

國立交通大學
工業工程與管理學系

博士論文

整合良率預估及缺陷樣式辨識之晶圓

缺陷診斷系統

Constructing a Wafer Defect Diagnostic System by
Integrating Yield Prediction and Defect Pattern
Recognition

研究生：趙豐昌

指導教授：唐麗英 教授

中華民國九十八年三月

整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統
Constructing a Wafer Defect Diagnostic System by Integrating
Yield Prediction and Defect Pattern Recognition

研究生：趙豐昌

Student : Li-Chang Chao

指導教授：唐麗英

Advisor : Lee-Ing Tong



Industrial Engineering and Management

March 2009

Hsinchu, Taiwan, Republic of China

中華民國九十八年三月

整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統

研究生：趙豐昌

指導教授：唐麗英 教授

國立交通大學工業工程與管理學系

摘要

積體電路的製作技術日趨精密，然無論如何改善製程技術，其產品的良率仍無法達到 100%。造成積體電路產品良率損失的最主要原因，乃是晶圓上產生的缺陷。隨著晶圓面積的增大，缺陷出現群聚現象，以往常用之卜瓦松良率模式 (Poisson Yield Model) 會因低估良率而不再適用。近年來，許多中外文獻針對晶圓之缺陷群聚現象而提出一些良率模式，然而這些良率模式各有不完善之處。如：負二項良率模式 (Negative Binomial Yield Model) 則因其中的缺陷群聚參數 α 值過於散亂，同時也可能為負值，而造成分析良率時的不便；複合卜瓦松良率模式 (Compound Poisson Yield Model) 雖然預測的良率較卜瓦松良率模式準確，但模式的建構卻相當複雜，較不易為業界所用。而 Jun *et al.* 以迴歸分析法所建構之良率模式，其模型之適配度與資料是否違反迴歸模型的假設都是值得考慮的。利用倒傳遞網路構建之良率模式則必需經由適當地設定網路參數才能得到較佳的預測結果。此外，目前半導體晶圓資料分析仍是以人工的方式分析晶圓上缺陷的空間樣式來找出製程變異的可能原因。然而人工判定除了費時外，亦可能因誤判進而影響偵測製程變異的準確性。有文獻提出一些晶圓缺陷樣式辨識方法來判定製程是否異常，但皆有其不完善之處。因此，本研究的主要目的是要發展一個整合良率預估模式及缺陷樣式辨識之晶圓缺陷診斷系統。本研究利用一般迴歸神經網路 (General Regression Neural Network, GRNN) 來構建良率模式，並利用多類別支撐向量機 (Multi-class Support Vector Machines, Multi-class SVM) 來構建一個晶圓缺陷樣式辨識系統；最後整合成一個良率預估模式及缺陷樣式辨識之晶圓缺陷診斷系統。本研究最後以模擬實驗來說明本研究所提之整合良率預估模式及缺陷樣式辨識之晶圓缺陷診斷系統的可行性；並進一步與現有文獻所提之良率預估模式與缺陷樣式辨識系統進行比較以驗證本研究的有效性與優越性。

【關鍵字】 積體電路、缺陷、群聚現象、良率模式、一般迴歸神經網路、樣式辨識、支撐向量機

Constructing a Wafer Defect Diagnostic System by Integrating Yield Prediction and Defect Pattern Recognition

Student : Li-Chang Chao

Advisor : Lee-Ing Tong

Department of Industrial Engineering and Management
National Chiao Tung University

Abstract

Wafer yield is a highly effective means of evaluating the process capability of integrated circuit manufacturers. The defect number and cluster intensity of defects on a wafer are two critical factors influence wafer yield. As wafer sizes increase, the clustering phenomenon of defects increases. Clustered defects cause the conventional Poisson yield model underestimate actual wafer yield. The cluster parameter α of the negative binomial model can be very scattered and negative when the model is applied to predict yield. Compound Poisson yield models are complicated. The degree of fitness for regression model must be considered when the regression methods are utilized to model the yield. Obtaining good prediction network requires substantial effort to identify the parameters of back-propagation neural network. Although some yield models consider the effects of defect clustering on yield prediction, these models have some drawbacks. Furthermore, the possible causes of process variation can be found out by operators through analyzing the defect pattern on a wafer. Judging the process by operators can waste time and the accuracy of process detecting can be influenced due to the erroneous conclusion. Although some recognizing methods for defect pattern on a wafer were proposed, these recognizing methods still have some flaws. This study presents a novel wafer defect diagnostic system that utilizes a general regression neural network integrating a multi-class support vector machines to predict the wafer yield and recognize the defect pattern on a wafer. A simulation study is utilized to demonstrate the effectiveness of the proposed method.

【Key Words】 Integrated circuit, defect, clustering phenomenon, yield model, general regression neural network, pattern recognition, support vector machines

誌謝

本篇論文得以順利完成，首先要感謝指導老師唐麗英教授的悉心指導，使我在研究的過程中受益良多，謹此致上無限的感激。此外，感謝本系梁高榮老師、科技管理研究所虞孝成老師、清華大學科技管理研究所黎正中老師以及國防大學理工學院動力及系統工程學系王春和老師等人對本論文提出之建議與指導，使本論文更臻完善，謹此致謝。

在博士班的生涯中，感謝與我一起分享喜悅，分擔憂愁的同窗好友，以及同門的學弟妹，我會永遠記得與你們相處的這段美好時光。同時也感謝我目前的服務學校致遠管理學院，能讓我在職進修以完成博士學位。

最後，僅以此論文獻給我的家人——偉大的母親，有您的付出才有今日的我，由衷地感謝您；給予我精神支柱的太太，有你在背後默默地支持，我才能無後顧之憂來完成學位；帶給我愉悅心情的兩個可愛小寶貝，你們讓我忘了一天的疲憊，也讓我覺得付出有所代價；還有大哥、大嫂、大姐、二姐以及兩位姪女，有你們大家的付出，我們家才能展現和樂融洽的氣氛，謝謝你們！

趙豐昌 謹誌於

交通大學工業工程與管理研究所

2009年3月

目 錄

第一章 緒論.....	1
1.1 研究背景與動機.....	1
1.2 研究目的.....	3
1.3 研究限制.....	3
1.4 研究架構.....	4
第二章 文獻探討.....	5
2.1 積體電路良率模式.....	5
2.1.1 卜瓦松良率模式.....	5
2.1.2 複合卜瓦松良率模式.....	5
2.1.3 負二項良率模式.....	6
2.1.4 使用迴歸分析法構建之良率模式.....	8
2.1.5 類神經網路良率模式.....	9
2.2 一般迴歸神經網路.....	10
2.2.1 GRNN 網路架構.....	10
2.2.2 GRNN 網路基本原理.....	11
2.2.3 GRNN 的優點.....	12
2.3 缺陷群聚指標.....	13
2.3.1 負二項良率模式的 α 群聚參數.....	13
2.3.2 空間統計學之缺陷群聚指標.....	13
2.3.3 無統計假設之缺陷群聚指標.....	14
2.3.4 其他缺陷群聚指標.....	15
2.4 缺陷樣式之辨識.....	16
2.4.1 倒傳遞網路(Back-Propagation Neural Network, BPNN).....	16
2.4.2 徑向基底函數(Radial Basis Function, RBF)神經網路.....	18
2.4.3 支撐向量機(Support Vector Machines, SVM).....	20
2.4.3.1 支撐向量分類(Support Vector Classify).....	20
2.4.3.2 核函數(Kernel Function).....	22
2.4.4 多類別支撐向量分類.....	23
2.4.4.1 一對多(One Versus the Rest)分類法.....	23
2.4.4.2 一對一(Pairwise)分類法.....	23
2.4.4.3 DAG(Directed Acyclic Graph)分類法.....	24
2.4.4.4 一次(Considering all Data at a Once)分類法.....	24
第三章 研究方法.....	25
3.1 本研究之特徵因子的選取.....	25

3.2 本研究構建之良率模式.....	28
3.3 本研究構建之晶圓缺陷樣式辨識系統.....	29
3.4 整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統.....	31
第四章 模擬驗證.....	33
4.1 模擬實驗.....	33
4.2 良率模式的比較.....	36
4.2.1 各良率模式與缺陷數因子水準之關係.....	39
4.2.2 各良率模式與群聚程度因子水準之關係.....	43
4.2.3 各良率模式與群聚分佈面積因子水準之關係.....	46
4.3 晶圓缺陷樣式辨識技術的比較.....	49
4.3.1 各樣式辨識技術與缺陷數因子水準之關係.....	51
4.3.2 各樣式辨識技術與缺陷群聚程度因子水準之關係.....	54
4.3.3 各樣式辨識技術與群聚分佈面積因子水準之關係.....	57
4.4 整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統.....	60
第五章 結論.....	65
參考文獻.....	66



表目錄

表 2.1	積體電路良率模式彙整表.....	7
表 2.2	不同 α 值之下，負二項良率模式近似之良率模式及群聚現象.....	8
表 2.3	α 值與 CI 值之比較.....	15
表 4.1	245 片測試樣本的良率預估值與真實良率間之變異數分析.....	38
表 4.2	5 種缺陷樣式下，各良率模式在 3 種設計因子的預估績效.....	49
表 4.3	5 種缺陷樣式下，各辨識技術在 3 種設計因子的辨識績效.....	60
表 4.4	10 片晶圓樣式之相關資訊.....	64
表 4.5	10 片晶圓之 GRNN 預估良率與良率程度.....	64
表 4.6	6 片晶圓之真實缺陷樣式與多類別 SVM 辨識之樣式.....	64



圖目錄

圖 2.1	GRNN 的網路架構.....	10
圖 2.2	缺陷分布圖與兩軸的投影.....	15
圖 2.3	BPNN 之網路架構.....	17
圖 2.4	RBF 神經網路.....	20
圖 2.5	最佳超平面.....	21
圖 2.6	DAG 分類法處理 3 類別的分類.....	24
圖 3.1	夾角與距離示意圖.....	27
圖 3.2	群聚指標示意圖.....	28
圖 3.3	本研究之 GRNN 良率模式架構圖.....	29
圖 3.4	常見的晶圓表面缺陷樣式.....	30
圖 3.5	本研究之多類別 SVM 晶圓缺陷樣式辨識系統架構圖.....	30
圖 3.6	本研究之整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統構建流程圖.....	32
圖 4.1-	(a) 隨機樣式.....	34
圖 4.1-	(b) 牛眼樣式.....	34
圖 4.1-	(c) 弦月樣式.....	35
圖 4.1-	(d) 底部樣式.....	35
圖 4.1-	(e) 環狀樣式.....	36
圖 4.2	245 片測試樣本的良率預估值與真實良率間之 RMSE 的比較.....	38
圖 4.3	245 片測試樣本的良率預估值與真實良率間之散佈圖.....	39
圖 4.4	隨機樣式時，各種良率模式之缺陷數因子水準與 RMSE 的關係.....	40
圖 4.5	牛眼樣式時，各種良率模式之缺陷數因子水準與 RMSE 的關係.....	40
圖 4.6	弦月樣式時，各種良率模式之缺陷數因子水準與 RMSE 的關係.....	41
圖 4.7	底部樣式時，各種良率模式之缺陷數因子水準與 RMSE 的關係.....	42
圖 4.8	環狀樣式時，各種良率模式之缺陷數因子水準與 RMSE 的關係.....	42
圖 4.9	牛眼樣式時，各種良率模式之群聚程度因子水準與 RMSE 的關係.....	43
圖 4.10	弦月樣式時，各種良率模式之群聚程度因子水準與 RMSE 的關係.....	44
圖 4.11	底部樣式時，各種良率模式之群聚程度因子水準與 RMSE 的關係.....	45
圖 4.12	環狀樣式時，各種良率模式之群聚程度因子水準與 RMSE 的關係.....	45
圖 4.13	牛眼樣式時，各種良率模式之群聚分佈面積因子水準與 RMSE 的關係.....	46
圖 4.14	弦月樣式時，各種良率模式之群聚分佈面積因子水準與 RMSE 的關係.....	47
圖 4.15	底部樣式時，各種良率模式之群聚分佈面積因子水準與 RMSE 的關係.....	48
圖 4.16	環狀樣式時，各種良率模式之群聚分佈面積因子水準與 RMSE 的關係.....	49
圖 4.17	隨機樣式時，各種辨識技術之缺陷數因子水準與正確辨識率的關係.....	52
圖 4.18	牛眼樣式時，各種辨識技術之缺陷數因子水準與正確辨識率的關係.....	52
圖 4.19	弦月樣式時，各種辨識技術之缺陷數因子水準與正確辨識率的關係.....	53
圖 4.20	底部樣式時，各種辨識技術之缺陷數因子水準與正確辨識率的關係.....	53

圖 4.21	環狀樣式時，各種辨識技術之缺陷數因子水準與正確辨識率的關係	54
圖 4.22	牛眼樣式時，各種辨識技術之群聚程度因子水準與正確辨識率的關係.....	55
圖 4.23	弦月樣式時，各種辨識技術之群聚程度因子水準與正確辨識率的關係.....	55
圖 4.24	底部樣式時，各種辨識技術之群聚程度因子水準與正確辨識率的關係.....	56
圖 4.25	環狀樣式時，各種辨識技術之群聚程度因子水準與正確辨識率的關係.....	57
圖 4.26	牛眼樣式時，各種辨識技術之群聚分佈面積因子水準與正確辨識率的關係....	57
圖 4.27	弦月樣式時，各種辨識技術之群聚分佈面積因子水準與正確辨識率的關係....	58
圖 4.28	底部樣式時，各種辨識技術之群聚分佈面積因子水準與正確辨識率的關係....	59
圖 4.29	環狀樣式時，各種辨識技術之群聚分佈面積因子水準與正確辨識率的關係....	60
圖 4.30- (a)	隨機樣式.....	62
圖 4.30- (b)	牛眼樣式.....	62
圖 4.30- (c)	弦月樣式.....	62
圖 4.30- (d)	底部樣式.....	63
圖 4.30- (e)	環狀樣式.....	63



第一章 緒論

1.1 研究背景與動機

半導體產業是我國目前重點發展的產業。依原料、生產/加工至產品產出，半導體產業大致可區分為半導體材料(含化學品)、光罩、設計(含 CAD 軟體)、製程、封裝、測試及設備等七個技術領域。半導體產品包括積體電路(Integrated Circuits, IC)、分離式(Discrete)元件和光電(Optoelectronic)元件等三大類，廣泛應用於資訊、通訊、消費性電子、工業儀器、運輸及國防太空等領域[2]。

然而，無論如何改善積體電路的製作技術，其產品的良率仍無法達到 100%。造成積體電路產品良率損失的最主要原因，乃是晶圓上所產生的缺陷與缺陷的群聚程度。這些缺陷一般可分為三類，即嚴重缺陷(gross defects)，參數缺陷(parameter defects)以及隨機缺陷(random defects)等[18]。嚴重缺陷是屬於相當大的缺陷。例如，刮痕(scratches)、處置不當或尚未完全移除的光阻劑等所造成的損壞；嚴重缺陷經常會造成良率的損失。參數缺陷是指會影響設備電氣參數(電流、電壓、電阻...等)的缺陷。例如，一個晶片上由於雜質(dopant)不均勻的集中，可能造成原先被視為相同的電阻器上卻有不同的電阻值；參數缺陷有時可能產生良率損失，但它們較常引起一些可靠度問題(例如，設備在使用一段時間後，失去了原先所設計的功能)。隨機缺陷乃是機遇性發生的缺陷，造成良率損失的隨機缺陷又稱之為致命缺陷(fatal defects)或失效(faults)。

並非所有的製程缺陷都會造成良率的損失；製程缺陷是否會造成產品失效乃是與產品的特性(例如，晶片面積、缺陷分布狀況、線寬大小、線路形狀...等)有關。由於線寬大小與線路形狀所引起的良率損失，不易用機率模式來描述，所以一般預測積體電路良率模式的文獻中，都著重於以晶片面積與缺陷分布狀況兩參數來預測良率。

藉由特定的機率函數通常可以描述缺陷的分布情形，並進而建立產品之良率與製程缺陷間的關係，即所謂的良率模式(Yield Model)。自 1960 年代開始就有學者提出各種不同的良率模式來預估良率[14][37][41][45]。其中卜瓦松良率模式(Poisson Yield Model)對於小晶片的良率預測效果較佳[6]。負二項良率模式(Negative Binomial Yield Model)則對大晶片的良率預測效果較好；目前的積體電

路工廠多採用負二項良率模式來預估良率[14]。複合卜瓦松良率模式[14]在良率預測上亦較卜瓦松良率模式準確。Jun *et al.* [25] 利用一缺陷群聚指標 CI 與晶片上的平均缺陷數建構出一個迴歸模式來估計晶圓良率；而其中之 CI 值能以均勻的數值範圍來量化不同嚴重程度的群聚現象。還有一些利用類神經網路(例如，倒傳遞網路)構建之良率模式也常被用來預估良率[1][4][42]。

Stapper[40]指出，當晶圓的面積增大時，晶圓上的缺陷並非呈現隨機性分布，而是有群聚現象發生。這些群聚的缺陷通常分布在晶圓的邊緣[38]。缺陷群聚現象使得卜瓦松良率模式會低估產品應有的良率。負二項良率模式中的缺陷群聚參數 α 值過於散亂，同時也可能為負值，這些都造成分析良率時的不便。複合卜瓦松良率模式雖然良率的預測較卜瓦松良率模式準確，但模式的建構卻相當複雜，較不易為業界所接受。Jun *et al.* [25] 以迴歸分析法所建構之良率模式，其模式之適配度以及資料是否違反迴歸模式的基本假設都是值得考慮的。利用倒傳遞網路構建之良率模式則必需經由適當地設定許多網路參數才能得到較佳的預測結果。

目前在半導體晶圓缺陷資料分析方法中，大多是以人工的方式分析晶圓上缺陷的空間樣式來找出製程變異的可能原因[28]。然而人工判定除了費時外，亦可能因誤判進而影響製程變異偵測上的準確性。因此，如何構建一個有效的晶圓缺陷樣式辨識系統一直受到學界與業界的重視。一般之晶圓缺陷樣式辨識主要可分成三個領域：統計分析法、啟發式演算法以及人工智慧學習[31]。統計分析法乃針對問題給定不同的分配假設以進行資料點的群聚現象分析，其目的在於如何將空間上的資料點進行切割，以達到樣式分類辨識的目的[23]。啟發式演算法通常利用柔性計算的方法(例如，基因演算法與模糊理論)來進行樣式辨識。然而，基因演算法一般必須付出相當多的努力於評估過程才能達到最佳解[7]。而模糊理論的主要限制之一乃是其語意控制規則相當難產生；再者，必須借重一些專家的知識與經驗才能設計出一完善的模糊邏輯控制器[12]。人工智慧學習乃藉由模仿生物神經系統的運算來達成資訊處理之目的，類神經網路即人工智慧學習中一項常用的技術。然而，類神經網路技術的主要限制之一乃是其各項網路參數必須適度的設定才能得到較佳的結果[19]。近年來支撐向量機(Support Vector Machines, SVM)已廣泛應用在樣式分類辨識上，而且有好的辨識結果[10][22]。例如：臉和手印身分驗證[49]、土地的覆蓋變化察覺[30]、國畫的辨識

與分類[24]、遺傳學的綜合症診斷分類[15]等。

缺陷之群聚現象可能造成良率模式低估晶圓良率；而晶圓的缺陷分布若呈現特定樣式(pattern)，亦可視為製程異常。因此，當一片晶圓出現中、低良率時，應進一步診斷該晶圓表面的缺陷是否存在某種特定的缺陷樣式。為適切地反應出產品的真實良率，並且能夠針對呈現特定樣式的晶圓做出正確的製程異常診斷，有必要構建出一個能考慮到缺陷群聚現象與缺陷分布呈現特殊樣式的一個較精密、正確但又計算簡便之整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統，使業界能有效地用其來做出正確的製程異常診斷。

1.2 研究目的

針對缺陷群聚現象，鑑於過去學者提出之良率預估模式各有其不完善之處；再者，為滿足學界與業界對建構一個有效的晶圓缺陷樣式辨識系統的需求。本研究的主要目的即是要發展一個整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統。本研究利用一般迴歸神經網路(General Regression Neural Network, GRNN)來構建良率模式，並利用多類別支撐向量機(Multi-class Support Vector Machines, Multi-class SVM)來構建缺陷樣式辨識系統；最後再將前述兩者整合成一個良率預估模式及缺陷樣式辨識之晶圓缺陷診斷系統。本研究以模擬之晶圓缺陷實驗來說明本研究所提之晶圓缺陷診斷系統的可行性；並進一步與中外文獻所提之良率預估模式與缺陷樣式辨識系統進行比較以驗證本研究的有效性。

1.3 研究限制

本論文的研究限制如下：

1. 晶圓上每個缺陷皆會影響產品的良率。缺陷數愈多即表示製程異常的狀況愈嚴重。
2. 不考慮缺陷的大小與型態，只針對晶圓上的總缺陷數分析其對良率的影響。
3. 當製程處於管制狀態時，缺陷應隨機分布於晶圓表面。當缺陷群聚程度愈大或缺陷分布所呈現的特定樣式愈明顯時，即表示製程異常的狀況愈嚴重。

1.4 研究架構

本論文共分為五個章節，第一章為緒論，說明本研究之背景與動機、研究目的、研究限制以及研究架構；第二章為文獻探討，介紹積體電路良率模式、一般迴歸神經網路、缺陷群聚指標以及缺陷樣式之辨識；第三章為研究方法，介紹本研究之特徵因子的選取、本研究構建之良率模式、本研究構建之晶圓缺陷數診斷系統以及整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統；第四章說明如何以模擬實驗來驗證本研究所發展的整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統的可行性；第五章為本論文的結論。



第二章 文獻探討

本章介紹各種積體電路良率模式、一般迴歸神經網路、缺陷群聚指標以及缺陷樣式辨識之相關文獻。

2.1 積體電路良率模式

自 1960 年代開始就有學者提出各種不同的良率模式來預估良率[14]。只要求得良率模式中之各項參數(例如，缺陷個數、缺陷密度、晶片面積...等等)，便能夠得到相對的預估良率。以下便對各種良率模式以及利用迴歸分析法與類神經網路所構建之良率模式做一介紹。

2.1.1 卜瓦松良率模式

卜瓦松良率模式[18]是由卜瓦松機率函數發展而來。使用此模式預估良率時必須符合下面兩個基本假設：

1. 缺陷必須呈均勻且隨機性分布。
2. 缺陷密度為一固定常數。

在此假設下，卜瓦松良率模式為：

$$Y = P(k = 0) = e^{-\lambda} \quad (1)$$

其中， k 表每個晶片上的缺陷數， λ 表每個晶片上的平均缺陷數。

由卜瓦松良率模式可知，只要求得每個晶片上的平均缺陷數 λ_0 ，便可得到良率。因此，卜瓦松良率模式擁有計算方便之特性。一般工業之品質管制的應用上，在處理缺陷資料時，通常均假設缺陷的分布為卜瓦松分配；即缺陷呈隨機分布。然而目前已有廣泛的報告指出，在 IC 製程中晶圓表面所產生的缺陷有群聚的傾向[40]。很明顯的，這種群聚現象的發生破壞了卜瓦松模式的前提假設。

Cunningham[14]指出，當晶片面積小於0.25平方公分時，卜瓦松良率模式既簡單又精確。但隨著晶片面積的增大，缺陷出現群聚現象的機率增加，卜瓦松良率模式出現低估良率的情形。因此在預測較大面積之晶片的良率時，目前業界多採用負二項良率模式。

2.1.2 複合卜瓦松良率模式

卜瓦松良率模式的基本假設是缺陷密度為一固定常數。但 Murphy[29]認為

缺陷密度在各晶片，甚至各晶圓之間並不相同。他主張必須以一個呈現常態的機率函數 $f(D)$ 來取代。因此 Murphy 所主張的良率模式如下：

$$Y = \int_0^{\infty} e^{-DA} f(D) dD \quad (2)$$

其中， D 表缺陷密度， A 表晶片面積， $f(D)$ 表缺陷機率密度函數。式子(2)稱為複合卜瓦松良率模式(Compound Poisson Yield Model)。

Murphy 認為利用高斯(Gaussian)分配來估計 $f(D)$ 相當合理。但由於無法整合表示式，所以利用式子(3)之 triangular form 來近似 $f(D)$ 。亦即

$$f(D) = \begin{cases} \frac{D}{D_0^2} & , 0 \leq D \leq D_0 \\ \frac{1}{D_0} \left(2 - \frac{D}{D_0}\right) & , D_0 \leq D \leq 2D_0 \\ 0 & , D \geq 2D_0 \end{cases} \quad (3)$$

其中， D 表缺陷密度， D_0 表平均缺陷密度。將上式(3)代入式子(2)，可得 Murphy 良率模式為：

$$Y = \left(\frac{1 - e^{-D_0 A}}{D_0 A}\right)^2 \quad (4)$$

自 Murphy 提出缺陷密度的觀念以來，陸續有學者以不同的機率密度函數 $f(D)$ 來建構積體電路良率模式，其相關資料彙整於表 2.1 [14][34]。

2.1.3 負二項良率模式

由 Stapper[39]所提出的負二項良率模式是業界常採用的良率模式。Stapper 以 Gamma 函數來描述式子(2)中之缺陷機率密度函數 $f(D)$ 。亦即將式子(5)之 $f(D)$ 代入式子(2)，

$$f(D) = \frac{D^{\alpha-1} e^{-D/\beta}}{\beta^{\alpha} \Gamma(\alpha)} \quad (5)$$

得良率模式如下：

$$Y = \frac{1}{(1 + D_0 A / \alpha)^{\alpha}} \quad (6)$$

其中， D_0 表每單位面積之平均缺陷數， A 表晶片面積， α 表群聚參數。

表 2.1 積體電路良率模式彙整表

模式名稱	良率模式	缺陷密度分配	提出學者及年代
Poisson	$Y = e^{-D_0 A}$	-	1960
Murphy (Poisson-Triangular mixture)	$Y = \left(\frac{1 - e^{-D_0 A}}{D_0 A} \right)^2$	Triangular	Murphy 1964
Poisson-Uniform mixture	$\frac{1 - e^{-2D_0 A}}{2D_0 A}$	Uniform	Murphy
Seeds (Poisson-Exponential mixture)	$Y = \frac{1}{1 + D_0 A}$	Exponential	Seeds 1967
Negative Binomial (Poisson-Gamma mixture)	$Y = \frac{1}{(1 + D_0 A / \alpha)^\alpha}$	Gamma	Stapper 1973
Half-Gaussian (Poisson-Half-Gaussian mixture)	$Y = \exp[(\pi/4)D_0^2 A] \operatorname{erfc}\left(\frac{\pi^{1/2} D_0 A}{2}\right)$	Half-Gaussian	Stapper 1991
Poisson-Weibull mixture	$\sum_{k=0}^{\infty} (-1)^k \frac{(D_0 A)^k}{k!} \frac{\Gamma(1 + k / \alpha)}{[\Gamma(1 + 1 / \alpha)]^k}$	Weibull	-
Poisson-Rayleigh mixture	$Y = 1 - D_0 A \exp[(D_0 A)^2 / \pi] \left(\operatorname{erfc}\left(\frac{D_0 A}{\pi^{1/2}}\right) \right)$	Rayleigh	Raghavachari <i>et al.</i> 1996
Poisson-Inverse Gaussian mixture	$Y = \exp\left[\phi \left(1 - (1 + (2D_0 A / \phi))^{1/2}\right)\right]$	Inverse Gaussian	Raghavachari <i>et al.</i> 1996

負二項良率模式之群聚參數 α 可以下列式子表示，亦即

$$\alpha = \frac{\lambda^2}{(\sigma^2 - \lambda)} \quad (7)$$

其中， λ 表每個晶片上的平均缺陷數， σ^2 表晶片上缺陷之變異數。藉由選擇不同的 α 值，Gunningham[14] 模擬了一些學者所提出的良率模式。表 2.2 列出在不同的 α 值之下，負二項良率模式近似的良率模式以及群聚情形。

負二項良率模式對大晶片的良率預測效果好；目前的積體電路工廠多採用

負二項良率模式來預估良率。然而，當利用式子(7)來求 α 值時， α 值會顯得過於散亂，同時也可能為負值(σ^2 可能小於 λ)，這些都會造成分析良率時的不便。

表 2.2 不同 α 值之下，負二項良率模式近似之良率模式及群聚現象

α 值	近似之良率模式	群聚現象
大約 10 ~ ∞	卜瓦松模式	無
4.2	Murphy	一些
3	Dingwall	一些
1	Seeds	很多

2.1.4 使用迴歸分析法構建之良率模式

Jun *et al.*[25]提出一個缺陷群聚指標(Cluster Index, CI)。 CI 可以式子(8)來計算。

$$CI = \min\left\{\frac{s_v^2}{\bar{v}}, \frac{s_w^2}{\bar{w}}\right\} \quad (8)$$

其中， \bar{v} 和 s_v^2 是 v_i 的樣本平均數與樣本變異數， \bar{w} 和 s_w^2 是 w_i 的樣本平均數與樣本變異數；而 v_i 和 w_i 是 x 軸和 y 軸上的一組缺陷區間(defect intervals)。 v_i 和 w_i 的定義如下：

$$v_i = x_{(i)} - x_{(i-1)}, i = 1, 2, \dots, n \quad (9)$$

$$w_i = y_{(i)} - y_{(i-1)}, i = 1, 2, \dots, n \quad (10)$$

其中， $x_{(i)}$ 和 $y_{(i)}$ 是 x 軸和 y 軸上第 i 小的缺陷座標。Jun *et al.*證明當缺陷的分布呈現隨機性時，群聚指標 CI 值會接近1；但假如缺陷的分布呈現群聚現象時，群聚指標 CI 值會大於1。

Jun *et al.*所構建之預測良率的迴歸模型如下：

$$Y^K = \beta_0 + \beta_1 e^{-\lambda} + \beta_2 \sqrt{CI} + \varepsilon \quad (11)$$

其中， λ 是每一晶片上的平均缺陷數， K 、 β_0 、 β_1 以及 β_2 是迴歸模型欲估計的未知參數， ε 為誤差項並假設服從於平均數為0以及變異數為 σ^2 的常態分配。Jun *et al.*將模擬資料透過Box-Cox最佳化程序[9]來估計迴歸模型的四項未知參數 K 、 β_0 、 β_1 以及 β_2 。在估計出迴歸模型的四項未知參數後，就可得到樣本的迴歸方程式，並進一步利用樣本的迴歸方程式來做預測。只要計算出晶圓的 CI 值

及 λ 值後，就可代入樣本的迴歸方程式來預測此晶圓的良率。

但利用迴歸模型來做預測時，必須先假設一個明確的函數式。而此函數式的適配度是值得考慮的。再者，資料是否違反迴歸分析的基本假設亦是個問題。最後，若資料不能能夠適配出一合適的迴歸方程式時，又必須將資料做進一步的轉換來做分析，而這些都造成在做資料分析時不便。

2.1.5 類神經網路良率模式

Tong *et al.*[42]以類神經網路的模糊自適應共振理論網路(Fuzzy ART)修正晶圓上的缺陷群聚現象，直到缺陷分佈符合卜瓦松分配才停止，最後計算缺陷密度，建立以卜瓦松分配為基礎的修正良率模式。此模式在預估良率上比卜瓦松或負二項良率模式更為精確。但由於反覆地對每片晶圓進行群聚分類及測試工作，仍有構建流程複雜之缺失。

李靜宜[1]以卜瓦松良率與Tong *et al.*[43]的一維轉軸群聚指標(CI_R)為輸入，利用倒傳遞網路(Back-Propagation Neural Network, BPNN)構建一個晶圓良率預估模式。此模式能解決傳統卜瓦松良率模式未考慮到缺陷群聚的問題，且不需複雜的統計模式即能精確的預測出晶圓的真實良率。但利用倒傳遞網路構建之良率模式一般必需經由適度地設定許多網路參數(例如，學習速率、隱藏層數、隱藏層單元數以及慣性項)才能得到較佳的預測結果。

Tsai *et al.*[44]考慮缺陷之夾角間距變異與距原點之距離間距變異而建構了一缺陷群聚指標 CI_T 來衡量晶圓上缺陷點的群聚程度；並將 CI_T 代入Jun *et al.*[25]所建構的良率模式來預估晶圓良率。藉由與Jun *et al.*所建構的 CI 良率模式做比較來驗證由 CI_T 所建構的良率模式能精確地預估晶圓良率。但仍有與Jun *et al.*相同之迴歸模型適配度問題以及資料是否違反迴歸模型的基本假設都是值得考慮的。

楊博欽[4]以缺陷總數與Tsai *et al.*[44]的群聚指標(CI_T)為輸入，應用自組性演算法(Group Method of Data Handling, GMDH)構建一個晶圓良率模式。此模式不需要任何統計假設，且可以構建出一個預測良率的數學方程式，實用價值甚高。而在良率預測上也較業界常用的負二項良率模式和利用倒傳遞網路構建之良率模式更為精確。但楊博欽所構建出的良率模式其數學方程式仍有太過冗長的缺失。

2.2 一般迴歸神經網路

一般迴歸神經網路(General Regression Neural Network, GRNN)是 Specht [36] 於 1991 年所提出。GRNN 克服了機率神經網路(Probabilistic Neural Network, PNN)[5]只能處理分類的問題，不但能處理分類的問題，亦具有處理連續變數的能力。GRNN 屬於監督式學習網路，因而適合預測、診斷等應用。

2.2.1 GRNN 網路架構

GRNN 是一個三層的網路，其網路架構如圖 2-1 所示，包括：

- (1) 輸入層：第一層為輸入層。輸入層的輸入單元(input unit)只負責分配變數給第二層樣式層的樣式單元(pattern unit)。當有新的向量變數 X 進入網路後，此新變數會減去原先內部儲存的訓練範例；然後將差量的平方值加總當成樣式層單元的淨輸入。
- (2) 樣式層(隱藏層)：第二層為樣式層。樣式層由訓練樣式(training pattern)的各個樣式所組成。樣式層單元的輸出使用非線性的轉換函數，一般為指數函數。樣式層單元的淨輸入經由非線性的轉換函數後可得樣式層單元的輸出值。該輸出值接著會被傳送到總和單元 (summation unit)。
- (3) 輸出層：第三層為輸出層。總和單元加總來自於樣式層單元的加權後輸出值以形成第一項總和單元輸出值；同時總和單元加總來自於樣式層單元的加權後輸出值與其所對應觀測值 Y 的乘積而形成第二項總和單元輸出值。輸出層單元的輸出值即為第二項總和單元輸出值除以第一項總和單元輸出值，而此亦為所對應觀測值 Y 的估計值。

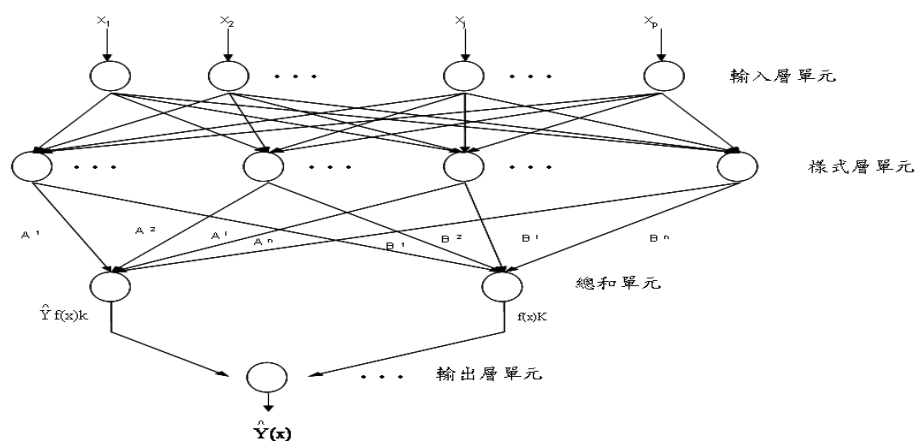


圖 2.1 GRNN 的網路架構

2.2.2 GRNN 網路基本原理

GRNN 與傳統的迴歸模式必須先假設一個明確的函數式不同。GRNN 可利用一個連續型的聯合機率密度函數(joint probability density function)來描述。假設 $f(\mathbf{x}, y)$ 代表一已知的連續型聯合機率密度函數，其中 \mathbf{x} 為一向量隨機變數， y 為一隨機變數。令 \mathbf{X} 為向量隨機變數 \mathbf{x} 的某特定觀測值，則在 $\mathbf{x}=\mathbf{X}$ 的條件下， y 的條件期望值(也可稱為 y 在 \mathbf{X} 上的迴歸)可表示如下：

$$E[y | \mathbf{x} = \mathbf{X}] = \frac{\int_{-\infty}^{\infty} y f(\mathbf{X}, y) dy}{\int_{-\infty}^{\infty} f(\mathbf{X}, y) dy} \quad (12)$$

當 $f(\mathbf{x}, y)$ 未知時，可以觀測值 (\mathbf{x}, y) 來估計。GRNN 利用無母數方法 Parzen[32] window 的方式來估計 $f(\mathbf{x}, y)$ 。其估計量 $\hat{f}(\mathbf{x}, y)$ 可表示如下：

$$\hat{f}(\mathbf{x}, y) = \frac{1}{(2\pi)^{(p+1)/2} \sigma^{(p+1)}} * \frac{1}{n} \sum_{i=1}^n \exp\left[-\frac{(\mathbf{x} - \mathbf{x}^i)^T (\mathbf{x} - \mathbf{x}^i)}{2\sigma^2}\right] * \exp\left[-\frac{(y - y^i)^2}{2\sigma^2}\right] \quad (13)$$

其中 p 是 \mathbf{x} 的維度， σ 是平滑因子參數， \mathbf{x}^i 和 y^i 是一些樣本觀測值 (\mathbf{x}, y) ， n 是樣本數。將式子(13)代入式子(12)可得式子(14)。

$$\hat{E}(y | \mathbf{X}) = \hat{Y}(\mathbf{X}) = \frac{\sum_{i=1}^n \exp\left[-\frac{(\mathbf{X} - \mathbf{x}^i)^T (\mathbf{X} - \mathbf{x}^i)}{2\sigma^2}\right] \int_{-\infty}^{\infty} y \exp\left[-\frac{(y - y^i)^2}{2\sigma^2}\right] dy}{\sum_{i=1}^n \exp\left[-\frac{(\mathbf{X} - \mathbf{x}^i)^T (\mathbf{X} - \mathbf{x}^i)}{2\sigma^2}\right] \int_{-\infty}^{\infty} \exp\left[-\frac{(y - y^i)^2}{2\sigma^2}\right] dy} \quad (14)$$

式子(14)可進一步簡化如下：

$$\hat{Y}(\mathbf{X}) = \frac{\sum_{i=1}^n y^i \exp\left[-\frac{D_i^2}{2\sigma^2}\right]}{\sum_{i=1}^n \exp\left[-\frac{D_i^2}{2\sigma^2}\right]} \quad (15)$$

其中， $D_i^2 = (\mathbf{X} - \mathbf{x}^i)^T (\mathbf{X} - \mathbf{x}^i)$ 。

GRNN 即是以式子(15)來求得觀測值 y 的估計值。

一般 GRNN 以指數分配為轉換函數，其轉換函數可表示如下：

$$f(D_i^2) = \exp\left[-\frac{D_i^2}{2\sigma^2}\right] \quad (16)$$

當一個新的向量變數 \mathbf{X} 進入網路後，會減去原先內部儲存的訓練範例；然後將差量的平方值加總當成樣式層單元的淨輸入。而樣式層單元的淨輸入經由式子

(16)之轉換函數後可得樣式層單元的輸出值。該輸出值接著會被傳送到總和單元。總和單元加總來自於樣式層單元的加權後輸出值以形成第一項總和單元輸出值，第一項總和單元輸出值可表示如下：

$$\sum_{i=1}^n \exp\left[-\frac{D_i^2}{2\sigma^2}\right] \quad (17)$$

。另一方面，總和單元加總來自於樣式層單元的加權後輸出值與其所對應觀測值 y^i 的乘積而形成第二項總和單元輸出值。第二項總和單元輸出值可表示如下：

$$\sum_{i=1}^n y^i \exp\left[-\frac{D_i^2}{2\sigma^2}\right] \quad (18)$$

。輸出層單元的輸出值即為第二項總和單元輸出值除以第一項總和單元輸出值，此同於式子(15)所示的對應觀測值 y 之估計值。

一般類神經網路的績效可以誤差項的均方根(root-mean squared error, RMSE)來衡量。RMSE 可以式子(19)來求得。

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (A_i - O_i)^2}{n}} \quad (19)$$

其中， n 表示訓練樣式的個數， A_i 表示實際的輸出值， O_i 表示實際輸出值的預測值。本研究亦以 RMSE 來衡量各模式預測的績效。

2.2.3 GRNN 的優點

GRNN 有以下的優點：

- (1) GRNN 的網路學習過程為一次學習。因此它主要的優點是網路能夠快速的學習，而且不需擔心類似 BPNN[5][17]過度訓練(overtraining)或訓練不足(undertraining)的問題。
- (2) 針對稀疏資料(sparse data)，當樣本數很大時，GRNN 亦能夠快速收斂到一個最佳的迴歸面。
- (3) 不像 BPNN 必需設定許多網路參數。例如，隱藏層數、隱藏層單元數、學習速率(learning rate)以及慣性項(momentum)。GRNN 唯一要設定的參數是平滑因子(smoothing factor) σ 值。如此便減少了類神經網路所為人詬病的參數設定問題。
- (4) GRNN 並不是利用最陡坡降法(gradient descent)來得到最佳解，因此減少了像

BPNN 可能只尋找到區域最佳解(local optimal)的機會。

因為 GRNN 有上述的優點，因此本研究利用 GRNN 來建構晶圓的良率預估模式。

2.3 缺陷群聚指標

晶圓表面上缺陷群聚的嚴重程度可利用群聚指標來描述。目前較常用來衡量缺陷群聚程度的群聚指標有負二項良率模式的 α 群聚參數、空間統計學的 V/M 群聚指標以及無統計假設之缺陷群聚指標(Cluster Index, CI)，以下將分別說明這些缺陷群聚指標以及一些國內文獻所提出的缺陷群聚指標。

2.3.1 負二項良率模式的 α 群聚參數

Stapper[39]以 Gamma 函數來描述 Murphy[29]良率模式中之缺陷機率密度函數所得的負二項良率模式如式子(6)所示。其中，負二項良率模式之群聚參數 α 可表示如下：

$$\alpha = \frac{\lambda^2}{(\sigma^2 - \lambda)}$$

其中， λ 表每個晶片上的平均缺陷數， σ^2 表晶片上缺陷之變異數。當利用上式求 α 值時， α 值會顯得過於散亂，同時也可能為負值(σ^2 可能小於 λ)，這些都會造成分析良率時的不便。因此，Cunningham[14]建議以式子(20)來計算 α 值。

$$\alpha = \left(\frac{\bar{\lambda}}{\sigma}\right)_{avg}^2 \frac{1}{1 - (\bar{\lambda}_{avg}/\sigma^2)} \quad (20)$$

其中， $\bar{\lambda}_{avg}$ 表平均每片晶圓上之每個晶片上的缺陷密度。

2.3.2 空間統計學之缺陷群聚指標

Tyagi與Bayoumi [45][46]應用空間統計學中之方格法來衡量晶圓上缺陷的群聚現象。方格法是將所研究的區域(在本研究中即為晶圓的表面)劃分成若干個面積相同的方塊，稱為方格(此處的「方格」在本研究中即為晶片)，接著計算出每一方塊內包含的「點」數(此處的「點」在本研究中即為缺陷)。最後經由分析便可得知所研究區域內的「點」是呈現隨機分布或者有群聚現象的發生。如果晶圓上缺陷呈現隨機分布，則表示其遵守卜瓦松分配。由於卜瓦松分配的變異數等於平均數，亦即變異數與平均數的比值(Variance-Mean Ratio, V/M)會等於1；當點出現群聚現象時，則 V/M 值會大於1。此指標的缺點是在相同的點

分布下，可能會因方格選取的方法、形狀及大小的不同，而使得 V/M 值有相當大的差異；再者，隨著晶圓缺陷密度的遞增， V/M 值並沒有呈規則性遞增的特性[45][46]。

2.3.3 無統計假設之缺陷群聚指標

Jun *et al.*[25]提出一個缺陷群聚指標(Cluster Index, CI)。群聚指標 CI 的計算是將一片晶圓上的 n 個缺陷的空間座標分別投影到X軸和Y軸上，如圖2.2所示。由圖2.2可得知，投影後在X軸和Y軸上的座標點同時出現群聚現象才表示該晶圓之缺陷具有群聚現象，如(d)所示；若投影後僅在Y軸或X軸上的座標點出現群聚現象，如(b)、(c)所示；亦或投影後不出現群聚現象，如(a)所示；此兩者之晶圓上的缺陷均不具群聚現象。

CI 可以下式表之：

$$CI = \min\left\{\frac{s_v^2}{\bar{v}}, \frac{s_w^2}{\bar{w}}\right\}$$

其中， \bar{v} 和 s_v^2 是 v_i 的樣本平均數與樣本變異數， \bar{w} 和 s_w^2 是 w_i 的樣本平均數與樣本變異數； v_i 和 w_i 是 x 軸和 y 軸上的一組缺陷區間(defect intervals)。 v_i 和 w_i 的定義如下：

$$v_i = x_{(i)} - x_{(i-1)}, i = 1, 2, \dots, n$$

$$w_i = y_{(i)} - y_{(i-1)}, i = 1, 2, \dots, n$$

其中， $x_{(i)}$ 和 $y_{(i)}$ 是 x 軸和 y 軸上第 i 小的缺陷座標。Jun *et al.*證明當缺陷的分布呈現隨機性時，群聚指標 CI 值會接近1；但假如缺陷的分布呈現群聚現象時，群聚指標 CI 值會大於1。Jun *et al.*透過模擬實驗得到群聚參數 α 值和群聚指標 CI 值之間的相對關係，如表2.3所示。 α 值愈小代表群聚程度愈嚴重，而 CI 值愈大代表群聚程度愈嚴重；與 α 值相較， CI 值有較均勻的數值範圍來量化不同嚴重程度的群聚現象，這是 CI 指標優於 α 指標的地方。然而經由模擬驗證發現， CI 指標針對某些不同之缺陷分佈樣式，其群聚指標 CI 值卻相同[7]。因此 CI 指標針對某些缺陷分佈樣式會有誤判的現象。

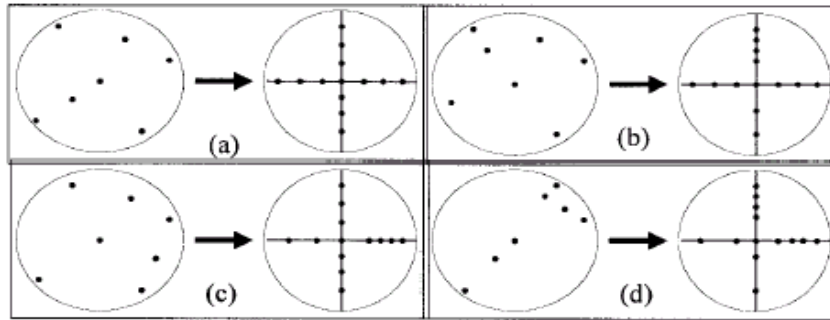


圖 2.2 缺陷分布圖與兩軸的投影

表 2.3 α 值與 CI 值之比較[25]

群聚現象	負二項良率模式的 群聚指標 α	無統計假設的 群聚指標 CI
無	10 或以上	0.85-1.00
些微	4.2	1.00-1.10
中等	3	1.10-1.25
嚴重	1	1.25-2.70

2.3.4 其他缺陷群聚指標

Tong *et al.*[43]將晶圓上的缺陷之二維座標投影至一維座標軸上。利用一維轉軸的方法，對所有的缺陷在某特定角度的一維轉軸上做投影；並藉由計算在該特定角度之軸上點的間距之變異係數的平方而得到一個指標值。然後改變轉軸角度以計算出180個不同轉軸角度下的指標值。最後，計算這180個指標值的平均值，即為一維轉軸群聚指標 CI_R 。經由模擬與實際晶圓資料的驗證， CI_R 指標確實較 CI 指標能更有效地偵測出不同的缺陷群聚程度。但 CI_R 指標仍有無法辨別環狀樣式之缺陷分佈型態的缺失[7]。

許志瑋[3]利用各個缺陷與X軸第一象限之正值夾角的變異係數，以及各個缺陷到座標原點之長度的變異係數，求得此兩變異係數的比值而定義一環狀群聚指標。環狀群聚指標雖可補 CI_R 指標無法辨認環狀缺陷群聚現象的缺失；然而環狀群聚指標只能偵測出環狀的缺陷群聚現象。若只看環狀群聚指標值將無法判斷晶圓上缺陷的群聚程度。

Tsai *et al.*[44]利用各缺陷與X軸第一象限正值夾角之間距的變異係數平方，以及各缺陷距原點座標距離之間距的變異係數平方，求得此兩變異係數平

方的最大值而發展出一缺陷群聚指標 CI_T 。 CI_T 群聚指標值越大，代表群聚的現象越嚴重。 CI_T 能以均勻的數值範圍來量化不同嚴重程度的群聚現象，更能偵測出環狀的缺陷群聚現象。

2.4 缺陷樣式之辨識

針對晶圓缺陷樣式辨識的方法，相關研究主要從統計分析法、啟發式演算法以及人工智慧學習的觀點來分析[31]。統計分析法是假設落於晶圓上的缺陷符合某種機率分配，進而以此分配的特性來進行資料點的群聚現象分析[23]。統計分析法的目的在於如何將空間上的資料點進行切割，以達到樣式分類辨識的目的。啟發式演算法是加入個人的專業知識來設定較佳的參數及演算程序，通常利用柔性計算的方法(例如，基因演算法與模糊理論)來進行樣式辨識。然而基因演算法一般必須付出相當多的努力於適配函數與較佳解的選取評估過程上才能達到最佳解 [7]。而模糊理論的主要限制之一乃是其語意控制規則相當難產生；再者，必須借重一些專家的知識與經驗才能設計出一完善的模糊邏輯控制器 [12]。人工智慧學習是以訓練樣本來建立網路模型，並以測試樣本來檢測所建之網路模型的穩健性。類神經網路即人工智慧學習中一項常用的技術。以類神經網路技術來做樣式辨識時，雖具有容錯能力的優點，但類神經網路技術的主要限制之一乃是其各項網路參數必須適度的設定才能得到較佳的結果 [19]。再者，所選取的樣本需足夠且具代表性則是利用類神經網路技術的關鍵成功因素。

針對晶圓缺陷樣式辨識的方法，已有許多以類神經網路技術的辨識工具被發展出來。其中，倒傳遞網路(Back-Propagation Neural Network, BPNN)是最具代表性且應用最普遍的網路模式[5]。再者，徑向基底函數(Radial Basis Function, RBF)神經網路是一個較新穎但較少被應用的技術，而且能夠避免網路模式訓練時間的冗長，以及增加網路學習的效率，並且已經成功的被應用在許多不同的領域中[5][16]。基於此，本研究利用 BPNN 與 RBF 神經網路所建構之分類模式與本研究所提之多類別 SVM 分類模式做晶圓缺陷樣式辨識準確率的比較。下面介紹 BPNN、RBF 神經網路以及 SVM。

2.4.1 倒傳遞網路(Back-Propagation Neural Network, BPNN) [5][17]

BPNN 是一個三層的網路，其網路架構如圖 2.3 所示，包括：

- (1) 輸入層：輸入層單元被用來表示網路的輸入變數，輸入變數的數目依問題而定，且輸入層單元使用線性轉換函數。
- (2) 隱藏層：隱藏層單元被用來表現與輸入層單元間的交互影響，隱藏層單元的數目通常以試誤法(trial and error)來決定。隱藏層單元使用非線性轉換函數，而最常使用的非線性轉換函數為雙彎曲函數(binary sigmoid function)，且可表示如下：

$$f(x) = \frac{1}{(1 + e^{-x})} \quad (21)$$

- (3) 輸出層：輸出層單元為網路輸出變數之表現。BPNN 利用網路輸出值與實際輸出目標值間的誤差，以回饋方式來調整網路加權值直到最佳狀態為止。輸出層單元的數目是依問題而定，並且使用非線性轉換函數。

BPNN 網路的計算過程分為前向(forward)與後向(backward)兩個階段。首先輸入層單元接收到輸入的樣本資料後直接將它傳給隱藏層單元；隱藏層單元會將加權後的資料加總並透過轉換函數而得隱藏層單元的輸出值；此輸出值進一步被傳給輸出層單元。輸出層單元再以同樣的方法得到最終的網路輸出值，此階段為網路的前向階段。而網路在比較網路輸出值與實際輸出目標值的差量值後，利用此差量值再倒回去進行權重的調整，此為網路的後向階段。

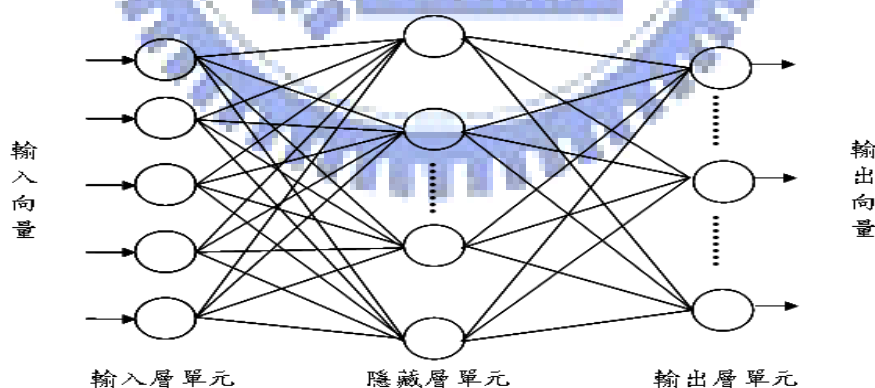


圖 2.3 BPNN 之網路架構

根據上述網路的計算過程，BPNN 網路演算法可描述如下：

- 步驟一：在某特定的範圍內將所有的權重值初始化成一個極小的隨機數值。
- 步驟二：以隨機方式選取一組訓練樣式($\mathbf{x}_p, \mathbf{t}_p$)，再利用前向階段方式計算各

層處理單元所對應的輸出值，並表示如下：

$$O_j^q = f\left(\sum_i O_i^{q-1} w_{ji}^q\right) \quad q = 1, 2, \dots, Q \quad (22)$$

其中， \mathbf{x}_p 表示第 p 個訓練樣式的輸入向量， \mathbf{t}_p 表示第 p 個訓練樣式的目標向量， O_j^q 表示第 q 層之第 j 個處理單元的輸出， f 表示非線性轉換函數，且第一層的輸入可表示如下：

$$O_j^0 = x_j \quad (23)$$

步驟三：利用網路最後一層（輸出層）之處理單元的輸出值 O_j^Q 與其所對應的實際輸出目標值 t_j 來計算輸出層各處理單元的差量值，此差量值如下：

$$\delta_j^Q = f'(H_j^{Q-1})(O_j^Q - t_j) \quad (24)$$

其中， δ_j^Q 表示輸出層（第 Q 層）之第 j 個處理單元的差量值， H_j^{Q-1} 表示第 $(Q-1)$ 層之第 j 個處理單元的輸出值。

步驟四：利用步驟三所計算之輸出層各處理單元的差量值來計算前一層隱藏層之各處理單元的差量值，此差量值如下：

$$\delta_j^{q-1} = f'(H_j^{q-1}) \sum_j \delta_j^q w_{ij}^q \quad q = Q, Q-1, \dots, 2. \quad (25)$$

其中， δ_j^{q-1} 表示隱藏層（第 $q-1$ 層）之第 j 個處理單元的差量值， w_{ij}^q 表示第 q 層之第 j 個處理單元與其前一層之第 i 個處理單元間的權重值。

步驟五：利用步驟三與步驟四所計算之各輸出層與隱藏層處理單元的差量值來調整輸出層與隱藏層處理單元間的權重值，此權重值之計算如下：

$$w_{ij}^{new} = w_{ij}^{old} + \Delta w_{ij}^q \quad q = Q, Q-1, \dots, 2. \quad (26)$$

其中， Δw_{ij}^q 表示第 q 層之第 j 個處理單元與其前一層之第 i 個處理單元間的權重變化值，且可表示如下：

$$\Delta w_{ij}^q = \eta O_i^{q-1} \delta_j^q \quad q = Q, Q-1, \dots, 2. \quad (27)$$

其中， η 表示學習速率。

步驟六：重複步驟二到步驟五直到誤差值之總和達到一可接受的水準為止。

2.4.2 徑向基底函數(Radial Basis Function, RBF)神經網路

RBF神經網路[5][16]是監督式學習(supervised learning)網路模式的一種。RBF神經網路之結構包含三層：輸入層(input layer)、隱藏層(hidden layer)及輸出

層(output layer)，如圖2.4所示。網路的層與層之間有連接鍵以順向(forward)方式相互連接，且各連接鍵皆含有一個權重值來代表輸入訊息的重要程度。在輸入層部分，神經元的數目即為所欲輸入的變數個數；在隱藏層部分，隱藏層數目固定為一層且一般以試誤法來求得神經元的最佳數目；在輸出層部分，神經元的數目即為所欲輸出的變數個數。RBF的學習可分為兩個階段：第一階段為非監督式學習，第二階段為監督式學習。在第一階段會先計算出訓練樣本與各個隱藏層結點中心的距離，並以擁有最小距離的結點做為該訓練樣本所歸屬的隱藏層結點，進而達成將訓練樣本聚類到隱藏層結點的目的。第二階段主要的目的是將隱藏層結點的輸出值進行線性轉換，進而產生輸出層結點的輸出值，藉此達成將訓練樣本對應到所屬輸出值的目的。

RBF神經網路之隱藏層結點被稱為kernel，並使用一組被稱為基底函數(basis function)的鐘型對稱函數作為轉換函數。最常被使用的基底函數為高斯核函數(Gaussian kernel function)，其基本型式可表示如下：

$$h_j = \exp\left[-\frac{(\mathbf{x}_p - \mathbf{c}_j)^T(\mathbf{x}_p - \mathbf{c}_j)}{2\sigma_j}\right] \quad j = 1, 2, \dots, N \quad (28)$$

其中， h_j 為第 j 個隱藏層結點之輸出值， \mathbf{x}_p 為輸入向量， \mathbf{c}_j 為第 j 個隱藏層結點的中心， σ_j 則為第 j 個隱藏層結點的平滑參數(smoothing parameter)， N 代表隱藏層結點的個數。由於高斯核函數的鐘型對稱特性，因此當輸入向量越靠近隱藏層結點的中心時，經高斯核函數轉換出來的值就愈大；反之則愈小。而隱藏層結點的平滑參數 σ 乃用來表示在歸屬於該隱藏層結點訓練樣本的分散狀況。

RBF神經網路之輸出層結點使用通用差距法則(general delta rule)[5]來調整連接鍵的權重值，進而求得輸出層結點的輸出值 o_k 。 o_k 可表示如下：

$$o_k = \sum_{j=1}^N (w_{jk} \times h_j) \quad (29)$$

其中， o_k 為第 k 個輸出層結點的輸出值， w_{jk} 為第 j 個隱藏層結點與第 k 個輸出層結點的連接鍵權重值， h_j 為第 j 個隱藏層結點之輸出值， N 代表隱藏層結點的個數。由式子(29)可知，輸出層結點的輸出值是由隱藏層與輸出層之間的連接鍵權重值與隱藏層結點之輸出值進行線性組合而得，並非透過非線性的轉換函數而得。

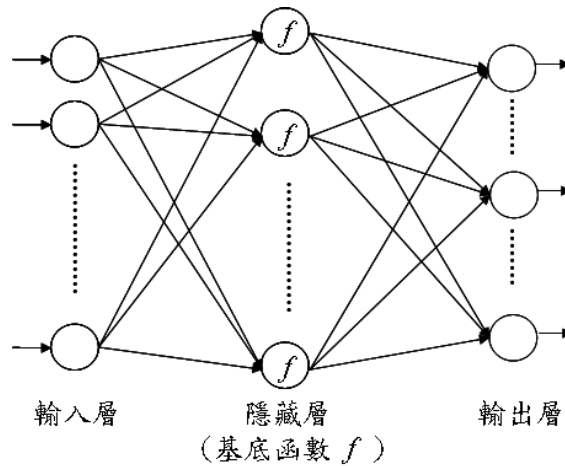


圖 2.4 RBF 神經網路

2.4.3 支撐向量機(Support Vector Machines, SVM)

SVM 的理論基礎來自於 Vapnik[13]等學者所提出的統計學習理論 (Statistical Learning Theory, SLT), SLT 是希望在學習上得到最佳的一般化 (generalization)能力。最佳的一般化能力是希望 SVM 能經由樣本資料找出資料間的規則, 進而學習出較平坦的曲線, 最後再透過資料所呈現的規則性來對未來可能發生的數據進行分類或迴歸預測。亦即, SVM 是以 SLT 為基礎而發展出來的機器學習系統。

2.4.3.1 支撐向量分類(Support Vector Classify)

針對支撐向量分類, Vapnik [13]提出分類最佳超平面(optimal hyperplane)的觀念。亦即假設一組訓練集如下:

$$(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_m, y_m), \quad \mathbf{x}_i \in R^n, y_i \in \{+1, -1\}。$$

此訓練集可被一個超平面區分成兩類。而超平面方程式可用兩個向量內積(dot product)的方式來呈現, 如下所示:

$$\langle \mathbf{w}, \mathbf{x} \rangle + b = 0 \tag{30}$$

其中, \langle, \rangle 為內積運算符號; \mathbf{w} 為一組與訓練樣式向量 \mathbf{x} 對應的權重向量; b 為一門閾值(threshold)。若一個超平面能正確地分類一組訓練集, 並且擁有離超平面最近的向量 \mathbf{x}_i 距離此超平面間的距離最大的特性者, 即為分類該組訓練集之最佳超平面, 如圖 2.5 所示。

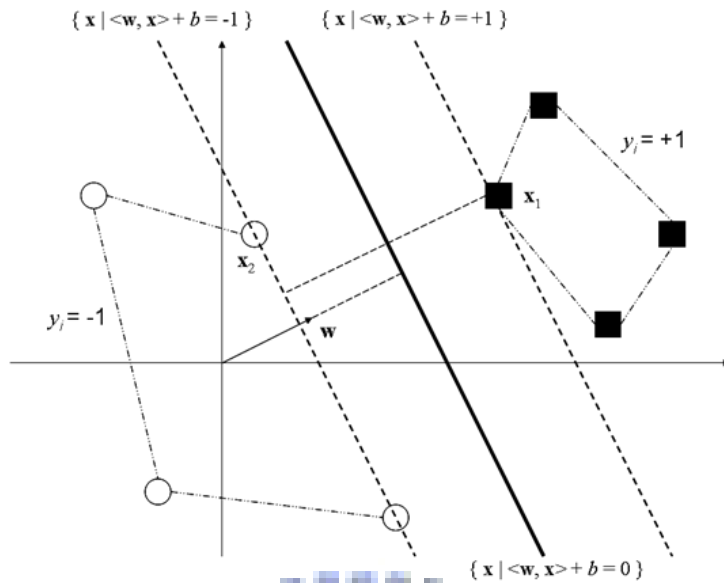


圖 2.5 最佳超平面[35]

由圖 2.5 可知，將 \mathbf{w} 與 b 在尺度上做適度的調整可使得距離超平面最近的訓練向量 \mathbf{x}_1 與 \mathbf{x}_2 滿足下列式子：

$$\langle \mathbf{w}, \mathbf{x}_1 \rangle + b = +1 \quad (31)$$

$$\langle \mathbf{w}, \mathbf{x}_2 \rangle + b = -1 \quad (32)$$

由式子(31)與(32)可得下式：

$$\left\langle \frac{\mathbf{w}}{\|\mathbf{w}\|}, (\mathbf{x}_1 - \mathbf{x}_2) \right\rangle = \frac{2}{\|\mathbf{w}\|} \quad (33)$$

其中， $\|\mathbf{w}\|$ 為權重向量 \mathbf{w} 的長度。由式子(33)可得知，離最佳超平面最近的向量與此超平面之間的距離為 $\frac{1}{\|\mathbf{w}\|}$ ，一般稱為差數(margin)。為了求得最佳超平面(即

最大化差數)，可藉由解下列二次規劃問題來達成。

$$\text{Min. } \tau(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|^2, \quad (34)$$

$$\text{s.t. } y_i (\langle \mathbf{x}_i, \mathbf{w} \rangle + b) \geq 1, \quad i = 1, \dots, m \quad (35)$$

此二次規劃問題可藉由解其對偶問題(dual problem)所形成的 Lagrangian 目標式來達成。其中，Lagrangian 目標式如下所示：

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_1^m \alpha_i (y_i (\langle \mathbf{x}_i, \mathbf{w} \rangle + b) - 1), \quad (36)$$

其中， α_i 為 Lagrange 乘數(Lagrange multipliers)且 $\alpha_i \geq 0$ 。為最大化 Lagrangian

目標式，可將式子(36)分別對 \mathbf{w} 與 b 做偏微分進而推導得到下列式子：

$$\sum_1^m \alpha_i y_i = 0, \quad (37)$$

$$\mathbf{w} = \sum_1^m \alpha_i y_i \mathbf{x}_i, \quad (38)$$

由式子(38)可知最佳超平面的向量解 \mathbf{w} 可由該組訓練集展開而得。其中，擁有 $\alpha_i > 0$ 的訓練樣式向量 \mathbf{x}_i 即稱為支撐向量。將式子(37)與式子(38)代入式子(36)可得對偶問題的最大化目標函式以及限制式如下：

$$Max. \quad W(\boldsymbol{\alpha}) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m \alpha_i \alpha_j y_i y_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle, \quad (39)$$

$$s.t. \quad \alpha_i \geq 0, \quad i = 1, \dots, m \quad (40)$$

$$\sum_1^m \alpha_i y_i = 0, \quad (41)$$

一旦求得目標函式的最佳解 α_i ，即可求得最佳超平面。而最佳超平面的決策函數可表示如下：

$$f(x) = \langle \sum_1^m \alpha_i y_i \mathbf{x}_i, \mathbf{x} \rangle + b \quad (42)$$

若進一步考量訓練範例違反式子(35)的可能性，則可加入寬鬆變數(slack variable, $\xi_i \geq 0$)來放寬限制式，並藉由解下列二次規劃問題來求得最佳超平面。

$$Min. \quad \tau(\mathbf{w}, \boldsymbol{\xi}) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^m \xi_i, \quad (43)$$

$$s.t. \quad y_i (\langle \mathbf{x}_i, \mathbf{w} \rangle + b) \geq 1 - \xi_i, \quad i = 1, \dots, m \quad (44)$$

$$\xi_i \geq 0 \quad (45)$$

其中， C 為大於零的常數，且被稱為誤差項的懲罰參數； C 決定最大化差數(margin)與最小化訓練誤差間的取捨。

2.4.3.2 核函數 (Kernel Function)

前一小節所介紹的支撐向量分類是利用SVM的概念來解決線性分類問題。然而，有些分類問題並非屬於線性可分問題，此時就必須利用核函數將訓練樣式向量 \mathbf{x} 映射到一個高維度的特徵空間(feature space)中，於特徵空間建構最佳超平面並將問題做線性的分類。目前常用的核函數為多項式(polynomial)核函數、徑

向基底(radial basis)核函數以及類神經網路(neural network)核函數[26]，如下所示：

$$(1) \text{多項式核函數：} K(\mathbf{x}, \mathbf{x}_i) = \langle \mathbf{x}, \mathbf{x}_i \rangle^d, \quad (46)$$

其中， d 表示多項式的次方。

$$(2) \text{徑向基底核函數：} K(\mathbf{x}, \mathbf{x}_i) = \exp(-\gamma \|\mathbf{x} - \mathbf{x}_i\|^2), \quad (47)$$

其中， γ 為參數且 $\gamma > 0$ 。

$$(3) \text{類神經網路核函數：} K(\mathbf{x}, \mathbf{x}_i) = \tanh(\kappa \langle \mathbf{x}, \mathbf{x}_i \rangle + \Theta), \quad (48)$$

其中， κ 為參數且 $\kappa > 0$ ； Θ 為水平位移且 $\Theta \in R$ 。

選擇適當的核函數 $K(\mathbf{x}, \mathbf{x}_i)$ 即可在特徵空間中訓練 SVM，並且求得最佳超平面的決策函數，如下所示：

$$f(x) = \sum_1^m \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) + b \quad (49)$$

針對核函數，通常無法以自動的方式來選取並調整其對應的參數。因此必須嘗試不同的核函數，並調整其對應的參數來得到最佳的結果。針對非線性問題，由於透過徑向基底核函數轉換所得的實驗結果很好。因此，非線性徑向基底核函數是最常被使用的核函數[26]；本研究亦以徑向基底核函數為 SVM 的核函數。當使用徑向基底核函數的 SVM 技術來做分類時，懲罰參數 C 與參數 γ 為最重要的參數。利用格子搜尋(grid search)法可於交叉驗證(cross validation)中選取最佳的 (C, γ) 參數組合，再將此參數組合用於訓練與測試資料中來得到最佳的結果 [11]。

2.4.4 多類別支撐向量分類

至目前為止所介紹的SVM是用於處理兩類別的分類問題。但許多實際問題上所需處理的是多類別的分類問題。由於本研究乃藉由多類別SVM來進行多類別晶圓缺陷樣式的分類，在此即介紹幾種多類別支撐向量分類法。

2.4.4.1 一對多(One Versus the Rest)分類法

一對多分類法[8]在處理類別問題時，會產生 m 個兩類別的SVM，其中第 i 個SVM的產生方式，是給定第 i 類資料的標註為 $\{+1\}$ ，而其他類別的所有資料的標註則為 $\{-1\}$ 。解一對多分類法之最佳化問題後可得到 m 個決策函數。而未知類別資料 \mathbf{x} 則被歸於具有最大決策函數值之類別。

2.4.4.2 一對一(Pairwise)分類法

一對一分類法[27]在處理 m 類別問題時，每兩類資料都會產生一個SVM。解一對一分類法之最佳化問題後可得到 $\frac{m(m-1)}{2}$ 個決策函數。而在決定未知類別資料 x 時，一對一分類法採用投票(vote)策略[35]。最後將未知類別資料 x 歸於得票數最高的類別；但若發生兩類別票數相等時，則任意選擇其中一類別為最後結果。

2.4.4.3 DAG(Directed Acyclic Graph)分類法

DAG分類法是學者Taylor所提出[33]。此方法在訓練階段與一對一分類法相同。不同處在於測試階段，DAG分類法建立一個二元且無循環的有向圖來做分類；其中包含 $\frac{m(m-1)}{2}$ 個內部節點(internal node)和 m 個葉節點(leaf node)，如圖2.6所示。每一個內部節點都表示一個SVM， m 個葉節點表示類別。而未知類別資料 x 乃由根節點(root node)開始測試，再依據該SVM來決定下一層應歸於左邊或右邊的內部節點，而最後到達的葉節點就是 x 的類別。

2.4.4.4 一次(Considering all Data at a Once)分類法

Vapnik[47]認為處理多類別的分類問題時，可將問題合併成單一的最佳化問題而一次求解，不用處理多個最佳化問題。一次分類法在處理 m 個類別資料時，會產生 m 個決策函數。其中第 i 個決策函數，可以區分類別為 i 的資料和其它所有類別的資料。然而決策函數並非由 m 個最佳化問題各別求得，而是結合成一個最佳化問題一次求解。在決定未知類別資料 x 時，一次分類法利用決策函數產生 m 個值， x 則被歸於具有最大決策函數值之類別。

由於一對一分類法的訓練時間較短[22]，所以本研究所採用之多類別SVM是利用一對一分類法的方式來進行多類別晶圓缺陷樣式的分類。

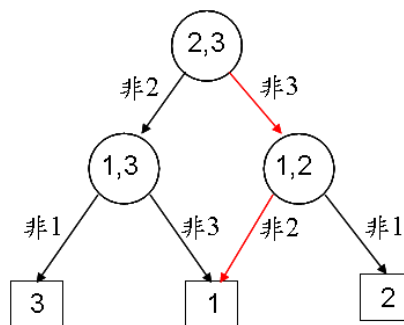


圖 2.6 DAG 分類法處理 3 類別的分類

第三章 研究方法

本研究之主要目的是要發展一個整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統。本研究利用一般迴歸神經網路(GRNN)來構建良率模式，並利用多類別支撐向量機(Multi-class SVM)來構建缺陷樣式辨識系統；最後整合成一個良率預估模式及缺陷樣式辨識之晶圓缺陷診斷系統。本研究以模擬之晶圓缺陷實驗來說明本研究所提之整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統的可行性；並進一步與中外文獻所提之良率預估模式與缺陷樣式辨識系統做比較以驗證本研究的有效性與優越性。

本章共分為四小節，第一小節介紹本研究之特徵因子的選取，第二小節介紹本研究構建之良率預估模式，第三小節介紹本研究構建之晶圓缺點樣式辨識系統，第四小節介紹本研究構建之整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統。

3.1 本研究之特徵因子的選取

由於本研究是要發展一個整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統。因此，當一片晶圓出現中、低良率時，應進一步診斷該晶圓表面的缺陷是否存在某種特定的缺陷樣式。為利用 GRNN 來構建良率模式，並利用多類別 SVM 來構建缺陷樣式辨識系統，本研究乃以對晶圓良率有所影響的因素做為本研究之特徵因子。

一般的良率模式為單位面積的缺陷密度與晶片面積的函數[14]，其公式可表示如下：

$$Y = f(D, A, K) \quad (50)$$

其中， D 表示單位面積的缺陷密度， A 表示晶片面積， K 為晶片面積 A 的修正因子。由式子(50)可知，單位面積的缺陷密度 D 為影響良率的因子之一，同時可以用來描述晶圓上缺陷的群聚程度。因此，本研究擷取單位面積的缺陷密度 D 做為本研究之特徵因子。

此外，本研究利用各個缺陷與 X 軸第一象限正值夾角的變異 CV_A ，以及各個缺陷距原點座標距離的變異 CV_D 做為本研究之特徵因子，圖 3.1 為上述夾角與

距離的示意圖。其中 CV_A 與 CV_D 的公式推導如下：

步驟一：計算晶圓上各缺陷座標與第一象限的 X 軸所成的正值夾角 θ_i ，且公式可表示如下：

$$\theta_i = \tan^{-1}\left(\frac{Y_i}{X_i}\right), i=1,2,\dots,n \quad (51)$$

其中， X_i 為第 i 個缺陷的 X 軸座標， Y_i 為第 i 個缺陷的 Y 軸座標。再將 θ_i 做遞增排序後可得 $\theta_{(i)}$ 。 $\theta_{(i)}$ 表示排序後第 i 小的夾角，而夾角的間距 A_i 可表示如下：

$$A_i = \theta_{(i)} - \theta_{(i-1)}, i=1,2,\dots,n \quad (52)$$

其中 $\theta_{(0)}$ 為 0。

步驟二：計算晶圓上各缺陷座標到原點的距離 L_i ，且公式可表示如下：

$$L_i = \sqrt{X_i^2 + Y_i^2}, i=1,2,\dots,n \quad (53)$$

其中， X_i 為第 i 個缺陷的 X 軸座標， Y_i 為第 i 個缺陷的 Y 軸座標。再將 L_i 做遞增排序後可得 $L_{(i)}$ 。 $L_{(i)}$ 表示排序後第 i 小的距離，而距離的間距 D_i 可表示如下：

$$D_i = L_{(i)} - L_{(i-1)}, i=1,2,\dots,n \quad (54)$$

其中 $L_{(0)}$ 為 0。

步驟三：計算 CV_A 與 CV_D ，且公式如下所示：

$$CV_A = \frac{S_A}{\bar{A}} \quad (55)$$

$$CV_D = \frac{S_D}{\bar{D}} \quad (56)$$

其中， \bar{A} 和 S_A 表示 A_i 的樣本平均數與樣本標準差； \bar{D} 和 S_D 表示 D_i 的樣本平均數與樣本標準差。

當缺陷分佈呈現隨機分佈時，不論是夾角的間距或是距離的間距都會比較一致，也就是說間距的變異會較小；當缺陷分佈有群聚現象時，不論是何種群聚樣式都會造成至少一種間距的變異會增大。因此只要其中一種間距的變異大，便認為此晶圓的缺陷呈現群聚現象。由此可知，夾角的變異 CV_A 與距離的變異 CV_D 皆能反應缺陷的群聚現象，並進而可能影響晶圓良率。基於此，本研究擷取各個缺陷與 X 軸第一象限正值夾角的變異 CV_A ，以及各個缺陷距原點座標距離的變

異 CV_D 做為本研究之特徵因子。

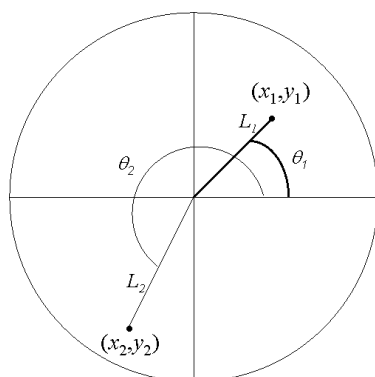


圖 3.1 夾角與距離示意圖

最後，本研究提出一個新的指標來描述缺陷的群聚現象，並做為本研究之特徵因子。以下即說明本研究所發展之群聚指標的建立過程及其優點。

本研究利用晶圓上缺陷的群聚分佈情形發展出新的缺陷點群聚指標 CI_E 。 CI_E 的計算公式如下：

$$CI_E(p_1, p_2, \dots, p_s) = \sum_{i=1}^s (p_i \times \log_2(\frac{1}{p_i})) \quad (57)$$

其中， s 表示缺陷的群聚數， p_i 表示第 i 群聚內的缺陷數佔所有缺陷數的比例。圖3.2為本研究之缺陷群聚指標 CI_E 示意圖。由圖3.2可知缺陷之群聚程度愈嚴重，則 CI_E 值愈大，如3.2-(a)所示；反之，隨著缺陷之群聚程度趨緩，則 CI_E 值逐漸變小，如3.2-(b)(c)所示。由此可知，本研究所發展之群聚指標 CI_E 具有可準確偵測出各種嚴重程度之缺陷群聚現象的優點。由於本研究所發展之新群聚指標 CI_E 可準確偵測出各種嚴重程度之缺陷群聚現象。基於此，本研究擷取新群聚指標 CI_E 做為本研究之特徵因子。

本研究利用上述所擷取的4個晶圓特徵因子：單位面積的缺陷密度 D 、夾角的變異 CV_A 、距離的變異 CV_D 以及新群聚指標 CI_E 當作輸入，並利用GRNN技術來構建良率模式，再配合多類別SVM技術來構建缺陷樣式辨識系統；最後整合成一個良率預估模式及缺陷樣式辨識之晶圓缺陷診斷系統以準確地估計晶圓良率及辨識晶圓缺陷樣式，提供工程人員改善晶圓品質之資訊。

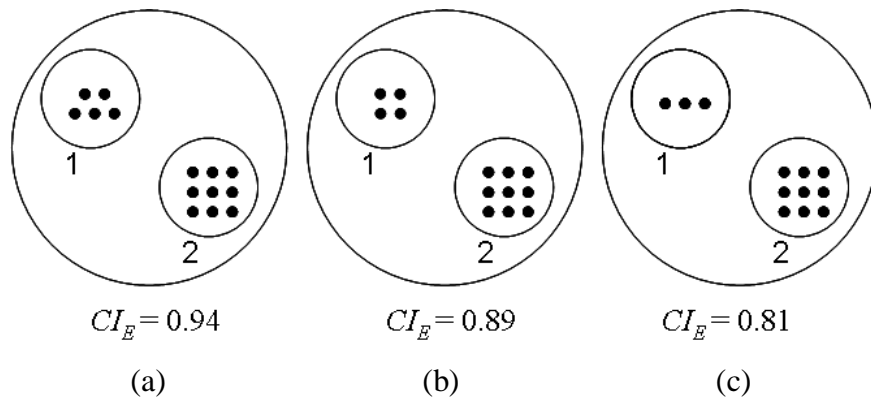


圖 3.2 群聚指標示意圖

3.2 本研究構建之良率模式

本研究是利用一般迴歸神經網路(GRNN)來建構良率模式，其建構流程如下：

步驟一：本研究利用程式語言 Borland Delphi 來模擬 8 吋晶圓的表面缺陷樣式，進而得到每一片晶圓之缺陷總數，以及缺陷的座標。

步驟二：利用步驟一所得到的晶圓缺陷數以及缺陷座標計算出 GRNN 網路輸入的 4 個晶圓特徵因子：單位面積的缺陷密度 D 、夾角的變異 CV_A 、距離的變異 CV_D 以及新群聚指標 CI_E 。

步驟三：利用步驟二所得到的 4 個晶圓特徵因子當輸入，以晶圓上完全沒有缺陷的晶片數除以晶圓上總晶片數所得的真實良率預估值(以下簡稱真實良率)當輸出，透過 GRNN 來建構良率模式。本研究以類神經網路套裝軟體 Neuroshell® 2 來訓練及測試 GRNN 網路，並利用訓練完成的網路來預估晶圓的良率。圖 3.3 為本研究構建之 GRNN 良率模式架構圖。

步驟四：利用本研究之 GRNN 良率模式與其他中、外學者提出的良率模式做比較。並利用各良率模式之預估良率分別與真實良率做比較；且以誤差項的均方根(RMSE)來比較上述良率預估模式的準確性。

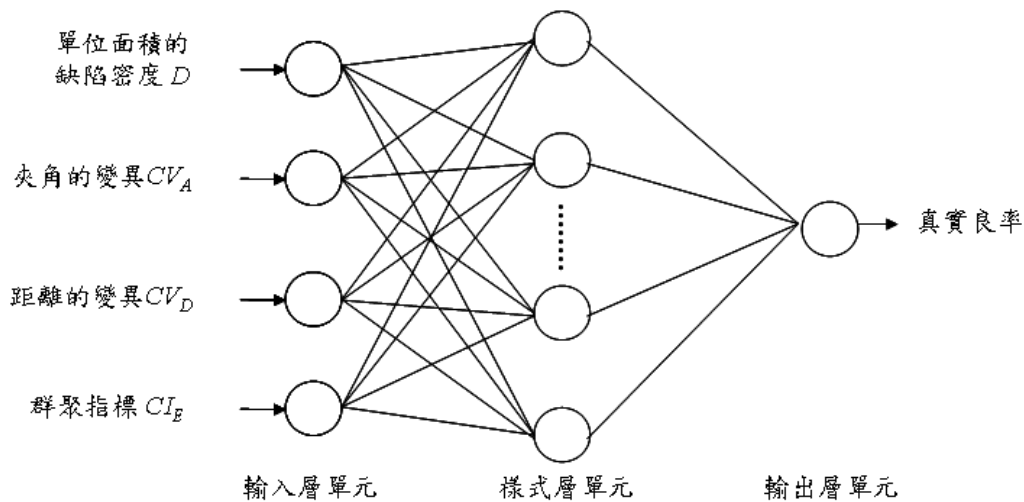


圖 3.3 本研究之 GRNN 良率模式架構圖

3.3 本研究構建之晶圓缺陷樣式辨識系統

晶圓表面上的缺陷呈現群聚現象是影響晶圓良率的主要因素之一。晶圓表面上缺陷的分佈除了呈現隨機(Random)分佈外，如圖 3.4-(a)所示；常見的缺陷群聚樣式主要有牛眼樣式(Bull's Eye Pattern)、弦月樣式(Crescent Moon Pattern)、底部樣式(Bottom Pattern)以及環狀樣式(Edge Pattern)，如圖 3.4-(b)~(e)所示[20]。當晶圓的良率呈現中、低良率時，有必要分析晶圓上的缺陷分佈是否呈現特定樣式，藉由缺陷樣式的辨識以做為製程變異之調整依據。此處所謂之中、低良率乃詢問 IC 製造業之製程工程師對晶圓良率於業界的一般性定義。一般定義良率在 70% 左右為中良率，而 50% 以下為低良率。本研究乃利用多類別 SVM 建構一個晶圓缺陷樣式辨識系統，以進行圖 3.4 中之缺陷樣式的辨識。其建構流程如下：

步驟一：本研究利用程式語言 Borland Delphi 來模擬 8 吋晶圓的表面缺陷樣式，進而得到每一片晶圓之缺陷總數，以及缺陷的座標。

步驟二：利用步驟一所得到的晶圓缺陷數以及缺陷座標計算出多類別 SVM 輸入的 4 個晶圓特徵因子：單位面積的缺陷密度 D 、夾角的變異 CV_A 、距離的變異 CV_D 以及新群聚指標 CI_E 。

步驟三：利用步驟二所得到的 4 個晶圓特徵因子當輸入，並以欲辨識的 5 種對應之晶圓缺陷樣式類別當作輸出，再配合多類別 SVM 來建構一個晶圓缺陷樣式辨識系統，以進行圖 3.4 中之 5 種缺陷樣式的辨識。本研究以 LIBSVM 軟體[11][22]

來訓練及測試多類別 SVM，並利用訓練完成的多類別 SVM 來對未來可能發生的晶圓缺陷樣式進行辨識分類，並進一步診斷以做為製程變異的調整依據。圖 3.5 為本研究所提的晶圓缺陷樣式辨識系統的架構圖。

步驟四：利用本研究之多類別 SVM 分類模式與利用 BPNN 網路以及 RBF 神經網路技術所建構之分類模式做晶圓缺陷樣式辨識準確率的比較。並以在相同的測試樣本中能夠被正確辨識其所對應的晶圓缺陷樣式者，亦即有較高的正確辨識率者，為較能夠正確地來辨識晶圓表面的缺陷樣式的辨識系統。

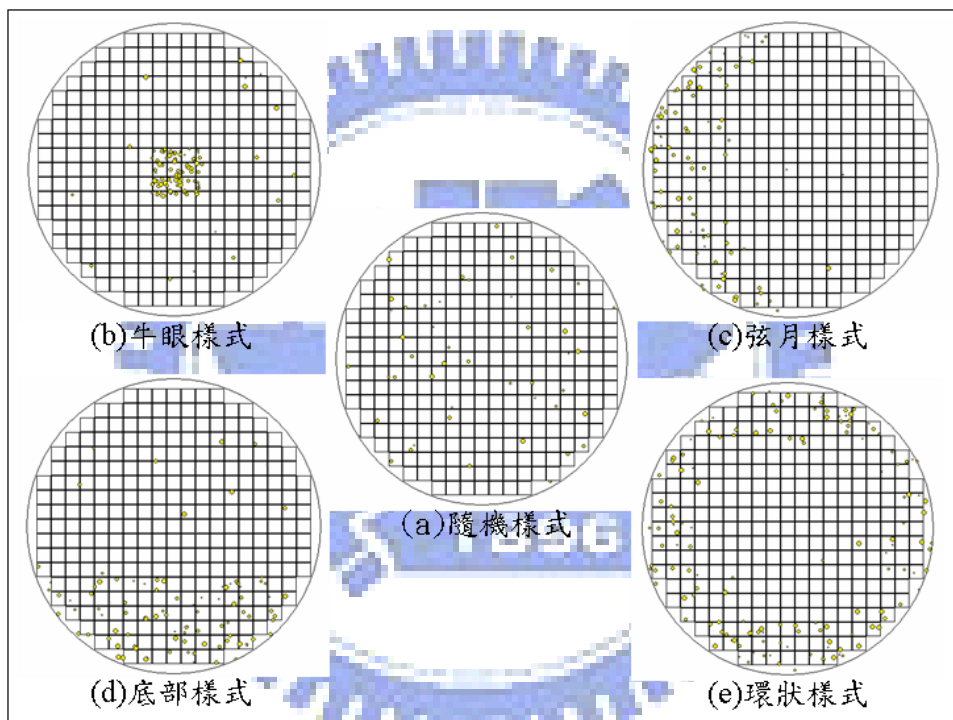


圖 3.4 常見的晶圓表面缺陷樣式



圖 3.5 本研究之多類別 SVM 晶圓缺陷樣式辨識系統架構圖

3.4 整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統

為適切地反應出產品的真實良率，並且能夠針對呈現特定樣式的晶圓做出正確的製程異常診斷。當一片晶圓出現中、低良率時，應進一步診斷該晶圓表面的缺陷是否存在某種特定的缺陷樣式。本研究整合 GRNN 良率預估模式與多類別 SVM 缺陷樣式辨識系統，以進行晶圓缺陷診斷。其建構流程如下：

步驟一：本研究利用程式語言 Borland Delphi 來模擬 8 吋晶圓的表面缺陷樣式，進而得到每一片晶圓之缺陷總數，以及缺陷的座標。

步驟二：利用步驟一所得到的晶圓缺陷數以及缺陷座標計算出 4 個晶圓特徵因子：單位面積的缺陷密度 D 、夾角的變異 CV_A 、距離的變異 CV_D 以及新群聚指標 CI_E 。

步驟三：利用步驟二所得到的 4 個晶圓特徵因子當輸入，以晶圓上完全沒有缺陷的晶片數除以晶圓上總晶片數所得的真實良率當輸出，透過 GRNN 來建構良率模式。本研究以類神經網路套裝軟體 Neuroshell® 2 來訓練及測試 GRNN 網路，並利用訓練完成的網路來預估晶圓的良率。

步驟四：利用步驟二所得到的 4 個晶圓特徵因子當輸入，並以欲辨識的 5 種對應之晶圓缺陷樣式類別當作輸出，再配合多類別 SVM 來建構一個晶圓缺陷樣式辨識系統。本研究以 LIBSVM 軟體來訓練及測試多類別 SVM，並利用訓練完成的多類別 SVM 來對未來可能發生的晶圓缺陷樣式進行辨識分類，並進一步診斷以做為製程變異的調整依據。

步驟五：若經由步驟三之已訓練完成的 GRNN 良率模式預估的晶圓是呈現中、低良率時，則將該晶圓的 4 個特徵因子當輸入，並利用步驟四之已訓練完成的多類別 SVM 來對該晶圓之缺陷樣式進行辨識分類及診斷。圖 3.6 為本研究之整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統的構建流程圖。

第四章 模擬驗證

本研究以一個模擬實驗來說明本研究所提之整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統的可行性與有效性。

4.1 模擬實驗

晶圓表面上的缺陷呈現群聚現象是影響晶圓良率的主要因素之一。晶圓表面上缺陷的分佈除了呈現隨機樣式外，常見的缺陷群聚樣式主要有牛眼、弦月、底部以及環狀樣式[20]。本研究根據上述所提之隨機樣式以及 4 種缺陷群聚樣式，以電腦模擬實驗來模擬 8 吋晶圓的 5 種缺陷分佈樣式。為探討在 5 種缺陷樣式下本研究之晶圓缺陷診斷系統與缺陷數、缺陷群聚程度以及群聚分佈面積之間的關係；本研究利用 3 個因子來設計模擬實驗。此三因子如下所述：

- (1)缺陷數：是指分布於晶圓表面的缺陷數目；其中缺陷數包含 5 種水準，即 25、50、100、200 以及 300。
- (2)群聚程度：是指晶圓表面上各種缺陷群聚樣式內的缺陷佔總缺陷數的百分比；其中缺陷群聚程度包含 4 種水準，即 80%、85%、90% 以及 95%；而其餘非群聚部份的缺陷則隨機分布於該晶圓表面。
- (3)群聚分佈面積：是指晶圓表面上缺陷可能群聚分佈的面積大小；其中群聚分佈面積包含 3 種水準，即 25、49 以及 81 平方公分。

本研究由上述所介紹之 5 種常見的晶圓表面缺陷樣式來建立一電腦模擬實驗。本模擬實驗以程式語言 Borland Delphi 來模擬 8 吋晶圓的表面缺陷樣式，並利用 3 個因子來設計晶圓樣式。圖 4.1 為本研究設計的部份晶圓缺陷樣式。其中，隨機樣式不考慮缺陷群聚程度與群聚分佈面積，且此處只列出缺陷數為 50 以及 100 的晶圓缺陷樣式各 3 片，如圖 4.1-(a)所示；其餘 4 種缺陷群聚樣式皆只列出缺陷數為 50 與 100，以及缺陷群聚程度為 90%的晶圓缺陷樣式各 3 片，如圖 4.1-(b)~(e)所示。本實驗除隨機樣式考慮缺陷數因子之 5 種水準組合外，其餘 4 種缺陷群聚樣式各有 60 種水準組合。在每一水準組合下重複試驗 5 次，因此共有 1225 次試驗，亦即可得 1225 片晶圓缺陷樣式。

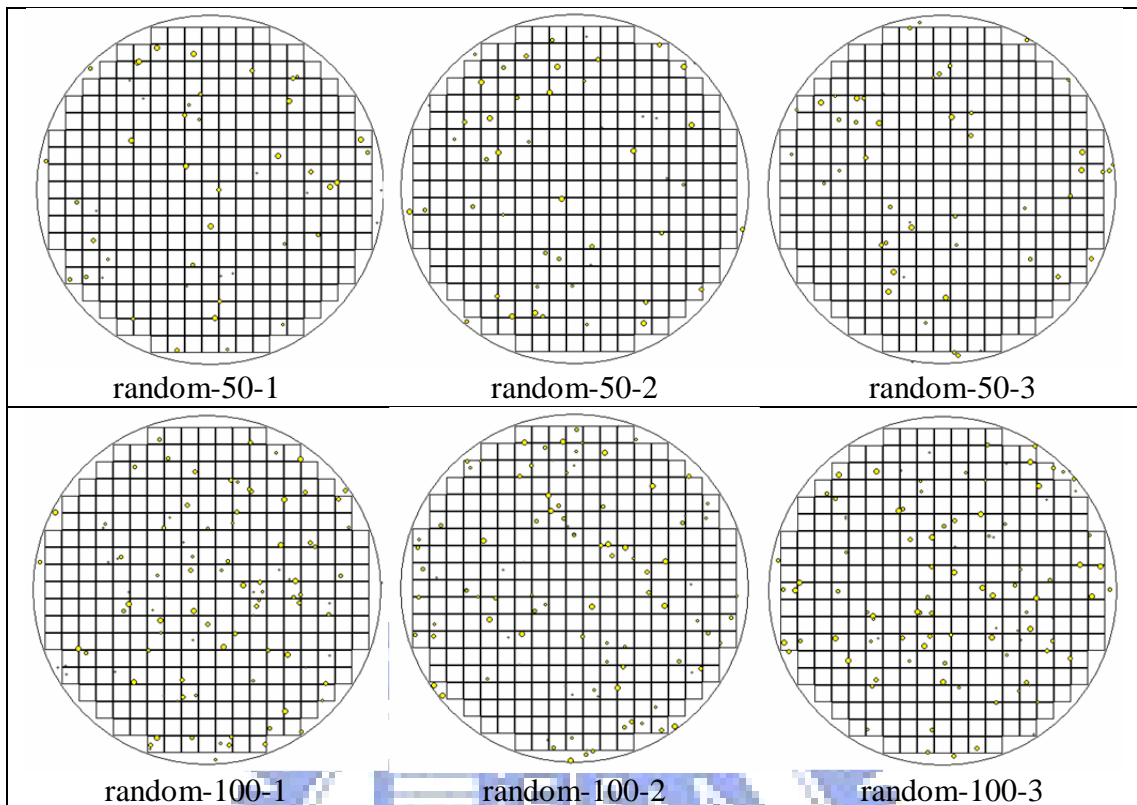


圖 4.1- (a) 隨機樣式

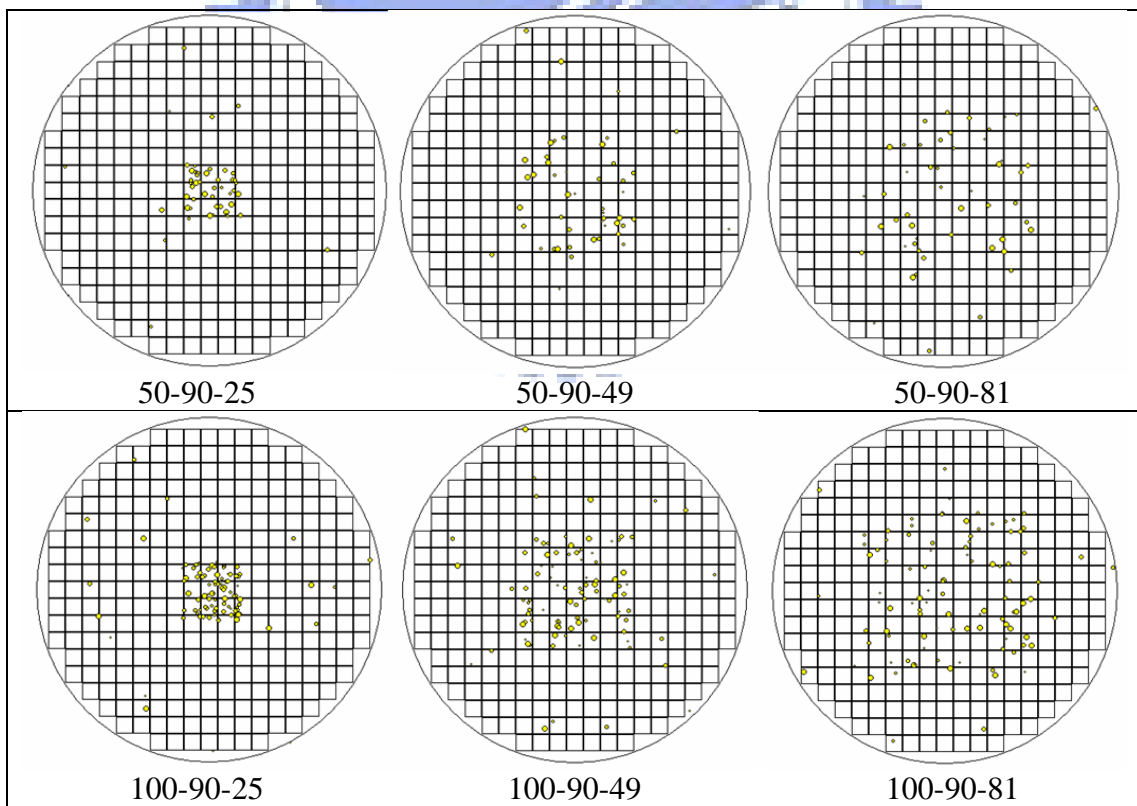


圖 4.1- (b) 牛眼樣式

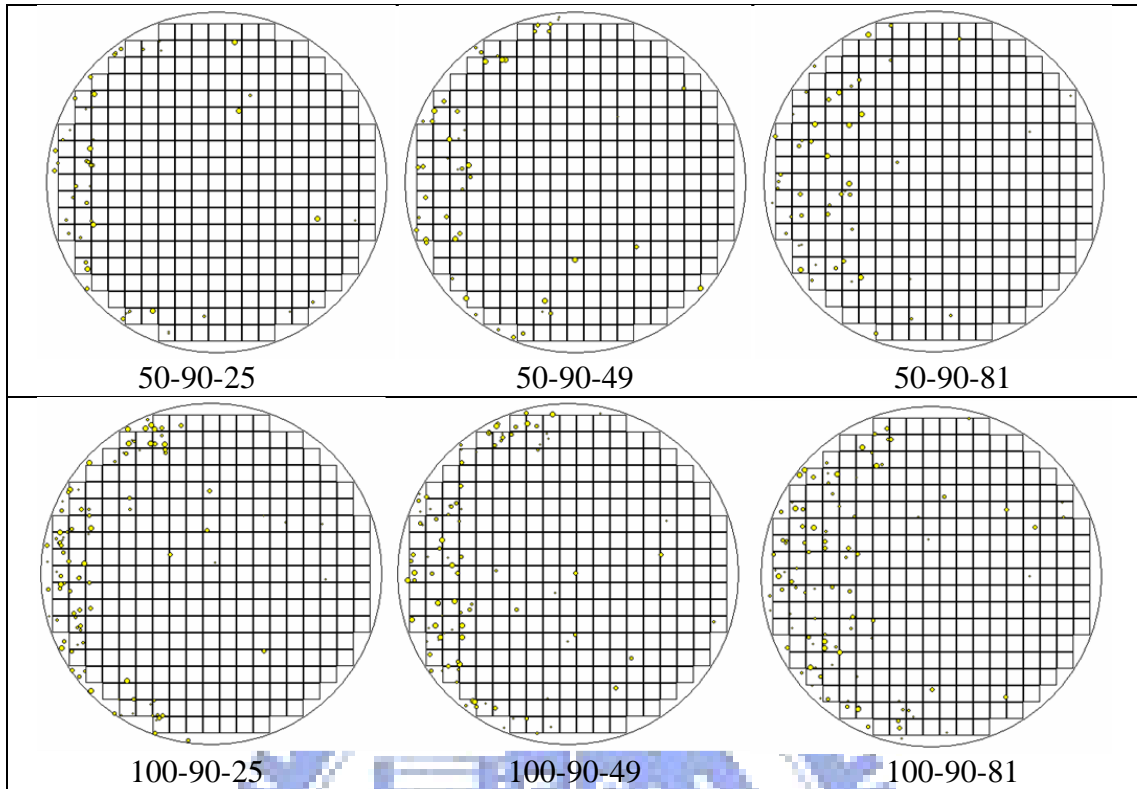


圖 4.1-(c) 弦月樣式

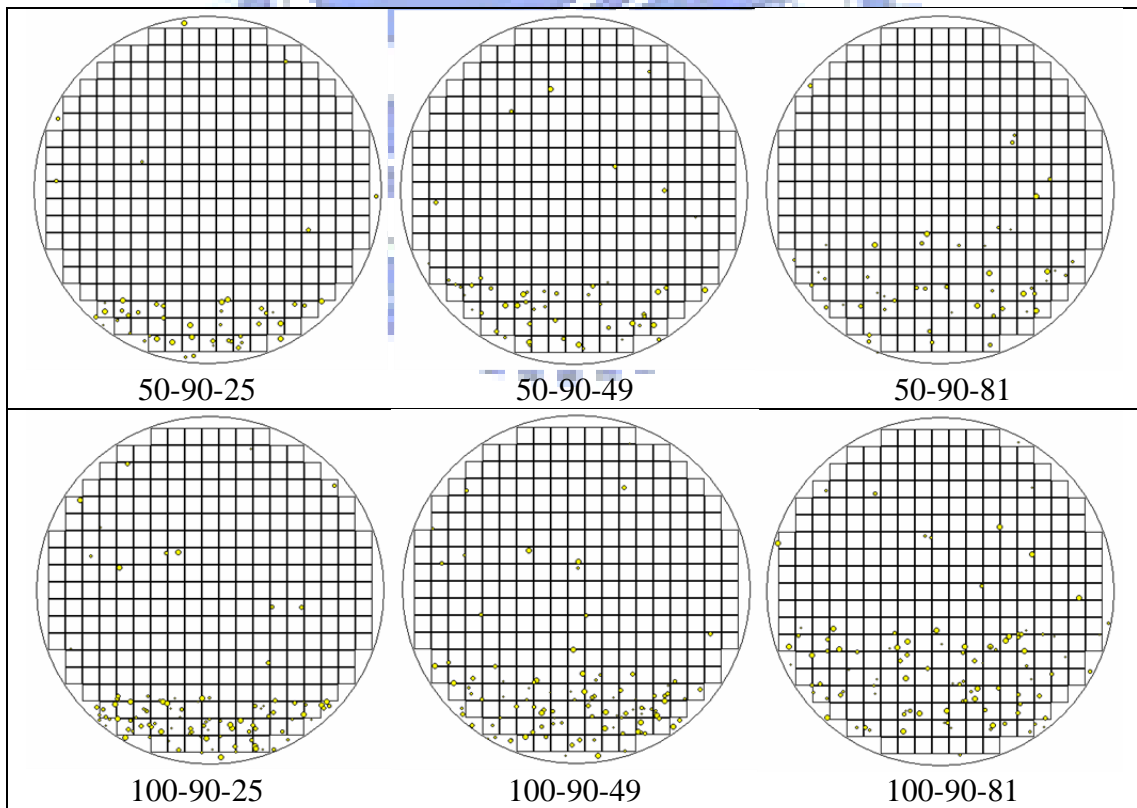


圖 4.1-(d) 底部樣式

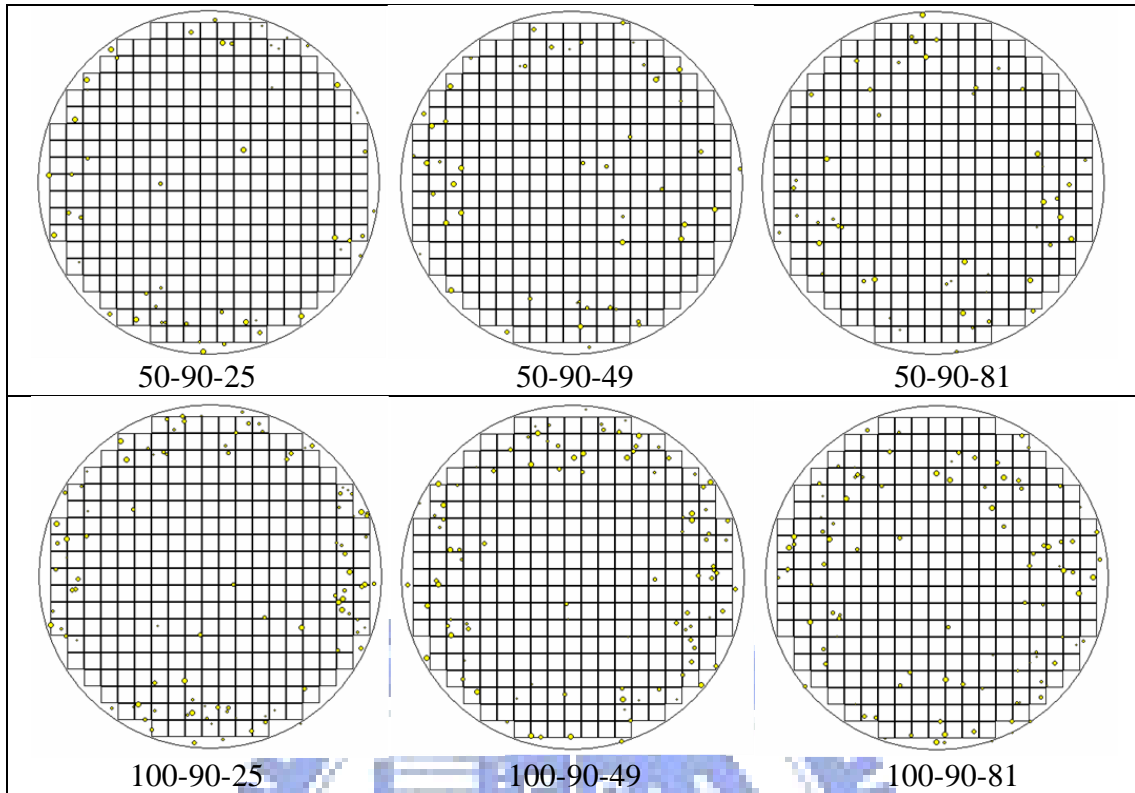


圖 4.1-(e) 環狀樣式

4.2 良率模式的比較

為探討本研究所提的 GRNN 良率模式與其他良率模式的差異性。本研究利用 4.1 小節模擬實驗所得的 1225 片 8 吋晶圓數據。其中，980 片作為訓練樣本，245 片作為測試樣本。在每次的模擬試驗中擷取 4 個晶圓特徵因子(D 、 CV_A 、 CV_D 、 CI_E) 當作輸入，並可得到晶圓上完全沒有缺陷的晶片數除以晶圓上總晶片數所得的真實良率當作輸出，以及用於其他良率模式的各項參數值。利用已訓練完成的 GRNN 良率模式所預估的良率值與其他良率模式的良率預估值做比較來探討良率模式的差異性。本研究利用各良率模式所預估的 245 片 8 吋晶圓之測試樣本的良率預估值與真實良率做比較，並以式子(19)之誤差均方根(Root Mean Squared Error, RMSE)來評估良率模式的準確性。RMSE 愈低則表示此良率模式愈準確。

本研究利用 4 個晶圓特徵因子當作輸入，並以晶圓的真實良率當作輸出，透過套裝軟體 NeuroShell® 2 來訓練及測試 GRNN 網路。最後利用已訓練完成的 GRNN 網路當作本研究之良率預估模式。再者，為比較本研究所提的 GRNN 良率模式與其它良率模式的差異性。本研究與傳統的 Poisson 良率模式、負二項良

率模式(Neg_Bio)、Jun *et al.* [25]所提的良率模式(Y_IC)、Tsai *et al.*[44]所提的良率模式(Y_ICT)以及李靜宜[1]所提的倒傳遞網路良率模式(BPNN)做比較，並以誤差均方根 RMSE 來評估各良率模式的準確性。

本研究透過套裝軟體 NeuroShell® 2 來訓練及測試本研究之 GRNN 網路以及李靜宜之倒傳遞 BPNN 網路。在網路參數設定中，設定網路層為 3 層[21]且隱藏層單元數為 25 個[48]。在設定最大訓練次數 10^5 個 epochs 或誤差均方 MSE 小於 10^{-6} 的停止條件下，以試誤法的方式找出具有最小 MSE 的網路設定值。其中，GRNN 網路唯一的網路參數為平滑因子(smoothing factor)且其值為 0.0722；而 BPNN 網路之學習速率(learning rate)為 0.6，慣性項因子(momentum factor)為 0.9。再者，在得到模擬晶圓的缺陷數以及缺陷的座標資料後，便可藉由簡單的計算求得用以計算 Poisson 良率、負二項良率(Neg_Bio)、Jun *et al.*所提良率(Y_IC)以及 Tsai *et al.*所提良率(Y_ICT)所需的各項參數值。表 4.1 為 245 片測試樣本的良率預估值與真實良率間之變異數分析。由表 4.1 之 ANOVA 表中的 P-值 $1.67E-76$ 可知各組間之良率值有顯著性的差異；再由摘要內容可知 true_yield 與本研究之 GRNN 良率預估值的平均值與變異數相當接近，可知本研究之 GRNN 良率模式較其他良率模式更能準確的預估晶圓良率。圖 4.2 為各良率模式所預估之 245 片測試樣本的良率預估值與真實良率間之 RMSE 的比較。其中，RMSE 愈低則表示此良率模式愈準確。圖 4.3 為各良率模式所預估之 245 片測試樣本的良率預估值與真實良率間之散佈圖。由圖 4.2 與圖 4.3 可知，本研究所提的 GRNN 良率模式具有最低的 RMSE 值 0.0405 與最高的相關係數 r 值 0.9705。由此可得知，本研究所提的 GRNN 良率模式在良率預估上的確比其他良率模式準確。

表 4.1 245 片測試樣本的良好率預估值與真實良好率間之變異數分析

ANOVA						
變源	SS	自由度	MS	F	P-值	臨界值
組間	14.7149	6	2.4525	68.4501	1.67E-76	2.1039
組內	61.1954	1708	0.0358			
總和	75.9103	1714				

摘要				
組	個數	總和	平均	變異數
true_yield	245	180.8562	0.7382	0.0250
Poisson	245	162.1836	0.6620	0.0448
Neg_Bio	245	203.9260	0.8324	0.0114
Y_CI	245	141.6824	0.5783	0.0591
Y_CIT	245	133.4493	0.5447	0.0616
BPNN	245	178.4414	0.7283	0.0234
GRNN	245	179.0357	0.7308	0.0255

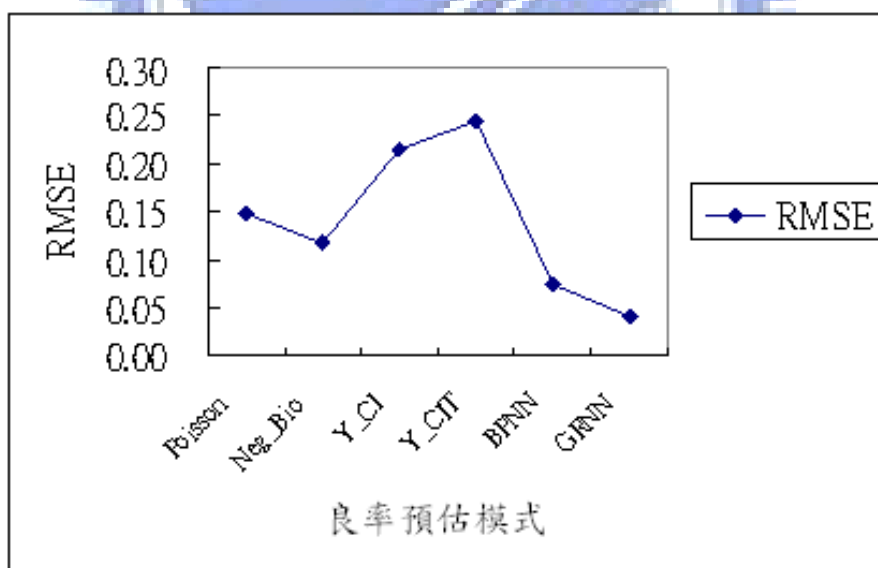
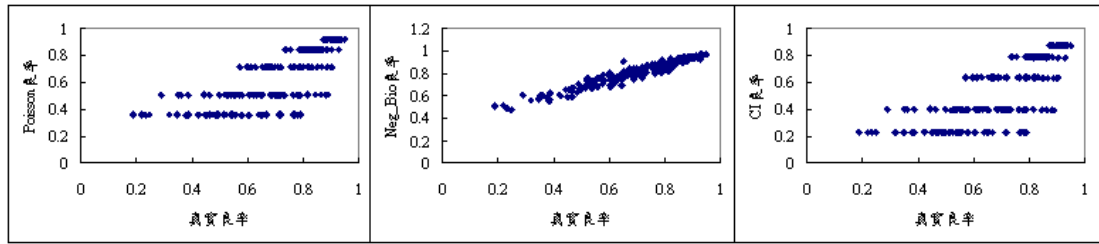


圖 4.2 245 片測試樣本的良好率預估值與真實良好率間之 RMSE 的比較



RMSE = 0.1489
r = 0.7789

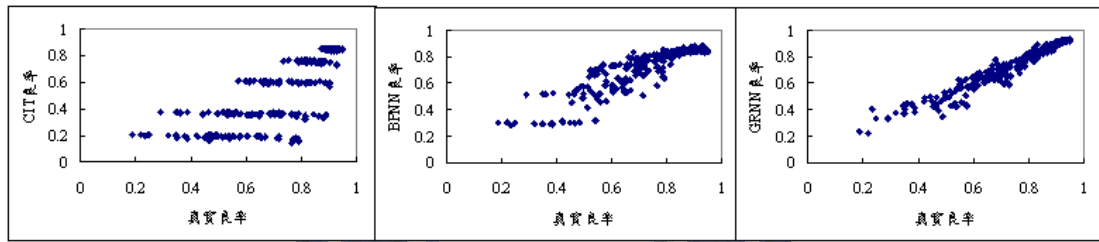
(a)

RMSE = 0.1171
r = 0.9704

(b)

RMSE = 0.2165
r = 0.7779

(c)



RMSE = 0.2454
r = 0.7656

(d)

RMSE = 0.0744
r = 0.8962

(e)

RMSE = 0.0405
r = 0.9705

(f)

圖 4.3 245 片測試樣本之良率預估值與真實良率間之散佈圖

為進一步探討在 5 種晶圓缺陷樣式之下，各良率模式與三個設計因子(缺陷數、缺陷群聚程度、以及群聚分佈面積)間的關係。底下即針對此 5 種晶圓缺陷樣式，探討各良率模式與 3 個設計因子水準的關係。

4.2.1 各良率模式與缺陷數因子水準之關係

本研究之模擬實驗所設計的晶圓表面缺陷數因子包含 5 種水準，即 25、50、100、200 以及 300。本研究利用各良率模式所預估的 1225 片 8 吋晶圓的良率預估值與真實良率做比較，探討在 5 種晶圓缺陷群聚樣式(隨機、牛眼、弦月、底部與環狀)下，各良率模式與缺陷數因子水準間的關係。本研究利用各良率模式所預估的 1225 片 8 吋晶圓的良率預估值與真實良率做比較，並以誤差均方根 RMSE 來評估良率模式的準確性。

(1) 隨機樣式

圖 4.4 為晶圓表面缺陷呈現隨機樣式分佈時，各種良率模式與缺陷數因子水準間的關係。藉由比較 RMSE 可得知各良率模式在不同的缺陷數之下，其良率預估的準確性。由圖 4.4 可知本研究之 GRNN 良率模式、Jun *et al.*所提的良率模式(Y_IC)以及 Tsai *et al.*所提的良率模式(Y ICT)在缺陷數為 25, 50, 100 以及 200 時之 RMSE 均呈現較佳的績效，亦即有較佳的良率預估值；而負二項良率模式

(Neg_Bio)則各種缺陷數下均呈現較差的績效，亦即有較差的良率預估值。由圖 4.4 可知本研究之 GRNN 良率模式在晶圓表面缺陷呈現隨機樣式分佈時，在多數的缺陷數下能比其他良率模式有較佳的良率預估值。

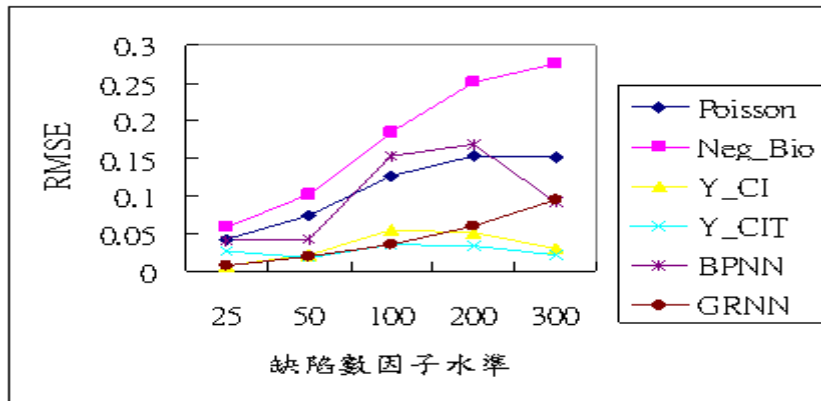


圖 4.4 隨機樣式時，各種良率模式之缺陷數因子水準與 RMSE 的關係

(2) 牛眼樣式

圖 4.5 為晶圓表面缺陷呈現牛眼樣式分佈時，各種良率模式與缺陷數因子水準間的關係。由圖 4.5 可知本研究之 GRNN 良率模式在各種缺陷數下之 RMSE 均呈現較佳的績效，亦即有較佳的良率預估值；而 Poisson 良率模式、Jun *et al.* 所提的良率模式(Y_IC)以及 Tsai *et al.*所提的良率模式(Y_ICT)則在缺陷數 100, 200 以及 300 下均呈現較差的績效，亦即有較差的良率預估值。由圖 4.5 可知本研究之 GRNN 良率模式在晶圓表面缺陷呈現牛眼樣式分佈時，在各種缺陷數下均能比其他良率模式有較佳的良率預估值。

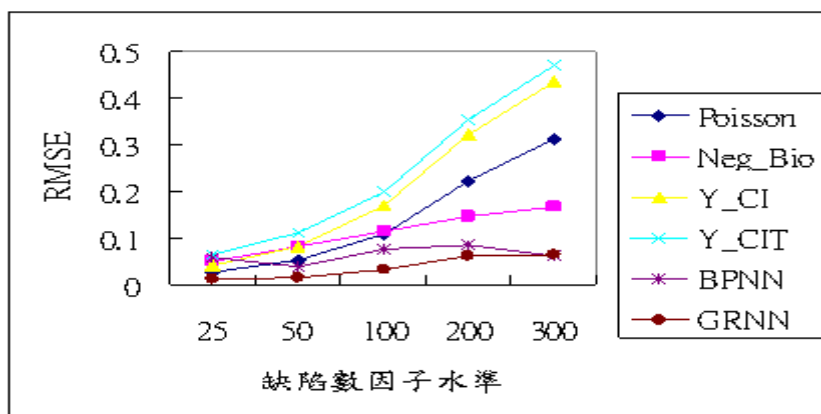


圖 4.5 牛眼樣式時，各種良率模式之缺陷數因子水準與 RMSE 的關係

(3) 弦月樣式

圖 4.6 為晶圓表面缺陷呈現弦月樣式分佈時，各種良率模式與缺陷數因子水準間的關係。由圖 4.6 可知本研究之 GRNN 良率模式在各種缺陷數下之 RMSE 均呈現較佳的績效，亦即有較佳的良率預估值；而 Poisson 良率模式、Jun *et al.* 所提的良率模式(Y_IC)以及 Tsai *et al.*所提的良率模式(Y ICT)則在缺陷數 100, 200 以及 300 下均呈現較差的績效，亦即有較差的良率預估值。由圖 4.6 可知本研究之 GRNN 良率模式在晶圓表面缺陷呈現弦月樣式分佈時，在各種缺陷數下均能比其他良率模式有較佳的良率預估績效。

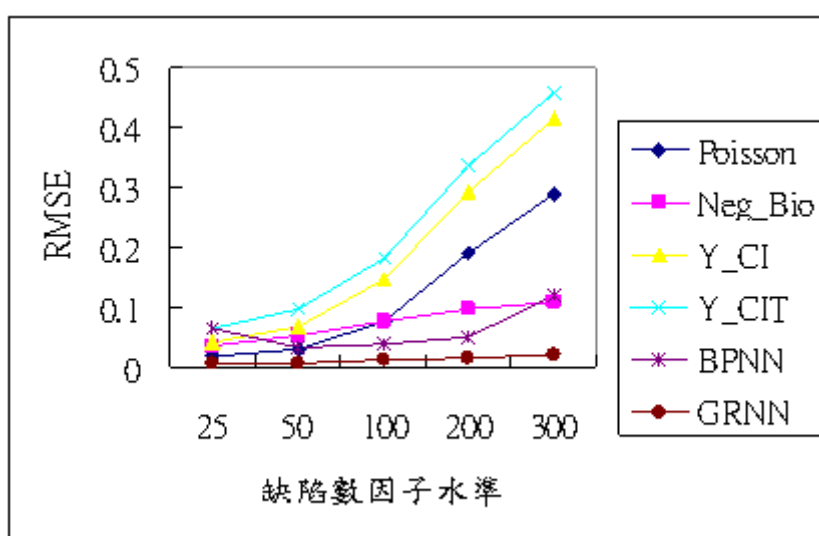


圖 4.6 弦月樣式時，各種良率模式之缺陷數因子水準與 RMSE 的關係

(4) 底部樣式

圖 4.7 為晶圓表面缺陷呈現底部樣式分佈時，各種良率模式與缺陷數因子水準間的關係。由圖 4.7 可知本研究之 GRNN 良率模式在各種缺陷數下之 RMSE 均呈現較佳的績效，亦即有較佳的良率預估值；而 Poisson 良率模式、Jun *et al.* 所提的良率模式(Y_IC)以及 Tsai *et al.*所提的良率模式(Y ICT)則在缺陷數 200 以及 300 下均呈現較差的績效，亦即有較差的良率預估值。由圖 4.7 可知本研究之 GRNN 良率模式在晶圓表面缺陷呈現底部樣式分佈時，在各種缺陷數下均能比其他良率模式有較佳的良率預估績效。

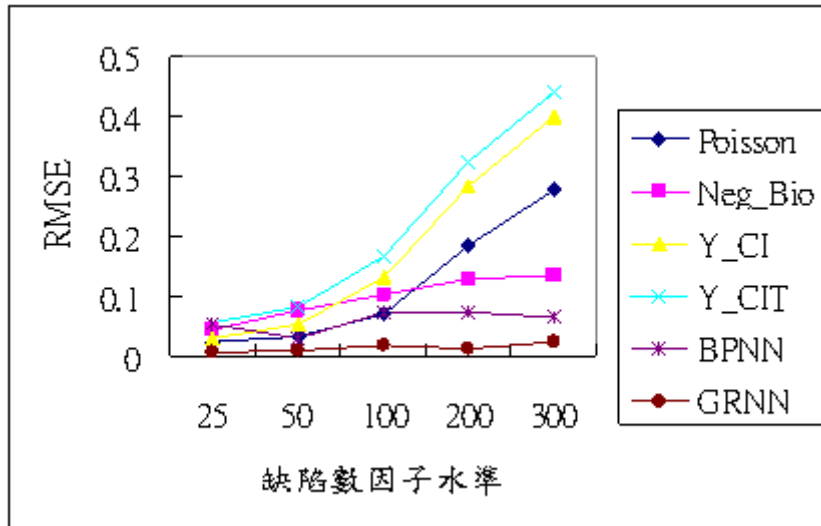


圖 4.7 底部樣式時，各種良率模式之缺陷數因子水準與 RMSE 的關係

(5) 環狀樣式

圖 4.8 為晶圓表面缺陷呈現環狀樣式分佈時，各種良率模式與缺陷數因子水準間的關係。由圖 4.8 可知本研究之 GRNN 良率模式在各種缺陷數下之 RMSE 均呈現較佳的績效，亦即有較佳的良率預估值；而負二項良率模式(Neg_Bio)在各種缺陷數下均呈現較差的績效，亦即有較差的良率預估值。由圖 4.8 可知本研究之 GRNN 良率模式在晶圓表面缺陷呈現環狀樣式分佈時，在各種缺陷數下均能比其他良率模式有較佳的良率預估績效。

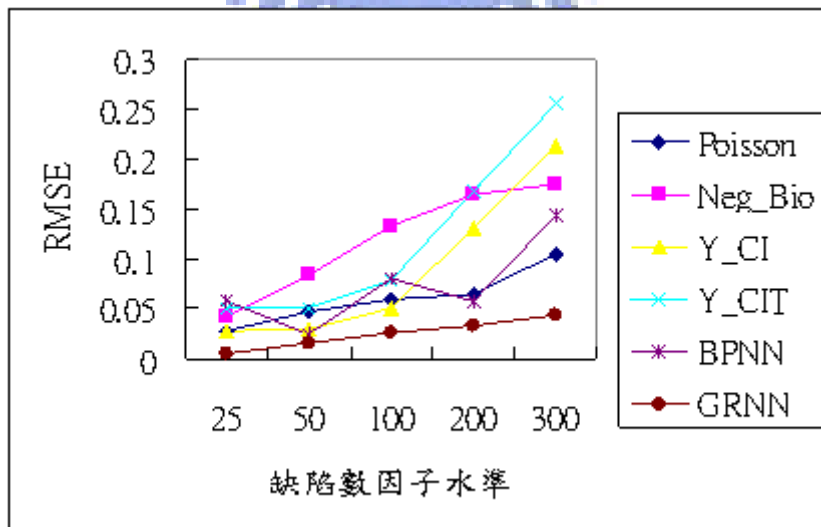


圖 4.8 環狀樣式時，各種良率模式之缺陷數因子水準與 RMSE 的關係

4.2.2 各良率模式與群聚程度因子水準之關係

本研究之模擬實驗所設計的晶圓表面缺陷群聚樣式之缺陷群聚程度因子包含 4 種水準，即 80%、85%、90% 以及 95%；而其餘非群聚部份的缺陷則呈隨機分佈。本研究利用各良率模式所預估的 1225 片 8 吋晶圓的良率預估值與真實良率做比較，來探討在 4 種晶圓缺陷群聚樣式(牛眼、弦月、底部與環狀)下，各良率模式與群聚程度因子水準間的關係。本研究利用各良率模式所預估的 1225 片 8 吋晶圓的良率預估值與真實良率做比較，並以誤差均方根 RMSE 來評估良率模式的準確性。

(1) 牛眼樣式

圖 4.9 為晶圓表面缺陷呈現牛眼樣式分佈時，各種良率模式與缺陷群聚程度因子水準間的關係。由圖 4.9 可知本研究之 GRNN 良率模式與 BPNN 良率模式在各種缺陷群聚程度下之 RMSE 均呈現較佳的績效，亦即有較佳的良率預估值；而 Jun *et al.*所提的良率模式(Y_IC)以及 Tsai *et al.*所提的良率模式(Y ICT)則在各種缺陷群聚程度下均呈現較差的績效，亦即有較差的良率預估值。由圖 4.9 可知本研究之 GRNN 良率模式在晶圓表面缺陷呈現牛眼樣式分佈時，在各種缺陷群聚程度下均能比其他良率模式有較佳的良率預估績效。

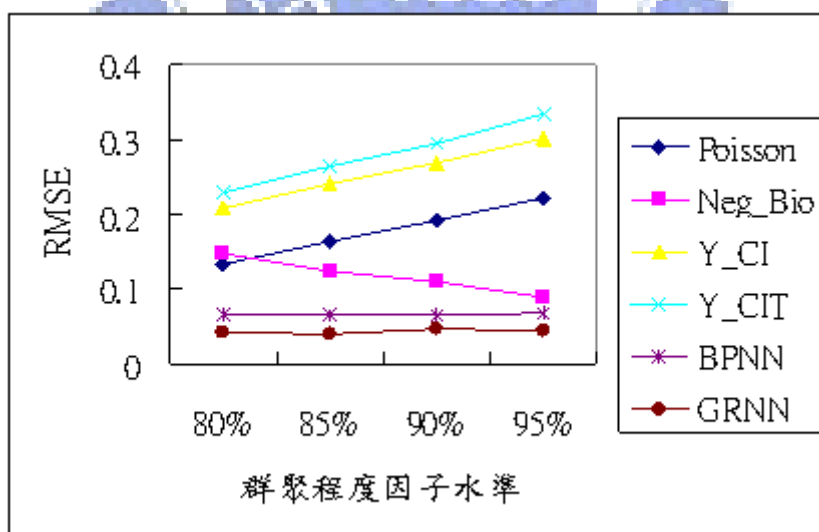


圖 4.9 牛眼樣式時，各種良率模式之群聚程度因子水準與 RMSE 的關係

(2) 弦月樣式

圖 4.10 為晶圓表面缺陷呈現弦月樣式分佈時，各種良率模式與缺陷群聚程度因子水準間的關係。由圖 4.10 可知本研究之 GRNN 良率模式在各種缺陷群聚程度下之 RMSE 均呈現較佳的績效，亦即有較佳的良率預估值；而 Poisson 良率模式、Jun 所提的良率模式(Y_IC)以及 Tsai *et al.*所提的良率模式(Y ICT)則在各種缺陷群聚程度下均呈現較差的績效，亦即有較差的良率預估值。由圖 4.10 可知本研究之 GRNN 良率模式在晶圓表面缺陷呈現弦月樣式分佈時，在各種缺陷群聚程度下均能比其他良率模式有較佳的良率預估績效。

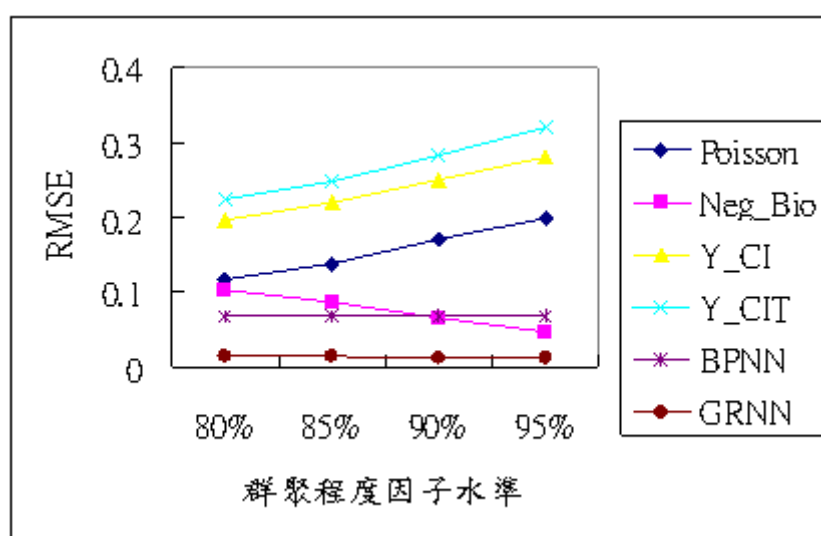


圖 4.10 弦月樣式時，各種良率模式之群聚程度因子水準與 RMSE 的關係

(3) 底部樣式

圖 4.11 為晶圓表面缺陷呈現底部樣式分佈時，各種良率模式與缺陷群聚程度因子水準間的關係。由圖 4.11 可知本研究之 GRNN 良率模式在各種缺陷數下之 RMSE 均呈現較佳的績效，亦即有較佳的良率預估值；而 Jun *et al.*所提的良率模式(Y_IC)以及 Tsai *et al.*所提的良率模式(Y ICT)則在各種缺陷群聚程度下均呈現較差的績效，亦即有較差的良率預估值。由圖 4.11 可知本研究之 GRNN 良率模式在晶圓表面缺陷呈現底部樣式分佈時，在各種缺陷群聚程度下均能比其他良率模式有較佳的良率預估績效。

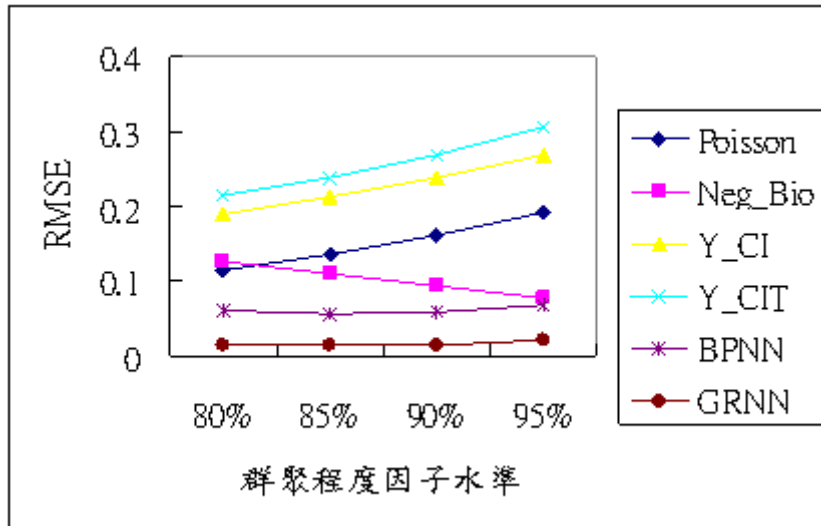


圖 4.11 底部樣式時，各種良率模式之群聚程度因子水準與 RMSE 的關係

(4) 環狀樣式

圖 4.12 為晶圓表面缺陷呈現環狀樣式分佈時，各種良率模式與缺陷群聚程度因子水準間的關係。由圖 4.12 可知本研究之 GRNN 良率模式在各種缺陷群聚程度下之 RMSE 均呈現較佳的績效，亦即有較佳的良率預估值；而負二項良率模式(Neg_Bio)在缺陷群聚程度為 80% 下呈現較差的績效，亦即有較差的良率預估值。由圖 4.12 可知本研究之 GRNN 良率模式在晶圓表面缺陷呈現環狀樣式分佈時，在各種缺陷群聚程度下均能比其他良率模式有較佳的良率預估值。

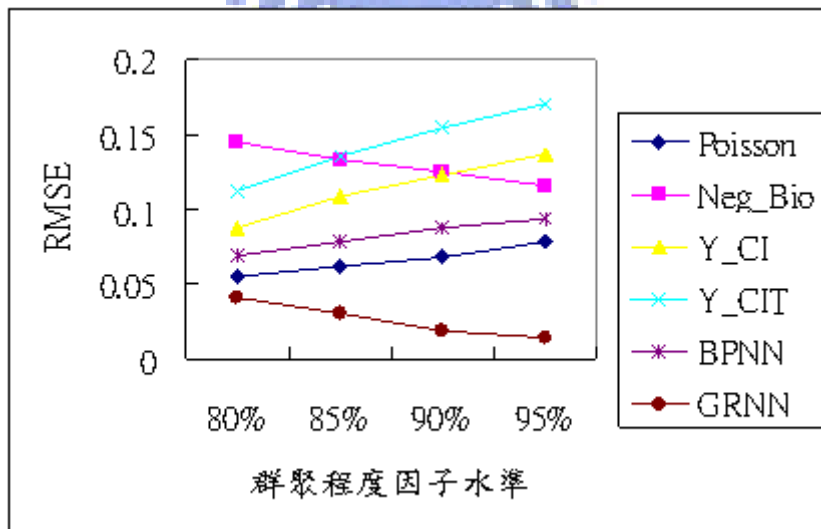


圖 4.12 環狀樣式時，各種良率模式之群聚程度因子水準與 RMSE 的關係

4.2.3 各良率模式與群聚分佈面積因子水準之關係

本研究之模擬實驗所設計的晶圓表面上缺陷可能群聚分佈的面積大小因子包含 3 種水準，即 25、49 以及 81 平方公分。本研究利用各良率模式所預估的 1225 片 8 吋晶圓的良率預估值與真實良率做比較，來探討在 4 種晶圓缺陷群聚樣式(牛眼、弦月、底部與環狀)下，各良率模式與群聚分佈面積因子水準間的關係。本研究利用各良率模式所預估的 1225 片 8 吋晶圓的良率預估值與真實良率做比較，並以誤差均方根 RMSE 來評估良率模式的準確性。

(1) 牛眼樣式

圖 4.13 為晶圓表面缺陷呈現牛眼樣式分佈時，各種良率模式與群聚分佈面積因子水準間的關係。由圖 4.13 可知本研究之 GRNN 良率模式在各種群聚分佈面積下之 RMSE 均呈現較佳的績效，亦即有較佳的良率預估值；而 Jun *et al.*所提的良率模式(Y_IC)以及 Tsai *et al.*所提的良率模式(Y ICT)則在各種群聚分佈面積下均呈現較差的績效，亦即有較差的良率預估值。由圖 4.13 可知本研究之 GRNN 良率模式在晶圓表面缺陷呈現牛眼樣式分佈時，在各種群聚分佈面積下均能比其他良率模式有較佳的良率預估績效。

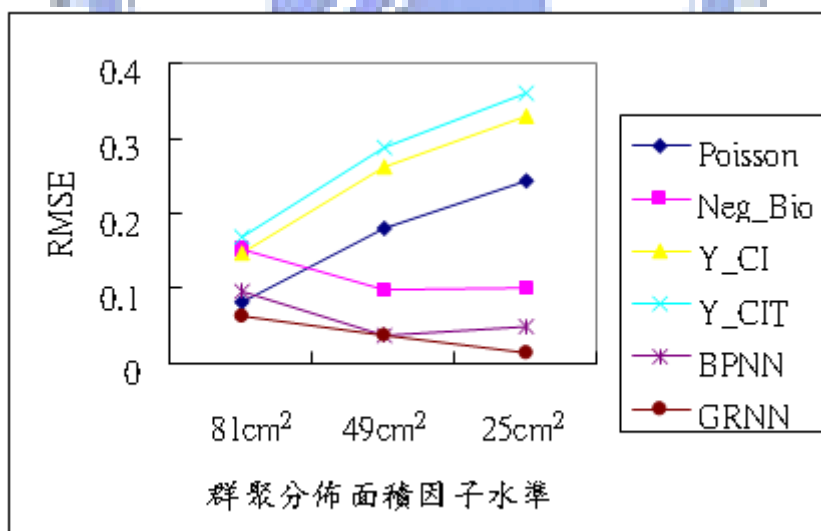


圖 4.13 牛眼樣式時，各種良率模式之群聚分佈面積因子水準與 RMSE 的關係

(2) 弦月樣式

圖 4.14 為晶圓表面缺陷呈現弦月樣式分佈時，各種良率模式與群聚分佈面積因子水準間的關係。由圖 4.14 可知本研究之 GRNN 良率模式在各種群聚分佈

面積下之 RMSE 均呈現較佳的績效，亦即有較佳的良率預估值；而 Poisson 良率模式、Jun 所提的良率模式(Y_IC)以及 Tsai *et al.*所提的良率模式(Y ICT)則在各種群聚分佈面積下均呈現較差的績效，亦即有較差的良率預估值。由圖 4.14 可知本研究之 GRNN 良率模式在晶圓表面缺陷呈現弦月樣式分佈時，在各種群聚分佈面積下均能比其他良率模式有較佳的良率預估績效。

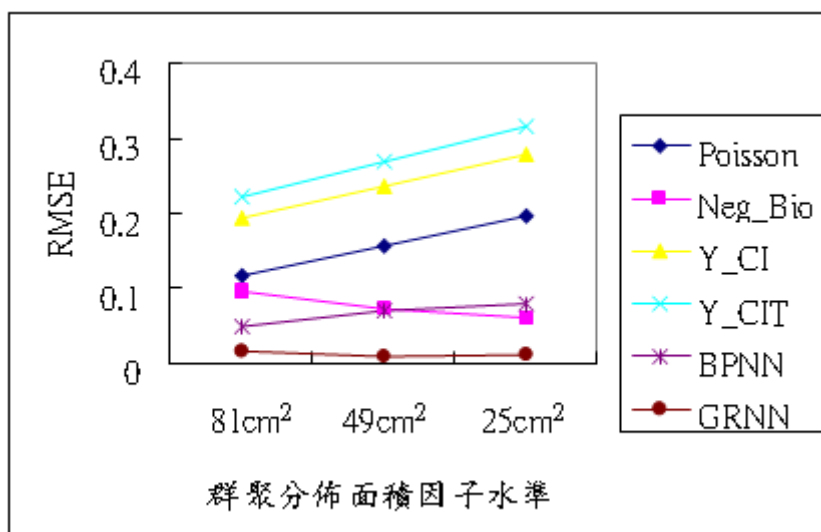


圖 4.14 弦月樣式時，各種良率模式之群聚分佈面積因子水準與 RMSE 的關係

(3) 底部樣式

圖 4.15 為晶圓表面缺陷呈現底部樣式分佈時，各種良率模式與群聚分佈面積因子水準間的關係。由圖 4.15 可知本研究之 GRNN 良率模式在各種缺陷群聚分佈面積下之 RMSE 均呈現較佳的績效，亦即有較佳的良率預估值；而 Jun *et al.*所提的良率模式(Y_IC)以及 Tsai *et al.*所提的良率模式(Y ICT)則在各種群聚分佈面積下均呈現較差的績效，亦即有較差的良率預估值。由圖 4.15 可知本研究之 GRNN 良率模式在晶圓表面缺陷呈現底部樣式分佈時，在各種群聚分佈面積下均比其他良率模式有較佳的良率預估績效。

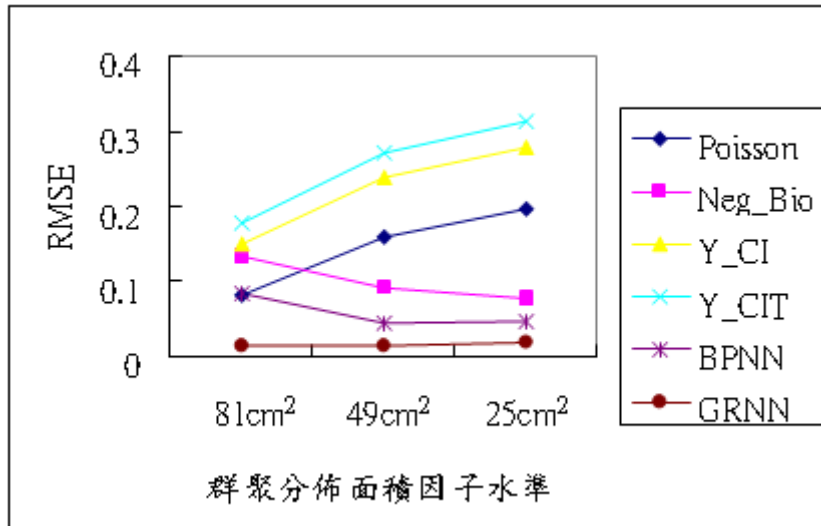


圖 4.15 底部樣式時，各種良率模式之群聚分佈面積因子水準與 RMSE 的關係

(4) 環狀樣式

圖 4.16 為晶圓表面缺陷呈現環狀樣式分佈時，各種良率模式與群聚分佈面積因子水準間的關係。由圖 4.16 可知本研究之 GRNN 良率模式在各種群聚分佈面積下之 RMSE 均呈現較佳的績效，亦即有較佳的良率預估值；而 Jun *et al.* 所提的良率模式(Y_IC)以及 Tsai *et al.* 所提的良率模式(Y ICT)則在群聚分佈面積 25cm² 下之 RMSE 呈現較差的績效，亦即有較差的良率預估值。由圖 4.16 可知本研究之 GRNN 良率模式在晶圓表面缺陷呈現環狀樣式分佈時，在各種群聚分佈面積下均能比其他良率模式有較佳的良率預估績效。

針對上述所探討之各良率模式與 3 個設計因子水準間的關係可知：在 5 種晶圓缺陷樣式下，本研究所發展之 GRNN 良率模式比其他良率模式更能準確預估各種缺陷數、群聚程度以及群聚分佈面積的晶圓良率。表 4.2 為 5 種缺陷樣式下，各良率模式在 3 種設計因子的良率預估績效彙整表。由表 4.2 可得知本研究之 GRNN 良率模式在各種晶圓缺陷樣式下均有很好的良率預估績效；BPNN 良率模式整體上也有好的良率預估績效；Jun *et al.* 所提的良率模式(Y_IC)以及 Tsai *et al.* 所提的良率模式(Y ICT)對隨機樣式有好的良率預估績效；而負二項良率模式(Neg_Bio)在缺陷樣式為牛眼、弦月以及底部時亦有好的良率預估績效。再者，針對缺陷樣式為弦月與環狀時，各良率模式有相似的良率預估值。

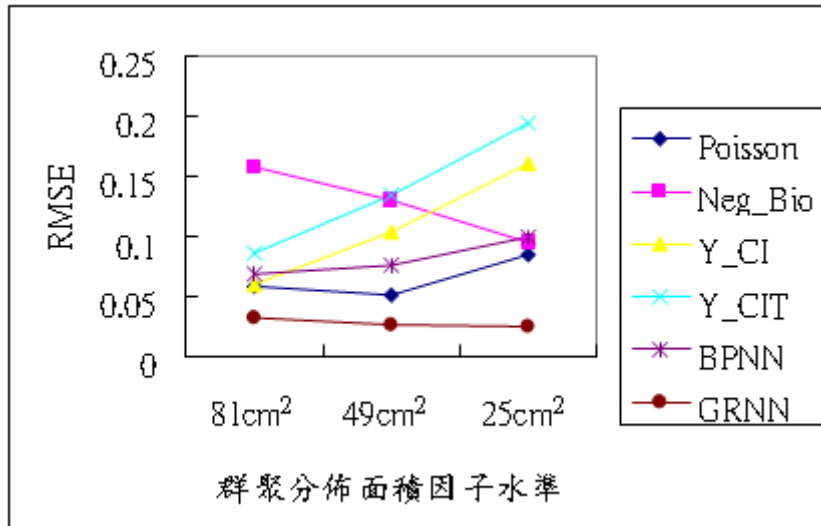


圖 4.16 環狀樣式時，各種良率模式之群聚分佈面積因子水準與 RMSE 的關係

表 4.2 5 種缺陷樣式下，各良率模式在 3 種設計因子的預估績效

缺陷樣式	良率模式 設計因子	Poisson	Neg_Bio	Y_IC	Y ICT	BPNN	GRNN
		良率模式	良率模式	良率模式	良率模式	良率模式	良率模式
隨機	缺陷數	○	△	●	●	○	●
牛眼	缺陷數	△	○	△	△	○	●
	缺陷群聚程度	○	○	△	△	●	●
	群聚分佈面積	○	○	△	△	●	●
弦月	缺陷數	△	○	△	△	○	●
	缺陷群聚程度	△	○	△	△	○	●
	群聚分佈面積	△	○	△	△	○	●
底部	缺陷數	△	○	△	△	○	●
	缺陷群聚程度	△	○	△	△	○	●
	群聚分佈面積	○	○	△	△	●	●
環狀	缺陷數	○	△	△	△	○	●
	缺陷群聚程度	○	△	△	△	○	●
	群聚分佈面積	○	△	△	△	○	●

●：優 ○：良 △：普通

4.3 晶圓缺陷樣式辨識技術的比較

當一片晶圓出現中、低良率時，應進一步診斷該晶圓表面的缺陷是否存在某種特定的缺陷群聚現象。此處所謂之中、低良率乃詢問 IC 製造業之製程工程

師對晶圓良率於業界的一般性定義。一般定義良率在 70% 左右為中良率，而 50% 以下為低良率。為利用多類別 SVM 來辨識晶圓表面之缺陷樣式，本研究利用 4.1 小節之模擬實驗所得的 1225 片 8 吋晶圓缺陷樣式作為建構多類別 SVM 的樣本。其中 980 片作為訓練樣本，245 片作為測試樣本。再者，為探討本研究所提的多類別 SVM 晶圓缺陷樣式辨識系統與其他缺陷樣式辨識技術的差異性。在每次的模擬試驗中擷取 4 個晶圓特徵因子(D 、 CV_A 、 CV_D 、 CI_E)當作輸入，並可得知欲辨識的 5 種對應之晶圓缺陷樣式類別當作輸出，以及用於其他缺陷樣式辨識技術的各項參數值。透過訓練完成之多類別 SVM 的缺陷樣式辨識正確率與其他樣式辨識技術的正確率做比較來探討各種缺陷樣式辨識技術的效果。正確辨識率愈高則表示此缺陷樣式辨識技術愈準確。

本研究利用 4 個晶圓特徵因子當作輸入，並以欲辨識的 5 種對應之晶圓缺陷樣式類別當作輸出，透過多類別 SVM 軟體 LIBSVM 來訓練及測試多類別 SVM。最後利用已訓練完成的多類別 SVM 當作本研究之晶圓缺陷樣式辨識系統。為比較本研究所提的多類別 SVM 晶圓缺陷樣式辨識系統與其它缺陷樣式辨識技術的差異性。本研究利用倒傳遞網路(BPNN)以及徑向基底函數網路(RBF)來辨識晶圓缺陷樣式，並以欲辨識的 5 種對應晶圓缺陷樣式類別之正確辨識率來評估本研究之多類別 SVM 晶圓缺陷樣式辨識系統與 BPNN 以及 RBF 網路技術在缺陷樣式辨識上的準確性。

本研究透過軟體 LIBSVM 來訓練及測試本研究之多類別 SVM。LIBSVM 利用格子搜尋(grid search)法來選取最佳的懲罰參數 C 與參數 γ 的參數組合。本研究將所得的(C, γ)參數組合再用於訓練與測試樣本以得到最佳的分類結果。在本研究中之所得之 γ 值為 0.125，而 C 值為 8192。在 245 片測試樣本中有 199 片能夠被正確的辨識其所對應的晶圓缺陷樣式，正確辨識率為 81.23%。而 BPNN 網路與 RBF 神經網路同樣以此 4 個晶圓特徵因子當作輸入，並以欲辨識的 5 種對應之晶圓缺陷樣式類別當作輸出，並利用 NeuroSolutions 5.0 套裝軟體來做訓練及測試。在 BPNN 網路與 RBF 神經網路之網路參數設定中，設定網路層為 3 層[21]且隱藏層單元數為 25 個[48]。並在設定最大訓練次數 10^5 個 epochs 或誤差均方 MSE 小於 10^{-6} 的停止條件下，以試誤法的方式找出具有最小 MSE 的網路設定值。在相同的 245 片測試樣本中有 125 片能夠透過學習完成的 BPNN 網路正確的辨識其所對應的晶圓缺陷樣式，正確辨識率為 51.02%；而有 100 片能夠透過

學習完成的 RBF 神經網路正確的辨識其所對應的晶圓缺陷樣式，正確辨識率為 40.82%。由上述之正確辨識率可得知本研究所提之多類別 SVM 在本模擬實驗中比利用 BPNN 網路以及 RBF 神經網路更能夠正確地來辨識晶圓表面的缺陷樣式。

為探討在 5 種晶圓缺陷樣式之下，各樣式辨識技術三個設計因子(缺陷數、缺陷群聚程度、以及群聚分佈面積)間的關係。本研究利用 4.1 小節之模擬實驗所得的 1225 片 8 吋晶圓數據，透過各種缺陷樣式辨識技術做辨識，並以正確辨識率來比較各種缺陷樣式辨識技術的辨識能力。正確辨識率愈高，代表該缺陷樣式辨識技術的辨識能力愈好。底下即針對此 5 種晶圓缺陷樣式，探討各辨識技術與 3 個因子水準的關係。

4.3.1 各樣式辨識技術與缺陷數因子水準之關係

本研究之模擬實驗所設計的晶圓表面缺陷數因子包含 5 種水準，即 25、50、100、200 以及 300。本研究利用各缺陷樣式辨識技術所辨識的 1225 片 8 吋晶圓的分類與其真實分類做比較，探討在 5 種晶圓缺陷群聚樣式(隨機、牛眼、弦月、底部與環狀)下，各缺陷樣式辨識技術與缺陷數因子水準間的關係，並以正確辨識率來比較各種缺陷樣式辨識技術的辨識能力。

(1) 隨機樣式

圖 4.17 為晶圓表面缺陷呈現隨機樣式分佈時，各種缺陷樣式辨識技術與缺陷數因子水準間的關係。藉由比較正確辨識率可得知各辨識技術在不同的缺陷數之下，其辨識技術的準確性。由圖 4.17 可知本研究之多類別 SVM 在缺陷數為 200 與 300 下之正確辨識率呈現較佳的績效，亦即有較佳的缺陷樣式辨識分類；而 BPNN 網路與 RBF 神經網路則在缺陷數為 25, 50, 100 以及 100 下均呈現較差的績效，亦即有較差的缺陷樣式辨識分類。由圖 4.17 可知本研究之多類別 SVM 在晶圓表面缺陷呈現隨機樣式分佈時，在各種缺陷數下均能比其他辨識技術有較佳的缺陷樣式辨識分類績效。

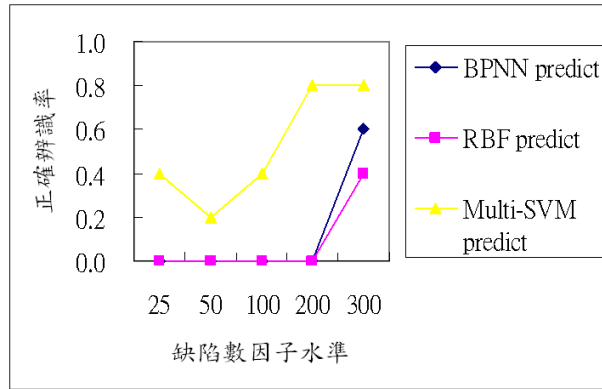


圖 4.17 隨機樣式時，各種辨識技術之缺陷數因子水準與正確辨識率的關係

(2) 牛眼樣式

圖 4.18 為晶圓表面缺陷呈現牛眼樣式分佈時，各種缺陷樣式辨識技術與缺陷數因子水準間的關係。藉由比較正確辨識率可得知各辨識技術在不同的缺陷數之下，其辨識技術的準確性。由圖 4.18 可知本研究之多類別 SVM 在各缺陷數因子水準下之正確辨識率均呈現較佳的績效，亦即有較佳的缺陷樣式辨識分類；而 RBF 神經網路則在缺陷數為 50, 100, 200 以及 300 下均呈現較差的績效，亦即有較差的缺陷樣式辨識分類。由圖 4.18 可知本研究之多類別 SVM 在晶圓表面缺陷呈現牛眼樣式分佈時，在各種缺陷數下均能比其他辨識技術有較佳的缺陷樣式辨識分類績效。

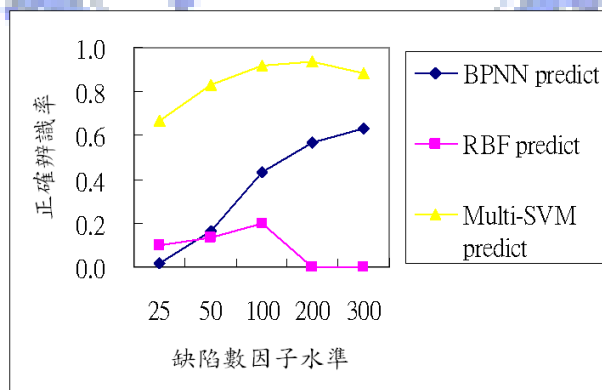


圖 4.18 牛眼樣式時，各種辨識技術之缺陷數因子水準與正確辨識率的關係

(3) 弦月樣式

圖 4.19 為晶圓表面缺陷呈現弦月樣式分佈時，各種缺陷樣式辨識技術與缺陷數因子水準間的關係。藉由比較正確辨識率可得知各辨識技術在不同的缺陷數之下，其辨識技術的準確性。由圖 4.19 可知本研究之多類別 SVM 在各缺陷數因

子水準下之正確辨識率均呈現較佳的績效，亦即有較佳的缺陷樣式辨識分類；而 RBF 神經網路則在缺陷數為 100, 200 以及 300 下均呈現較差的績效，亦即有較差的缺陷樣式辨識分類。由圖 4.19 可知本研究之多類別 SVM 在晶圓表面缺陷呈現弦月樣式分佈時，在各種缺陷數下均能比其他辨識技術有較佳的缺陷樣式辨識分類績效。

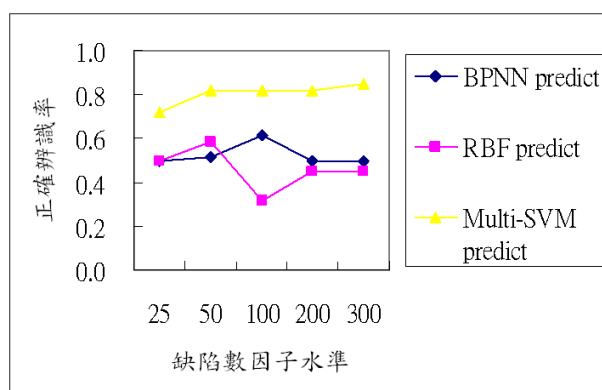


圖 4.19 弦月樣式時，各種辨識技術之缺陷數因子水準與正確辨識率的關係

(4) 底部樣式

圖 4.20 為晶圓表面缺陷呈現底部樣式分佈時，各種缺陷樣式辨識技術與缺陷數因子水準間的關係。藉由比較正確辨識率可得知各辨識技術在不同的缺陷數之下，其辨識技術的準確性。由圖 4.20 可知本研究之多類別 SVM 在缺陷數 200 與 300 下之正確辨識率呈現較佳的績效。RBF 神經網路則在缺陷數為 25 與 100 下亦呈現較佳的績效；而 RBF 神經網路則在缺陷數為 300 下呈現較差的績效，亦即有較差的缺陷樣式辨識分類。由圖 4.20 可知本研究之多類別 SVM 在晶圓表面缺陷呈現底部樣式分佈時，整體上在缺陷數因子水準亦有好的缺陷樣式辨識分類績效。

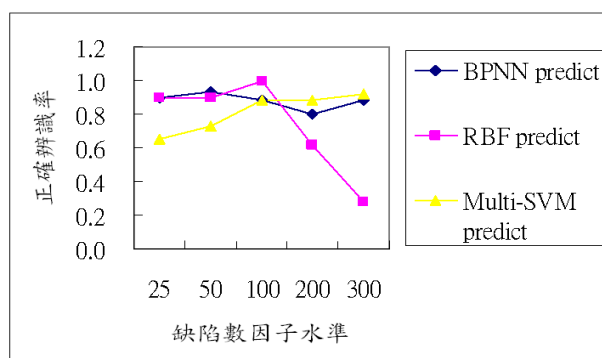


圖 4.20 底部樣式時，各種辨識技術之缺陷數因子水準與正確辨識率的關係

(5) 環狀樣式

圖 4.21 為晶圓表面缺陷呈現環狀樣式分佈時，各種缺陷樣式辨識技術與缺陷數因子水準間的關係。由圖 4.21 可知本研究之多類別 SVM 在各缺陷數因子水準下之正確辨識率均呈現較佳的績效。BPNN 網路在缺陷數為 25, 50 以及 100 下均呈現較差的績效；而 RBF 神經網路則在缺陷數為 200 與 300 下呈現較差的績效，亦即有較差的缺陷樣式辨識分類。由圖 4.21 可知本研究之多類別 SVM 在晶圓表面缺陷呈現環狀樣式分佈時，在各種缺陷數下均能比其他辨識技術有較佳的缺陷樣式辨識分類績效。

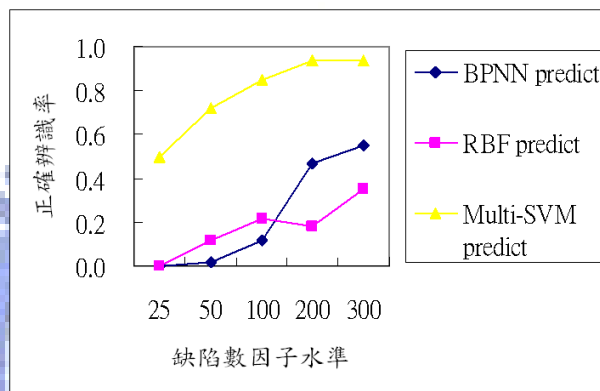


圖 4.21 環狀樣式時，各種辨識技術之缺陷數因子水準與正確辨識率的關係

4.3.2 各樣式辨識技術與缺陷群聚程度因子水準之關係

本研究之模擬實驗所設計的晶圓表面缺陷群聚樣式之缺陷群聚程度因子包含 4 種水準，即 80%、85%、90% 以及 95%；而其餘非群聚部份的缺陷則呈隨機分佈。本研究利用各缺陷樣式辨識技術所辨識的 1225 片 8 吋晶圓的分類與其真實分類做比較，探討在 4 種晶圓缺陷群聚樣式(牛眼、弦月、底部與環狀)下，各缺陷樣式辨識技術與缺陷數因子水準間的關係，並以正確辨識率來比較各種缺陷樣式辨識技術的辨識能力。

(1) 牛眼樣式

圖 4.22 為晶圓表面缺陷呈現牛眼樣式分佈時，各種缺陷樣式辨識技術與缺陷群聚程度因子水準間的關係。藉由比較正確辨識率可得知各辨識技術在不同的缺陷群聚程度之下，其辨識技術的準確性。由圖 4.22 可知本研究之多類別 SVM 在各缺陷群聚程度因子水準下之正確辨識率均呈現較佳的績效，亦即有較佳的缺

陷樣式辨識分類；而 RBF 神經網路則在各種缺陷群聚程度下均呈現較差的績效，亦即有較差的缺陷樣式辨識分類。由圖 4.22 可知本研究之多類別 SVM 在晶圓表面缺陷呈現牛眼樣式分佈時，在各種缺陷群聚程度下均能比其他辨識技術有較佳的缺陷樣式辨識分類績效。

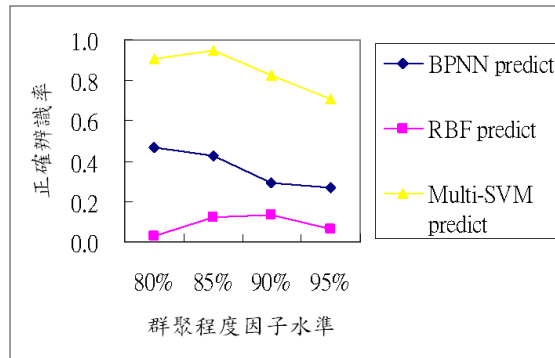


圖 4.22 牛眼樣式時，各種辨識技術之群聚程度因子水準與正確辨識率的關係

(2) 弦月樣式

圖 4.23 為晶圓表面缺陷呈現弦月樣式分佈時，各種缺陷樣式辨識技術與缺陷群聚程度因子水準間的關係。藉由比較正確辨識率可得知各辨識技術在不同的缺陷群聚程度之下，其辨識技術的準確性。由圖 4.23 可知本研究之多類別 SVM 在各缺陷群聚程度因子水準下之正確辨識率均呈現較佳的績效，亦即有較佳的缺陷樣式辨識分類；而 RBF 神經網路則在缺陷群聚程度為 85%, 90% 以及 95% 下均呈現較差的績效，亦即有較差的缺陷樣式辨識分類。由圖 4.23 可知本研究之多類別 SVM 在晶圓表面缺陷呈現弦月樣式分佈時，在各種缺陷群聚程度下均能比其他辨識技術有較佳的缺陷樣式辨識分類績效。

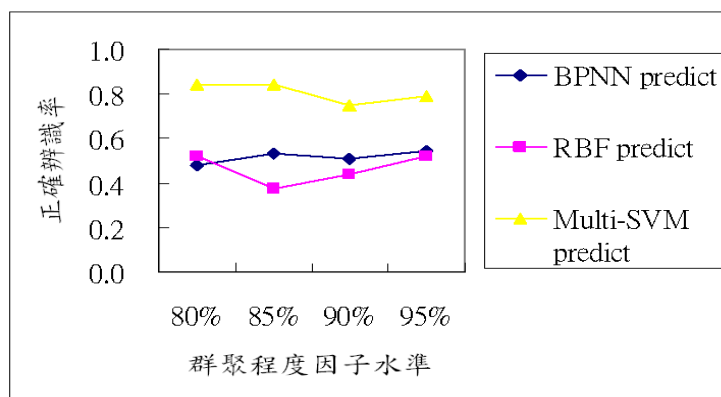


圖 4.23 弦月樣式時，各種辨識技術之群聚程度因子水準與正確辨識率的關係

(3) 底部樣式

圖 4.24 為晶圓表面缺陷呈現底部樣式分佈時，各種缺陷樣式辨識技術與缺陷群聚程度因子水準間的關係。藉由比較正確辨識率可得知各辨識技術在不同的缺陷群聚程度之下，其辨識技術的準確性。由圖 4.24 可知本研究之多類別 SVM 在缺陷群聚程度為 90% 下之正確辨識率呈現較佳的績效。BPNN 神經網路則在缺陷群聚程度為 80%、85% 以及 95% 下亦呈現較佳的績效；而 RBF 神經網路則在缺陷群聚程度為 85%、90% 以及 95% 下呈現較差的績效，亦即有較差的缺陷樣式辨識分類。由圖 4.24 可知本研究之多類別 SVM 在晶圓表面缺陷呈現底部樣式分佈時，整體上在缺陷群聚程度因子水準亦有好的缺陷樣式辨識分類績效。

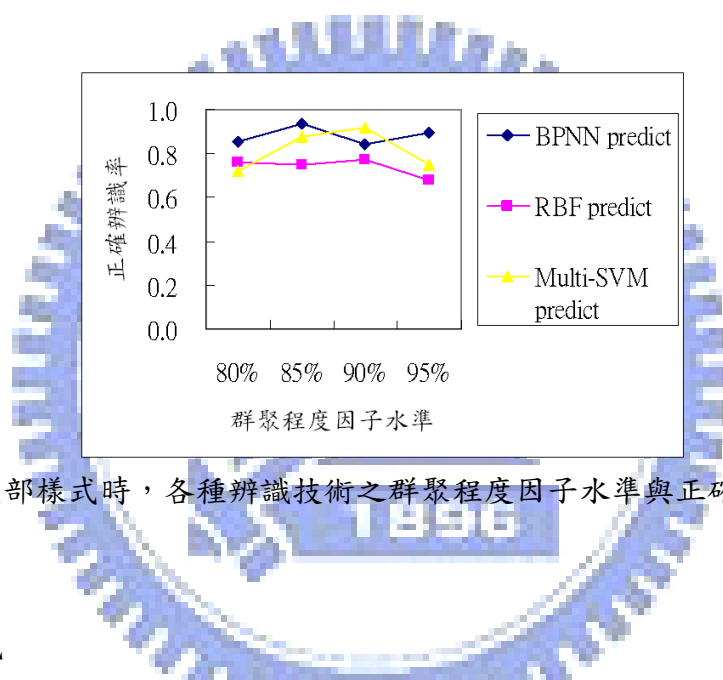


圖 4.24 底部樣式時，各種辨識技術之群聚程度因子水準與正確辨識率的關係

(4) 環狀樣式

圖 4.25 為晶圓表面缺陷呈現環狀樣式分佈時，各種缺陷樣式辨識技術與缺陷群聚程度因子水準間的關係。由圖 4.25 可知本研究之多類別 SVM 在各缺陷群聚程度因子水準下之正確辨識率均呈現較佳的績效。而 RBF 神經網路則在各缺陷群聚程度下均呈現較差的績效，亦即有較差的缺陷樣式辨識分類。由圖 4.25 可知本研究之多類別 SVM 在晶圓表面缺陷呈現環狀樣式分佈時，在各種缺陷群聚程度下均能比其他辨識技術有較佳的缺陷樣式辨識分類績效。

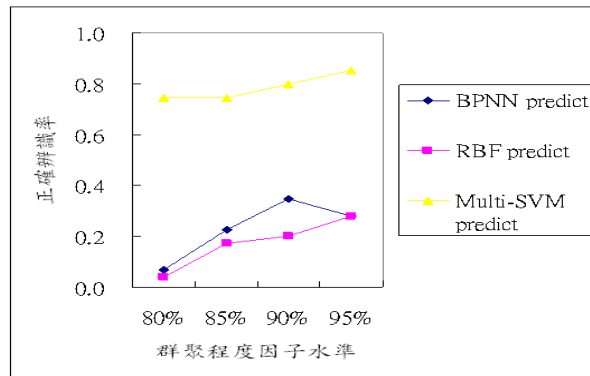


圖 4.25 環狀樣式時，各種辨識技術之群聚程度因子水準與正確辨識率的關係

4.3.3 各樣式辨識技術與群聚分佈面積因子水準之關係

本研究之模擬實驗所設計的晶圓表面上缺陷可能群聚分佈的面積大小因子包含 3 種水準，即 25、49 以及 81 平方公分。本研究利用各缺陷樣式辨識技術所辨識的 1225 片 8 吋晶圓的分類與其真實分類做比較，來探討在 4 種晶圓缺陷群聚樣式(牛眼、弦月、底部與環狀)下，各缺陷樣式辨識技術與群聚分佈面積因子水準間的關係，並以正確辨識率來比較各種缺陷樣式辨識技術的辨識能力。

(1) 牛眼樣式

圖 4.26 為晶圓表面缺陷呈現牛眼樣式分佈時，各種缺陷樣式辨識技術與群聚分佈面積因子水準間的關係。藉由比較正確辨識率可得知各辨識技術在不同的群聚分佈面積之下，其辨識技術的準確性。由圖 4.26 可知本研究之多類別 SVM 在各缺陷群聚分佈面積因子水準下之正確辨識率均呈現較佳的績效；而 RBF 神經網路則在各種缺陷群聚程度下均呈現較差的辨識績效。由圖 4.26 可知本研究之多類別 SVM 在晶圓表面缺陷呈現牛眼樣式分佈時，在各種缺陷群聚分佈面積下均能比其他辨識技術有較佳的缺陷樣式辨識分類績效。

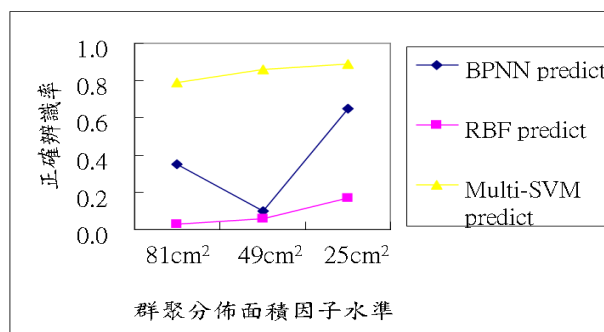


圖 4.26 牛眼樣式時，各種辨識技術之群聚分佈面積因子水準與正確辨識率的關係

(2) 弦月樣式

圖 4.27 為晶圓表面缺陷呈現弦月樣式分佈時，各種缺陷樣式辨識技術與群聚分佈面積因子水準間的關係。藉由比較正確辨識率可得知各辨識技術在不同的群聚分佈面積之下，其辨識技術的準確性。由圖 4.27 可知本研究之多類別 SVM 在各缺陷群聚程度因子水準下之正確辨識率均呈現較佳的績效；而 RBF 神經網路則在缺陷群聚分佈面積為 49cm^2 與 25cm^2 下均呈現較差的績效；而 BPNN 網路則在缺陷群聚分佈面積為 81cm^2 下呈現較差的缺陷樣式辨識分類。由圖 4.27 可知本研究之多類別 SVM 在晶圓表面缺陷呈現弦月樣式分佈時，在各種缺陷群聚分佈面積下均能比其他辨識技術有較佳的缺陷樣式辨識分類績效。

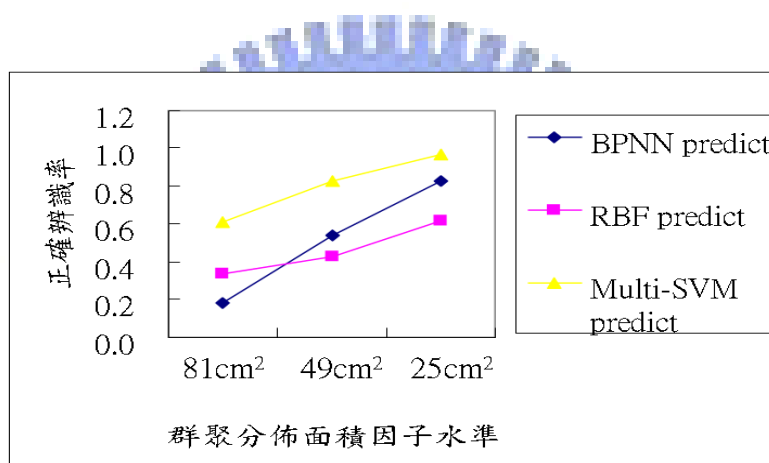


圖 4.27 弦月樣式時，各種辨識技術之群聚分佈面積因子水準與正確辨識率的關係

(3) 底部樣式

圖 4.28 為晶圓表面缺陷呈現底部樣式分佈時，各種缺陷樣式辨識技術與缺陷群聚分佈面積因子水準間的關係。藉由比較正確辨識率可得知各辨識技術在不同的缺陷群聚分佈面積之下，其辨識技術的準確性。由圖 4.28 可知本研究之多類別 SVM 在缺陷群聚分佈面積 81cm^2 下有好的缺陷樣式辨識分類績效；BPNN 網路在各缺陷群聚分佈面積下之正確辨識率均呈現較佳的績效。而 RBF 神經網路則在缺陷群聚分佈面積 81cm^2 下呈現較差的分類績效。由圖 4.28 可知本研究之多類別 SVM 在晶圓表面缺陷呈現底部樣式分佈時，整體上在缺陷群聚分佈面積因子水準亦有好的缺陷樣式辨識分類績效。

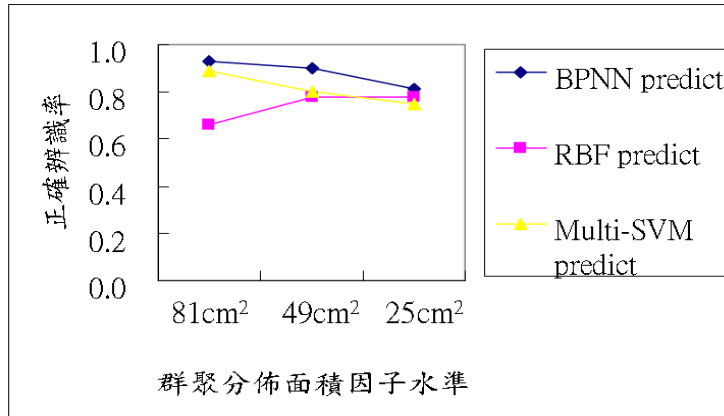


圖 4.28 底部樣式時，各種辨識技術之群聚分佈面積因子水準與正確辨識率的關係

(4) 環狀樣式

圖 4.29 為晶圓表面缺陷呈現環狀樣式分佈時，各種缺陷樣式辨識技術與缺陷群聚分佈面積因子水準間的關係。由圖 4.29 可知本研究之多類別 SVM 在各缺陷群聚分佈面積因子水準下之正確辨識率均呈現較佳的績效。RBF 神經網路則在缺陷群聚分佈面積為 81cm² 與 49cm² 下均呈現較差的績效；而 BPNN 網路則在缺陷群聚分佈面積為 25cm² 下呈現較差的缺陷樣式辨識分類。由圖 4.29 可知本研究之多類別 SVM 在晶圓表面缺陷呈現環狀樣式分佈時，在各種缺陷群聚分佈面積下均能比其他辨識技術有較佳的缺陷樣式辨識分類績效。

針對上述所探討之各辨識技術與 3 個因子水準間的關係可知：在 5 種晶圓缺陷樣式下，本研究所發展之多類別 SVM 比其他辨識技術更能準確辨識各種缺陷數、群聚程度以及群聚分佈面積的晶圓缺陷樣式的類別。表 4.3 為 5 種缺陷樣式下，各辨識技術在 3 種設計因子的缺陷樣式辨識績效彙整表。由表 4.3 可得知本研究之多類別 SVM 在各種晶圓缺陷樣式下均有好的辨識績效；BPNN 與 RBF 辨識技術在缺陷樣式為底部時有好的辨識績效，但隨機、牛眼、以及環狀樣式之辨識上的績效較差。

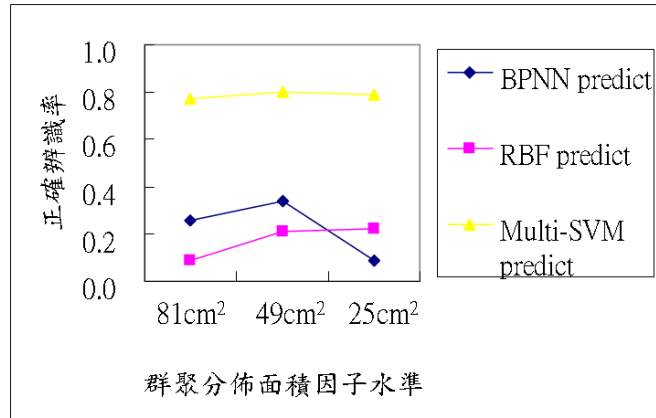


圖 4.29 環狀樣式時，各種辨識技術之群聚分佈面積因子水準與正確辨識率的關係

表 4.3 5 種缺陷樣式下，各辨識技術在 3 種設計因子的辨識績效

缺陷樣式	辨識技術 設計因子	BPNN	RBF	多類別
		神經網路	神經網路	SVM
隨機	缺陷數	▲	▲	△
	缺陷群聚程度	△	▲	○
	群聚分佈面積	▲	▲	○
牛眼	缺陷數	△	△	○
	缺陷群聚程度	△	△	○
	群聚分佈面積	△	△	○
弦月	缺陷數	○	△	○
	缺陷群聚程度	○	○	○
	群聚分佈面積	○	○	○
底部	缺陷數	▲	▲	○
	缺陷群聚程度	▲	▲	○
	群聚分佈面積	▲	▲	○

○：良 △：普通 ▲：差

4.4 整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統

為適切地反應出產品的真實良率，並且能夠針對呈現特定樣式的晶圓做出正確的製程異常診斷。當一片晶圓出現中、低良率時，應進一步診斷該晶圓表面的缺陷是否存在某種特定的缺陷樣式。而 IC 製造業一般定義良率在 70% 左右為

中良率，而 50% 以下為低良率。本研究利用 3.4 小節之晶圓缺陷診斷系統的構建流程圖來整合 GRNN 良率預估模式與多類別 SVM 缺陷樣式辨識系統，以進行晶圓缺陷診斷。

為說明本研究之晶圓缺陷診斷系統之構建流程，本研究進一步以電腦模擬 8 吋晶圓的 5 種表面缺陷樣式各 2 片。圖 4.30 為此 10 片晶圓之樣式，而該晶圓的相關資訊如表 4.4 所示。將此 10 片晶圓之 4 個晶圓特徵因子(D 、 CV_A 、 CV_D 、 CI_E)當作輸入，再利用 4.2 小節之已訓練完成的 GRNN 網路來預估此 10 片晶圓的良率。表 4.5 為 10 片晶圓之 GRNN 預估良率與其對應的良率程度。由表 4.5 可得知本研究之 GRNN 良率模式能準確地預估此 10 片晶圓的良率。針對表 4.5 之中、低良率晶圓(編號：1, 2, 4, 6, 8, 10)需再辨識此 6 片晶圓表面是否呈現特定的缺陷樣式。接著將此 6 片晶圓的 4 個晶圓特徵因子(D 、 CV_A 、 CV_D 、 CI_E)當作輸入，並利用 4.3 小節之已訓練完成的多類別 SVM 來辨識此 6 片晶圓的類別。表 4.6 為此 6 片晶圓之真實缺陷樣式與利用多類別 SVM 所辨識之樣式。由表 4.6 可知呈現中、低良率的晶圓均能夠被正確的辨識其所屬的缺陷樣式。由以上之晶圓缺陷診斷系統的構建流程可得知，本研究之整合 GRNN 良率預估模式與多類別 SVM 缺陷樣式辨識系統的確能對中、低良率的晶圓做出正確的診斷。

利用本研究之晶圓缺陷樣式診斷系統的構建流程，可使 IC 製造業之製程工程師在每個關鍵製程利用晶圓缺陷檢測機台(例如，KLA)所蒐集的晶圓表面缺陷資訊做即時的製程診斷與晶圓最終良率的預估，並可視診斷結果進一步做適當的製程參數調整以提高良率。再者，製程工程師亦可在晶圓完成製造並經由檢測後，藉由位元圖(bitmap)得到的不良晶片資訊與本研究於每個關鍵製程的製程診斷做比對，以分析可能導致該晶片不良的原因。

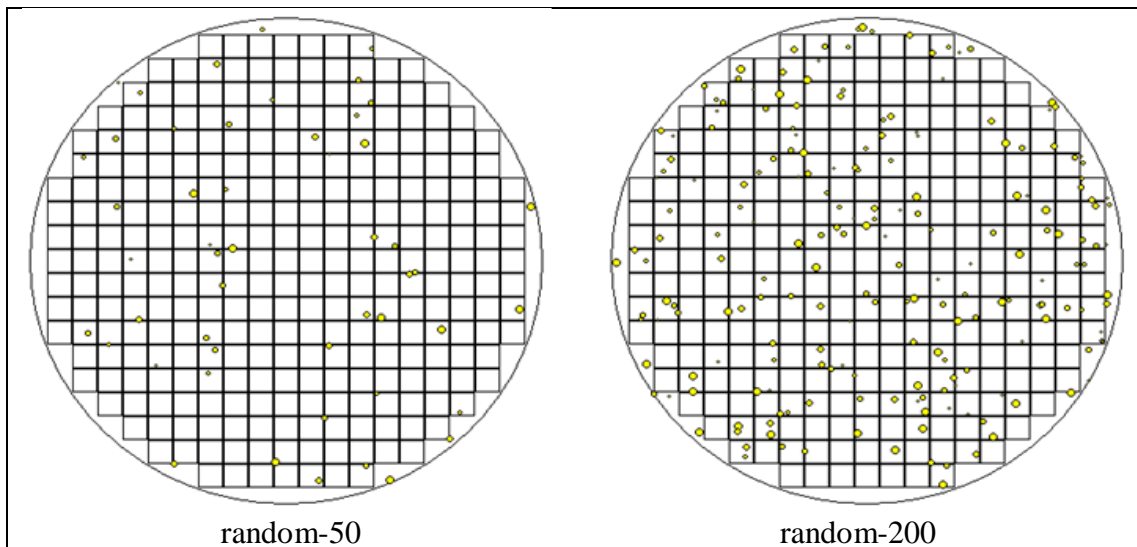


圖 4.30- (a) 隨機樣式

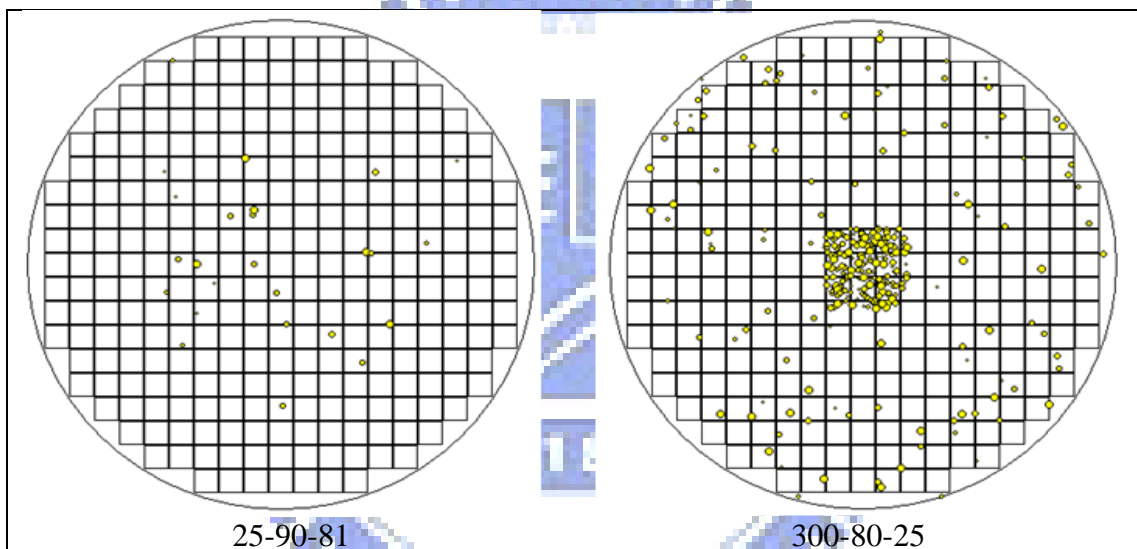


圖 4.30- (b) 牛眼樣式

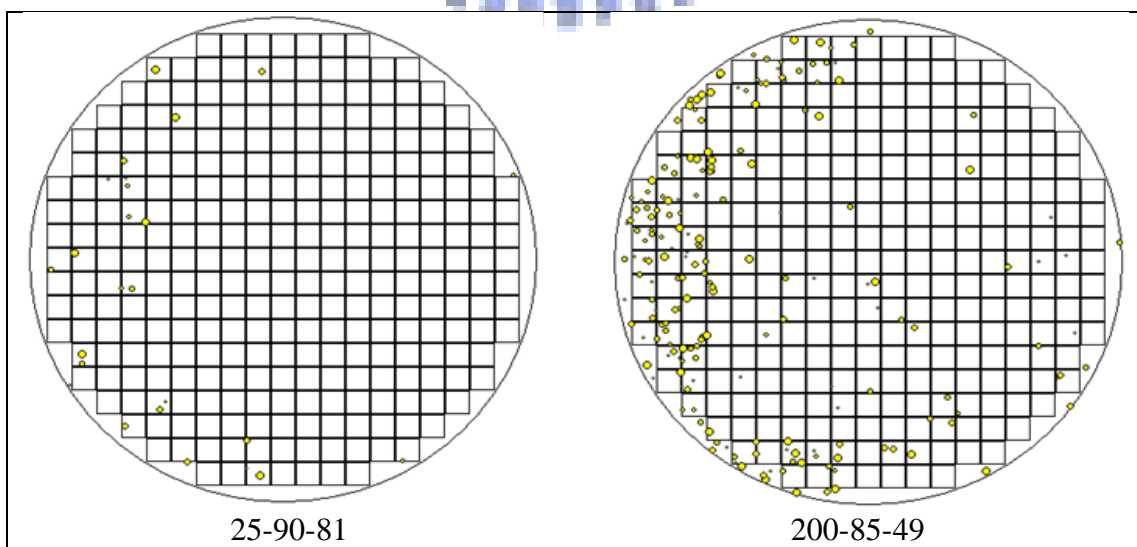


圖 4.30- (c) 弦月樣式

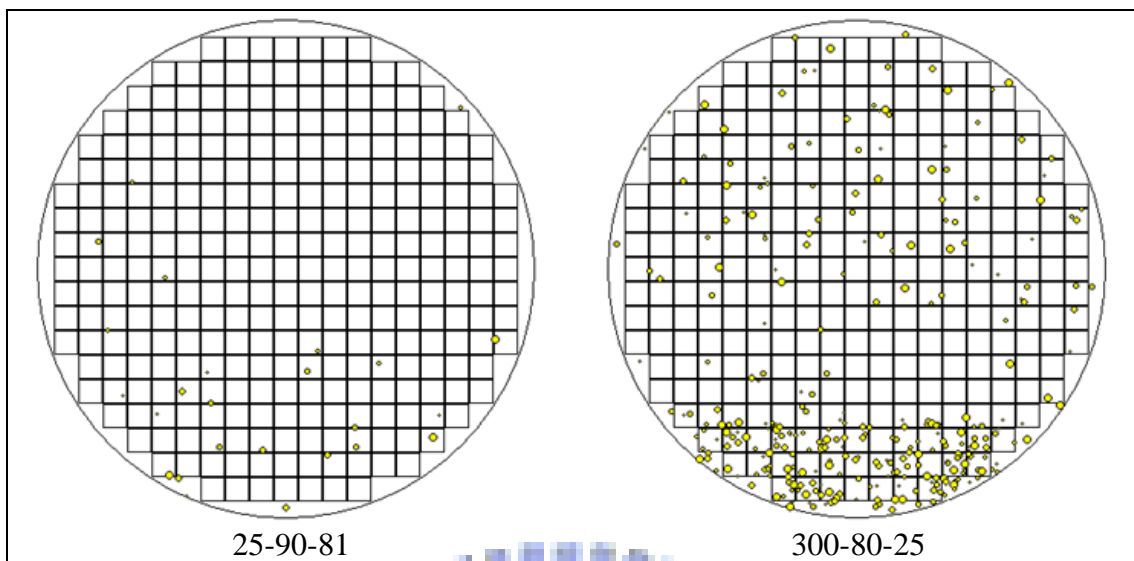


圖 4.30- (d) 底部樣式

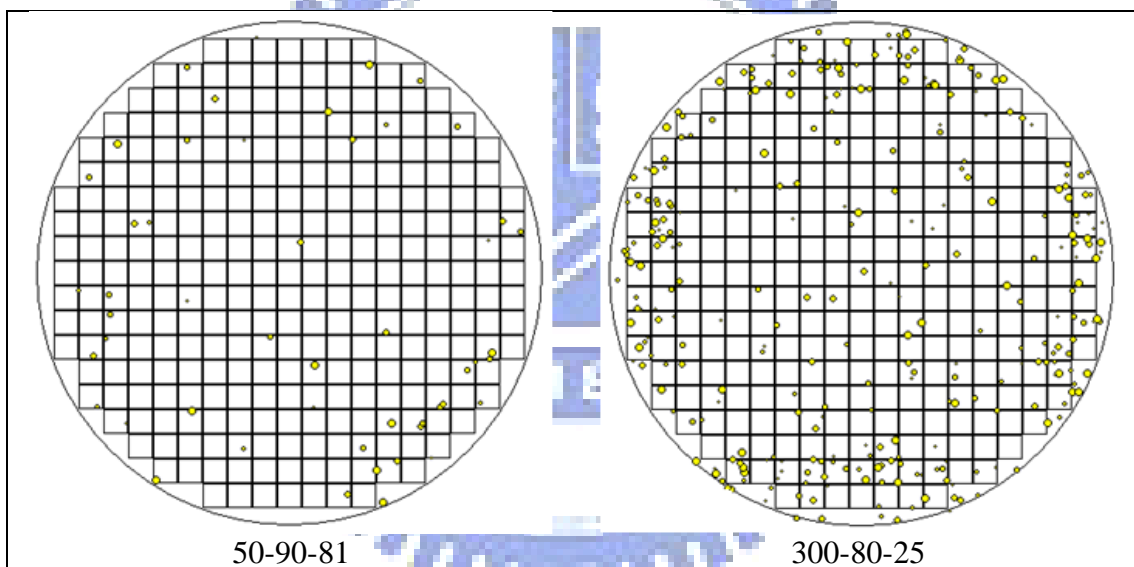


圖 4.30- (e) 環狀樣式

表 4.4 10 片晶圓樣式之相關資訊

編號	缺陷樣式	缺陷數	群聚程度 (%)	群聚分佈面積 (cm ²)	D	CV_A	CV_D	CI_E	真實良率
1	隨機	50	-	-	0.1712	1.1418	0.8087	8.4471	0.7671
2	隨機	200	-	-	0.6849	0.9982	1.0875	10.0615	0.3630
3	牛眼	25	90	81	0.0856	0.7647	1.8420	6.6800	0.8938
4	牛眼	300	80	25	1.0274	1.1146	2.5495	7.2216	0.6678
5	弦月	25	90	81	0.0856	1.1456	1.2198	5.5984	0.9007
6	弦月	200	85	49	0.6849	2.6426	3.7178	7.5699	0.7055
7	底部	25	90	81	0.0856	2.7074	0.8017	6.5272	0.8836
8	底部	300	80	25	1.0274	3.0236	2.4663	8.2907	0.6336
9	環狀	50	90	81	0.1712	0.9681	0.9719	6.8748	0.8219
10	環狀	300	80	25	1.0274	1.3103	2.7298	9.1304	0.4589

表 4.5 10 片晶圓之 GRNN 預估良率與良率程度

編號	缺陷樣式	真實良率	GRNN 預估良率	良率程度
1	隨機	0.7671	0.7816	中良率
2	隨機	0.3630	0.4089	低良率
3	牛眼	0.8938	0.8905	高良率
4	牛眼	0.6678	0.6850	中良率
5	弦月	0.9007	0.9051	高良率
6	弦月	0.7055	0.7067	中良率
7	底部	0.8836	0.8878	高良率
8	底部	0.6336	0.6350	中良率
9	環狀	0.8219	0.8242	高良率
10	環狀	0.4589	0.4629	低良率

表 4.6 6 片晶圓之真實缺陷樣式與多類別 SVM 辨識之樣式

編號	真實缺陷樣式	良率程度	多類別 SVM 辨識樣式
1	隨機	中良率	隨機
2	隨機	低良率	隨機
4	牛眼	中良率	牛眼
6	弦月	中良率	弦月
8	底部	中良率	底部
10	環狀	低良率	環狀

第五章 結論

當晶圓的面積增大時，晶圓上的缺陷並非呈現隨機性分布，而是有群聚現象發生。缺陷群聚現象使得卜瓦松良率模式低估產品的良率；負二項良率模式中的 α 值過於散亂，也可能為負值；複合卜瓦松良率模式的建構複雜，較不易為業界所接受。Jun *et al.* 所建構之良率模式，必需考慮迴歸模型之適配度與基本假設是否違反；利用倒傳遞網路構建良率模式則需考慮網路參數設定問題。此外，目前是以人工的方式分析晶圓上缺陷的空間樣式以找出製程變異的可能原因。然而，人工判定除了費時外，亦可能因誤判進而影響偵測製程變異的準確性。因此，構建一個有效的晶圓缺陷樣式辨識系統為學界與業界重視。許多文獻提出一些晶圓缺陷樣式辨識方法來判定製程是否異常，但亦皆有其不完善之處。

本研究利用一般迴歸神經網路(GRNN)來構建良率模式，並利用多類別支撐向量機(Multi-class SVM)來構建缺陷樣式辨識系統；最後整合成一個良率預估模式及缺陷樣式辨識之晶圓缺陷診斷系統。本研究以模擬實驗來說明本研究所提之整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統的可行性；並進一步與學者所提之良率預估模式與缺陷樣式辨識系統做比較以驗證本研究的有效性與優越性。本研究所提之整合良率預估及缺陷樣式辨識之晶圓缺陷診斷系統的優點如下：

1. 本研究能找到四個切題的晶圓特徵因子(D 、 CV_A 、 CV_D 以及 CI_E)當作 GRNN 良率預估模式之輸入變數來預估晶圓良率；比其他中外文獻所提的良率模式只能找到一個或兩個參數來預估晶圓良率有更準確的預估效果。
2. 本研究所提之多類別 SVM 晶圓缺陷樣式辨識系統在本模擬實驗之個案中，亦比利用 BPNN 網路以及 RBF 神經網路之辨識技術更能夠正確地來辨識晶圓表面的缺陷樣式。

本研究構建之晶圓缺陷樣式診斷系統未來可與晶圓缺陷檢測掃瞄器 KLA 機台整合應用於 IC 生產線上。當製程呈現中、低良率時，可即時追蹤發生原因，以進一步做診斷改善。再者，製程工程師亦可在晶圓完成製造並經由檢測後，藉由位元圖得到的不良晶片資訊與本研究於每個關鍵製程的診斷做比對，以分析可能導致該晶片不良的原因，儘速進行改善措施。

參考文獻

- [1] 李靜宜,「應用倒傳遞類神經網路構建積體電路之良率模式」,國立交通大學工業工程研究所碩士論文,2003。
- [2] 莊達人,『VLSI 製造技術』,高立,2006。
- [3] 許志璋,「整合缺陷點數與群聚指標之積體電路良率模式」,國立交通大學工業工程研究所碩士論文,2003。
- [4] 楊博欽,「應用自組性演算法建構積體電路之良率模式」,國立交通大學工業工程研究所碩士論文,2005。
- [5] 葉怡成,『類神經網路模式應用與實作』,儒林,2003。
- [6] Albin, S.L., Friedman, D.J., “Clustered defects in IC Fabrication: Impact on Process Control Charts.” *IEEE Trans. on Semiconductor Manufacturing*, Vol. 4, No. 1, pp. 36-42, 1991.
- [7] Bhanu, B., Lee, S. and Ming, J., “Adaptive Image Segmentation Using a Genetic Algorithm.” *IEEE Transactions on Systems Man Cybernetics*, Vol. 25, No. 12, pp. 1543-1567, 1995.
- [8] Bottou, L., Cortes, C., Denker, J., Drucker, H., Guyon, I., Jackel, L., LeCun, Y., Muller, U., Sackinger, E., Simard, P. and Vapnik, V., “Comparison of Classifier Methods: A Case Study in Handwritten Digit Recognition.” *Proceedings of the International Conference on Pattern Recognition*. Los Alamitos, CA : IEEE Computer Society Press, 1994.
- [9] Box, G. E. P., Cox, D. R., “An Analysis of Transformations.” *Journal of the Royal Statistical Society, Series B*, Vol. 26, pp. 211-43, 1964.
- [10] Burges, C. J. C., “A Tutorial on Support Vector Machines for Pattern Recognition.” *Data Mining and Knowledge Discovery*, Vol. 2, No. 2, pp. 121-167, 1998.
- [11] Chang, C. C., Lin, C. J., “LIBSVM: A Library for Support Vector Machines.” 2004. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.
- [12] Chen, C. L., Chang, M. H., “Optimal Design of Fuzzy Sliding-Mode Control: A Comparative Study.” *Fuzzy Sets and Systems*, Vol. 93, No. 1, pp. 37-48, 1998.
- [13] Cortes, C., Vapnik, V., “Support Vector Networks.” *Machine Learning*, Vol. 20, No. 3, pp. 273-297, 1995.

- [14] Cunningham, J.A., "The Use and Evaluation of Yield Models in Integrated Circuit Manufacturing." *IEEE Trans. on Semiconductor Manufacturing*, Vol. 3, No. 2, pp. 60-71, 1990.
- [15] David, A., Lerner, B., "Support Vector Machine-based Image Classification for Genetic Syndrome Diagnosis." *Pattern Recognition Letters*, Vol. 26, No. 8, pp. 1029-1038, 2005.
- [16] Duda, R. O., Hart, P. E., *Pattern Classification and Scene Analysis*. John Wiley & Sons, New York, 1973.
- [17] Fausett, L., "*Fundamentals of Neural Networks Architectures, Algorithms, and Applications*." Prentice Hall, 1994.
- [18] Ferris-Prabhu, Alber.V., *Introduction to Semiconductor Device Yield Modeling*. Boston : Artech House, 1992.
- [19] Fiesler, E., "Comparative Bibliography of Ontogenic Neural Networks." *Proceedings of the International Conference on Artificial Neural Networks*, pp. 793-796, Sorrento, Italy, 1994.
- [20] Friedman, D. J., Hansen, M. H., Nair, V. N. and James, D. A., "Model-free Estimation of Defect Clustering in Integrated Circuit Fabrication." *IEEE Transactions on Semiconductor Manufacturing*, Vol. 10, No. 3, pp. 344-359, 1997.
- [21] Hornick, K., Stinchcombe, M. and White, H., "Universal Approximation of An Unknown Mapping and its Derivatives Using Multilayer Feedforward Networks." *Neural Networks*, Vol. 3, No. 5, pp. 551-560, 1990.
- [22] Hsu, C. W., Lin, C. J., "A Comparison of Methods for Multi-class Support Vector Machines." *IEEE Transactions on Neural Networks*, Vol. 13, No. 2, pp. 415-425, 2002.
- [23] Jain, A. K., Duin, R. P. W. and Mao, J., "Statistical Pattern Recognition: A Review." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, pp. 4-37, 2000.
- [24] Jiang, S., Huang, Q., Ye, Q. and Gao, W., "An Effective Method to Detect and Categorize Digitized Traditional Chinese Paintings." *Pattern Recognition Letters*, Vol. 27, No. 7, pp. 734-746, 2006.
- [25] Jun, C.H., Hong, Y., Kim, S.Y., Park, K.S. and Park, H., "A Simulation-based Semiconductor Chip Yield Model Incorporating a New Defect Cluster Index." *Microelectronics Reliability*, Vol. 39, No. 4, pp. 451-456, 1999.

- [26] Keerthi, S. S., Lin, C. J., "Asymptotic Behaviors of Support Vector Machines with Gaussian Kernel." *Neural Computation*, Vol. 15, No. 7, pp. 1667-1689, 2003.
- [27] KreBel, U., "Pairwise Classification and Support Vector Machines." *Advances in kernel methods: Support Vector Learning*, pp. 255-268. Cambridge, MA: MIT Press, 1999.
- [28] Merino, M. A., Cruceta, S., Garcia, A. and Recio, M., "SmartBitTM: Bitmap to Defect Correlation Software for Yield Improvement." *Advanced Semiconductor Manufacturing Conference and Workshop, 2000 IEEE/SEMI*, pp. 194-198. Boston, MA, 2000.
- [29] Murphy, B. T., "Cost-size Optimum of Monolithic Integrated Circuits." *Proc. IEEE*, Vol. 52, pp. 1537-1545, 1964.
- [30] Nemmour, H., and Chibani, Y., "Multiple Support Vector Machines for Land Cover Change Detection: An Application for Mapping Urban Extensions." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 61, No. 2, pp. 125-133, 2006.
- [31] Nieddu, L., Patrizi, G., "Formal Methods in Pattern Recognition." *European Journal of Operation Research*, Vol. 120, No. 3, pp. 459-495, 2000.
- [32] Parzen, E., "On Estimation of a Probability Density Function and Mode." *The Annals of Mathematical Statistics*, Vol. 33, No. 3, pp. 1065-1076, 1962.
- [33] Platt, J. C., Cristianini, N. and Shawe-Taylor, J., "Large Margin DAGs for Multiclass Classification." In: *Advances in neural information processing systems*, pp. 547-553. Cambridge, MA: MIT Press, 2000.
- [34] Raghavachari, M., Srinivasan, A. and Sullo, P., "Poisson Mixture Yield Models for Integrated Circuits: A Critical Review," *Microelectronics Reliability*, Vol. 37, No. 4, pp. 565-580, 1997.
- [35] Schölkopf, B., Smola, A. J., *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA: MIT Press, 2002.
- [36] Specht, D. F., "A General Regression Neural Network," *IEEE Transactions on Neural Networks*, pp. 568-576, 1991.
- [37] Stapper, C. H., and Rosner, R. J., "Integrated Circuit Yield Management and Yield Analysis: Development and Implementation." *IEEE Transactions on Semiconductor Manufacturing*, Vol. 8, No. 2, pp. 95-102, 1995.
- [38] Stapper, C. H., Armstrong, F. M. and Saji, K., "Integrated Circuit Yield

- Statistics,” *Proceedings of The IEEE*, Vol. 71, pp. 453-470, 1983.
- [39] Stapper, C. H., “Defect Density Distribution for LSI Yield Calculations.” *IEEE Transaction on Electron Devices (Correspondence)*, Vol. 20, No. 7, pp. 655-657, 1973.
- [40] Stapper, C. H., “The Effects of Wafer to Wafer Defect Density Variations on Integrated Circuit Defect and Fault Distributions.” *IBM Journal of Research Development*, Vol. 29, No. 1, pp. 87-97, 1985.
- [41] Stapper, C. H., “On Murphy’s Yield Integral.” *IEEE Trans. on Semiconductor Manufacturing*, Vol. 4, No. 4, pp. 294-297, 1991.
- [42] Tong, L. I., Lee, W. I. and Su, C. T., “Using a Neural Network-Based Approach to Predict the Wafer Yield in Integrated Circuit Manufacturing.” *IEEE Trans. On Components, Packaging, and Manufacturing Technology-Port C*, Vol. 20, Iss. 4, pp.288-294, 1997.
- [43] Tong, L. I., Wang, C. H. and Chen, D. L., “Development of a New Cluster Index for Wafer Defects.” *The International Journal of Advanced Manufacturing Technology*, Vol. 31, No. 7-8, pp. 705-715, 2007.
- [44] Tsai, W. J., Tong, L. I. and Wang, C. H., “Developing a New Defect Cluster Index.” *Journal of the Chinese Institute of Industrial Engineers*, Vol. 25, No. 1, pp. 18-30, 2008.
- [45] Tyagi, A., Bayoumi, A. M., “Defect Clustering Viewed Through Generalized Poisson Distribution.” *IEEE Trans. on Semiconductor Manufacturing*, Vol. 5, No. 3, pp. 196-206, 1992.
- [46] Tyagi, A., Bayoumi, M. A., “The Nature of Defect Patterns on Integrated-circuit Wafer Maps.” *IEEE Transactions on Reliability*, Vol. 43, No. 1, pp. 22-29, 1994.
- [47] Vapnik, V., *Statistical learning theory*. New York: Wiley, 1998.
- [48] Widrow, B., Winter, R. G. and Baxter, R. A., “Learning Phenomena in Layered Neural Networks.” *Proceedings of the First IEEE International Conference on Neural Networks*, pp. 411-429. San Diego, 1987.
- [49] Zhou, J., Su, G., Jiang, C., Deng, Y. and Li, C., “A face and Fingerprint Identity Authentication System Based on Multi-route Detection.” *Neurocomputing*, Vol. 70, No. 4-6, pp. 922-931, 2007.