

# 104年天氣分析與預報研討會 溫度資料檢覈技術改進之研究

多采科技有限公司 Manysplendid Infotech,Ltd.

作者:陳怡彤、鄭安孺、李天浩、顧欣怡、黃于盈、張育璋

報告者: 陳怡彤

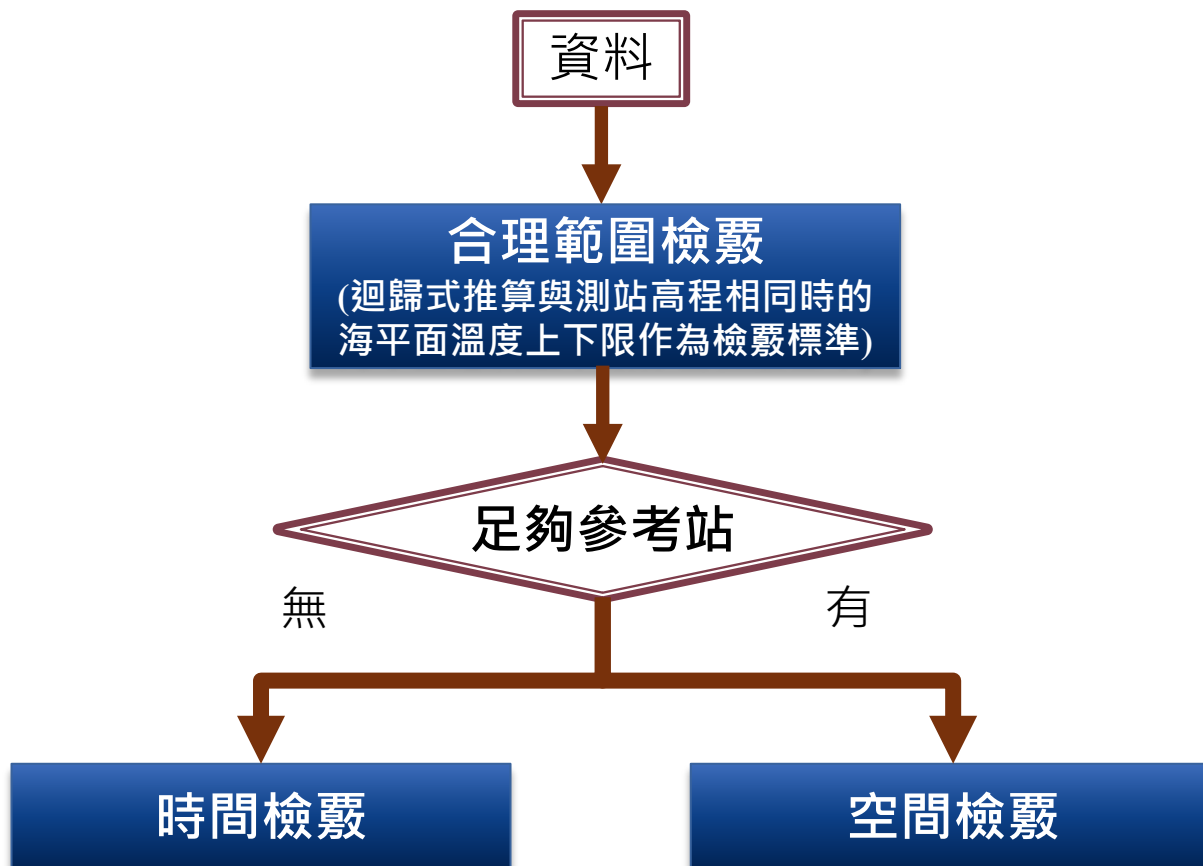
# 前言

- ▶ 溫度是影響人類活動的重要指標。
- ▶ 鄭等(2012)提出的溫度資料檢覈機制已可有效檢覈出異常溫度資料，但部分資料於清晨或劇烈天氣(午後雷陣雨)發生時，雖溫度變化劇烈，但觀測資料應為正確，卻未能通過檢覈。
- ▶ 本研究考慮測站間日、夜溫度變化特性不同，及劇烈天氣變化發生狀況，提出溫度資料檢覈技術改進方法，提升檢覈成效降低誤判機率。

# 方法改進項目

	鄭等(2012)方法	本研究改進
合理範圍檢覈	-15~45°C	<ul style="list-style-type: none"> <li>考慮到<b>高程</b>對溫度的影響，利用迴歸式<math>T_s = T_o + \gamma \cdot h</math>推算與測站高程相同時的海平面溫度上下限作為檢覈標準</li> <li>各站各期合理範圍不同</li> </ul>
空間檢覈	<ul style="list-style-type: none"> <li>利用前三年歷史觀測資料<b>逐期</b>建置參考站</li> </ul>	<ul style="list-style-type: none"> <li>利用前三年歷史觀測資料，<b>逐期逐時</b>建置參考站</li> </ul>
時間檢覈	<ul style="list-style-type: none"> <li>利用前三年歷史觀測資料逐站統計連續時間的溫差並選取累積機率<math>\leq 0.05\%</math>的溫差作為檢覈門檻值</li> </ul>	<ul style="list-style-type: none"> <li>統計2003-2012年的歷史溫差資料，訂定<b>各期逐時有雨、無雨時溫升及溫降門檻值</b></li> <li>一致性檢覈</li> </ul>

# 溫度資料檢覈流程



# 溫度資料檢覈法-合理範圍檢覈

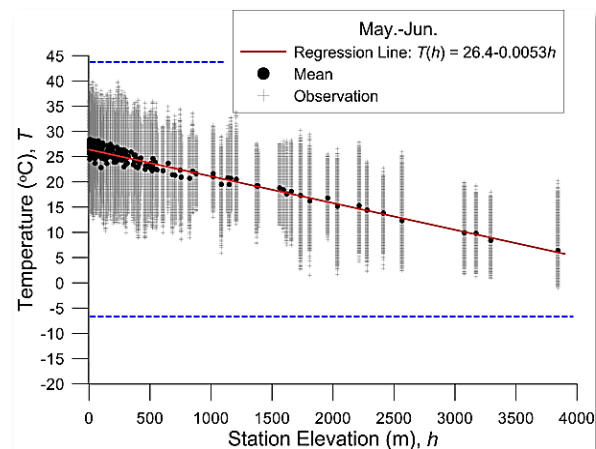
- ▶ 依據台灣氣候特性將一年分為**五期** (12-2月、3-4月、5-6月、7-9月、10-11月)
- ▶ 計算檢覈合理範圍

- 將所有測站置於同一高程後，以歷史資料統計出各期海平面溫度的最高及最低溫。

- 迴歸式  $T_s = T_o + \gamma \cdot h$  推算與測站高程相同時的海平面溫度上下限作為檢覈標準。  $T_s$  為海平面溫度， $T_o$  為測站溫度

$$T_{s.min} - \gamma \cdot h \leq T_o \leq T_{s.max} - \gamma \cdot h$$

- 若測站溫度落於上下限範圍外則視為錯誤資料，反之，則繼續檢覈。



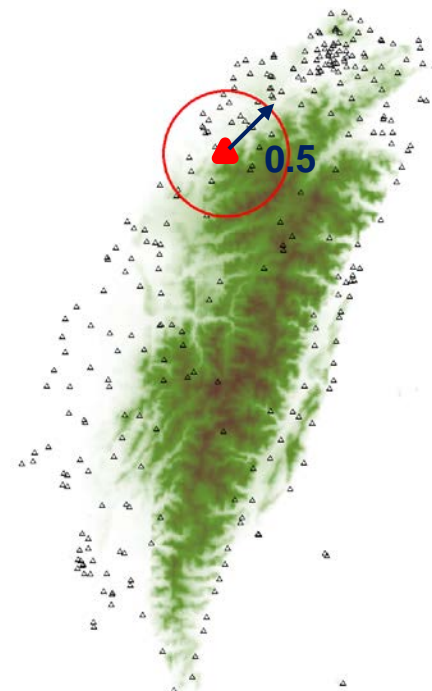
2003至2012年第三期測站溫度與高程散布圖

期	$\gamma$ 平均 遞減率	海平面 溫度上限	海平面 溫度下限
第一期	0.0046	40.0	-3.0
第二期	0.0049	43.5	-2.0
第三期	0.0053	44.5	7.0
第四期	0.0056	46.0	9.0
第五期	0.0050	45.5	0.0

# 溫度資料檢覈法-空間檢覈

## ▶ 篩選參考站

- 將一年分為五期，逐時逐期建置參考站資料
- 使用前三年歷史小時觀測資料
- 有效資料筆數>90筆
- 篩選半徑<0.5經緯度，相關係數>0.707
  - 先從待檢測站周圍半徑0.1經緯度的範圍開始搜尋
  - 選取3至5個最相關且有足夠資料數的候選站為參考站
  - 若無法選取到足夠的參考站，則逐次擴大半徑0.05經緯度直到挑選到足夠參考站或半徑達0.5經緯度為止。



# 溫度資料檢覈法-空間檢覈

- ▶ 建立參考站與待檢測站的線性迴歸式

$$y = b_0 + b_1 x \quad b_0 \text{為斜率} \cdot b_1 \text{為截距}$$

- ▶ 估計權重係數

- $\hat{y} = \sum_{j=1}^n w_j y_j$   $\hat{y}$ : 權重估計溫度,  $y_j$ : 第j個參考站溫度,  $w_j$ : 第j個參考站權重

- 應用最小估計誤差與不偏估條件

$$\sum_{j=1}^n w_j = 1$$

$$\sum_{i=1}^{n-1} w_j (cov_{ij} - cov_{in} - cov_{jn} + \sigma_n^2) = \sigma_n^2 - cov_{jn}$$

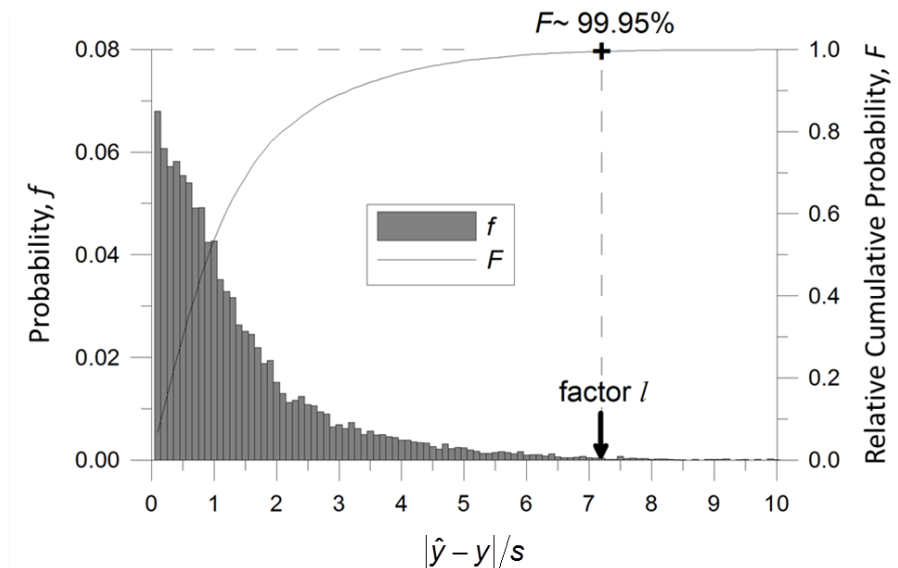
$cov_{ij}$  參考站i與參考站j對測站估計誤差協變異數

# 溫度資料檢覈法-空間檢覈

## ▶ 設定信賴區間

- $\hat{y} - 7.2 \times s \leq y \leq \hat{y} + 7.2 \times s$   
7.2為反映當估計誤差累積機率達99.95%信心區間時的倍數因子

## ▶ 當觀測值是落在信賴區間內時，接受觀測值為真



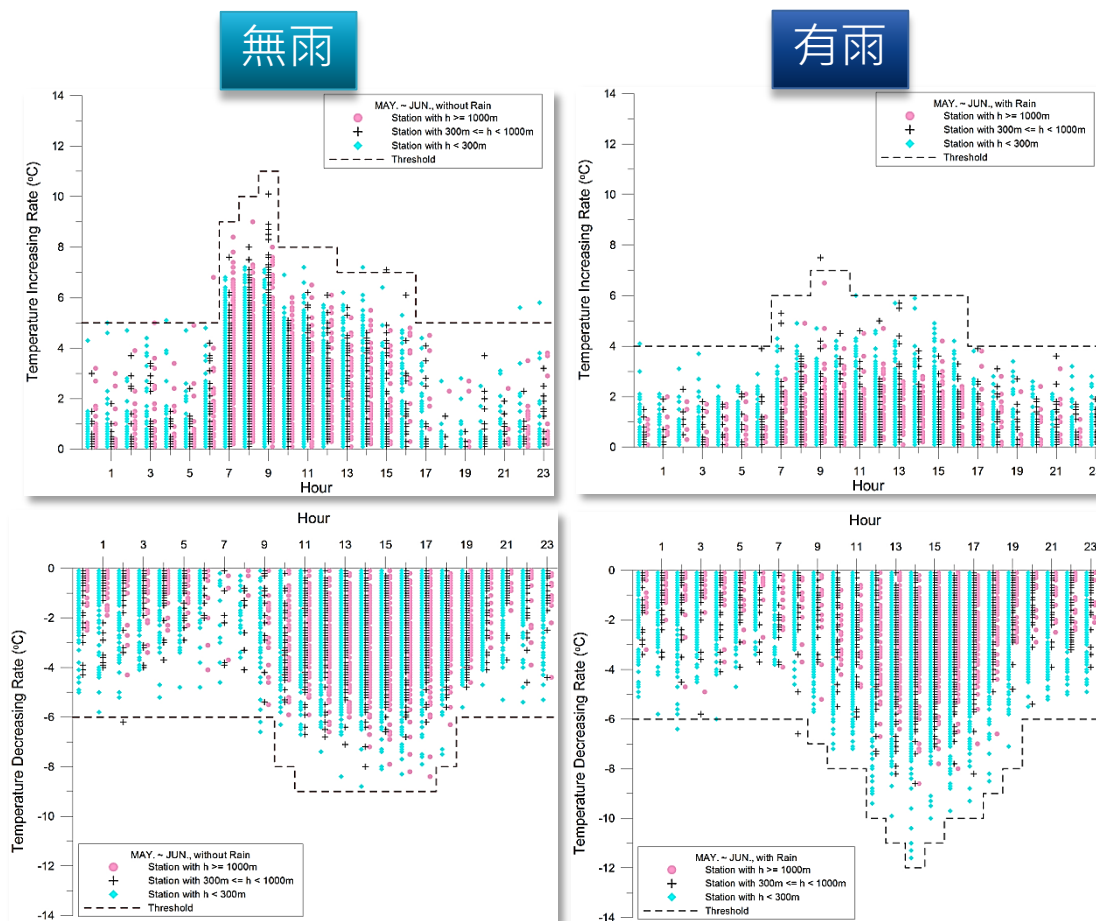
2003-2012年 $|\hat{y} - y|/s$ 機率分布及累機率圖

- 實際在進行資料檢覈時，是對所有的站先逐站計算估計誤差比值  
 $d = |\text{觀測值} - \text{估計值}| / \text{估計誤差門檻值}(7.2s)$ 
  - 若 $d > 1$ 則先將該站標示可能異常站，當所有的站都測試完畢，比較各異常站中，將d值最大的異常站剔除，重複檢覈直到異常站數量為0為止



# 溫度資料檢覈法-時間檢覈

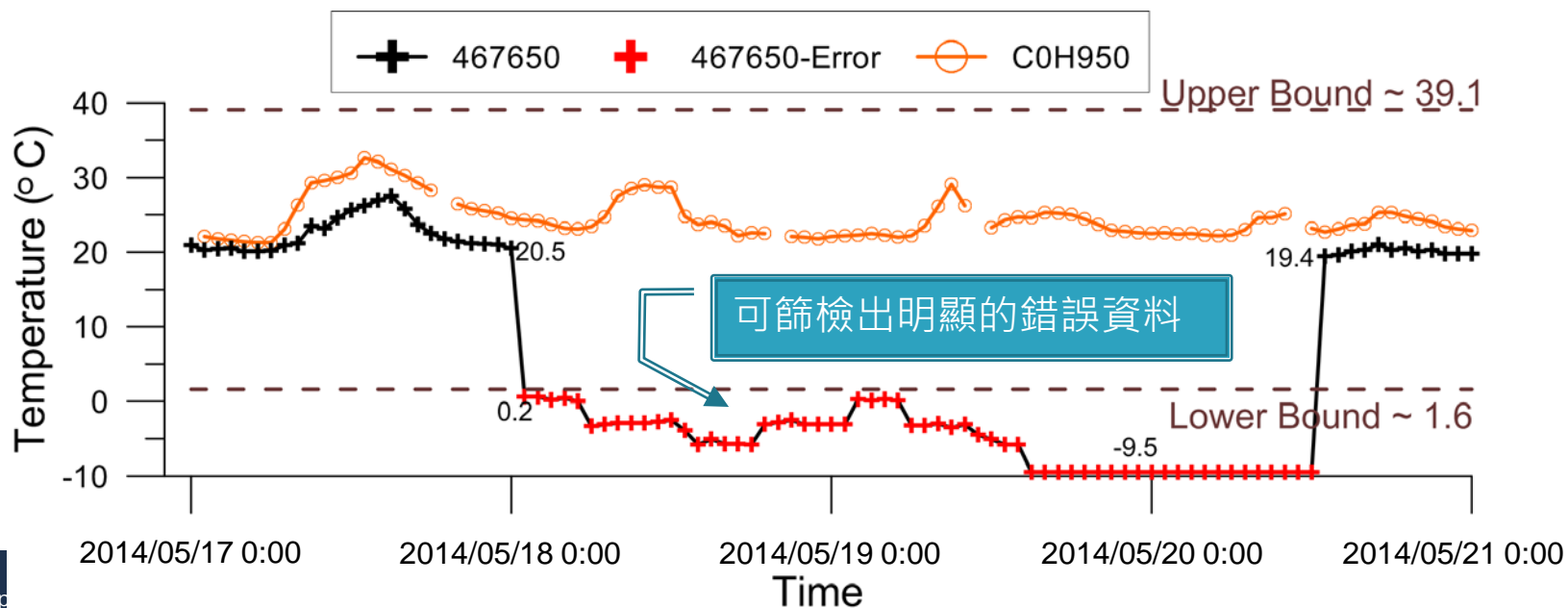
- ▶ 訂定溫升/溫降門檻值
  - 統計2003-2012年的歷史溫差資料
  - 訂定各期逐時有雨、無雨時溫升及溫降門檻值
- ▶ 溫降門檻值 < 溫差 < 溫升門檻值
  - 滿足條件則繼續檢覈，反之，視為資料有誤
- ▶ 一致性檢覈
  - 理論上待檢測站及其周圍鄰近測站溫度應有相同的變化趨勢。若差異過大，則資料有誤。



第三期(5-6月) 每小時溫升/溫降散布圖

# 個案分析-467650日月潭

- ▶ 2014年5月17日0時至21日0時467650日月潭站及其鄰近站C0H950中寮站時間序列圖
  - 鄭等(2012)合理範圍檢覈 $-15\sim 45^{\circ}\text{C}$ ，通過檢覈
  - 日月潭高度1017.5公尺，海平面上下限 $7.0\sim 44.5^{\circ}\text{C}$
  - 本研究合理範圍檢覈 $1.6\sim 39.1^{\circ}\text{C}$ ，未通過檢覈



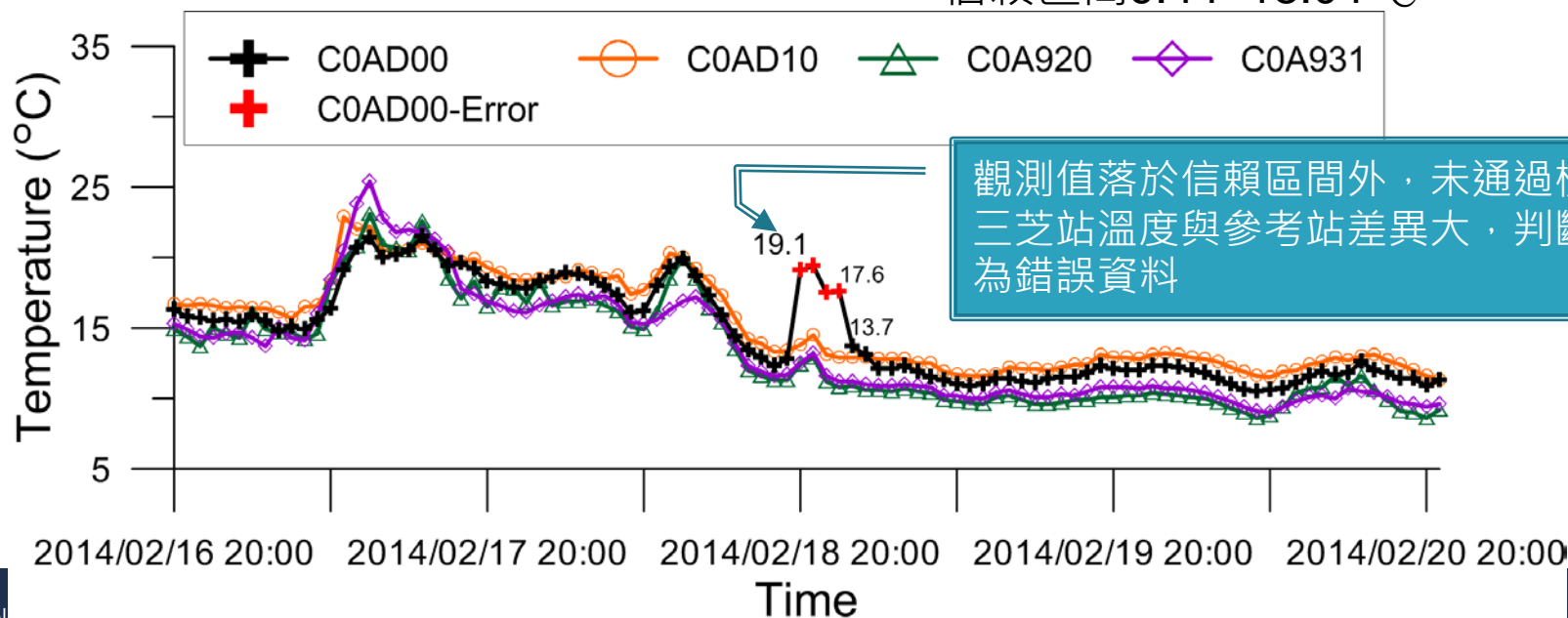
# 個案分析-C0AD00三芝

- ▶ 2014年2月16日20時至20日20時C0AD00三芝站及其參考站C0A931三和站、C0AD10八里站、及C0A920富貴角溫度時序圖

- 18日20時三芝站溫度19.1°C

參考站站碼	觀測	估計	權重
C0AD10	13.8	13.46	0.562
C0A920	12.4	13.99	0.358
C0A931	12.5	14.63	0.080

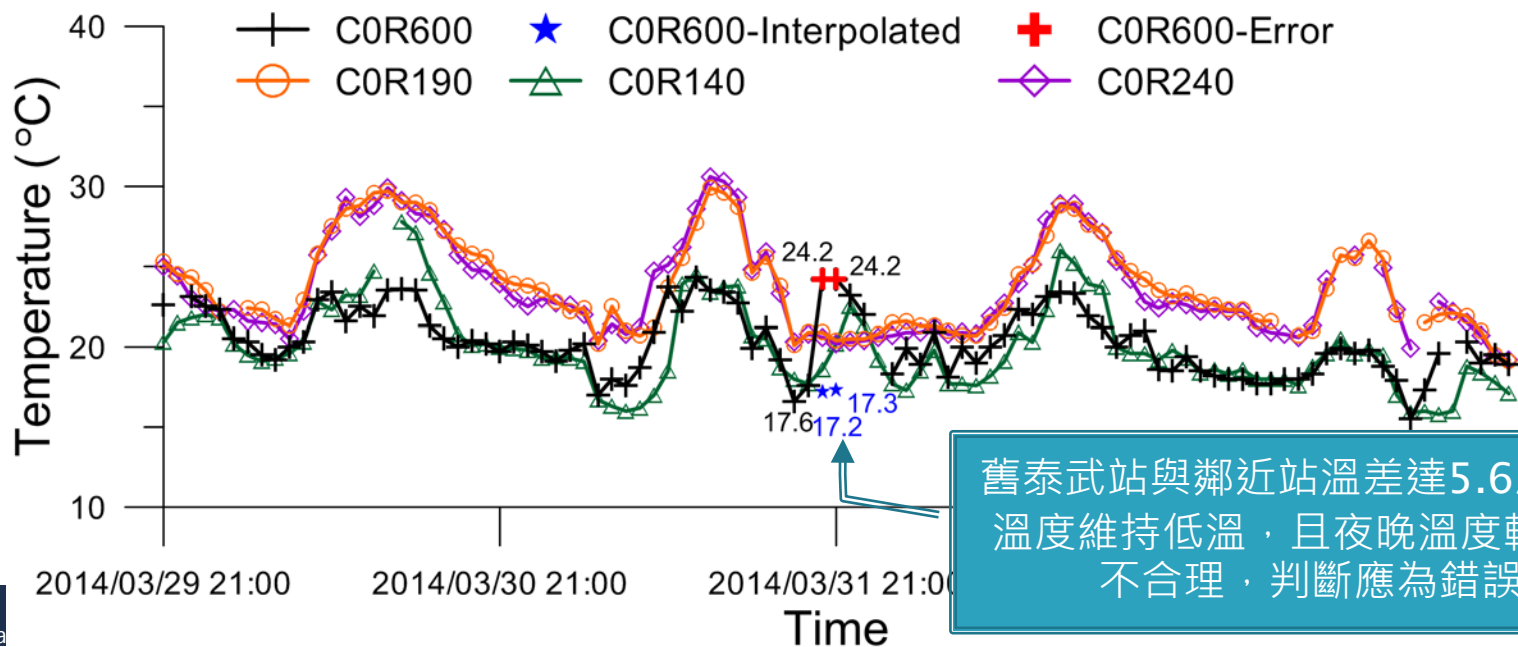
- 權重估計溫度13.74°C
- 信賴區間9.44~18.04 °C



觀測值落於信賴區間外，未通過檢覈  
三芝站溫度與參考站差異大，判斷應  
為錯誤資料

# 個案分析-COR600舊泰武

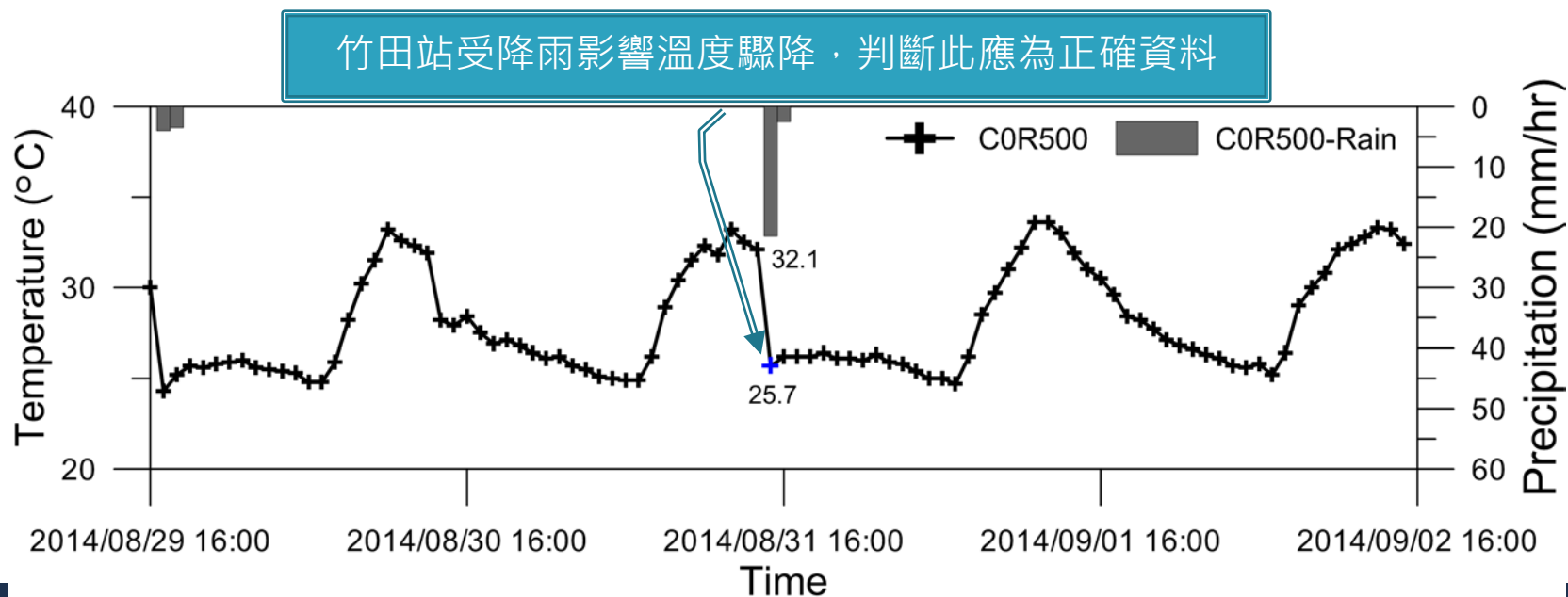
- ▶ 2014年3月29日21時至4月2日21時 COR600舊泰武站及其鄰近站COR190赤山站、COR140瑪家站及COR240來義站溫度時序圖
  - 舊泰武站31日20時溫升6.6 °C。
  - 本研究20時無降雨溫升門檻值6 °C，未通過檢覈。



舊泰武站與鄰近站溫差達5.6度，鄰近站溫度維持低溫，且夜晚溫度較白天高較不合理，判斷應為錯誤資料

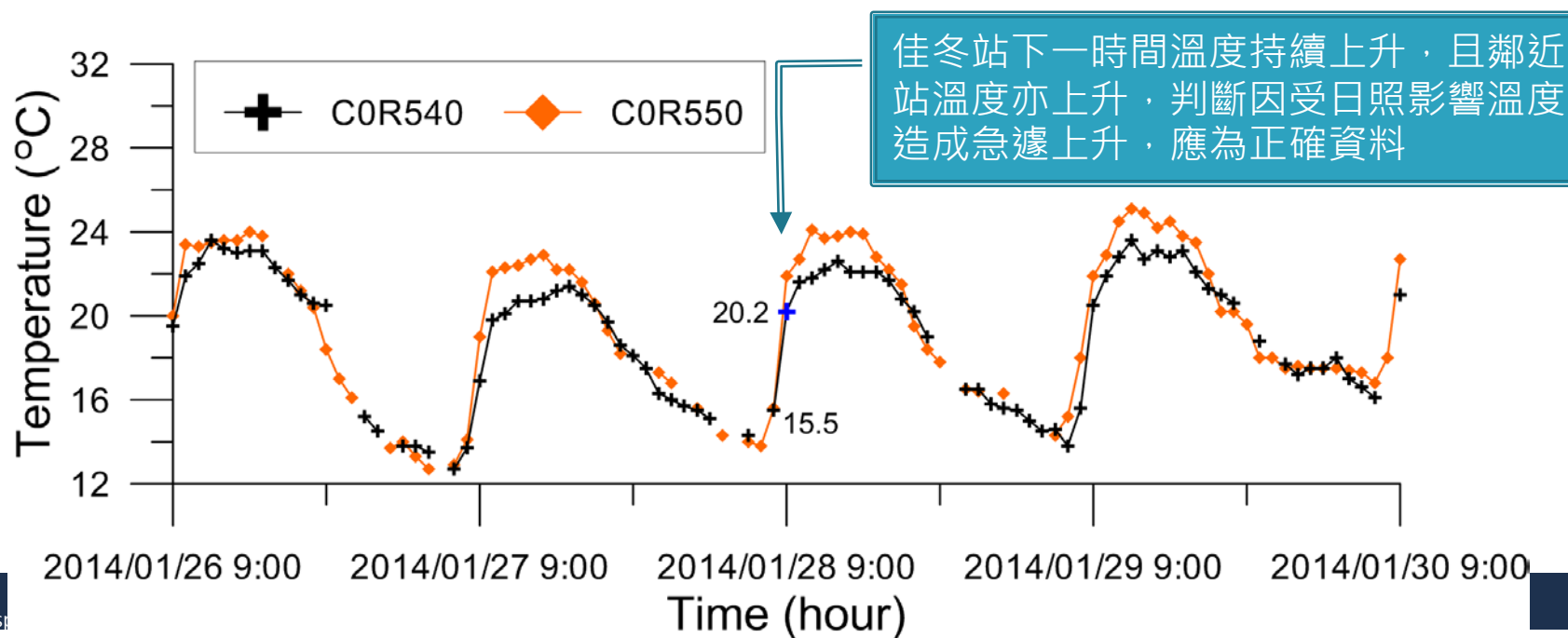
# 個案分析-COR500竹田

- ▶ 2014年8月29日16時至9月2日16時COR500竹田站溫度、降雨時序圖
  - 31日16時降雨21.5mm/hr，造成溫降6.4°C
  - 鄭等(2012)檢覈法時間溫差門檻值-5.2°C，未通過檢覈。
  - 本研究9時溫降門檻值-11°C，通過檢覈。



# 個案分析-COR540佳冬

- ▶ 2014年1月26日9時至30日9時COR540佳冬站及其鄰近站COR550新埤站溫度時序圖
  - 佳冬站28日9時溫升4.7 °C。
  - 鄭等(2012)檢覈法時間溫差門檻值4.3°C，未通過檢覈。
  - 本研究9時溫升門檻值11 °C，通過檢覈



# 結論

- ▶ 由上述個案可知，本研究提出的檢覈方法可檢覈出異常溫度資料，並有效降低劇烈天氣變化發生時對溫度資料的誤判，並提升整體檢覈成效。
- ▶ 本研究提出之溫度資料檢覈技術改進方法，為長期分析台灣測站特殊之環境背景所發展出來，故只適用於台灣測站的資料檢覈

# 報告結束

多采科技有限公司 Manysplendid Infotech,Ltd.