



網格化觀測溫度 V2 版資料 資料生產履歷



2024 年 1 月 12 日

臺灣氣候變遷推估資訊與調適知識平台

Taiwan Climate Change Projection Information and Adaptation Knowledge Platform

網格化觀測溫度 V2 版資料生產履歷

1. 資料紀錄

2019.10(5km 資料上架)

2020.10([更新至 2018 年](#))

2021.03([更新至 2019 年](#))

2021.09(1km 資料上架)

2022.03([更新至 2020 年](#))

2023.01([更新至 2021 年，網格經緯度改版為 0.01°和 0.05°](#))

2024.01([更新至 2022 年](#))

2. 產製目的

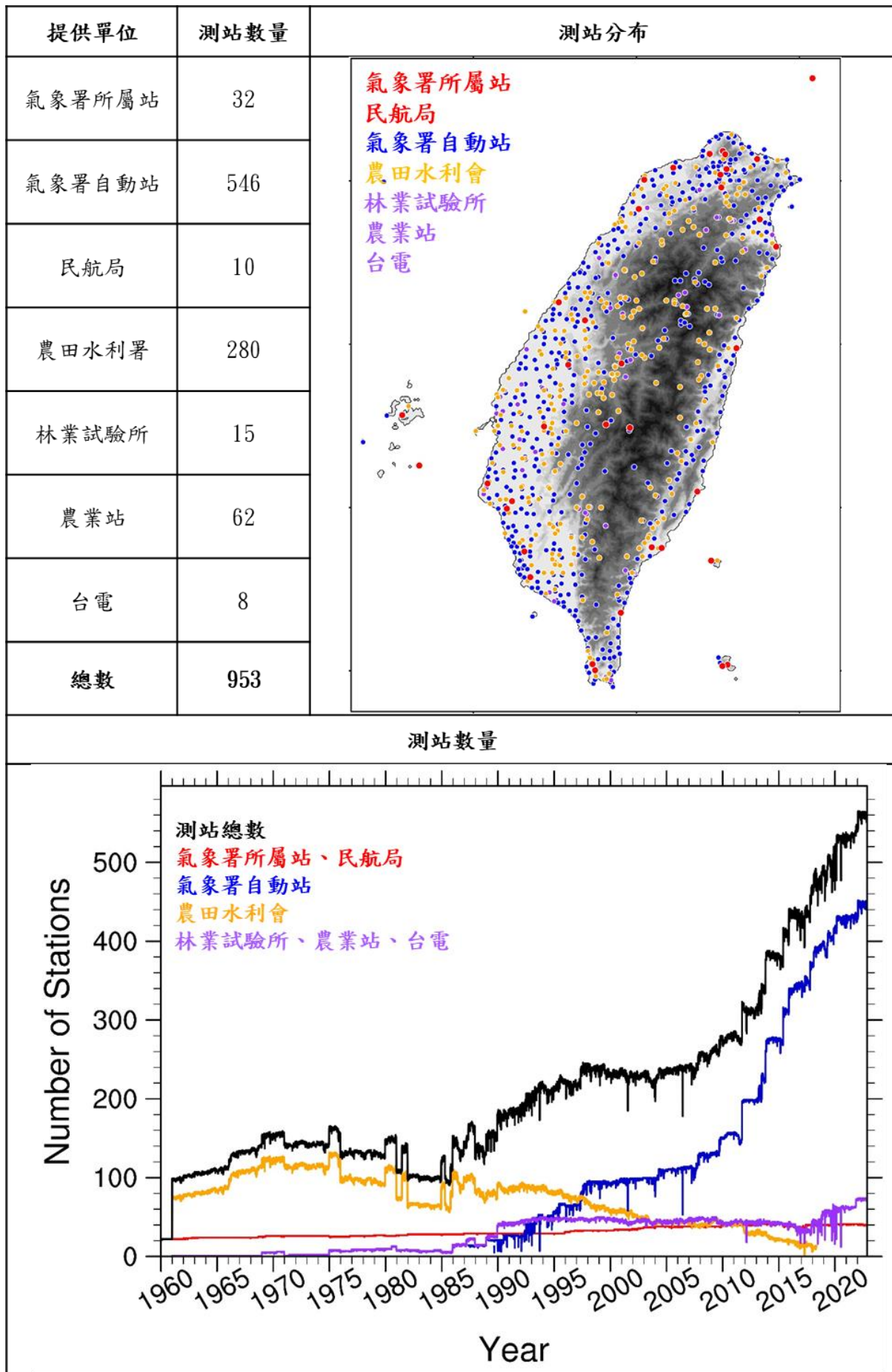
為因應全球氣候變遷下威脅日趨加劇的區域極端天候帶來的社經衝擊，國科會政策推動的臺灣氣候變遷推估資訊與調適知識平台計畫(Taiwan Climate Change Projection Information and Adaptation Knowledge Platform, TCCIP)的任務之一就是蒐集整合來自交通部氣象署、經濟部水利署、行政院環保署、行政院農業委員會農業試驗所和林業試驗所等單位的測站資料，將測站點的資料面化為網格資料，建立臺灣的氣候變遷觀測資料庫。建置長期且高時空解析度的歷史網格化觀測資料，除了可偵測氣候變遷訊號並提出預警，強化診斷天候事件之時空演化的能力、亦可應用於氣候模式調校模式偏差、訓練統計降尺度模型，以及量化災害的衝擊評估，更能進一步提供不同研究領域，例如水文災害、公共衛生、植物生態等相關研究的需求。

3. 資料來源

■ 測站資料

使用氣象署、農田水利署、林試所、農試所、民航局及台電等多個單位所提供之測站資料，總時間長度為 1960~2022 年，並且持續逐年蒐集、更新資料，由於來自不同單位的測站時空分布有顯著的變化，測站空間分布及測站數量逐年變化呈現如表 1。

表 1、測站資料來源



4. 產製流程

■ 圖示:

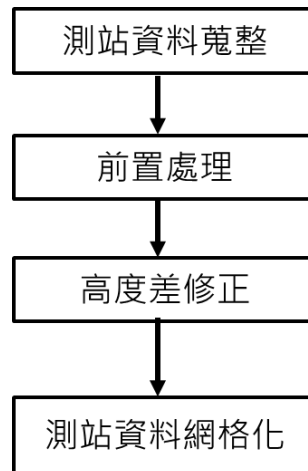


圖 1、溫度日資料產製流程

■ 文字敘述:

步驟一、

測站資料網格化前，必須先進行資料前處理，包含測站資料檢核與篩選(如附件一)，確認測站三組溫度(平均溫、最高溫、最低溫)皆有資料，才能當作參考測站用來估計網格點的溫度資料，若有缺少任何一組溫度資料即視為無法使用，不另外進行測站資料補遺。

步驟二、

因目標網格點和參考測站間海拔高度有所差異，必須考慮高度差影響造成的溫度差異，在資料網格化前須進行測站溫度修正。依據所有測站在不同日期(365 天)，取每一天包含前後各 15 天，共 31 天窗區的溫度資料的氣候平均值，利用溫度氣候值與測站高度進行迴歸分析，計算每一天的溫度遞減率經驗值，即平均溫、最高溫、最低溫三組溫度資料，在不同日期，皆有各自的溫度遞減率經驗值，並適用於所有測站。最後依據參考測站與網格點的高度差異，代入溫度遞減率公式進行測站溫度修正。

步驟三、

資料網格化方法：以資料補遺方法為基礎，將網格點當作資料缺失，直接用網格點附近參考測站(權重最大的五個測站)的溫度資料，估計出網格點上的溫度值(如附件一)。

步驟四、

產製輸出為 0.01° 解析度的網格日資料，再以區域平均的方式轉換為 0.05° ，即 0.05° 資料是由網格面積內 25 個 0.01° 網格計算平均值，月資料則是由日資料進行轉換。

5. 資料不確定性

■ Leave-one-out 交叉驗證：

交叉驗證(Cross-Validation)為統計學上，用來驗證「統計模型」好壞的方法。在本研究中即用來評估溫度網格化方法的不確定性，其方法為” leave-one-out”。本研究欲評估溫度網格化方法在 1 月份的不確定性，先令 1960 年 1 月特定測站的可用觀測缺失，利用附近參考測站相對應日期的溫度資料，估計出該特定測站 1960 年 1 月的溫度缺失值，依此類推到其它年、月的資料，再重覆逐一執行到其它測站，得到一組推估的全部測站日溫度序列資料，即可評估其與實際觀測值間的誤差量。

透過交叉驗證(cross validation)實驗，評估產製網格化溫度日資料過程中的不確定性量化分析。

- 參考測站的溫度資料標準差：圖 2 是產製網格化溫度資料庫時，用來估計目標網格點溫度的五個參考測站溫度資料之間的標準差平均值。相較於平地區域而言，在山區的標準差明顯較大，特別是中部山區，代表該區域的目標網格點使用的參考測站之間資料性質差異較大(由於測站稀少，參考測站之間的溫度及高度差異大)，同樣是山區，但在雪山山脈區域和高雄、屏東南部山區的參考測站之間標準差相對較小，代表測站之間資料性質較一致。

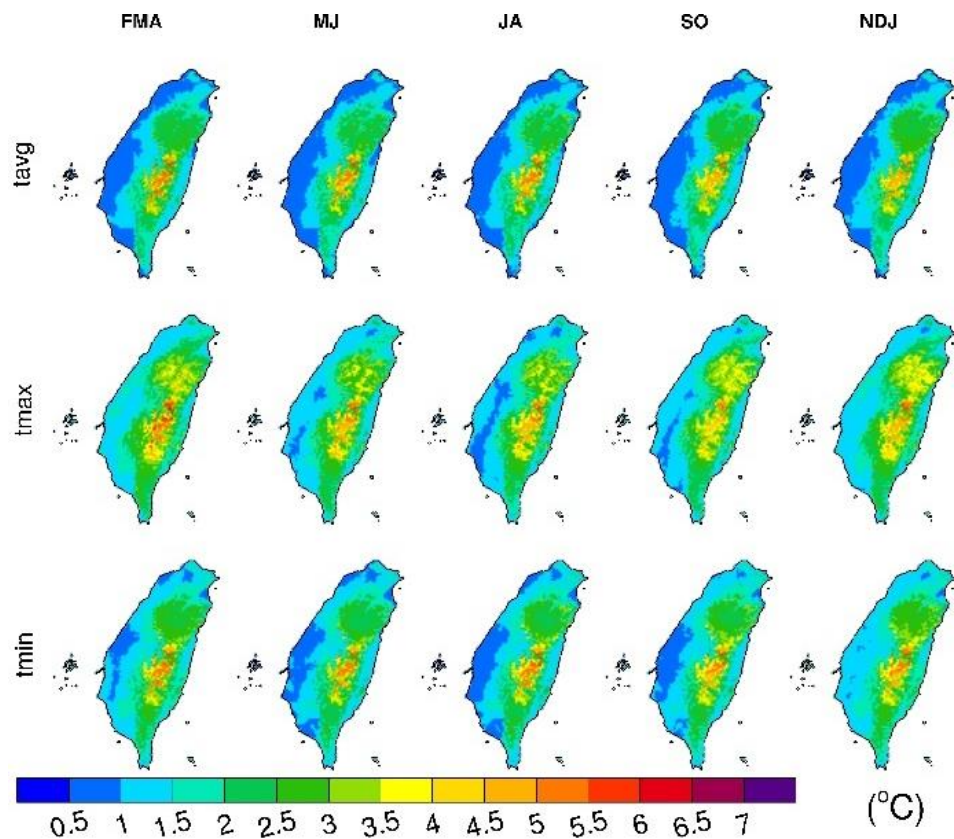


圖 2、估計目標網格點溫度的參考測站溫度資料，在不同季節的標準差平均值(單位: °C)

- 參考測站的平均距離：由參考測站相距目標網格點的平均距離(圖 3)，分析發現在山區的參考測站平均距離比平地大，特別是花東縱谷附近區域與兩側中央山脈、海岸山脈交界處，因測站密度較低或完整可用的溫度資料較少，參考測站距離目標網格點的平均距離可達 50km 左右。

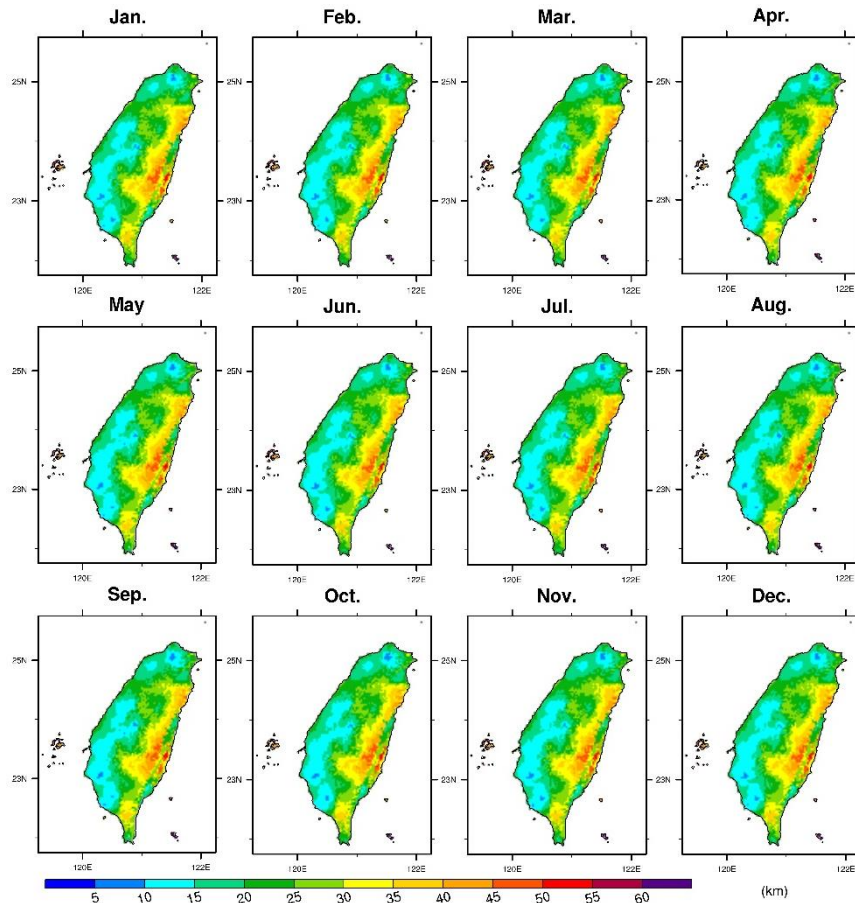


圖 3、參考測站相距目標格點的平均距離(單位:公里)

- 資料網格化方法的不確定性：圖 4 是三組溫度資料經過交叉驗證實驗後的測站估計值與實際觀測值的均方根誤差(root mean square error, RMSE)，內插至網格點上的結果。平地的誤差大都在 1°C 以下，山區誤差較大，最大可達 4°C 左右，特別是中央山脈部分區域，誤差超過 4°C ，代表測站豐度愈低，資料不確定性愈高，與圖 1、圖 2 中參考測站的溫度標準差及平均距離較大的區域吻合。

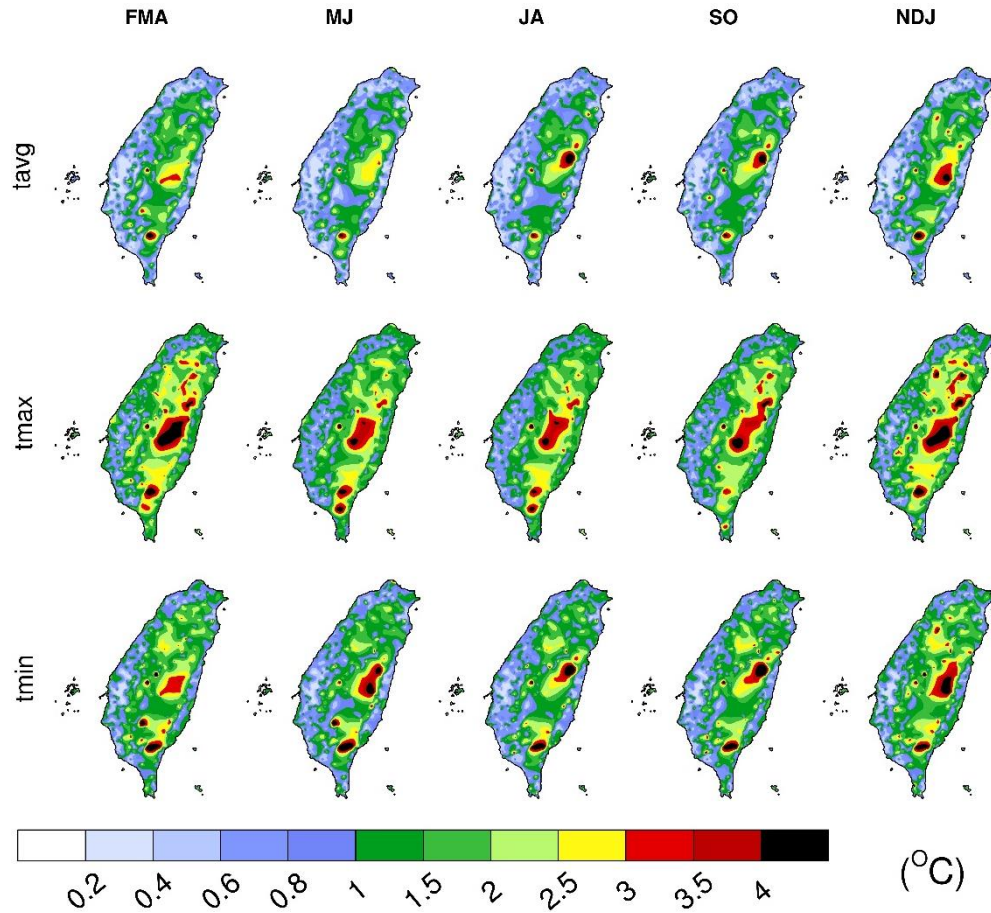


圖 4、透過交叉驗證實驗計算測站估計值與實際觀測值的均方根誤差。(單位: $^{\circ}\text{C}$)

6. 參考文獻

- Shepard, D. (1968). A Two-Dimensional Interpolation Function for Irregularly-Spaced Data. Proceedings of the 1968 ACM National Conference, New York, 27-29 August 1968, 517-524.
<http://dx.doi.org/10.1145/800186.810616>
- Shepard, D. S. (1984). Computer Mapping: The SYMAP Interpolation Algorithm, 133–145 pp., Springer, New York.
- Simolo C., Brunetti M., Maugeri M., Nanni T. (2010). Improving estimation of missing values in daily precipitation series by a probability density functionpreserving approach. International Journal of Climatology, 30, 1564-1576.
- Willmott C. J., Rowe C. M. & Philpot W. D. (1985). Small-Scale Climate Maps: A Sensitivity Analysis of Some Common Assumptions Associated with Grid-Point Interpolation and Contouring, The American Cartographer, 12:1, 5-16, DOI: 10.1559/152304085783914686

7. 發表文章

- 翁叔平、楊承道 (2012)。台灣地區月降雨及溫度 1 公里網格資料庫之建立(1960-2009)及其在近未來(2015-2039)的氣候推估應用，大氣科學，第 40 期，349-370。
- 翁叔平、楊承道 (2018)。臺灣地區日降雨網格化資料庫(1960~2015)之建置與驗證，台灣水利，第 66 卷(第 4 期)，33-52。
- 楊承道、翁叔平 (2022)。臺灣網格化觀測日溫度資料之驗證與分析，臺灣氣候變遷推估資訊與調適知識平台電子報，第 55 期。
https://tccip.ncdr.nat.gov.tw/km_newsletter_one.aspx?nid=20220124154325

8. 文件引用

- 楊承道，王俊寓，林士堯 (民 113 年 1 月 12 日)。網格化觀測溫度 V2 版資料生產履歷(4.2.1 版)。[擷取日期]，取自臺灣氣候變遷推估資訊與調適知識平台：
https://tccip.ncdr.nat.gov.tw/upload/data_profile/20220706114121.pdf

附件一、篩選和補遺

■ 除了氣象署自動站資料來源為 24 小時資料，統計得到日平均溫、最高溫與最低溫，其他測站資料來源皆使用日資料(民航局測站資料在 2017 年以前亦為時資料)。

■ 測站資料篩選流程：

相較與雨量測站數量，溫度測站數量較少，像是中央氣象署的自動站，有觀測雨量的測站有705站，但有觀測溫度的測站只有487站。同樣的情況也發生在其它單位的測站，如農業站和台電測站。加上中央氣象署的所屬測站、民航局測站、農田水利署測站和林業試驗所測站，合計約884站。

■ 測站資料網格化方法：

利用緯度、經度、高度、彼此間之夾角等測站地理資訊，估算距離目標測站一定範圍內，若干參考測站的距離、高度、方位角等三個權重參數，對目標測站的缺失資料進行補遺 (Shepard1968, 1984; Willmott et al.1985; Simolo et al.2010)。

計算周圍測站權重的方式：

a、決定參考測站。先計算目標測站與其它測站的距離，並訂下在不同時間點下，目標測站之搜尋半徑內至少要有 10 個測站有資料(即同一目標測站在不同時間點會有不同的搜尋半徑)，這些搜尋半徑內的測站稱為目標測站周圍的參考測站。

b、計算距離、高度、方位角等三個權重參數。先利用公式(a~c)算出各個參考測站的三個權重參數。

$$w_i^d(x, y) = \exp \frac{d_i^2(x, y)}{c_d} \quad (a)$$

其中 $c_d = \frac{\overline{d^2}}{\ln 2}$ 。 $w_i^d(x, y)$ 代表參考測站 i 的距離權重參數； $d_i(x, y)$ 為參考測站 i 與目標測站的距離； c_d 為常數(隨搜尋半徑 d 不同而不同)； $\overline{d^2}$ 為目標測站在不同時間點的搜尋半徑開平方。式(a)的物理意義為：參考測站離目標測站的距離愈遠，距離權重參數就愈小。

$$w_i^h(x, y) = \exp \frac{\Delta h_i^2(x, y)}{c_h} \quad (b)$$

其中 $c_h = \frac{\overline{h^2}}{\ln 2}$ 。 $w_i^h(x, y)$ 代表參考測站 i 的高度權重參數； $\Delta h_i^2(x, y)$ 為參考測站 i 與目標測站的高度差(km)； c_h 為常數； $\overline{h^2}$ 為常數 0.35(km)開平方(cf.Simolo et al.2010)。式(b)的物理意義為：參考測站離目標測站的高度差愈大，高度權重參數就小。

$$w_i^{ang}(x, y) = 1 + \frac{\sum_{j \neq i} w_j^d(x, y) w_j^h(x, y) (1 - \cos \theta_{(x, y)}(j, i))}{\sum_{j \neq i} w_j^d(x, y) w_j^h(x, y)} \quad (c)$$

式(c)內的 i 和 j 皆代表參考測站； $w_i^{ang}(x, y)$ 是參考測站 i 相對於目標測站的方位角權重參數；而 $\sum_{j \neq i} w_j^d(x, y) w_j^h(x, y) (1 - \cos \theta_{(x, y)}(j, i))$ 代表除參考測站 i 以外，其餘參考測站 j 的距離權重參數、高度權重參數以及參考測站 j 與參考測站 i 之間夾角的乘積去做累加。至於 $\sum_{j \neq i} w_j^d(x, y) w_j^h(x, y)$ 則表示除參考測站 i 以外，其餘參考測站 j 的距離權重參數和高度權重參數的乘積做累加。式(c)的物理意義為：參考測站間的方位角愈小，代表性質越相近，方位角權重參數就愈小。

c、利用式(d)獲得參考測站的乘積權重。

$$w_i(x, y) = w_i^d(x, y) w_i^h(x, y) w_i^{ang}(x, y) \quad (d)$$

d、從這 10 幾站的參考測站中，挑出乘積權重最大的前 5 站，再從這五站中，按其所占比率，重新分配權重。

e、有了五個參考測站的乘積權重後，個別參考測站的溫度值，經過高度差修正後，先乘以該站的乘積權重，而後累加之。累加後的總值即為目標網格點的估計值。

■ 進行高度差修正時，溫度遞減率的計算方法：

在一年 365 天中，依據各測站不同日期，取前、後 15 天，共 31 天窗區的溫度資料平均的氣候平均值與測站高度做回歸分析計算出溫度遞減率經驗值。使得資料特性不盡相同的三組溫度資料在不同日期，有不同的溫度遞減率經驗值，各有 365 組溫度遞減率經驗值。在溫度資料網格化過程中，使用三組溫度資料各自的溫度遞減率經驗值，對目標網格點和參考測站間的高度差做溫度修正，當參考測站的溫度資料經高度修正後，出現溫度資料異常時，例如均溫大於最高溫等情況，則排除該參考測站，最後從中選取權重最大的 5 個測站的溫度資料，進行目標網格點的溫度估計。