



Manufacturing Data Science

Prognostics and Health Management (第 13 章 故障預測與健康管理)

Chia-Yen Lee, Ph.D. (李家岩 博士)

Department of Information Management (資訊管理學系)
National Taiwan University (國立臺灣大學)

- 第一章 製造數據科學
- 第二章 製造系統分析與管理
- 第三章 數據科學基礎與模型評估
- 第四章 數據科學分析架構與系統運算決策
- 第五章 數據預處理與製造數據特性
- 第六章 線性分類器
- 第七章 無母數迴歸與分類
- 第八章 決策樹與集成學習
- 第九章 特徵挑選與維度縮減
- 第十章 類神經網路與深度學習
- 第十一章 集群分析
- 第十二章 特徵工程、數據增強與數據平衡
- **第十三章 故障預測與健康管理**
- 第十四章 可解釋人工智慧
- 第十五章 概念漂移
- 第十六章 元啟發式演算法
- 第十七章 強化學習

藍：老師課堂講授
綠：學生自學

- 附錄A 線性迴歸
- 附錄B 支持向量機
- 附錄C 統計製程管制與先進製程控制
- 附錄D 超參數最佳化

- 應用涵蓋
產能規劃、瑕疵檢測、製程監控與診斷、機台保養、需求預測、生產排程、電腦視覺、自動光學檢測、原料價格預測與採購等

□ 本章介紹「**故障預測與健康管理**」(prognostics and health management, PHM) 使我們得以將數據轉換成量化且可視化的資訊，增進我們對系統健康狀況的知識理解（降低不確定性），從而強化對系統維護與保養的策略，減低系統故障所產生的風險

- 然而，此方法所以複雜在於其涉及了橫跨感測技術（sensing technology）、控制理論（control theory）、機械工程（mechanical engineering）、物理學（physics）以及可靠度分析（reliability analysis）等多項科學領域，整合不易。
- 幸運的是，透過資訊技術演進與數據科學，得以進一步地整合上述領域，以預測減低對健康狀態的不確定性，並考慮保養決策最小化故障風險所帶來的損失。

□ 保養策略的定義

- 在製造業裡，系統、設備、機台以及元件的保養（maintenance）是維持穩定且高品質生產的一大關鍵。為何需要保養呢？因設備可能發生失效（failure），也就是設備未能發揮預期的作業機能狀態。理論上，在可靠度工程中假定一般設備其失效率（failure rate）滿足「浴缸曲線」（bathtub curve），如圖所示（Klutke et al., 2003），以下分別說明三個階段。

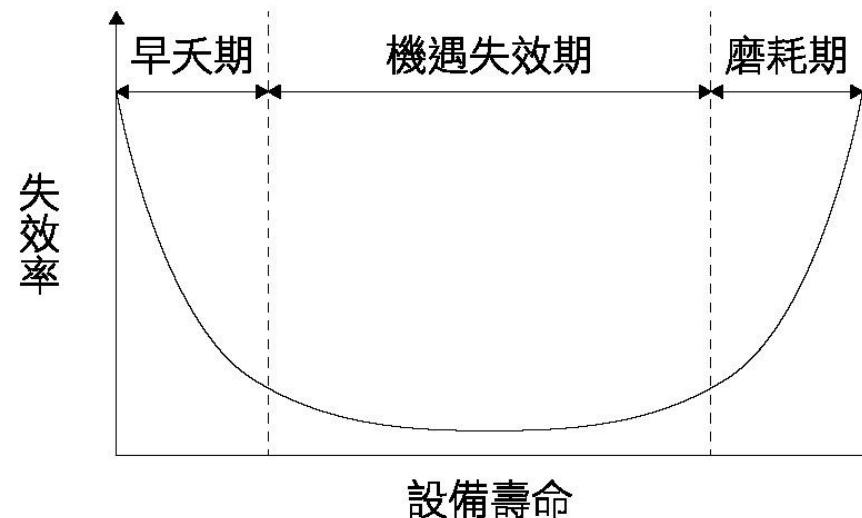
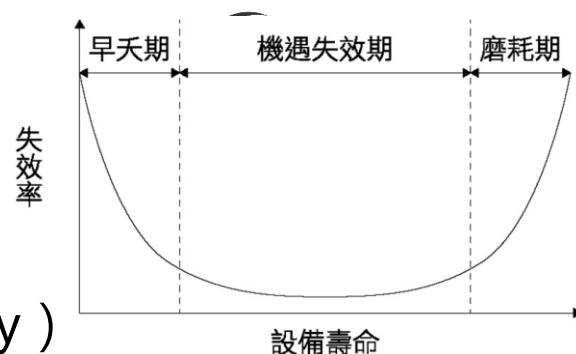


圖 13.1 沐缸曲線

保養策略的演進

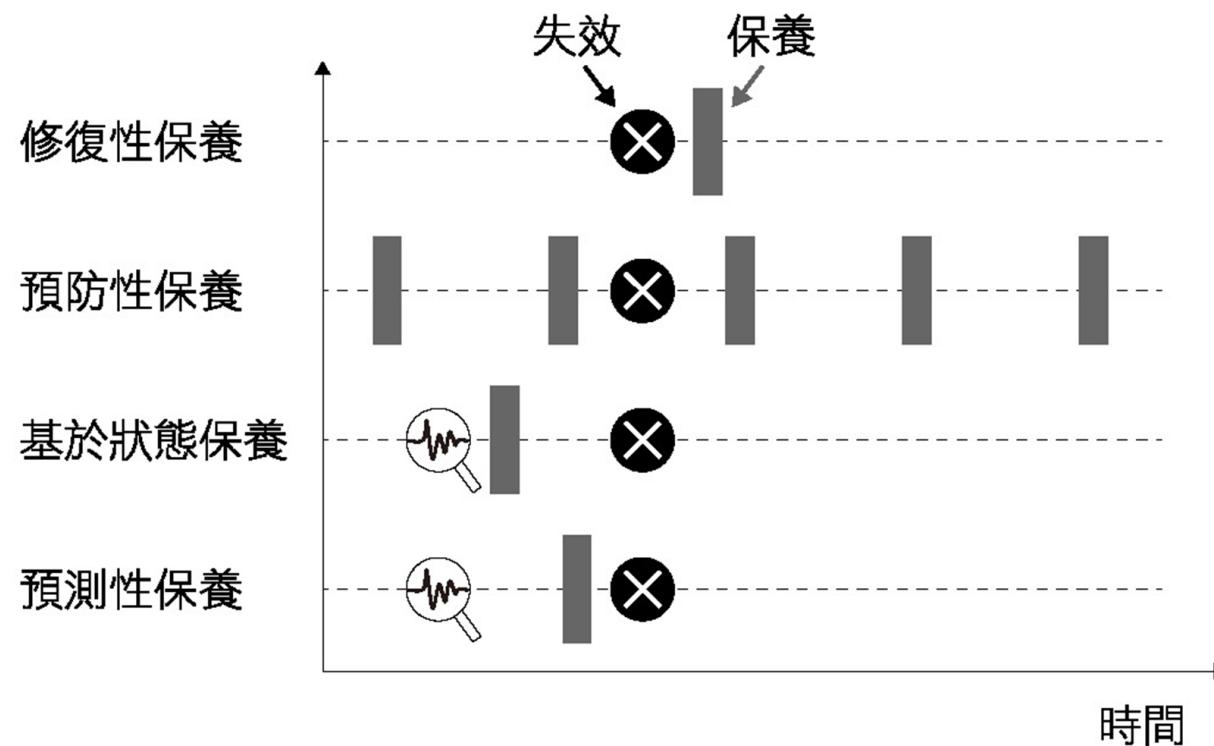
□ 浴缸曲線

- 「早夭期」 (early failure , 或稱嬰兒期 infant mortality)
 - 其失效主要原因來自於**先天缺陷準備不足**，例如設備或元件設計缺陷、材料或製程檢驗不符合標準、安裝或使用的不當、儲存或運輸上的損毀等。換言之，一般在收到產品後，應先按照**使用手冊** (user guide) 內的說明，檢驗產品外觀、功能、安裝、與操作流程，來確認裝置或產品在啟動時的符合規範與安全性。
- 「機遇失效期」 (chance failure , 或稱使用期 useful life)
 - 其失效主要原因來自於**環境突發狀況**，例如不當外力或汙染、環境工況變化、人為失誤**MO** (missing operation , 例如未按照SOP施行) 、設備未有足夠的容錯能力、天災或意外等。因此，平日應做好人為教育訓練，並適當地對環境工況加以管制，以維持適當的生產條件，避免減少設備或元件的使用壽命。
- 「磨耗期」 (wear-out , 或稱老化期 aging)
 - 其失效主要原因來自於**長期累積與使用**，例如磨耗、疲乏、腐蝕、龜裂、老化、不當替換與翻修等。一般來說，在適當的時機透過各類的維修保養，來延緩其失效發生的時間。



□ 保養的演進 (Ran et al., 2019)

- 由傳統被動的「修復性保養」(corrected maintenance, CM)、「預防性保養」(preventive maintenance, PM)、「基於狀態保養」(conditional-based maintenance, CBM)、發展至「預測性保養」(predictive maintenance, PdM)。以下將依序說明此四種保養策略的特性以及其各自的優劣。



□ 修復性保養 (CM)

- 是一種「運行到故障」(run-to-failure) 時才進行的維修策略，也就是在故障時後才採取相對應的措施進行補救，是一種被動策略。雖然此策略是最大限度地運作以最大化產能，並且不花費任何時間與成本進行事前的保養。然而，其缺點在於**運行到故障**所產生的通常是龐大且嚴重的損失，可能一個元件導致整個機台或相關設備的損毀，抑或是系統重製導致一段時間生產線的停擺，這樣的損失代價是一般產業不期望發生的。

□ 預防性保養 (PM)

- 則是一種週期性、計畫性的保養策略，也就是對不同的系統、設備、機台以及元件以領域知識（或經驗法則）訂定相對應的保養週期或指標，從而依據保養週期或指標進行保養。此策略相對於**CM**盡可能減少故障發生的可能，然而其缺點一方面是即使處在正常的運行狀態下也會週期性的執行維修保養，因而過度地保養將產生額外的停機時間（產能損失）和保養成本，另一方面依舊無法避免潛在的突發故障。事實上，**PM**主要是針對浴缸曲線中間較為穩定的**使用期**做出合適的保養策略。

□ 基於狀態保養 (CBM)

- 主要是透過監控的手法，可以透過感測器（或驅動器、控制器）隨時掌握設備狀態，及時發現異常並立刻維護。當出現徵兆在發生故障前，便有效預防加以控制及排除，減少嚴重故障事件。**CBM**需要安裝即時監控的裝置，隨時收集狀態的相關數據，後端也需要建置龐大的數據庫系統與對應的計算資源，以利監控。根據**CBM**再進一步地延伸，透過數據預測設備或其零組件狀態，便演進到**預測性保養**。

□ 預測性保養 (PdM)

- 則是預測設備或元件的狀況，直接預測何時可能發生故障並提前執行維修保養的策略，使得我們能在**維護及保養成本**（保養頻率）與**維修 (repair) 成本**間取得權衡最小化耗費的成本。基於數據科學的流程架構，由感測器的數據收集進行**健康指標**（health index）的建構，再階段性區分設備健康狀態，以至預測**剩餘可用壽命**（remaining useful life, RUL）的模型建構，模型也包含預測何時發生故障、發生可能的故障類型、剩餘壽命、以及故障可能影響的層面與風險損失等。值得留意的是，**PdM**需考量額外建構與維護資訊系統的成本，其包含在上述維護及保養成本中。

表 13.1 保養策略的比較

保養策略	利益與好處	挑戰與困難	使用時機
CM	<ul style="list-style-type: none"> - 最大產能使用率 - 無保養成本 	<ul style="list-style-type: none"> - 未預期的停機 - 較高的備件存貨成本 - 設備潛在的損害 - 較高的維修成本 	<ul style="list-style-type: none"> - 重複性、標準性、或非關鍵的設備 - 機台當機後維修成本較低
PM	<ul style="list-style-type: none"> - 減少設備故障與非預期停機 - 較低維修成本 	<ul style="list-style-type: none"> - 有一定的存貨 - 預期內停機時間上升 - 機況良好的設備依然保養（不修沒事、一修反而有問題） - 保養成本增加 	<ul style="list-style-type: none"> - 隨著週期時間或使用時間（里程數）發生故障的可能性會上升
CBM	<ul style="list-style-type: none"> - 立即偵測設備異常與維護處理，降低非預期停機 - 該保養時才保養，增加設備使用率，減少不必要的保養與維修成本 	<ul style="list-style-type: none"> - IT基礎建設的建置，包含設備端的感測器與後端的數據庫與分析系統 - 在工況或環境噪音較多的情況下，常造成假警報（false alarm）過多 	<ul style="list-style-type: none"> - 十分關鍵性的設備
PdM	<ul style="list-style-type: none"> - 全面性的設備健康 - 建立分析架構，潛在改善整體工廠效益（包含浪費排除、生產力提升、品質改善等） - 避免運行到故障 - 減少不必要的零組件更換 	<ul style="list-style-type: none"> - 包含 CBM 的挑戰與困難 - 有前端感測後端數據分析，工廠運行系統將更為複雜 	<ul style="list-style-type: none"> - 透過監控與數據收集能有成本效益的故障模式預測

□ PHM問題層面與系統整合

- 以更全面且系統化的方式強化上述提及PdM策略。若深入檢視保養問題，我們欲解決以下的幾個疑問 (Lee et al., 2014)
 - 機台此刻的運行狀態是什麼？(What is the machine operating status now?)
 - 「健康評估」 (health assessment)
 - 機台主要可能發生的故障是何種類型？是什麼原因造成該故障的？ (Which primary faults happened to the machine? Why does the fault occur?)
 - 「故障診斷」 (diagnosis)
 - 機台發生嚴重故障會在什麼時候？是什麼原因造成衰退的速度加快？ (When will the machine break down? Why does the degradation accelerate?)
 - 「故障預測」 (prognosis)
 - 機台應如何建立保養策略？ (How should we conduct the maintenance strategy?)
 - 「保養決策最佳化」 (maintenance strategy optimization)

□ 健康評估

- 是藉由數據收集、預處理與特徵萃取並執行CBM。比較現況感測器所收集的特徵以及經萃取後的特徵（例如：時域頻域轉換、統計量轉換等）的分布與樣型是否與期望穩定的狀態有所差異，以管制的方式進行監控並於異常時發出警報。並可同時對多個機台或多個特徵進行健康指標的整合計算。

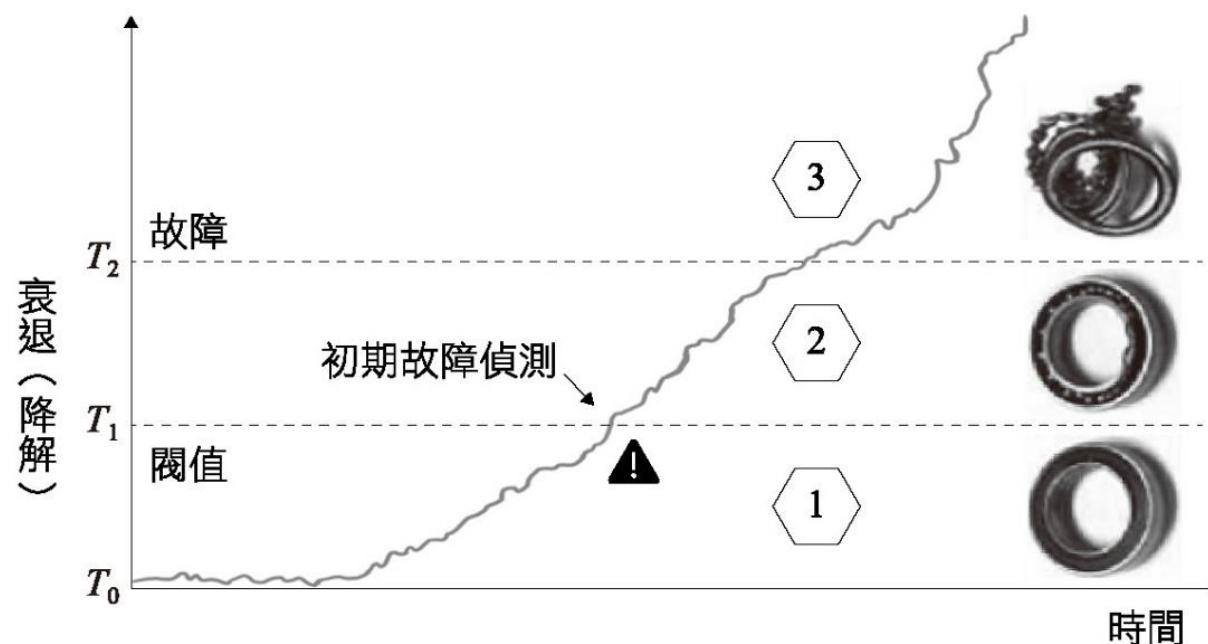


圖 13.3 設備與其零組件的階段性衰退 (Atamuradov et al., 2017)

□ 故障診斷

- 可藉由**數據科學**的分類模型與**特徵工程鑑別關鍵因子**來找出故障的種類與原因，如圖所示，例如可由加速規所收集的振動訊號剖析出軸承損耗可能是發生在**內環**（inner race）、**外環**（outer race）或是**滾珠**（rolling element）等異常。

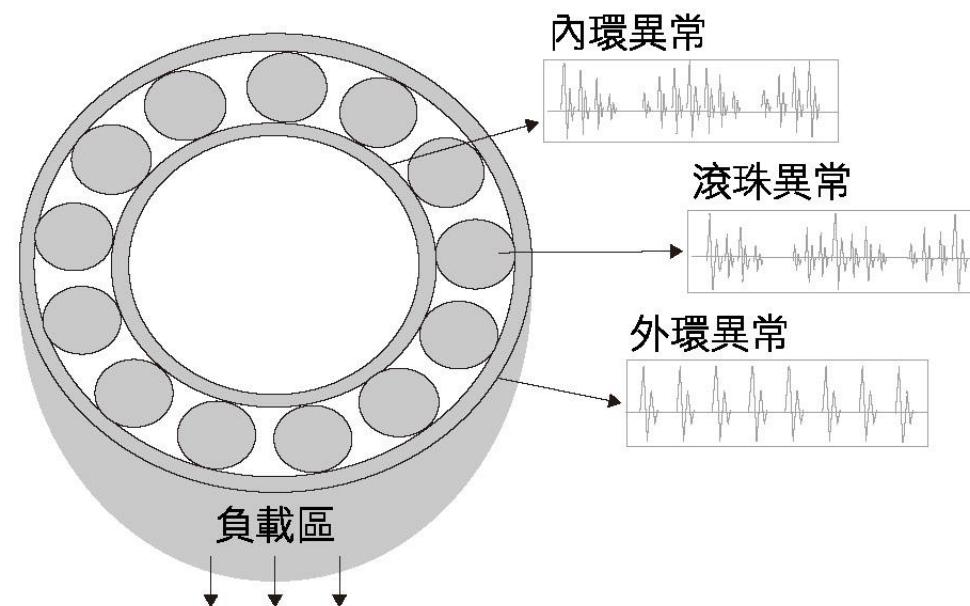


圖 13.4 軸承的故障診斷

□ 故障預測

- 可藉由迴歸模型來預測**剩餘可用壽命** (Gouriveau et al., 2016)。為了簡化問題，我們假設某一軸承的「衰退狀態」 (degradation state) 為一維度的量化數值，因此隨著時間的流逝，衰退狀態的數值將逐漸越大直到軸承完全損毀。
- 因此我們可估計機台的剩餘可用壽命為某個時間點 tc 直到「故障閾值」 tf (failure threshold)的時間長度， $RUL = tf - tc$
 - 以統計信賴區間而言而，隨著時間的推移我們對軸承剩餘可用壽命的信賴區間會逐漸由大到小，也就是越接近故障時我們所估計的剩餘可用壽命會更加精確

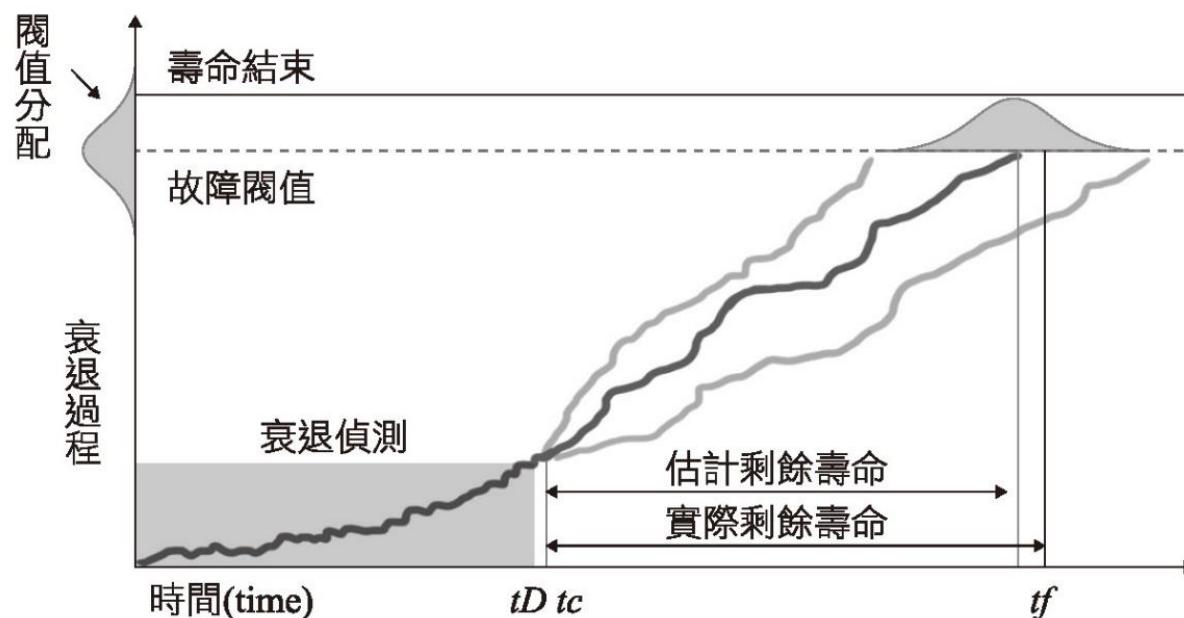


圖 13.5 軸承的故障預測 (Gouriveau et al., 2016)

□ 保養決策最佳化

- 是藉由前述三個面向瞭解機台狀態並透過數據收集降低部分不確定性後，以最佳化的保養決策降低**總成本**以及增加**可靠度與可用性** (reliability and availability)。一般來說，需決定哪一個設備要優先保養（設備維修排序）以及該設備需要維修的大約時間點。
- 決定哪一個設備要優先保養(設備維修排序)
 - (1)從上次保養到目前已使用時間長度
 - (2)從上次保養到目前已加工的批量數量(接近里程數的概念)
 - (3)加工品質變異(良率)
 - (4)WIP堆貨狀況
 - (5)需求目標(某段時間內規劃應該要完成的產品數量)等
- 該設備需要維修的大約時間點
 - 決定保養時間是除了考慮上述部分的因子外，需要權衡兩種主要風險
 - 產能損失風險(提早保養會損失產能而來不及出貨)
 - 良率損失風險(延後保養需承擔產品品質變異而有不良品的狀況)
 - 在各別計算成本後使總成本最小化。

□ 若將上述PHM的四個面向**健康評估**、**故障診斷**、**故障預測**以及**保養決策最佳化**以系統的視角呈現 (Gouriveau et al., 2016)

- 在一個運行的系統下，對設備規律地進行**健康評估**，透過對數據量化、特徵萃取、監控與視覺化當下的健康狀態。當故障被偵測時，同時進行**故障診斷**與**故障預測**來判別是哪個類型的故障以及其即將發生的時間，從而提供**保養決策最佳化**最充分的資訊並對系統執行實際的保養。

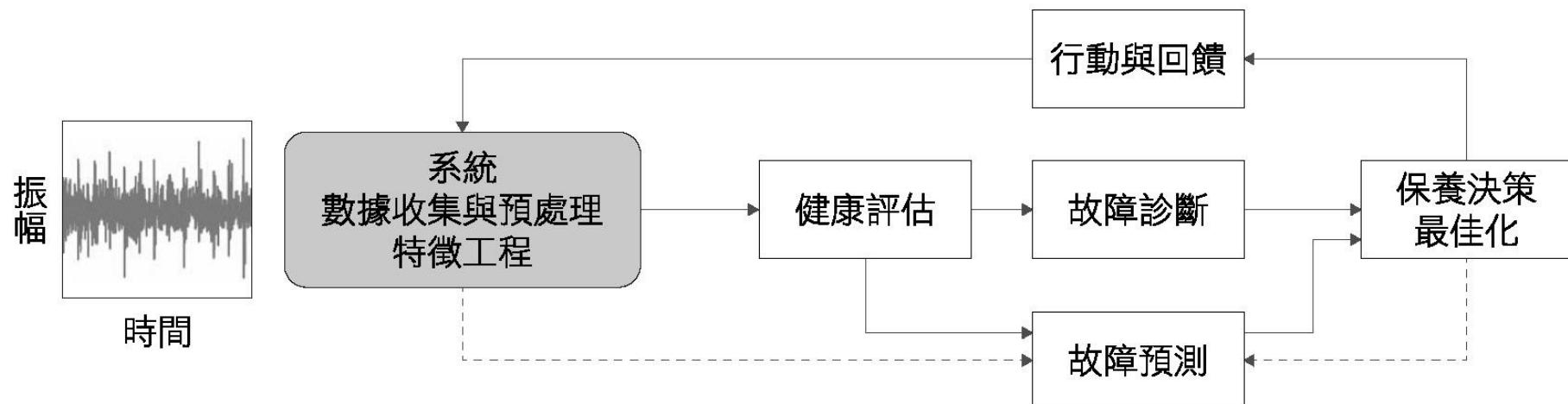


圖 13.6 健康評估、故障診斷、故障預測與保養決策最佳化 (Gouriveau et al., 2016)

□ 數據收集與預處理

- 直接 (direct) 且通常為離線 (off-line) 的方法〔亦多為基於模型 (model-based) 的方法〕
 - 對設備、機構本身物理特性直接地量測，例如與材料特性有關的剛性 (stiffness) 、阻尼 (damping) 、慣量 (mass) 、摩擦力 (friction) 、齒隙 (backlash) 等。
 - 具備充分的領域知識(針對不同設備對應不同的物理特性)與離線量測的成本(量測工具與時間)
- 間接且通常為線上 (on-line) 的方法〔亦多為基於數據 (databased or model-free) 的方法〕
 - 是以感測器或內建的驅動器存取數據，例如電流 (current) 、振動 (vibration) 、聲發射 (acoustic emission) 、油屑 (oil debris) 、影像 (image) 、溫度 (temperature) 、壓力 (pressure) 等與設備間接相關的特性，須由特徵工程進一步對特徵進行萃取與轉換才能得出與設備有關的健康狀態。
 - 需在設備上增加額外的量測裝置(例如加速規)以及透過龐大的運算力以合適的特徵工程萃取出與健康有關的特徵。

□ 數據收集與預處理

- 理論上，若能同時以**直接**與**間接**方法一併收集數據最為理想
 - 實務上須考量資源的限制以選擇相對**成本效益高**(cost-effective)方法進行數據收集與計算，考量數據收集在時間上(取樣頻率)與空間上(感測器的數量與位置)的解析度。
 - 例如在時間上，長期的故障預測可以較低取樣頻率的方式收集數據，而當接近故障時則逐步提高數據收集的頻率
 - 在空間上，可依據機台加工的軸向(x、y、z軸)安裝對應的量測裝置(加速規)，而安裝位置的選擇十分重要，通常直接的決定了數據的品質。
 - 可依照機台的作動，例如前進、後退、上升、下降、旋轉等切割對應的振動訊號，健康狀態的評估經常與某幾項作動息息相關，例如上升作動可因負載重量的大小(工況不同)而呈現出不同的老化狀況。
- **數據預處理**中，當數據訊號的採樣頻率相對高數據量過大時，可對時域訊號使用**滑動時窗** (sliding window)，選取資料範圍並取出該時窗內訊號的統計量 (例如平均值、最大值) 以簡化數據。

□ 特徵工程

- 以感測器與驅動器所收集的數據經過數據預處理後，可進行特徵工程
 - 第一，可使用時域轉換的方法，包含了滑動時窗特徵的時域統計量或以時間序列分解後再做時域統計量的轉換。
 - 第二，可使用頻域轉換以及時頻域轉換的方法，包含了傅立葉轉換、短時距傅立葉轉換、小波轉換以及希爾伯特-黃轉換 (Hilbert-Huang transform) 等，以頻域或時頻域的視角萃取特徵。
 - 抑或在轉動設備進行進階的頻譜分析，例如階次追蹤(order tracking) (Wang et al., 2019) 可考慮轉速不固定下(變速)，在恆定角增加量在原始信號中將時域中的非平穩振動信號轉換為角域(angular domain)中的平穩振動信號，以分辨轉速的分數與倍數的變化問題，以萃取特徵。
 - 在收集了各式各樣的特徵之後，可透過特徵挑選或維度縮減方法以找出對機台健康有代表性的重要特徵。

表 13.2 轉動機械系統的相關元件、故障缺陷與其特徵 (Lee et al., 2014)

元件	故障與缺陷	特性	常見的測度	常用的特徵
軸承 (bearing)	<ul style="list-style-type: none"> 外環缺陷 (outer-race defect) 內環缺陷 (inner-race defect) 滾珠缺陷 (roller defect) 保持架缺陷 (cage defect) 	<ul style="list-style-type: none"> 低資訊量 低振幅 高噪音 	<ul style="list-style-type: none"> 振動 油屑 聲發射 	<ul style="list-style-type: none"> 時域統計量 振動頻率特性 應力波、脈衝與金屬碎片的特性
齒輪 (gear) 皮帶 (belt)	<ul style="list-style-type: none"> 脫落 (spall) 斷裂 (crack) 鬆弛、磨損 (fatigue, wear) 	<ul style="list-style-type: none"> 高噪音 易受其他元件影響 (軸承、軸) 	<ul style="list-style-type: none"> 振動 油屑 聲發射 	<ul style="list-style-type: none"> 時域統計量 振動頻率特性 油屑與機構系統的特性
軸 (shaft)	<ul style="list-style-type: none"> 不平衡 (unbalance) 彎曲 (bend) 裂紋 (crack) 不對心 (misalignment) 磨損 (rub) 	<ul style="list-style-type: none"> 振動訊號資訊充足 缺陷發生時能由轉速的諧波頻率明顯看出 	<ul style="list-style-type: none"> 振動 	<ul style="list-style-type: none"> 時域統計量 振動頻率特性 機構系統的特性
幫浦 (pump)	<ul style="list-style-type: none"> 閥門衝擊 (valve impact) 活塞撞擊 (piston slap) 液壓問題 (hydraulic problem) 	<ul style="list-style-type: none"> 可能的脈衝源相當複雜 非線性與時變行為 	<ul style="list-style-type: none"> 振動 聲發射 壓力 	<ul style="list-style-type: none"> 時域統計量 振動頻率特性 應力波與脈衝的特性
交流發電機 (alternator)	<ul style="list-style-type: none"> 定子故障 (stator faults) 轉子電氣故障 (rotor electrical faults) 轉子機械故障 (rotor mechanical faults) 	<ul style="list-style-type: none"> 電壓電流有利於非破壞性且低成本的收集 	<ul style="list-style-type: none"> 定子 振動 磁場 電流電壓 	<ul style="list-style-type: none"> 時域統計量 振動頻率特性 特定諧波頻率成分

□ 第一階段健康評估

- **狀態監控**是對機台的參數(或數據轉換成特徵)進行監控，經由比較特徵在現況下與標準或穩定的情況差異為何
- **健康指標**是綜合所有特徵的狀態整合出一個單一指標對整體設備系統完整的評估，從而理解老化衰退階段為何，以利**故障診斷與預測**的建模。

□ 狀態監控

- **狀態監控**是用以監控此刻機台參數或重要特徵是否異常發生。一方面可採用統計思維的**統計品質管制** (statistical process control, SPC)，也就是假設參數變異服從某個分配，建構上下管制界線來偵測異常；抑或是採用機器學習思維的**非監督式學習**，透過分群進行離群值偵測，或以一分類學習來建構**健康基底** (healthy baseline) 。

□ 健康指標

- **健康指標**牽涉的是**融合** (fusion) 思維，也就是進行**整合**。其中融合又可被分為三種層級，分別是「**數據層級融合**」 (data-level fusion) 、「**特徵層級融合**」 (feature-level fusion) 以及「**決策層級融合**」 (decision-level fusion) (Liu et al., 2013) 。

□ 健康指標

- 數據層級融合是將多個感測器收集的數據整合成單一特徵
- 特徵層級融合需掌握衰退機制與物理特性的先驗知識從而創造特徵
- 決策層級融合則是依據特徵與衰退過程分析建立相對應的決策
- 後兩者的融合依賴於特徵的領域知識以及預處理方式，若沒有足夠的背景知識，可考慮以**數據層級融合**是最直接且可簡化後續複雜度的方法。
- 健康指標三特性
 - 單調性 (monotonicity) 指的是衰退的過程必然是一個不可逆的現象
 - 相關性 (correlation) 指的是衰退過程與時間的相關性
 - 穩健性 (robustness) 指的是(1)特徵包含變異與雜訊的程度，過高的雜訊則使得該特徵的品質較差；(2)該指標放到其他相似機台上也適用/通用

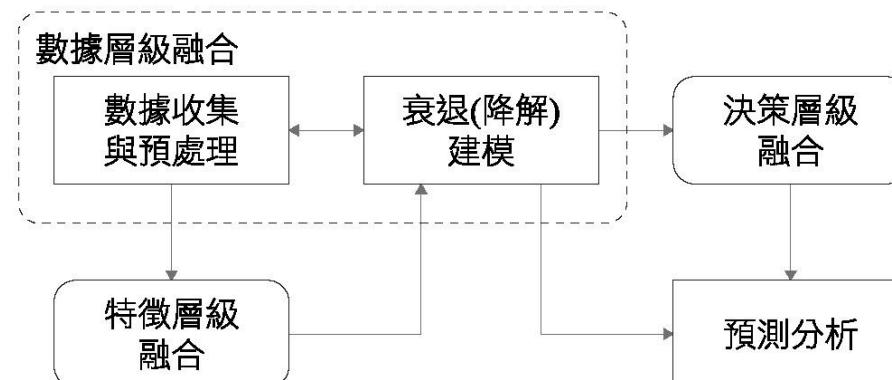


圖 13.8 數據層級融合、特徵層級融合與決策層級融合 (Liu et al., 2013)

□ 健康指標

- 數據層級融合主要有兩種形式

- 兩階段的方式

- 事前評估特徵的衰退特性(上述的三個特性)挑選出重要的特徵，再將這些特徵以權重的方式整合成單一特徵，而權重可由個特徵與正常健康狀態平均值距離的倒數設定(可以迭代的方式逐步更新權重)，或以假設服從多變量常態分配的霍特林T平方(Hotelling's T^2)的統計量(常用於多變量管制圖考慮因子間的相關性)

- 一階段的方式

- 融合時同步考量衰退特性以**數學規劃**的形式求解權重，並將衰退特性列於目標式或限制式中考量(Liu et al., 2013)。

Equipment Health Index (EHI)

❑ Hotelling's T-squared distribution (HTI)

- Multivariate Control Chart (statistical-based)
- Mahalanobis distance (馬氏距離)

$$t^2 = (\bar{\mathbf{x}} - \boldsymbol{\mu})' \hat{\Sigma}_{\bar{\mathbf{x}}}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu})$$

Also, from the [distribution](#),

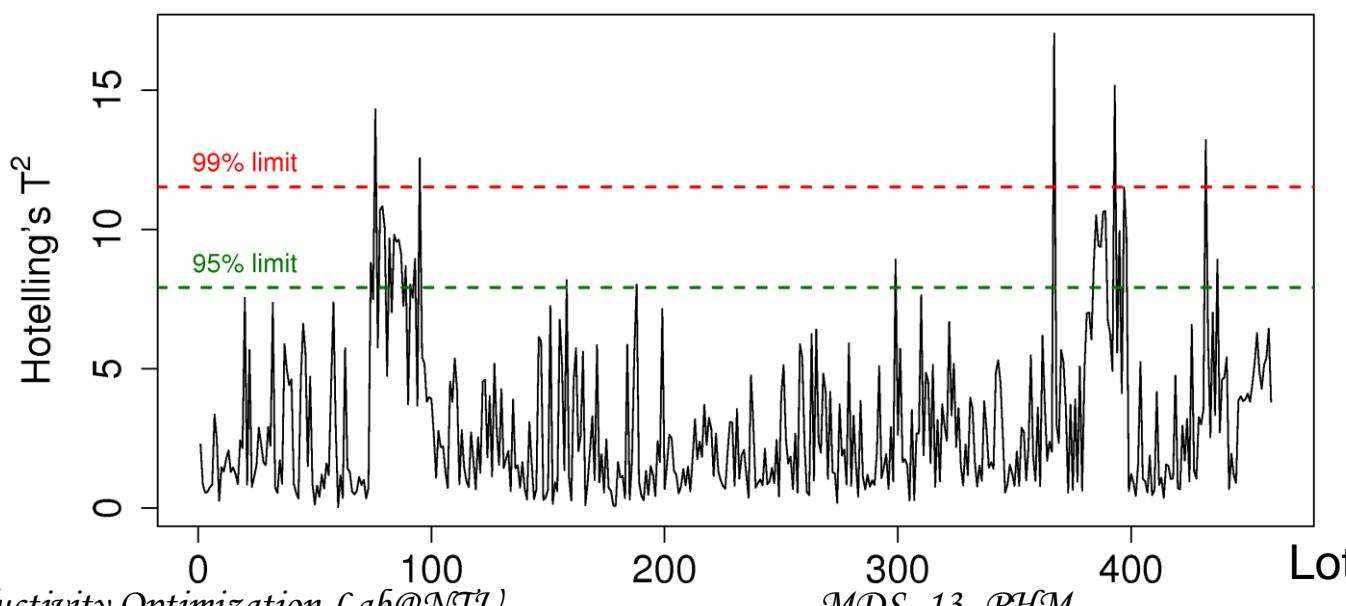
$$t^2 \sim T_{p,n-1}^2 = \frac{p(n-1)}{n-p} F_{p,n-p},$$

where $F_{p,n-p}$ is the [F-distribution](#) with parameters p and $n - p$.

Sample covariance

$$\hat{\Sigma} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'$$

The sample covariance matrix of the mean $\hat{\Sigma}_{\bar{\mathbf{x}}} = \hat{\Sigma}/n$



<https://learnche.org/pid/latent-variable-modelling/principal-component-analysis/hotellings-t2-statistic>

□ 第二階段故障診斷

- 實際是傳統機械工程領域中常見的議題，特別在轉動機械的診斷發展相當完善，其中常以振動訊號作為主要特徵，這是由於它直接反映這些轉動機械中元件的狀態。典型的分析手法也相當全面多元包含了
 - 「頻譜分析」 (spectral analysis)
 - 「波形分析」 (waveform analysis)
 - 「相位分析」 (phase analysis)
 - 「共振分析」 (resonance analysis)
- 正如同前述的**特徵工程**將這些頻域特徵萃取出來，這些頻域特徵可作為**根本原因分析**(root cause analysis)的診斷手法。
- 實際上，現今的的故障診斷則進一步由數據科學的**分類模型**與**特徵挑選**找尋關鍵因子來找出故障的種類與原因，不僅能將前述對頻域萃取出的特徵納入模型外，更能考慮與其他特徵複雜的非線性與交互作用關係。

□ 故障診斷

● 以滾動軸承為例

- 滾動軸承主要由內環、外環、滾珠以及保持器(retainer)所組成，四個元件各有其特有頻率，如表所示，可透過計算公式藉由觀察這些特定頻率或倍頻(又稱諧波(harmonic))關係的變化來對滾動軸承進行診斷
- 其中RPM為軸的轉速、N為軸承滾珠數目、Pd為軸承節徑(也就是(內徑+外徑/2))、Bd為軸承滾珠直徑、 ψ 為滾珠之接觸角。
- 要留意的是，這些公式是理論上推導的結果，實際狀況還需根據損壞程度、負載變化、工況環境等而有所差異。

表 13.3 軸承元件損壞的對應頻率

軸承損壞的元件	對應的頻率計算公式
內環 (ball pass frequency inner race, BPFI)	$1/2 \times RPM \times N \times (1 - Bd / Pd \times \cos \psi)$
外環 (ball pass frequency outer race, BPFO)	$1/2 \times RPM \times N \times (1 + Bd / Pd \times \cos \psi)$
滾珠 (ball spin frequency, BSF)	$1/2 \times RPM \times Pd/Bd \times [1 - (Bd / Pd \times \cos \psi)^2]$
保持器 (fundamental train frequency, FTF)	$1/2 \times RPM \times (1 \times Bd / Pd \times \cos \psi)$

□ 第三階段故障預測

- 又稱為剩餘可用壽命預測(RUL prediction)
- 由於剩餘可用壽命是時間單位為連續的目標值，因而故障預測所使用的是迴歸方法。
- 其方法主要可分為三種模型(Lei et al., 2018)
 - 「物理模型方法」(physics model approaches)
 - 「數據驅動方法」(data-driven approaches)
 - 「混合方法」(hybrid approaches)

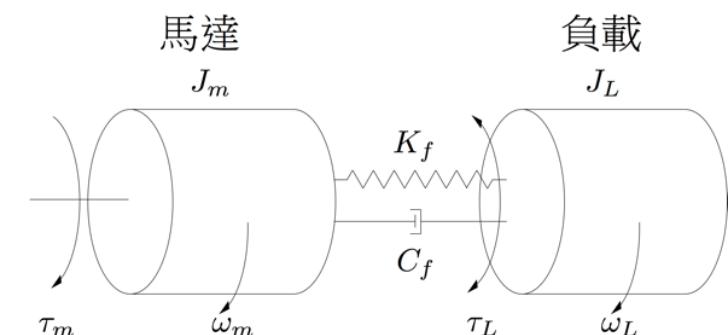
□ 物理模型方法

- 指的是將機台或元件的衰退過程基於物理特性（例如材料特性）所建構的數學模型。這類模型通常需對機構交互的物理特性有全面的理解，透過特定的實驗與分析技術，例如：「掃頻實驗」（frequency sweep experiment）或「有限元素分析」（finite element analysis），來精確地估計這些參數。
- 另一方面，同時瞭解由於現場工況持續改變，對於這些參數需要更新或重新估計，常見的參數更新方法包括「貝氏更新」(Bayesian update)、「卡爾曼濾波器」(Kalman filter)以及「粒子濾波器」(particle filter)等。
- 雖然物理模型方法對於複雜機械系統的應用有受到相當大的限制(例如先驗的物理知識、特定實驗的成本、參數更新的成本、複雜物理模型推導不易等)，然而這些估計出的物理參數是具備實際物理意義的，因此在模型解釋能力上有較突出的表現。

□ 物理模型方法

- 舉例一個「雙質點撓性系統」，如圖所示(Hung et al., 2021)，可透過以下方式來估計其馬達與負載之間的剛性與阻尼，以瞭解機構的衰退狀態
 - 剛性材料結構抵抗應力對其產生變形的能力，跟材料特性有關，參數 K_f 表示
 - 阻尼是振動系統中受外界作用(例如摩擦力)或系統本身固有阻滯而使振動幅度(能量)隨時間下降的物理現象，與速度呈正比，其參數以 C_f 表示
 - 令 τ_m 為馬達轉矩(motor torque)、 ω_m 為馬達速度(speed)、 J_m 為馬達慣量(inertia)、 τ_L 為負載轉矩、 ω_L 為負載速度、及 J_L 為負載慣量，如公式所示呈現各參數之間的非線性關係。

$$\left\{ \begin{array}{l} \frac{d}{dt}\omega_m = \frac{1}{J_m} \cdot (\tau_m - \tau_L) \\ \frac{d}{dt}\omega_L = \frac{1}{J_L} \cdot (\tau_L) \\ \frac{d}{dt}\tau_L = K_f \cdot (\omega_m - \omega_L) + C_f \cdot \frac{d}{dt}(\omega_m - \omega_L) \end{array} \right.$$



- 一般來說，此系統被安裝後會由掃頻進行初始狀態的「系統鑑別」(system identification)，估計出上述參數，經由長時間的運作後，可藉由再一次離線(off-line)的系統鑑別或線上(on-line)的參數估計方法(例如元啟發式演算法)來估計這些係數，萃取出當下具有解釋力的機構物理特性，提供機構衰退(老化)程度的相關資訊。

故障預測

□ 數據驅動方法

- 指的是基於數據先前的狀態去預測未來可能的狀態或與歷史相似樣型以估計剩餘可用壽命，其中又可分為統計模型方法與機器學習方法。

□ 統計模型方法

- 又被稱為「實證模型方法」(empirical model-based approaches)，是基於實證性知識以特定的統計模型或分配進行剩餘可用壽命的估計。此類方法通常是將數據配適於具有隨機性的「隨機係數模型」(random coefficient model)或是「隨機過程模型」(stochastic process model)（一般不涉及物理模型相關知識），這樣的估計可伴隨著信賴區間的建構來描述對故障預測的不確定性。
- 常見的方法包含
 - 「維納過程模型」(Wiener process model)
 - 「伽瑪過程模型」(gamma process model)
 - 「逆高斯過程模型」(inverse Gaussian process model)
 - 「隱藏式馬可夫模型」(hidden Markov model, HMM)
 - 「比例風險模型」(proportional hazards model)

□ 維納過程模型

- 若令某一量測(例如振動訊號的振幅) $\{X(t), t \geq 0\}$ 是個連續隨機過程滿足 $X(0) = 0$ ，其中 t 是時間。若任何兩個互斥(disjoint)的時間區間呈現**穩定獨立增加量** $\Delta X(t) = X(t + \Delta t) - X(t)$ 遵循常態分配 $\Delta X(t) \sim N(\mu \Delta t, \sigma^2 \Delta t)$ ，其中 Δt 是時間區間，則**維納過程模型**定義如公式
 - $X(t) = \mu t + \sigma B(t)$
- 其中 μ 是飄移(drift)參數， $\sigma > 0$ 是擴散(diffusion)參數， $B(\cdot)$ 是**標準布朗運動**(Brownian motion)函數。接著我們令 θ 代表失效閾值(threshold)， ξ 為設備或元件第一次到達失效閾值的壽命，則可定義壽命 $\xi = \inf\{t | X(t) \geq \theta\}$ 為一個**逆高斯分布**(inverse Gaussian distribution)的隨機變數，其**機率密度函數與累積分布函數**如公式所示
 - pdf: $f(t; \mu, \sigma, \theta) = \frac{l}{\sqrt{2\pi\sigma^2 t^3}} \exp\left(-\frac{(\theta-\mu t)^2}{2\sigma^2 t}\right)$
 - cdf: $F(t; \mu, \sigma, \theta) = \Phi\left(\frac{\mu t - \theta}{\sigma\sqrt{t}}\right) + \exp\left(\frac{2\mu\theta}{\sigma^2}\right) \Phi\left(-\frac{\mu t + \theta}{\sigma\sqrt{t}}\right)$
- 其中 $\Phi(\cdot)$ 為標準常態累積分布函數。若可收集衰退數據 $X(t_k) = x_k$ ，可表示**RUL**在時間點 t_k 其剩餘壽命為 ξ ，則條件累積分布函數如公式
 - $F_\xi(t | X_1, X_2, \dots, X_k) = P(\xi \leq t | X_1, X_2, \dots, X_k) = P(X(t + t_k) - X(t_k) \geq \theta - X(t_k))$

□ 維納過程模型

- 根據逆高斯分布，可推得壽命的期望值與變異數為 $E[\xi] = \frac{\theta}{\mu}$ ； $Var[\xi] = \frac{\theta\sigma^2}{\mu^3}$ 。為了估計維納過程的參數 μ 與 σ^2 ，常用方法為**最大概似估計(MLE)**
- 令量測值 X_{ij} 代表一個設備或元件保養更換後的第 i 次衰退過程於第 j 個量測樣本點。 $\Delta X_{ij} = X_{i(j+1)} - X_{ij}$ 表示在時間間隔 $\Delta t_{ij} = t_{i(j+1)} - t_{ij}$ 的衰退增加量，其中 $i = 1, \dots, n$ ， $j = 1, \dots, m_i$ 。為簡化分析，我們固定相等時間間隔 $\Delta t_{ij} = \Delta t_i$ 。因此，維納過程的獨立增加量可以表示為 $\Delta X_{ij} \sim N(\mu_i \Delta t_i, \sigma_i^2 \Delta t_i)$ ，其概似函數 $L(\cdot)$ 可表示如公式所示。

$$- L(\mu_i, \sigma_i^2) = \prod_{j=1}^{m_i} \frac{l}{\sqrt{2\pi\sigma_i^2 \Delta t_i}} \exp\left(-\frac{(\Delta X_{ij} - \mu_i \Delta t_i)^2}{2\sigma_i^2 \Delta t_i}\right)$$

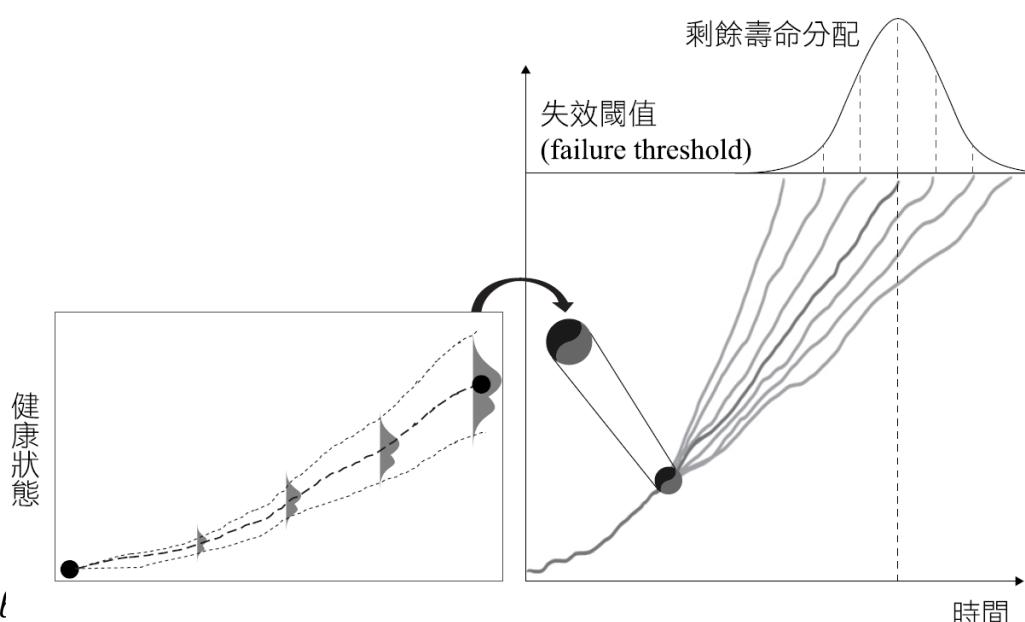
- 因此，對概似函數取一階導數為零(分別對參數 μ 與 σ^2)，再對其作不偏估計量的修正，可推得其 μ_i 與 σ_i^2 的不偏估計式如公式所示

$$- \begin{cases} \hat{\mu}_i = \frac{\sum_{j=1}^{m_i} \Delta X_{ij}}{\sum_{j=1}^{m_i} \Delta t_i} \\ \hat{\sigma}_i^2 = \frac{1}{m_i - 1} \sum_{j=1}^{m_i} \frac{(\Delta X_{ij} - \hat{\mu}_i \Delta t_i)^2}{\Delta t_i} \end{cases}$$

□ 維納過程模型

- 根據 n 次的衰退過程，可推得 μ 與 σ^2 的估計式如公式

$$\begin{cases} \hat{\mu} = \frac{\sum_{i=1}^n \hat{\mu}_i}{n} \\ \hat{\sigma}^2 = \frac{\sum_{i=1}^n \hat{\sigma}_i^2}{n} \end{cases}$$
- 實務上， $X(t)$ 可以是**任何能呈現機構衰退的量測**，例如振動訊號或健康指標，再根據上述的維納過程推導，可以進行壽命模擬或預測。
- 如圖所示(Gouriveau et al., 2016)，考量數據的隨機性後，剩餘壽命呈現一隨機分布(常態分布)，因此也常估計剩餘壽命的**信賴區間**。
- 值得注意的是，維納過程中的**參數**並非在隨時間的推論過程中固定不變，透過新數據的收集，可藉由例如**粒子濾波器(particle filter)**的**貝氏推論**，持續性或週期性的更新參數，以校正維納過程模型的預測。



□ 隱藏式馬可夫模型 (HMM)

- 是馬可夫鏈 (Markov chain) 的延伸。馬可夫鏈描述隨機變數序列 (隨機過程) 的機率，其中以狀態 (state) 來表示我們有興趣且可量測的特性，通常以集合來收集可能發生的狀態，例如設備狀態的集合 ={健康、異常、損壞} 。
- 馬可夫鏈有個強假設，未來的狀態推估只跟當前現在狀態有關，跟過去所處狀態沒有關係。例如預測明天的天氣狀態只跟今天的天氣狀態有關，跟昨天的天氣狀態沒有關聯。這又稱為 **馬可夫特性** (Markovian property) 或 **無記憶**(memoryless) 特性。
- 若一隨機過程 $\{X_t\}$ 滿足馬可夫特性，則我們稱為 **馬可夫鏈** 。令一個狀態變數具有 n 種狀態並其序列为 $\{k_1, k_1, \dots, k_{t-1}, k_t = i, k_{t+1} = j\}$ ，其中 t 為時間，則馬可夫假設可以表示如公式所示。
 - $P\{X_{t+1} = j | X_1 = k_1, X_2 = k_2, \dots, X_{t-1} = k_{t-1}, X_t = i\} = P\{X_{t+1} = j | X_t = i\}$

□ 隱藏式馬可夫模型 (HMM)

- 因此可定義**轉移機率**(transition probability)為 $a_{ij} = P\{X_{t+1} = j|X_t = i\}$ ，代表在時間點 t 由狀態*i*轉移到時間點 $t + 1$ 狀態為*j*的機率，並且離開狀態*i*的所有轉移機率加總為1，也就是 $\sum_{i=1}^n a_{ij} = 1$ 。例如圖描述一個馬可夫鏈，為簡化問題將連續時間的剩餘壽命轉換成離散的設備狀態，用節點表示並分別為{健康、異常、損壞}，實線的邊(arc)代表轉移機率 a_{ij} 。例如設備從這一期為**異常**轉換到下一期為**損壞**的機率為0.2。

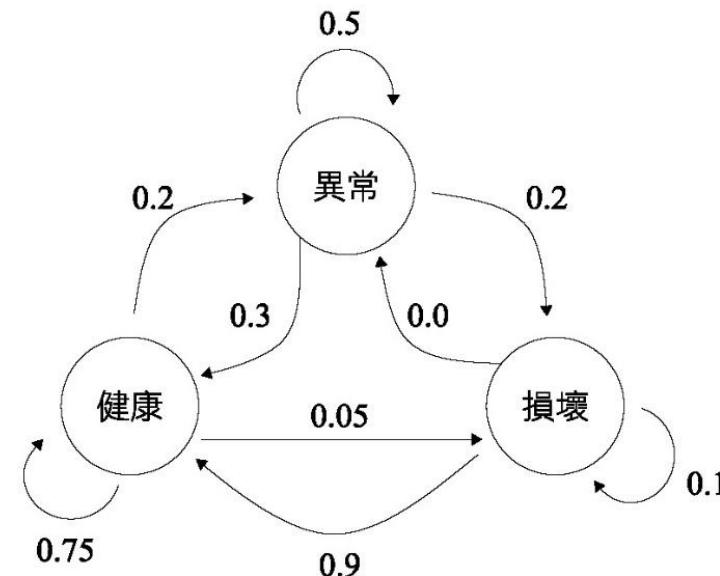
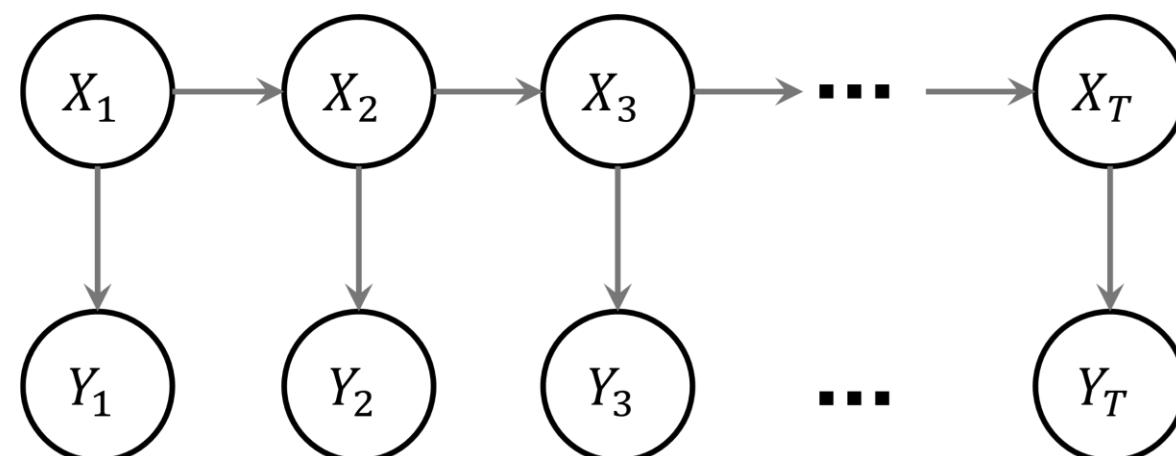


圖 13.11 設備狀態的馬可夫轉移機率圖

□ 隱藏式馬可夫模型 (HMM)

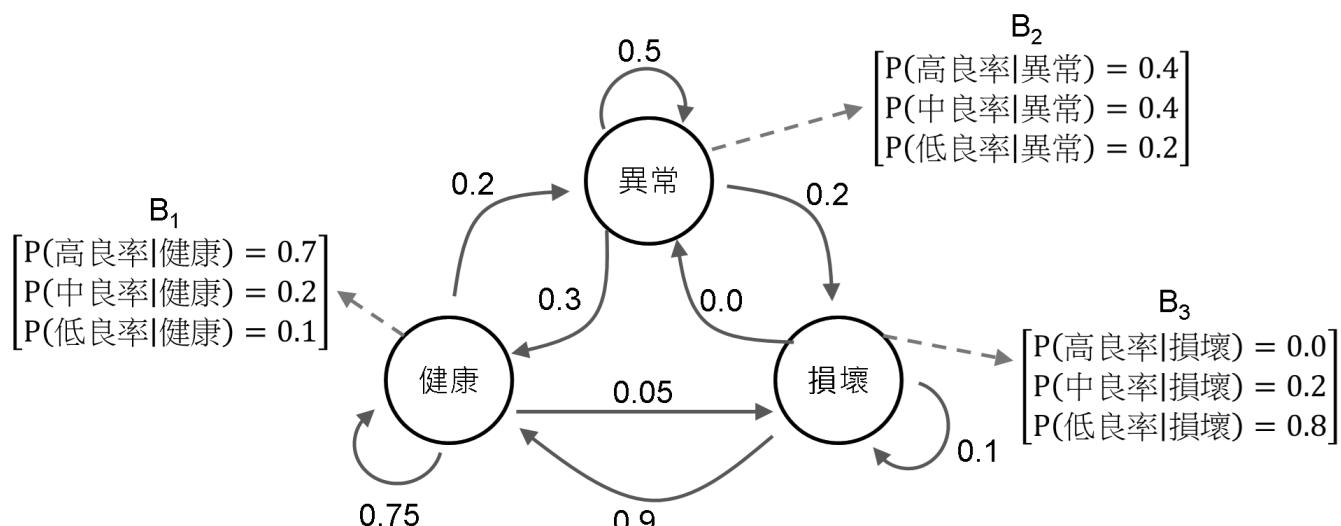
- 當計算可觀測事件之時間序列的機率時，馬可夫鏈是常用的方法之一
 - 狀態對於觀察者來說是直接可見的，狀態的轉移機率便是相關的參數。
 - 但在許多情況下我們無法直接觀察到感興趣的事件(被隱藏(hidden))，但受狀態影響的某些變數則是可見的，透過其它與之相關的資訊來推論。
- 隱藏式馬可夫模型用來描述一個含有隱含未知參數的馬可夫過程
 - 其從可觀察的參數中確定該過程的隱含參數，然後用這些參數作進一步分析
 - 舉例來說，一隨機過程 $\{X_t\}$ 為設備實際的健康狀態，也就是剩餘可用壽命」然而我們無法直接測量，因為它們沒有被觀察到。因此，我們透過可觀測的特徵 $\{Y_t\}$ 來協助推估 $\{X_t\}$ ，如圖所示。隱藏式馬可夫模型使我們可以建立可觀察到(observed)事件和隱藏(hidden)事件，並視為因果因子。



□ 隱藏式馬可夫模型 (HMM)

- 一階(first-order)隱藏式馬可夫模型有兩個假設
 - 馬可夫假設：任一期的隱藏狀態只跟上一期的隱藏狀態有關。
 - 觀測獨立假設(observation independence)：輸出觀測值 Y_t 的機率僅取決於產生該觀測的狀態 X_t ，而跟任何其他狀態或任何其他觀測無關，如公式

$$P\{Y_t|X_1, X_2, \dots, X_T, Y_1, Y_2, \dots, Y_T\} = P\{Y_t|X_t\}$$
- 舉例而言，若以該設備生產產品品質作為觀測對象
 - 令觀測值 $\{Y\}=\{\text{低良率}、\text{中良率}、\text{高良率}\}$ ，在給定觀測序列 $\{Y_t\}$ 下，欲找出導致設備健康狀況造成良率變化的序列 $\{X_t\}$ ，如圖所示，其中 $B_i = b_i(Y_t)$ 為觀測可能性(observation likelihoods)的序列，也稱為排放機率(emission probabilities)或輸出機率，每個表示觀測值 Y_t 從狀態*i* ∈ {X}產生的機率。



□ 隱藏式馬可夫模型 (HMM)

- Rabiner(1989)提出了隱藏式馬可夫模型可使用於三大基本問題的處理
 - 問題1(可能性)：給定HMM參數組合 (A, B) 和一個觀測序列 $\{Y_t\}$ ，則可估計觀測序列的發生機率 $P(Y_t | (A, B))$ 。例如可根據不同保養策略下可能產生的觀測序列，並評估期望效益或損失。
 - 問題2(解碼)：給定HMM參數組合 (A, B) 和一個觀測序列 $\{Y_t\}$ ，找出**最佳的隱藏狀態序列** $\{X_t\}$ 。例如根據良率推估最可能的設備健康狀態，根據價格波動來推估最可能的產地原料供應狀況等。
 - 問題3(學習)：給定一個觀測序列 $\{Y_t\}$ 和狀態集合，學習HMM的參數組合 (A, B) 使得該**序列發生的條件機率最大**。例如已知良率的時間序列與設備狀態，估計健康狀態衰退到異常狀態的機率為何(也就是找出圖的轉移機率)。
- 在對於求解此三大問題上，問題1常用「前向演算法」(forward algorithm)求解；問題2常用「維特比演算法」(Viterbi algorithm)；問題3常用「前向後向演算法」(forward-backward algorithm)。

□ 機器學習方法

- 以監督式學習方法來學習機台或元件衰退（降解degradation）的樣型
 - 非物理模型或統計模型的建構。此類方法能在複雜機械系統中有效地進行故障預測，可建構出複雜的非線性降解關係，而這是典型的物理模型或統計模型不易做到的，常用方法包含支援向量迴歸、梯度提升機及深度學習等。
- 為了訓練機器學習模型
 - 大量的數據與標籤是必要的
 - 除了數據品質與預處理方式顯著影響預測績效
 - 特徵工程轉換出來新特徵也可能提供潛在有用的資訊來提升預測準確度
- 特徵工程可使用滑動時窗選取資料範圍並從不同的視角來萃取相關特徵
 - 例如從時域、時域轉折點(可用線性分段(linear segmentation)找出轉折點)、頻域、時頻域等，來萃取相關的特徵 (參閱「**特徵工程、數據增強與數據平衡**」)
 - **時域**：最大值、平均值、標準差、偏態、峰態、均方根、線性迴歸的斜率係數、截距、均方誤差等。
 - **時域轉折點**：第一轉折時間點出現的位置、所有轉折時間點的標準差、偏態與峰態等。
 - **頻域**：最大振幅的能量與頻率、第二大振幅的能量與頻率、所有振幅的平均值、標準差、偏態、峰態等。
 - 同時可將這些特徵，透過例如**自編碼器(autoencoder)**或**生成對抗網路(GAN)**等非監督式學習方法來萃取非線性特徵。

□ Feature Engineering (Time Domain, Frequency Domain, Time-frequency Domain)

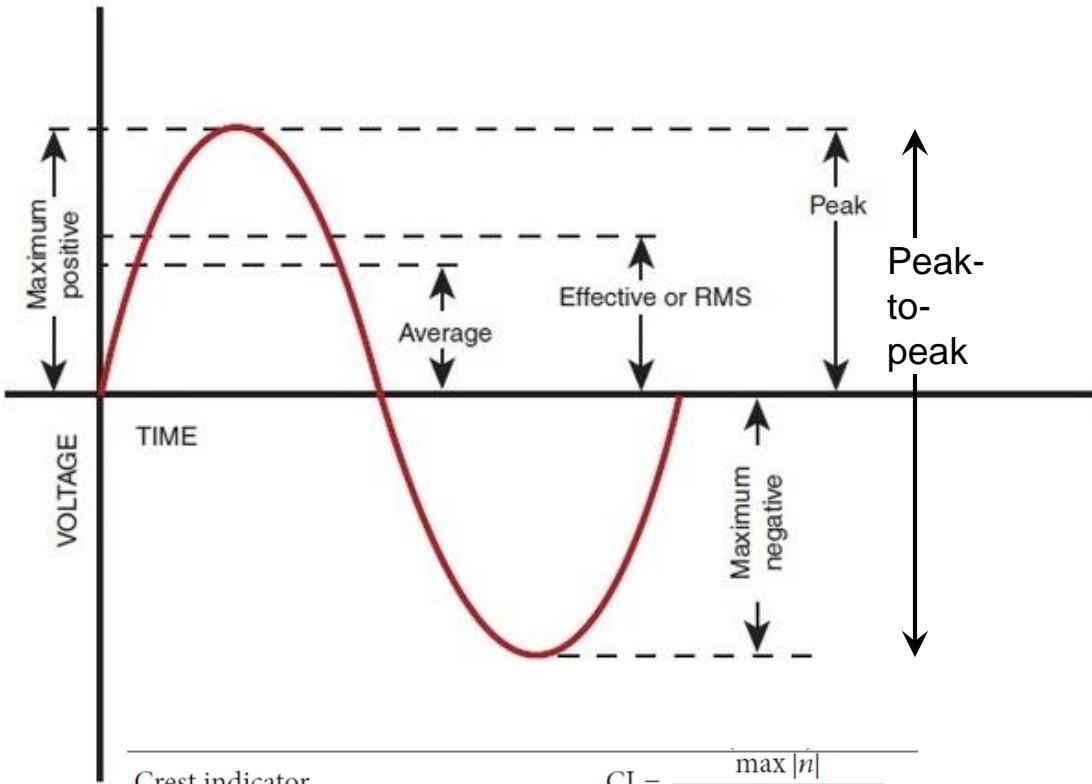
- Root Mean Square (RMS), Mean, Variance, Skewness, Kurtosis, and Crest Factor/Max/Min, ...
- Time synchronous averaged (TSA) signal, **sliding window**, residual signal (RES), difference signal (DIF), ...
- Time series decomposition, **Hilbert vibration decomposition**, empirical mode decomposition (EMD), intrinsic mode functions (IMF), ...
- Regression (intercept, slope), **exponential degradation $\theta \exp(\beta t)$** (amplitude θ , slope β), autoregression, ...

Component	Issue & failure	Characteristic	Common measures	Common features	Common algorithms used
Bearing	Outer-race, inner-race, roller, and cage failures	Raw data does not contain insightful information; low amplitude; high noise	Vibration, oil debris, acoustic emission	Vibration characteristic frequency, time domain statistical characteristics, metallic debris shape, size, quantity, sharp pulses and rate of development of stress-waves propagation	Fourier Transform (FT) [18,19], Short Time Frequency Transform (STFT) [20], Wavelet Transform (WT) [21], Empirical Mode Decomposition (EMD) [22], Bispectrum [23], Autoregression (AR) Frequency Spectra [24], Hilbert Spectrum [25], Instantaneous Power Spectrum [26], Hilbert–Huang Transform (HHT) [27], Neural

Lee, et al. (2014). Prognostics and health management design for rotary machinery systems- Reviews, methodology and applications. Mechanical Systems and Signal Processing, 42, 314-334.

□ Feature Engineering

Feature	Equation
Maximum Value	$f_{1i} = \max(X(i))$
Mean Value	$f_{2i} = \frac{1}{N} \sum_{i=1}^N X(i)$
Minimum Value	$f_{3i} = \min(X(i))$
Standard Value	$f_{4i} = \sqrt{\frac{1}{N} \sum_{i=1}^N (X(i) - f_{1i})^2}$
Peak to Peak Value	$f_{5i} = f_{1i} - f_{3i}$
Mean Amplitude	$f_{6i} = \frac{1}{N} \sum_{i=1}^N X(i) $
Root Mean Square Value	$f_{7i} = \sqrt{\frac{1}{N} \sum_{i=1}^N X(i)^2}$
Skewness Value	$f_{8i} = \frac{1}{N} \sum_{i=1}^N X(i)^3$
Kurtosis Value	$f_{9i} = \frac{1}{N} \sum_{i=1}^N X(i)^4$
Waveform Indicator	$f_{10i} = \frac{f_{7i}}{f_{6i}}$
Pulse Indicator	$f_{11i} = \frac{f_{1i}}{f_{6i}}$
Kurtosis Index	$f_{12i} = \frac{f_{9i}}{f_{7i}}$
Peak Index	$f_{13i} = \frac{f_{1i}}{f_{7i}}$
Square Root Amplitude	$f_{14i} = (\frac{1}{N} \sum_{i=1}^N \sqrt{ X(i) })^2$
Margin Indicator	$f_{15i} = \frac{f_{1i}}{f_{14i}}$
Skewness Indicator	$f_{16i} = \frac{f_{8i}}{f_{7i}}$



Crest indicator

$$CI = \frac{\max |n|}{\sqrt{(1/N) \sum_{j=1}^N (n_j)^2}}$$

Clearance indicator

$$CLI = \frac{\max |n|}{\left((1/N) \sum_{j=1}^N \sqrt{|n_j|} \right)^2}$$

Shape indicator

$$SI = \frac{\sqrt{(1/N) \sum_{j=1}^N (n_j)^2}}{(1/N) \sum_{j=1}^N |n_j|}$$

Impulse indicator

$$MI = \frac{\max |n|}{(1/N) \sum_{j=1}^N |n_j|}$$

W. Jiang, Y. Hong, B. Zhou, X. He, and C. Cheng, “A gan-based anomaly detection approach for imbalanced industrial time series,” IEEE Access, vol. 7, pp. 143 608–143 619, 2019.

<https://www.quora.com/Is-the-DC-voltage-the-same-in-the-RMS-voltage>

□ Feature Extraction

Index	Feature	Sensor	Index	Feature	Sensor
1	RMS		45	RMS	
2	Standard Deviation		46	Standard Deviation	
3	Mean value		47	Mean value	
4	Variation		48	Variation	
5	Skewness		49	Skewness	
6	Kurtosis		50	Kurtosis	Spindle bearing temperature
7	Crest factor		51	Crest factor	
8	Peak value		52	Peak value	
9-16	Fundamental and harmonics		53	RMS	
17-19	BPFO (1-3x) energies		54	Standard Deviation	
20 - 22	BPFI (1-3x) energies		55	Mean value	Ambient temperature
23-25	BSF (1-3x) energies		56	Variation	
26-28	FTF energies		57	Skewness	
29	RMS		58	Kurtosis	
30	Standard Deviation		59	Crest factor	
31	Mean value		60	Peak value	
32	Variation	Current	61	RMS	
33	Skewness		62	Standard Deviation	
34	Kurtosis		63	Mean value	Loading deck
35	Crest factor		64	Variation	temperature
36	Peak value		65	Skewness	
37	RMS		66	Kurtosis	
38	Standard Deviation		67	Crest factor	
39	Mean value	Motor temperature	68	Peak value	
40	Variation				
41	Skewness				
42	Kurtosis				
43	Crest factor				
44	Peak value				

Feature	Definition
Peak value	$\max = \max_{j=1,\dots,N} n_j$
Mean	$u = \frac{1}{N} \sum_{j=1}^N n_j$
Standard deviation	$\sigma = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (n_j - u)^2}$
Root mean square	$\text{RMS} = \sqrt{\frac{1}{N} \sum_{j=1}^N (n_j)^2}$
Skewness	$SK = \frac{\sum_{j=1}^N (n_j - u)^3}{(N-1)\sigma^3}$
Kurtosis	$KU = \frac{\sum_{j=1}^N (n_j - u)^4}{(N-1)\sigma^4}$
Crest indicator	$CI = \frac{\max n }{\sqrt{(1/N) \sum_{j=1}^N (n_j)^2}}$
Clearance indicator	$CLI = \frac{\max n }{\left((1/N) \sum_{j=1}^N \sqrt{ n_j } \right)^2}$
Shape indicator	$SI = \frac{\sqrt{(1/N) \sum_{j=1}^N (n_j)^2}}{(1/N) \sum_{j=1}^N n_j }$
Impulse indicator	$MI = \frac{\max n }{(1/N) \sum_{j=1}^N n_j }$
SD of asinh(X)	$\sigma \left(\log \left[x_j + (x_j^2 + 1)^{1/2} \right] \right)$
SD of atan(X)	$\sigma \left(\frac{i}{2} \log \left(\frac{i+x_j}{i-x_j} \right) \right)$
Entropy (threshold)	$E = \sum_{j=1}^n E(x_j)$
Energy	$e = \sum_{j=1}^n x_j^2$
Root mean square (RMS)	$RMS = \sqrt{\frac{1}{n} (x_1^2 + \dots + x_n^2)}$
Kurtosis	$\frac{\sum_{j=1}^n (x_j - \bar{X})^4}{(n-1)\sigma^4}$
Upper bound	$\max(X) + \frac{1}{2} \frac{\max(X) - \min(X)}{n-1}$

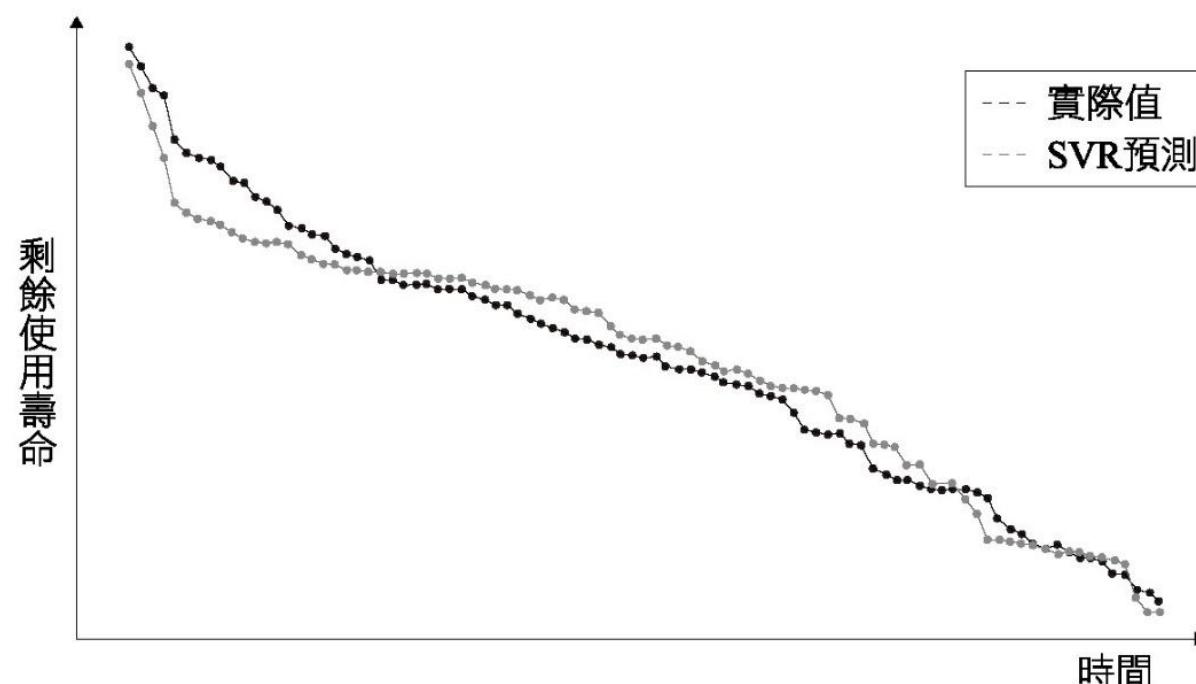
Javed et al. (2015). Enabling health monitoring approach based on vibration data for accurate prognostics. IEEE Transactions on Industrial Electronics, 62(1), 647-656.

Liao, L. (2014). "Discovering prognostic features using genetic programming in remaining useful life prediction," IEEE Transactions on Industrial Electronics, vol. 61, no. 5, pp. 2464-2472.

Tsui et al. (2015). Prognostics and health management: a review on data driven approaches. Mathematical Problems in Engineering, 2015, Article ID 793161, 17 pages.

□ 機器學習方法

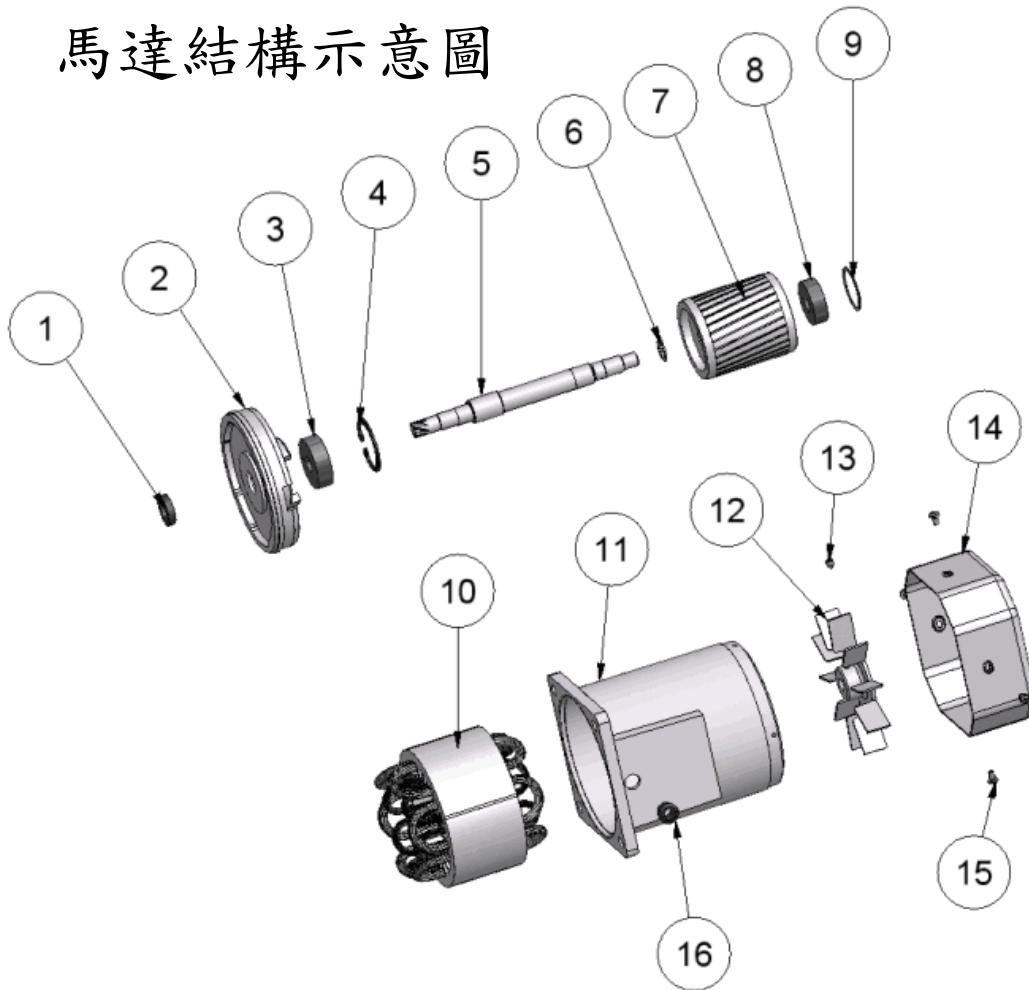
- 最後，透過重要特徵的挑選建立剩餘壽命預測模型。圖呈現用**支援向量迴歸(support vector regression, SVR)**預測剩餘可用壽命，該圖x軸是設備運行的週期取樣時間點(每個週期取樣20秒的訊號數據並轉為特徵)，y軸是剩餘可用壽命。



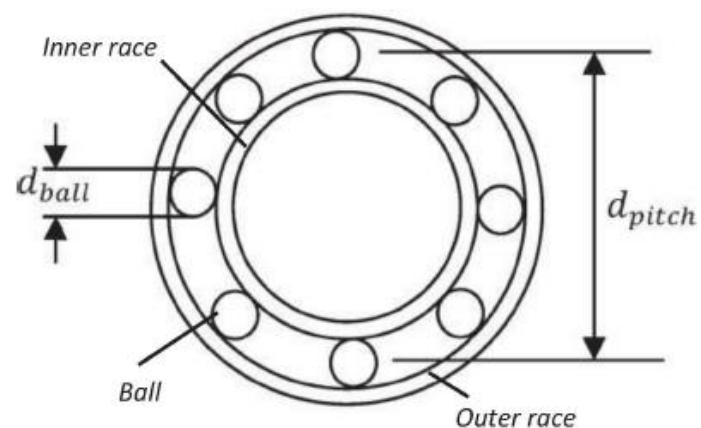
13.14 機器學習的剩餘壽命預測圖

□ 馬達軸承壽命預測 Remaining Useful Life (RUL) Prediction

馬達結構示意圖



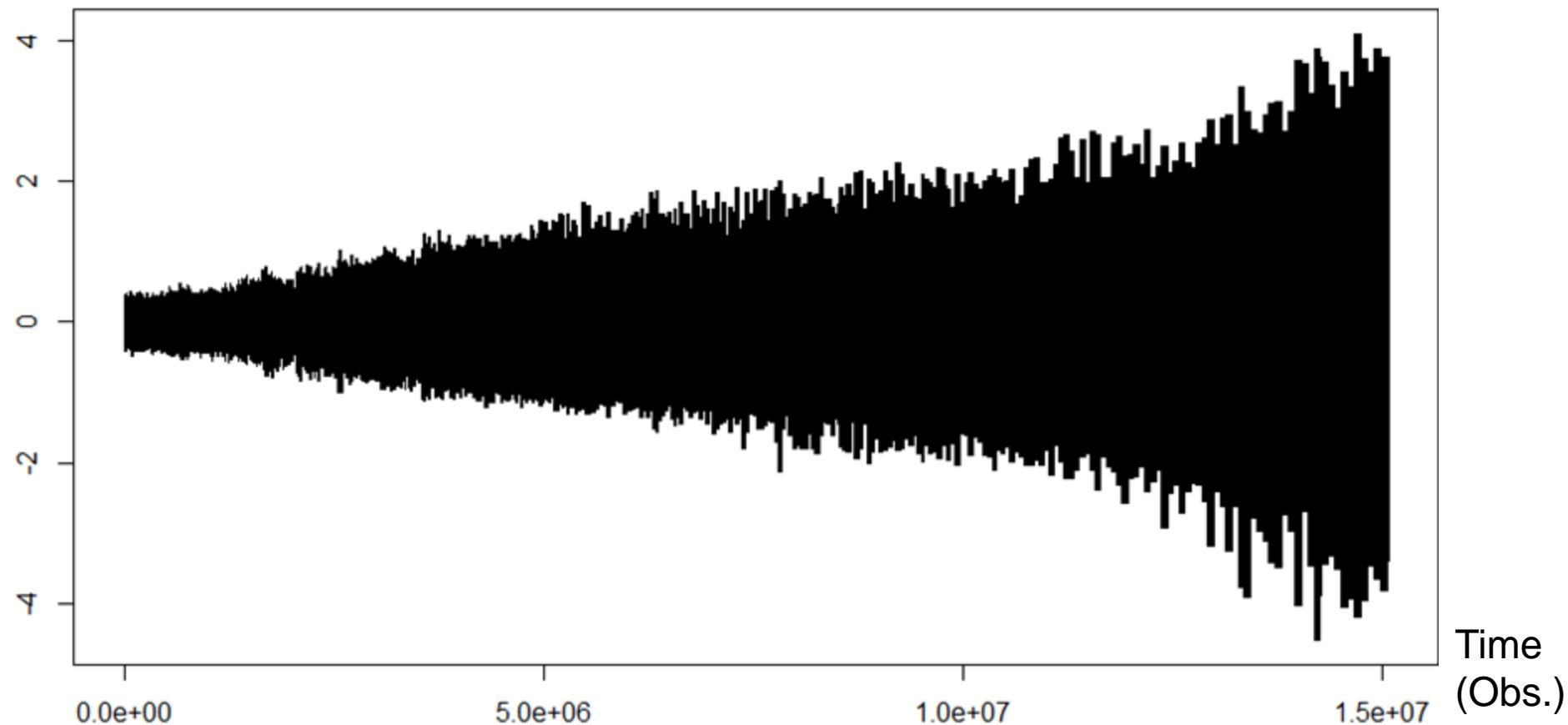
No	名稱	No	名稱
1	油封	9	波型華司
2	馬達前蓋	10	定子線圈
3	滾珠軸承	11	馬達機殼
4	R扣環	12	風扇
5	馬達齒軸	13	固定螺絲
6	S扣環	14	風罩
7	轉子	15	十字頭螺絲
8	滾珠軸承	16	護線扣環



http://www.li-ming.com.tw/index.php?option=com_content&view=article&id=225&Itemid=351&lang=

Vibration Dataset

Amplitude

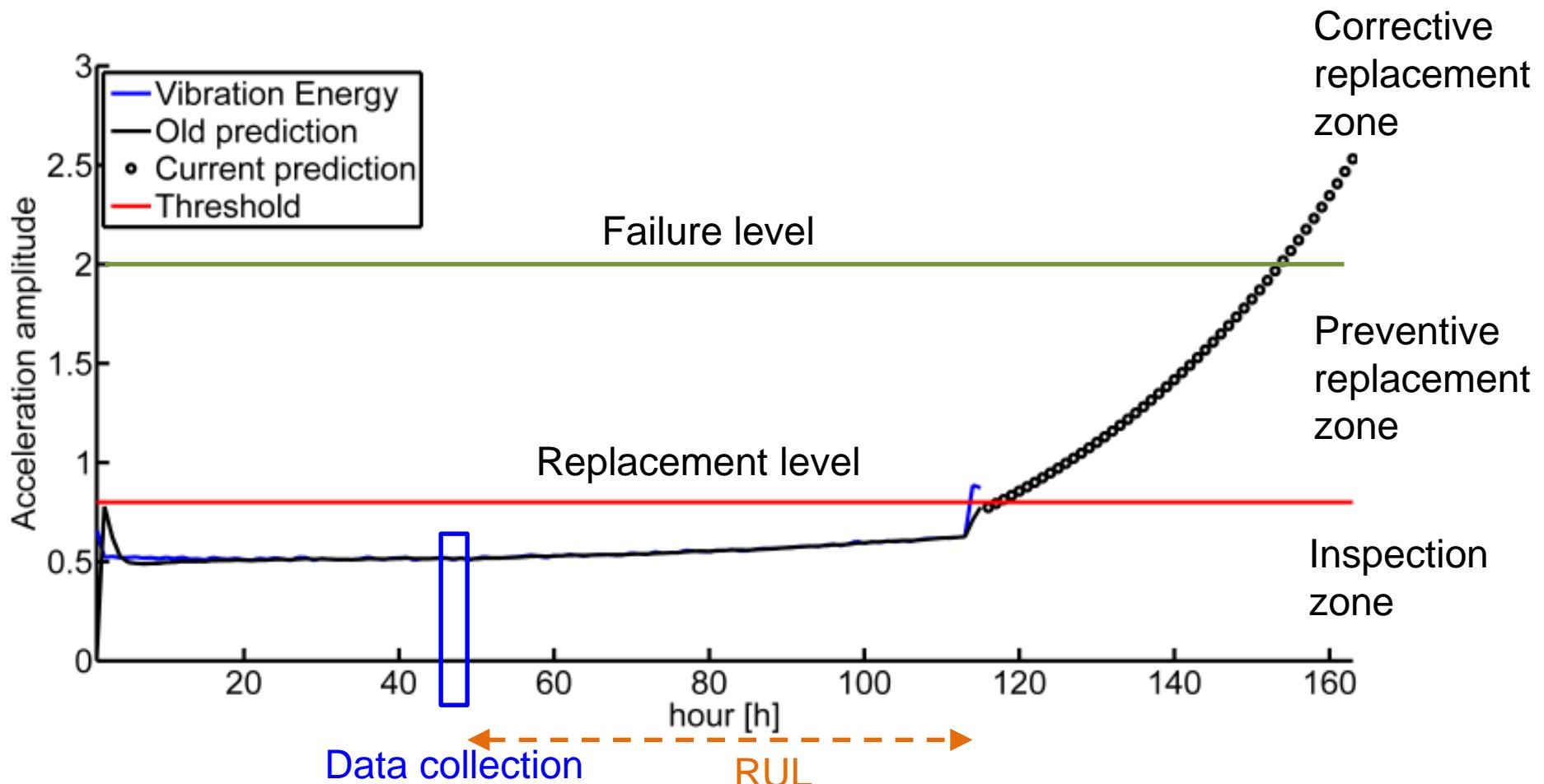


Lee et al. (2019)

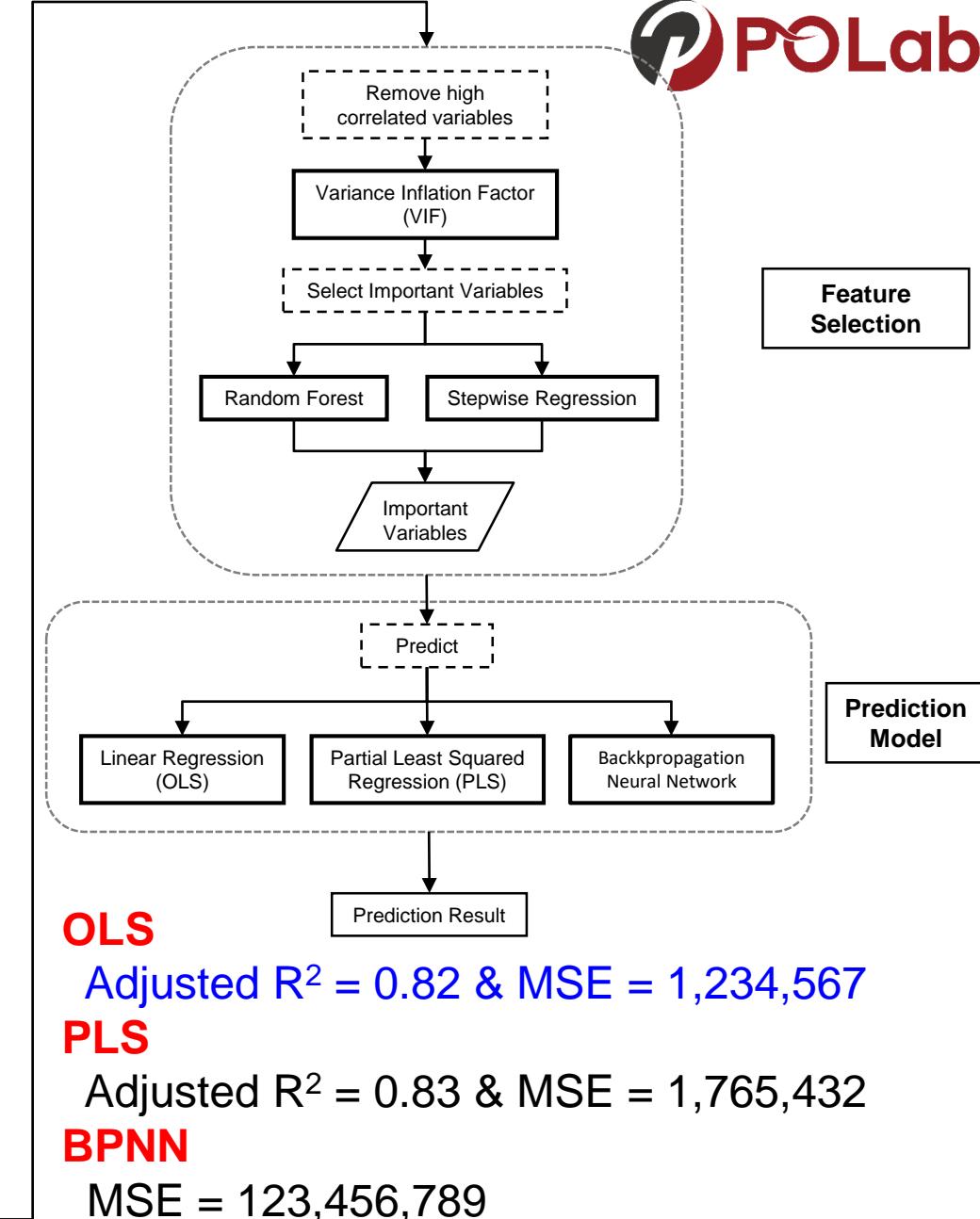
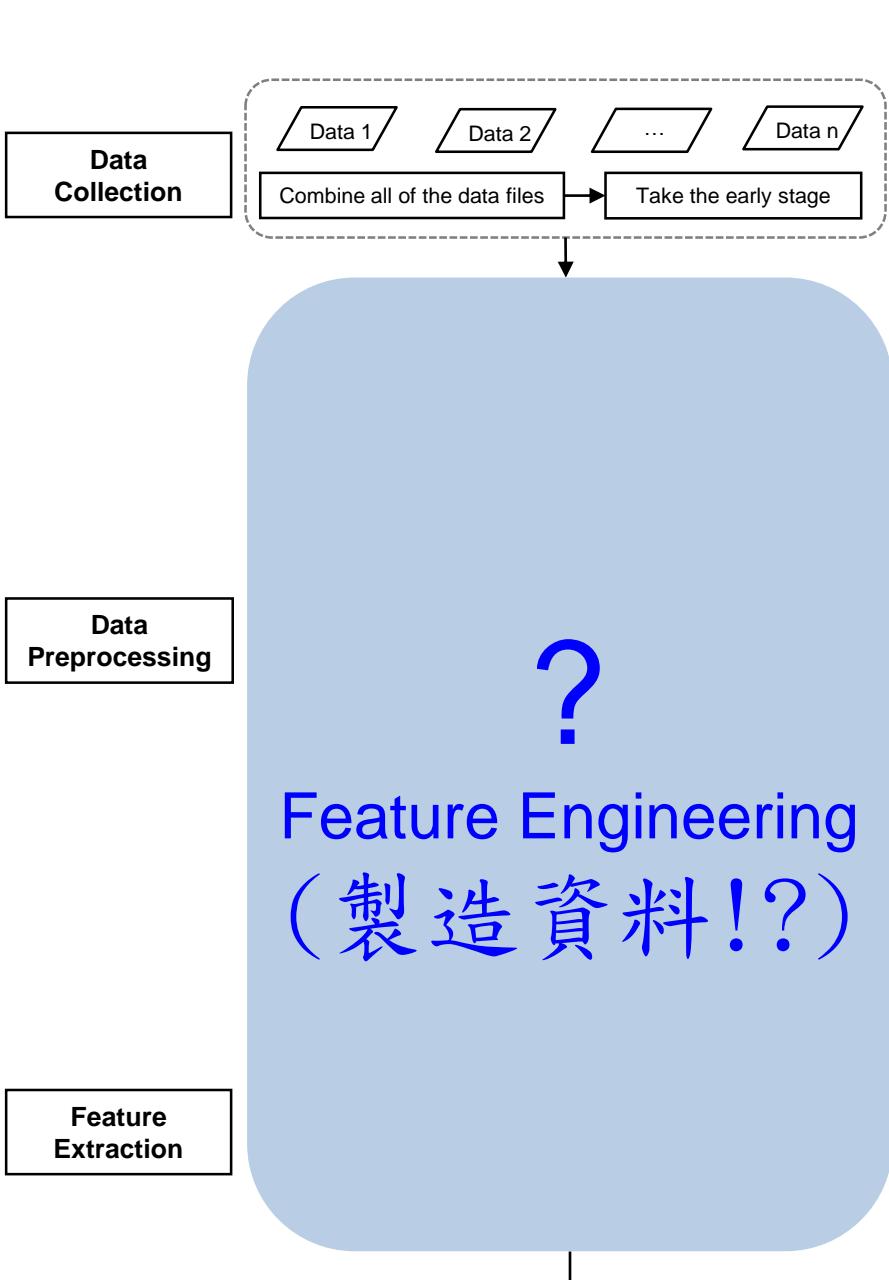
Lee, C.-Y., T.-S. Huang, M.-K. Liu, and C.-Y. Lan. 2019. Data Science for Vibration Heteroscedasticity and Predictive Maintenance of Rotary Bearings. *Energies.*, 12 (5), 801.

□ Prediction of the Deterioration

- Vibration acceleration signal



revised from Grall et al. (2002) and Rocchi et al. (2014)



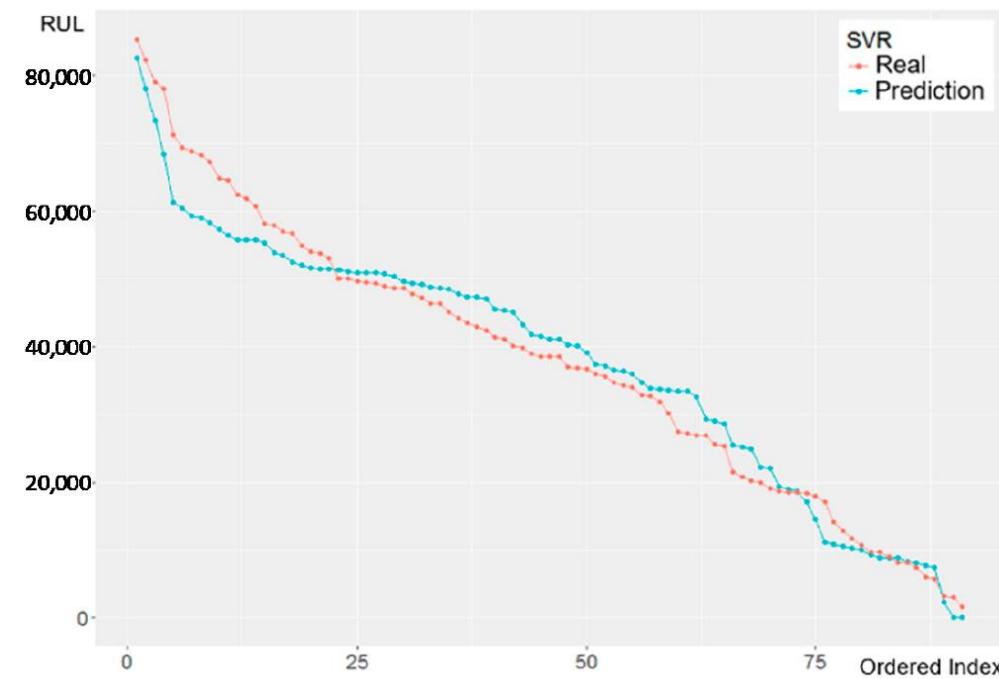
Lee, C.-Y., T.-S. Huang, M.-K. Liu, and C.-Y. Lan. 2019. Data Science for Vibration Heteroscedasticity and Predictive Maintenance of Rotary Bearings. *Energies.*, 12 (5), 801.

Method/ Dimensions	Time-Series Dimension	Change-Point Dimension	Frequency Dimension
Method	OLS	Piecewise linear segmentation	FFT
Features	(1) Mse.ts (2) Slope.ts (3) Intercept.ts (4) Skewness.ts (5) Kurtosis.ts (6) Max.ts	(7) Sd.cp (8) First-point.cp (9) Skewness.cp (10) Kurtosis.cp	(11) Ampl1.f (12) Ampl1-freq.f (13) Ampl2.f (14) Ampl2-freq.f (15) Ampl-mean.f (16) Ampl-var.f (17) Ampl-skewness.f (18) Ampl-kurtosis.f

Feature Selection

RUL Prediction

VIF Test		Stepwise Regression	
Features	VIF	Features	State
Mse.ts	>10		
Slope.ts		Slope.ts	out
Intercept.ts	>10		
Skewness.ts		Skewness.ts	out
Kurtosis.ts		Kurtosis.ts	in
Max.ts		Max.ts	in
Sd.cp		Sd.cp	out
First-point.cp	>10		
Skewness.cp		Skewness.cp	out
Kurtosis.cp		Kurtosis.cp	out
Ampl1.f	>10		
Ampl1-freq.f		Ampl1-freq.f	in
Ampl2.f		Ampl2.f	in
Ampl2-freq.f		Ampl2-freq.f	in
Ampl-mean.f	>10		
Ampl-var.f		Ampl-var.f	out
Ampl-skewness.f		Ampl-skewness.f	out
Ampl-kurtosis.f	>10		



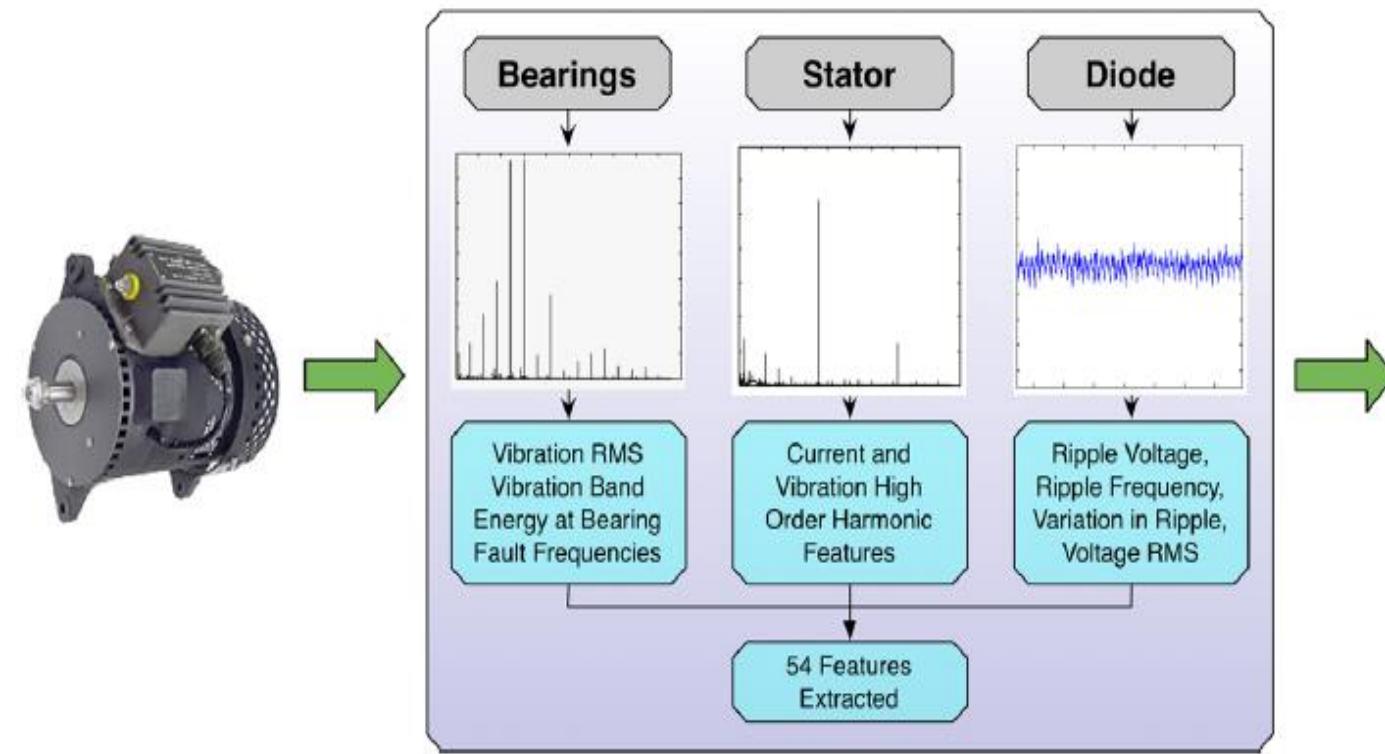
Lee, C.-Y., T.-S. Huang, M.-K. Liu, and C.-Y. Lan. 2019. Data Science for Vibration Heteroscedasticity and Predictive Maintenance of Rotary Bearings. *Energies.*, 12 (5), 801.

Equipment

Sensor Data Collection

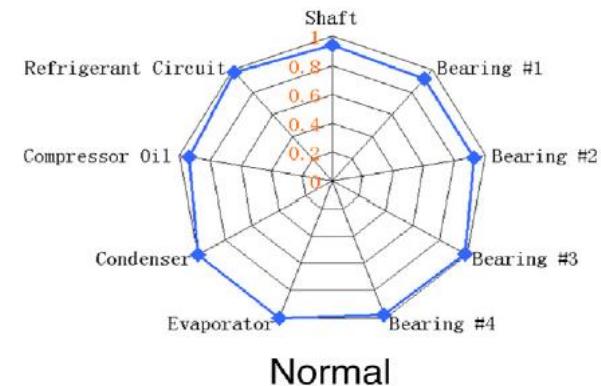
Modelling

Visualization

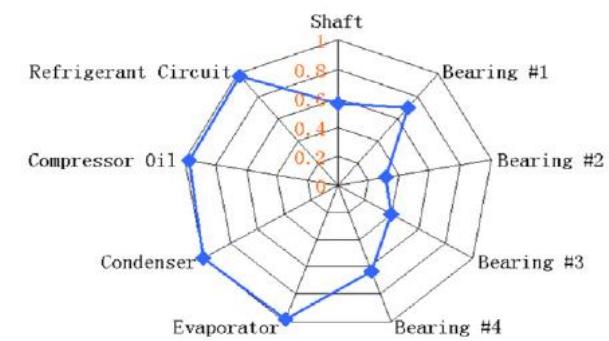


(Lee et al., 2014)

Radar chart for health monitoring



Normal

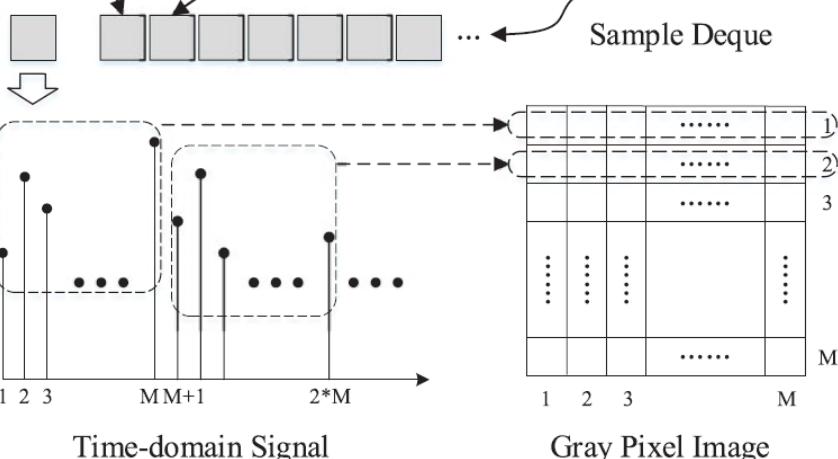
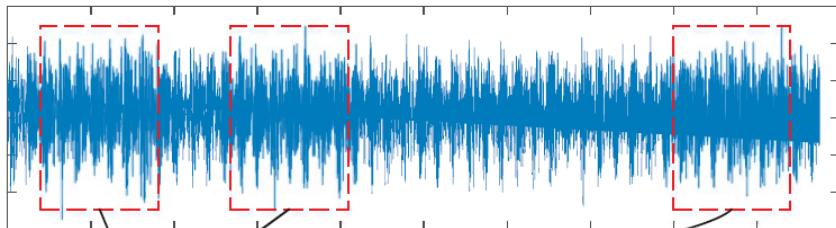


Abnormal

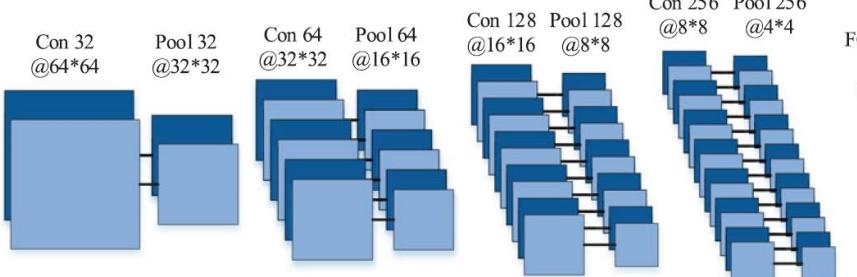
Lee, et al. (2014). Prognostics and health management design for rotary machinery systems- Reviews, methodology and applications. Mechanical Systems and Signal Processing, 42, 314-334.

Convolutional Neural Network (CNN) Fault Diagnosis

Random Segment of Raw Signal

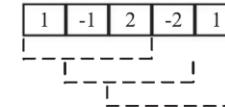


$$P(j, k) = \text{round} \left\{ \frac{L((j-1) \times M + k) - \text{Min}(L)}{\text{Max}(L) - \text{Min}(L)} \times 255 \right\}$$



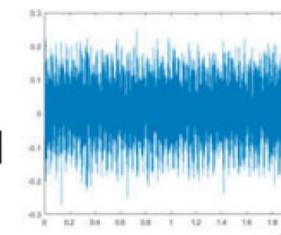
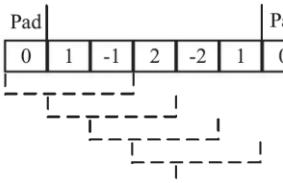
No Padding

Dimension loss:
from 5 to 3

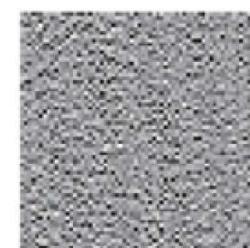


Zero-Padding

No loss on dimension



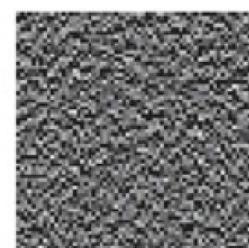
Padding for control the size of feature dimension and the zero-padding for preventing the dimension loss



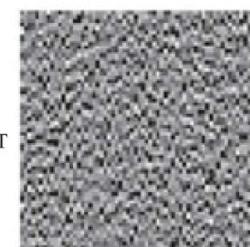
inner race wearing



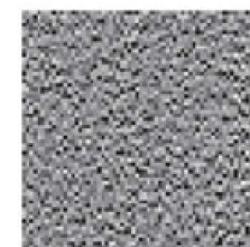
impeller wearing



normal



outer race wearing



bearing roller wearing

**Self-Priming
Centrifugal Pump**

Wen, L., Li, X., Gao, L., and Zhang, Y. 2018. A new convolutional neural network-based data-driven fault diagnosis method. IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, 65 (7), 5990-5998.

□ 混合方法

- 數據驅動方法一方面統計模型能賦予故障預測的隨機性
- 機器學習能更準確地預測複雜的機械系統
- 物理模型方法在於我們可萃取出具備有實質物理意義的物理模型參數
- **混合方法**是基於上述兩種方法各自所具備的使用時機與優勢，建構出更具解釋性與準確性的故障預測模型，其常見兩種方法如圖所示（Gouriveau et al., 2016）。

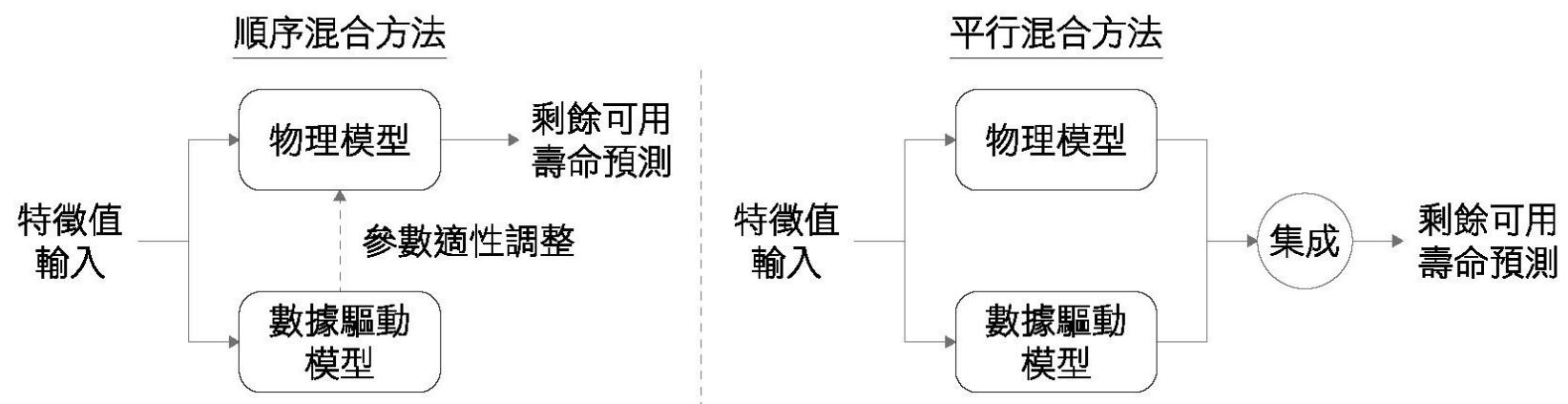


圖 13.15 故障預測的兩種混合方法 (Gouriveau et al., 2016)

□ 現況困難與挑戰

- PHM 面臨的挑戰主要涵蓋四大面向 (Ran et al., 2019)

— 數據質與量

- 然而對於故障數據的收集而言，實務上長時間的數據收集不僅無法快速累積數據量，還會面臨製程技術、工況變更、與產品組合的改變等，數據收集常不穩定（例如長時間的遺漏值）或包含雜訊與噪音。

— 數據不平衡

- 對於製造現場在故障事件發生前就已先事前進行維護與保養，使得真正發生故障的真實數據是十分稀少。另一方面，即使有存留保養相關的數據，往往只有保養後的紀錄，但沒有測量保養前的狀態，因此在建模上能觀察到的數據仍十分偏頗

— 多元件系統 (multicomponent systems)

- 隨著製造技術迅速地發展與增長，製造系統變得越加複雜並涉及更大量的設備與相關機台元件，然而現有的大多數PHM方法僅專注於特定的元件，同時涉及多個元件將增加PHM的交互作用分析與複雜度。

— 數位孿生 (digital twin)

- 數位孿生通常包含一個**模擬系統**（其架構包含多個模型或模組），該系統將不斷更新以反映其真實情形中的孿生狀態，因而該模擬系統不僅須考量原有生產系統外，還需納入PHM系統。這使得我們能取得大量的故障數據，從而不斷改善精進改善數位孿生，朝向智慧製造的虛擬化工廠邁進。

□ 保養定位與策略發展

- 保養能以**複雜度** (complexity) 與**不確定性** (uncertainty) 兩面向探討
 - 複雜度可進一步以「侵入/非侵入式」 (intrusive/nonintrusive) 以及「少/多變量」 (low/high variables) 區分
 - 不確定性可進一步以「靜態/動態」 (static/dynamic) 與「確定性 / 隨機性」 (deterministic/probabilistic) 區分

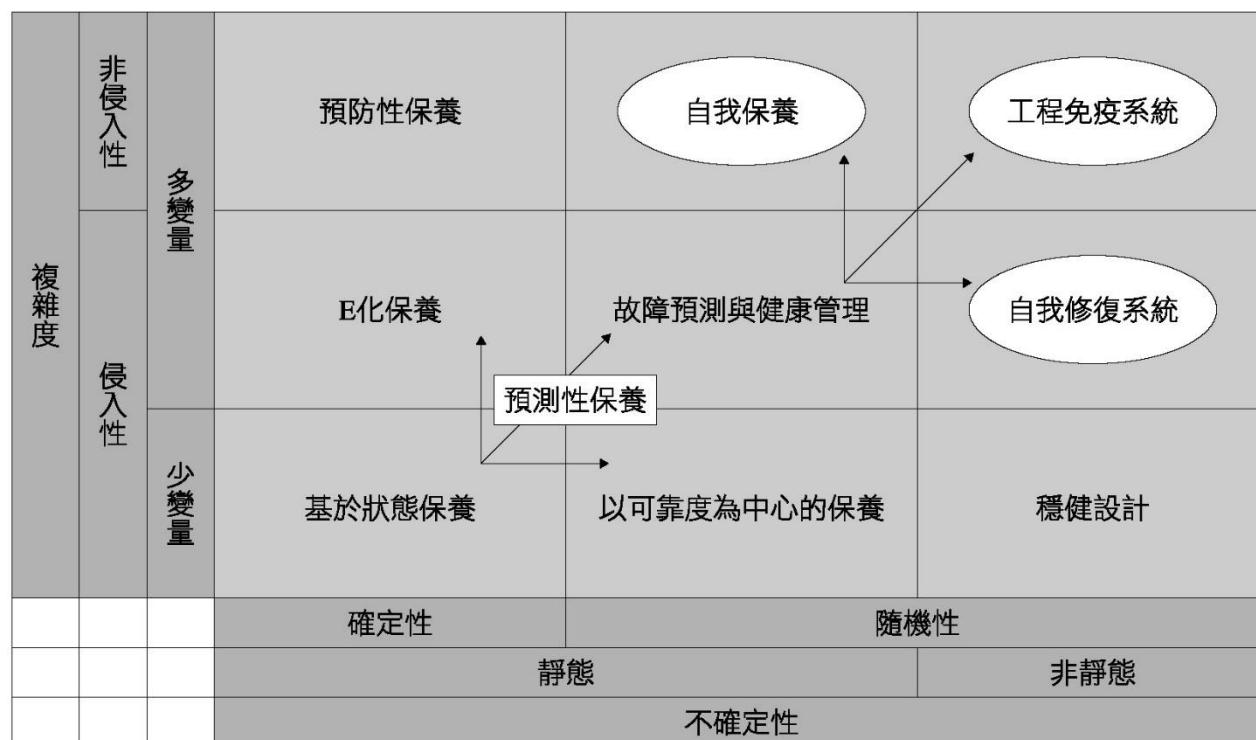


圖 13.16 PHM 保養定位與策略 (Lee et al., 2014)

Thanks for your attention



NTU Dept. of Information Management
name: 李家岩 (FB: Chia-Yen Lee)
phone: 886-2-33661206
email: chiayenlee@ntu.edu.tw
web: <https://polab.im.ntu.edu.tw/>