

臺北測站之統計降尺度定量降水研究

姚春仔 林沛練 曾仁佑

中央大學大氣物理研究所

摘要

為了統計降尺度定量降水推估的改進和未來朝向氣候變遷降水研究之目的，本研究利用綜觀天氣分型結合邏輯和非線性迴歸，建立統計降尺度降水模式並評估其在臺北測站之效用與仍需改進的問題和目標。

本研究使用了臺北測站的地面觀測資料和 NCEP 高空再分析資料來建構降水模式，研究包含下列兩個主要部分: (1)自動天氣分型，及(2)發展降水天氣型之降水推估模式。天氣分型的結果可以歸納出 7 類和降水有關的天氣型態，分別為冷季的冷鋒、滯留鋒 I、滯留鋒 II、暖季滯留鋒、熱帶低壓、局地熱對流及颱風。結果顯示，有進行天氣分型且搭配邏輯迴歸的機率預報和非線性之降水量推估與不進行天氣分型的降水推估結果相比是有改進的。天氣類型中暖季滯留鋒為間歇性的中尺度對流降水，統計模式對於大部分的移動性對流降水較難以掌握，因測站還未能反應移動的鋒面對流系統天氣變量時，對流系統已帶來降水，導致迴歸效果不佳。而檢驗後發現，熱帶低壓容易與輕颱混淆，且颱風不適用此天氣分類方法，導致檢驗時的降水推估表現不佳，除此之外其他天氣類型都有良好的改善。為改善颱風類型的降水推估，最後獨立挑出颱風個案改以路徑分類做測試，結果有很好的改善。

關鍵字: 自動天氣分型、邏輯迴歸、統計降尺度。

一、前言

目前大部分的全球環流模式(Global climate models; GCM) 解析度大約在幾百公里而區域氣候模式(Regional climate models: RCM)解析度可達數十公里。但在研究氣候變遷下的影響問題時(例如:空氣汙染、極端降雨或凍雨...等)，我們需要相當於單點測站的氣候資訊，因此需要比全球環流模式和區域氣候模式更精細的空間解析度的氣候資訊，這時就需要透過降尺度的方法來得到。降尺度方法的發展上，分成動力降尺度(dynamic downscaling)和統計降尺度(statistical downscaling)兩種主要方法。雖然數值模式可提供大部份的天氣資訊但是對於小尺度訊息提供仍然不足，尤其對於未來的全球氣候模擬，需要相當龐大的計算資源與空間，且要提高全球模式的解析度，其代價甚是昂貴，因此對於未來氣候變遷的影響研究，許多科學家使用統計降尺度的方法來得到單點測站的氣候資訊並做進一步的分析。

目前統計降尺度的研究，主要是對大尺度的天氣變數(預報因子)和局地或是單點測站的地面變數(預報量)，兩者之間的關係做量化處理。而大多的統計降尺度研究，著重在單點測站的日降水量預報，是因為降水量一直都是各個模式中最主要的預報項目，且降水量雖能從氣候模式中輸出得到，但其準確度相較於其他氣象變數都來的低(Wilby, 2004)，這也是本研究做降雨量預報的原因。

為了統計降尺度定量降水模擬的改進和未來朝向氣候變遷降水研究之目的，本研究選用全球模式輸出之氣象場資料，用於建立大尺度氣象變數與測站點降水量之間的關係。加上綜觀天氣分型結合降水模擬目前在台灣尚未有相關的研究討論，於是本研究利用台北測站做新方法的測試，希望藉此能對統計降水量的預報有所改進。因時間的緣故，目前本研究只是初步地將降水模式建構起來並評估其效用，除了天氣分型的方法改進之外，還有多種不同的方法，像是不同變數的選擇和加入不同地區的變數，這些方法應該會

對降水推估有更好的改進，未來可以多做測試，得到較好的結果後，之後便可運用於氣候變遷的降水研究。

二、 資料及分析方法

2.1 資料來源與處理

本研究的統計分析資料時間為 1992 年至 2012 年。其中因 1997 年測站資料嚴重缺失，扣除 1997 年後，資料總共有 20 年。所使用的資料分為地面觀測資料及高空的再分析資料。地面觀測資料來自中央氣象局 (CWB) 臺北局屬測站，每小時的海平面氣壓、地面溫度、風向、風速、總雲量、露點溫度，還有每日的累積降水量。高空資料使用美國國家氣象局國家環境預測中心 (NOAA National Centers for Environmental Prediction ; NCEP/NCAR Reanalysis 1) 每 6 小時一筆的全球再分析資料，網格解析度為 $2.5^{\circ} \times 2.5^{\circ}$ ，全球網格點數 144×73 ，垂直解析度為 1000hPa 到 10hPa 共 17 層，本研究所使用到的高空氣象資料包含了氣溫、相對溼度和 U、V 風速，高度只取 5 層，分別為 925、800、700、600 和 500hPa，因為降水的生成與型態主要是由 500hPa 以下的大氣參數來決定 (Cheng et al. 2004)。

選取這些每小時氣象變數是基於氣團性質的不同而做的綜觀天氣系統分類，此分類法根據熱力學和流體力學的變量 (氣溫、水汽、壓力、風向和風速、雲覆蓋量即能見度等等) 來判斷 (Davis and Walker 1992)。最後氣象變數是參考 Cheng et al. (2010) 的研究所決定的，這些變數所分類出的綜觀天氣型態適合研究評估氣候變遷對各種環境問題的影響 (例如：空氣品質和人類健康) (Cheng et al., 2007a)。

2.2 分析方法

整體來說建構統計降水模式有兩個主要部分，依實驗步驟為：(一) 自動天氣分類，並定義和降水有關的天氣類型，及(二) 發展降水類型內的建模，及驗證降水模型。

其分析過程中使用到許多統計的計算和方法，這部分都是使用 SAS (Statistics Analysis System) 軟體來處理。此 SAS 統計分析系統，由北卡羅來納州立大學兩位統計系任教的教授開始發展的，最早只是一個數

學統計軟體，於 1976 年由 Jim Goodnight 及 John Sall 博士等人成立統計分析系統公司，並且正式推出相關軟體。目前 SAS 最新的版本發展到 9.4 版。

本研究首先是對資料的前置處理，也就是進行自動天氣分型，其目的是先對資料做系統性的整理，讓迴歸效果有所提升。此部分由下面三個步驟來完成：(1) 主成分分析，(2) 分層聚類過程，(3) 不分層聚類過程 - 判別分析法。主成分分析屬於多變量分析方法中的一種，目的是化繁為簡，分析出變數的主要成分作為座標軸，並抽取出真正的重要資料，類似濾波的想法。經過主成分分析後，原臺北測站每日 144 個成份，縮減為冷季 14 個主成分和暖季 14 個主成分，這些主成份可解釋暖季原始數據的 87% 和冷季的 89% 訊息。其中，144 個天氣變量為每日 6 個天氣要素 (海平面氣壓、溫度、露點溫度、總雲量、U 風場和 V 風場)，且為 24 小時資料，因此一天有 144 個變量。

分析過程中會將一年分為冷、暖兩季，主要是為了將秋、冬轉換時的鋒面和春季的梅雨鋒面做個良好的區分。而冷、暖季的 14 個主成份和原始資料的線性組合後，得到新的變量稱為主成分分數，之後便以主成分分數來做為分類條件，分別對冷季和暖季進行天氣分類。經主成分分析得到主成分分數 (print component scores) 後，接下來便是利用 SAS 統計軟體進行群集分析，將比較相似的樣本聚集在一起，形成群集 (cluster)。群集分析可分成分層法 (Hierarchical) 和不分層法 (Nonhierarchical)。其中，分層聚類法計算距離時可分為『點間距離』和『群間距離』。本研究是使用群間距離裡的『平均法』 (Average Linkage) 來計算不同群間的距離。

群集分析的結果出來後，類型數目一開始會有很多，每個群集內的天數有多有少，當群集內樣本數太過少 (少於 50 天) 並沒有很大的代表性和意義，會導致進行迴歸分析時，統計的顯著性不大，所以將這些小的群組進行歸併的動作。歸併的原則是以幾個較大群體為基礎，以及歸併的小群體天氣屬性一定要和這些較大的群體相近，而且歸併以後，這些較大的天氣類型屬性不能改變。本研究的歸併動作，是使用人工去看每日的東亞地面天氣圖來判斷，隨機挑選群集內的幾十天來看其大致的天氣型態為何。天氣圖來源為中

央氣象局的地面東亞天氣圖和日本氣象廳的東亞地面天氣圖。

經過分層聚類過後已確定類型數目，將進行下一個步驟-不分層聚類過程(判別分析)。不分層群集的分群法則有很多種，本研究採用 K-中心法(K-means method)，這個方法也是最普遍被採用的。多做這一步的好處是可以使(1)群內變異最小；(2)群間變異最大(Cheng et al. 2007a)，得到最佳的聚類。

判別分析完成後得到了數個主要的天氣型態，而接下來我們的目的是降水迴歸建模，因此我們只需分析跟降水有關的天氣類型即可，於是接下來找出跟降水有關的天氣類型。為了確定與降水事件最有相關的天氣型，這裡有兩個部分:(1)以在各個類別內的日平均降水量大於 1.0 毫米為基準，日平均降水量大於 1 毫米的屬於會降水的類別，小於平均日降水量 1.0 毫米的則屬於不會降水的類型。(2)而除了分析各個類別內的日降水量外，還要計算各個類別內日降水事件的發生頻率(actual frequency)和期望頻率(expected frequency)的比值，如式(2.12)，若是該天氣類型內的降水日和總降水日數的比例(即實際頻率)，與該天氣型的天數和整體資料總天數的比值(即期望頻率)，兩個頻率值的比值大於 1.0，則表示該天氣型除了是會降水的天氣類型，還是個降水事件發生頻率比較高的天氣類型，足以選擇為和降水相關的天氣型。

$$\frac{\text{actual frequency}}{\text{expected frequency}} = \frac{\text{Rain days of type/total rain days}}{\text{Days of type/total days}} > 1.0 \quad (2.2)$$

找出會降水的天氣類型後，接下來就是針對這些會降水的類型建立各自的降水迴歸模式。降水建模包含兩個部分:(1)逐步邏輯迴歸，用來推估日降水事件的發生機率；(2)找出邏輯迴歸的機率值與真實測站降水之間的關係，利用非線性迴歸方法推估日降水量。

檢驗的部分，一開始整個氣象資料和降水資料會被分成兩個部分，一部分是用來建立降水預報模型，另一部分資料用來做檢驗使用。檢驗的數據從 20 年中隨機挑選 5 年(1992 到 2012 年中隨機挑選的五年分別為: 1995、1999、2002、2005、2011 年。)，也就是所有天氣資料年份的 1/4 年會用來進行驗證，這樣的話檢驗數據集將會完全獨立於發展建模的部分。

三、 結果

利用地面六個氣象要素經過聚類分析得到綜觀天氣分類後，再根據每個天氣類型內的平均降水量是否大於 1mm 之標準，歸納出 7 種和降水相關的天氣型態。分別為:冷鋒、冷季的滯留鋒 I、冷季的滯留鋒 II(西南雲系的移入)、暖季的滯留鋒 III、熱帶低壓和熱帶擾動、副高邊緣(局地熱對流降水)以及颱風。

從 1992 到 2012 年期間的臺北測站，本研究使用邏輯迴歸所建的降水機率模型的和諧性分數都有在 0.97~0.98 之間，表示模型能力是可信的。這裡也測試了三組不同的分類降水事件(0.2、10 和 25mm)，結果顯示使用邏輯迴歸來模擬降水機率的能力不錯，因為都有相當高的 ROC 得分。且建模數據集和檢驗數據集的 ROC 得分都有大於 0.92(任和降水閾值)。接著使用可信度圖(圖 3-1)來對降水機率模型進行評估，結果顯示，建模數據集和檢驗數據集的機率模型蠻接近對角線的分布，這也說明了邏輯迴歸的降水機率模型結果是可信的。

日降水量的推估，這裡將利用表 3-1 的標準對降水量推估結果進行評估

圖 3-2 為 7 種天氣類型的推估結果為優、良、普通和差的百分比表現，其中，Type5(熱帶低壓)建模(圖 3-2 e)的時後有進行天氣分類的優加良的百分比數比不進行天氣分類的百分比數高，但是在檢驗的時候，檢驗的優加良百分比數比起建模的百分比數低，於是查看了檢驗數據集中 Type5 的天氣圖，發現此天氣分類容易和一些臺灣附近有颱風的天氣狀態混合或是有輕度颱風也會被判別為熱帶低壓的天氣類型，表示做天氣分類時，六個變數對於颱風和熱帶低氣壓系統有時難以分辨。Type4(暖季滯留鋒)(圖 3-2 d)的推估結果看到，此方法對於暖季滯留鋒降水推估改進不彰，有進行分類和沒進行分類的降水推估結果是差不多的，主要是春季梅雨鋒面移入的一波波鋒面對流系統無法適時的在台北測站上空被量測到，又降水變異度大，因此天氣分類後沒有辦法得到改善。圖 3-2 a、b、c 分別為冷季的三個天氣類型(冷鋒、滯留鋒 I、滯留鋒 II(西南雲系移入))和暖季的類型，有進行天氣分類比沒有進行天氣分類的降水推估好，建模和檢驗數據集約

有 10%~20%的改進，雖然 Type1(冷鋒)檢驗時其優加良的百分比數差不多，但是 Type1(冷鋒)推估結果為差的百分比數有下降 5%。此外，可以看到不論是建模還是檢驗數據集的颱風類型表現的都是最差。結果發現颱風的天氣類型(圖 3-2 g)使用這六個每日天氣要素去分類的效果不理想，因為這六個天氣變量(海平面氣壓、溫度、露點溫度、U 風場、V 風場以及總雲覆蓋量)主要是對綜觀尺度的系統來做判別，但是對於颱風比較屬於中尺度的系統無法做出良好的分類，分類的時候颱風類型多半為單獨獨立的個案類型，天氣狀態變異度大，再加上颱風的降水相較於其他天氣系統變異度大，導致其效果不甚理想。

因為颱風用這六個天氣變量來做分類效果不好，所以將颱風獨立出來另外嘗試使用路徑分類，再使用同樣的邏輯迴歸和非線性迴歸得到降水量的推估。結果顯示，颱風降水事件單獨挑出來，並且使用路徑來做分類的效果比用氣團的特性來分類效果要好。對於颱風從北部經過的第一個路徑分類(Track1+Track2)和經過臺灣中部(Track3)的颱風，降水推估都有明顯的改進，Track1+Track2 的降水推估結果，優和良百分比增加 20%，Track3 的優和良百分比數增加 10%，但對於從臺灣東部由南往北經過(Track6)的颱風降水量推估，優和良的百分比數則下降 10%。分析其原因後發現颱風歷史路徑 6 的颱風路徑變異度很大，造成此路徑的天氣變量和降水情況複雜，因此迴歸建模的預報結果不好。

四、 結論與討論

本研究的自動天氣分型過程包含了主成分分析、分層聚類、判別分析(不分層聚類)，經由這些步驟後得到台灣臺北測站的天氣分類，而天氣分類的結果中有七種是和降水有關的天氣類型，之後以這七種降水天氣類型做逐步累積邏輯迴歸得到降水機率，再將降水機率和實際降水量以非線性迴歸建模得到最終的降水推估。

從整個降水迴歸模式的推估結果可以看到，執行天氣分型比未執行天氣分型確實能夠改善定量降水推估，七種與降水有關的天氣類型降水推估中，冷季的三種天氣類型不論是建模還是檢驗數據的降水量推估

結果，其為優和良的百分比數有 10~20%的改進。冷季降水型態較穩定，降水推估的改進效果較暖季多。暖季建模部分，對於熱帶低壓和局地熱對流降水，執行天氣分型的優良百分比數比未執行天氣分型的百分比數改進了約 10%，至於暖季的滯留鋒面和颱風降水則是沒有甚麼改進，因為暖季滯留鋒的鋒面對流胞從測站外移入，天氣資訊在台北測站上空未能掌握但已有強降水的發生，而颱風降水受颱風位置與測站間距離遠近影響，且外圍雲帶的移入也容易帶來降水，可是短暫快速的雲帶強降水天氣資訊也難以在測站上空被捕捉到，這些都是綜觀網格點的資料對於這些小尺度系統天氣訊息提供不足所造成的問題。此外，熱帶低壓、滯留鋒面和颱風的天氣狀態，這六個天氣變數狀態有某種程度的相似，因此天氣分類時可能會有些許混淆的狀況，尤其是熱帶低壓和熱帶擾動的天氣系統再分類時，容易和一些輕度颱風和颱風在臺灣附近的時候混淆。這問題經由獨立抽出處理和以路徑做分類測試過後，颱風這類型的降水推估有良好的改善。雖然颱風路徑六的優和良的百分比數下降，但這是因為颱風路徑六的路徑差異太大的問題，也許未來使用颱風中心位置的經緯度來做更精細的分類或許可以得到更佳的改善。整體來說，天氣分型結合降水迴歸模式的表現，對降水量推估是有改進的表現，但仍需更進一步的調整與改進。

與 Cheng et al.(2010)的研究結果相比，使用天氣分類結合邏輯迴歸和非線性迴歸模式的降水量推估，不論有或無天氣分類下，臺灣的改善程度沒有加拿大地區來的好。因為加拿大的天氣特性相較台灣地區單純，地處廣大的內陸，也無颱風天氣類型降水。另外劉等人(2010)的邏輯斯迴歸分析在降雨機率預報上的應用及校驗分析研究中說到，邏輯迴歸方法屬於迴歸分析的一種，迴歸分析針對穩定的資料有較強的估計能力。這也是冷季降水量推估普遍表現的比暖季降水量推估好的原因，因為冷季環境較穩定且降水變異度小，暖季環境不穩定度大且局地對流系統常帶來強降水但又無法掌握中小尺度移入的天氣系統訊息，讓降水迴歸推估表現不佳。

本研究的天氣分型結合邏輯迴歸和非線性迴歸建模對降水推估結果稍有進步，但仍有很多需進一步

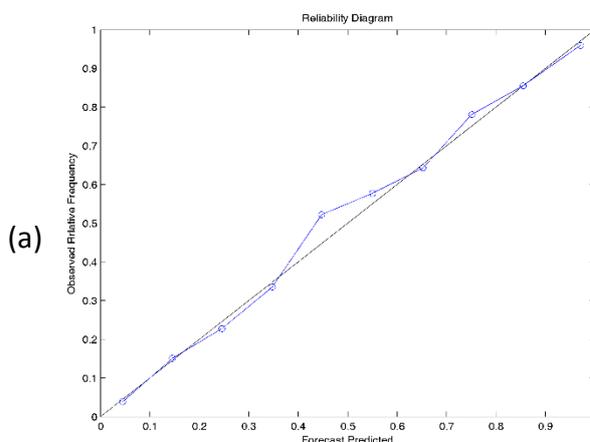
探討改進的空間。對於大雨和局地的熱帶對流降水和滯留鋒的鋒面對流降水掌握不佳的問題，也許可以透過增加天氣變量或是分析較大區域的其他測站來改善，而不是只有臺北測站之天氣訊息，例如可以考慮加入鋒面南下會經過的中國東南沿海一帶的測站，像是香港測站，如此可以掌握更多鋒面的天氣資訊。另外就是未來可以考慮加入其他的天氣變量(例如:大氣不穩定指數 LI、K、 γ ...等參數)做不一樣的分類結果測試。颱風的部分可以參考李和盧(2012)的研究，找出延臺灣海岸線的 300 公里內的颱風中心與測站距離 r 的關係，然後以距離 r 做分類，應該可以使降水推估會有更好的表現。預報因子的部分，未來可以參考中央氣象局梅雨季豪雨預報指引，加入 850hPa 相對溼度、850hPa 的 24 小時高度差、700hPa 相對溼度、700hPa 垂直速度和 200hPa 輻合輻散度，應該有比較好的降水推估結果，因為如此更能確實掌握相對於中、高緯度(例如:加拿大)地區更複雜的天氣系統，並統計出更顯著有效的天氣狀態與降水量之間的關係，提升降水迴歸模型對於暖季梅雨鋒面的推估能力。另外，因為台灣地區受季風影響很大，或許也可以考慮加入季風指數(Monsoon Index)當做預報因子。

除此之外，統計迴歸與資料的準確也有很密切的關係，Wilby et al.(2004)指出若要更進一步得讓統計降尺度的模擬提升的話，是需要較好的全球環流模式(GCM)次網格資訊，尤其是對於極端事件的模擬，例如：熱浪、強降水或是局地的洪災研究，因此未來也可以嘗試使用不同的、更好的 GCM 模式輸出資料來測試。針對研究結果後續的改進，在檢驗的時候可重複做 4 次，如此一來就可以累積出 20 年之檢驗資料，有較足夠和可以信服的資料量來做驗證比較會更好。

如果天氣分類結合統計迴歸模式經由上述的不同方法得到良好的改進，未來這方法可進一步運用在氣候變遷下的降水評估與研究，將不同暖化情境或不同年代的未來全球大氣環流模式(GCM)輸出結果帶入統計模式，分析其推估之降水並做比較和討論。當然這前提是假設在觀測預報量與預報因子之間的關係在氣候變遷下仍然保持不變(Yarnal et al.,2001)。

參考文獻

- Wilby, R.L., S.P. Charles, E. Zorita, B. Timbal, P. Whetton, and L.O. Mearns, 2004. Guidelines for use of climate scenarios developed from statistical downscaling methods, TGICA IPCC August 8, 1-27.
- Cheng, C. S., H. Auld, G. Li, J. Klaassen, B. Tugwood, and Q. Li, 2004: An automated synoptic typing procedure to predict freezing rain: An application to Ottawa, Ontario. *Wea. Forecasting*, 19, 751–768.
- , and Coauthors, 2007a: A synoptic climatological approach to assess climatic impact on air quality in south-central Canada. Part I: Historical analysis. *Water Air Soil Pollut.*, 182, 131–148.
- Cheng, Chad Shouquan, Guilong Li, Qian Li, Heather Auld, 2010: A Synoptic Weather Typing Approach to Simulate Daily Rainfall and Extremes in Ontario, Canada: Potential for Climate Change Projections. *J. Appl. Meteor. Climatol.*, 49, 845–866.
- Davis RE and Walker DR, 1992: An upper-air synoptic climatology of the western United States. *Journal of Climate*, 5: 1449–1467.
- Yarnal B, Comrie AC, Frakes B, Brown DP., 2001.: Developments and prospects in synoptic climatology. *International Journal of Climatology* 21: 1923–1950.
- 劉家豪、張庭槐與洪忠和，2010：「邏輯斯迴歸分析在降雨機率預報上的應用及校驗分析」，天氣分析與預報研討會論文彙編，中央氣象局，臺灣，臺北，258-263 頁。
- 李清騰、盧孟明，2012：「從氣候觀點探討影響臺灣颱風的定義問題」。氣象學報，第四十八卷，第四期，25-38 頁。



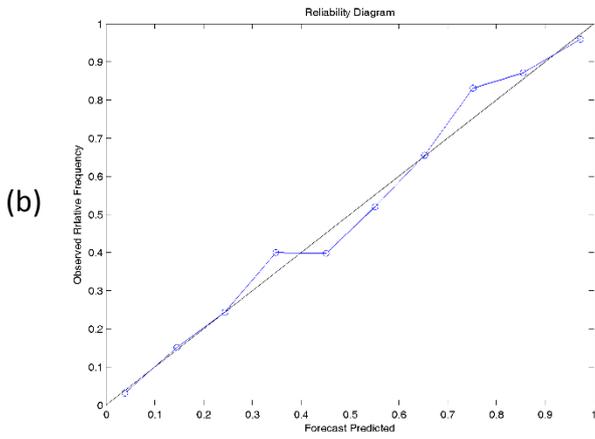
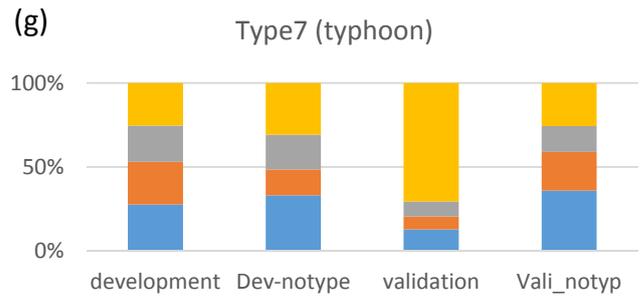
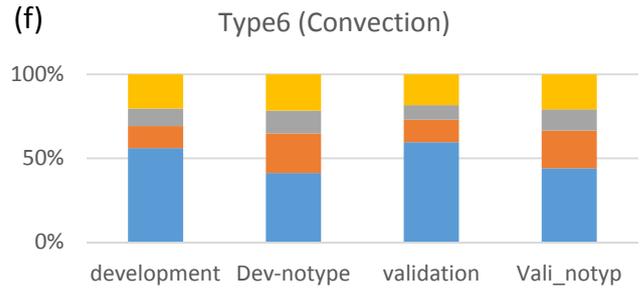
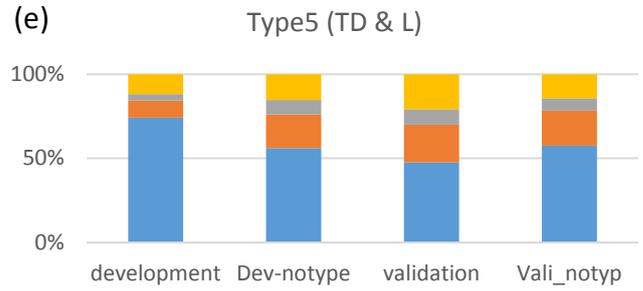
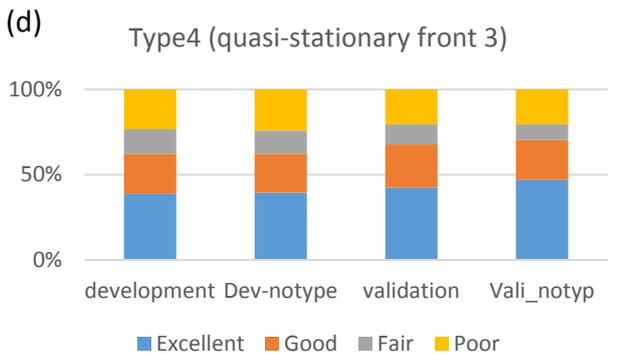
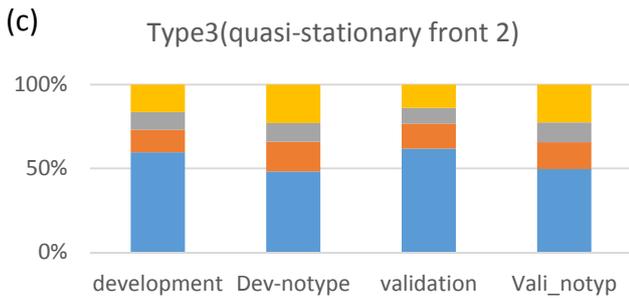
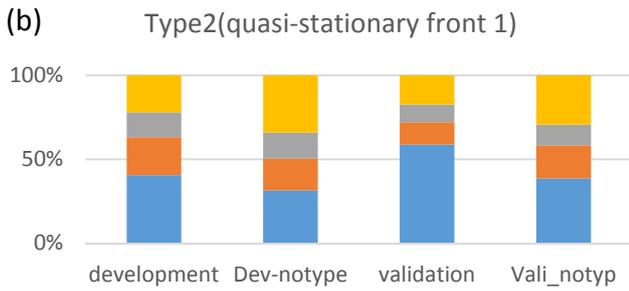
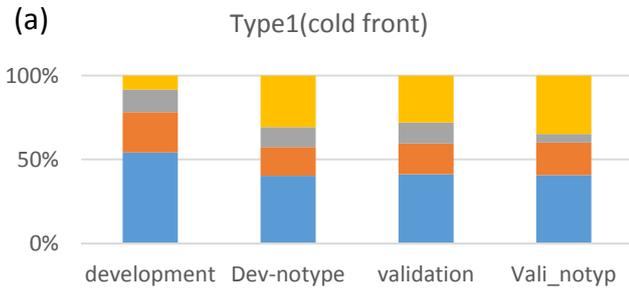


圖 3-1 建模數據集和檢驗數據集於 1992-2012 年的邏輯迴歸降水機率模式可信度圖。a 建模數據集>10mm 機率之可信度圖，b 檢驗數據集>10mm 機率可信度圖。

表 3-1 降水評估標準。Diff=觀測和推估的降水量之絕對值差異(mm)； σ =觀測降水的標準偏差 (mm)。

分類指標	觀測的降水量<5mm	觀測的降水量>5mm
優	Diff \leq 1.5mm	Diff $\leq \sigma/4$ of obs
良	1.5mm < Diff \leq 3.0mm	$\sigma/4$ of obs < Diff $\leq \sigma/2$ of obs
普通	3.0mm < Diff \leq 4.0mm	$\sigma/2$ of obs < Diff $\leq \sigma$ of obs
差	Diff > 4.0mm	Diff > σ of obs



■ Excellent ■ Good ■ Fair ■ Poor

圖 3-2 建模/檢驗數據集與有/無天氣分類的各個天氣類型推估表現圖。a 冷鋒、b 冷季滯留鋒、c 冷季滯留鋒-西南雲系移入、d 暖季滯留鋒、e 熱帶低壓&TD、f 局地熱對流、g 颱風。圖由左至右依序為建模有天氣分類、建模無天氣分類、檢驗有天氣分類、檢驗無天氣分類。藍色代表優、橘色代表良、灰色代表普通、黃色代表差，縱軸為所占百分比數。