

高密度氣溫與雨量站觀測資料齊整率 與合理性查核分析

呂致穎¹ 林淑卿² 陳品妤¹ 楊啟瑞² 沈里音¹
中央氣象局氣象科技研究中心¹ 中央氣象局第二組²

大綱

- 前言
- 資料來源及研究方法
- 個案分析結果
- 小結與未來工作

前言

1. 中央氣象局自動氣象站及雨量站在近期的資料齊整率與早期建立時相比，有明顯的改善情形，其原因有二：(1)氣象局第二組對於自動觀測儀器的汰換及新增作業；(2)透過進行資料補整作業程序，將正確可用資料填補回去，改善了整體資料齊整率。
2. 2014年對於氣溫資料缺遺與合理性查核問題，已發展了合理的補整與偵錯方法（UK去一法、氣候離群值**Climate Outlier**）。
3. 2015年針對雨量資料特殊碼問題，陳等（2015）發展了合理處理方法；另氣溫資料的合理性查核分析，陳等（2015）提出透過觀測值與估計參考值的迴歸關係產出的迴歸離群值（**Regression Outlier**）來協助資料正確性檢測。
4. 2016年針對雨量資料合理性查核分析，應用了陳等（2016）採用超越機率值（Exceedance Probability，**EP**）來協助檢測，藉由比對六個不同降水延時下，EP值為1%、5%的雨量閾值（threshold），作為判斷氣候極端降雨事件的標準，並配合比對天氣系統，判斷是否為合理值。

資料來源

1. 為符合全臺測站分佈密度的均勻性，因此設定測站選取條件有二：1.1998年至2012年共15年；2.資料齊整率需 $\geq 85\%$ ，選取符合條件之測站有**110個氣溫目標站**及**318個雨量目標站**。
2. **氣溫資料**採用於即時接收階段即已經過簡易檢核處理，依物理特性給定合理性範圍($T < -20^{\circ}\text{C}$ 或 $T > 50^{\circ}\text{C}$)進行大型錯誤檢核 (Gross error check) 的原始資料，資料長度為1998年至2016年等近19年的逐時資料。
3. **雨量資料**則採用經過處理雨量特殊碼的QC (Quality Control) 資料。參考陳等 (2015) 的處理方法，將累積於後的雨量值為正的資料視為缺值；並將累積於後的雨量值為0的連續特殊碼筆數，小於24筆的皆修正為0。資料長度同為1998年至2016年等近19年的逐時資料。

研究方法

➤ 氣溫資料查核

1. 首先採用氣候背景範圍進行氣候檢驗 (Climatological test) ，即是以氣候平均背景值來提取氣候離群值 (Climate Outlier , CO) 。
2. 參考陳等 (2015) 藉由具可信度的參考推估值(UK值)來檢測觀測數據的合理性，並運用迴歸方法由迴歸離群值 (Regression Outlier , RO) 來尋找可疑的錯誤數據。
3. 上述兩個方法找出的可疑資料，再經由天氣系統比對，判斷是否為合理值。

➤ 雨量資料查核

1. 參考陳等 (2016) 先將資料分為夏、冬兩季，其中夏季月份為4至9月，冬季月份為10月至次年的3月，計算318個雨量測站的觀測值超越機率值 (Exceedance Probability ，簡稱EP值) 。
2. 將經處理後的雨量資料以箱型計數 (box-counting) 方式重新組成為六種延時的降雨事件 (1hr、3hr、6hr、12 hr、24 hr和48 hr) ，根據線性動差法的推估結果採用皮爾森第三型分布，模擬降雨事件雨量的母體機率分布，並計算EP值為1%和5%的雨量閾值作為判斷氣候極端降雨事件的標準，再經由天氣系統比對，判斷是否為合理值。

個案分析統計-氣溫(CO方法)

➤ 氣溫資料查核-氣候離群值(CO)方法

- 2011年至2016年氣候離群值筆數統計 (紅字為錯誤筆數)

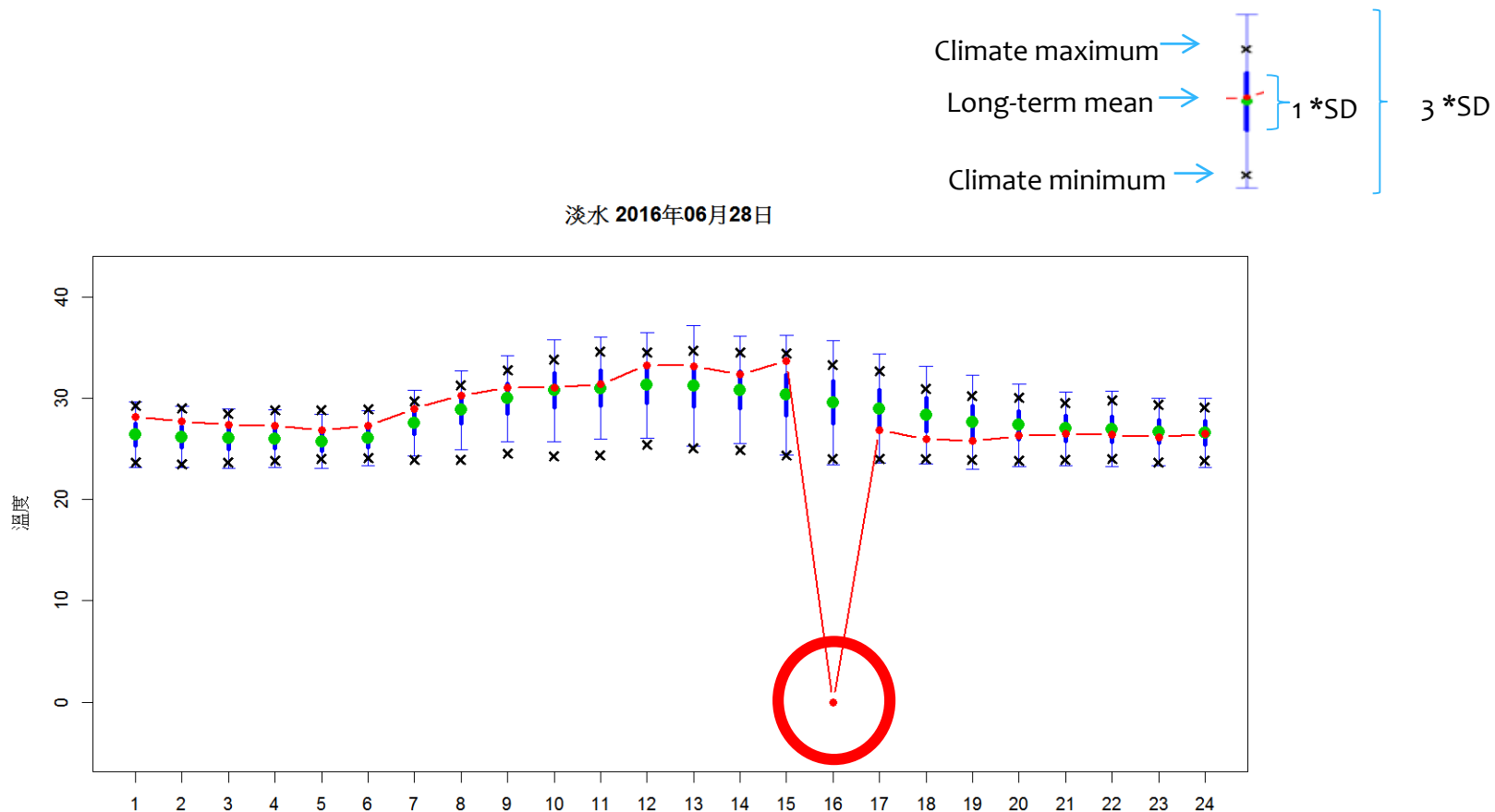
	> 5SD	> 6SD	> 7SD	> 8SD	> 9SD	> 10SD
2011	17	0	0	0	0	0
2012	24	5	0	0	0	0
2013	165	45	21	9	5	3
2014	45	11	5	1	1	1
2015	233	95	33	7	1	0
2016	82	12	1	2	1	1

PS. 2016年資料僅統計至8/31止

- ✓ 於近6年的566筆可疑資料中，大部份多為合理值，
僅有3筆為錯誤值。

錯誤資料個案-氣溫(CO方法)

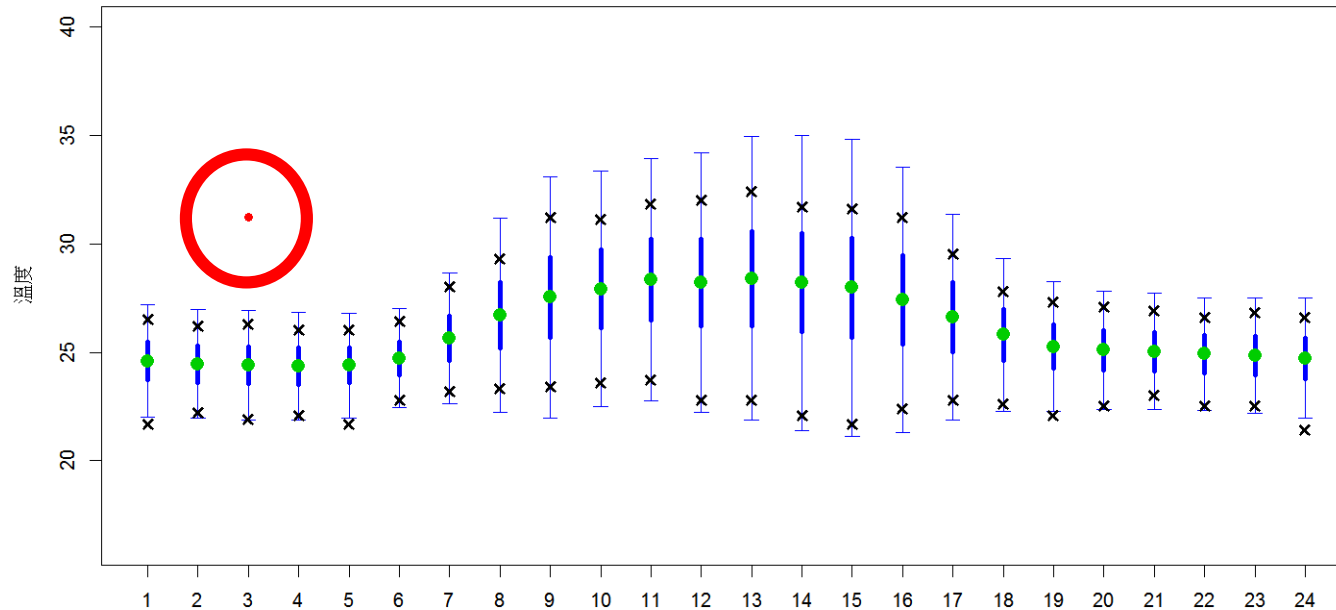
➤ 觀測值為0，資料有明顯不連續情形，判斷為錯誤值。



錯誤資料個案-氣溫(CO方法)

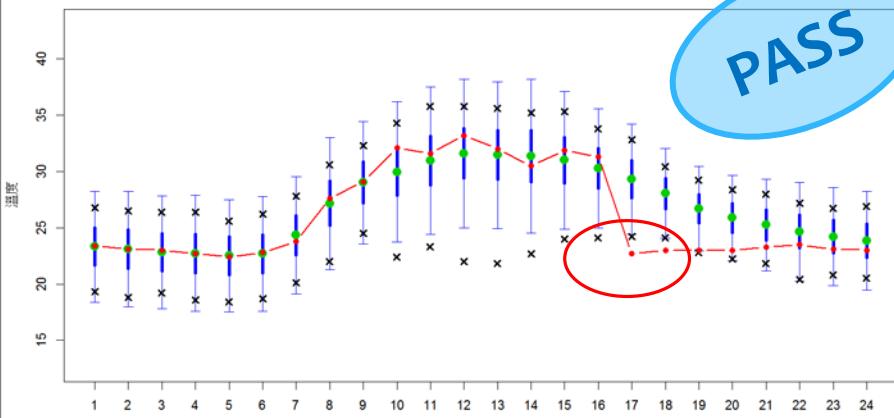
- 資料傳輸過程中受電波干擾，導致出現錯誤溫度值，該段受干擾期間資料應皆設為故障值。

檳榔 2016年06月20日



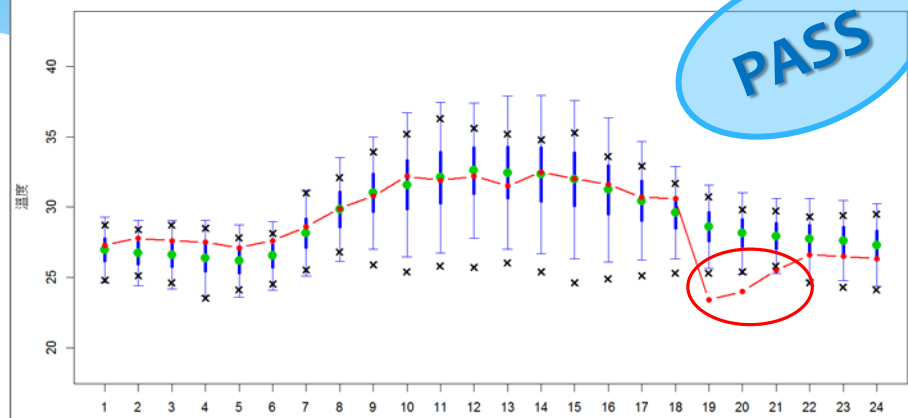
合理資料個案-氣溫(CO方法)

古亭坑 2016年05月13日



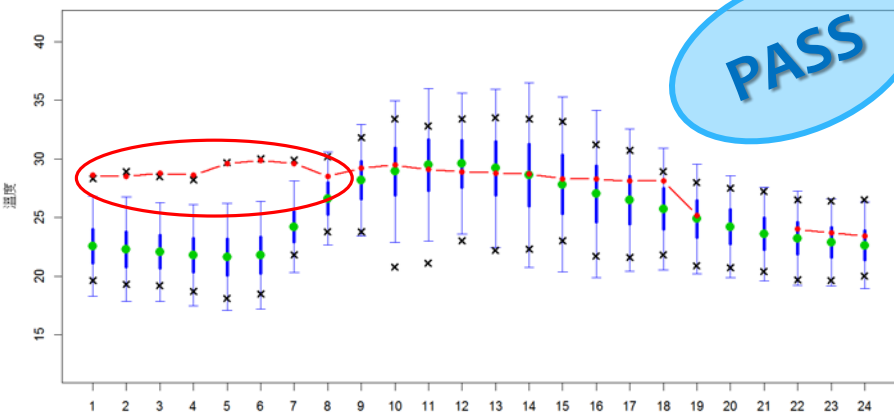
通過檢查原因：
5/13 17L 該站時雨量41mm，造成迅速降溫現象。

鹿港 2016年07月3日



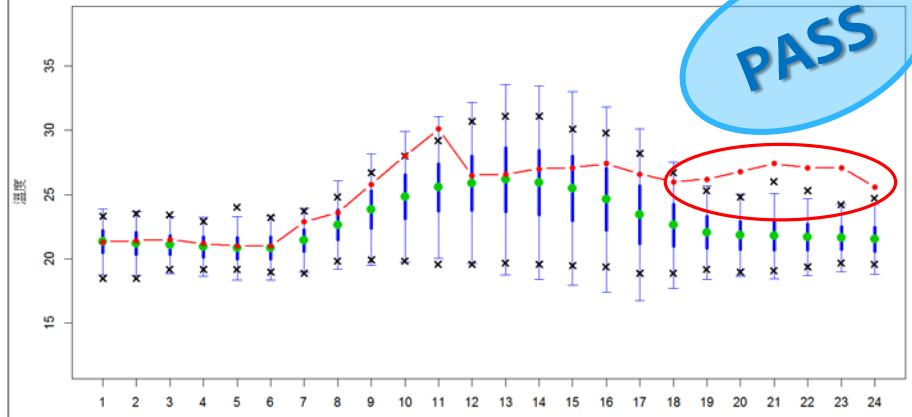
通過檢查原因：
7/3 19L~21L 連續3小時有降雨，導致氣溫迅速下降。

梅花 2016年07月8日



通過檢查原因：
7/8 受尼伯特颱風過山沉降增溫，夜晚出現較高氣溫值。

日月潭 2016年07月31日



通過檢查原因：
7/31 受妮妲颱風外圍環流過山沉降增溫，夜晚有高氣溫值。

合理資料分析(CO)-天氣系統統計

➤ 統計造成近6年較大氣候離群值(CO>5)的特殊天氣現象，說明如下：

■ 高溫(離群值為正值)：

- 颱風(熱帶系統)或其外圍環流造成氣流過山沉降，產生增溫作用。
- 西南風增強，造成氣流過山沉降，產生增溫作用。
- 焚風現象。
- 太平洋高壓增強。
- 鋒面前緣暖區增溫作用。

■ 低溫(離群值為負值)：

- 寒流、強烈大陸冷氣團、大陸冷氣團、東北季風。
- 輻射冷卻效應
- 颱風或其外圍環流造成的降雨。
- 西南氣流降雨。
- 鋒面降雨。
- 午後熱對流降雨。

個案分析統計-氣溫(RO方法)

➤ 氣溫資料查核-迴歸離群值(RO)方法

- 2011年1月至2016年8月迴歸離群值>10筆數統計 (紅字為錯誤值)

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
2011	0	0	0	0	0	0	0	0	0	0	0	0
2012	0	0	0	0	0	0	0	0	0	0	0	0
2013	0	0	0	2	0	0	0	0	0	0	0	0
2014	0	0	7	0	0	0	0	0	1	1	0	0
2015	0	0	0	0	0	2	0	1	0	0	0	0
2016	0	0	0	0	1	2	0	0	—	—	—	—

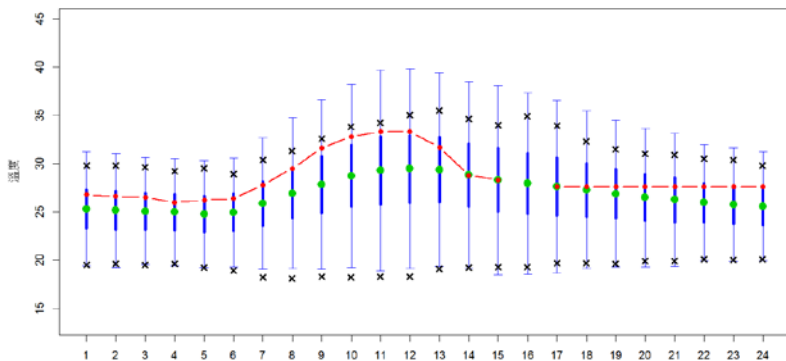
PS. 2016年資料僅統計至8/31止

- ✓ 於近6年的17筆RO>10可疑資料中，判斷有5筆為合理資料，主要為特殊天氣系統造成UK參考值的高估或低估，導致RO值有偏大的狀況；另有12筆為錯誤資料。

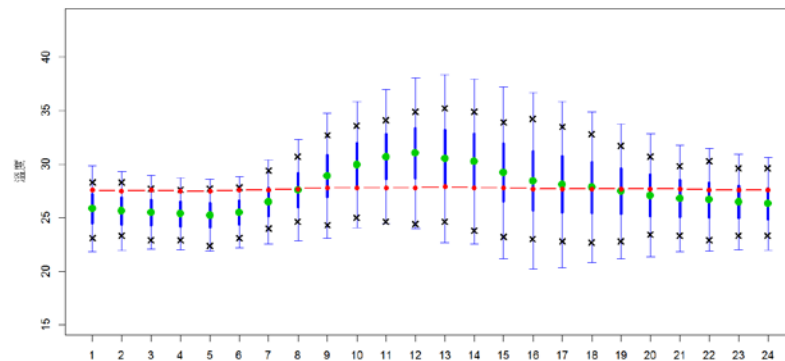
錯誤資料個案-氣溫(RO方法)

➤ COA9G0南港站於2015年6月14日的17時至17日的12時有連續資料錯誤的情形。

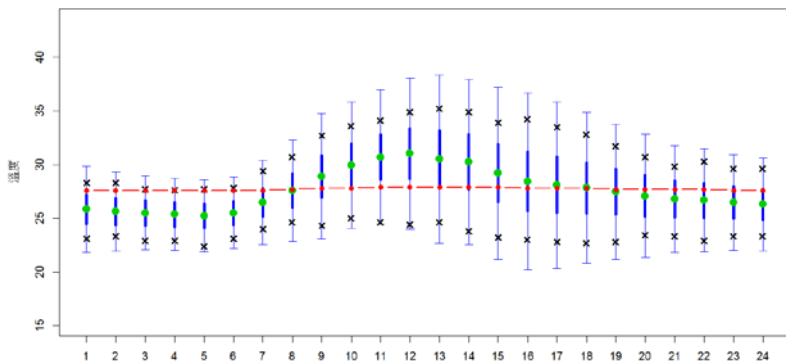
南港 2015年06月14日



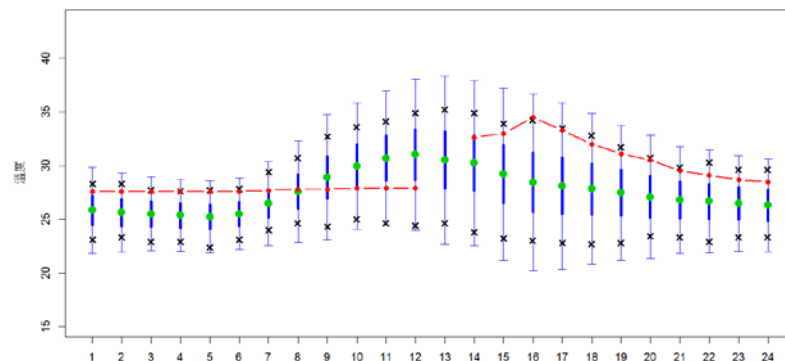
南港 2015年06月15日



南港 2015年06月16日



南港 2015年06月17日



合理資料個案-氣溫(RO方法)

➤ C0A9F0內湖站於2016年5月9日15時的RO值為-10.664。

1	date	TT	UK	dist	stud	abs_stud	hatV	hr	stname	dist99	corr	
2	20160509	23.1	28.4	0.089	-10.664	10.664	0.002	15	C0A9F0	Cook_99%=	0.013	0.988

C0A9F0 UK參考值及其附近參考站實際觀測值如下所列：

datehh	TT	UK_T1O	C0A9A0	C0A9G0	466920
2016050915	23.1	28.4	26.8	23.9	27.2

✓ 判斷為合理值

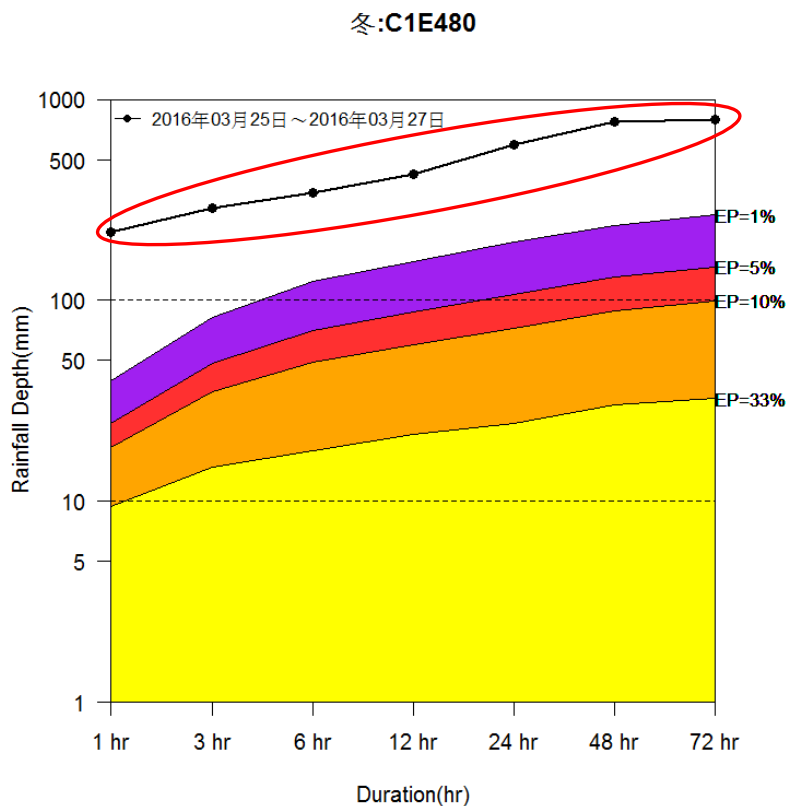
原因：2016年5月9日各地晴朗炎熱，北臺灣、花蓮及中南部山區因午後對流旺盛，局部地區有大雨發生。C0A9F0內湖站於5/9 15時有局部降雨15.5mm，但附近的參考站並無降雨，造成UK參考值有高估情形。

合理資料分析(RO)-天氣系統統計

- 統計造成近6年較大迴歸離群值的特殊天氣現象，說明如下：
 - 日夜溫差大造成UK參考值的高估或低估
 - 颱風環流過山沉降增溫（焚風）造成UK參考值的低估
 - 局部降雨造成UK參考值的高估。

錯誤資料個案-雨量(EP值方法)

➤ 2016年3月26日至27日鳳美站有出現異常大值的逐時雨量，最大值為217.5mm。



➤ 2016年3月C1E480鳳美站超越EP值=5%的時雨量值

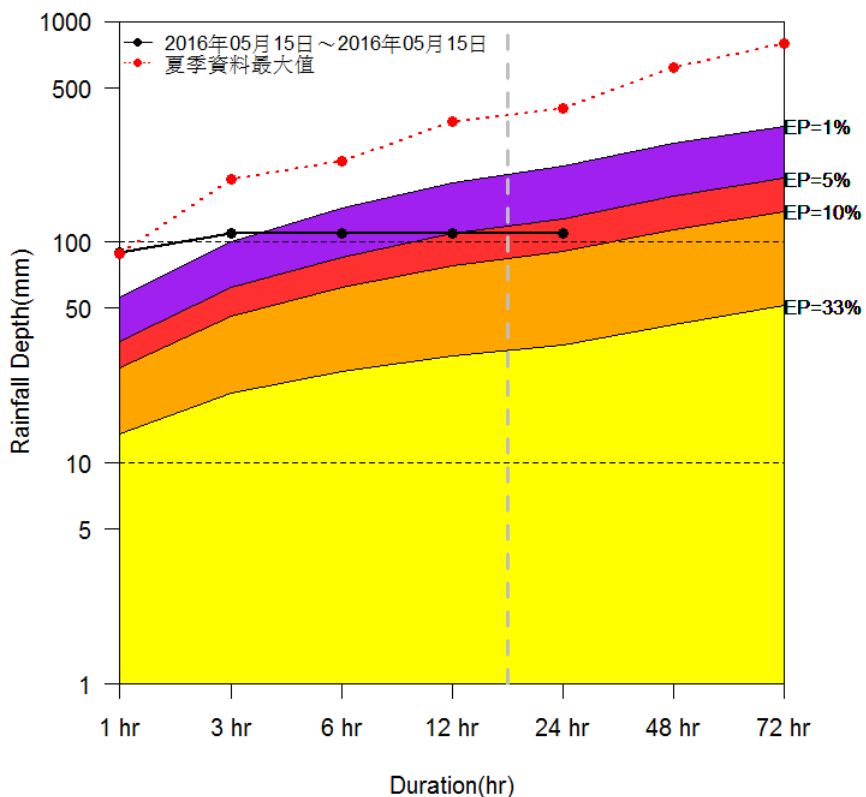
C1E480	133/26-03	136/26-04	74/26-08	70.5/26-12	43/26-19
C1E480	217.5/27-03	47.5/27-04			

■ 此段期間並無明顯的降雨天氣系統，因此判斷該期間資料為錯誤值；另查詢二組該站的維護紀錄後，確認於3月22日至3月27日期間，**自動雨量儀的磁簧開關故障**，造成有異常的雨量值出現，此紀錄支持其雨量觀測值有高度異常的可能性，因此該站於此期間的資料應皆改為故障值-9999。

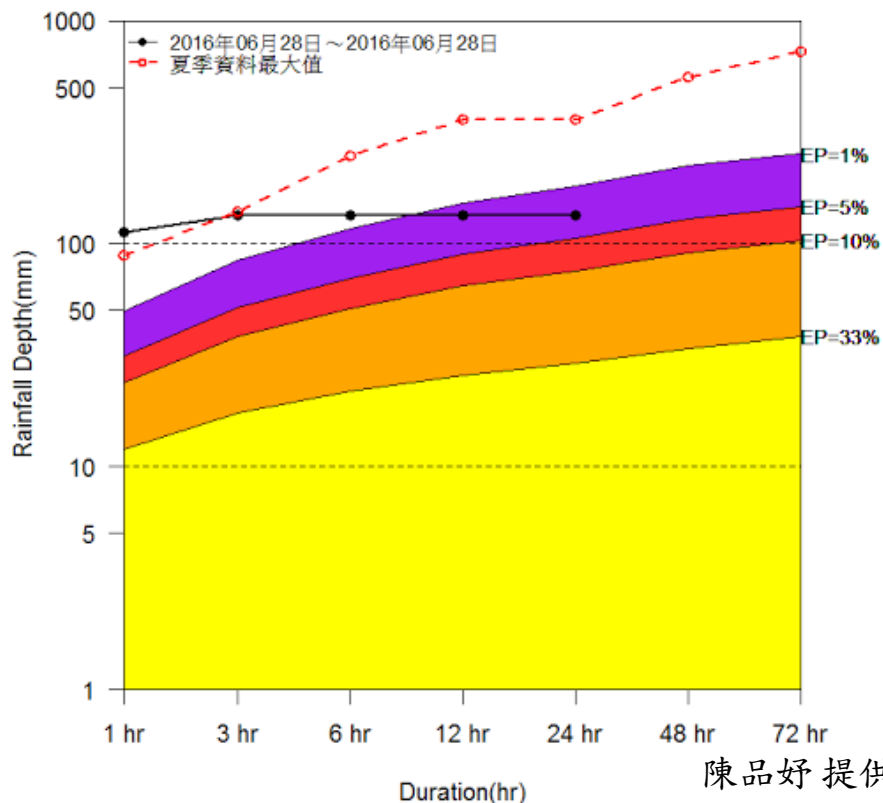
合理資料個案-雨量(EP值方法)

- 2016年5月15日屈尺站於15時的雨量值為89mm及2016年6月28日八德站於16時的雨量值為112.5mm，兩者皆創下該站設站以來的整點雨量新紀錄。

夏: C0A580屈尺



夏: C1C490八德



小結與未來工作

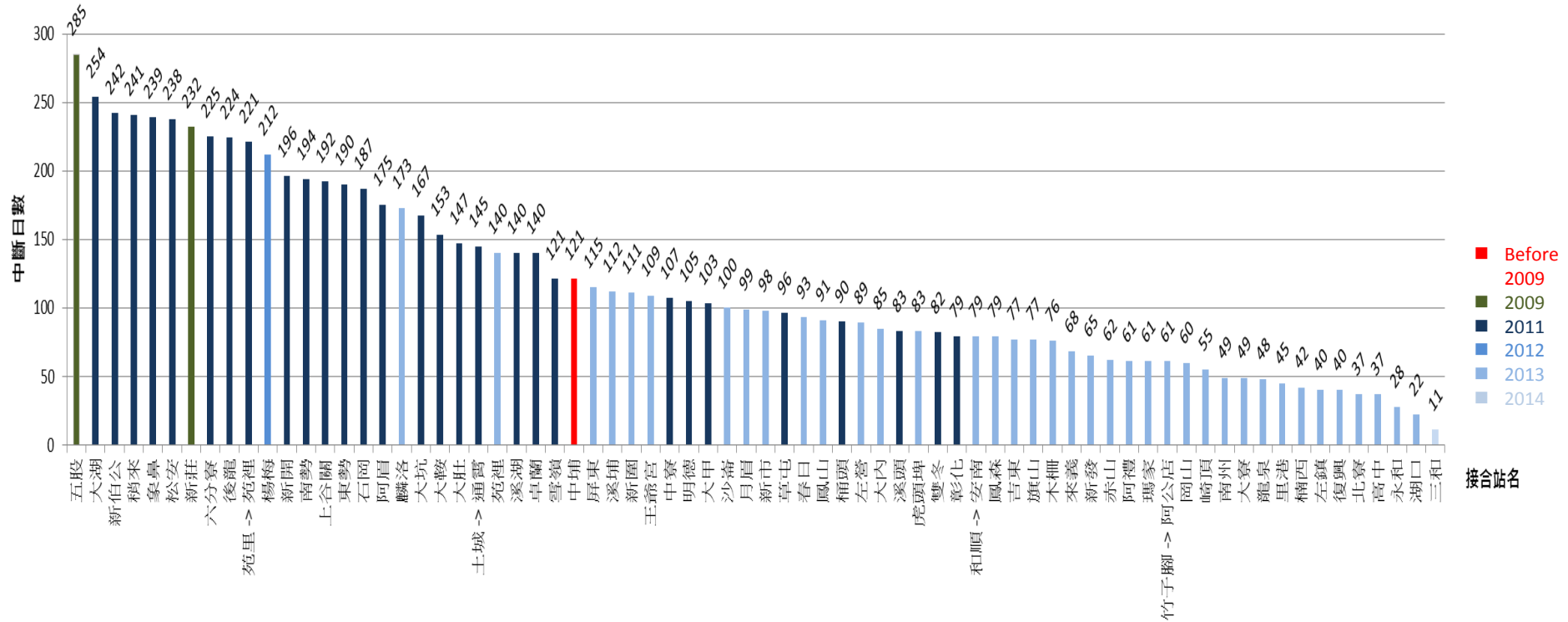
- 透過分析1998年至2016年等近19年的110個氣溫目標站及318個雨量目標站之逐時資料齊整率，發現於近期的資料齊整率與早期建立時相比，有明顯的改善情形。
- 氣溫資料及雨量資料合理性查核分析，透過以上3種查核方法與天氣系統的比對，增加對於可疑資料合理性判斷的經驗累積，對於特殊天氣系統造成的極端事件可更加準確的判斷合理性，以利後續資料品管流程的改進；並可將偵測出的錯誤資料即時修正原始資料，以增進自動觀測資料的正確性。
- 未來工作為持續透過其他方法，例如Local Outlier Factor (LOF) 的多變數分析來協助進行合理性查核，並朝能夠應用於即時資料QC為目標。

報告完畢
謝謝聆聽

附錄

雨量目標清單接合站資料實際中斷時間排序

資料中斷日數測站排序



資料中斷日數前後期差異原因

- 2011年前的汰換案為整批驗收，
- 2013年後的汰換案為分批驗收。

2015年上半年目標站齊整率(二組補整前後差異)

補整前

	SN	STO	1月	2月	3月	4月	5月	6月	1~6月總齊整率	備註
大尖山	29	C0A590	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	撤站
新屋	48	C0C450	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	撤站
紅葉山	83	C0S680	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2014/11/12之後無資料
宜梧	66	C0K291	32.3%	35.0%	37.1%	0.0%	0.0%	0.0%	17.4%	C0K290 介接站 (2015/3/31短距離移位汰換)
關子嶺	106	C0X020	98.8%	99.4%	99.2%	49.2%	0.0%	0.0%	57.8%	2015/4/16遷站，更換站碼為C0X240
虎尾	68	C0K330	99.1%	99.3%	99.3%	99.2%	19.4%	0.0%	69.4%	2015/5/7原址汰換
佳里	107	C0X080	96.6%	96.6%	96.6%	97.4%	31.5%	0.0%	69.8%	2015/5/11原址汰換
台西	58	C0G710	93.3%	78.6%	99.2%	99.7%	78.5%	0.0%	74.9%	2015/5/25遷站，更名為大城C0G740
富貴角	34	C0A920	37.5%	64.6%	80.8%	88.2%	98.1%	98.3%	77.9%	
大埔	67	C0K300	93.8%	94.0%	93.7%	92.9%	93.5%	0.0%	78.0%	2015/6/6遷站，更名為古坑C0K490
新營	73	C0O910	96.1%	96.4%	98.1%	96.3%	96.2%	9.6%	82.1%	2015/6/4遷站，更改站碼為C0X250
善化	72	C0O900	71.5%	78.1%	80.0%	93.5%	79.0%	91.1%	82.2%	
天母	42	C0A9C0	94.4%	81.3%	40.7%	94.4%	98.3%	93.9%	83.8%	
社子	38	C0A980	90.6%	94.0%	84.7%	94.4%	97.2%	80.7%	90.3%	
梅花	49	C0D360	97.7%	99.6%	98.8%	87.5%	68.1%	96.0%	91.3%	
四湖	65	C0K280	95.6%	96.6%	96.6%	95.1%	96.0%	76.8%	92.8%	2015/7/18原址汰換
內湖	45	C0A9F0	98.5%	95.8%	84.9%	84.7%	98.1%	98.1%	93.4%	
美濃	104	C0V310	93.7%	95.7%	97.6%	92.5%	93.4%	95.7%	94.8%	

補整後

善化	72	C0O900	77.3	100.0	99.9	99.7	96.0	91.1	94.0	
社子	38	C0A980	99.7	100.0	99.9	100.0	100.0	81.0	96.8	
梅花	49	C0D360	97.7	99.6	98.8	99.4	80.8	99.9	96.0	
美濃	104	C0V310	96.0	100.0	99.9	100.0	100.0	98.2	99.0	

氣候離群值分析(>5SD)：2016年7月TT目標站(110站)(1/2)

TT Climate Outlier Items : 9

1	pt	sto_id	yyyymmddhh	L_Bound	Y_org	U_Bound	Xm	Sigma	(X-Xm)/Sigma
2	43	467650	2016073121	16.2550	27.4000	27.3150	21.7850	1.1060	5.0769
3	43	467650	2016073122	16.7050	27.1000	26.7050	21.7050	1.0000	5.3950
4	43	467650	2016073123	17.2670	27.1000	26.0070	21.6370	0.8740	6.2506
5	37	467770	2016070320	24.1460	23.2000	33.7260	28.9360	0.9580	-5.9875
6	38	C0D360	2016070805	14.0400	29.6000	29.2300	21.6350	1.5190	5.2436
7	38	C0D360	2016070806	14.1590	29.9000	29.4090	21.7840	1.5250	5.3220
8	43	C0D360	2016073124	16.4400	30.9000	28.9200	22.6800	1.2480	6.5865
9	39	C0E430	2016071422	20.4060	29.8000	29.5260	24.9660	0.9120	5.3004
10	37	C0G640	2016070319	23.6250	23.4000	33.5750	28.6000	0.9950	-5.2261

1998~2016 7月Climate Outlier>5SD 筆數

1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
0	0	2	10	14	10	3	10	17	0	4	2
2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
0	2	0	42	15	30	9	-	-	-	-	-

氣候離群值分析(>5SD)：2016年7月TT目標站(110站)(2/2)

➤ 本月所找到>5SD的氣候離群值共有9筆，判斷皆為合理值。

詳細分析如下：

1. 467650日月潭(3筆)：判斷應為合理值。
 - ✓ 原因為7/31傍晚起受妮妲颱風外圍環流過山沉降增溫，夜晚出現較高氣溫值。
2. 467770梧棲(1筆)：判斷應為合理值。
 - ✓ 原因為7/3 19L~21L 連續3小時有降雨，導致迅速氣溫下降。
3. C0D360梅花(2筆)：判斷應為合理值。
 - ✓ 原因為7/8 受尼伯特颱風過山沉降增溫，夜晚出現較高氣溫值。
4. C0D360梅花(1筆)：判斷應為合理值。
 - ✓ 原因為7/31 傍晚起受妮妲颱風外圍環流過山沉降增溫，夜晚出現較高氣溫值。
5. C0E430南庄(1筆)：判斷應為合理值。
 - ✓ 原因為7/14 太平洋高壓影響，白天晴朗炎熱，晚上降溫較慢。
6. C0G640鹿港(1筆)：判斷應為合理值。
 - ✓ 原因為7/3 19L~21L 連續3小時有降雨，導致迅速氣溫下降。

空間一致性檢查：2016年7月TT目標站氣候離群值分佈圖(>4SD)

➤ QC合理性判斷：經空間一致性檢查，均認為合理。

- ✓ 於7/3臺中及彰化地區午後熱對流旺盛， $<-4SD$ 的氣候離群值有4筆，從19L~21L連續3小時降雨，導致降溫較為明顯。
- ✓ 於7/6太平洋高壓影響期間， $>4SD$ 的氣候離群值有4筆，分布區域為大臺北地區，顯示這些區域增溫較為明顯。
- ✓ 於7/7~7/9尼伯特颱風環流過山沉降增溫， $>4SD$ 的氣候離群值有17筆，分布區域為中部以北及東部地區，顯示這些區域增溫較為明顯。
- ✓ 於7/9尼伯特颱風影響期間， $<-4SD$ 的的氣候離群值有3筆(四湖)，顯示該站受降雨影響，氣溫下降明顯。
- ✓ 於7/11西南風偏強期間， $<-4SD$ 的的氣候離群值有1筆(四湖)，顯示該站受降雨影響，氣溫下降明顯。
- ✓ 於7/14太平洋高壓增強期間， $>4SD$ 的氣候離群值有1筆(三義)，顯示該站增溫較為明顯。
- ✓ 於7/31妮妲颱風外圍環流過山沉降增溫期間， $>4SD$ 的氣候離群值有8筆，分布區域為中部地區，顯示這些區域增溫較為明顯。

