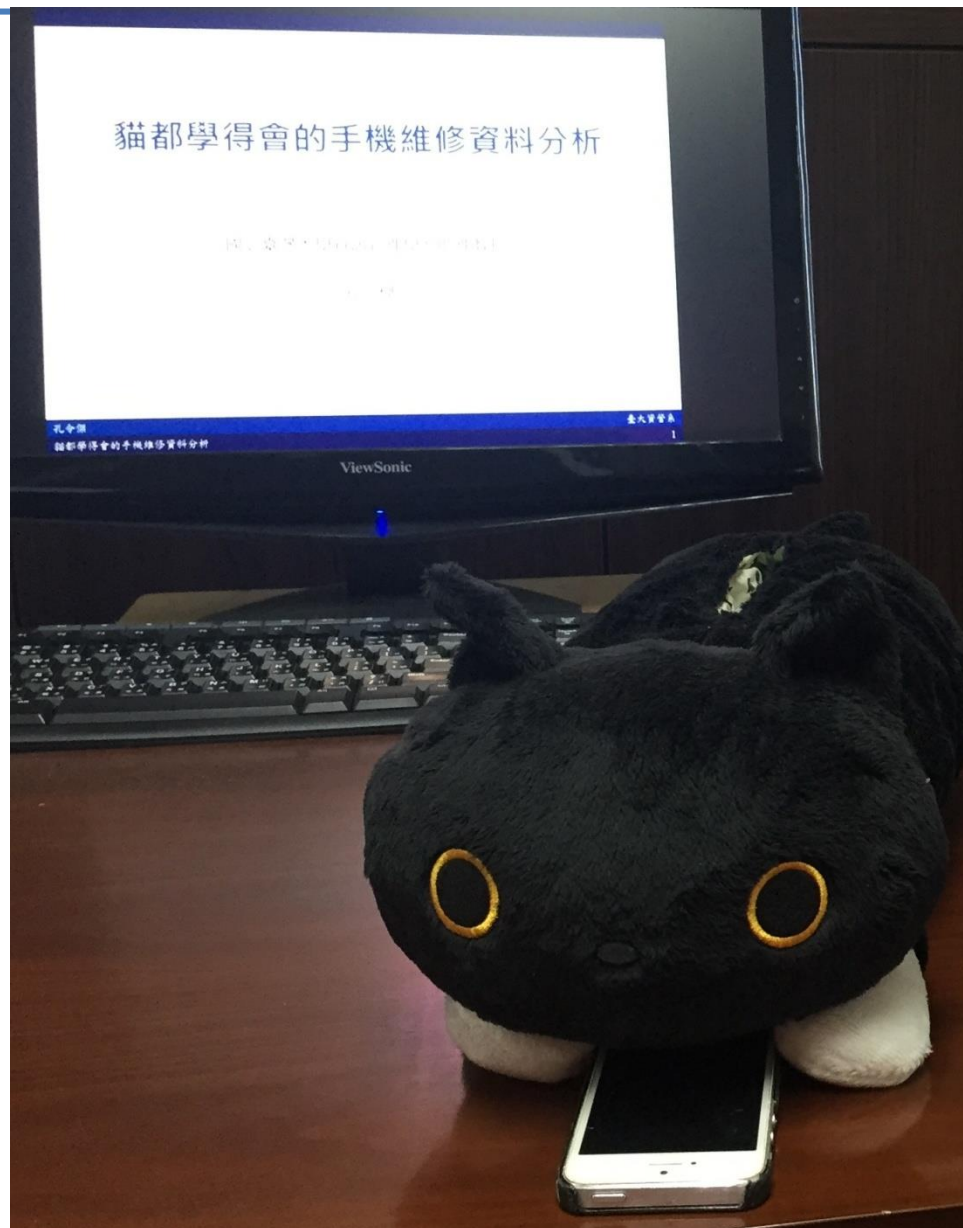


貓都學得會的 手機維修資料分析

國立臺灣大學資訊管理學系
孔令傑



手機維修

- 臺灣最大的**手機**零售商
 - 超過 600 家實體零售店
- 接受各種手機**維修**需求
 - 不論是否由其售出
 - 不論是否在保固內
- 以其**快速**且**高品質**的維修服務自豪

料件耗用預測

- 要「快」就要把料件備好！
 - 維修用的料件要預先跟原廠採購並儲存好
- 如何減低存貨成本？
 - 需要好的訂補貨政策
 - 需要好的需求預測
- 本專案：做手機料件的維修耗用預測，愈準愈好！

資料

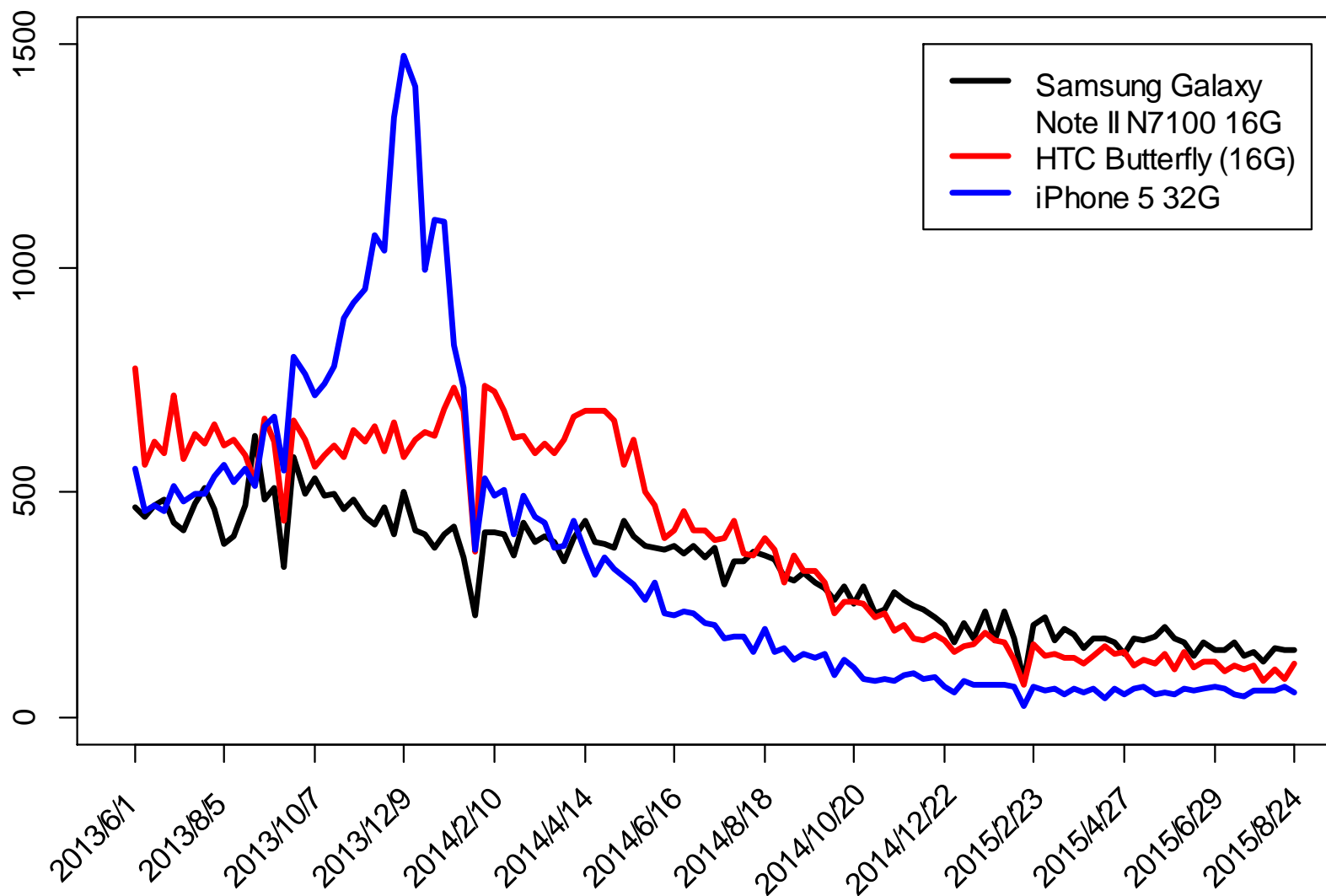
- 資料起迄：2013/6/1 到 2015/8/31
- 產品主檔（大約 2,500 筆）
 - 產品編號、產品品牌、產品名稱
- 料件主檔（大約 7,000 筆）
 - 料件編號、料件名稱
- 手機送修交易檔（大約 2,000,000 筆）
 - 維修單編號、送修日期、產品編號
 - 購買日期、保固狀態
- 料件耗用交易檔（大約 500,000 筆）
 - 維修單編號、料件編號、耗用數量

手機送修交易檔

- 每支手機都有歷史維修記錄
 - 逐週累積起來，就是個標準的**時間序列**資料

Week ID	Begin Date	End Date	Samsung Galaxy Note II 16G	HTC Butterfly	iPhone 5 32GB
1	2013/6/1	2013/6/9	529	776	555
2	2013/6/10	2013/6/16	486	563	458
3	2013/6/17	2013/6/23	518	614	473
4	2013/6/24	2013/6/30	514	588	461
5	2013/7/1	2013/7/7	490	719	516
6	2013/7/8	2013/7/14	456	577	480

Historical product repair consumption



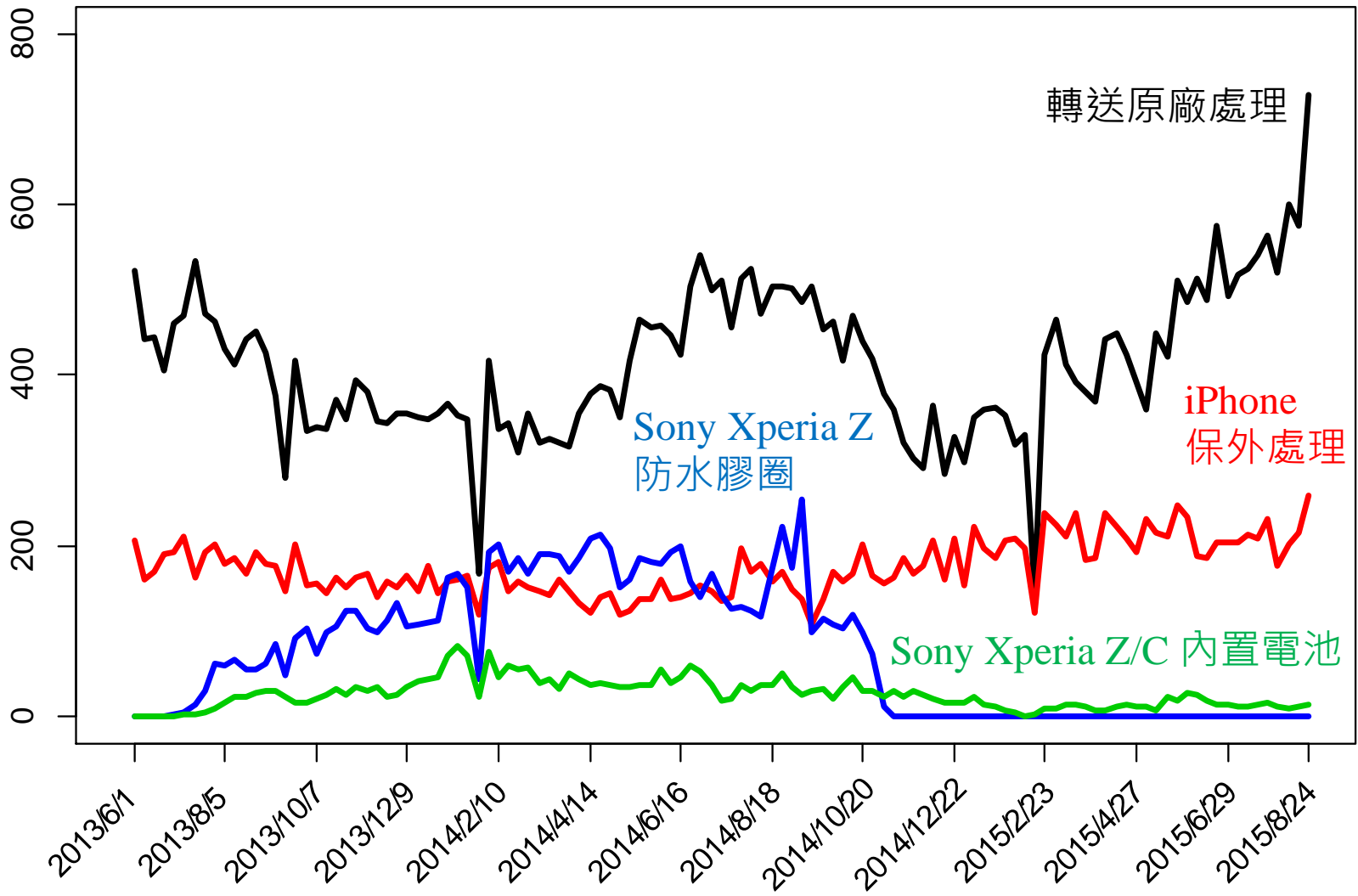
手機送修交易檔

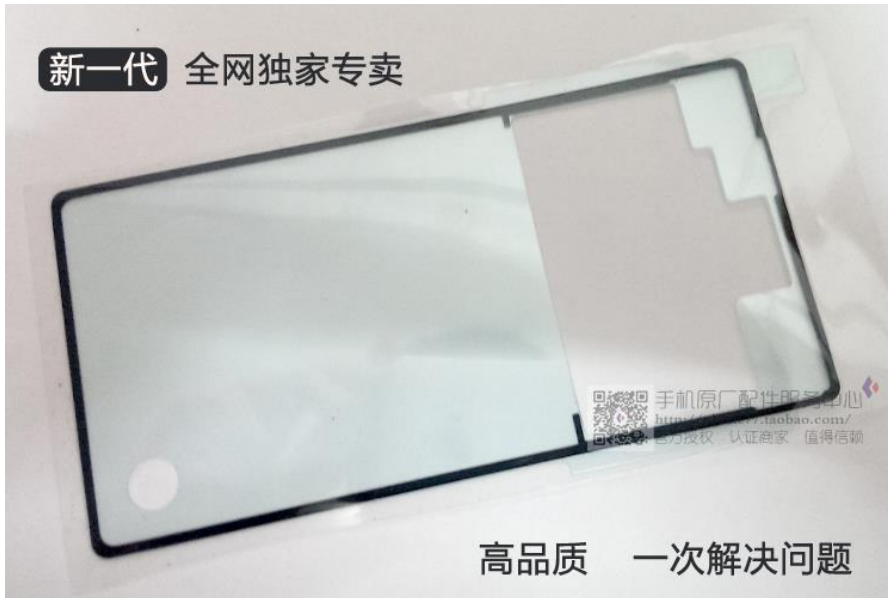
- 看起來挺酷的！
- 不過我們要準備的是「料件」的存貨
- 我們要做的是料件耗用的預測

料件耗用交易檔

- 每個料件都有歷史耗用記錄
 - 逐週累積起來，就是個標準的**時間序列**資料
- 在那兩年間，什麼料件被耗用的最多？

Historical item repair consumption





<https://world.taobao.com/item/38994644175.htm?spm=a1z3p.7664257.0.0.HrMKCk>



<https://tw.bid.yahoo.com/item/Sony%E7%B4%A2%E5%B0%BCc6602-c6603-S39h-C2305%E6%89%8B%E6%A9%9F%E9%9B%BB%E6%B1%A0-L36H%E5%85%A7%E7%BD%AE%E9%9B%BB%E6%B1%A0-100112339956>

現行方法 (Current Method)

- 要改進預測準確度，當然要先知道現在用的是什麼方法
- 針對一個料件，考慮其過去六週的耗用量，以如下方法預測此週的耗用量
- 令 x_i 為第 i 週的耗用量， $i = 1, \dots, 6$.

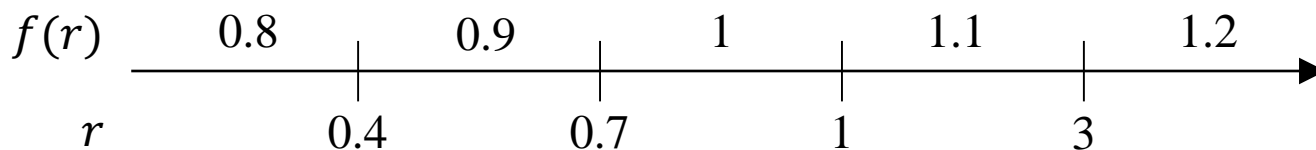
i	1	2	3	4	5	6	7
x_i	94	93	98	118	113	140	?

- 目標：預測 x_7

現行方法 (Current Method)

- 步驟一：找出「趨勢」

- 令 $A = x_1 + x_2 + x_3$ 、 $B = x_4 + x_5 + x_6$ ，以及 $r = \frac{B}{A}$ 。
- 根據 r 查表求出 $f(r)$ ：



- 步驟二：乘上「趨勢」

- 令 $C = \frac{x_3 + x_4 + x_5 + x_6}{4}$ 。
- 預測此週的耗用量 x_7 為 $C \times f(r)$

現行方法 (Current Method)

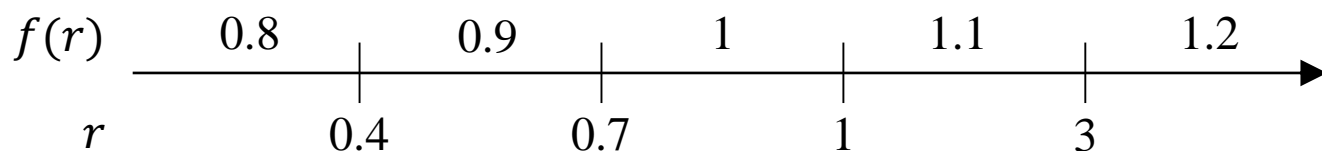
- 舉例來說：

i	1	2	3	4	5	6	7
x_i	94	93	98	118	113	140	?

$$- r = \frac{118+113+140}{94+93+98} = 1.3.$$

$$- f(r) = 1.1.$$

$$- x_7 = 1.1 \times \left(\frac{98+118+113+140}{4} \right) = 1.1 \times 117.25 \approx 129.$$



關於現行方法

- 公司：「有什麼方法可以讓我們改進預測準確度嗎？」
- 我：「請問現在有多不準呢？」
- 公司：「嗯……我們也不知道……」

大綱

- 前言
- 料件時間序列法
- 產品時間序列法
- 單品預測彙總法
- 比較與結論

料件時間序列法

- 從料件耗用的歷史資料出發
- 考慮三種直觀的時間序列方法
 - HW 指數平滑法 (Holt-Winters Exponential smoothing)
 - 迴歸 (Regression)
 - 現行改良法 (Modified Current Method)
- 對每一個料件，以歷史資料「學習」最佳方法：
 - 對每個方法，以 MAE (mean absolute error) 找尋其最佳參數
 - 比較 (採用最佳參數的) 三種方法，以 MAE 找尋其最佳方法
 - 以該方法及其最佳參數做預測

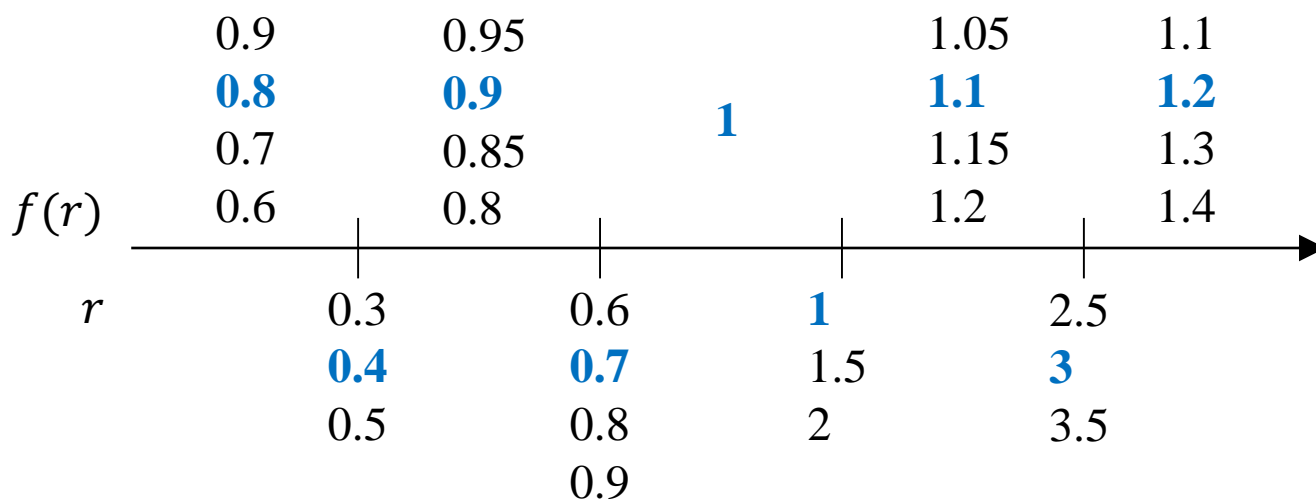
HW 指數平滑法的參數組合

- 參數：
 - $\alpha \in \{0.1, 0.2, \dots, 0.9\}$
 - $\beta \in \{0.1, 0.2, \dots, 0.9\}$
- 預測公式：
 - $s_t = \alpha x_t + (1 - \alpha)(s_{t-1} + b_{t-1})$
 - $b_t = \beta(s_t - s_{t-1}) + (1 - \beta)b_{t-1}$
 - $x_{t+m} = x_t + mb_t$

迴歸的參數組合

- 參數：
 - 以過去 n 期歷史資料做預測， $n \in \{2, 3, \dots, 10\}$
 - 以過去 m 期歷史資料做自變數， $m \in \{0, 1, \dots, n - 2\}$
- 預測公式：
 - $x_t = \alpha + \beta_0 t + \beta_1 x_{t-1} + \dots + \beta_m x_{t-m}$

現行改良法的參數組合



誤差衡量

- 三種常見的誤差衡量公式：

- $$MSE = \frac{\sum_{t=1}^n (f_t - x_t)^2}{n}$$

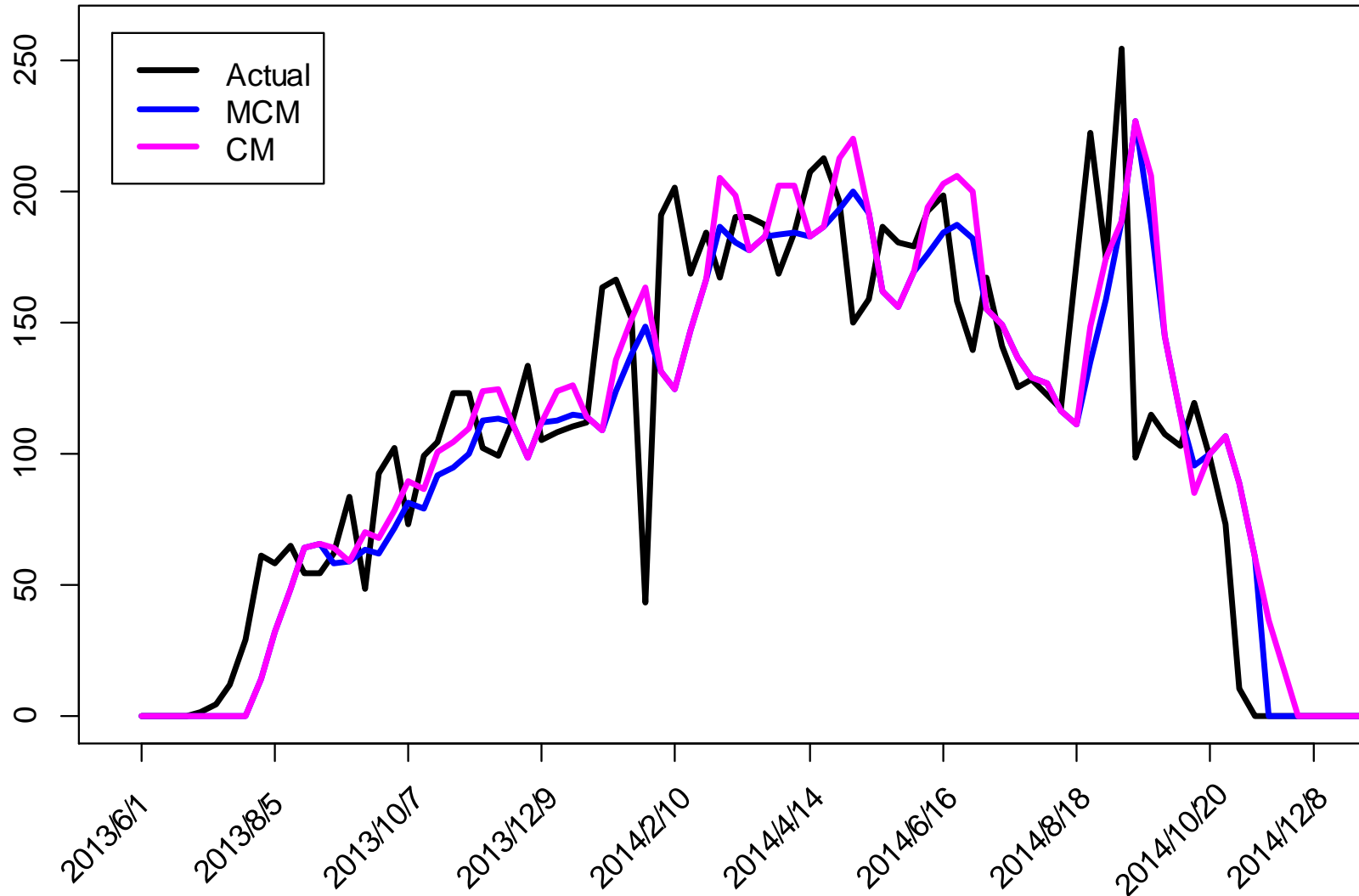
- $$MAE = \frac{\sum_{t=1}^n |f_t - x_t|}{n}$$

- $$MAPE = \frac{\sum_{t=1}^n |f_t - x_t| / x_t}{n}$$

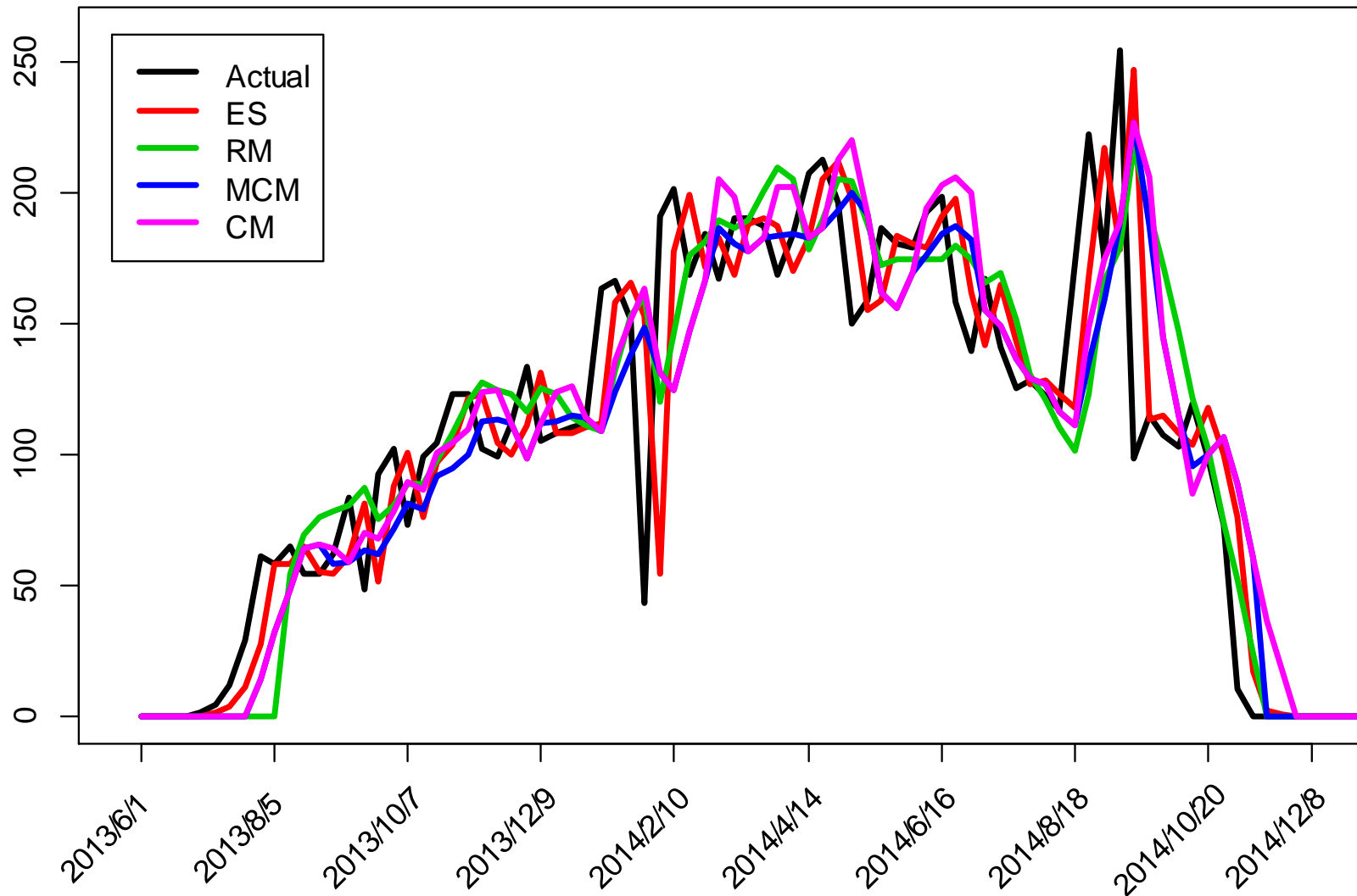
- 公式選擇：

- 單一料件內的方法比較用 MAE
 - 料件間比較用 MAPE

Learning outcome for Sony Xperia Z 防水膠圈



Learning outcome for Sony Xperia Z 防水膠圈



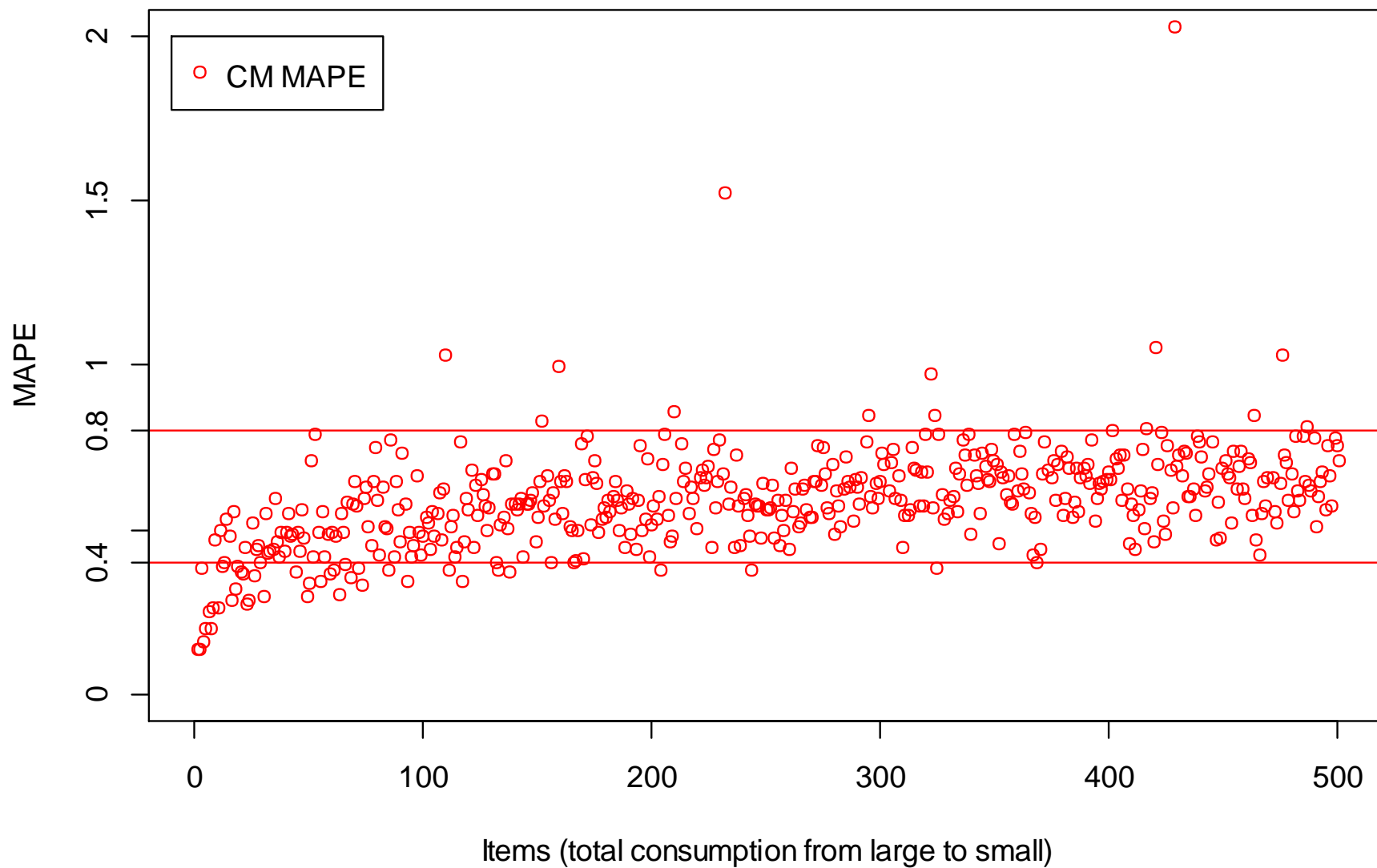
學習結果

- 最佳方法分佈：

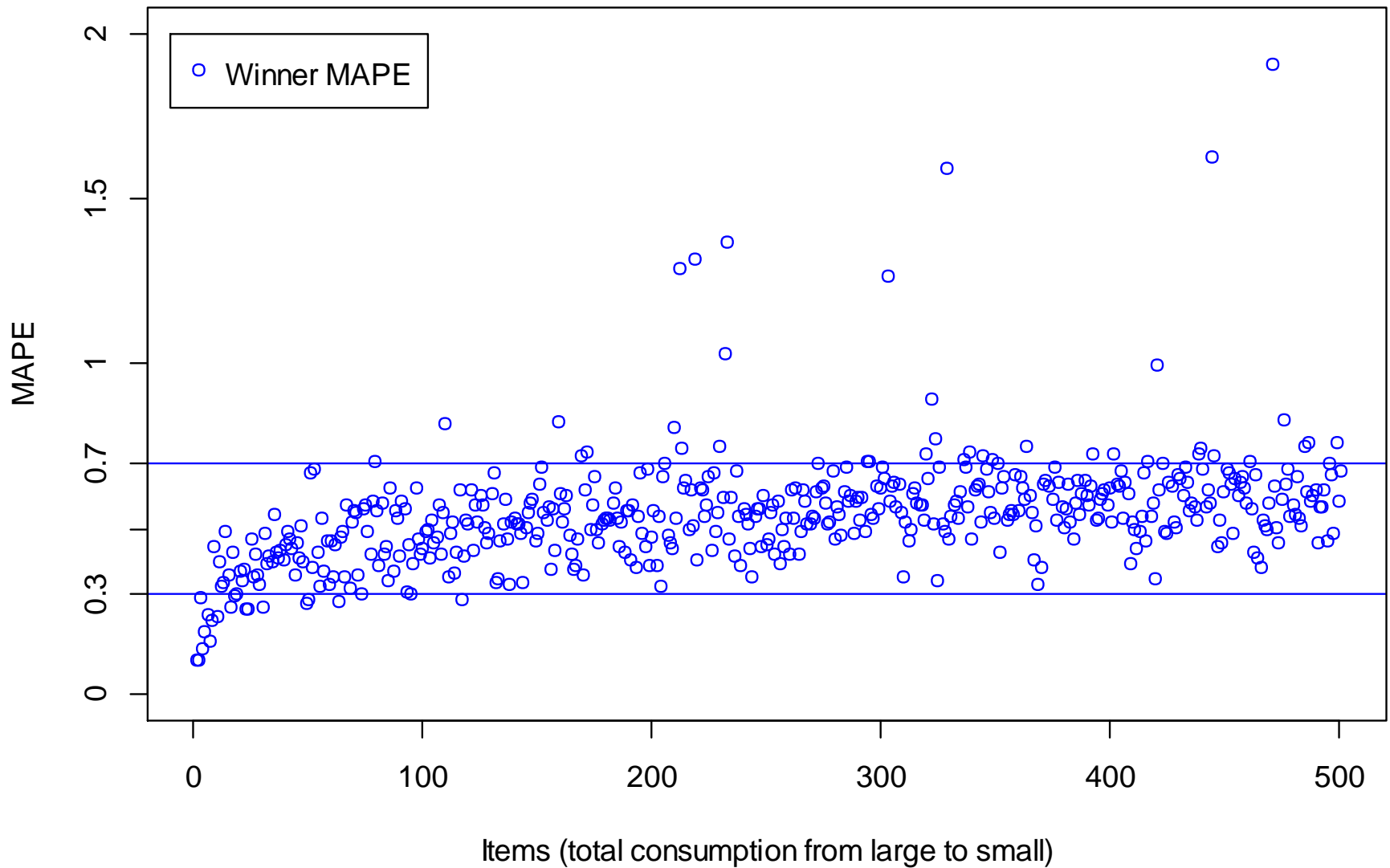
最佳方法	歷史總料件耗用量		
	前 10	前 50	前 500
ES	5	35	334
RM	2	5	42
MCM	3	9	113
CM	0	1	11

- 簡單的方法有時很有效！

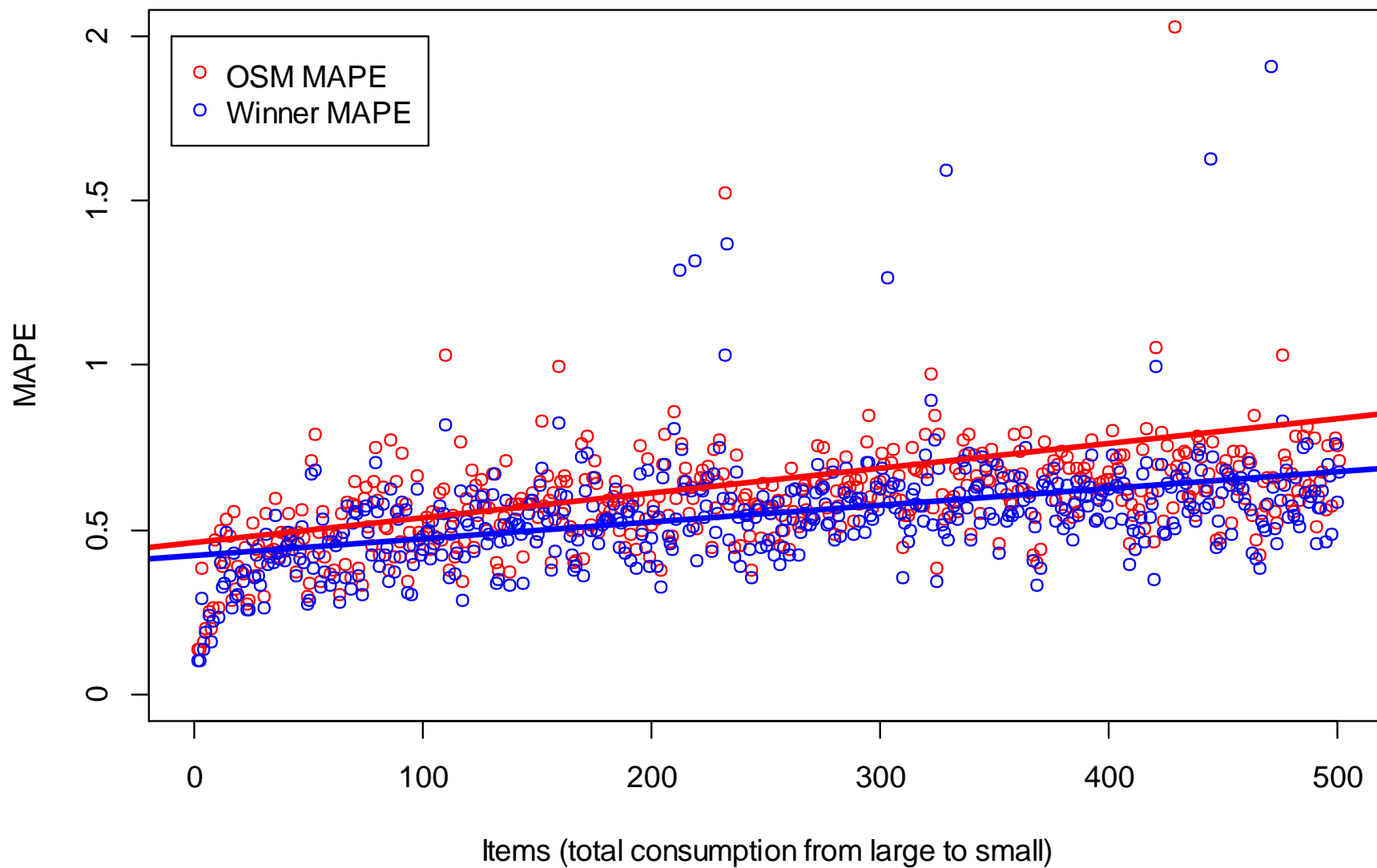
CM MAPE



Winner MAPE



MAPE Improvement



下一步？

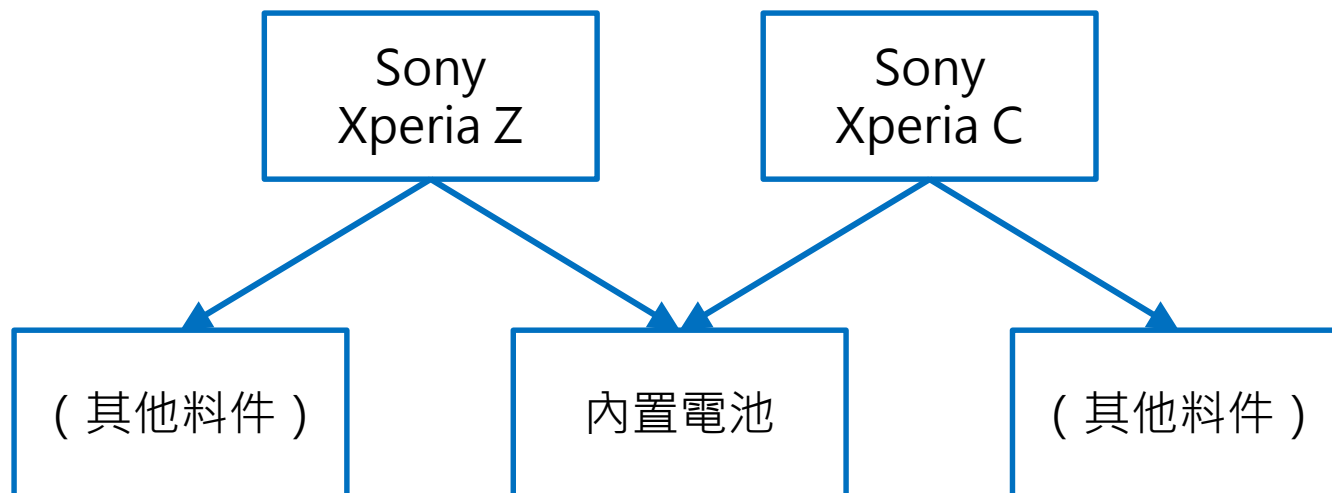
- 「料件時間序列法」簡單，也有一定效果
- 不過：
 - 未考慮料件缺貨的影響
 - 未考慮產品送修資訊
 - 未考慮共用料資訊（ Sony Xperia Z 和 Sony Xperia C 的電池 ）
- 何不用產品送修的時間序列預測送修，再計算料件耗用？

大綱

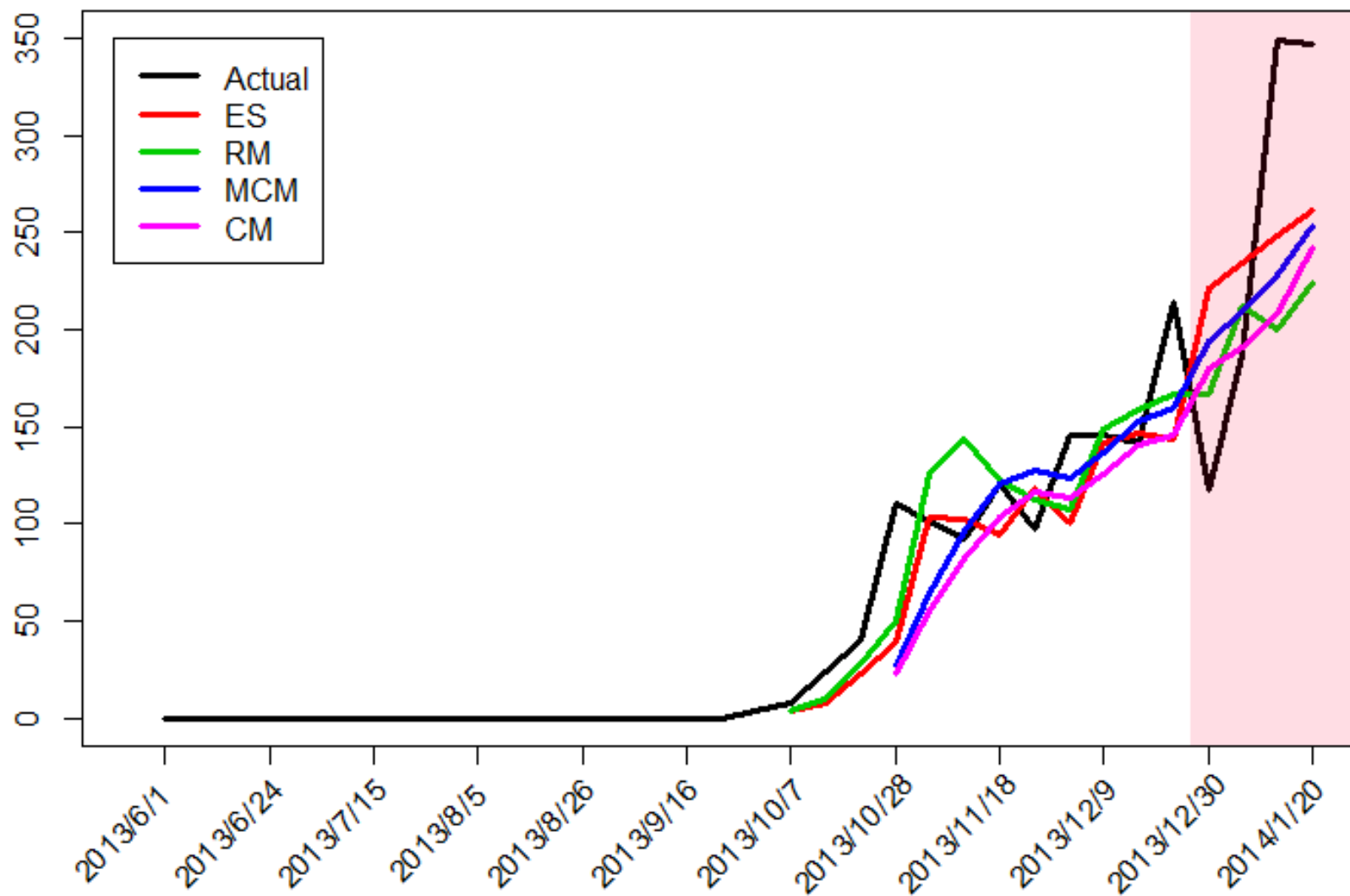
- 前言
- 料件時間序列法
- **產品時間序列法**
- 單品預測彙總法
- 比較與結論

產品時間序列法

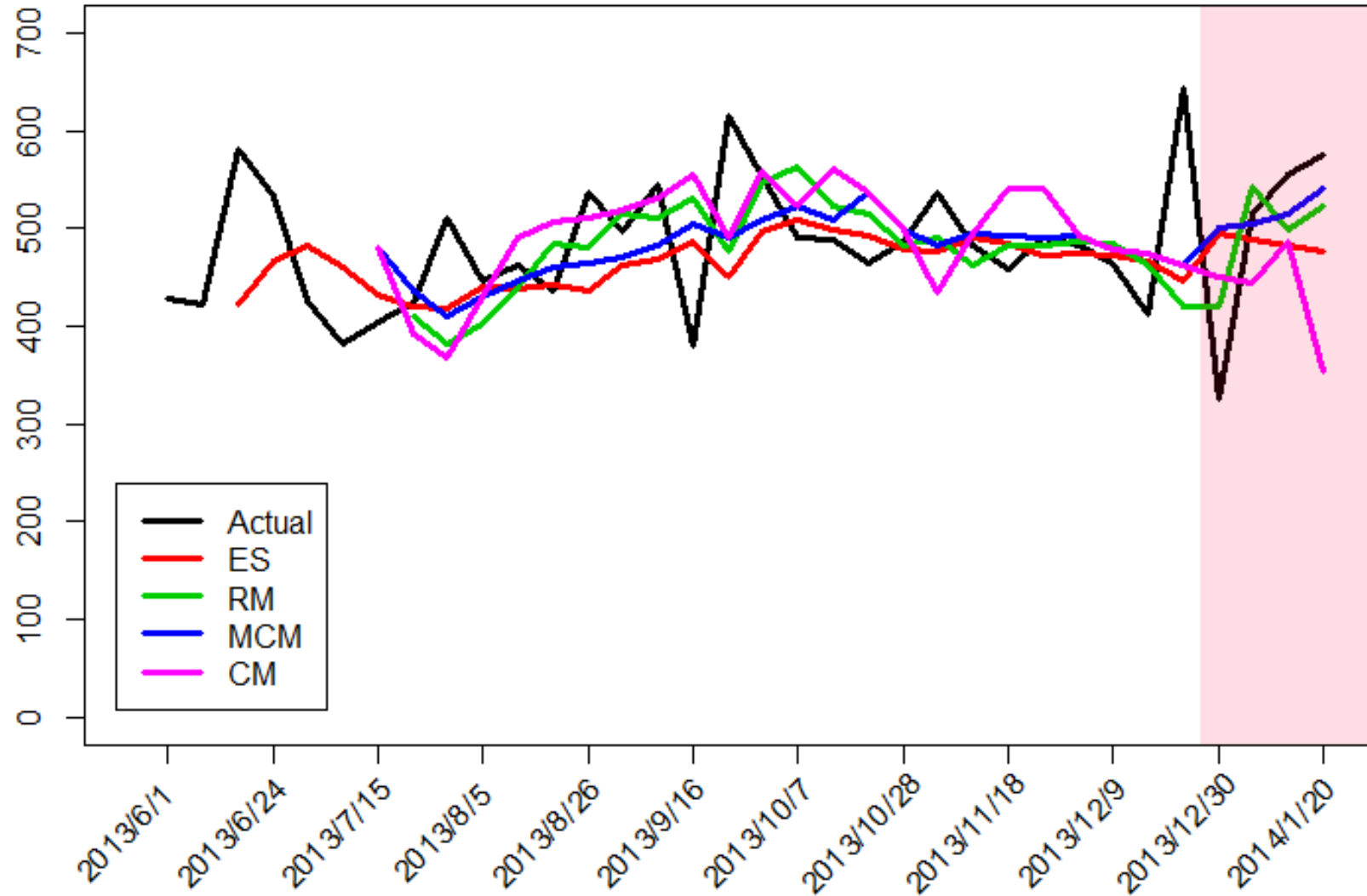
- 「產品時間序列法」：
 - 先根據產品歷史送修時間序列資料，**預測產品送修量**。
 - 統計各產品的**料件耗用歷史分佈**
 - 料件耗用預測值 = \sum 產品送修預測值 \times 料件耗用分佈



Learning and forecasting for Sony Xperia C



Learning and forecasting for Sony Xperia Z



Sony xPeria C

送修次數	料件名稱	料件耗用	耗用機率
1245	C2305/C6602 內置電池*	60	0.048
1245	C2305 前殼組件 - 黑/紫 (含液晶及觸控面板)	59	0.047
1245	轉送原廠處理	31	0.025
1245	C2305 前殼組件 - 白(含液晶及觸控面板)	25	0.021
1245	C2305 主排線	11	0.009

Sony xPeria Z

送修次數	料件名稱	料件耗用	耗用機率
14492	C6602 防水膠圈	2028	0.140
14492	轉送原廠處理	716	0.049
14492	C2305/C6602 內置電池*	520	0.036
14492	C6602 相機模組	350	0.024
14492	C6602 主排線	306	0.021

再下一步？

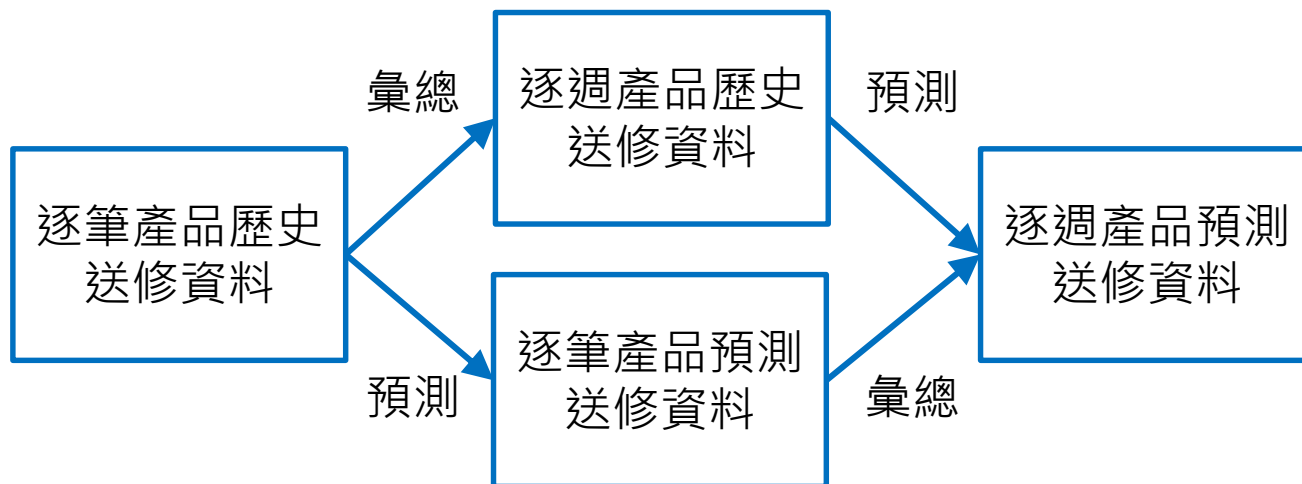
- 「產品時間序列法」簡單，也有一定效果
- 不過：
 - 未考慮產品銷售資訊
- 何不先看看目前賣掉多少，然後估算每一個流通在外的產品被送修的機率？

大綱

- 前言
- 料件時間序列法
- 產品時間序列法
- **單品預測彙總法**
- 比較與結論

單品預測彙總法

- 「單品預測彙總法」：
 - 統計各產品的銷售後送修日數歷史分佈
 - 針對每一個產品，計算其每一筆歷史銷售在下週的送修機率，合計為此產品在下週的預期送修量
 - 料件耗用預測值 = \sum 產品送修預測值 \times 料件耗用分佈



單品預測彙總法：範例

- 假設某產品過往四週的銷售與送修記錄如下：

週次	1	2	3	4	5
銷售	100	80	70	90	?
送修	5	10	25	60	?

- 要如何預測第五週的送修量？

單品預測彙總法：範例

- 先拆解送修記錄：

週次	銷售量	送修量			
		1	2	3	4
1	100	5	10	20	15
2	80	-	0	5	15
3	70	-	-	0	20
4	90	-	-	-	10

- 售出後當週送修機率 = $\frac{5+0+0+10}{100+80+70+90} \approx 4.4\%$
- 售出後隔週送修機率 = $\frac{10+5+20}{100+80+70} = 14\%$
- 售出後隔兩週送修機率 = $\frac{20+15}{100+80} \approx 20.6\%$
- 售出後隔三週送修機率 = $\frac{15}{100} \approx 15\%$

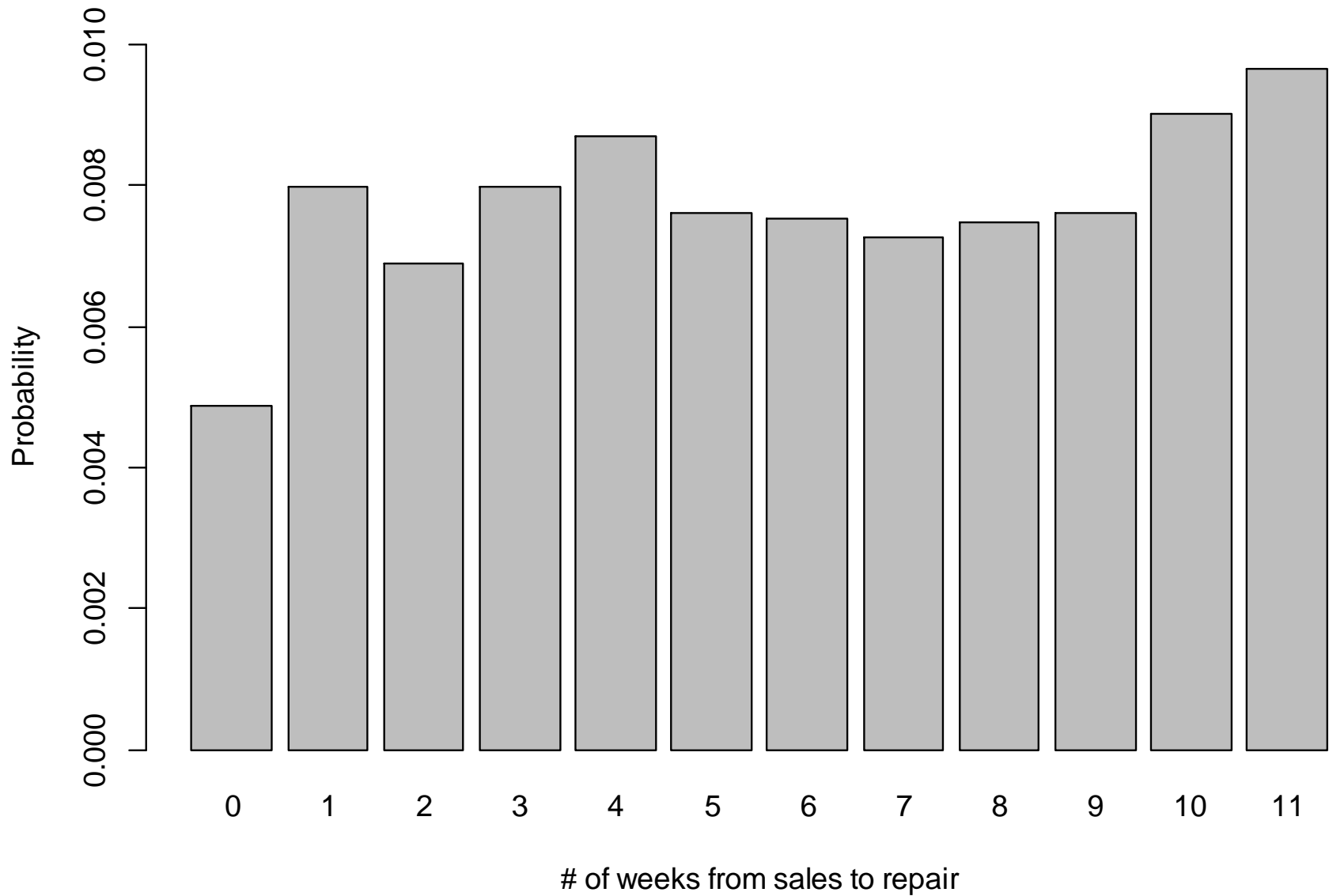
單品預測彙總法：範例

- 計算來自各週銷售的下週預測送修量

週次	銷售量	送修量				下週預測量
		1	2	3	4	
1	100	5	10	20	15	-
2	80	-	0	5	15	$80 \times 15\% = 12$
3	70	-	-	0	20	$70 \times 20.6\% = 14.4$
4	90	-	-	-	10	$90 \times 14\% = 12.6$
5	F	-	-	-	-	$F \times 4.4\%$

- 用你喜歡的方法（例如產品時間序列法）預測 F 。
- 下週預測送修量 $= 12 + 14.4 + 12.6 + 0.044F = 39 + 0.044F$

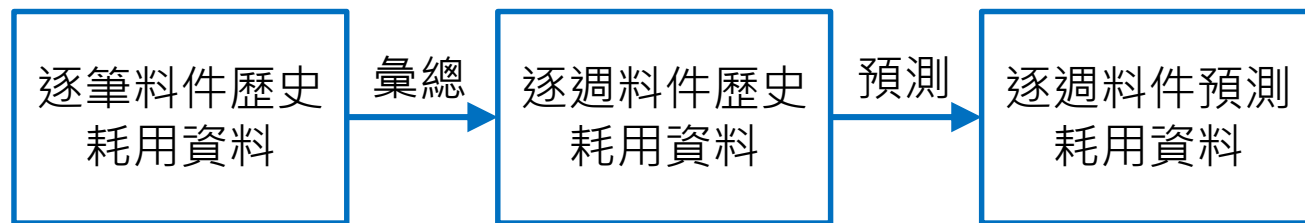
Repair distribution of Xperia C



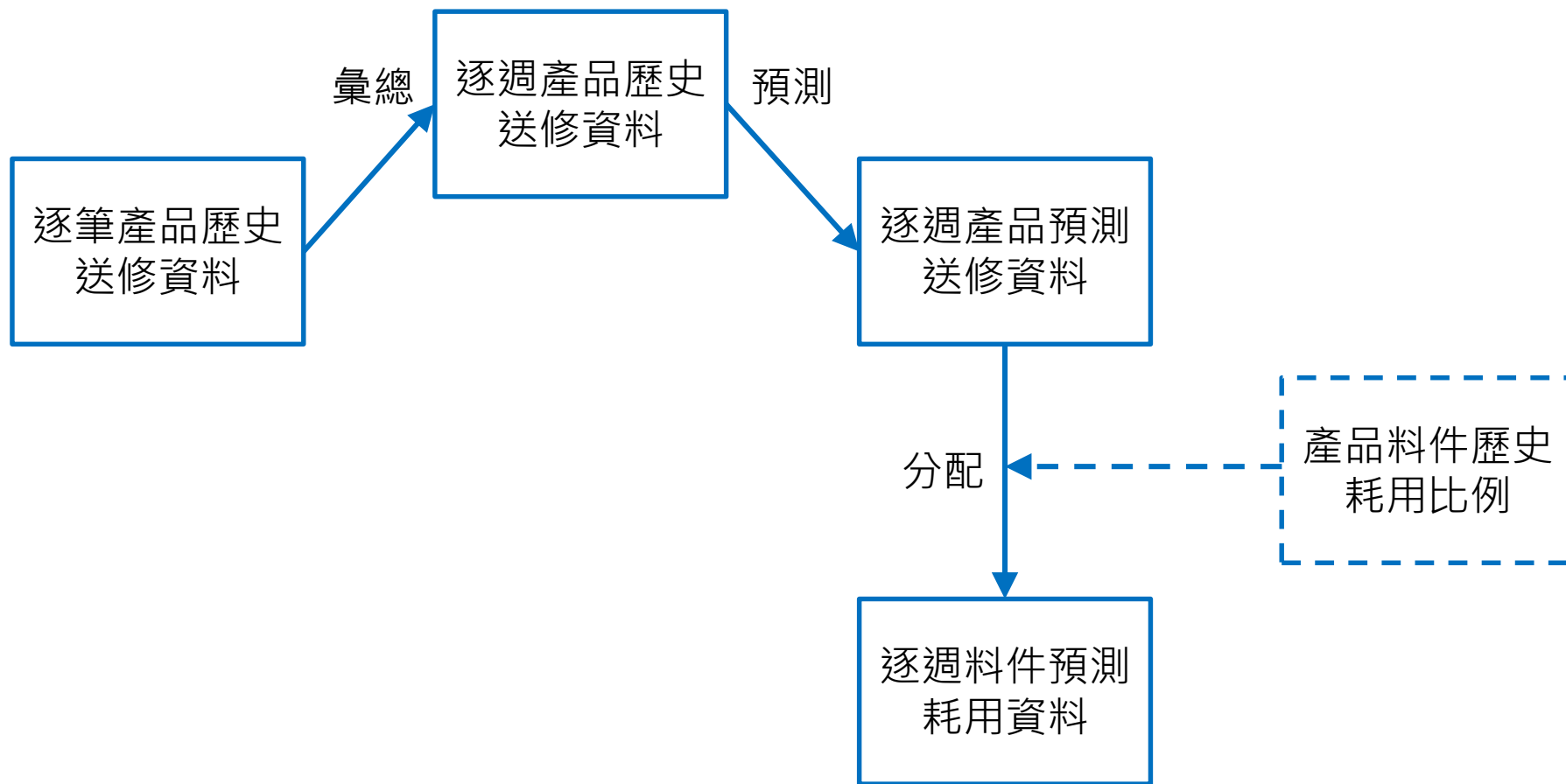
大綱

- 前言
- 料件時間序列法
- 產品時間序列法
- 單品預測彙總法
- 比較與結論

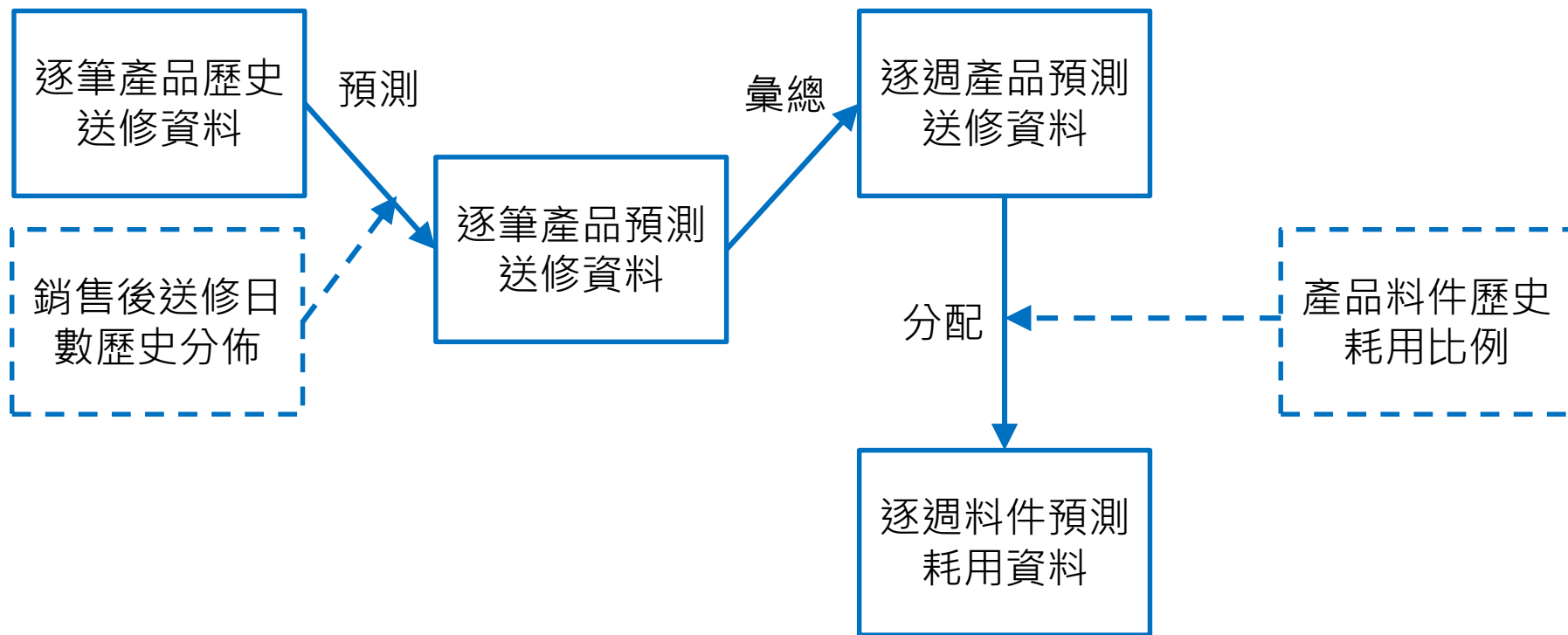
三種方法：料件時間序列法



三種方法：產品時間序列法



三種方法：單品預測彙總法



三種方法比較

- 在 2013 年上市的产品共 141 項：
 - 耗用 515 種料件或服務
 - 其中 384 種料件在 2014 年一月至少耗用一個
- 針對每一個料件：
 - 考慮 2013/1/1 到 2013/12/31 的銷售資料
 - 考慮 2013/6/1 到 2013/12/31 的維修資料
 - 以三種方法個別產生對 2014/1 的四週預測值
- 以四週平均 MAE 最小者為最佳方法

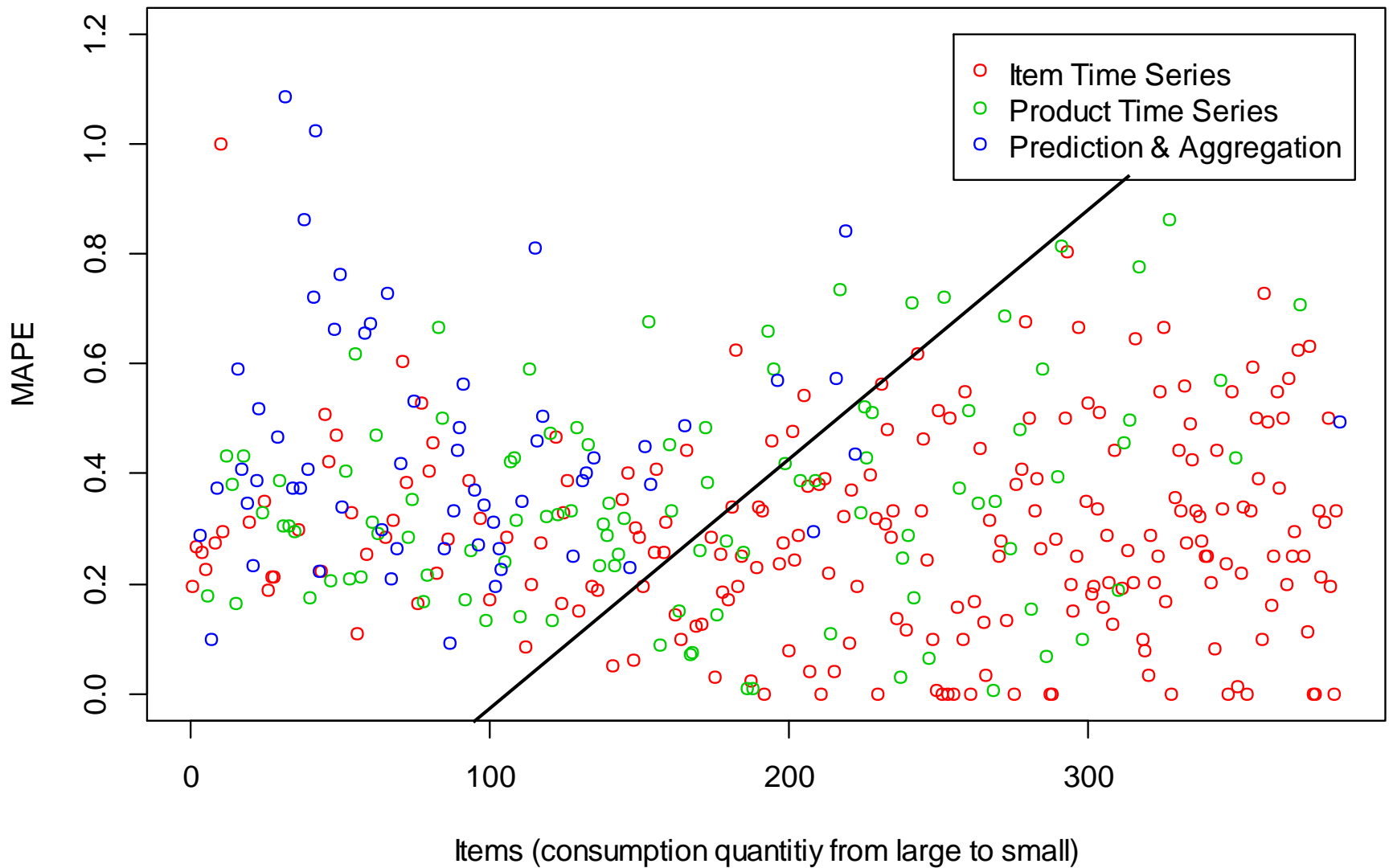
三種方法比較：最佳方法分佈

- 三種方法各適合一定比例的料件

方法	成為最佳方法的料件數	成為最佳方法的料件比例
料件時間序列法	220	57.3%
產品時間序列法	103	26.8%
單品預測彙總法	61	15.9%

- 簡單的方法有時很有效！
- 怎樣的料件適合哪個方法？

Comparisons of the three methods



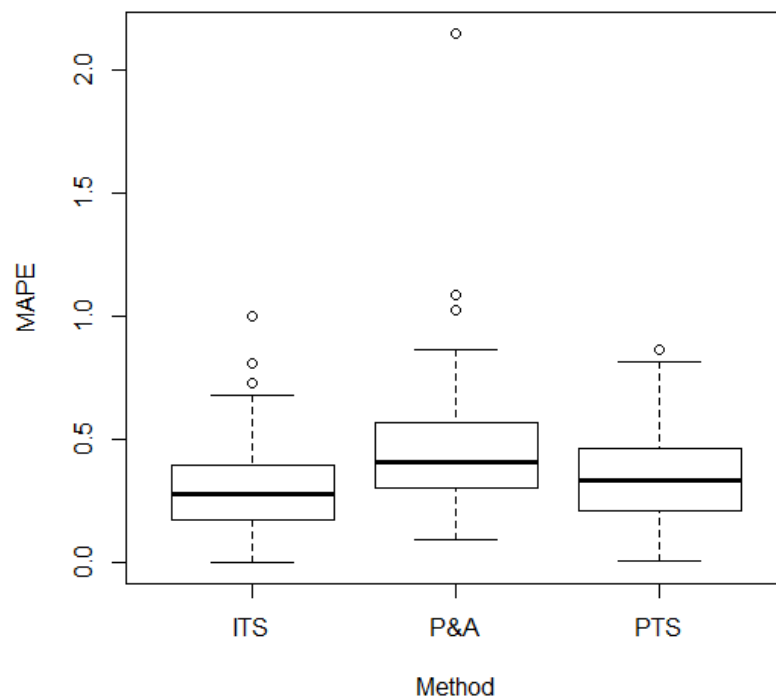
三種方法比較：預測困難度的影響

- 簡單的迴歸分析得到：

Coefficients:

	Estimate	Pr(> t)	
(Intercept)	0.28778	< 2e-16	***
P&A	0.18774	1.02e-09	***
PTS	0.06323	0.011	*

- 單品預測彙總法適合用在**比較難預測**的料件
 - 簡單的交給時間序列方法就好



三種方法比較：耗用數量的影響

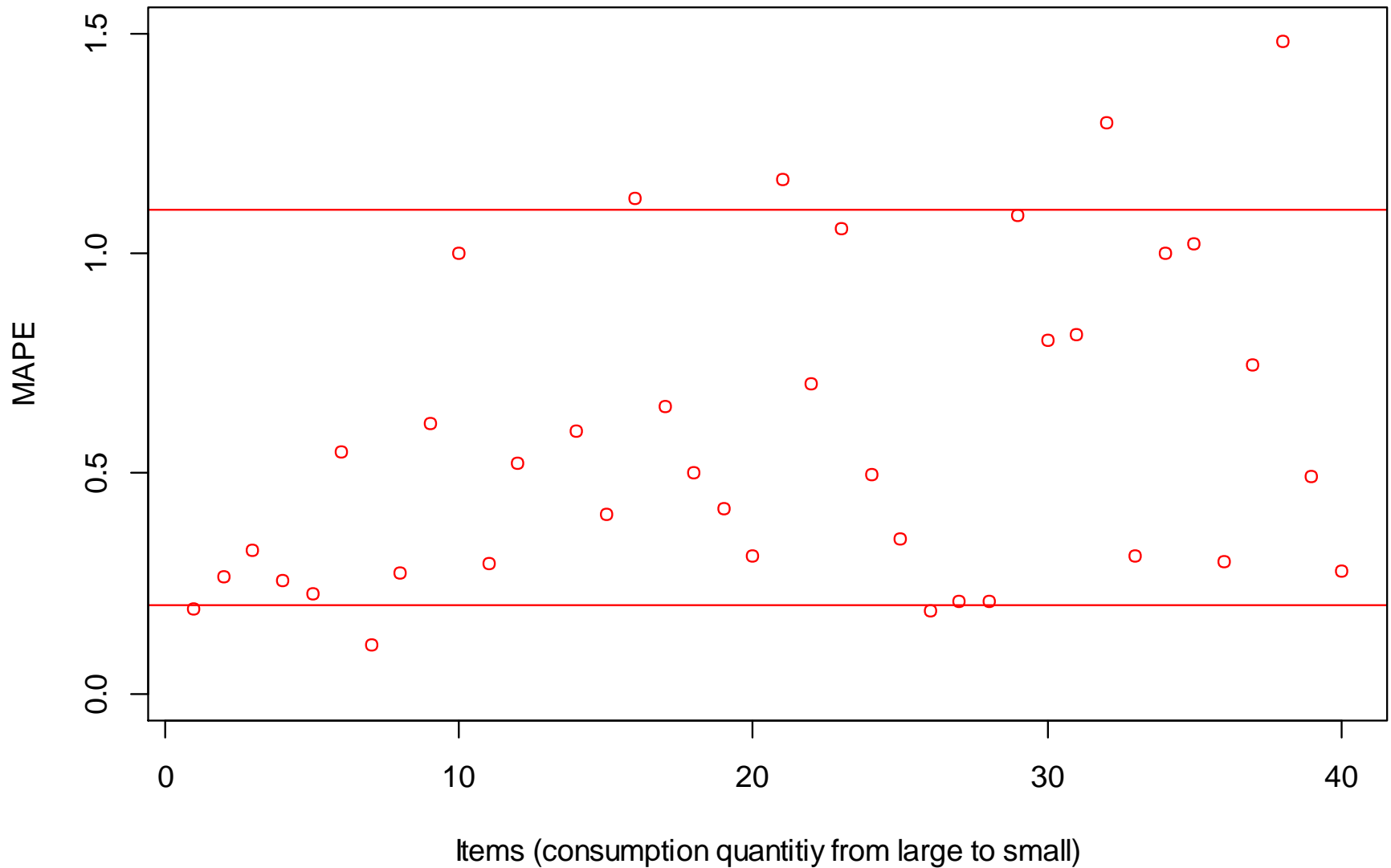
- 如果我們只看耗用至少 100 個的 40 個料件：

方法	成為最佳方法的料件數	成為最佳方法的料件比例
料件時間序列法	13	0.325
產品時間序列法	11	0.275
單品預測彙總法	16	0.4

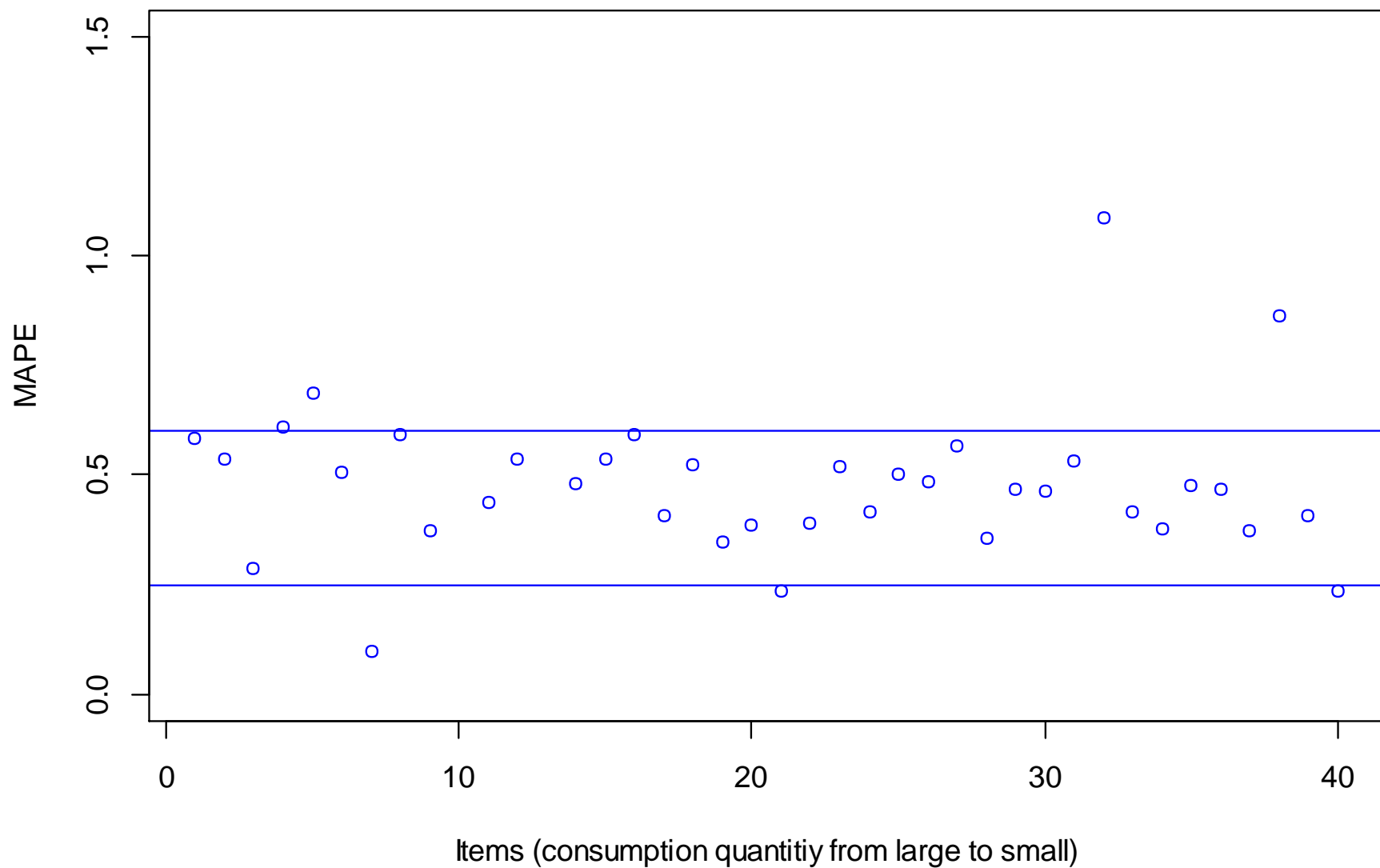
- 對**耗用數量**大的料件，考慮單品預測彙總法效益更大

- 有多大？

ITS MAPE



P&A MAPE



結論

- 手機維修料件耗用預測
- 料件時間序列法
 - 簡單、有一定效果
- 產品時間序列法
 - 考慮料件缺貨、產品送修、共用料資訊
- 單品預測彙總法
 - 考慮產品銷售
 - 適用於耗用量大、較難預測的產品
- **簡單的資料科學**就可以很有用！

謝謝！

`lckung@ntu.edu.tw`