

## 科技部新聞稿

### 巨量資料分析在高科技產業品質管理之應用

103.11.19

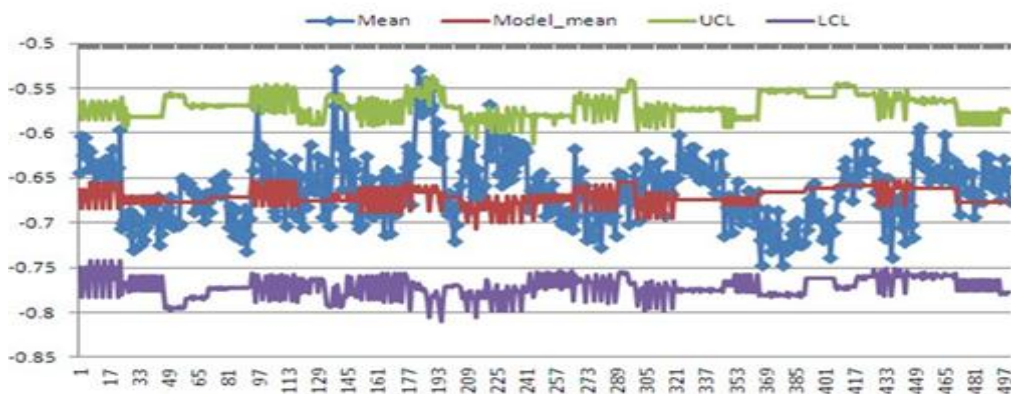
晶圓的製造過程非常繁複，需經大量機台層層加工塗料，才會產生一個合格的晶圓成品。因此，若是想要找出製程中產生問題的機台，進一步降低不良率，就成了高科技產業品質管理上的一大困難。更加困難的是，在製程初期，晶圓成品仍少(即資料樣本數少)，此時如何在大量機台中，想挑選出少數真正有影響力的機台(即影響成品的變數)，往往如大海撈針般，困難度極高。因此如何利用少數成品在大量的機台中，尋找出製程中關鍵的機台，以提升良率，更進一步提升產能是高科技產業非常關心的議題。

晶圓製程中所面臨的問題，也就是學術上所謂『高維度迴歸模型選擇』的問題，也就是所謂  $p$  大  $n$  小的問題。所謂“ $p$  大  $n$  小”中的  $p$  是指模型中解釋變數之個數，也就是晶圓製程中通過機台的數量，而  $n$  則為樣本數，也就是晶圓製程中成品的數量，如何透過學術上高維度模型選取的方法，將影響成品品質的關鍵性機台找出來，是重要議題；而此議題即是近年來大家非常關注所謂巨量資料(Big Data)的議題。

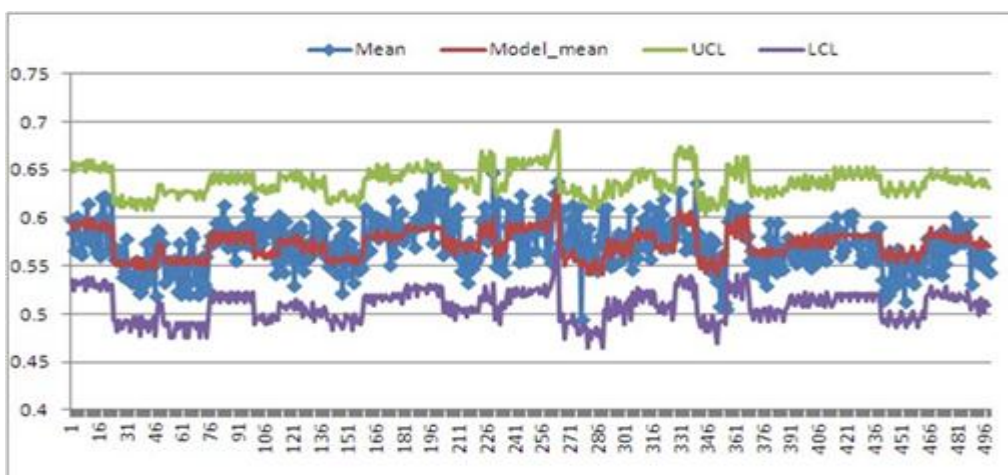
銀慶剛教授，是國內率先研究高維度迴歸模型的統計學家之一，早於 2006 年即預見前述問題的重要性，在當時國科會(現科技部)的支持之下，開始深入探討。早期，他在模型選擇理論上的成就經報導後，獲得了台積電統計發展與應用部的重視，也因此開啟了合作的端緒。國際上一份極重要的統計期刊 Statistica Sinica，於 2011 年底，將銀教授與中央研究院院士美國史丹福大學統計系的黎子良教授在  $p$  大  $n$  小工作上的合作成果，以主要文章(leading article)的形式刊登。在這篇總結數年研究心得長達 41 頁的論文中，他們提出一個遠較其他方法優越的三階段選模法，能在許多極端困難的情況下，挑選出重要變數。

台積電面臨的  $p$  大  $n$  小問題，經雙方初步討論指出，其複雜度遠超過銀教授及黎院士論文中所設定的情況，頗具挑戰性。因此雙方在銀教授現有的理論基礎上，進行更大的突破及創新，以期提高晶圓良率，進而達到穩定量產的最終目的。

合作初期，由於銀教授團隊的結果明顯優於早期公司內部所建立之統計模型，因此正式展開雙方合作計畫。如下圖一、圖二所示，銀教授團隊重新建立的統計模型，可以充分掌握資料波動的特性，深獲肯定。



圖一：上圖中藍色為實際資料，紅色為早期建立之統計模型的預測值。可以看出其模型無法反映資料實際之起伏變動，故遭受質疑。



圖二：上圖中藍色為實際資料，紅色為重新建立之統計模型的預測值。可以看出其模型隨資料波動而變動，各項數據亦顯示銀教授模型的優點，並深獲台積電肯定。

其後，合作期間，銀教授團隊模型的優勢一再在實際生產線上得到驗證。例如，與工廠實際配合時，銀教授依據模型建議可能產生問題的機台，停止生產，並測試產品不良率，在數次實驗中，不良率分別下降 11%~14%。此外亦依據工廠具專業背景之線上工程師建議，停止使用可能產生問題的機台，並測試產品不良率，測試結果顯示，多數不良率下降在 3% 以內，有時不良率甚至會上升，也因此更加深工廠對銀教授團隊的信任，建立其後良好的合作關係。有了這些經驗，工廠也對模型的預測表現感到訝異，特別是銀教授僅以數據分析，並未親臨工廠，卻可以找出影響成品的機台，也因此更加深雙方合作的信任基礎，建立良好的合作關係。

此後，銀教授亦進一步替工廠解決顧客所提出『檢驗代表性的問題』，提出檢定方法、理論背景，並以實際資料加以驗證，亦得到相當成功的結果。

**這個過程中主要的成就可以總結為以下五點：**

- (1) 協助解決高科技產品品質管理問題，找出製程中影響成品品質的機台，有效降低成品不良率。
- (2) 學術上創新之研究課題：在學界業界的合作下，開創出許多創新的研究議題，其後陸續發表於國際一流期刊上，目前仍在持續進行中。
- (3) 學生透過合作議題，參與計畫，學以致用。
- (4) 學生獲得就業機會。
- (5) 學界業界雙邊合作良好範例。