

# 測站即時溫度觀測檢核流程與案例分析

陳奕翰<sup>1</sup> 高裕哲<sup>1</sup> 馮智勇<sup>2</sup> 沈里音<sup>2</sup>  
多采科技有限公司<sup>1</sup> 中央氣象局科技中心<sup>2</sup>

## 摘要

測站即時溫度觀測資料檢核往往受限於作業時間，除了依賴比較測站自身觀測時序變化是否合理外，常見藉由比較同時其他測站觀測資料是否發生相同變化特性或者是應用空間統計方法進行遮蔽測站實驗加以判讀。然而，由於空間統計方法的估計值不易反應測站局部差異產生的影響，因此需要先透過歷史資料分析掌握系統性偏差特徵後再行修正，而不便於即時檢核作業。

本研究為此結合通用克利金法(Universal Kriging)與誤差衰退平均法(Decaying average)發展即時溫度觀測資料檢核流程，將通用克利金法之溫度估計值視為一種預報值，並以誤差衰退平均法所找出的系統性偏差對此估計值進行修正，再利用此偏差修正後的估計值與估計誤差標準差，建構區間進行溫度資料之檢驗。而估計誤差包含(1)UK法的理論估計誤差、(2)估計值與觀測值的實際誤差，建構檢核區間時，將擇兩種估計誤差標準差中數值較大者。

關鍵字：Decaying average、Universal Kriging、QC、溫度檢核

## 一、前言

正確的觀測記錄，可供許多學術或研究單位進行後續分析，但不正確之觀測記錄可能造成錯誤的推論或作出不可靠的分析結果，故發展有效的資料檢核方法是相當重要的議題。

目前臺灣包含綜觀氣象站、合作氣象站以及地面自動氣象站約有500個觀測站點，但臺灣地形多變且複雜，溫度觀測會因地形、地貌、地點、風場通風性、太陽是否容易直射的不同而有其獨特的變化特性；本研究針對測站溫度的檢核流程中，首先將各個測站溫度以通用克利金法(Universal Kriging，簡稱UK)估計出該測站之溫度估計值，並建立溫度的檢核區間進行判斷，如超出檢核區間則提列為錯誤觀測值，詳細的檢核流程將於本文第二段進行說明，本文第三段探討此檢核方法之成效，最後則為研究總結與討論。

## 二、研究方法

溫度檢核流程中，將通用克利金法(Universal Kriging，簡稱UK)之溫度估計值視為一種預報值，並以誤差衰退平均法所找出的系統性偏差對此估計值進行修正，誤差衰退平均法為美國國家環境預報中心(National Centers for Environmental Prediction，NCEP)發展的方法，原理係以卡門濾波(Kalman Filter)為基礎，無須在作業主機中保有大量歷史資料即可逐步找出預報值與觀測值的系統性偏差，具有簡便且高運作效率的特性。再利用此偏差修正後的估計

值與估計誤差標準差，建構檢核區間進行溫度資料之檢驗。以下將對通用克利金法與誤差衰退平均法進行說明：

(一) 通用克利金(Universal Kriging):

測站溫度估計值  $f_0$  是利用通用克利金法進行估算，原理為根據觀測資料進行的結構分析、最佳線性不偏估計(Best Linear Unbiased Estimation，BLUE)，以及103年度氣候資料整集分析系統計畫成果可知臺灣測站溫度觀測值具有隨高度( $h$ )與緯度( $y$ )變化的線性趨勢等原則，再利用拉格朗日乘子(Lagrange multiplier)  $v_1, v_2, v_3$ ，結合3個不偏估條件式，得下列  $n+3$  組聯立方程式

$$\begin{cases} \sum_{j=1}^n \lambda_j \cdot cov(d_{ij}) + v_1 + v_2 y_i + v_3 h_i = cov(d_{i0}) & \text{for } i=1, 2, \dots, n \\ \sum_{i=1}^n \lambda_i = 1 & \sum_{i=1}^n \lambda_i h_i = h_0 & \sum_{i=1}^n \lambda_i y_i = y_0 \end{cases} \quad (1)$$

解  $n+3$  個未知數  $\lambda_1, \lambda_2, \dots, \lambda_n, v_1, v_2, v_3$ ，其中  $d_{ij}$  為測站  $i, j$  之距離； $h_i, y_i$  分別為測站  $i$  之高度與緯度； $\lambda_i$  為測站  $i$  之權重係數，則測站溫度估計值  $f_0$  為其餘測站觀測值之權重平均

$$f_0 = \sum_{i=1}^n \lambda_i a_i \quad (2)$$

而估計誤差可以下式計算：

$$sd = \sqrt{\sigma_z^2 - v_1 - v_2 y_0 - v_3 h_0 - \sum_{i=1}^n \lambda_i \cdot cov(d_{i0})} \quad (3)$$

(二) 誤差衰退平均法(Decaying average)

因溫度呈現季節性的變化，故將分為不同月份不同小時，各自進行系統性偏差之估計，而以誤差衰退平均法進行系統性偏差之估算流程可分為以下幾個步驟：

(1) 誤差估計:

以  $b^\tau(t)$  代表於  $t$  時間(例如20180704 00:00)

可估計之  $\tau$  時間(在此表示檢核當下的時刻)預報場誤差值，採下式進行計算。其中， $f^\tau(t')$  為  $t'$  時間(例如20180704 00:00)預報  $\tau$  時間的預報值，而  $a(t'+\tau)$  為  $t'+\tau$  時間已知真實值。

$$b^\tau(t) = f^\tau(t') - a(t'+\tau) \quad (4)$$

(2) 更新系統性偏差值:

以  $B^\tau(t)$  代表於  $t$  時間可使用之  $\tau$  時間(在此表示檢核當下的時刻)預報場系統偏差值，採下式進行計算：

$$B^\tau(t) = (1-w) \cdot B^\tau(t^*) + w \cdot b^\tau(t) \quad (5)$$

其中， $B^\tau(t^*)$  為最近一次更新的系統性偏差值，以上述的  $t$  時間為例，假設逐10分鐘檢核，則  $B^\tau(t^*)$  為20180703 00:00的結果；而權重係數  $w$  代表採用近期幾次誤差資訊，權重係數0.05約代表利用最近40-60次(天)(權重係數倒數的二至三倍)之估計值與真實場所計算的系統性偏差。

(3) 以系統偏差值更新值修正估計值:

將  $t$  時間的  $\tau$  時間預報場預報減去系統偏差值  $B^\tau(t)$ ，即可得到估計校正值  $F^\tau(t)$ ，即：

$$F^\tau(t) = f^\tau(t) - B^\tau(t) \quad (6)$$

除利用衰退平均法動態計算估計值的系統性偏差外，也可將此方法應用在計算估計值的估計誤差標準差上，假設某測站之觀測值序列為  $O_1, O_2, \dots, O_t$ ，而對應之估計值為  $E_1, E_2, \dots, E_t$ ，令估計誤差為  $D_1, D_2, \dots, D_t$ ，其中， $D_i = E_i - O_i$ ， $i = 1, 2, \dots, t$ ，則估計誤差的平均值  $\bar{D}_t$  以下式動態計算

$$\bar{D}_t = \bar{D}_{t-1} \cdot (1-\omega) - D_t \cdot \omega \quad (7)$$

估計誤差標準差  $STD_{Est-Obs,t}$  以下式動態計算

$$STD_{Est-Obs,t} = STD_{Est-Obs,t-1} \cdot (1-\omega) - S_t \cdot \omega \quad (8)$$

其中， $S_t = (D_t - \bar{D}_t)^2$ ； $\omega = 0.05$ 。

(三) 空間檢核流程

利用通用克利金法與誤差衰退平均法找出溫度估計值，再以估計誤差標準差作為估計值的可能誤差範圍建構一檢核區間，若溫度觀測值落在此檢核

區間外，則視為潛在異常值，如圖1，檢核流程為以下幾個步驟：

- (1) 利用UK法找出溫度估計值  $f_0(t)$  與估計誤差標準差  $STD_{UK}$ ，如式2與式3；
- (2) 利用衰退平均法動態計算前一時間修正後估計值與觀測值的估計誤差標準差  $STD_{Est-Obs}$ ，如式8；
- (3) 利用衰退平均法動態計算前一時間估計值的系統性偏差值  $B_0(t)$  (平均偏差值及偏差估計值的權重平均)，如式5；
- (4) 利用  $B_0(t)$  修正溫度估計值得  $F_0(t)$ ，如式6；
- (5) 利用  $F_0(t)$  加減3.5倍估計誤差標準差，建立檢核區間  $[F_0 - 3.5 \cdot STD, F_0 + 3.5 \cdot STD]$  (10)
- (6) 將落在檢核區間外的溫度觀測值標記為潛在異常值，否則視為合理的觀測值，若為第二次進入步驟(5)，則結束檢核流程，否則進入步驟(6)
- (7) 若無溫度觀測值被標記為潛在異常值，則結束檢核流程，並進入步驟(7)；
- (8) 將步驟(5)檢核並判斷為合理的觀測值作為新的觀測資料集，針對步驟(5)檢核並標記為潛在異常值的觀測值重新進行檢核，即回到步驟(1)

檢核步驟(4)的估計誤差，評估方式包含UK法的理論估計誤差、估計值與觀測值的實際誤差，建構檢核區間時，將擇兩種估計誤差標準差中較大者，但當建構檢核區間的估計誤差標準差，乘以3.5倍率後，仍小於1時，為避免檢核結果過於敏感，將直接設定估計值可能的誤差範圍為1，故檢核步驟(4)的估計誤差標準差等於

$$STD = \text{Max}(STD_{Kriging}, STD_{Est-Obs}) \quad (11)$$

$$STD = \text{Max}(STD, 1) \quad (12)$$

為避免異常的觀測值影響UK法的估計結果，溫度觀測值經檢核後，將初步分為正常或潛在異常值，此時，針對潛在異常值將以正常觀測值利用UK法重新進行估計並建構檢核區間進行檢核，經再次檢核的潛在異常值仍被判定為不合理，才進行潛在異常值的提列，否則將被調整為正常觀測值。

以修正後UK估計值加減其3.5倍估計誤差標準差所建構之區間進行資料之檢核，為預期該修正後估計值具有一定程度的參考性，殘差值序列在某範圍內震盪且期望值接近0，因此，若觀測紀錄不落在該區間內，則有很大可能為異常值，其概念如一標準

常態分布序列，約99%數值散佈在距離平均值3個標準差之內的範圍。

#### (四) 輔助研判機制:

檢核中以空間統計方法進行溫度值的估計，但在觀測資料不齊全、估計目標處在離島或高山地區等情況下，溫度估計值的可信度將下降，為避免誤判狀況發生，檢核流程中加入下列輔助研判機制:

##### (1) 參考測站時間序列:

若溫度觀測值一小時內的變化小於0.5度，將視為合理的觀測值。

##### (2) 雨量觀測資料輔助研判:

當零星降雨時，可能發生某測站降雨而鄰近測站卻無降雨的狀況，此時該站的估計值可能因參考了鄰近測站的資訊而高估，故當測站發生降雨，該站觀測值落於檢核區間外但小於4倍估計誤差標準差，將調整為合理觀測值。

### 三、結果分析

本研究以綜觀氣象站、合作氣象站以及地面自動氣象站約有500個觀測站點，即時檢核2018年2月1日到現在的逐10分鐘溫度觀測資料進行資料品質檢核，以下將列舉數個案分析檢核方法的能力與難處:

A2K570為古坑合作站，圖2顯示古坑合作站於白天升溫時以及夜晚降溫時容易因速率較鄰近站快(局屬古坑站C0K490)，因此檢核系統將正確的觀測值提列為嫌疑犯，造成此原因為，兩站雖為鄰近站，但是擺放位置有明顯地差異，古坑合作站擺放在地面，而古坑局屬站則是擺放於屋頂，因此，環境上有根本的差異造成升溫降溫的速率不同。

圖3顯示，466910(鞍部測站)於2018年6月11日23點有夜晚升溫的現象，因此，466930站(竹子湖站)的克利金估計值受到466910站的影響出現溫度上升的趨勢，導致466930站正確的觀測值落在檢核區間之外提列為嫌疑犯。

B2Q810(恆春畜產測站)於2018年6月23日12點20分被提列為錯誤觀測值，從圖X也可以很明顯看到該測站後面的溫度觀測值也是明顯小於估計值，溫度觀測值很明顯地異常，但是卻未被檢核系統提列為錯誤觀測值，從圖4(b) (c) (d) (e)可以看到B2Q810的鄰近站:C0R350(貓鼻頭)、C0R360(恆春)、C0R370(佳樂水)、C0R620(墾丁雷達)於紅框的時間點，皆無紀錄到觀測資料，導致克利金估計估計值時只能拿相對較B2Q810遠的467590(圖(f)恆春站)來計算估計值，造成克利金的估計誤差標準差出現異常地放大而將一些錯誤的觀測值包進檢核區間內通過檢核。

### 四、總結與討論

目前的溫度檢核流程中使用通用克利金法搭配誤差衰退平均法以及輔助研判機制已經能夠有效且正確地檢核大部分的溫度觀測資料，但在日間升溫以及夜間降溫時，檢核系統容易出現誤判，未來將自身的溫度變化速率加入檢核系統中，期望能降低誤判頻率。

而在夜間升溫的部分，從每天的即時觀測資料發現這類的溫度變化現象相當地頻繁出現，但又沒有一定的發生頻率，目前對這類的現象只能持續地觀察追蹤，並思考解決的方式。

### 五、參考文獻

1. 交通部中央氣象局，2014:「氣候變遷應用服務能力發展計畫103年氣候資料整及分析系統發展」。
2. 交通部中央氣象局，2015:「氣候變遷應用服務能力發展計畫104年氣候資料整及分析系統發展」。
3. 交通部中央氣象局，2016:「氣候變遷應用服務能力發展計畫105年氣候資料整及分析系統發展」。
4. 交通部中央氣象局，2017:「氣候變遷應用服務能力發展計畫106年氣候資料整及分析系統發展」。
5. 高裕哲, 馮智勇, 詹智雄, 沈里音 2016: 應用Decaying Average發展測站即時溫度觀測檢核技術,105年天氣分析與預報研討會。

## 六、附圖

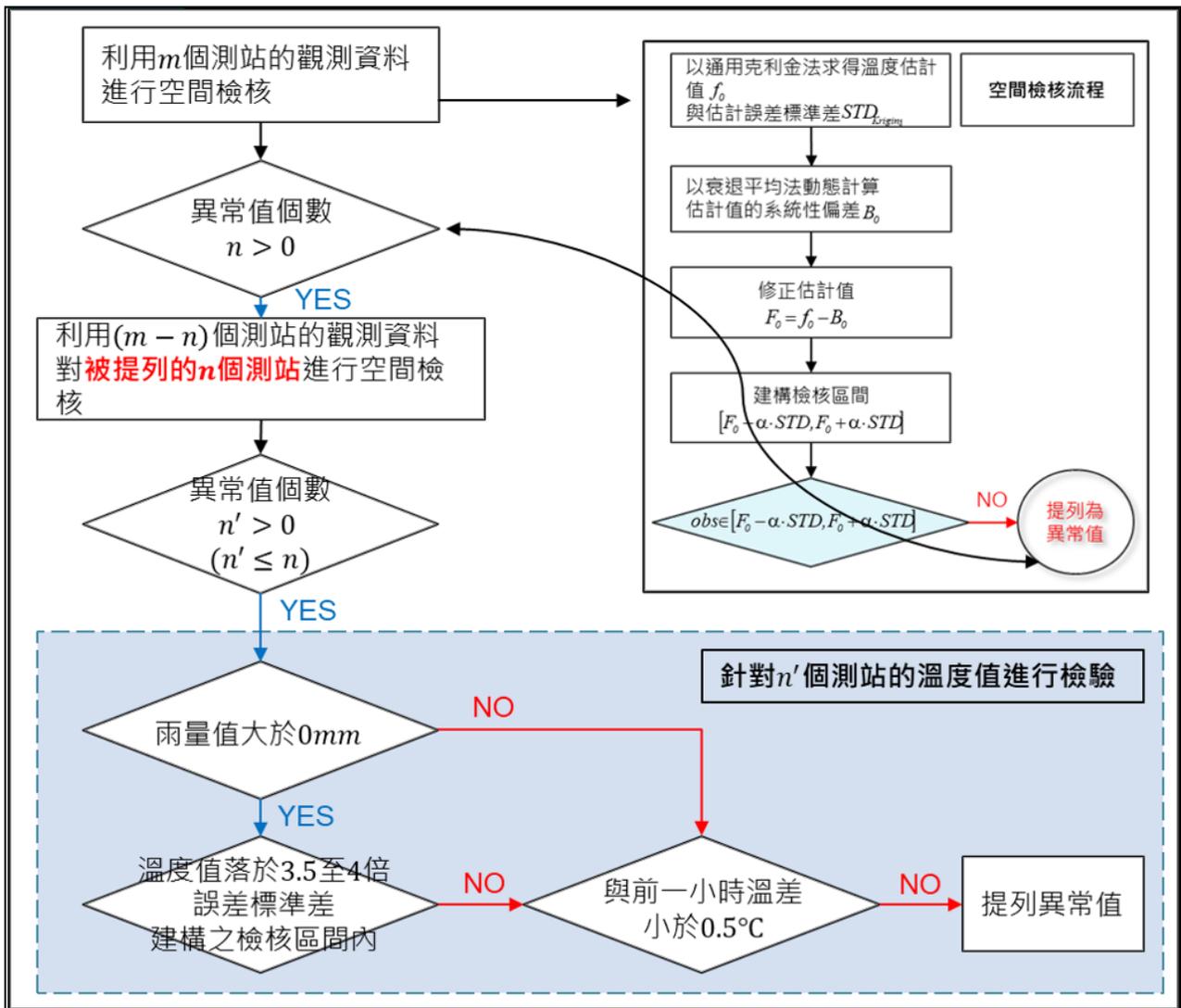


圖 1 溫度檢核流程圖

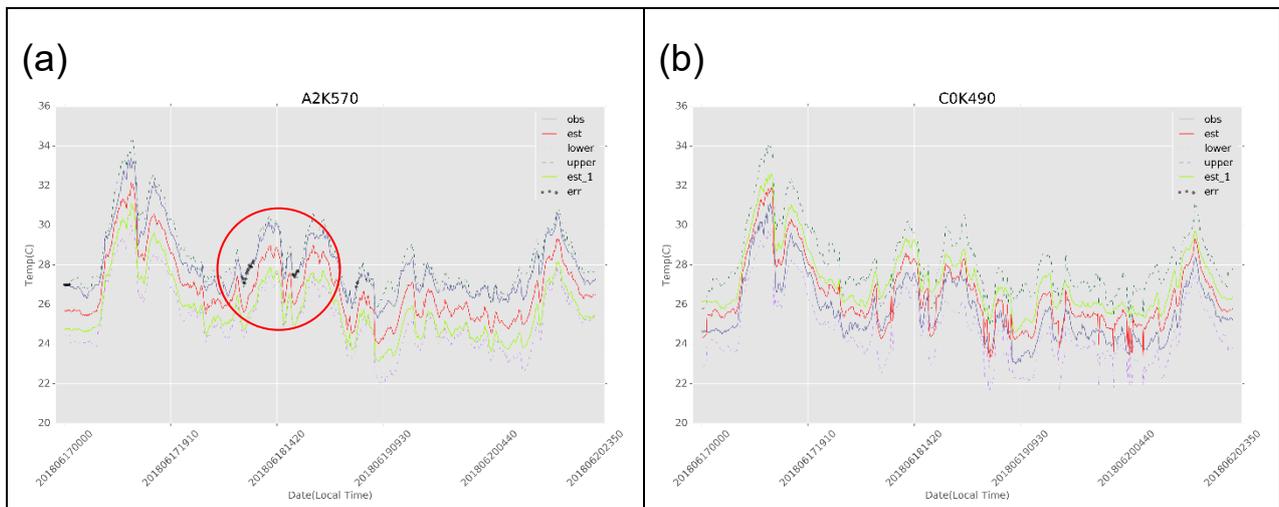


圖 2 (a)古坑合作站(A2K570)及(b)古坑局屬站(C0K490)時間序列，黑點代表資料檢核出的錯誤資料，藍線代表觀測值，紅線代表克利金估計值

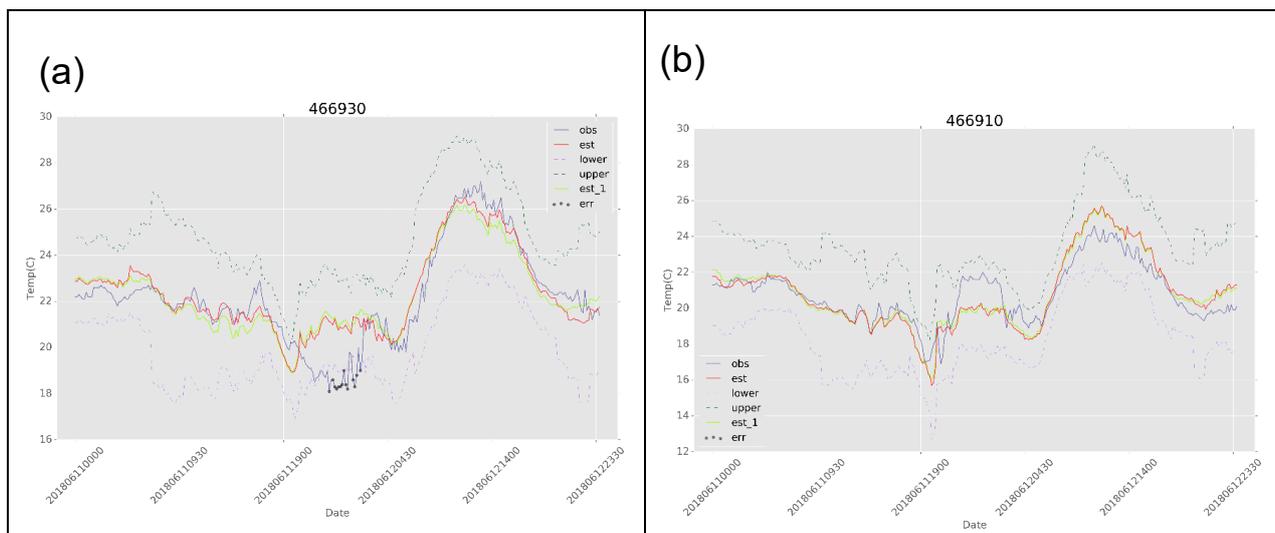


圖 3 (a)竹子湖站(466930)及(b)鞍部站(466910)時間序列，黑點代表資料檢核出的錯誤資料，藍線代表觀測值，紅線代表克利金估計值

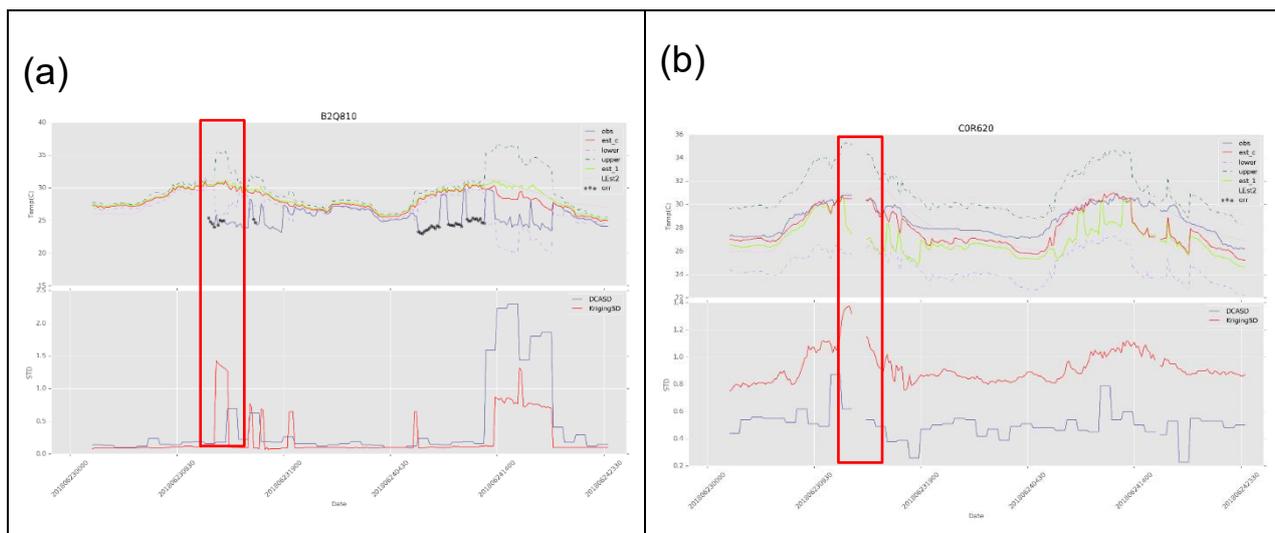


圖 4 (a)恆春畜產站(b)墾丁雷達站(c)墾丁站(d)佳樂水站(e)貓鼻頭站(f)恆春站時間序列圖

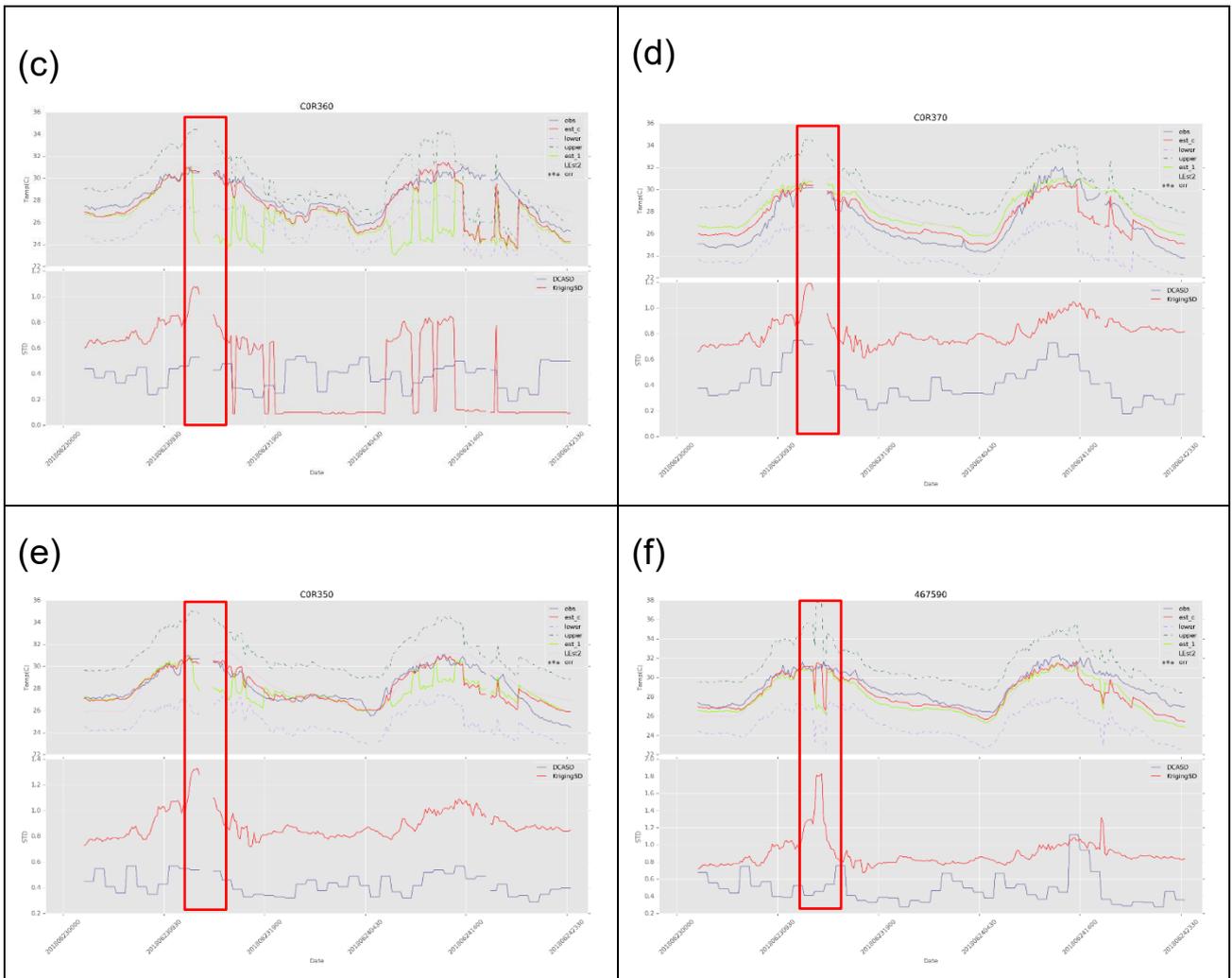


圖 4 時間序列圖(續)