

# 中央氣象局所屬雨量站群集分析

陳品妤 沈里音  
中央氣象局氣象科技研究中心

## 摘要

臺灣地區的降水特性由於天氣系統與地理條件的不同存在顯著的差異性，本研究挑選中央氣象局(以下簡稱氣象局)全臺有人氣象站和自動雨量站資料量達80%以上且均勻分布在全臺共318個測站作為分析對象，使用1998至2017年共20年的日雨量資料進行分析。

本研究使用K-means分群演算法，將318站分為六個群集，利用雨量在不同降水類型上的第一四分位數和第三四分位數以及該季的最大日雨量的第一四分位數、第三四分位數和第90百分數共五種的特性來看降水分區的結果。其中降水類型細分為五季：12月跨年到2月定為第一季(DJF)、3-4月為第二季(MA)、5-6月為第三季(MJ)、7-8月為第四季(JA)、9-11月為第五季(SON)。

分類的結果顯示318個測站依臺灣地區的降水特性分為6個區域，東部分成2區，中央山脈以西分4區。一、東北角和宜蘭地區；二、花蓮和臺東南至太麻里溪為止的區域；三、臺北盆地延著平地區域往南至臺南的鹽水溪以北為一區；四、南部高屏600公尺以下的區域；五、南部高屏600公尺以上的區域；六、中、北部與南投地區600公尺以上的區域。詳細分群參見圖1。

關鍵字：氣候分區、降水量、K-means

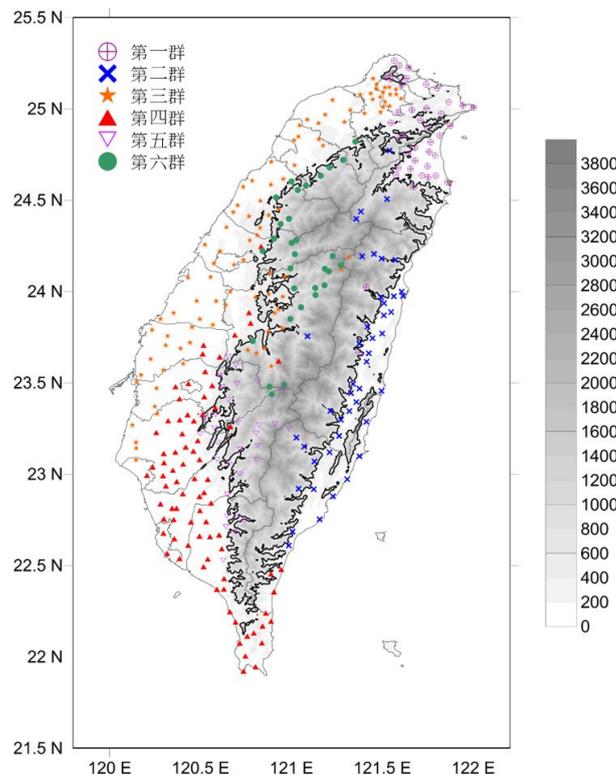


圖1 318站依雨量特性分六個區域。(圖中黑色線段為高程600公尺等高線)

## 一、前言

隨著各地觀測測站不斷增多，測站密集度增加的情況下更有利於對於各地的氣候特性做探討，其中雨量特性更是最難以捉摸，大尺度的氣候特性大致方向已趨於確定，但較細部的天氣系統所帶來的影響範圍，到如今仍有很大的探究空間。

隨著測站數量增加，各種應用需求也快速的湧現，然而對於資料如何轉換成有用的資訊，這需要各界多方的溝通與觀念交流，期望以此研究帶出更多的想法，達到拋磚引玉的效果。

臺灣地區的降水特性受到天氣系統與地理環境交互影響，存在顯著的差異性，無法單用一個值或一個準則可以通用全國，需發展出適合各地的界定值，所以將已往常用的均值的概念擴充到各分區。本研究使用K-means分群法，將大量的測站資料壓縮成少數代表點來代表各群的資料，進一步探討各地的特性差異，同時也想了解各分群的降水特性為何，試著歸納出可用的資訊。

## 二、資料及研究方法

### (一) 資料

本研究所使用的資料為中央氣象局氣候變遷應用服務能力發展計畫所整集的雨量資料，將雨量測站所回傳的雨量特殊碼累計於後(-9996)經過處理後，在時、日、月雨量資料皆有較精準的計算結果。共計有318個雨量目標測站作為分析使用。

資料年份為1998年至2017年共20年的資料長度，包含當日24小時觀測資料筆數超過80%所計算的日雨量資料與當月的所有小時雨量觀測資料量超過80%所累積計算的月雨量。

依降雨特性分五個季節：第一季12~2月[DJF]、第二季3~4月[MA]、第三季5~6月[MJ]、第四季7~8月[JA]、第五季9~11月[SON]。

### (二) 方法

本研究使用K-means分群演算法，此由J.B. MacQueen於1967年提出的演算法，須事先設定群集的數量k，隨機找出以k個測站作為初始的中心位置，計算所有的測站到此k個測站的距離，依照最短的距離分為同一群。初始分群的結果會產生各自的質量中心，此質量中心將取代一開始隨機挑選的k個測站，再重新計算所有測站到此質量中心的距離，

重新分配各群集中的測站分子。重複上述步驟直到各群之間的分子不再改變為止。詳細數學模式不在此介紹。

利用K-means的分群演算理論，利用最短距離作為最佳化的分區依據，在本研究中嘗試不同降雨資料的統計特性的組合作為雨量分區的依據。在嘗試近十種的統計特性組合後，選出最佳結果為五種統計特性：當季所有月雨量資料集合的第1四分位(Q1)和第3四分位數(Q3)，以及當季中的最大日雨量值所形成的年最大值序列的Q1、Q3和第90百分位數(P90)，配合五個降雨特性季節，共有25個影響因子作為K-means的分類資料。

## 三、分析結果

本研究設定分6群，比現行的四分區再多兩個，不分更多區的原因在於在使用上以及結果分析上的便利性。分類後的結果如圖2。K-means法透過25個影響因子分類，將318個雨量目標站分為六群，第一群有39站，分布在東北角一帶及宜蘭地區。第二群有44站，中央山脈以東，蘭陽平原以南的花東一帶。第三群有94站，最多站的一群，廣布在中央山脈以西的沿海平原一帶，北從臺北市南至臺南市中心。第四群有80站，次多的一群，嘉南平原沿阿里山山脈丘陵一帶以南至屏東平原，繞過鵝鑾鼻至臺東的太麻里溪以南。第五群和第六群屬山地地區，分別是31站和30站，測站數最少的兩群，分界大致以濁水溪為界，以南是第五群的範圍，以北則是第六群的範圍。

將六群的五種統計特性分別歸納並分析，將各群取平均，先比較群與群之間在降雨季節特性的差異點。最後再比較統計特性所代表的意義為何。

### (一) 特性一 月雨量第1四分位數

各季節下月雨量的第1四分位，即歷年月資料排序後，第25百分位的數值，意義上為雨量較小的數值，但又不至於為0。一樣是降雨量的Q1，在不同的季節中的差異又是為何？以第一群來說，屬東北角宜蘭地區，在季節的排序從小至大依序是JA、MA、DJF、MJ、SON。將各群的平均結果整理在表1，並繪製成曲線圖，如圖3所示。第四群和第五群的季節特性一樣，從小至大依序是DJF、MA、SON、MJ、JA。差別在於值的大小，第五群的平均日雨量在各季節皆會大於第四群。另一個相似的兩群是第三群和第六群，差異在第1季和第5季的順位，其次依序是MA、JA、MJ。其值為第六群大於第三群。

## (二)特性二 月雨量第3四分位數

各季節下月雨量的第3四分位，即歷年月資料排序後，第75百分位的數值，意義上為雨量較大的數值，但又不至於太受極值影響。將各群的平均結果整理在表2和圖4。除了第四群和第五群仍維持和Q1相同特性外，其他各群皆有些微變動。各季相對較大的月雨量值，第一群在MA為最小，SON為最大，其他群皆在DJF為最小，JA最大。

## (三)特性三 最大日雨量

每年各季節下最大日雨量所形成的序列的第1四分位，即極值中的較小值，時間尺度降為日來看。意義上為每年日雨量較極端的數值中，其Q1值會是多少？將各群的平均結果整理在表3和圖5。各群平均最大日雨量的值域是12.06 ~ 186.83 mm/day。日尺度的特性應該會與月尺度有所差異，從結果來看第一群和第二群特性相似，第三群到第六群的特性相似，這反映出臺灣地區東部和西部的日雨量極端特性有所不同。西部在DJF最小，西北部(第三群和第六群)最大在MJ，西南部(第四群和第五群)最大在JA。

每年各季節下最大日雨量所形成的序列的第3四分位，即極值中的較大值，各群平均最大日雨量的值域是32.97 ~ 424.18 mm/day。各群的平均結果整理在表4和圖6。各群在最大日雨量的特性基本上明顯可以分為西部和東部，西部在DJF最小，在JA最大；東部在MA最小，在SON最大。

每年各季節下最大日雨量所形成的序列的第90百分位數，即極值中的極值，各群平均最大日雨量的值域是50.34 ~ 546.00 mm/day。各群的平均結果整理在表5和圖7。西部在DJF最小，在JA最大；東部在MA最小，在SON最大。

## 四、結語與討論

將318個雨量目標站利用K-means法分為六群，將目標縮小至小群來探討區域特性，期望能發展出不同降雨氣候特性的應用。本研究結果綜合整理為以下結果：

- (一) 在各季的月雨量特性當中，看出第一群和第二群有明顯異於其他四群的特性，由此可知臺灣大致以中央山脈為界，西半部和東半部有不同的降雨特性，其東半部可再細分為宜蘭和花東地區。
- (二) 在各季的月雨量特性中，第一群明顯和其他各群特性不同，它在JA時是最小，其他群幾乎都是在JA時最大。

(三) 第五群不管在哪種特性下，其各季節雨量排序幾乎不變，除了在較極端的兩個特性中是SON大於MJ，基本上在各季中由小至大的排序為DJF < MA < SON < MJ < JA。

(四) 在最大日雨量的特性下，很明顯第一群和第二群為一組，第三到第六群為一組，差別在於值的大小，以P90來看在MA時，第一群大於第二群；第六群>第五群>第三群>第四群。也就是高山區的第五群小於第六群，以及在地平的第四群小於第三群。在JA時，第二群大於第一群；第五群>第六群>第四群>第三群。順位剛好和MA相反。

K-means法為非監督式學習，主要根據資料本身的特性，來進行資料分群的一種方法。當我們對資料還不夠深入了解時，便可先用K-means法，觀察潛藏在資料中的特性，再擬定後續各式分析的應用。

此研究結果也讓我們對臺灣地區降雨型態更加理解，整體而言臺灣在冬季受到東北季風影響，降雨以東北角為最，這種降雨型態在東半部大至上可以延伸至宜蘭花蓮交界之處。隨著季節交替，春雨、梅雨等南方氣流所造成的降雨，影響的範圍以西南側迎風面為主，與冬季東北季風為主的長時間降雨型態並不相同；夏季則為熱對流所造成的降雨型態為主，加上颱風帶來的雨勢。本研究利用不同降雨延時(日、月)資料的特性進行降雨型態分成6個分群，可做為日後研究降雨型態之參考。

## 參考文獻

陳品好、沈里音、陳雲蘭，2015：“「臺灣長期氣候資料整合分析」計畫研究(2) 一自動站雨量累計於後記錄值的問題分析及處理”，天氣分析與預報研討會，中央氣象局。

MacQueen, J., 1967: "Some Methods for Classification and Analysis of Multivariate Observations." Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, 1, 281-297.

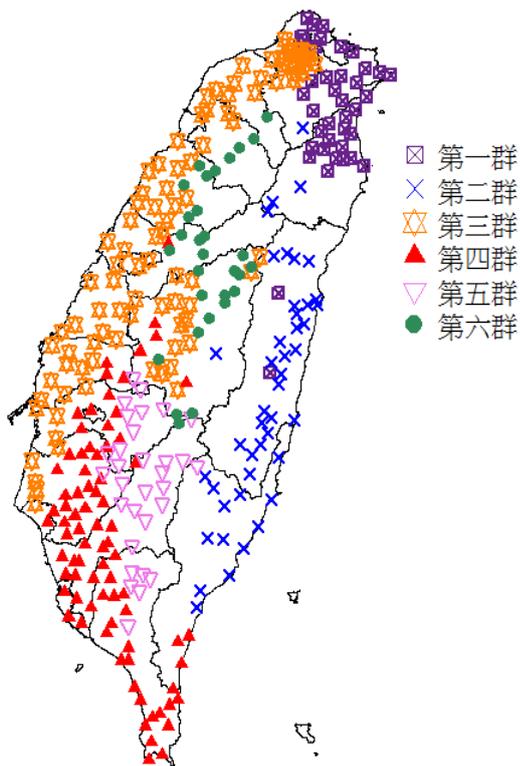


圖2 K-means法分為六群的結果。

表1 各分群平均月雨量在不同季節上的Q1的排序(小至大, 1→5)

	DJF	MA	MJ	JA	SON
第一群	3	2	4	1	5
第二群	1	2	4	5	3
第三群	2	3	5	4	1
第四群	1	2	4	5	3
第五群	1	2	4	5	3
第六群	1	3	5	4	2

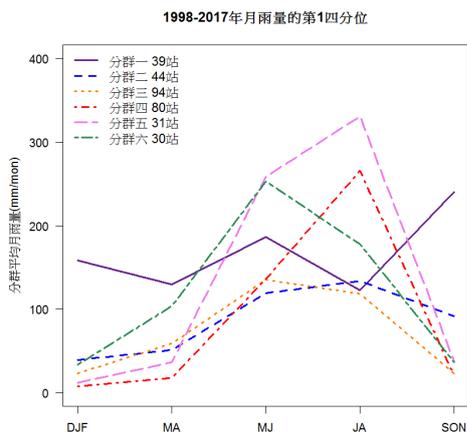


圖3 各分群在五季中的月雨量在20年資料中的Q1的平均表現。

表2 各分群平均月雨量在不同季節上的Q3的排序(小至大, 1→5)

	DJF	MA	MJ	JA	SON
第一群	3	1	4	2	5
第二群	1	2	3	5	4
第三群	1	3	4	5	2
第四群	1	2	4	5	3
第五群	1	2	4	5	3
第六群	1	3	4	5	2

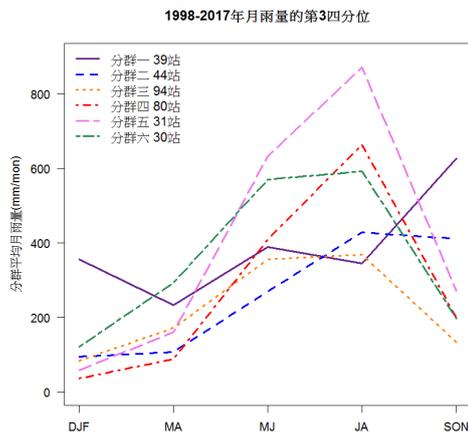


圖4 各分群在五季中的月雨量在20年資料中的Q3的平均表現。

表3 各分群平均最大日雨量在不同季節上的Q1的排序(小至大, 1→5)

	DJF	MA	MJ	JA	SON
第一群	2	1	3	4	5
第二群	2	1	3	5	4
第三群	1	2	5	4	3
第四群	1	2	4	5	3
第五群	1	2	4	5	3
第六群	1	2	5	4	3

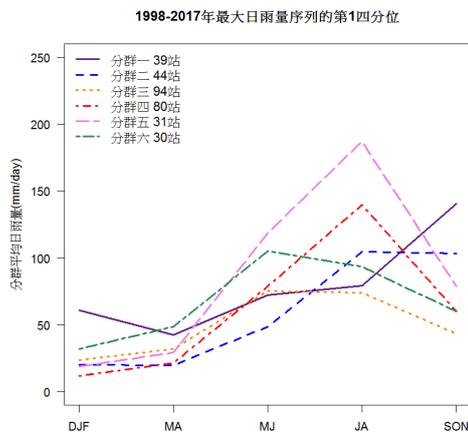


圖5 各分群在五季中的20年最大日雨量的Q1的平均表現。

表4 各分群平均最大日雨量在不同季節上的Q3的排序(小至大，1→5)

	DJF	MA	MJ	JA	SON
第一群	2	1	3	4	5
第二群	2	1	3	4	5
第三群	1	2	4	5	3
第四群	1	2	4	5	3
第五群	1	2	3	5	4
第六群	1	2	4	5	3

1998-2017年最大日雨量序列的第3四分位

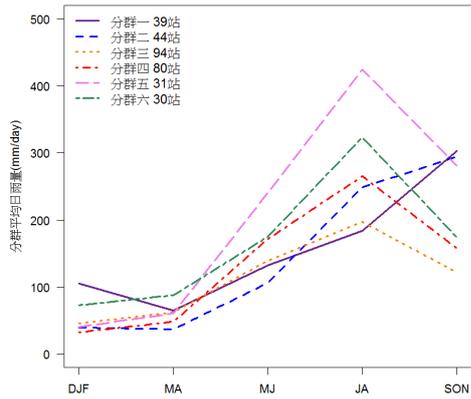


圖6 各分群在五季中的20年最大日雨量的Q3的平均表現。

表5 各分群平均最大日雨量在不同季節上的P90的排序(小至大，1→5)

	DJF	MA	MJ	JA	SON
第一群	2	1	3	4	5
第二群	2	1	3	4	5
第三群	1	2	3	5	4
第四群	1	2	3	5	4
第五群	1	2	3	5	4
第六群	1	2	4	5	3

1998-2017年最大日雨量序列的第90百分位

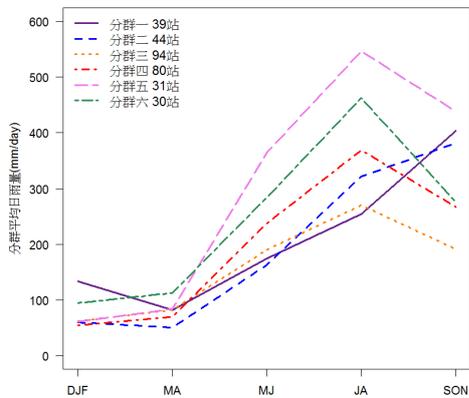


圖7 各分群在五季中的20年最大日雨量的P90的平均表現。