

“Ensemble-based adaptive learning for remaining useful life (RUL) prediction of equipment components” (應用整體式適應性學習方法於設備元件剩餘壽命之預測)

PI: Guan-Hua Huang (National Chiao Tung University)

Role: Principal investigator

Industrial Technology Research Institute: 12/01/2015-12/31/2016

This project will use some innovative statistical methods, such as nonparametric regression degradation model and ensemble-based adaptive learning, to predict the remaining useful life (RUL) of equipment components.

1. 研究動機

國際半導體技術組織 ISMI 將機台預診斷 (Prognostics and Health Management, PHM) 列為 “Next Generation Factory Realization Roadmap” 中的重要計畫。國際半導體設備大廠 Applied Materials 於 2013 年發表研究報告，針對安裝於美光、英特爾的設備機台以及製程數據進行故障預測的研究，單一零件故障預警可節省 USD\$ 13K/年。機台預診斷中的一個主要任務，是預估零件的剩餘壽命 (Remaining Useful Life, RUL)。RUL 為執行狀態基準維護 (Condition Based Maintenance, CBM) 策略時的重要依據。利用巨量資料 (Big Data) 分析技術，從機台的製造資料中，萃取出與產品品質和機台健康狀態相關的資訊，將其用於 RUL 的估計，以發展可以提升產能與產品品質的機台預診斷技術，同時創造機台使用者與製造者之產業競爭力，將在未來技術發展上能充分發揮與產業界結合及協助產業升級之效益。

2. 國內外技術概況、差異比較

對於現今高可靠度產品而言，往往無法在有限的試驗時間內收集到足夠的產品失效資料來預測產品壽命。此時，若存在適當的品質特性 (Quality Characteristic, QC) 其衰變量與產品壽命息息相關，則可利用衰變資料來預測產品壽命。利用衰變資料來預測產品壽命，首先必須建立適當的衰變模型 (Degradation Model)。文獻上衰變路徑之模型常用 Wiener 過程 (Wiener Process) 或高斯過程 (Gaussian Process)。當衰變路徑為單調 (Monotone) 時，Gamma 和逆高斯 (Inverse Gaussian, IG) 過程則是常被使用的衰變模型。這些參數衰變模型 (Parametric Degradation Model) 雖較簡單但因侷限於必須事先選定一固定函數形式之模型，較無彈性。提供先進電子設備預診斷解決方案的國際大廠 Ridgetop Group, Inc.，利用 QC 的參數衰變模型，加上適應性的學習方法 (Adaptive Learning)，建購了他們的 RUL 估計軟體 ARULETM Reasoner，為現在光電、半導體業界應用最廣的 RUL 估計軟體。

本計畫將運用具創新性的統計方法，建立資料自己決定 (Data-Driven) 的無母數迴歸衰變模型 (Nonparametric Regression Degradation Model)。相較於傳統的參數衰變模型，在衰變路徑之函數形式不確定的情形下，我們提出的無母數方法是較佳的選擇。又由於機台的故障模式 (Failure Mode) 可能不只一種、生命週期間可能有概念漂移 (Concept Drift) 的現象、不同機台/腔體之間會有機差 (Cross-Tool/Cross-Chamber) 的問題、不同機型之間會有機構設計不同 (Cross-Tooltype) 的問題，單一的衰變模型將無法正確地描述我們的機台資料。因此我們將採用整體式適應性的學習方法 (Ensemble-based Adaptive Learning)，結合數個不同的衰變模型，針對不同的機台/腔體/機型，動態地調整模型參數與結合權重。預計採用的衰變模型包括傳統的參數衰變模型、我們所提出的無母數迴歸衰變模型、工研院巨資中心研究團隊所發展的機器學習模型 (Machine Learning Model)。我們寄望這個整體式適應性的 RUL 預測，將可大幅改進預測的誤差。

3. 計畫目標

現存大部分的 RUL 估計，多採用單一參數模型來描述 QC 的衰變量，進而預測產品壽命。然而，隨著製程設備複雜，使用現有之方式進行 RUL 估計，因無考慮到設備零件異常模式的差異與機台/腔體/機型間的不同，使得估計準確度偏低。為求更進一步地提升預診斷的精度，就必須研發一個智慧型的 RUL 估計機制，來結合數個不同的衰變模型，並能動態地調整模型參數與結合權重，以提升預測的準確性。

本計畫將以巨量資料分析技術，研發整體式適應性的學習方法，以改善硬體零件 RUL 估計效能與設計錯誤修正。未來透過與關鍵組件老化訊號監控與老化訊號預警 (Predictive Maintenance) 系統技術的結合，可有效掌握機台運轉性能，減少無預警提醒之情況，提高機台整體之稼動率，以期達到機台近零失效 (Near-Zero Downtime) 的維運目標，間接提高機台製造商其產品之附加價值與智慧化能力。

本計畫將針對前期的 LED MOCVD 過濾器 (Particle Filter, PF) 機台零件資料與 IEEE PHM 2014 Data Challenge 的資料進行剩餘壽命的預測。計畫的預期目標為：(1) 針對機台 PF 零件，所研發的剩餘壽命整體式適應性的學習方法，其預測正確率的改善率 >20%、mean squared error (MSE) 的改善率 >20%；(2) 針對 IEEE PHM 2014 Data Challenge，所研發的剩餘壽命整體式適應性的學習方法，預測的分數可提升至少 10%。

4. 計畫內容

本計畫將運用許多具創新性的統計方法。對於能搜集到「錯誤事件資料」(Failure

Event Data) 的系統，進行存活分析 (Survival Analysis)，其中可利用 Kaplan-Meier Estimator for Survival Function 來估計 (預測) 設備元件壽命會大於某一觀察時間的機率，利用 Cox Proportional Hazards Model 來建立存活壽命與元件不同特徵值 (Features) 間的關係。存活分析內包含豐富的模型，可處理許多不同型態的存活資料，包括當我們要分析經由整體式學習 (Ensemble Learning) 所得到的特徵集合 (Feature Pool) 中所有零件的存活壽命時，多 (高) 維度存活資料分析也可利用。

對於僅能搜集到「與錯誤事件相關的狀況監視資料」(Condition Monitoring Data Related to Failure Events) 的系統，可靠度分析 (Reliability Analysis) 中常以衰變模型 (Degradation Model) 來描繪這類型的資料，而常用的衰變模型包含：Regression-Based Model、Wiener Processes、Gamma Processes、Stochastic Filtering-Based Model、Hidden Markov Model。本計畫也預計發展資料自己決定 (Data-Driven) 的無母數迴歸衰變模型 (Nonparametric Regression Degradation Model)，相較於傳統的參數衰變模型，在衰變路徑之函數形式不確定的情形下，我們提出的無母數方法將會是較佳的選擇。

由於機台的故障模式 (Failure Mode) 可能不只一種、生命週期間可能有概念漂移 (Concept Drift) 的現象、不同機台/腔體之間會有機差 (Cross-Tool/Cross-Chamber) 的問題、不同機型之間會有機構設計不同 (Cross-Tooltype) 的問題，單一的衰變模型將無法正確地描述我們的機台資料。因此我們將採用整體式適應性的學習方法 (Ensemble-based Adaptive Learning)，結合傳統的參數衰變模型、我們所提出的無母數迴歸衰變模型、工研院巨資中心研究團隊所發展的機器學習模型 (Machine Learning Model)，針對不同的機台/腔體/機型，動態地調整模型參數與結合權重，以提升 RUL 預測的準確性。