 0.50 = 0.20\ \mathrm{rad/s}$$

7.5.1 残差与基线：正常残差基线

正常数据先校准偏差和波动尺度

教学算例：正常残差均值 0.02 rad/s，标准差 0.05 rad/s；均由训练侧正常记录估计。

$$z_k = \frac{r_k - \mu_0}{\sigma_0}$$

μ_0 是正常均值；σ_0 是正常波动尺度；z_k 无量纲。
新样本沿用训练侧这组统计量；测试段不参与拟合归一化参数。

$$z_k = \frac{0.20 - 0.02}{0.05} = 3.6$$

给定判据 $|z_k| > 3$：该样本异常；故障类别还需要其他证据。
实际选阈值时比较独立验证集中的误报、漏报和告警时间。

7.5.1 残差与基线：持续确认的逐点更新

持续确认减少瞬时误报，也增加等待

教学序列：$|z_k| > 3$ 连续两点才告警；采样间隔 0.2 s，故障从 0.75 s 开始。

单点越线不确认；连续两点才告警

t = 1.20 s

同输入 · 同时刻

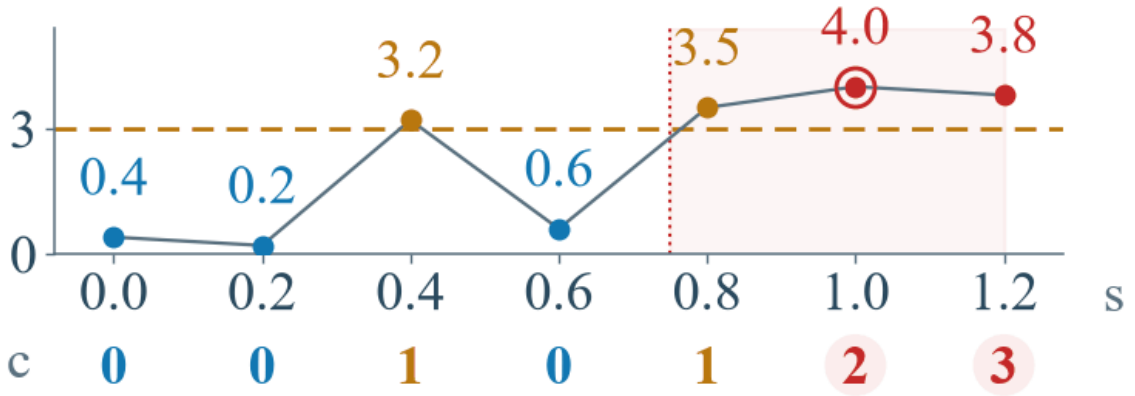


$r = 0.710 - 0.500$

$= \mathbf{0.210\ rad/s}$

$t_k = 1.2$ s

标准化残差 |z|



连续计数 c

3

$|z| > 3$

连续两点

告警

延迟 0.25 s

故障0.75 s → 告警1.00 s；等待0.25 s

$$\Delta t_{\mathrm{det}} = 1.00 - 0.75 = 0.25\ \mathrm{s}$$

7.5.1 残差与基线：增强参考模型的关系

校正正常模型误差，才能保留故障偏离

配对真实与虚拟响应
计算正常误差 → 学习误差行为
得到误差重构 → 加回物理模型输出
得到校正参考

$$y_{\text{ref}} = y_{\text{phys}} + \hat{e}_{\text{normal}}$$

y_{phys} 为物理模型输出；$\hat{e}_{\text{normal}}$ 为正常工况的误差重构；二者与参考输出单位一致。
即课件7.2「增强数字孪生案例」页中的$X_{\text{comp}} = X_{\text{virt}} + \Delta$。

诊断应用：先验收正常参考，运行中保留可疑偏离，再做故障确认。
未知故障期间若继续学习残差，故障偏离会被当作模型误差吸收。

图：依据增强数字孪生论文 Fig.1(b) 整理（出处见课件7.2「增强数字孪生案例」页）

本章目录

故障诊断与健康预测

第七章 健康管理与安全评估

7.1 健康管理与安全评估基础
7.2 统一故障建模框架
7.3 虚实安全强化与故障数据方法
7.4 故障危害与安全评估
7.5 故障诊断与健康预测
7.6 安全决策与整机改进
7.7 故障注入接口与算法开发框架
7.8 例程实验与虚实验证
7.9 习题与课堂讨论

本节 4 个教学单元

7.5.1 残差与基线
教材对应：第7.1.2.1、7.1.5.4节

7.5.2 融合诊断
教材对应：第7.1.5.4节

7.5.3 评价与跨工况
教材对应：第7.1.5.4节

7.5.4 健康预测
教材对应：第7.1.2.1节（电池退化预测）

7.5.2 融合诊断：故障诊断方法

三类方法用不同依据解释异常

方法	从哪里获得判断依据	主要依赖与取舍
模型驱动	物理模型、观测器、解析冗余 比较预期与实测响应	解释明确； 模型偏差会进入残差
数据驱动	主成分分析、支持向量机、 神经网络学习正常与故障模式	需要覆盖工况和类别； 未知故障仍难识别
融合方法	残差学习、贝叶斯融合、 模型误差补偿与多源观测	同时核对模型、数据 和接口的一致性

诊断方法的适用范围，由模型和数据的覆盖共同决定。

7.5.2 融合诊断：模型与数据融合框架

训练阶段有标签，应用阶段只有观测

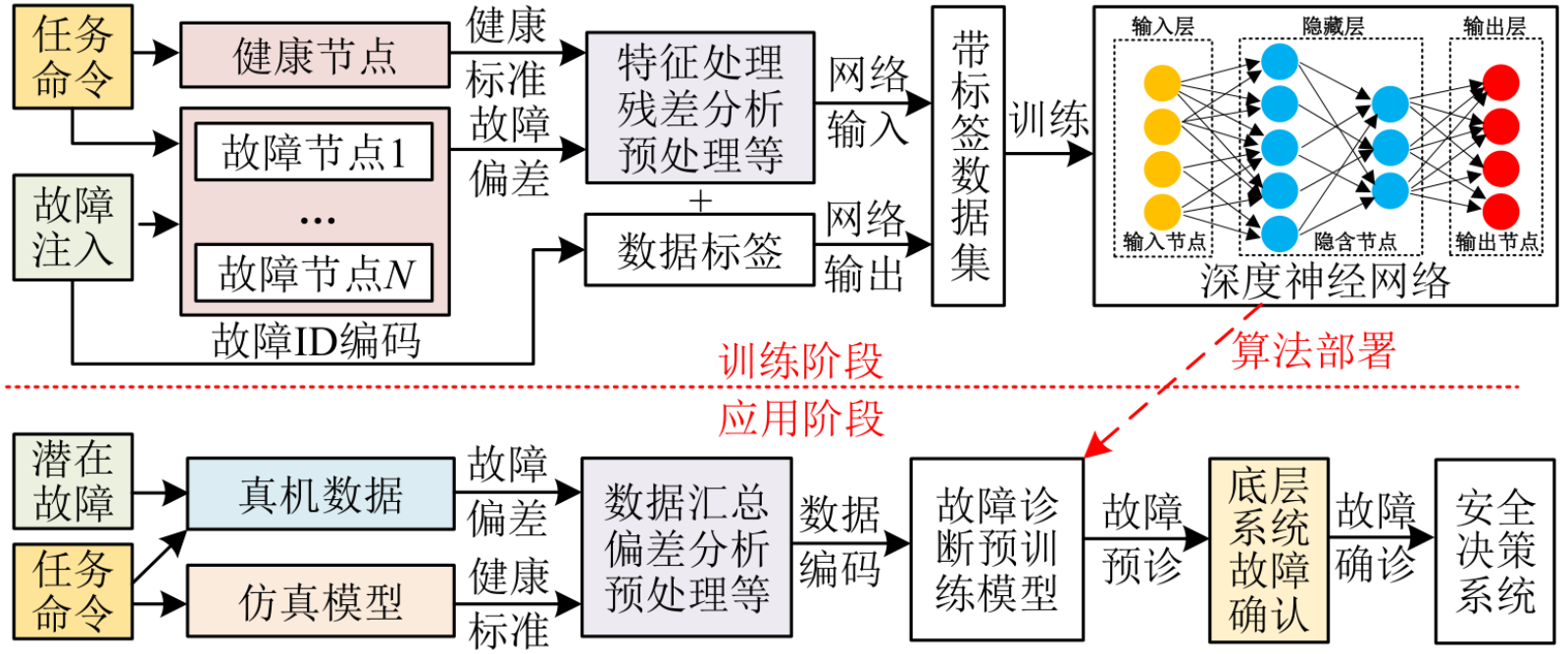


图7.10 模型与数据融合故障诊断训练框架

训练建立输入与标签的映射；应用以健康参考辅助识别，再交故障确认。

7.5.2 融合诊断：训练样本构成

残差是网络输入，注入信息是训练标签

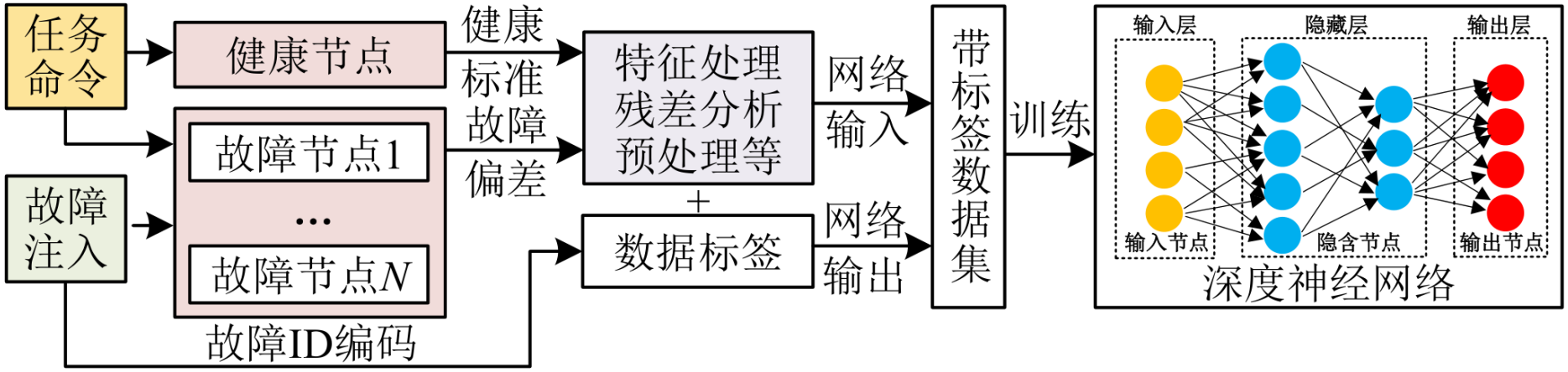


图7.10（训练局部）输入特征与标签走不同支路

节点：软件在环（SITL）、硬件在环（HITL）或真机，见课件7.3.1「虚实节点」页。

训练时：健康节点与故障节点并行执行同一任务，故障节点按ID注入。

特征与标签：残差对齐、归一化、切窗作输入；注入类别与时刻即“天然标签”。

7.5.2 融合诊断：运行时的数据与输出

应用中的类别分数还需要故障确认

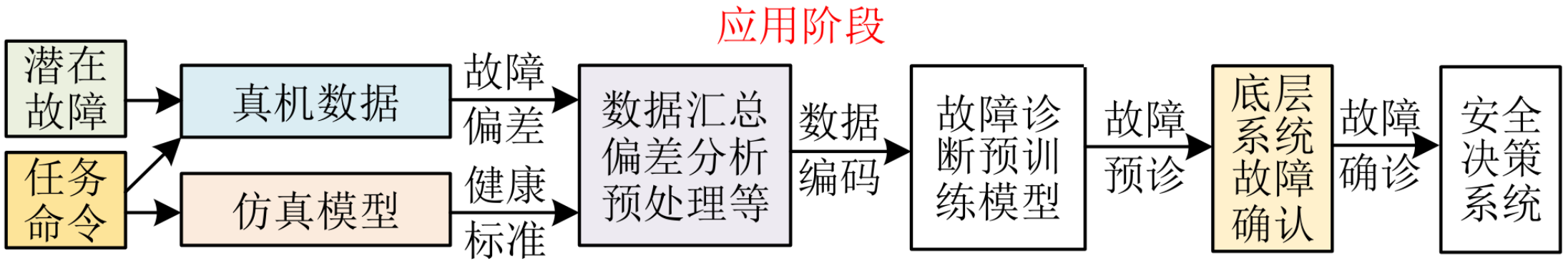


图7.10（应用局部）实时特征进入已训练模型

接口一致： 通道顺序、符号方向、单位、窗口与归一化统计沿用训练配置。

分数解释： 类别分数是模型的相对判断；概率解释依赖独立校准。

系统确认： 连续证据、冗余观测与任务状态汇总后，再交安全决策。

7.5.2 融合诊断：论文来源与图符对照

先认出处与入口，再对照论文图中的符号

提高内容：至「从关系图到类别输出」页为论文图网络；主线可跳到7.5.3。

论文：Tu J, Nian X, Dai X. Digital Twin-Guided Spatiotemporal Graph Representation Learning for Reliable UAV Fault Diagnosis under Complex Flight Conditions. IEEE Internet of Things Journal, 2026.

DOI：https://doi.org/10.1109/JIOT.2026.3709310

目录：[安装目录]/RflySimAPIs/PaperRes/6.IoT-DT-Diagnosis/paper/

Real Node：实机

Virtual Node：健康数字孪生

V^{RE}：实测节点

V^{VI}：孪生的虚拟输出

⊖：二者相减，得残差节点 V^{ER}

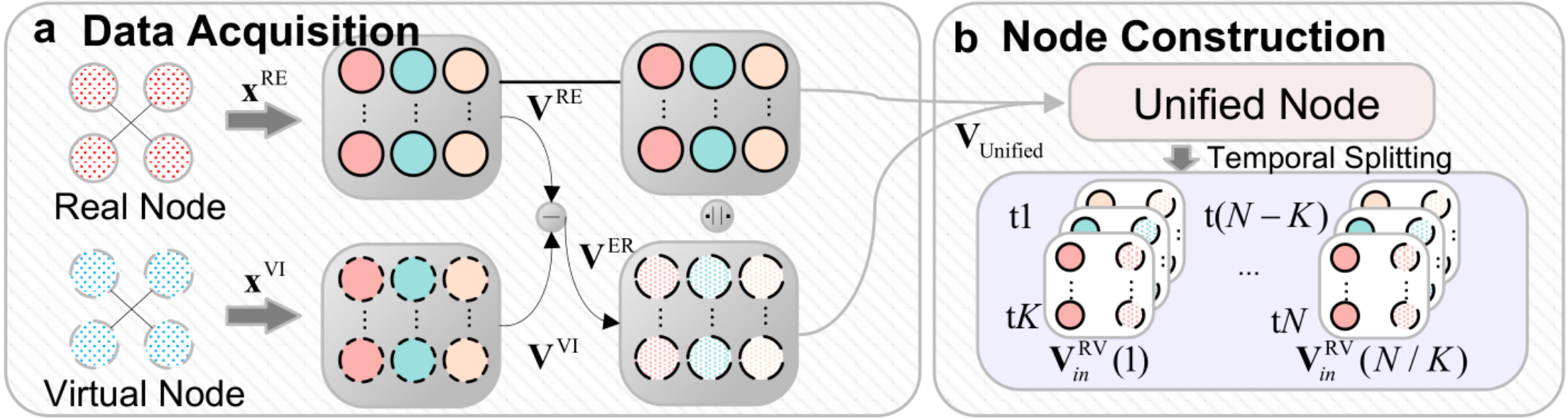
V_{Unified}：实测与残差合并

Temporal Splitting：按K个时刻切分

节点：图中的特征单元，不是仿真节点

7.5.2 融合诊断：论文中的节点构造

实测节点与残差节点保留不同信息

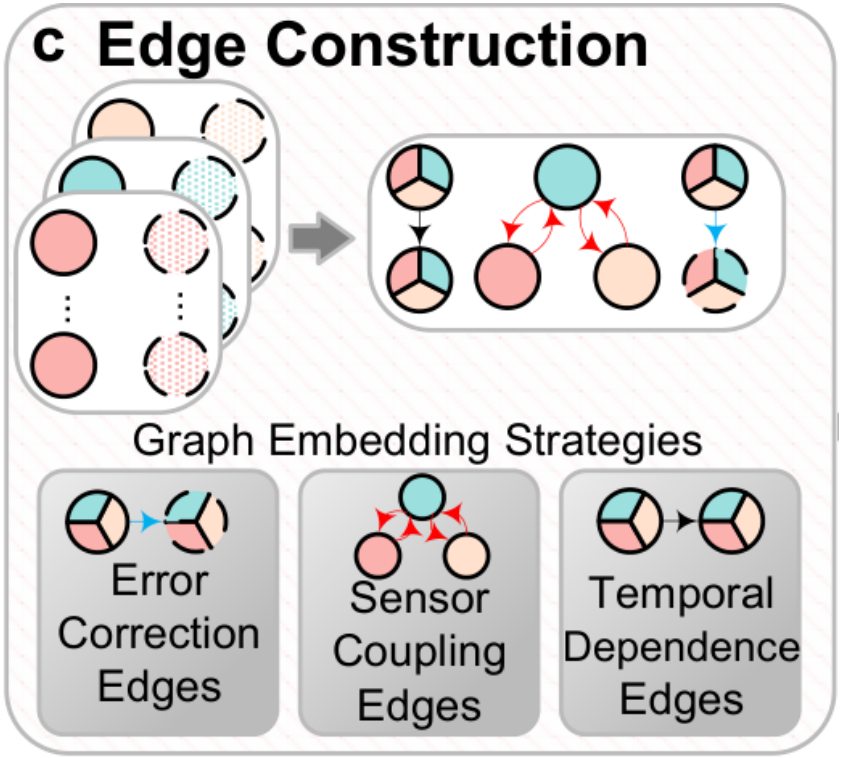


图：同上论文 Fig.3(a)(b) 节点构造

实测节点保留当前响应；残差节点表达相对健康参考的偏离。

7.5.2 融合诊断：物理关系与图连接

三类关系边传递不同的异常证据



图：同上论文 Fig.3(c) 三类关系边

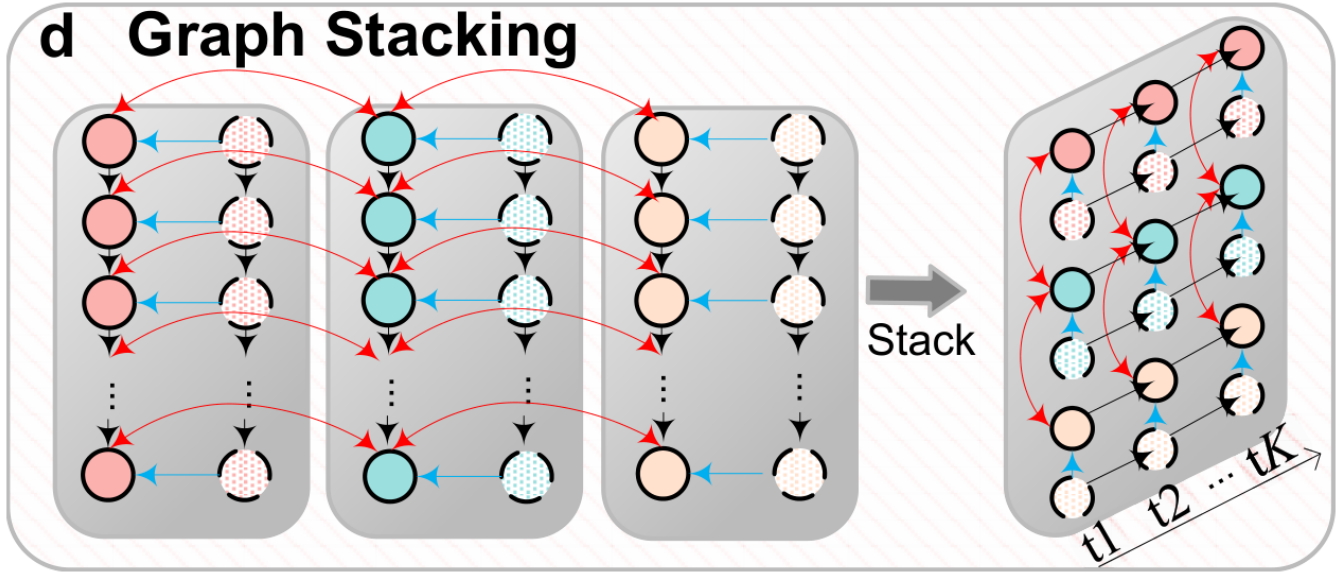
误差校正边：实测与对应残差联系，提供健康参考一致性信息。

传感器耦合边：同一时刻不同通道联系，检查共同运动中的一致性。

时间依赖边：前后时刻同类节点联系，保留故障发展的连续证据。

7.5.2 融合诊断：从关系图到类别输出

图序列把窗口内的时间变化保留下来



图：同上论文 Fig.3(d) 时间图序列

孪生一致时空图诊断网络（TC-SGDN）

图注意力：按相关程度加权汇聚相邻节点特征。

门控循环单元：逐时间步保存并更新序列信息。

在线窗口截至当前时刻，输出当前类别分数。

观测＋健康残差 ➡ 关系图＋时间序列 ➡ 已知故障类别分数

本章目录

故障诊断与健康预测

第七章 健康管理与安全评估

7.1 健康管理与安全评估基础
7.2 统一故障建模框架
7.3 虚实安全强化与故障数据方法
7.4 故障危害与安全评估
7.5 故障诊断与健康预测
7.6 安全决策与整机改进
7.7 故障注入接口与算法开发框架
7.8 例程实验与虚实验证
7.9 习题与课堂讨论

本节 4 个教学单元

7.5.1 残差与基线
教材对应：第7.1.2.1、7.1.5.4节

7.5.2 融合诊断
教材对应：第7.1.5.4节

7.5.3 评价与跨工况
教材对应：第7.1.5.4节

7.5.4 健康预测
教材对应：第7.1.2.1节（电池退化预测）

7.5.3 评价与跨工况：混淆矩阵

总体正确率会掩盖少数故障样本

教学构造：200个窗口，正常180、故障20；故障为正类，分数阈值0.5。

真实状态	预测正常	预测故障
正常 180	172：真负（TN）	8：假正（FP） 误报
故障 20	5：假负（FN） 漏报	15：真正（TP） 检出

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{187}{200} = 93.5\%$$

7.5.3 评价与跨工况：诊断评价指标

漏报看故障分母，误报看正常分母

沿用教学矩阵：TP=15，FN=5，FP=8，TN=172。

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{15}{20} = 75\%$$

$$\text{Miss} = \frac{FN}{TP + FN} = \frac{5}{20} = 25\%$$

$$\text{FPR} = \frac{FP}{FP + TN} = \frac{8}{180} \approx 4.4\%$$

$$\text{Precision} = \frac{TP}{TP + FP} \approx 65.2\%$$

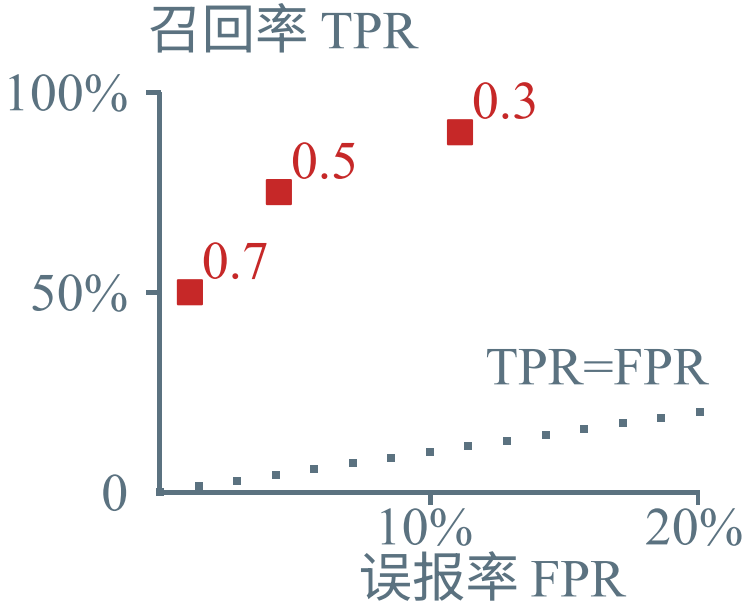
召回率：故障被检出的比例；漏报率：故障被漏掉的比例。
误报率（FPR）：正常被误报警的比例；查准率：报警中真故障的比例。
统计单位也要写明：窗口、整次飞行与每小时告警数具有不同含义。

7.5.3 评价与跨工况：分数阈值与ROC

阈值决定误报和漏报的取舍

接收者操作特性（ROC）：横轴为误报率 FPR。
纵轴为真阳性率 TPR，即故障召回率。

分数阈值	检出 / 漏报	误报 / 正常	TPR；FPR
0.3	18 / 2	20 / 180	90%；11.1%
0.5	15 / 5	8 / 180	75%；4.4%
0.7	10 / 10	2 / 180	50%；1.1%



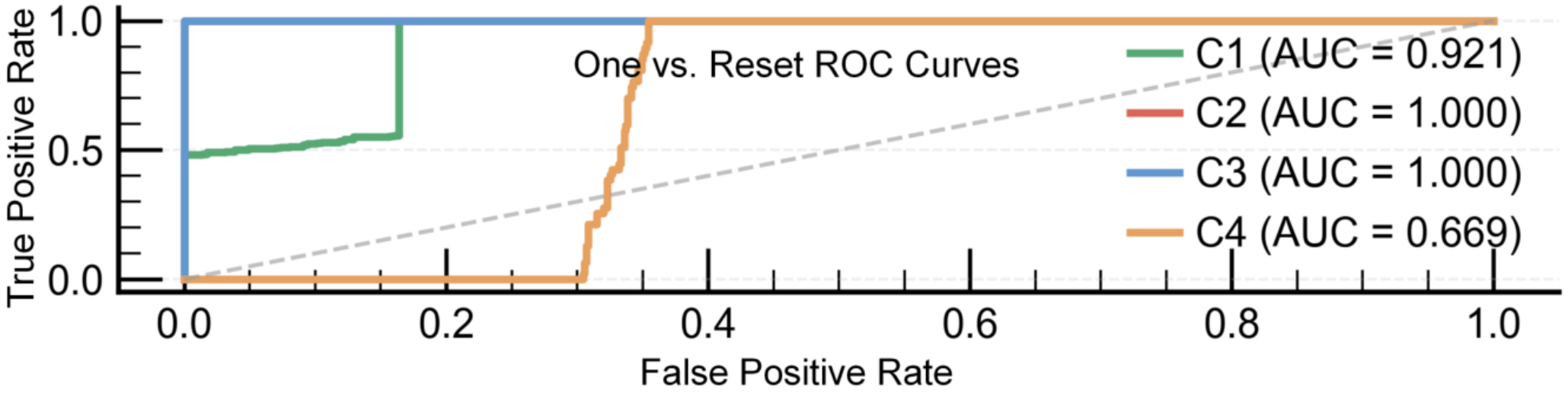
图：教学算例的ROC工作点

本表由同一组200个教学分数复算；每个阈值是右图ROC上的一个工作点。

7.5.3 评价与跨工况：类别ROC与AUC

曲线面积衡量区分能力，含义不同于准确率

曲线下面积（AUC）：汇总多个阈值的区分能力；这里按每个类别对其余类别评价。
类别对照：C1正常，C2加速度计噪声，C3单电机信号跳变，C4电机3效能缩放0.85。

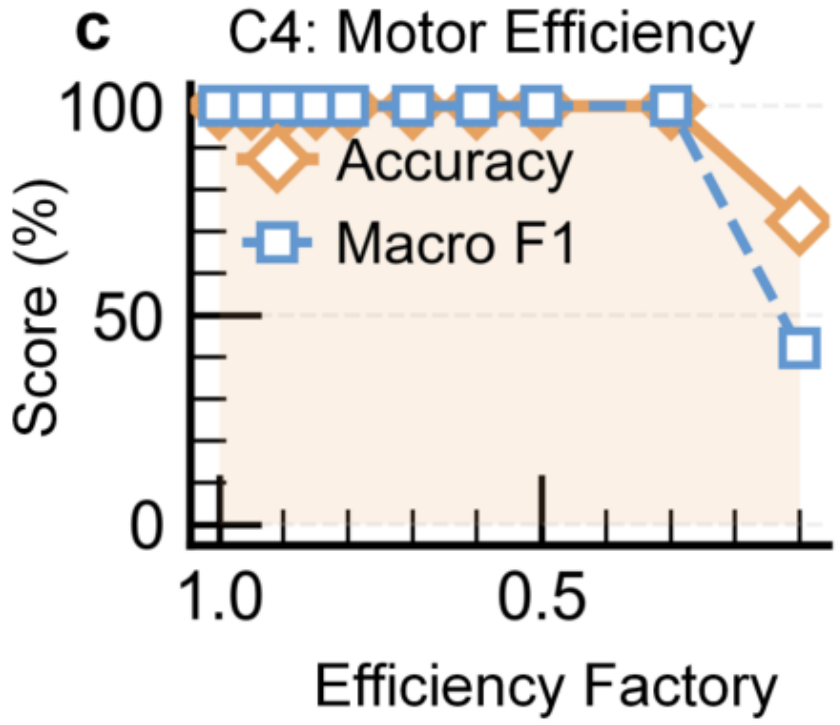


图：同上论文 Fig.11(d) 基准测试集的类别ROC

读图：越靠左上越好；C2与C3贴左上边界；C4在FPR<0.3时几乎检不出。

7.5.3 评价与跨工况：故障强度扫描

退化最强时，准确率与宏F1下降



图：同上论文 Fig.11(c) 强度扫描

条件：正常基础样本上对电机3效能因子η做乘性缩放；
与「类别ROC与AUC」页基准测试集（η=0.85，AUC=0.669）的测试条件不同。

橙线为准确率；蓝线为宏平均F1，
即各类别F1的等权平均。

横轴η向右减小，退化加重；
分数只在最强缩放端下降。

η是仿真对电机输出的乘性系数；
退化程度需回归或健康指标估计。

7.5.3 评价与跨工况：跨工况与数据泄漏

论文代码入口的相邻窗口重叠达95%

整次飞行 / 工况分组 → 组内切窗口 → 训练侧拟合尺度 → 独立测试

窗口 i：第1–80点 → 训练

窗口 i+1：第5–84点 → 测试

共享76点
泄漏反例

$$\text{overlap} = \frac{80 - 4}{80} = 95\%$$

划分原则： 见课件7.3.5「滑窗数据泄漏」与「训练验证测试划分」页。

代码入口： 论文notebook一处设置为80点窗口、步长4点，相邻窗口共享76点。

跨工况： 按载荷、机动条件把完整飞行分组，测试组逐类工况报告表现。

读代码时： 先查拆分在切窗之前还是之后；论文其他验证的划分以原文为准。

7.5.3 评价与跨工况：窗口、流水线与告警

检测延迟从故障生效时刻开始计时

时刻 / s	已知事件	告警状态
10.0	故障实际生效	起算点
10.4	首次超阈值	等待连续确认
10.6	第二次连续超阈值	告警确认

$$\Delta t_{\mathrm{det}} = 10.6 - 10.0 = 0.6\ \mathrm{s}$$

同上论文 Table VII：前向中位数（P50）10.13 ms，采集到输出流水线43–54 ms。
条件：惯性测量单元（IMU）50 Hz，80点窗口、480个图节点，地面RTX3060。
前向与流水线耗时不同；告警还要等待连续证据，延迟从故障生效起算。

本章目录

故障诊断与健康预测

第七章 健康管理与安全评估

7.1 健康管理与安全评估基础
7.2 统一故障建模框架
7.3 虚实安全强化与故障数据方法
7.4 故障危害与安全评估
7.5 故障诊断与健康预测
7.6 安全决策与整机改进
7.7 故障注入接口与算法开发框架
7.8 例程实验与虚实验证
7.9 习题与课堂讨论

本节 4 个教学单元

7.5.1 残差与基线
教材对应：第7.1.2.1、7.1.5.4节

7.5.2 融合诊断
教材对应：第7.1.5.4节

7.5.3 评价与跨工况
教材对应：第7.1.5.4节

7.5.4 健康预测
教材对应：第7.1.2.1节（电池退化预测）

7.5.4 健康预测：电池状态与寿命

三个量的定义：SOC、SOH看当前，RUL看未来

荷电状态（SOC）	容量健康状态（SOH）	剩余可用寿命（RUL）
剩余电荷 / 当前可用满充电荷	当前满充容量 / 额定容量	到指定阈值的剩余循环次数或时间
概念算例	概念算例	概念算例
$1.6/1.6 = 100\%$	$1.6/2.0 = 80\%$	还须给定退化规律和未来使用条件

算例的1.6 Ah为教学给定值；SOC、SOH描述当前状态，RUL描述未来。

7.5.4 健康预测：电池状态与寿命

同样充满电，老电池为什么更早耗尽？

同为 SOC 100%，满充容量却不同

新电池：额定 2 Ah

满充容量 2.00 Ah



已充满

剩余电荷 2.00 Ah

容量型 SOH 100%

旧电池：额定 2 Ah

满充容量 1.60 Ah



已充满

剩余电荷 1.60 Ah

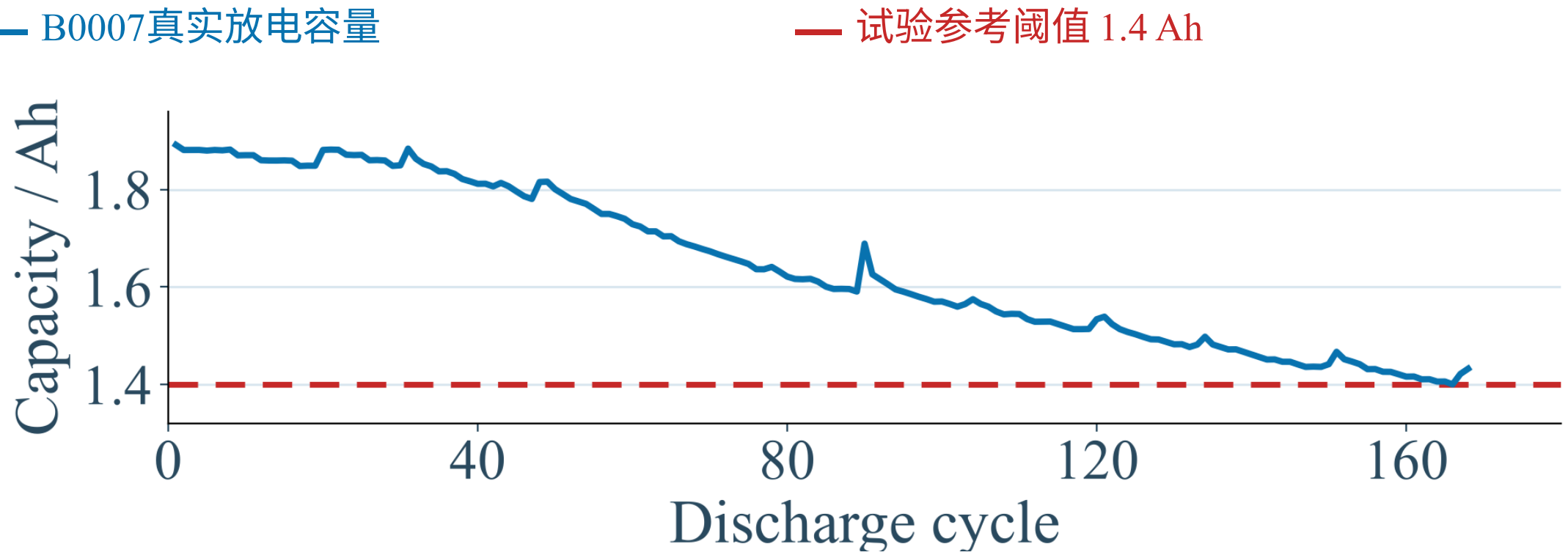
容量型 SOH 80%

充满不等于恢复健康；RUL 还要看退化历史

图：相同电流下的容量对照；充电恢复SOC，不能恢复已损失的容量。
容量健康状态是本例的SOH指标；RUL还需要退化历史与未来使用条件。

7.5.4 健康预测：NASA电芯记录

真实容量记录显示退化，也存在短期回升



图：NASA PCoE电池老化数据第5项，B0007电芯放电容量（环境24 °C）

7.5.4 健康预测：趋势外推的一步计算

只用第100次以前的数据预测阈值时刻

预测时刻：第100次；拟合范围：第71–100次；单位：容量Ah，循环编号无量纲。

只取第71–100次 → **最小二乘直线拟合** → **外推到1.4 Ah**

寿命终止（EOL）：容量首次≤1.4 Ah的循环；假设近期斜率与使用条件不变。

$$\hat{Q}(n) = 1.8500 - 0.002731n$$

$$\hat{n}_{\mathrm{EOL}} = \frac{1.8500 - 1.4}{0.002731} \approx 164.8$$

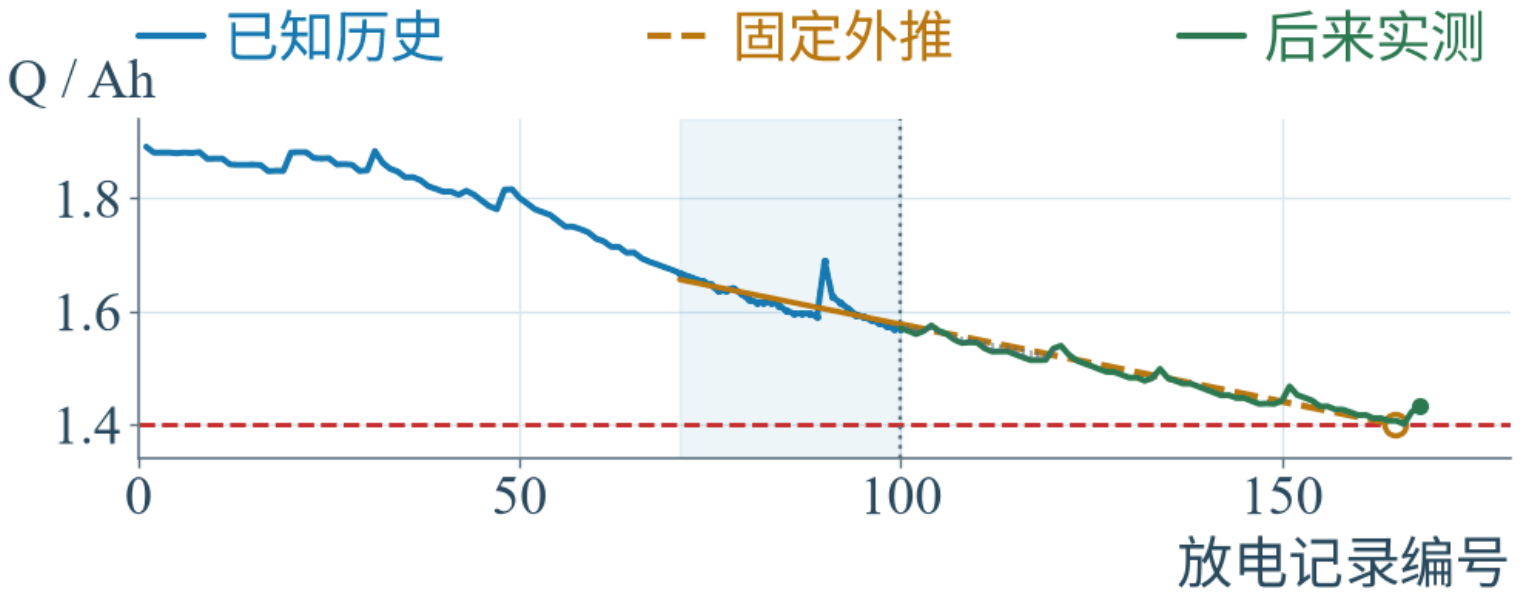
$$\widehat{\mathrm{RUL}} = \hat{n}_{\mathrm{EOL}} - 100 \approx 64.8 \text{ cycles}$$

7.5.4 健康预测：历史与未来的隔离

先冻结历史做预测，再揭示后来发生了什么

预测已冻结；后来观测用于检验

NASA B0007



预测 RUL
64.8 次循环
预测触阈 164.8

168 点内未越线
最低 1.4005 Ah

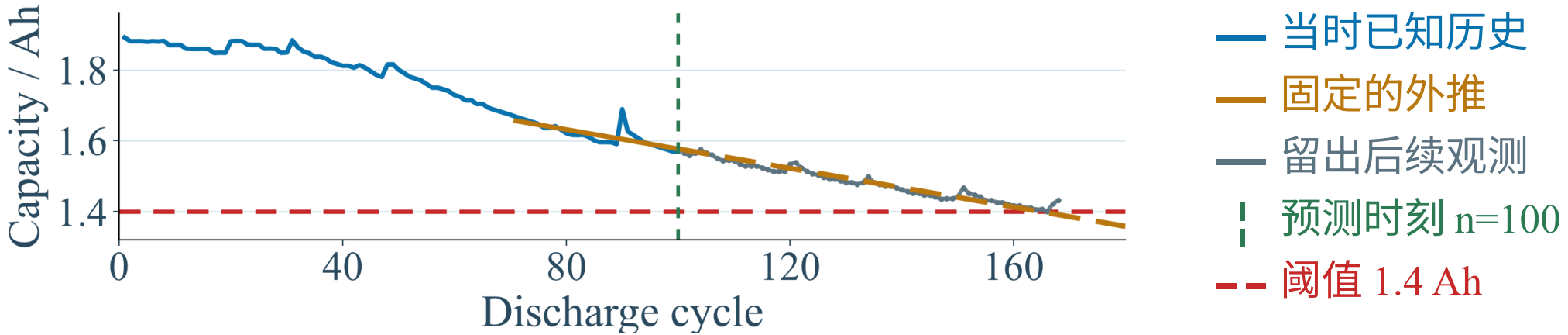
阈值 1.4 Ah

后20点 MAE = 0.0099 Ah；真实 RUL 仍未知

图：NASA B0007实验室容量记录；第100次冻结预测，后续观测用于验证。
记录期内尚未达到1.4 Ah，真实剩余寿命未知。

7.5.4 健康预测：历史与未来的隔离

后续观测只用于检验，不能回流进预测



图：B0007历史、固定外推与留出观测

平均绝对误差（MAE）：预测容量 $\hat{Q}(n)$ 与实测容量 $Q(n)$ 的绝对差平均值。

$$\mathrm{MAE}_{101:120} = \frac{1}{20}\sum_{n=101}^{120} |\hat{Q}(n) - Q(n)| \approx 0.0099\ \mathrm{Ah}$$

7.5.4 健康预测：误差与区间的来源

退化率情景范围不是统计置信区间

拟合当前容量 $\hat{Q}(100)$=1.5769 Ah，阈值 Q_{lim}=1.4 Ah；端点退化率 d 为教学设定。

$$\widehat{\mathrm{RUL}} = \frac{\hat{Q}(100) - Q_{\mathrm{lim}}}{d}, \quad d > 0$$

容量降幅 / (Ah/循环)	估计RUL / 循环	含义
0.0020	88.4	较慢退化情景
0.002731	64.8	第71–100次斜率的绝对值
0.0040	44.2	较快退化情景

统计区间须有误差模型或独立校准集，并报告覆盖率；本例仅做参数敏感性分析。

7.5.4 健康预测：预测验证与记录截尾

曲线未越过阈值，真实寿命尚未被观测

预测截止点	仅过去30点拟合	预测RUL / 循环	后20次MAE / Ah
60	31–60	82.5	0.0259
80	51–80	38.5	0.0431
100	71–100	64.8	0.0099

右截尾：记录在第168次结束，按容量≤1.4 Ah判据尚未越线。
最低容量约1.4005 Ah；记录终点是观测停止，寿命终点尚未出现。
删失只给出界：真实EOL>168，第100次预测的164.8至少偏早3.2次。

7.5.4 健康预测：NASA入口与离线任务

读准数据结构，再选择预测工具

仓库第5项MAT：cycle → 筛选type为discharge → data.Capacity（Ah）。
离线任务：给定容量表，截断历史、拟合趋势、检验后续20次误差。

NASA PCoE预测数据仓库：老化数据结构与试验说明

https://data.nasa.gov/dataset/li-ion-battery-aging-datasets

ProgPy模型与数据加载代码

https://github.com/nasa/progpy

ProgPy的nasa_battery面向仓库第11项随机负载数据，电芯名为RW1–RW28。
B0007属仓库第5项（Battery Data Set），按其MAT结构读取。
放电终止（EOD）是一段放电结束；寿命终止（EOL）是退化阈值到达。

练习与讨论

残差多大才告警，哪些量不能作输入

问题1：正常残差 $\mu_0 = 0.04$ rad/s，$\sigma_0 = 0.05$ rad/s；$|z_k| > 3$ 连续2点才告警。

0.0–0.5 s 每0.1 s 的残差为 0.10、0.18、0.21、0.12、0.20、0.22 rad/s，故障在0.35 s生效。求各 z_k、告警时刻和检测延迟。

问题2：用注入故障的仿真日志训练诊断模型，候选输入有四项：
①残差窗口；②日志中的故障ID；③故障注入时刻；④电机输出指令。

哪些输入会造成标签泄漏？说明判断依据。

参考要点

先减均值再除尺度，标签只作训练目标

$$z_k = \frac{r_k - \mu_0}{\sigma_0} = \frac{r_k - 0.04}{0.05}$$

z_k：1.2、2.8、3.4、1.6、3.2、3.6

计数：0、0、1、0、1、2，0.5 s 确认告警；漏减 μ_0 得2.0、3.6、4.2…，0.2 s 即误报。

$$\Delta t_{\mathrm{det}} = 0.5 - 0.35 = 0.15\ \mathrm{s}$$

问题2：②③由标签或试验设定生成，部署时得不到，属标签泄漏；①④运行时可得，可作输入，仍按完整飞行划分评价。

天然标签：注入参数自动转成训练标签，只作监督目标，不进输入。

练习与讨论

总体分数更高，就应该提高阈值吗

沿用200个教学窗口：正常180、故障20。阈值0.7时TP=10、FN=10、FP=2、TN=178。

问题1：计算阈值0.7时的Accuracy、查准率和F1（查准率与召回率的调和平均）。

问题2：与阈值0.5的Accuracy=93.5%、召回率75%、误报率4.4%比较，怎样选择？

问题3：故障2.3 s生效，$|z_k| > 3$ 连续3点确认；2.0–3.2 s每0.2 s的 z_k 为0.5、3.3、3.5、2.8、3.2、3.6、4.1。求检测延迟。

参考要点

提高阈值减少误报，却增加故障漏检

$$\text{Accuracy} = \frac{10 + 178}{200} = 94\%$$

$$\text{Recall} = \frac{10}{20} = 50\% \qquad \text{FPR} = \frac{2}{180} \approx 1.1\%$$

查准率：10/12≈83.3%（阈值0.5时65.2%）；F1=2TP/(2TP+FP+FN)=20/32=62.5%。
取舍：误报减少，漏检由25%增至50%；按任务后果与处置成本选择。
问题3：2.6 s计数归零，3.2 s第3点确认；起点取生效时刻2.3 s。

$$\Delta t_{\text{det}} = 3.2 - 2.3 = 0.9\ \text{s}$$

练习与讨论

这条容量曲线能证明RUL预测正确吗

问题1：在第100次预测时，把第101–120次加入拟合或选窗，是否仍是历史预测？

问题2：记录到第168次结束但未越过1.4 Ah，能否直接取168作为寿命终点？

问题3：后20次容量MAE约0.0099 Ah，能否据此宣布RUL误差很小？

问题4：44.2–88.4次来自退化率0.004–0.002，为什么不能标95%置信区间？

参考要点

有容量误差证据，仍缺真实终止时刻

历史隔离： 第100次预测只用截至100次的观测；未来仅用于冻结模型后的验证。

记录截尾： 第168次未达到给定阈值，记录终点不等于寿命终点，真实RUL未观测。

评价范围： 0.0099 Ah是未来20次容量MAE；删失下RUL误差只能给出界。

区间含义： 44.2–88.4次是手选参数敏感性范围；95%区间须给出生成方法与覆盖率。

本节小结

诊断与预测要把判断依据一起交付

残差与基线：同输入、时间、坐标和单位下比较；先校准正常偏差，保留故障偏离。

融合与评价：残差、传感关系与时序共同判断；标签不进输入，按独立工况评价。

误报与延迟：报告明确的分母、阈值和计时起点；
AUC不同于Accuracy，网络前向推理时间不同于检测延迟。

健康预测：SOC、SOH、RUL分开；只用历史，写明阈值、区间来源与截尾。