



網格化觀測雨量資料 資料生產履歷



2022 年 7 月 25 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

網格化觀測雨量資料生產履歷

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)

2. 產製目的

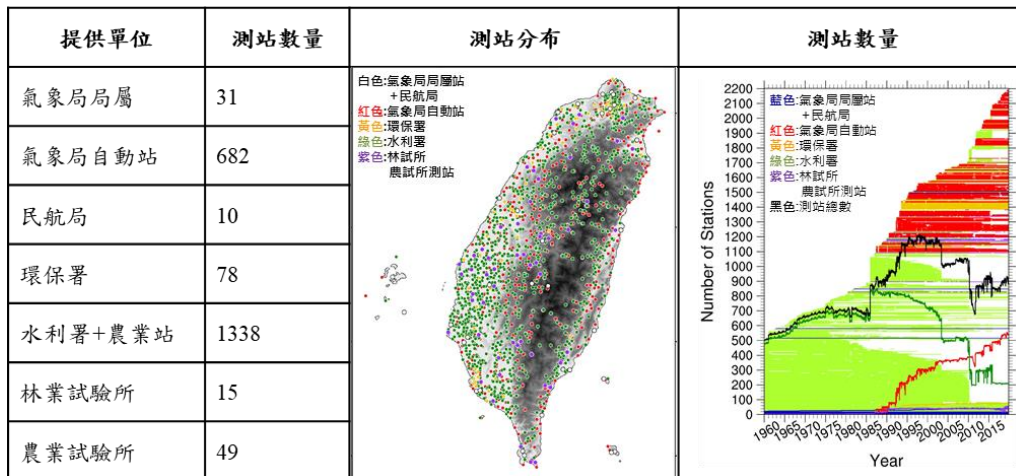
為因應全球氣候變遷下威脅日趨加劇的區域極端天候帶來的社經衝擊，科技部政策推動的臺灣氣候變遷推估資訊與調適知識平台計畫(Taiwan Climate Change Projection Information and Adaptation Knowledge Platform; TCCIP)的任務之一就是蒐集整合散佈在交通部氣象局、經濟部水利署、行政院環保署、行政院農業委員會農業試驗所和林業試驗所等單位的測站資料，將測站點的資料面化為網格資料，建立臺灣的氣候變遷資料庫。建置長期且高時空解析度的歷史觀測網格資料，除了可偵測氣候變遷訊號並提出預警，強化診斷天候事件之時空演化的能力、亦可應用於區域氣候模式(RCM)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

使用氣象局、水利署、民航局、環保署及林試所等多個單位所提供之測站資料，總時間長度為 1960~2019 年，測站分布及逐年數量呈現如表 1。

表 1、資料來源列表



4. 產製流程

■ 圖示：

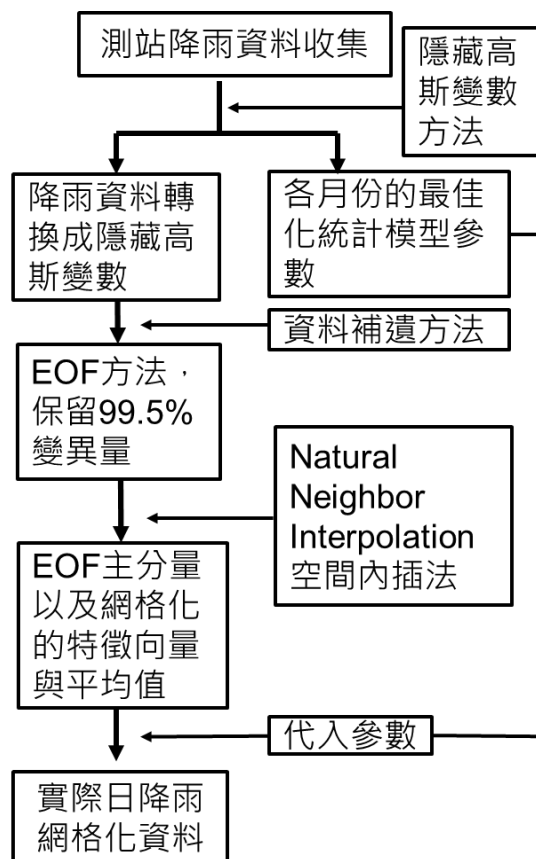


圖 1、日降雨資料產製流程

■ 文字敘述：

步驟一、

將蒐整的測站資料利用隱藏式高斯分配(Latent Gaussian variable, LGV)，對測站日降雨資料，分月份做標準化與參數化，轉換成隱藏高斯變數(資料為常態分佈)，並找出各月份的最

佳化統計模型與參數。

步驟二、

將經過 LGV 轉換後的測站降雨資料，利用資料補遺方法，將缺資料的地方補齊。

步驟三、

再用經驗正交函數(Empirical Orthogonal Function, EOF)分析方法，把資料拆解成時間和空間兩部分，用分析得到的特徵向量(eigenvector)，保留 99.5%的變異量後，利用 Natural Neighbor Interpolation 空間內插法，對特徵向量和 EOF 裡扣掉的平均值做網格化。

步驟四、

利用 EOF 分析得到的主分量(Principle Component, PC)乘回去相對應模組(mode)的網格化特徵向量值做累加，並加上各網格的平均值，得到標準化降雨網格資料，最後再代入，之前計算得到的各個月份的最佳化統計模型參數，反算出實際降雨量，得到日降雨網格化資料庫。

步驟五、

產製輸出為 1 公里解析度的網格日資料，再以區域平均的方式轉換為 0.05 度(約 5 公里)，月資料則是由日資料進行轉換。

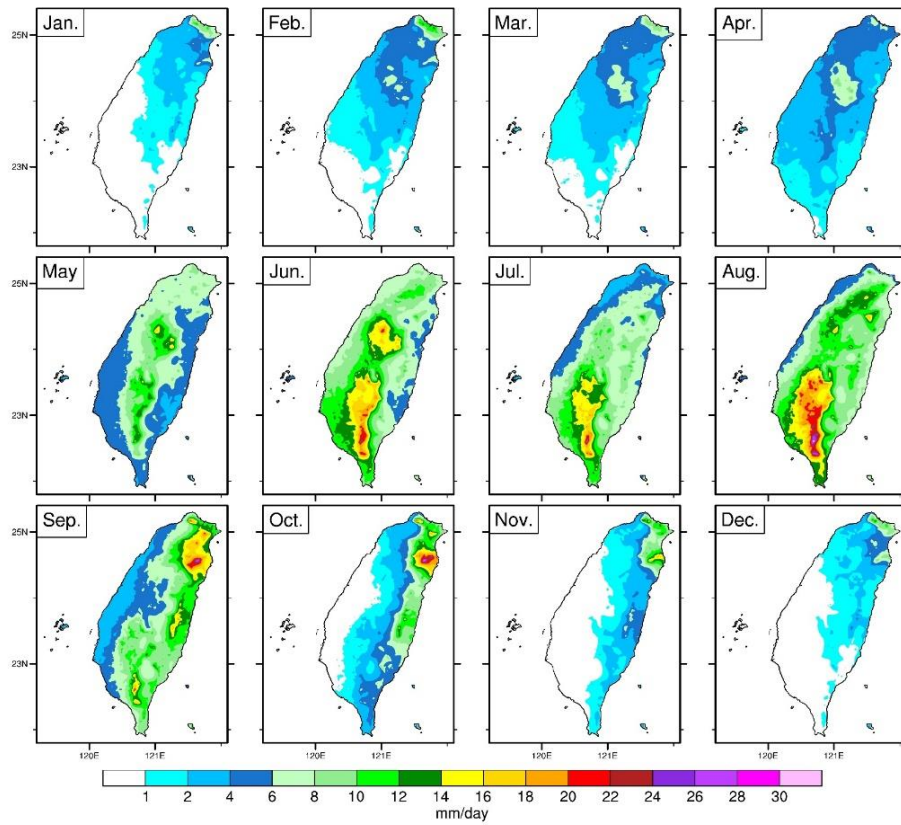
5. 資料不確定性

誤差來源有三個部分：

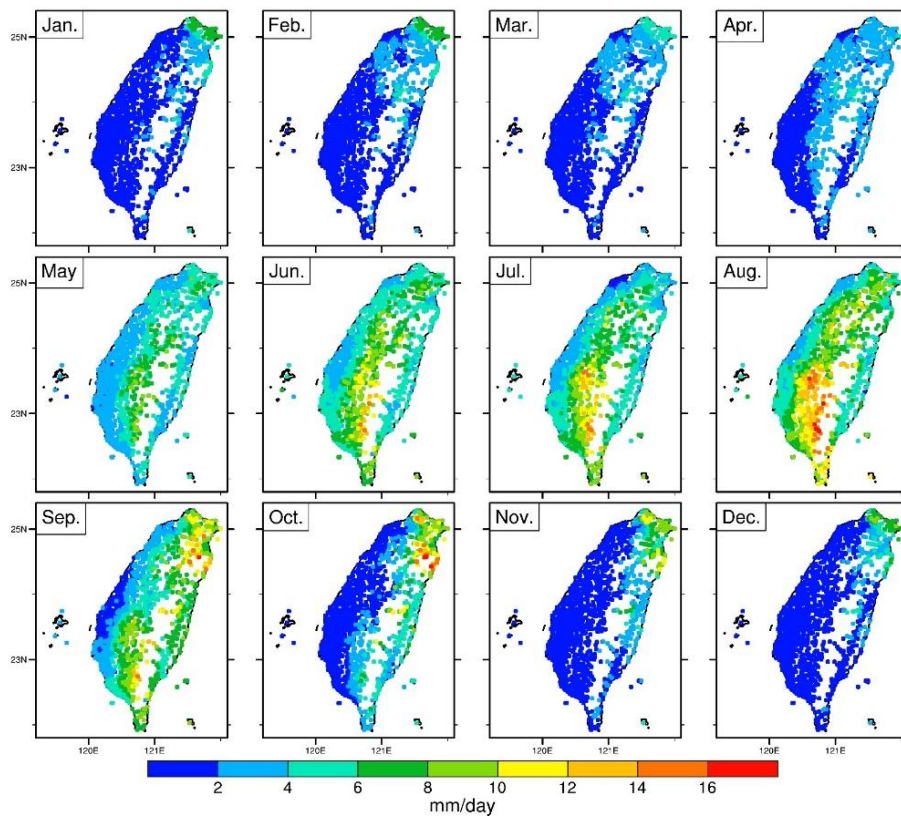
- **最佳化的統計模型不確定性：**不同月份的測站資料做參數化時，得到的參數和標準化降雨資料，反算實際降雨量和觀測值做比較，此部分誤差很小。
- **測站資料補遺的不確定性：**當測站資料缺失時，計算附近參考測站的實際降雨觀測資料標準差，並計算不同月份的標準差平均值。結果顯示降雨量越大的地區，標準差越大，且山區比平地大。
- **網格化時的不確定性：**結果顯示，降雨量越大的地區，誤差越大，山區比平地大，並顯示出誤差和測站豐富度負相關。

分析結果：

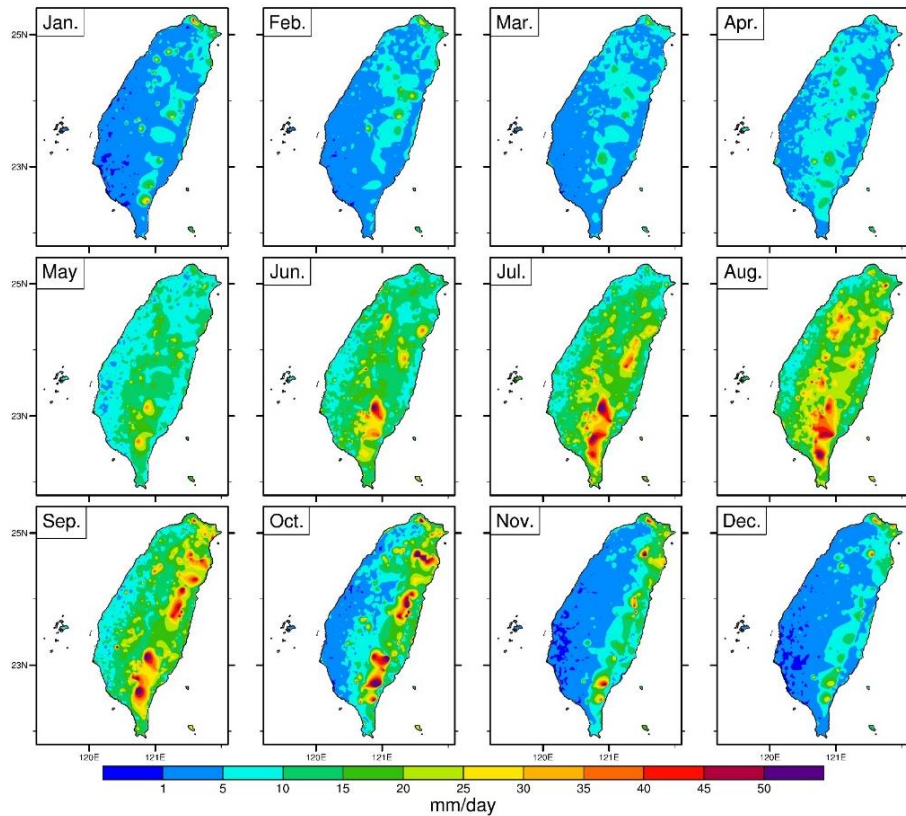
月平均氣候值：



圖二、日降雨觀測資料氣候值(單位 mm/day)



圖三、資料補遺誤差分布(標準差，單位 mm/day)



圖四、網格化不確定性分析(均方根誤差，單位 mm/day)

6. 參考文獻

Simolo C, Brunetti M, Maugeri M, Nanni T, 2010, Improving estimation of missing values in daily precipitation series by a probability density function preserving approach. International Journal of Climatology ,30,1564-1576.

Watson D, 1994, Nngridr: An implementation of natural neighbor interpolation., Dave Watson Publisher, Claremont, Australia, 170pp.

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。

翁叔平、楊承道(2020)，臺灣日降雨網格化觀測資料之建置與驗證，臺灣氣候變遷推估資訊與調適知識平台電子報，第 37 期。

https://tccip.ncdr.nat.gov.tw/km_newsletter_one.aspx?nid=20200515170742

8. 文件引用

楊承道，林士堯（民 110 年 9 月 15 日）。網格化觀測雨量日資料生產履歷(4.0 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/20200117105621.pdf

附件一、篩選和補遺

■ 測站資料篩選流程:

a、刪除缺乏測站資訊之測站：獲得之原始資料共有 1,497 站，經查發現其中 14 站缺乏測站資訊(經度、緯度或高度位置資料)，故剔除這些測站後，獲得完整測站資訊之測站共有 1,483 站。

b、刪除無數據資料之測站：1,483 站中，只有站名站號但無觀測數據資料的測站有 217 站，故刪除這些有名無實之測站後，有資料之測站數減為 1,266 站。

c、刪除經緯度錯誤與資料錯誤或不齊之測站：1,266 站中，發現火燒寮測站資料實為十日累積雨量而非日雨量，故在日資料處理時刪除該站，而臺中港、池上、外傘頂測站，其經緯度錯標至海上，因資料的正確經緯度無法從資料提供者處獲得，故只能刪除，剩下可用的 1,262 站。

d、經緯度重複測站之處理：在剩下的 1,262 站中，發現有 146 站經緯度重複。不少經緯度重複的測站，起因於單位名稱或直屬單位的更迭，導致重新改變站名或站號，因此經過一一比對，有的測站以合併成一站的方式來解決；有些則因其中一站的數據長度非常短(可能是測試性質之測站)，故刪除極短測站而保留較長紀錄一站的方式處理。在交叉比對後，146 站中經緯度兩兩相同的有 134 站(即 67 組)，三站經緯度相同的有 12 站(即 4 組)。在一一考慮各組測站地理位置、資料長度、資料重疊情形後，分別將各組合併或刪除該組其中一站，最後的處理結果是：經緯度兩兩重複之測站(67 組)：兩站合併成一站的有 49 組(從 98 站併為 49 站)，刪除一站而保留一站的有 18 組(從 36 站併為 18 站)；而三站經緯度重複的 4 組，則合併為 4 站。

■ 測站資料補遺流程:

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 $c_d = \frac{\bar{d}^2}{\ln 2}$ 。 $w_i^d(x, y)$ 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，可利用參考測站的雨量 LGV 值，先乘以該站的乘積權重，而後累加之。累加後的總值就是目標測站缺失資料的補遺值。