

# 半導體製程晶圓圖像

Big Data Analysis

不是這一組

組長：廖佳馨 H24041228

組員：何欣育 H24041202

廖怡華 H24041147

徐綺蔚 H24041058

王思涵 H24046032

# Content

---

01 研究目標

02 資料介紹

03 發現問題

04 資料特性

05 資料分析

5.1 前期處理

5.2 判斷隨機非隨機

5.3 除噪

5.4 SOM分群 I

5.5 Random forest I

5.6 SOM分群 II

5.7 Random forest II

06 總結

07 參考資料

08 開會紀錄

09 文獻回顧

10 心得

# 01 研究目標

# 1 研究目標與方向

(1) 判斷晶圓圖像為隨機或非隨機

(2) 晶圓缺陷圖分群

透過 SOM 有效地節省人力成本，將資料的樣型分群。

(3) 找出製程問題

利用晶圓圖模型之分類結果，找出各個可能的製程問題。

# 02 資料介紹

## 2 資料介紹

本資料(Data/Wafer\_2012\_Product\_A.txt)是有關於在西元2012年的晶圓片及其晶圓粒測試數據。  
總共有23,115,394筆

原變數	變數名稱	註解	
V1	Lot ID	晶圓片批數	V1 = 1,2,3,4,.....,3447 (除了1724)
V2	Wafer ID	晶圓片批數中第幾片	V2 = 1,2,3,4,.....,25
V3	X	晶圓片上的晶圓粒之X座標	V3 = 11,12,.....,41
V4	Y	晶圓片上的晶圓粒之Y座標	V4 = 11,12,.....,24
V5	Bin Code	測試晶圓粒失敗的測試項目之編碼 (若無失敗項目亦有編碼)	V5 = 1,2,4,8,9,10,11,13,14,15,16,17,19, 20,53,54,80,90,91,92
V6	Test Name	以測試晶圓粒失敗的測試項目之編碼做命名	V6 = "WaferTestBinCodeClass-Good" .....
V7	Group	將晶圓粒分成三類: -1 (TestName為Open的類型) 1 (TestName為Good的類型) 0 (other其他測試失敗的類型)	V7 = -1,0,1
V8	Date	紀錄的日期 (年/月/日)	V8 = "2012-05-01" .....
V9	Time	紀錄的時間 (時/分/秒)	V9 = "00:52:57" .....

## 2 資料介紹

Bin Code , Test Name, Group之對照表

Bin Code	1	2	4	8	9	11	12
Test Name	good	good	Scol	Isb1	Isb2	leakL	leakH
Group	1	1	0	0	0	0	0
Bin Code	13	14	15	16	17	19	20
Test Name	Open	Short	Iccg	Iccn	Iccs	Brow0	Brow1
Group	-1	0	0	0	0	0	0
Bin Code	53	54	80	90	91	92	
Test Name	Srow0	Srow1	MCC	CRC16	CRC16Lo	CRC16Hi	
Group	0	0	0	0	0	0	

# 03 發現問題

# 3 發現問題

## 資料遺失

初步檢視，發現第93004、13398800筆資料只有紀錄錯誤的LotID，其他變數都為遺漏值，因此應將這兩筆資料予以刪除。

### 第93004、13398800筆資料

```
[> data[c(93004,13398800),]  
    LotID WaferID  X  Y BinCode TestName Group Date Time  
1: 2012-10-02      NA NA NA       NA           NA  
2: 2012-05-14      NA NA NA       NA           NA
```

# 3 發現問題

各變數問題檢查

LotID (V1)

a. 資料的命名疏失：

由紀錄的時間、片數與位置我們可以得知：

第93003筆資料的LotID14264，應改成264。

第11650715筆資料的LotID15988，應改成1988。

```
[> data[c(93003,11650715),]  
  LotID WaferID X Y BinCode          TestName Group      Date     Time  
1: 14264        16 12 19           9 WaferTestBinCodeClass-Isb2    0 2012-10-02 08:23:12  
2: 15988        16 12 19           9 WaferTestBinCodeClass-Isb2    0 2012-05-02 08:23:12
```

b. 資料的長度檢視：

經由檢視發現：

LotID的長度為3446，而實際編號到第3447批，再將LotID做排序檢視，發現沒有第1724批。

# 3 發現問題

各變數問題檢查

X, Y (V3, V4)

## 晶圓粒測試的缺失與重疊

藉由程式，我們得出大部分的晶圓片上所測試的晶圓粒個數為344個，但是也有少數幾個是不同的，如以下表格：

LotID	5	6	14	264	1729	1730	1738	1988
WaferID	21	21	23	16	21	21	23	16
Die numbers	364	324	329	358	364	324	329	358

# 3 發現問題

各變數問題檢查

X, Y (V3, V4)

a. 資料的重疊：

先觀察晶圓粒測試個數比344還要多的晶圓片。

利用程式檢查是否有找出測試的晶圓粒是否有重複測試的紀錄。

以Lot 5, Wafer 21為例

No.	Lot ID	Wafer ID	x	y	Bin Code	Test Name	Group	Date	Time
23600	5	21	29	11	9	lsb2	0	2012-10-07	05:03:50
32200	5	21	29	11	1	Good	1	2012-10-06	15:50:10

應對方法：比較同一個lot、同一個wafer中的其他數據之檢測時間，判斷該留下何者

# 3 發現問題

各變數問題檢查

X, Y (V3, V4)

a. 資料的重疊：以Lot 5, Wafer 21為例

```
[> data[c(23590:23605),]  
  V1 V2 V3 V4 V5          V6 V7      V8      V9  
1: 5 21 28 15 1 WaferTestBinCodeClass-Good 1 2012-10-07 05:03:50  
2: 5 21 28 16 1 WaferTestBinCodeClass-Good 1 2012-10-07 05:03:50  
3: 5 21 28 17 1 WaferTestBinCodeClass-Good 1 2012-10-07 05:03:50  
4: 5 21 28 18 53 WaferTestBinCodeClass-Srow0 0 2012-10-07 05:03:50 [> data[c(32195:32203),]  
5: 5 21 28 19 1 WaferTestBinCodeClass-Good 1 2012-10-07 05:03:50      V1 V2 V3 V4 V5          V6 V7      V8      V9  
6: 5 21 28 20 1 WaferTestBinCodeClass-Good 1 2012-10-07 05:03:50 1: 6 21 28 20 90 WaferTestBinCodeClass-CRC16 0 2012-10-06 15:50:11  
7: 5 21 28 21 1 WaferTestBinCodeClass-Good 1 2012-10-07 05:03:50 2: 6 21 28 21 1 WaferTestBinCodeClass-Good 1 2012-10-06 15:50:11  
8: 5 21 28 22 90 WaferTestBinCodeClass-CRC16 0 2012-10-07 05:03:50 3: 6 21 28 22 20 WaferTestBinCodeClass-Brow1 0 2012-10-06 15:50:11  
9: 5 21 28 23 1 WaferTestBinCodeClass-Good 1 2012-10-07 05:03:50 4: 6 21 28 23 53 WaferTestBinCodeClass-Srow0 0 2012-10-06 15:50:11  
10: 5 21 28 24 90 WaferTestBinCodeClass-CRC16 0 2012-10-07 05:03:50 5: 6 21 28 24 1 WaferTestBinCodeClass-Good 1 2012-10-06 15:50:11  
11: 5 21 29 11 9 WaferTestBinCodeClass-Isb2 0 2012-10-07 05:03:50 6: 5 21 29 11 1 WaferTestBinCodeClass-Good 1 2012-10-06 15:50:11  
12: 5 21 29 12 20 WaferTestBinCodeClass-Brow1 0 2012-10-07 05:03:50 7: 5 21 29 12 1 WaferTestBinCodeClass-Good 1 2012-10-06 15:50:11  
13: 5 21 29 13 1 WaferTestBinCodeClass-Good 1 2012-10-07 05:03:50 8: 5 21 29 13 1 WaferTestBinCodeClass-Good 1 2012-10-06 15:50:11  
14: 5 21 29 14 1 WaferTestBinCodeClass-Good 1 2012-10-07 05:03:50 9: 5 21 29 14 1 WaferTestBinCodeClass-Good 1 2012-10-06 15:50:11  
15: 5 21 29 15 1 WaferTestBinCodeClass-Good 1 2012-10-07 05:03:50  
16: 5 21 29 16 1 WaferTestBinCodeClass-Good 1 2012-10-07 05:03:50
```

針對此例，應留下左邊的數據，而右邊的數據先放一邊。

# 3 發現問題

## 各變數問題檢查

X, Y (V3, V4)

### b. 資料的缺失：

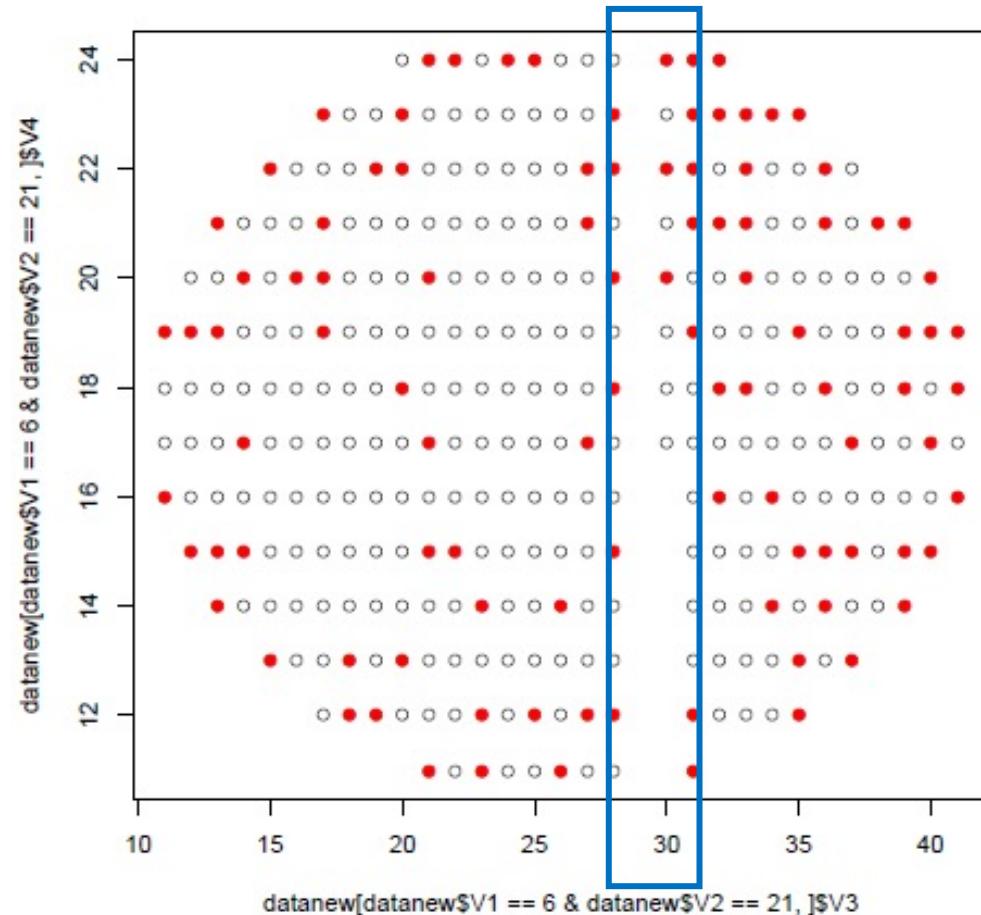
觀察晶圓粒測試個數比344還要少的晶圓片。

#### 以Lot 6 Wafer 21為例

晶圓粒測試個數為324個，  
缺失的20個點如右圖所示。

#### 應對方法：

1. 對照測試日期與時間之後，發現缺失的點恰好為Lot 5, Wafer21重複的晶圓粒，所以將資料合併起來。
2. 其他無法填補的晶圓片，則暫時放一邊。  
(Lot 14 Wafer 23 ; Lot 1738 Wafer 23)



# 3 發現問題

資料大面積重複

Lot1 ~ 1723 與 Lot1725 ~ 3447

舉例來說，LotID 1 與 1725的前十項對比：

Lot 1, Wafer 1

```
[> datanew[datanew$V1==1,][1:10,]
  V1 V2 V3 V4 V5          V6 V7     V8          V9
1:  1  1 11 16  9 WaferTestBinCodeClass-Isb2 0 2012-10-13 23:22:16
2:  1  1 11 17  9 WaferTestBinCodeClass-Isb2 0 2012-10-13 23:22:16
3:  1  1 11 18  9 WaferTestBinCodeClass-Isb2 0 2012-10-13 23:22:16
4:  1  1 11 19 11 WaferTestBinCodeClass-leakL 0 2012-10-13 23:22:16
5:  1  1 12 15  9 WaferTestBinCodeClass-Isb2 0 2012-10-13 23:22:16
6:  1  1 12 16 90 WaferTestBinCodeClass-CRC16 0 2012-10-13 23:22:16
7:  1  1 12 17 53 WaferTestBinCodeClass-Srow0 0 2012-10-13 23:22:16
8:  1  1 12 18  1 WaferTestBinCodeClass-Good  1 2012-10-13 23:22:16
9:  1  1 12 19 90 WaferTestBinCodeClass-CRC16 0 2012-10-13 23:22:16
10: 1  1 12 20  9 WaferTestBinCodeClass-Isb2 0 2012-10-13 23:22:16
```

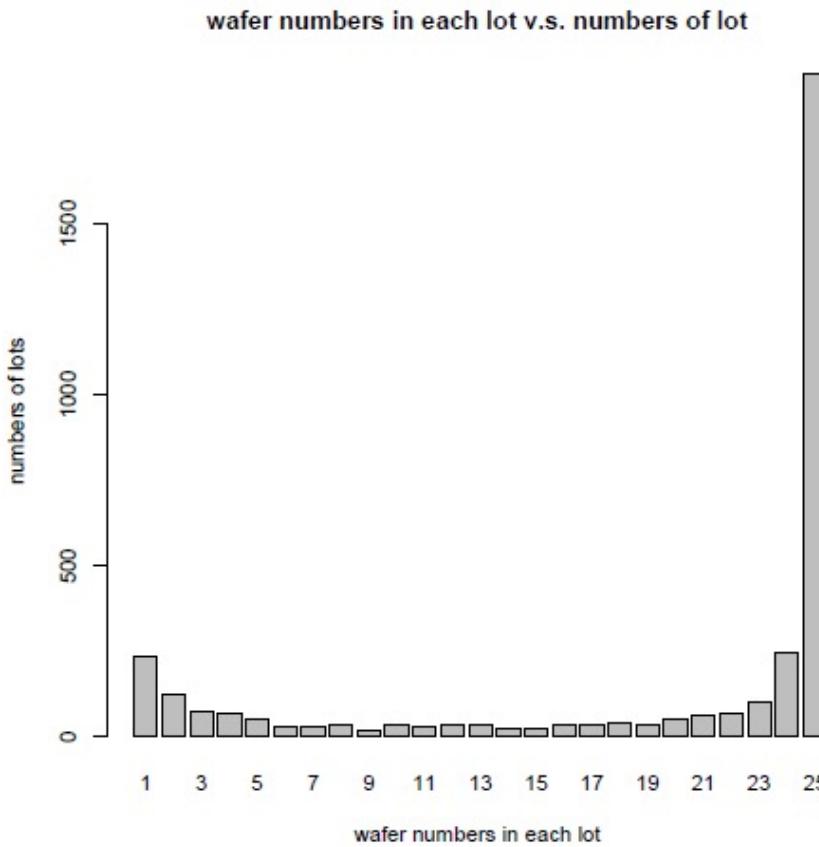
Lot 1725, Wafer 1

```
[> datanew[datanew$V1==1725,][1:10,]
  V1 V2 V3 V4 V5          V6 V7     V8          V9
1:  1  1 11 16  9 WaferTestBinCodeClass-Isb2 0 2012-05-13 23:22:16
2:  1  1 11 17  9 WaferTestBinCodeClass-Isb2 0 2012-05-13 23:22:16
3:  1  1 11 18  9 WaferTestBinCodeClass-Isb2 0 2012-05-13 23:22:16
4:  1  1 11 19 11 WaferTestBinCodeClass-leakL 0 2012-05-13 23:22:16
5:  1  1 12 15  9 WaferTestBinCodeClass-Isb2 0 2012-05-13 23:22:16
6:  1  1 12 16 90 WaferTestBinCodeClass-CRC16 0 2012-05-13 23:22:16
7:  1  1 12 17 53 WaferTestBinCodeClass-Srow0 0 2012-05-13 23:22:16
8:  1  1 12 18  2 WaferTestBinCodeClass-Good  1 2012-05-13 23:22:16
9:  1  1 12 19 90 WaferTestBinCodeClass-CRC16 0 2012-05-13 23:22:16
10: 1  1 12 20  9 WaferTestBinCodeClass-Isb2 0 2012-05-13 23:22:16
```

# 04 資料特性

# 4 發現的資料特性

## 每批晶圓片數之長條圖



總共晶圓片數 = 67196 片

Wafer Number	Lot Number	Wafer Number	Lot Number
1	236	14	24
2	124	15	22
3	72	16	32
4	68	17	36
5	52	18	38
6	28	19	34
7	30	20	50
8	32	21	60
9	20	22	68
10	34	23	100
11	30	24	246
12	34	25	1942
13	34		

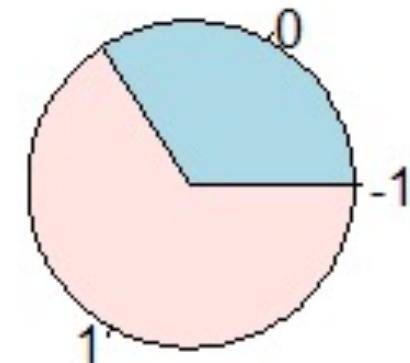
# 4 發現的資料特性

晶圓片測試結果的類型比例

a. 各種Group之比例

ClassName	Open	Other(failed)	Good	Total
Group	-1	0	1	
Number	17566	7852496	15245332	23115394
Percentage	0.08%	33.97%	65.95%	100%

group pie chart



b. 探討Group = 1中各種bincode的比例

Bincode	1	2	Total
	good	good	
ClassName			
Number	14245334	999998	15245332
Percentage	93.44%	6.56%	100%

# 4 發現的資料特性

晶圓片測試結果的類型比例

c. 探討Group = 0中各種失敗的比例

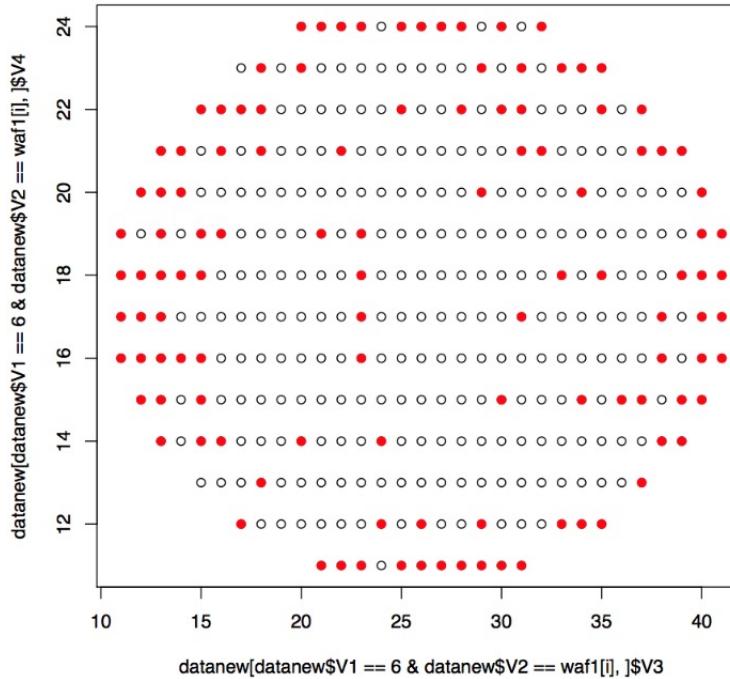
Bincode	90	53	4	9	20	11
ClassName	CRC16	Srow0	Scol	lsb2	Brow1	leakL
Number	2965030	1369458	1277152	718724	464026	304110
Percentage	37.76%	17.44%	16.26%	9.15%	5.90%	3.87%
Bincode	19	54	12	14	91	8
ClassName	Brow0	Srow1	leakH	Short	CRC16Lo	lsb1
Number	264754	206464	108306	73150	39146	33838
Percentage	3.37%	2.63%	1.38%	0.93%	0.5%	0.43%
Bincode	80	92	16	15	17	
ClassName	MCC	CRC16Hi	Iccn	Iccg	Iccs	Total
Number	19080	8930	98	52	178	7852496
Percentage	0.24%	0.11%	0%	0%	0%	100%

# 4 發現的資料特性

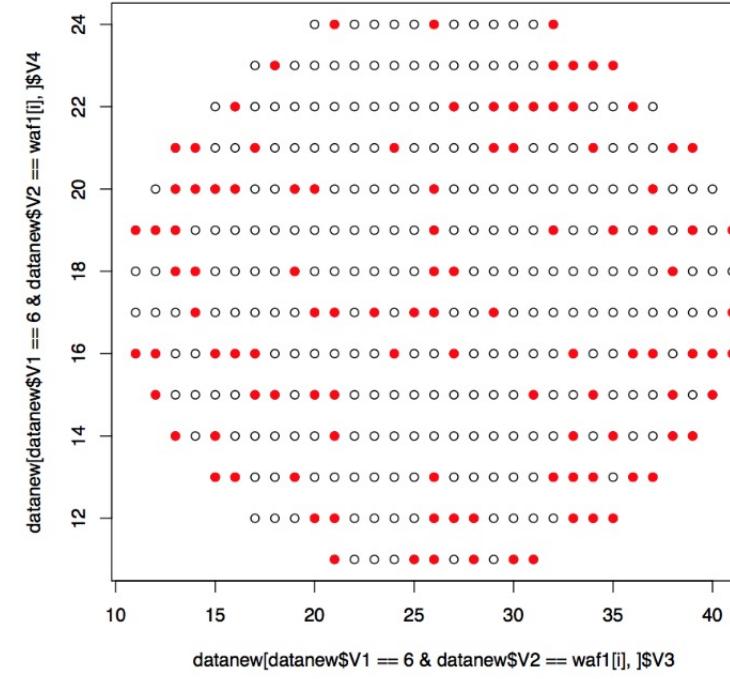
晶圓片測試結果圖

Defect Map 缺陷圖

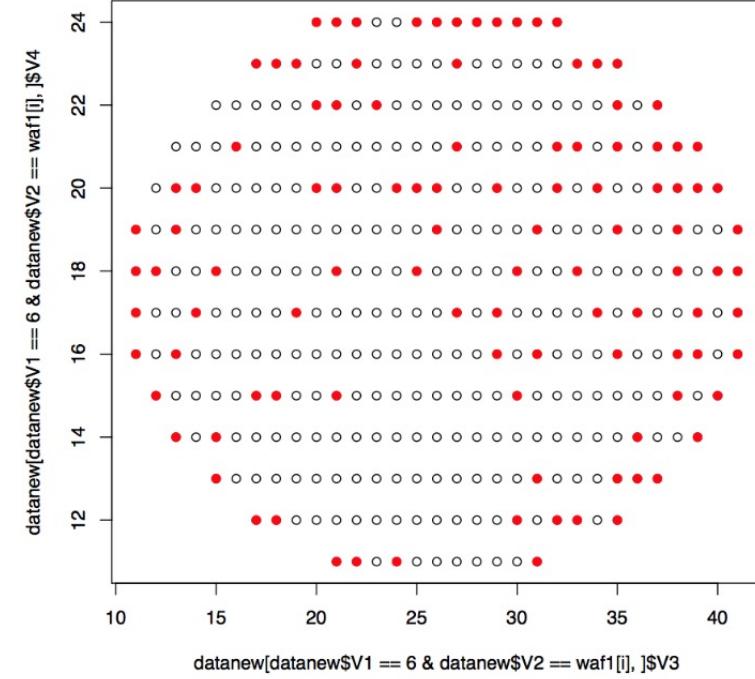
Lot 6, Wafer 1



Lot 6, Wafer 2



Lot 6, Wafer 3

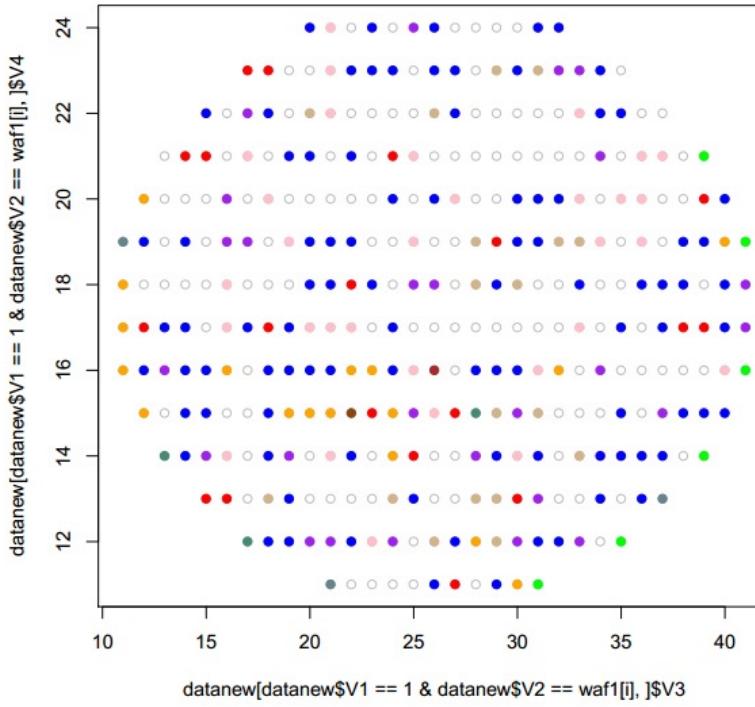


# 4 發現的資料特性

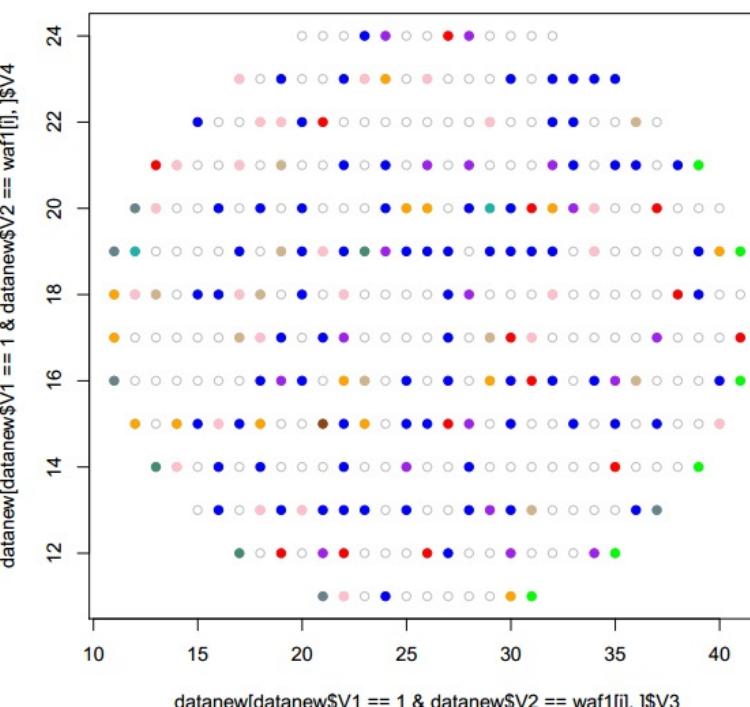
晶圓片測試結果圖

BIN Map 針測圖

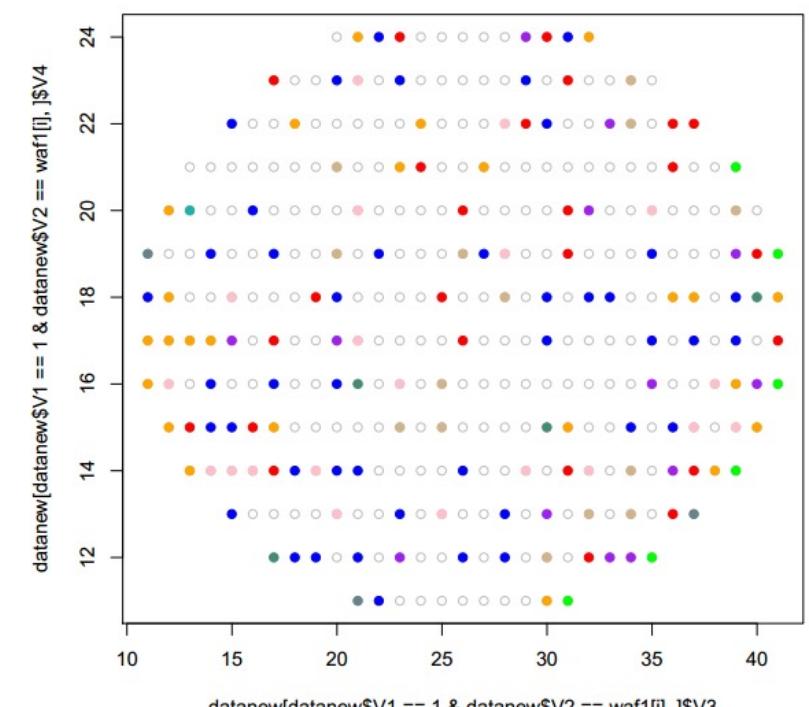
Lot 1, Wafer 1



Lot 1, Wafer 2



Lot 1, Wafer 3



# 05 資料分析

— 前期處理

# 5 資料分析 —— 前期處理

## 統整

1. 刪除：第93004、13398800筆資料
2. 更改：
  - 第93003筆，LotID14264改成264。
  - 第11650715筆，LotID15988改成1988。
  - 第32200~32219筆，LotID5改成6。
  - 第11589911~11589930筆，LotID1729改成1730
3. 放一邊：
  - 不足一片(344個) Lot14 Wafer 23、Lot1738 Wafer 23
  - 超過一片            Lot264 Wafer 16、Lot1988 Wafer 16 多出來的點放一邊

# 5 資料分析 —— 前期處理

## 轉換成矩陣形式

可更好地記錄晶圓片對應位置的晶圓粒狀況 (good die or bad die)

## 以Lot 1, Wafer 1為例

Good記錄0，Bad則記錄1 (Open也視同Bad)

	X	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	V22	V23	V24	V25	V26	V27	V28	V29	V30	V31			
1	1	NA	1	1	1	1	1	1	0	0	0	1	1	1	NA																				
2	2	NA	1	1	1	1	1	1	0	0	1	1	1	0	0	0	1	1	1	1	1	NA	NA	NA	NA	NA	NA								
3	3	NA	NA	NA	NA	NA	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	NA	NA	NA	NA	NA	
4	4	NA	NA	0	0	1	1	1	1	1	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	1	NA	NA	NA		
5	5	NA	1	0	0	1	1	1	1	1	1	1	0	0	0	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	NA		
6	6	1	1	0	0	0	0	1	0	0	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
7	7	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	0	1	1	1	0	1	1	1	1	1	1		
8	8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	1	1		
9	9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	1	1	1	1		
10	10	NA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	NA	NA		
11	11	NA	NA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	NA	NA		
12	12	NA	NA	NA	NA	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	NA	NA	NA	NA	NA
13	13	NA	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	0	1	NA	NA	NA	NA	NA	NA							
14	14	NA	NA	NA	NA	1	1	1	1	0	0	0	0	1	1	1	1	1	1	NA	NA	NA	NA	NA	NA										

# 05 資料分析

## —— 判斷隨機或非隨機

### Spatial Randomness test

# 5 資料分析 —— 判斷隨機或非隨機

Spatial Randomness test

Log Odds Ratio Test

採用King-move neighborhood rule

	i	j
j	j	j

$$Y_i = 0, \text{Good die}$$

$$Y_i = 1, \text{Bad die}$$

$$\delta_{ij} = \begin{cases} 1 & , \text{two dies } i \text{ and } j \text{ are neighbors} \\ 0 & , \text{otherwise} \end{cases}$$

i j

$$N_{GG} = \sum_{i<j} \sum_{\substack{i \\ \downarrow \\ j}} \delta_{ij} (1 - Y_i)(1 - Y_j)$$

$$N_{GB} = \sum_{i<j} \sum_{\substack{i \\ \downarrow \\ j}} \delta_{ij} (1 - Y_i) Y_j$$

$$N_{BG} = \sum_{i<j} \sum_{\substack{i \\ \downarrow \\ j}} \delta_{ij} Y_i (1 - Y_j)$$

$$N_{BB} = \sum_{i<j} \sum_{\substack{i \\ \downarrow \\ j}} \delta_{ij} Y_i Y_j$$

# 5 資料分析 —— 判斷隨機或非隨機

Spatial Randomness test

Log Odds Ratio Test

bad		
good	i	

$H_0$  : there is a random pattern

$H_1$  : there is a non-random pattern

$$\hat{\theta} = \log\left(\frac{(N_{GG} + 0.5)(N_{BB} + 0.5)}{(N_{GB} + 0.5)(N_{BG} + 0.5)}\right)$$

$$\hat{\theta} \sim N(0, \sigma^2) \text{ as } n \rightarrow \infty$$

$$\sigma^2 = \left( \frac{1}{N_{GG} + 0.5} + \frac{1}{N_{BG} + 0.5} + \frac{1}{N_{GB} + 0.5} + \frac{1}{N_{BB} + 0.5} \right)$$

若只看中間晶圓粒 i

$$\begin{aligned} N_{GG} &= 1 & N_{BG} &= 0 \\ N_{GB} &= 3 & N_{BB} &= 0 \end{aligned}$$

# 5 資料分析 —— 判斷隨機或非隨機

Spatial Randomness test

Log Odds Ratio Test

## 判定標準

選取標準常態累積機率值介於0.4~0.6之間的晶圓圖  
為隨機性故障之圖形

## 判定結果

隨機性圖形片數：1864片

非隨機性圖形片數：65330 片 +2片 = 65332片

不足一片的缺陷晶圓片，主觀判斷

# 5 資料分析 —— 判斷隨機或非隨機

Spatial Randomness test

Log Odds Ratio Test

以Lot 6, Wafer 1為例

$H_0$  : there is a random pattern

$H_1$  : there is a non-random pattern

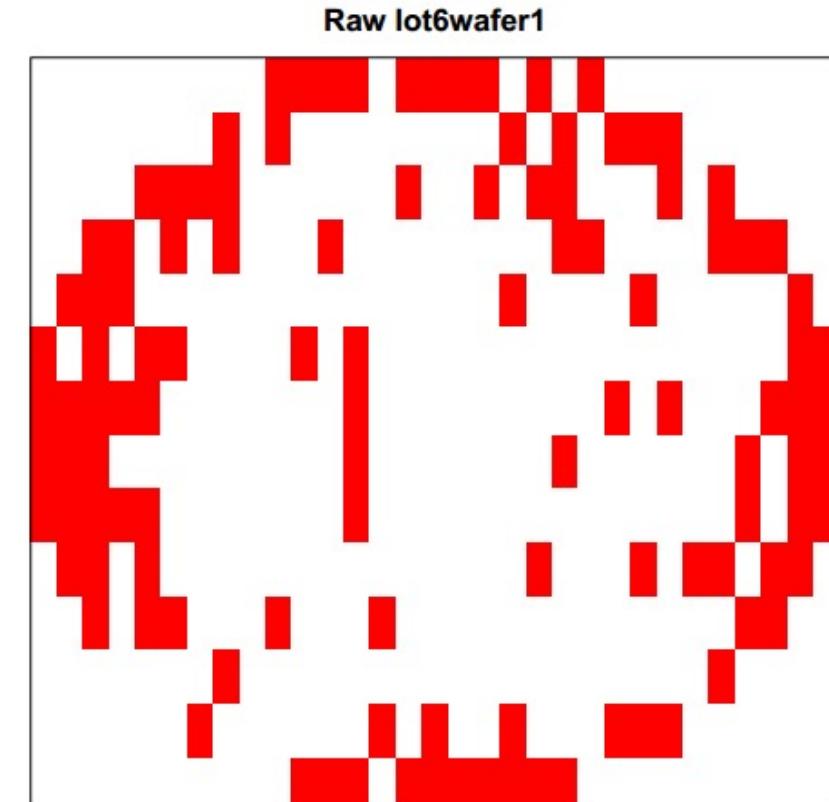
$$N_{GG} = 690 \quad N_{BG} = 208$$

$$N_{GB} = 203 \quad N_{BB} = 162$$

$$\hat{\theta} = \log \left( \frac{(N_{GG} + 0.5)(N_{BB} + 0.5)}{(N_{GB} + 0.5)(N_{BG} + 0.5)} \right) = 0.972489 \quad \hat{\theta} \sim N(0, \sigma^2) \text{ as } n \rightarrow \infty$$

$$\sigma^2 = \left( \frac{1}{N_{GG} + 0.5} + \frac{1}{N_{BG} + 0.5} + \frac{1}{N_{GB} + 0.5} + \frac{1}{N_{BB} + 0.5} \right) = 0.017312$$

$Z = 7.391082, \Phi(Z) = 1 > 0.6 \rightarrow$  判定為不隨機



# 5 資料分析 —— 判斷隨機或非隨機

Spatial Randomness test

Log Odds Ratio Test

以Lot 2, Wafer 8為例

$H_0$  : there is a random pattern

$H_1$  : there is a non-random pattern

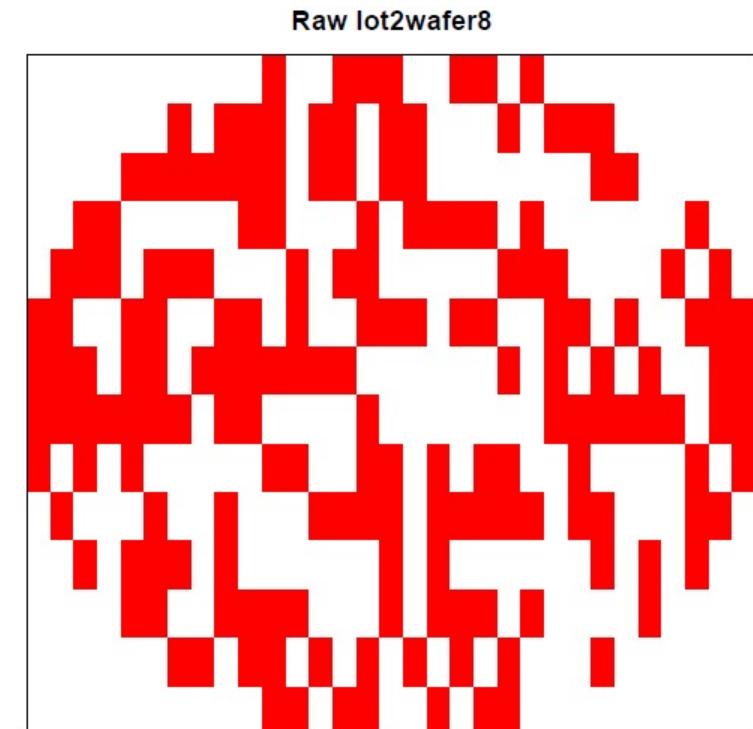
$$N_{GG} = 313 \quad N_{BG} = 316$$

$$N_{GB} = 312 \quad N_{BB} = 322$$

$$\hat{\theta} = \log \left( \frac{(N_{GG} + 0.5)(N_{BB} + 0.5)}{(N_{GB} + 0.5)(N_{BG} + 0.5)} \right) = 0.0220 \quad \hat{\theta} \sim N(0, \sigma^2) \text{ as } n \rightarrow \infty$$

$$\sigma^2 = \left( \frac{1}{N_{GG} + 0.5} + \frac{1}{N_{BG} + 0.5} + \frac{1}{N_{GB} + 0.5} + \frac{1}{N_{BB} + 0.5} \right) = 0.01$$

$Z = 0.1954, 0.4 \leq \Phi(Z) = 0.5774 \leq 0.6 \rightarrow \text{判定為隨機}$



# 05 資料分析

# —— 除噪 (1) 進退化法則 (2) Median Filter

# 5 資料分析 —— 除噪

Noise Reduction

進退化法則

↓ 晶圓粒i的鄰居位置權重 ↓

0.5	1	0.5
1	i	1
0.5	1	0.5

$$Y_i = 0, \text{Good die}$$

$$Y_i = 1, \text{Bad die}$$

$$\delta_{ij} = \begin{cases} 1 & , \text{two dies } i \text{ and } j \text{ are neighbors} \\ 0 & , \text{otherwise} \end{cases}$$

$$\omega_{ij} = \begin{cases} 1 & , \text{上下左右} \\ 0.5 & , \text{斜對角} \end{cases}$$

$$score_i = \sum \delta_{ij} \omega_{ij} Y_j$$

(a)4-1

$$Y'_i = \begin{cases} 1, & \text{if } score_i \geq 4 \\ Y_i, & \text{if } 4 > score_i \geq 1 \\ 0, & \text{if } score_i < 1 \end{cases}$$

(b)5-2

$$Y'_i = \begin{cases} 1, & \text{if } score_i \geq 5 \\ Y_i, & \text{if } 5 > score_i \geq 2 \\ 0, & \text{if } score_i < 2 \end{cases}$$

# 5 資料分析 —— 除噪

Noise Reduction

- Bad die
- Good die

Median Filter

將落在*i*附近的die(包含自己)寫成 $Y_{ij}$ 的序列，找出 $\text{median}(m_i)$

$$Y_i = 0, \text{Good die}$$
$$Y_i = 1, \text{Bad die}$$

$$Y'_i = \begin{cases} 1, m_i = 1 \\ Y_i, m_i = 0.5 \\ 0, m_i = 0 \end{cases}$$

濾噪

0	0	0
0	1(i)	0
0	1	1

$$\{0,0,0,0,1,0,0,1,1\} \quad m_i = 0$$

$$Y'_i = 0$$

補值

1	1	1
0	0(i)	1
0	1	1

$$\{1,1,1,1,0,0,0,1,1\} \quad m_i = 1$$

$$Y'_i = 1$$

一樣時

	1	
0(i)	1	1
0	0	1

$$\{1,1,0,0,0,1\} \quad m_i = 0.5$$

$$Y'_i = Y_i = 0$$

# 5 資料分析 —— 除噪

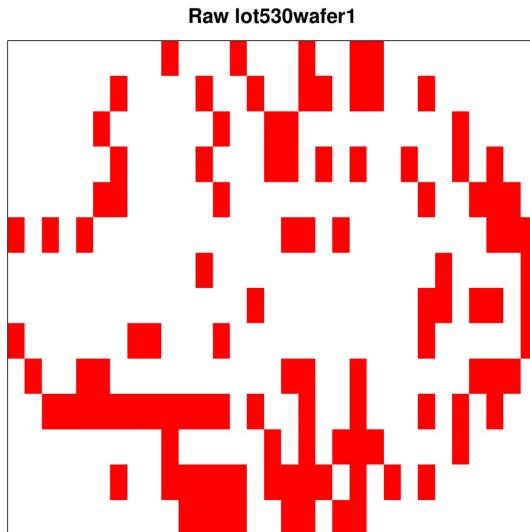
Noise Reduction

除噪選擇

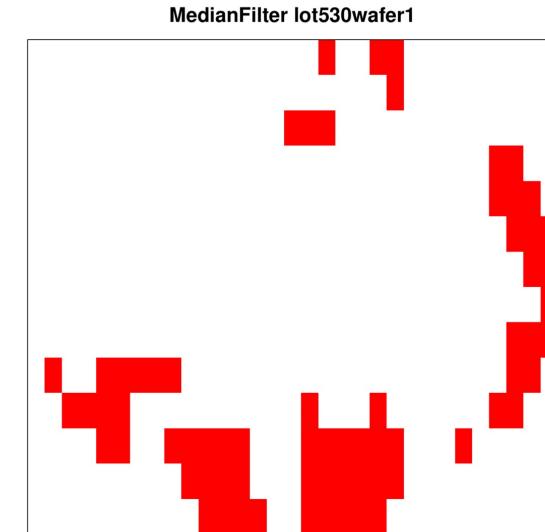
以Lot 530為例

由於Median Filter的效果較佳，因此我們最終決定使用經過Median Filter除噪的資料進行分析。

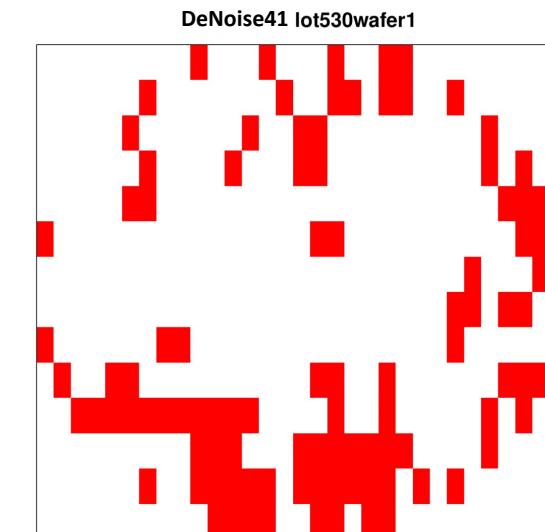
原始圖



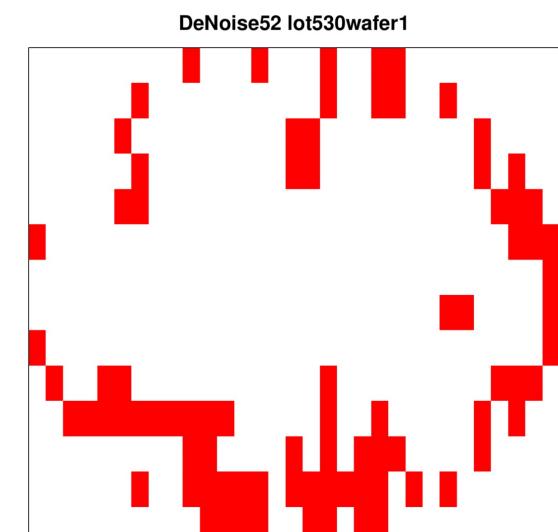
Median filter



4-1



5-2



# 5 資料分析 —— 除噪

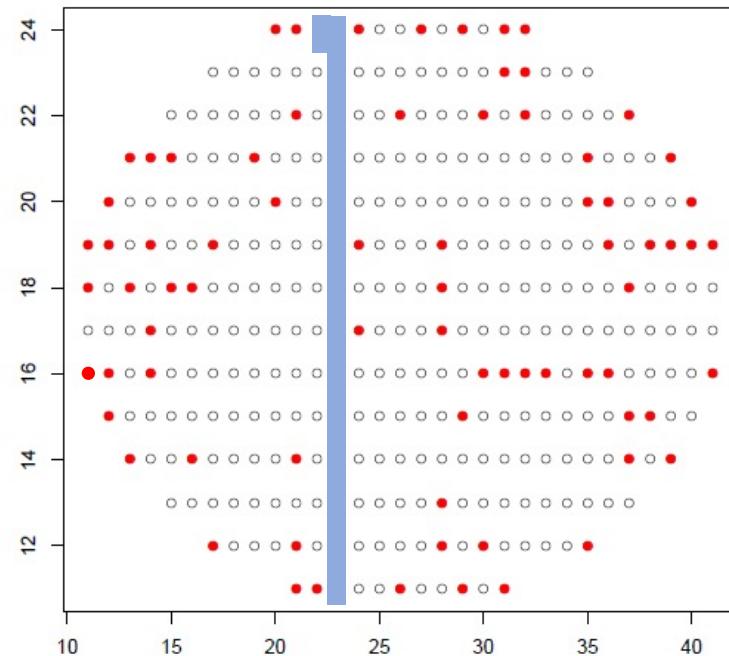
Noise Reduction

Median Filter

缺陷晶圓片處理

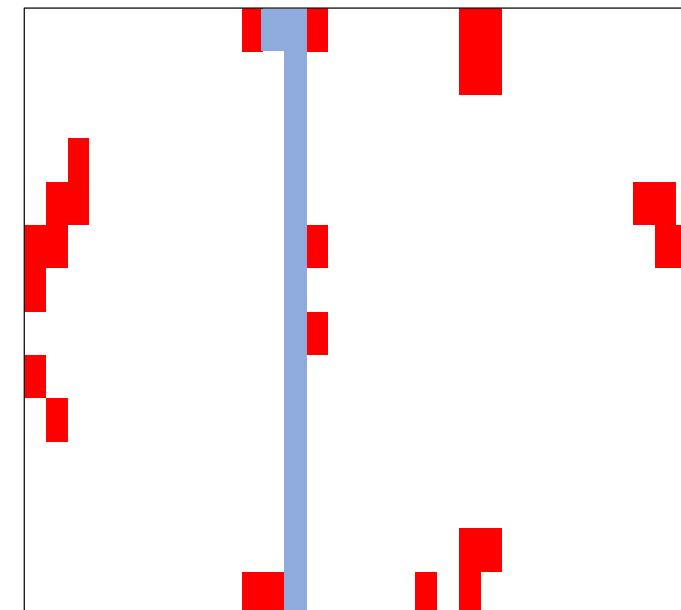
缺陷的晶圓片為lot 14 wafer 23與lot 1738 wafer 23  
將遺失晶粒周圍維持原狀，其他部分一樣用median除噪。

原始圖



Median filter

MedianFilter Lot14 Wafer23



# 05 資料分析

— SOM分群

# 5 資料分析 —— SOM分群

## 目的

將良率異常類別有效率分群

可透過 SOM 有效地節省人力成本、減少肉眼的判斷失誤，  
將資料的樣型分群，並標上label。

## 想法

除Random外之配適

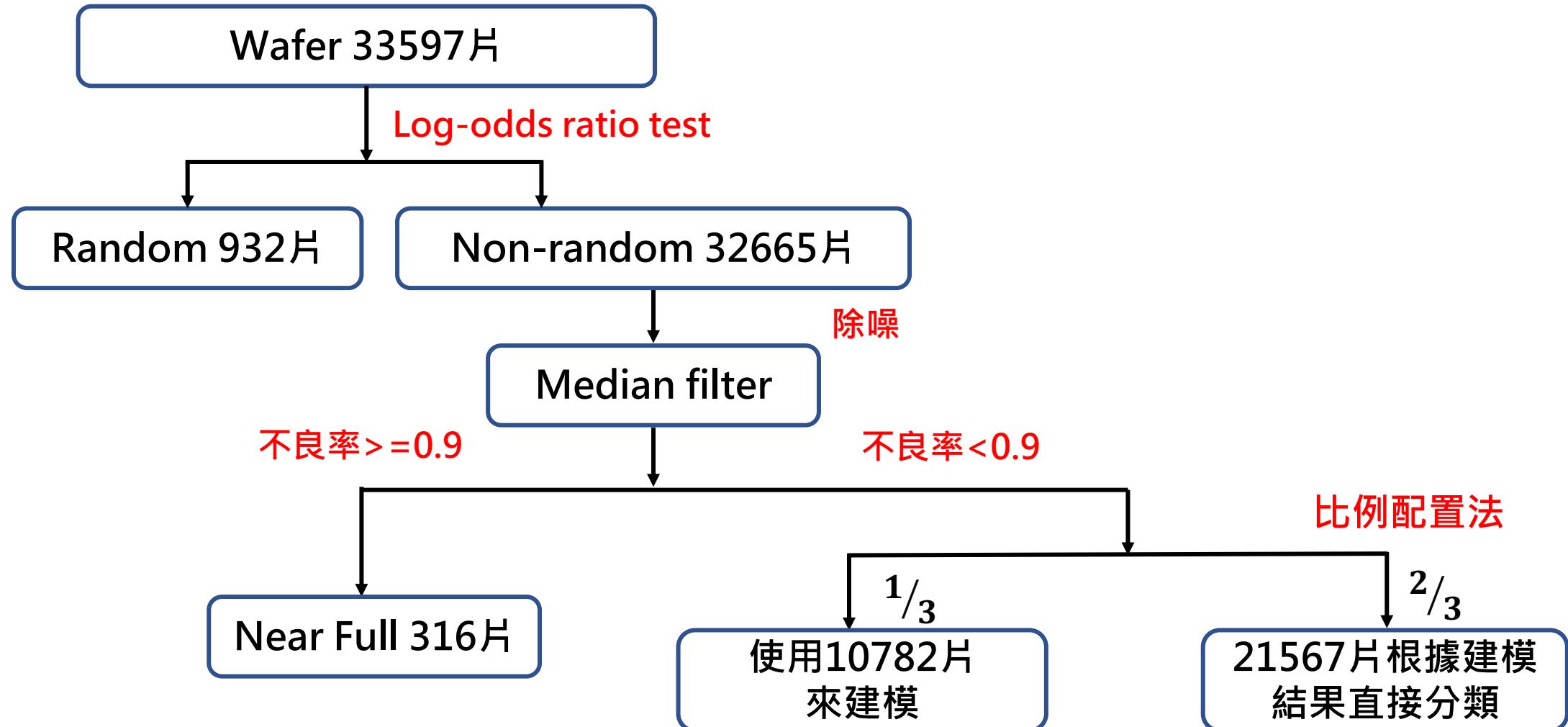
因為在前期處理，已經有良好分辨隨機與非隨機。  
因此，想嘗試將除Random外之晶圓片進行配適，提升分群效果。

## 使用資料集合

Lot1 ~ Lot 1723

# 5 資料分析 —— SOM分群

## 分群流程圖



# 5 資料分析 —— SOM分群

## Median-filter 各區間抽樣數

### 比例配置法

只針對排除Random、Near full( $>=0.9$ )  
後剩下的晶圓片

每個區間抽取三分之一樣本數來建模

不良率	建模	不建模	total
$p \leq 0.1$	1574	3147	4721
$0.1 < p \leq 0.2$	4958	9916	14874
$0.2 < p \leq 0.3$	2529	5059	7588
$0.3 < p \leq 0.4$	878	1757	2635
$0.4 < p \leq 0.5$	382	765	1147
$0.5 < p \leq 0.6$	216	432	648
$0.6 < p \leq 0.7$	116	232	348
$0.7 < p \leq 0.8$	77	154	231
$0.8 < p < 0.9$	52	105	157
total	10782	21567	32349

# 5 資料分析 —— SOM分群

## SOM 設定參數

$n$  : training 的晶圓片片數 = 10782片

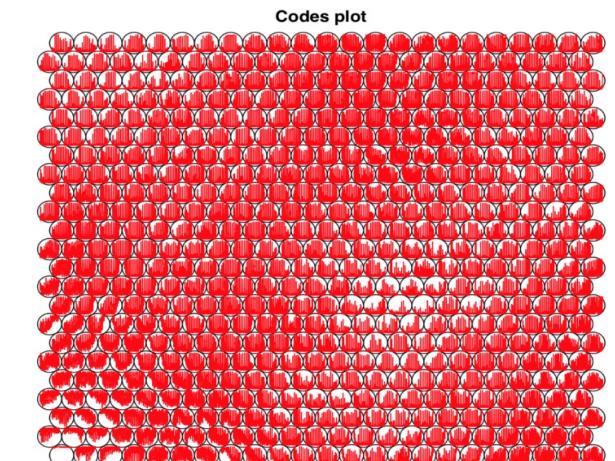
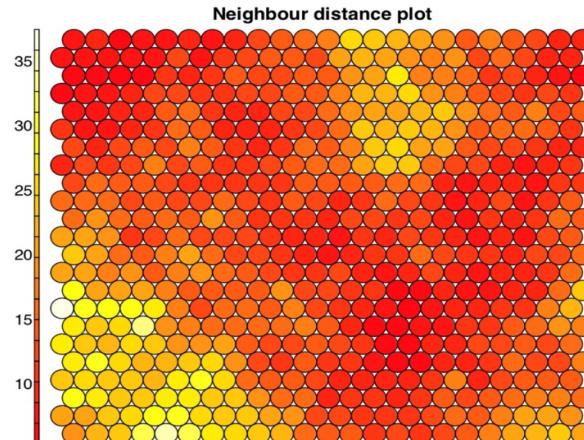
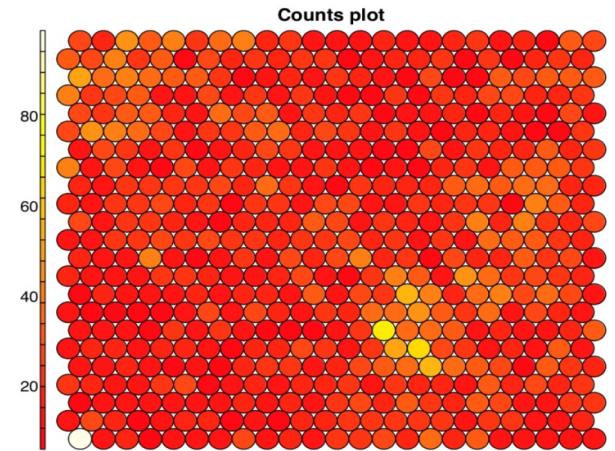
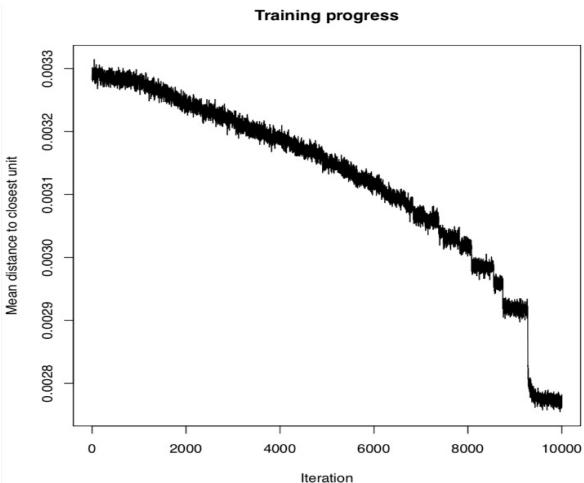
輸出層大小 :  $\sqrt{5 * n} * \sqrt{5 * n} = 23 * 23$

網路拓樸結構 : hexagonal

最大迭代次數 : 10000

學習率 : 0.05~0.01

## 分群結果



# 5 資料分析 —— SOM分群

分群結果

分群準則

Step 1.

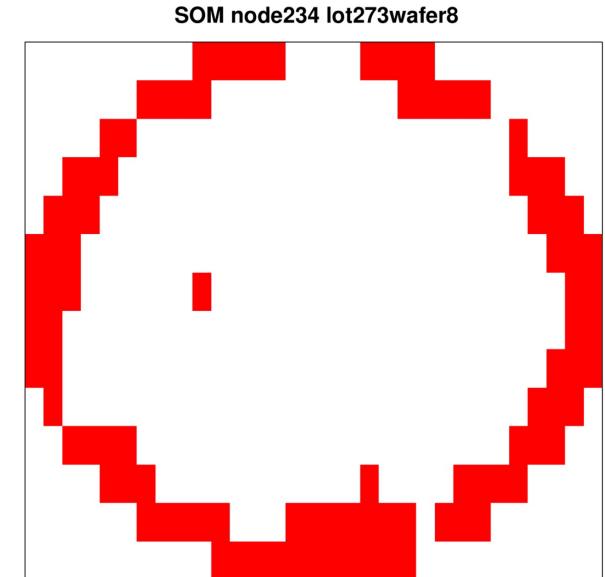
根據SOM建模結果，以 node 為單位，人工檢視一個node的所有晶圓片，採多數決的方式決定此 node 的 pattern (一個node只有一種pattern)，以此類推，將所有node都標好。

※ Pattern可選擇種類(8類)，是根據之前大致瀏覽的結果決定。

Step 2.

將其餘2/3之晶圓片放入模型做預測，記錄分到的node，並根據對應node的pattern，標上label。

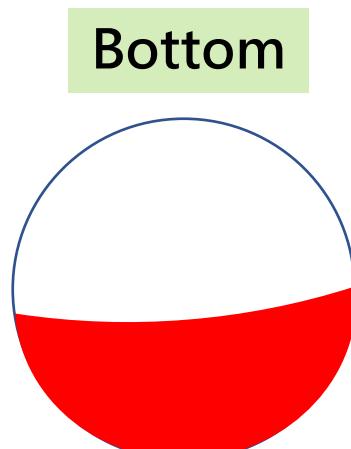
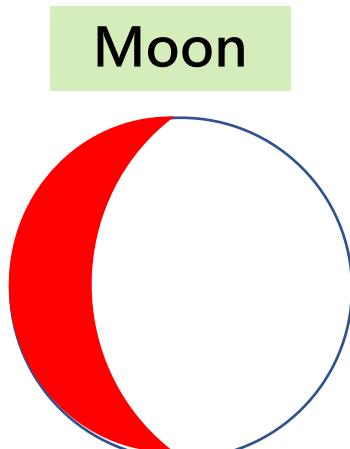
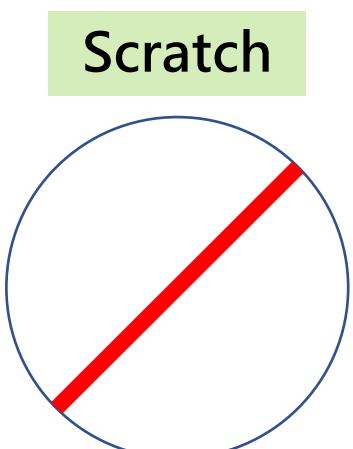
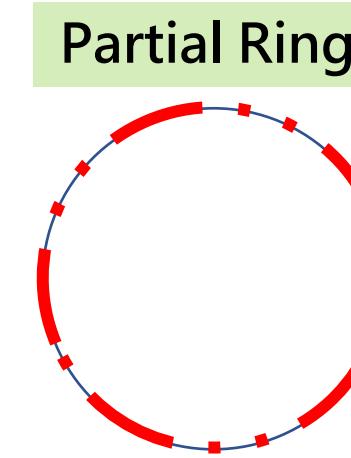
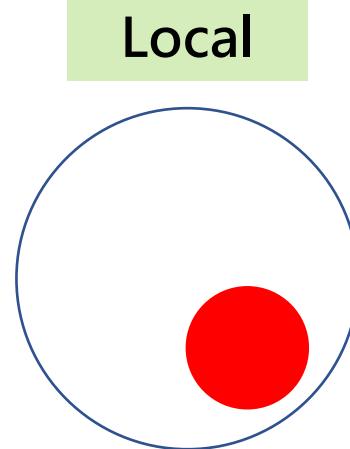
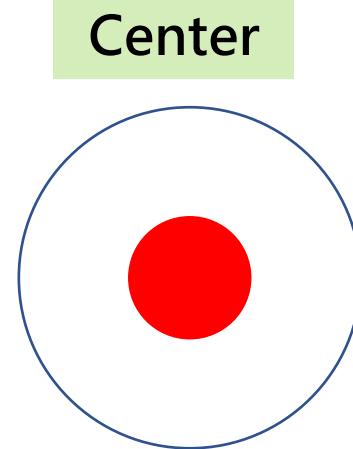
以node234為例



# 5 資料分析—— SOM分群

分群結果

SOM分群結果共八類



# 5 資料分析 —— SOM分群

## 分群結果

各pattern個  
數

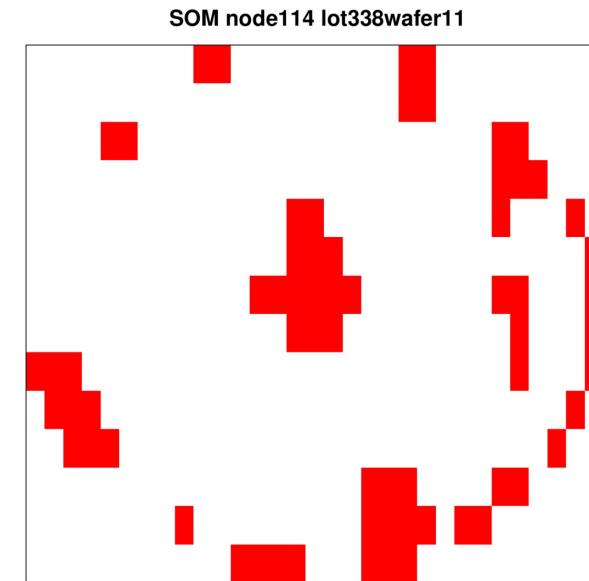
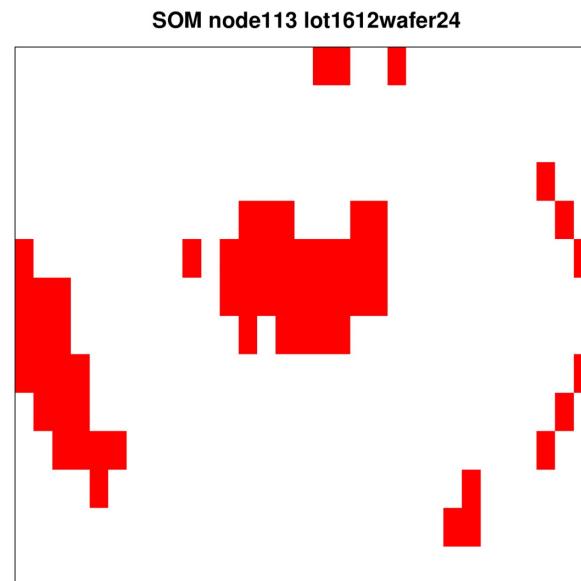
SOM分群8群 + Near Full + Random，總共10群

pattern	Center	Bottom	Local	Ring	Partial Ring
number	1033	651	4950	1282	23723
pattern	Moon	Scratch	Non pattern	Near Full	Random
number	339	47	324	316	932

# 5 資料分析 —— SOM分群

Pattern

Center

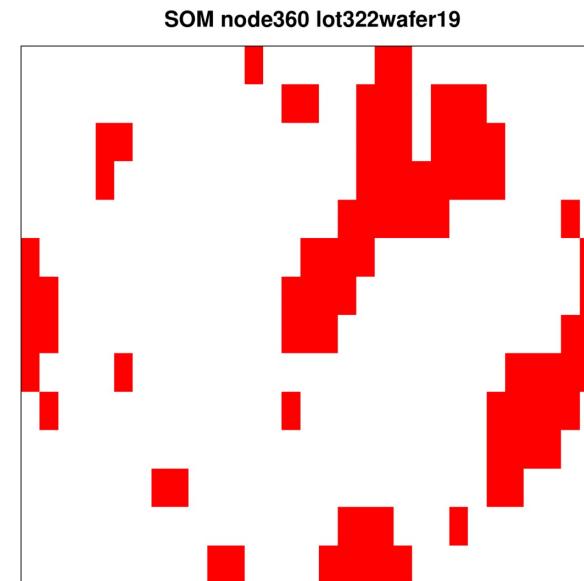
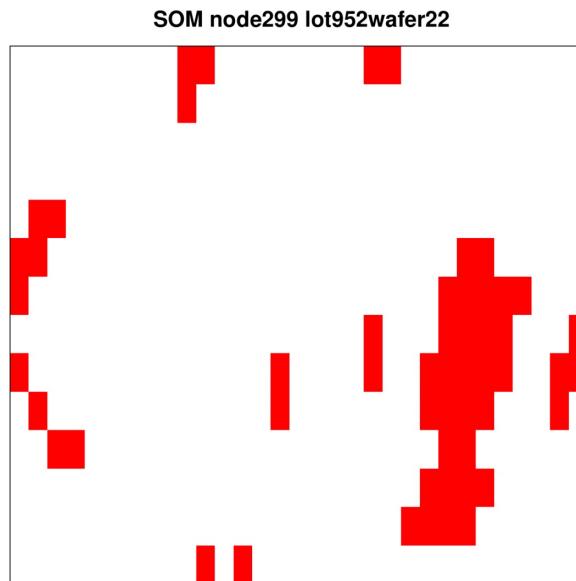


# 5 資料分析 —— SOM分群

Pattern

Local

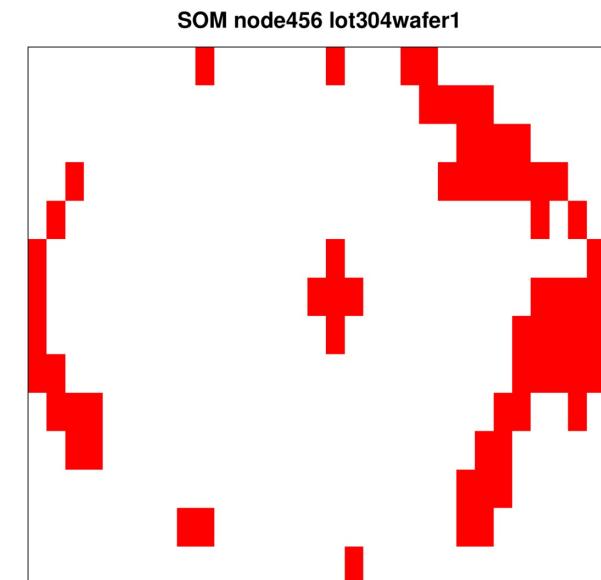
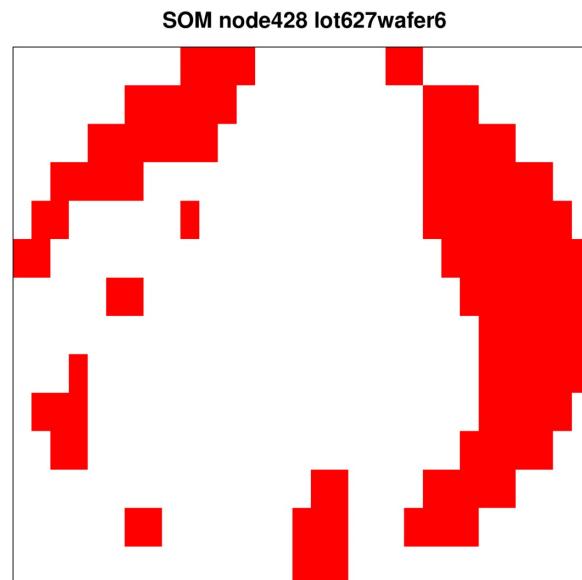
Scratch



# 5 資料分析 —— SOM分群

Pattern

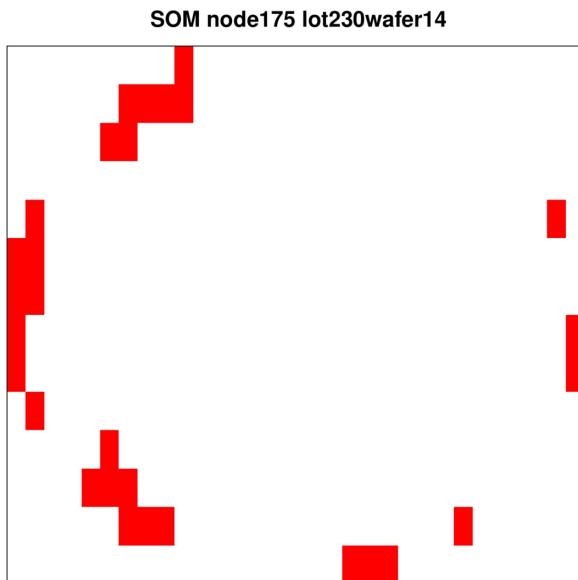
Moon



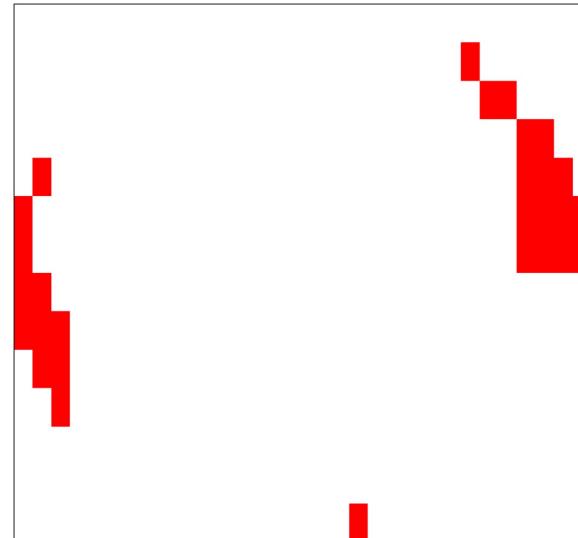
# 5 資料分析 —— SOM分群

Pattern

Partial Ring



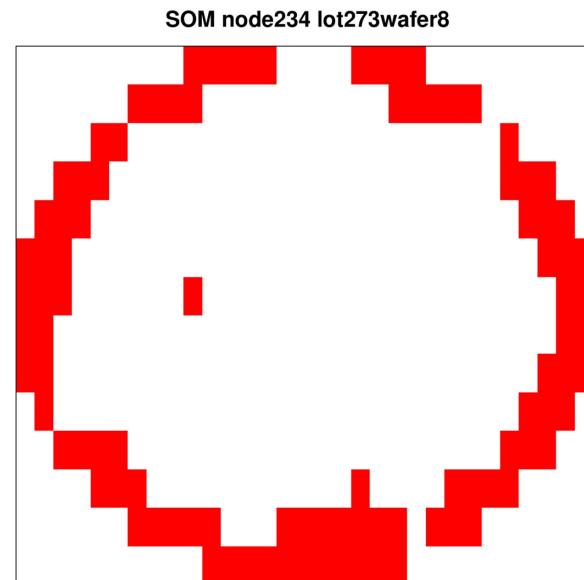
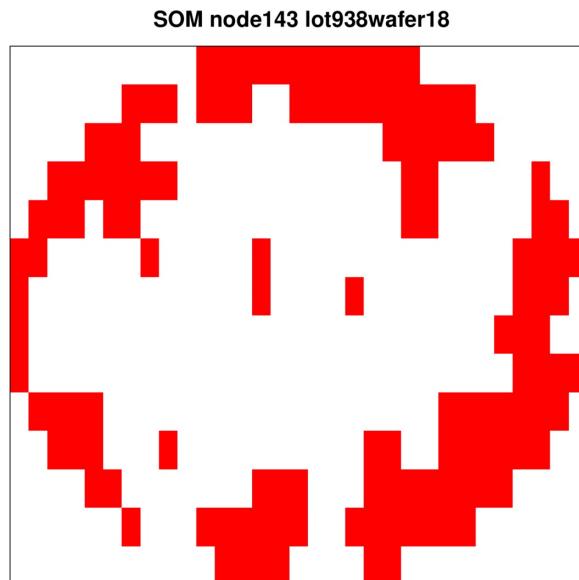
SOM node481 lot1517wafer3



# 5 資料分析 —— SOM分群

Pattern

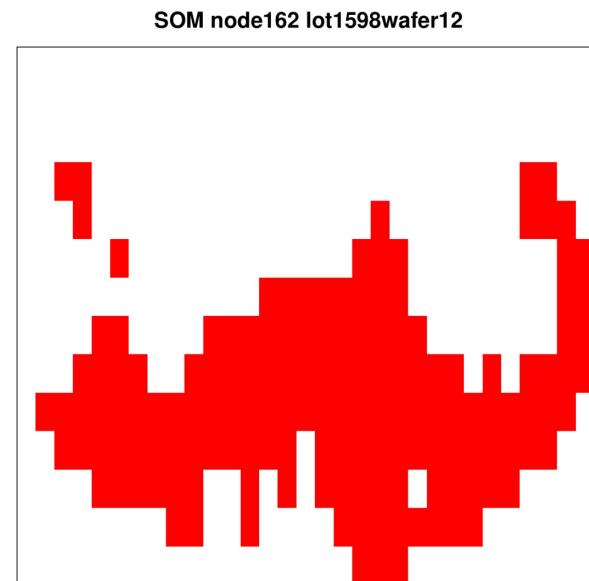
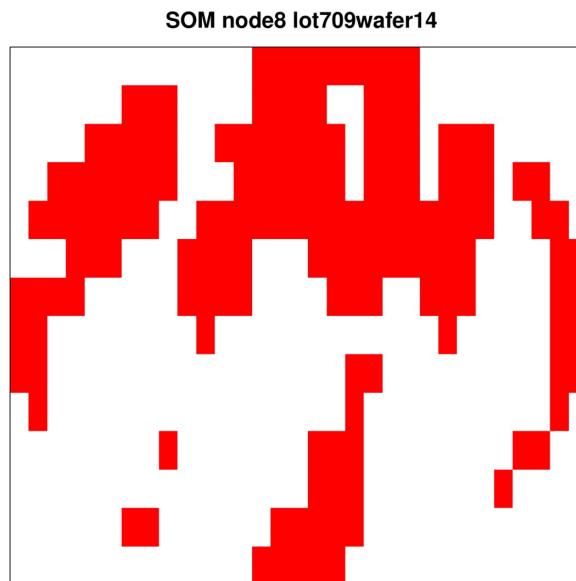
Ring



# 5 資料分析 —— SOM分群

Pattern

Bottom



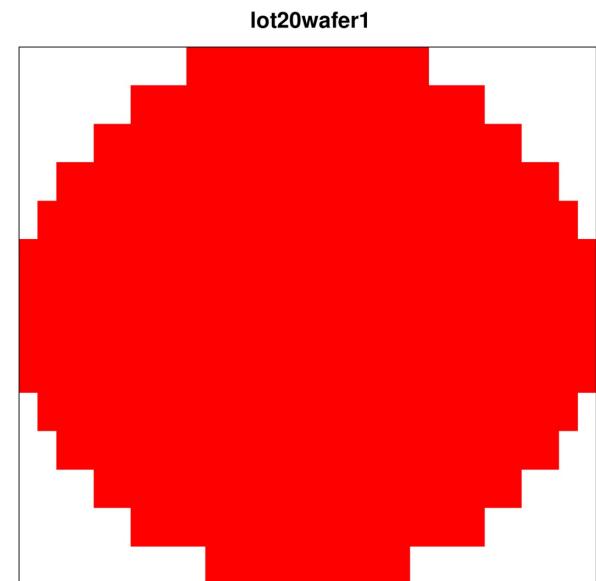
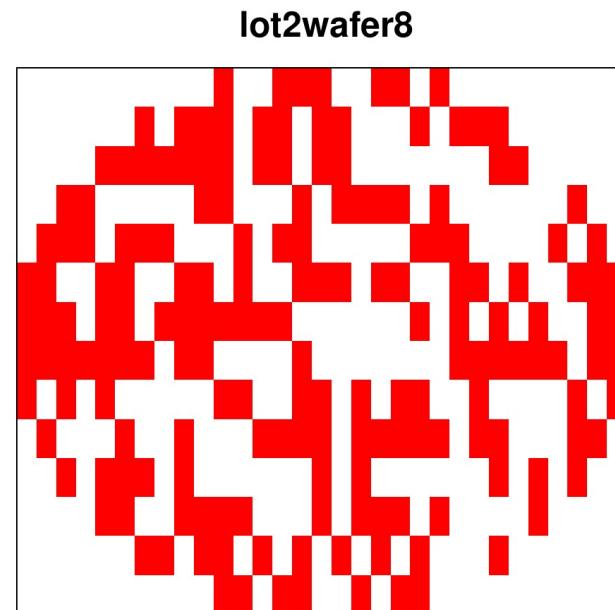
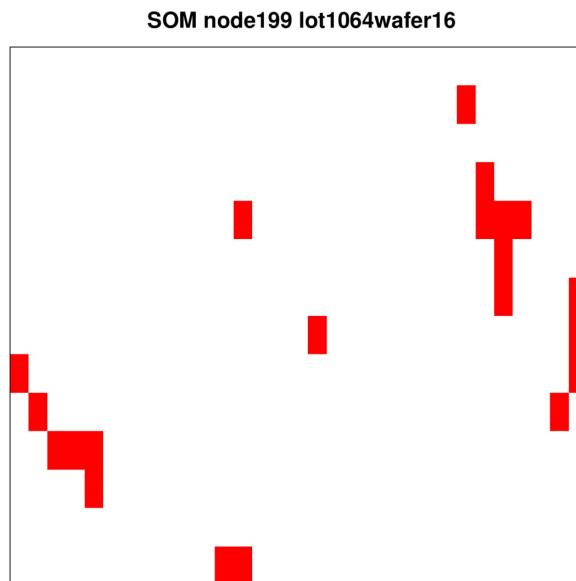
# 5 資料分析 —— SOM分群

Pattern

Nonpattern

Random

Near full



05

# 資料分析

— Random Forest分類

# 5 資料分析 —— Random Forest分類

## 目的

檢測目前建立的Label(良率異常類別分類)，是否有良好的效果。

## 使用資料集合

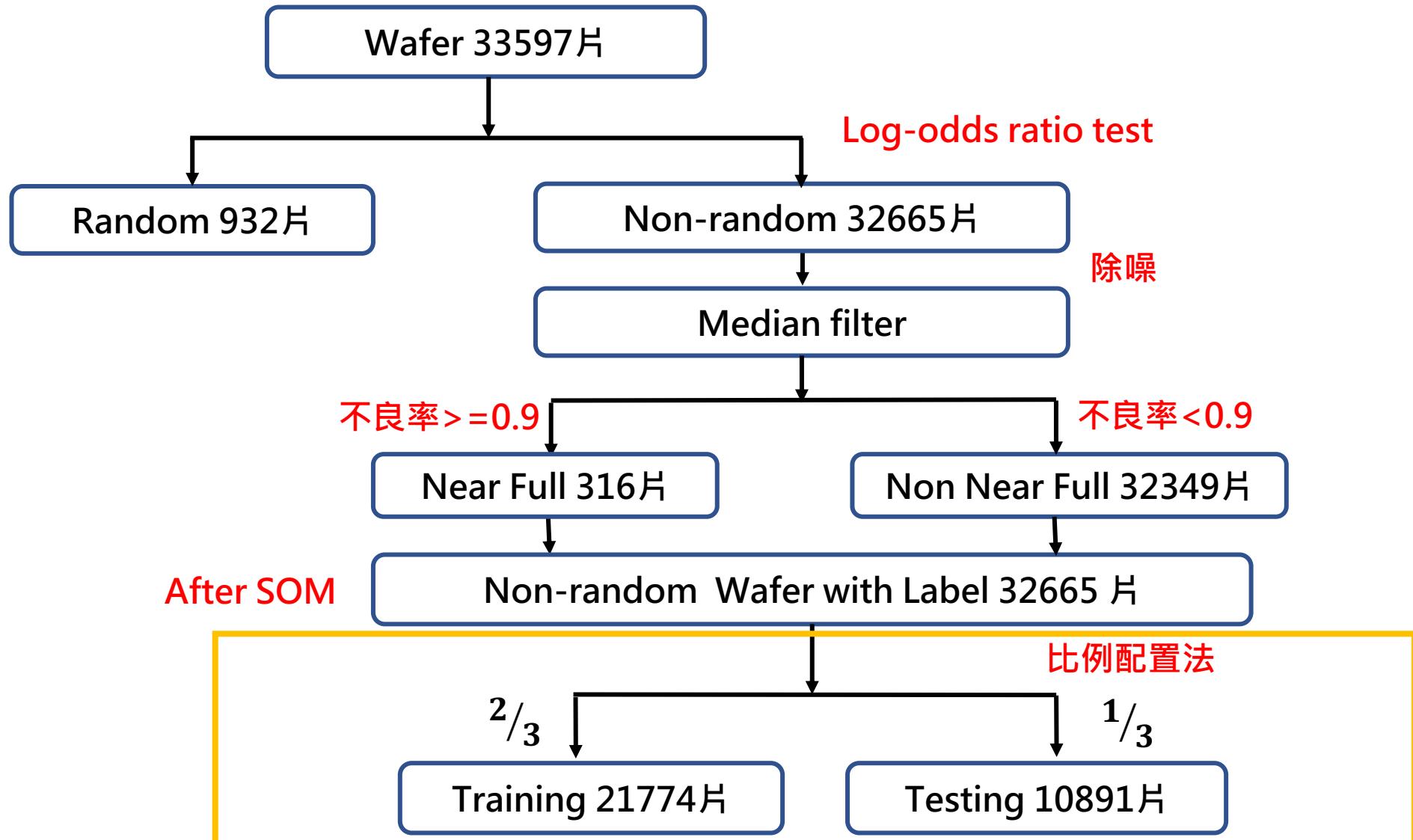
Lot1 ~ Lot 1723

## 建模流程

1. 決定 Training 抽樣方法：依照label採比例配置法
2. 嘗試不同模型參數組合：mtry & ntree
3. 移除Random後的分類結果

# 5 資料分析 —— Random Forest分類

## 流程圖



# 5 資料分析 —— Random Forest分類

決定抽樣數

比例配置法

只針對 “排除Random後” 剩下的晶圓片

依照pattern切割區域

每個區間抽取三分之二樣本數來建模

Pattern	Training	Testing	Total
Center	688	345	1033
Bottom	434	217	651
Partial Ring	15815	7908	23723
Ring	854	428	1282
Moon	226	113	339
Local	3300	1650	4950
Nonpattern	216	108	324
Scratch	31	16	47
Near full	210	106	316
total	21774	10891	32665

# 5 資料分析 —— Random Forest分類

Random Forest 分類

嘗試不同組合 mtry & ntree

mtry : Number of variables tried at each split

ntree : Number of tree

(1) 使用R內建default值 : ntree: 500 mtry : 18

(2) 使用以下排列組合

ntree: {100, 300, 500, 700}

→ 總共20種可能

mtry : {22, 43, 86, 172, 344}

找尋最佳參數

# 5 資料分析 —— Random Forest分類

分類結果

移除Random

(1) 使用R內建default值 : Number of trees: 500  
Number of variables tried at each split: 18

Confusion matrix:

	bottom	center	local	moon	near full	nonpattern	partial ring	ring	scratch	class.error
bottom	193	0	188	2	7	0	43	1	0	0.5553
center	0	143	83	0	0	0	462	0	0	0.792151
local	31	7	1345	0	7	0	1910	0	0	0.592424
moon	1	0	61	8	0	0	155	1	0	0.964602
near full	5	0	0	0	205	0	0	0	0	0.02381
nonpattern	0	1	0	0	0	17	198	0	0	0.921296
partial ring	3	33	205	0	0	0	15571	3	0	0.015428
ring	0	0	70	0	0	0	762	22	0	0.974239
scratch	0	0	23	0	0	0	8	0	0	1

花費時間：48分鐘

OOB estimate of error rate: 19.61%

y.test	bottom	center	local	moon	near full	nonpattern	partial ring	ring	scratch
bottom	103	0	17	1	3	0	1	2	0
center	0	77	4	0	0	0	14	0	1
local	94	50	668	31	1	0	101	44	9
moon	0	0	0	3	0	0	0	0	0
near full	5	0	5	0	102	0	0	0	0
nonpattern	0	0	0	0	0	6	0	0	0
partial ring	15	218	956	78	0	102	7790	372	6
ring	0	0	0	0	0	0	2	10	0
scratch	0	0	0	0	0	0	0	0	0

error rate: 19.58%

# 5 資料分析 —— Random Forest分類

分類結果

移除Random

(2)排列組合中最佳解：

Number of trees: 500

Number of variables tried at each split: 172

error rate		Ntree			
		100	300	500	700
mtry	22	19.72%	19.39%	19.38%	19.39%
	43	19.41%	19.32%	19.27%	19.19%
	86	19.37%	19.06%	19.09%	18.85%
	172	19.53%	19.09%	18.83%	18.86%
	344	19.34%	19.05%	18.91%	18.84%

# 05 資料分析

— 第二次SOM分群

# 5 資料分析 —— 第二次SOM分群

## 目的

### 將良率異常類別有效率分群

透過兩次 SOM 有效地節省人力成本、時間以及減少肉眼的判斷失誤，將資料的樣型分群，並標上label。

### 將誤判晶圓再次進行配適

我們發現若只有一次SOM分群，以多數決來決定node樣型，對於其他非多數之晶圓片來說，判定不佳。

因此，想嘗試將隨機森林中被同樣誤判之晶圓片進行第二次配適。

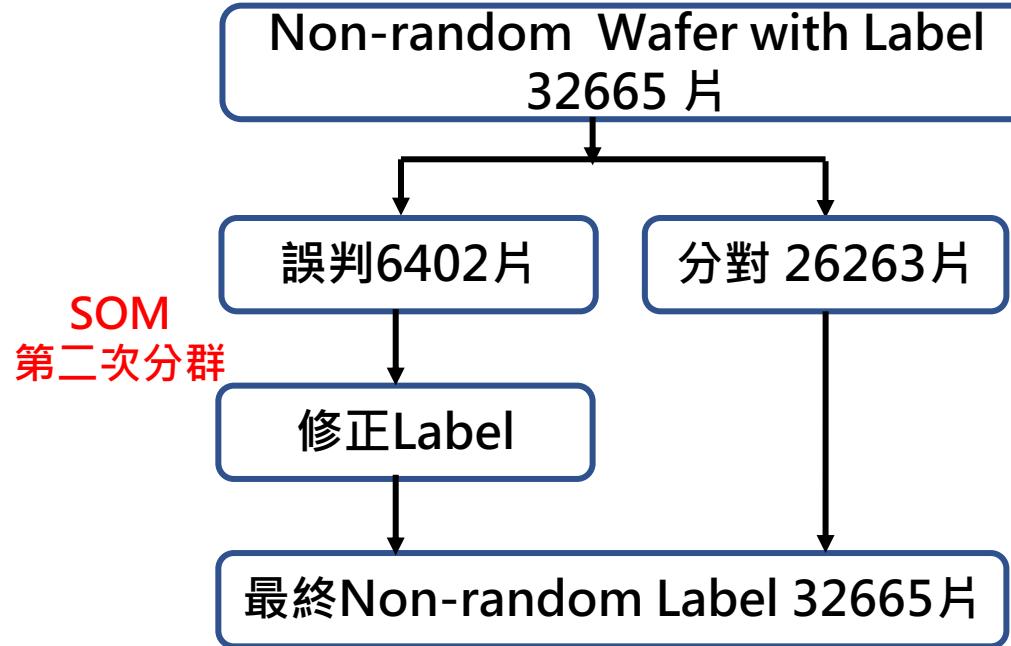
## 使用資料集合

Lot1 ~ Lot 1723 中 Random Forest 誤判之晶圓片(共6402片)

※ Random Forest 以 default 值做為參數( $n\text{tree} = 500, m\text{try} = 18$ )

# 5 資料分析 —— 第二次SOM分群

## 分群流程圖



SOM  
第二次分群

## SOM 設定參數

$n$  : 6402 片

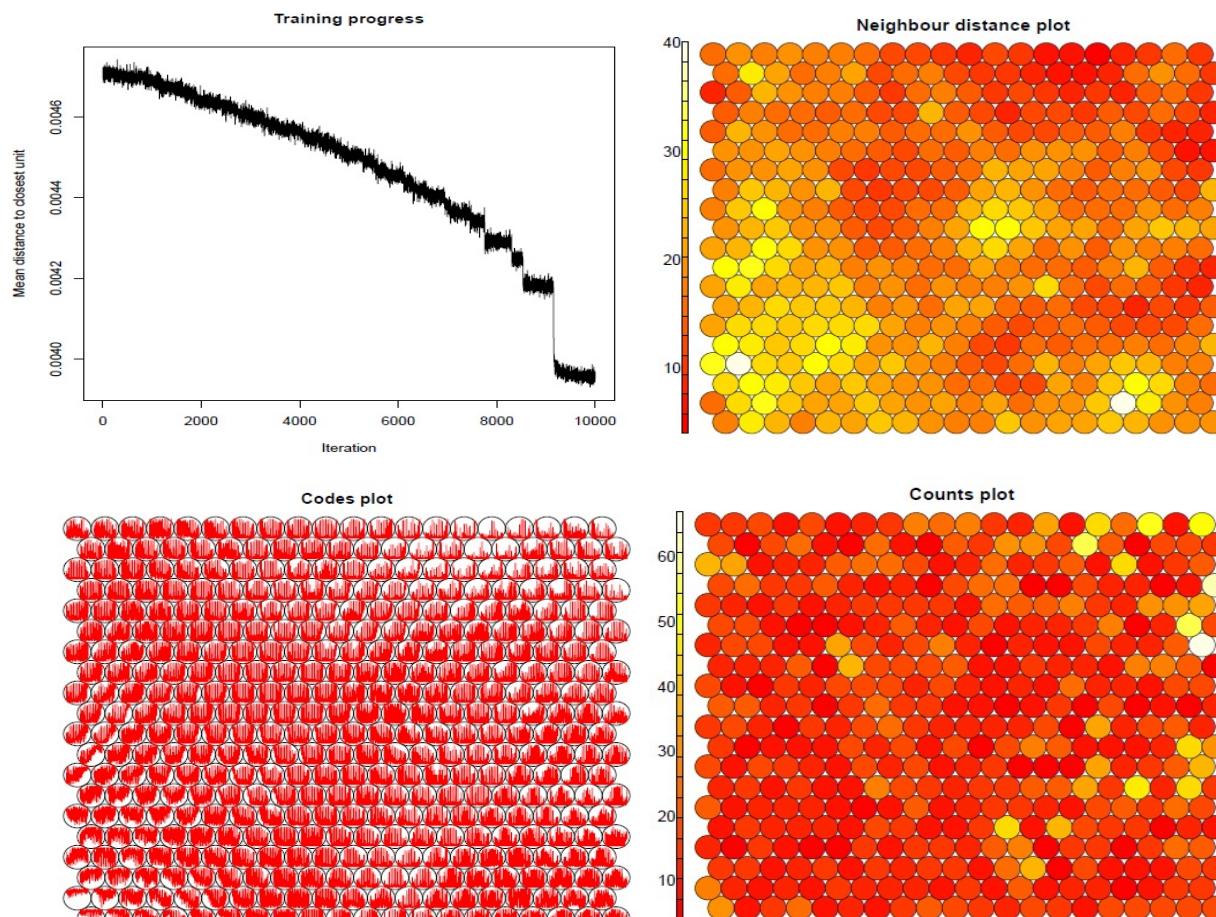
最大迭代次數 : 10000

輸出層大小 :  $20 * 20$

學習率 : 0.05~0.01

網路拓樸結構 : hexagonal

## 分群結果



# 5 資料分析 —— 第二次SOM分群

## 分群結果

各pattern個數

SOM分群8群 + Near Full + Random，總共10群

pattern	Center	Bottom	Local	Ring	Partial Ring
Number in SOM I	1033	651	4950	1282	23723
Number in SOM II	934	409	5161	486	25066
pattern	Moon	Scratch	Non pattern	Near Full	Random
Number in SOM I	339	47	324	316	932
Number in SOM II	93	34	166	316	932

# 05 資料分析

— 第二次Random Forest分類

# 5 資料分析 —— 第二次Random Forest分類

## 目的

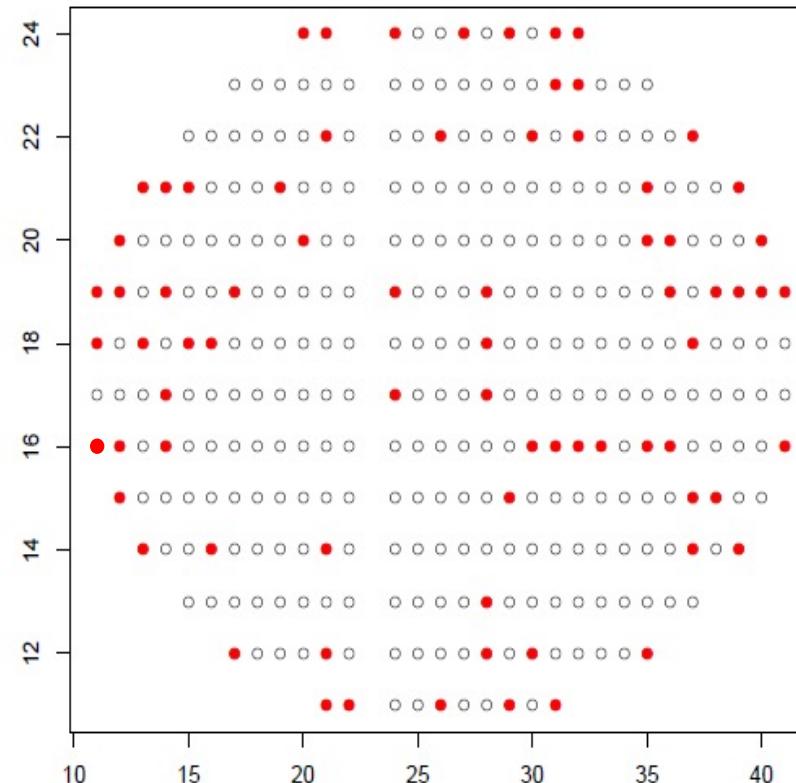
檢測修正後的Label(良率異常類別分類)，是否有很好的效果

## 使用資料集合

Lot 1 ~ 1723

Lot 1725 ~ 3447

原先放到一邊的晶圓片  
( lot 14 wafer 23,  
lot 1738 wafer 23)

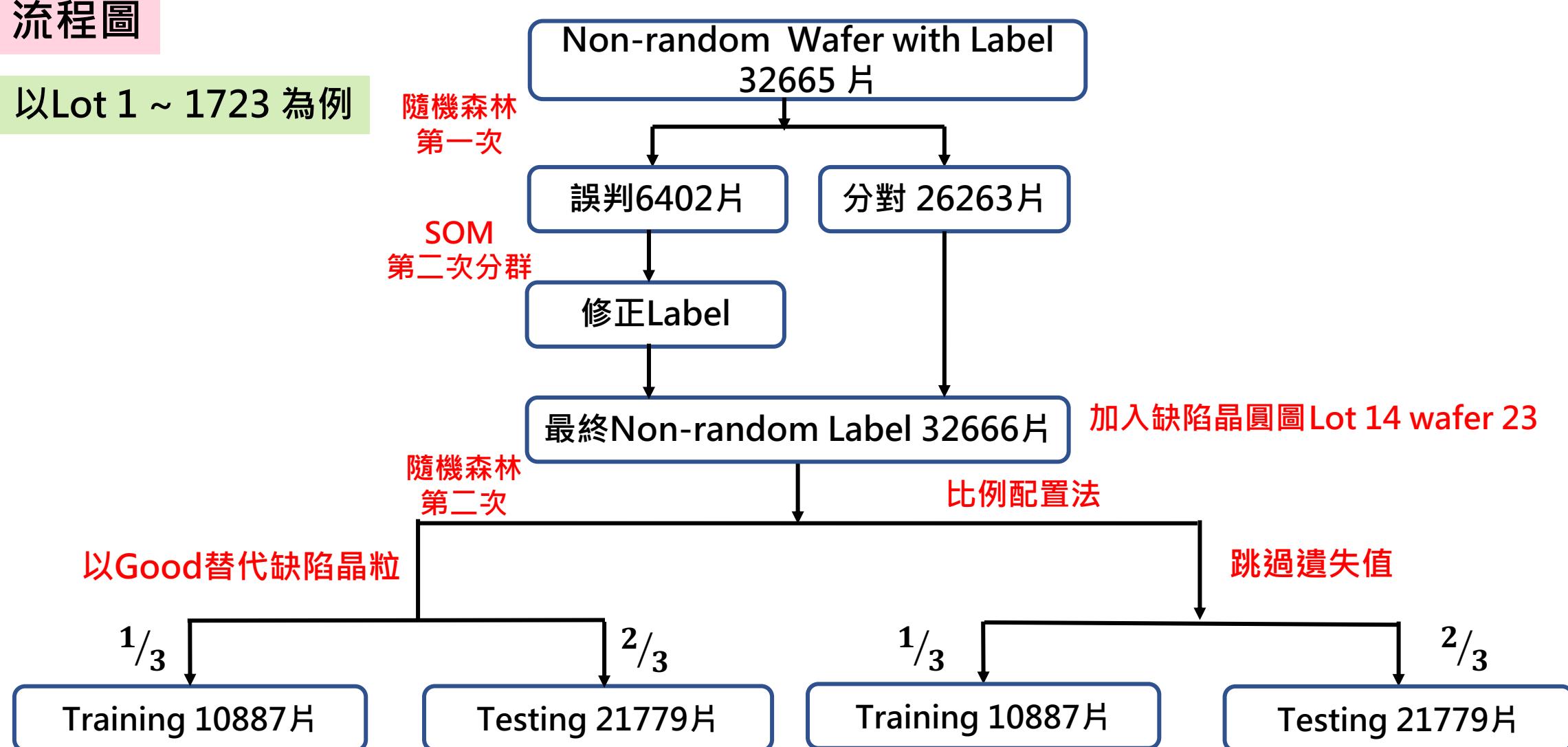


模型	缺陷晶圓處理方式	使用資料集合
A	以Good替代	Lot1~1723
B	以Good替代	Lot1~3447 (without Lot1724)
C	跳過遺失值	Lot1~1723
D	跳過遺失值	Lot1~3447 (without Lot1724)

# 5 資料分析 —— 第二次Random Forest分類

## 流程圖

以Lot 1 ~ 1723 為例



# 5 資料分析 —— 第二次Random Forest分類

分類結果

移除Random

Default : Number of trees: 500  
Number of variables tried at each split: 18

模型	缺陷晶圓處理方式	變數個數	使用資料集合	Training 片數	Testing 片數	參數	OOB(%)	誤差率(%)	時間 (分)
A	以Good 替代	344	Lot1~1723	10887	21779	Default	13.81	13.64	11.884
B	以Good 替代	344	Lot1~3447 (without Lot1724)	21779	43553	Default	8.92	8.93	45.889 (Putty)
C	跳過遺失值	329	Lot1~1723	10887	21779	Default	13.84	13.70	11.042
D	跳過遺失值	329	Lot1~3447 (without Lot1724)	21779	43553	Default	8.97	8.95	46.625 (Putty)

# 5 資料分析 ——第二次Random Forest分類

## 分類結果

各pattern個數

使用B模型預測結果，作為最終label (9種) + Random，總共10種，67196片

pattern	Center	Bottom	Local	Ring	Partial Ring
number	958	575	7918	326	54708
pattern	Moon	Scratch	Non pattern	Near Full	Random
number	62	23	133	629	1864

# 06 總結

# 6 總結

判斷隨機非隨機

Log Odds Ratio Test

判定標準

選取標準常態累積機率值介於0.4~0.6之間的晶圓圖  
為隨機性故障之圖形

判定結果

隨機性圖形片數：1864片

非隨機性圖形片數：65330 片 +2片 = 65332片

不足一片的缺陷晶圓片，主觀判斷

# 6 總結

晶圓缺陷圖分群及建模

SOM + Random Forest

## 想法

1. 除去Random後，我們使用SOM分群，並以多數決來決定node樣型，但可能因此忽略少數族群，使準確率降低。
2. 將初步隨機森林分類後誤判的晶圓片進行第二次SOM，期望能增加最終分類之準確率。

## 結論

1. 透過SOM擷取出共同的特徵，將全部的晶圓片分成若干個node，大幅減少時間成本。
2. 第二次Random Forest 的錯誤率下降不少，我們認為反覆的將誤判的晶圓片重新分群，對我們的建模有很大的幫助。
3. 最終模型預測準確率為91%，未來可直接將未知pattern的晶圓片丟入模型中預測label。

# 6 總結

## 製程問題

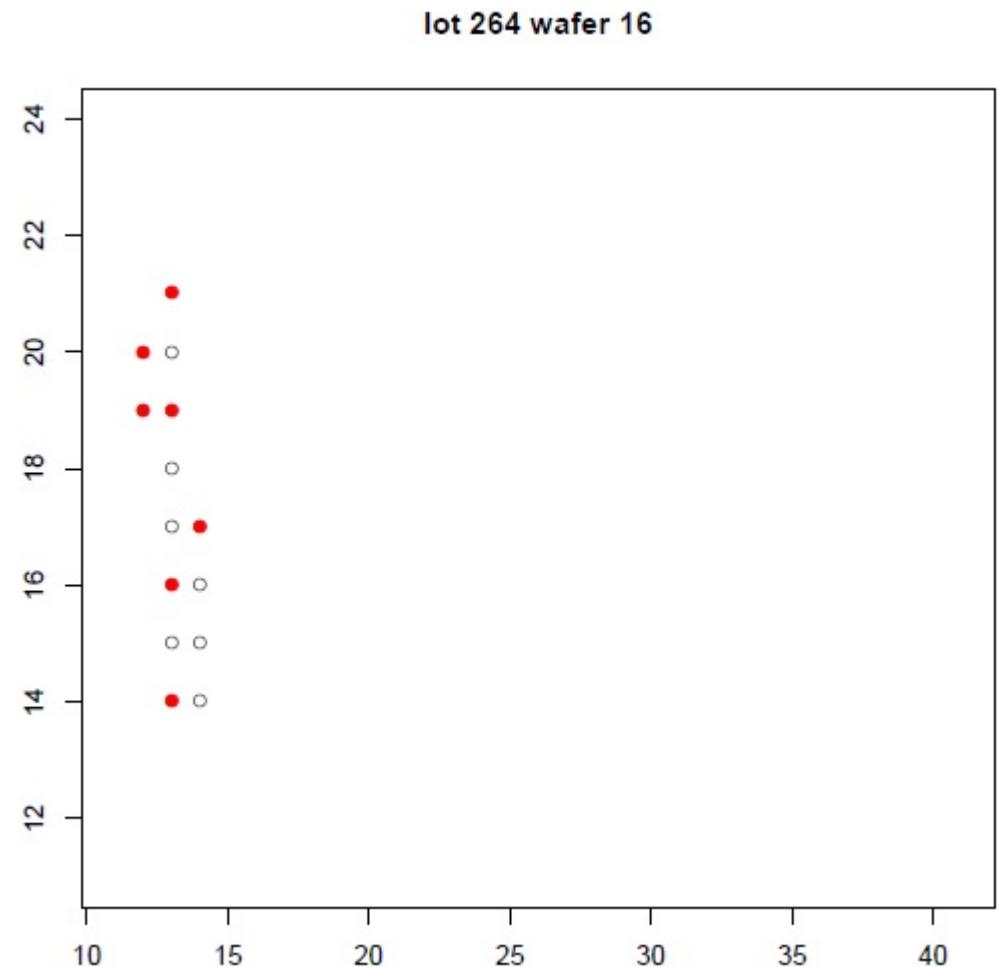
錯誤樣型	發生原因	RF predict
Center	蝕刻階段的電漿區出現問題。	958
Bottom	擴散階段加熱過程不均勻。	575
Local	探針移動至晶圓邊緣時，由於定位不準或接觸不良引起。	7918
Moon	薄膜階段的快速退火過程出現異常。	62
Scratch	晶圓表面在製造過程導致的摩擦或刮傷，或是本身材質上的缺陷。	23
Partial ring	人為疏失和晶圓本身不平坦佔主要因素。	54708
Ring		326
Non pattern	不屬於任何特定錯誤樣式，是對機台分析沒幫助的錯誤。	133
Near full	九成以上佈滿壞晶粒。	629
Random	隨機分布的壞晶粒，無任何顯著樣型存在。	1864

# 6 總結

檢視分析時被放一邊的資料

Lot 264 wafer 16 / Lot 1988 wafer 16

多出來的晶粒過於稀少，且本來的晶圓片也已完整，  
我們認為不足以判斷pattern類型。



# 6 總結

## 未來展望與檢討

1. 可以嘗試只用重要變數建立Random Forest模型，也許效果比較好。
2. 這次未針對Bin code 與測試時間做分析，未來有時間可以再嘗試。
3. 文獻中提到的特徵萃取方法，因能力與時間有限未被採用，十分可惜。
4. 文獻中曾經有提過利用baseline來處理複合型缺陷圖晶圓片，這也是一個值得被探討的問題。

# 6 總結

Time summary table

		變數個數	使用資料集合	Training 片數	時間 (分)	程式系統
SOM	第一次SOM	344	Lot1~1723	10782	1194.37	Rstudio
	第二次SOM	344	誤判晶圓片	6402	214.132	Rstudio
Random Forest	模型	缺陷晶圓處理方式	變數個數	使用資料集合	Training 片數	時間 (分)
	第一次	無放入	344	Lot1~1723	21774	48
	A	以Good替代	344	Lot1~1723	10887	11.884
	B	以Good替代	344	Lot1~3447 (without Lot1724)	21779	45.889
	C	跳過遺失值	329	Lot1~1723	10887	11.042
	D	跳過遺失值	329	Lot1~3447 (without Lot1724)	21779	46.625

# 07 參考資料

# 7 參考資料

1. 張喬凱，利用類神經方法建構晶圓缺陷點群聚圖案之辨識系統，國立交通大學，2006
2. 黃美瑄，WMG：經由多方塊分解之晶圓圖產生器，中華大學，2000
3. 莊銘弘，考量半導體製程能力限制下之晶圓圖隨機性辨識法及應用，國立交通大學，2010
4. Ming-Ju Wu, Wafer Maps Failure Pattern Recognition and Similarity Ranking for Large-Scale Data Sets,  
IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING, VOL. 28, NO. 1, 2015
5. 簡禎富，資料挖礦與大數據分析，前程文化事業股份有限公司，2017