

世界氣象預報公司的一週預報與校驗

黃俊翰 林重光 林秉煜 黃椿喜 呂國臣

氣象預報中心

中央氣象局

摘要

在現今資訊爆炸時代，民眾得知氣象資訊的管道不再只是官方氣象預報，而是來自於各樣式APP與民間氣象網站。因此，本文主要在分析現今世界知名四家氣象預報公司：AccuWeather、tenki（日本氣象協會）、Weather.com、WeatherRisk（天氣風險）與 Central Weather Bureau（中央氣象局）所提供的一週溫度預報與降雨預報於台灣四個縣市區域（台北、台中、高雄、花蓮）上的表現。透過將逐日或逐月之高溫、低溫、降雨預報個別分析校驗的方式，再綜合各項特性與相關性，以了解在不同條件下，各家氣象公司的預報在區域與時間上的整體表現情形。

經分析校驗各家氣象公司自2020年二月到七月期間的一週溫度與降雨預報表現後，發現不論在逐日或是逐月的表現上都有著很不一樣的特性。首先，在溫度預報方面：各家氣象公司在低溫的掌握度普遍比高溫來得好，且AccuWeather、WeatherCom、WeatherRisk、CWB都能維持小於2度的MAE。而在降雨預報方面：首先，部分氣象公司的一週預報於前三天與後四天預報表現上有著顯著落差，並且有不同預報模式揉合的現象。其次，在實際降雨日數較多的月份中（如：梅雨季，約三月到五月），普遍有過度預報（over-forecast）的情形發生。另外，透過新導出的「預報穩定準確率（Accuracy of Data Stability Score）」來檢視，可發現各家預報對於不穩定天候（如晴雨交替）的穩定性掌握程度也有所不同。搭配計算另一新導出的「連續性波動平均散度（Mean-Fluctuation-Continuation divergence）」得到月天候狀態的熵，可以更明確掌握日間的天气變化與季節變換下的不穩定性。

透過這樣的綜合診斷與分析各家氣象公司表現，除了藉此建立適當的校驗機制之外，更讓我們得以自我檢視，並更進一步地提升在一週預報上的精準度與穩定性。

關鍵字：氣象預報公司、預報、溫度、降雨、校驗、穩定性、熵

一、前言

本研究中選取現今四家較常被使用在手機 APP 或網頁上的天氣預報公司：AccuWeather、tenki（日本氣象協會）、WeatherCom（WeatherChannel）、WeatherRisk（天氣風險），與中央氣象局（Central Weather Bureau, CWB）來進行一週預報的比較與校驗。針對這五間公司的產品，選擇具有共通的三個指標：高溫、低溫、以及降雨，來作為校驗項目。在地點為代表台灣四大地區的台北（Taipei）、台中（Taichung）、高雄（Kaohsiung）、與花蓮（Hualien）來探討各家預報產品在不同地域上的表現。而用來校驗的實際觀測值為該區域的中央氣象局人工氣象站之資料。但要注意的有四點：1. 由於資料是直接取自於網路，僅作為學術研究之用；2. 這五家氣象預報公司在這四個地區的預報點位是與測站位置之距離不盡相同，約 1 ~ 8 公里不等；3. 所取得的天氣風險公司僅為地區城市尺度，未有鄉鎮資料，

且預報長度為未來六天；4. 未能取得日本氣象協會對於花蓮地區的預報資料。

二、使用資料處理與方法

由於資料是透過網路搜集，且各家的一週預報產出時間不同，預報時間長度也不盡相同。因此，在資料的整集流程上，是選定每日 UTC 上午十點來定時擷取各家氣象預報公司在不同地區之一週預報。擷取後的原始資料，會先整理為該日對未來七天的日預報。其中，礙於並非每家氣象預報公司都有提供完整七日的降雨機率，因此這裡的降雨日預報，是採用判斷各家氣象預報公司的天氣描述內是否有預報會降雨所成立。

此外，將要做校驗的日期之前七天內日預報中對於今日預報，與該校驗日期的觀測值做對齊校驗，此為日校驗資料。如每日的高溫校驗圖（圖 1）所示，若校驗的日期為

06/08，則將前七天（06/01 至 06/07）內對 06/08 的預報取七天溫度值，以倒序的方式與觀測值做比較。

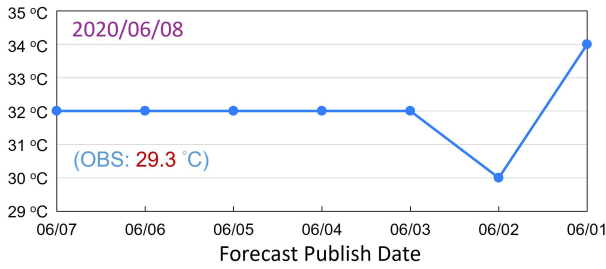


圖 1，每日的高溫校驗圖。將預報資料與實際觀測值做比較，圖中橫軸為了突顯校驗，而將時間倒序呈現。

該流程在高溫、低溫、與降雨預報的校驗是一樣的。而每月會綜合月內的日校驗資料，依照不同的預報項目而去做各自的分析與技術得分評比。最後再用所得到的日預報資料、日校驗資料、以及月校驗資料，來檢視過去預報上的參考。

三、溫度預報校驗

關於溫度預報的校驗，由於四個預報地區彼此的距離大且環境也有所不同，所以會分為高溫與低溫這兩個項目來分別分析各家預報在不同預報天數下於精準度、精確度的表現趨勢。

首先透過觀察日預報校驗，綜合各家表現，發現預報誤差會隨著預報日期的推移而有不定量增加之趨勢。即預報第一天的誤差會比預報第二天的誤差來得大，以後類推。其次，在透過觀察日預報校驗或月預報校驗中的偏差值、平均值與中位數等指標後，都可以發現高溫預報要比低溫預報出現更多被低估的情況。其意味著各氣象公司在低溫的掌握度上，普遍優於對高溫的掌握。詳細將在後面做進一步的介紹。

(一) 高溫預報校驗

在高溫預報校驗的部分，我們可以透過幾個主要指標來分析。首先，從 MAE(平均絕對誤差, Mean Absolute Error) 的角度來看月預報校驗的各個預報日期。如 2020 年 06 月台北地區的高溫預報校驗分布圖(圖 2)中，會發現各家公司在各個預報日期有著穩定性的表現，沒有明顯遞增的趨勢。

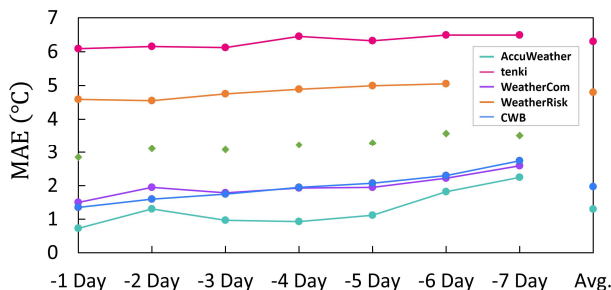


圖 2，各家氣象預報公司在 2020 年 06 月台北地區的高溫預報校驗之 MAE 分布。

綜合表現上，AccuWeather 的表現在每日的平均、或是預報第一天至第七天的平均，都是領先的。

若從月的時間尺度來看的話，以 2020 年 02 至 06 月台北地區為例(圖 3)，各家在 MAE 之於預報第一天至第七天的平均，顯示出不同地區特性。如，在台北、台中、高雄這三個地區上，各家的表現在預報前三天的情況差異不大；各家在高雄地區的表現都很接近，MAE 只小於 2°C。

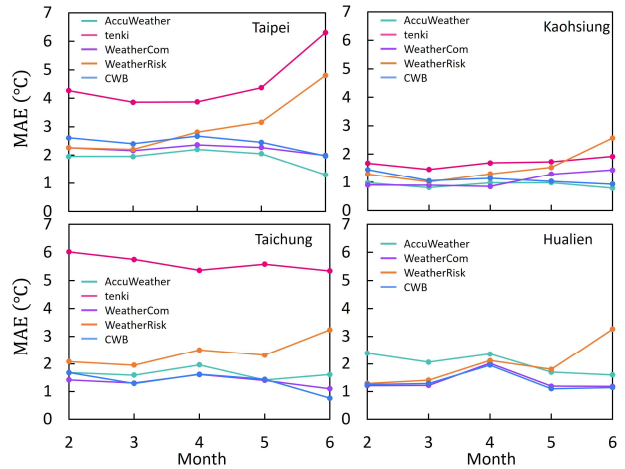


圖 3，各家氣象預報公司於 2020 年 02 至 06 月在四個地區之於預報第一天至第七天的平均分析之高溫預報校驗 MAE。

進一步地拆解來看，在預報誤差小於等於攝氏 2 度 (Error <= 2 degree Celsius) 的比例之七日平均值結果(圖 4)中，除了可以發現剛才從 MAE 中發現的地區特性之外，更反映出各公司 MAE 的組成成分。這可以讓我們更瞭解各家公司的準確性 (Error <= 2 degree Celsius 的部分的比例) 為何。其中，CWB 與 WeatherCom 在不同月分中，Error <= 2 degree Celsius 的比例普遍是較為多的。

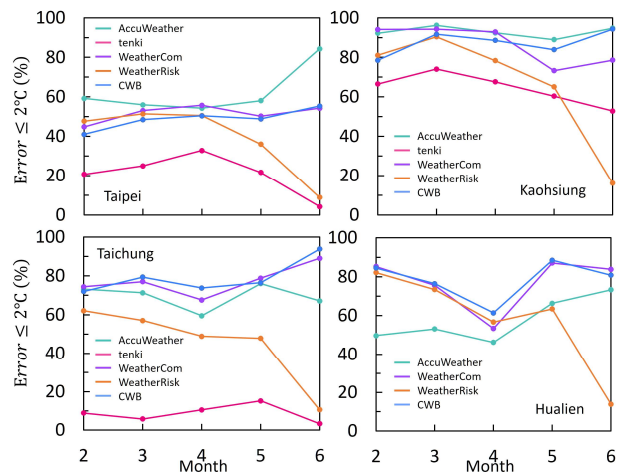


圖 4，各家氣象預報公司於 2020 年 02 至 06 月間在四個地區上，高溫預報於預報誤差小於等於 2 度在預報第一天至第七天的平均分析。

透過上面兩種分析指標：Error <= 2 degree Celsius 反映出了各家公司在日的時間尺度與月的時間尺度下有著不同的相對誤差表現，以及不同的地域特性。然而，這兩種分析指標是以平均的角度來看，所以可以透過搭配「偏差值分析」進一步瞭解各家公司的預報資料特性，以避免只看平均數與單純指標的盲點。

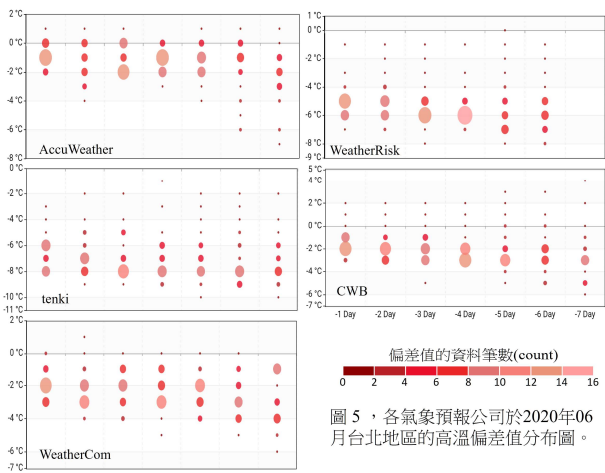


圖 5，各氣象預報公司於2020年06月台北地區的高溫偏差值分布圖。

顏色越淺表示該誤差值的發生次數越多，圖形的半徑也越大；反之，顏色越深表示該誤差值的發生次數越少，圖形的半徑也越小。以 2020 年 06 月台北地區的高溫偏差值分布圖(圖 5)為例，除了可以透過觀察資料點所代表的圓形大小來看資料密度，也可以從資料點的分布寬度來得知預報的誤差範圍。可以瞭解在 MAE 背後真實情況，而不受平均所影響判斷。此外，在前面分析中所觀察到的「預報誤差與預報日期有正相關」的現象，可以理解為模式預報在時間上的誤差累積。

綜合 MAE、Error ≤ 2 degree Celsius 的比例、偏差值分布這三樣指標工具，讓預報員可以瞭解高溫預報在不同預報日期天數、不同氣象公司、以及不同地區上的表現與特性，兼具廣度與深度。

(二) 低溫預報校驗

在低溫預報的表現方面，雖然整體趨勢上仍屬於低估的情況，但偏差值則較高溫預報的偏差值小。可以從 MAE 與預報日期的關係圖(圖 6)與(圖 2)比較得知，

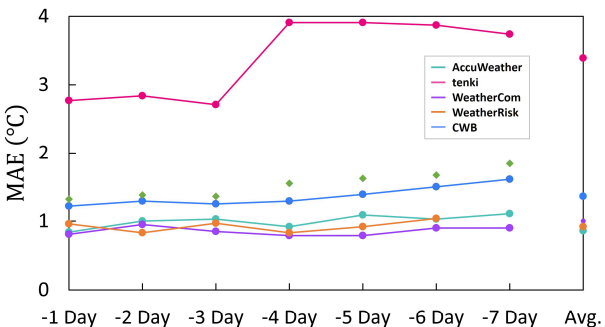


圖 6，各家氣象預報公司在2020年06月台北地區的低溫預報校驗之 MAE 分布圖。

以 2020 年 06 月台北為例，各家公司七日的平均表現都優於高溫預報。其中，AccuWeather、WeatherCom、WeatherRisk 的表現都很接近，且大多只有接近 1 度 MAE。

在不同月份的 MAE 方面，以 2020 年 06 月的預報之 MAE 校驗(圖 7)為例。各家預報在台北地區的表現很相近且不超過 2 度 MAE。在這裡 tenki 的表現也比在高溫部分來的好。其他四家，在其他三個地區上各家的預報表現，雖然有些起伏但誤差小於兩度居多。

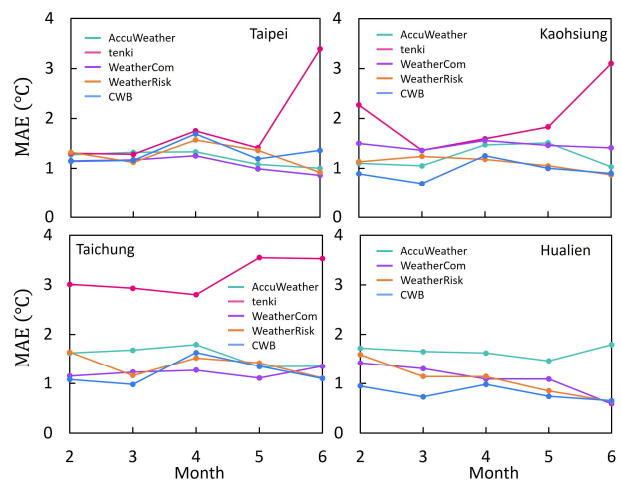


圖 7，各家氣象預報公司於2020年02至06月在四個地區之於預報第一天至第七天的平均分析之低溫預報校驗 MAE。

因為各家公司普遍在 MAE 上的表現顯示多小於 2 度，所以在各月份之誤差小於等於兩度的比例圖(圖 8)中，各家的表現在 80% 以上的比例也較高溫預報來得多。

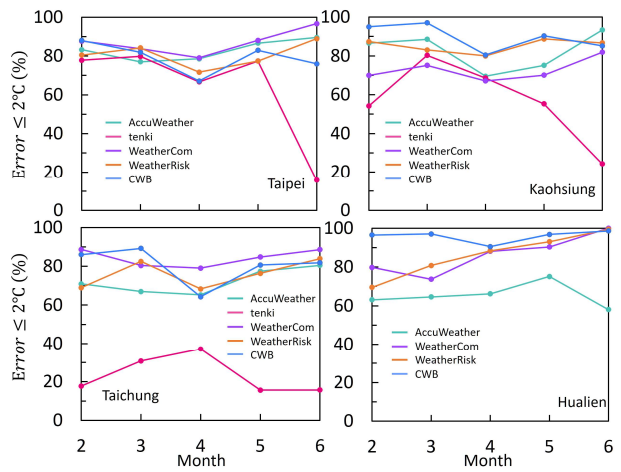


圖 8，各家氣象預報公司於 2020 年 02 至 06 月間在四個地區上，低溫預報於預報誤差小於等於 2 度在預報第一天至第七天的平均分析。

在進一步地透過偏差值分析查看各家預報資料後可以發現到，低估的情況較高溫預報的程度明顯來得輕，且偏差的廣度也較為減少。但也發現，部分公司在預報日期的前三天與後四天是有明顯的表現差距的。這個現象也出現在高溫預報的部分，推測與多模式揉合有關。

四、 降雨預報校驗

在現行作業與校驗分析上，對於降雨預報的處理主要可分為指標分數和格黑校驗這兩大類。在指標分數的部分，常用的有列聯表(表 1, Contingency table; Finley 1884)中的延伸應用: Threat Score (TS)、Equitable Threat Score (ETS)、Probability of Detection (POD)、False Rate (FAR)、Missed Rate (MS)、Bias Score (BS)，來做綜合判斷。在這，我們選用了其中兩項指標分數來協助做最初的校驗: Bias Score (BS) 來判斷是否有過度預報降雨 (over-forecast) 的

狀況；和 Threat Score (TS) 來協助判斷行預報技術的準確度。
註：Threat Score (TS) 是等價於 Critical Success Index (CSI)。

舉例來說，以 2020 年 06 月的各氣象公司對於台北的一週預報，如下圖（圖 9）所示。

圖 9，各家氣象公司於 2020 年 06 月台北地區的 Bias Score 狀況。

Company	-1 Day	-2 Day	-3 Day	-4 Day	-5 Day	-6 Day	-7 Day	Avg.
tenki	3.00	3.12	3.00	3.38	3.25	3.00	3.12	3.12
WeatherCom	2.62	2.25	2.75	2.88	3.00	3.12	3.12	2.82
CWB	3.75	3.75	3.75	1.50	1.75	1.75	1.75	2.57
AccuWeather	1.00	0.88	0.50	0.50	0.62	0.62	0.62	0.68
WeatherRisk	0.12	0.12	0.12	0.25	0.25	0.38		0.21
Avg per Day	2.10	2.02	2.02	1.70	1.77	1.77	2.15	



以預報七天的平均來說，AccuWeather 在 Bias Score 的分數上表現最好，分數為 0.68 最為接近 1.00。而包含 CWB 在內的 WeatherCom 與 tenki 都有過度預報降雨的情況。然而，在計算 Threat Score 時，這五家氣象預報公司的排名則有截然不同的變化，WeatherCom 與 CWB 是為領先。

對於這樣的結果，其實並不感到意外。畢竟降雨預報以目前的技術來說，仍是一項很有挑戰性的部分。但這已經讓我們重新思考，是否有更適合的方式讓預報員在做校驗的時候，能更簡單地瞭解整體狀況。況且，在這裡，我們所要進行的校驗目標除了有不同的氣象公司和地點以外，還多了一個時間維度：預報日期。

(一) Mean-Fluctuation-Continuation divergence

為了找出潛在的適用校驗指標，我們回到了資料面去檢視我們的預報與觀測資料。試著將不同的預報日期和校驗日期，以列表結果做整個月的展開。如下圖（圖 10）2020 年 06 月的各氣象公司對於台北的一週降雨預報中所示：

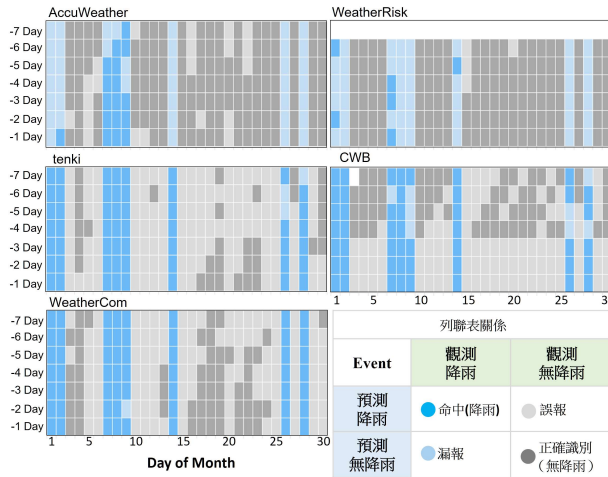


圖 10，將 2020 年 06 月台北站的各家氣象預報公司之降雨狀況，與預報日期作為 Y 軸，該月的每一天為 X 軸。

除了可以透過表格顏色的色系來區分是否觀測有雨以外，也發現到不同的氣象公司，在預報日期七日的表現呈現並非都是連續的現象。另外，也可以瞭解 Bias Score 的表現是受到什麼樣的預報所影響，如：連續的誤報或是漏報的情況。透過這樣的圖表展開，可以讓預報員更容易地去對照與掌握前述的指標分數。

除了前面這三項工具來幫助校驗與判斷之外，在此我們將更進一步地去瞭解降雨的天氣狀況以及更能代表預報降雨的指標。首先，為了瞭解天氣狀況的複雜度，我們先從簡單的天候狀況下出發。若要告訴民眾現在的天氣且天氣型態只有這晴天跟雨天這兩種的話，那在資訊傳遞的角度來說我們只需要傳遞 0 或是 1 即可代表這兩種天氣型態，也可以是為 1 bit。那如果要預報明天的天氣，而晴天與降雨的機率各為 50% 的話，那在這裡我們可以將這樣的不穩定性（uncertainty）與方才的資訊傳遞做結合。也就是說，對於民眾來說實際有用的訊息是為不穩定性的一半，即為 2 以 2 為底的對數。也就是「資訊熵（Information Entropy，後簡稱為熵或 Entropy）」

$$\log_2 2 = 1 \text{ bit} \quad (1)$$

這樣的觀念即為由 Shannon 所提出的「Information Entropy」（Shannon, 1948）。

$$H(P) = -P \log_2(P) \quad (2)$$

這裡 P 表示為機率函數。若有多項狀態時，平均可表示為 (Aurélien Géron, 1995)：

$$H(P) = -\sum_i^N P_i \log_2(p_i) \quad (3)$$

若今天有另外一套系統想一起來描述這天候狀況時後，則可以改用 Cross-Entropy (Peter Dayan, 1995) 來描述：

$$H(P, Q) = -P \log_2(Q) \quad (4)$$

換句話說，也就是從辨識事件機率函數 P 到機率函數 Q 所需要的花費。例如：將機率函數 P 當作是地面真值 (Ground Truth)，而機率函數 Q 可為預報。如此，接下來我們可以得到機率函數 P 與機率函數 Q 之間相對的熵值 (relative entropy)，也就是 Kullback-Leibler divergence (KL-divergence, KL 散度) (Kullback, S.; Leibler, R. A., 1951) 寫為

$$\begin{aligned} D_{KL}(P \parallel Q) &= \text{Cross-Entropy} - \text{Shannon-Entropy} \\ &= H(P, Q) - H(P) \\ &= \sum_i^N P_i \log_2 \left(\frac{P_i}{Q_i} \right) \end{aligned} \quad (5)$$

也就是機率函數 Q 到機率函數 P 的相對熵，或稱為基於機率函數 Q 的分布來敘述機率函數 P 分布的樣本所需要的額外平均熵，或是稱為機率函數 Q 到機率函數 P 的距離。相對熵可以幫助我們從另一個不同的角度去看觀測與預報。但這樣的 KL-divergence 卻存在有一個缺點：

$$D_{KL}(P \parallel Q) \neq D_{KL}(Q \parallel P) \quad (6)$$

也就是 Kullback-Leibler divergence 不具有對稱性，無法針對有序的觀測或是預報所反映。而另一個理論被提出來解決 KL-divergence 不具有對稱性的問題，是由 Jianhua Lin (1991) 提出的 Jensen-Shannon divergence (JS 散度)：

$$\frac{1}{2} D_{KL}(P \parallel M) + \frac{1}{2} D_{KL}(M \parallel P), \text{ where } M = \frac{P+Q}{2} \quad (7)$$

這樣的觀念類似於用資訊半徑 (information radius) 或總熵度的平均 (total divergence to average) 量測兩個機率函數與相似函數 M 的關係。雖然 JS-divergence 的概念解決了原先 KL-divergence 的對稱性問題。但是這樣的相似函數 M 的定義，卻仍不足以讓我們來表達整天天氣狀況的複雜度。

在此，我們參考 KL-divergence 與 JS-divergence 的描述法，針對自身所需重新定義機率函數 Q ，並改進在 JS-divergence 中的相似函數 M 之定義。從 Apollonius's theorem (Godfrey Charles, 1908) 出發，

$$\overline{PQ}^2 + \overline{PS}^2 = 2(\overline{SZ}^2 + \overline{PZ}^2), \quad (8)$$

延伸導出同時符合晴雨狀況、且具有對稱性的描述法，稱為「連續性波動平均散度」(Mean-Fluctuation-Continuation divergence, MFC-divergence)

$$\begin{aligned} &= \frac{1}{\sqrt{2}} \sqrt{D_{KL}(P \parallel Q)^2 + D_{KL}(P \parallel S)^2 - 2D_{KL}(Q \parallel Z)^2}, \\ &= \frac{1}{\sqrt{2}} \sqrt{D_{KL}(P \parallel Q)^2 + D_{KL}(P \parallel S)^2 - 2 \left\{ \frac{D_{KL}(Q \parallel S) + D_{KL}(S \parallel Q)}{2} \right\}^2} \\ &= D_{MFC}(P \parallel Z), \end{aligned} \quad (9)$$

這裡， Q 與 S 為和機率函數 P 等大小的機率函數，且 Q 與 S 的值互為相反。彼此關係如下圖 (圖 11) 所示。

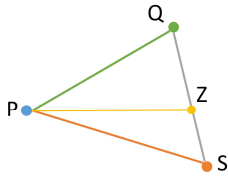


圖 11，機率函數 P, Q, S 與 \overline{PZ} 的關係式。

利用 KL-divergence 中的不對稱性

$$D_{KL}(Q \parallel S) \neq D_{KL}(S \parallel Q), \quad (10)$$

以及 JS-divergence 的概念。求得 $D_{MFC}(P \parallel Z)$ 即為我們所想知道的機率函數 P 於這樣月分布之天候複雜度。其極值如 (11) 式

$$D_{MFC}(R \parallel Z) \leq D_{MFC}(P \parallel Z) \quad (11)$$

此處 $D_{MFC}(R \parallel Z)$ 內的 R 表示值為全有或全無的機率函數。而使用上則會觀察減去 $D_{MFC}(R \parallel Z)$ 的部分，即稱之為「連續性波動平均散度的偏差值」：

$$= D_{MFC}(P \parallel Z) - D_{MFC}(R \parallel Z). \quad (12)$$

舉例來說，將 2020 年 06 月的台北站雨量觀測資料先做簡單的降雨狀況分類。展開後可以圖 (12) 這樣的開式表示：表格中，藍色代表當天有下雨；灰色則代表當天沒下雨；圖中機率函數 P 即為實際觀測降雨狀況。機率函數 Q 與機率函數 S 即為所設計的降雨分布。

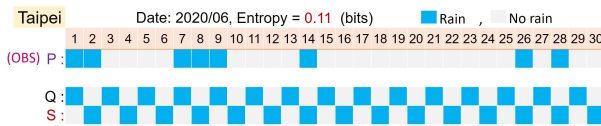


圖 12，將 2020 年 06 月台北站的降雨狀況記錄為 P 。並設計互為相同大小的 Q 與 S 來計算此降雨狀況的連續性波動平均散度之偏差值，為 0.11 (bits) 即 $D_{MFC}(P \parallel Z) - D_{MFC}(R \parallel Z)$ 。

從熵的角度來看，機率函數 Q 與機率函數 S 的值會是很低的。利用找尋機率函數 P 到機率函數 Q 及機率函數 S 的平均波動距離，來表示 機率函數 P 在系統中的複雜度。在這個例子中，所計算出來的 Entropy 是為 0.11 (bits)。而這個值所表示出，在 2020 年 06 月的台北站之天候複雜度在歷史紀錄來說是偏高的，如下圖 (圖 13) 所示。

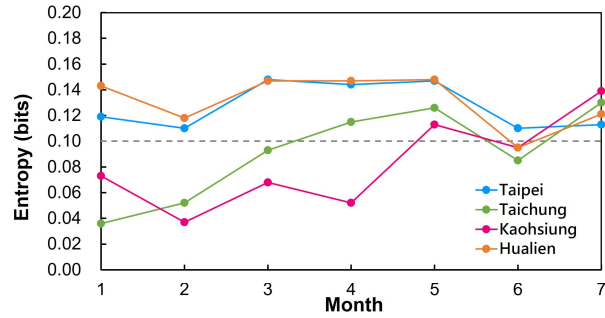


圖 13，計算觀測降雨在 2020 年 01 月至 07 月期間四個地區的連續性波動平均散度之偏差值，即 $D_{MFC}(P \parallel Z) - D_{MFC}(R \parallel Z)$ 。

(二) Accuracy of Data Stability Score

在有了量化整月降雨狀況的資訊後，現在可以開始來檢視降雨預報是否夠準確了。在普遍的使用上就有關於準確度 (Accuracy) 的定義：命中的次數之於總次數。而在針對整月的降雨預報交驗時，我們在此基礎上做了些調整，讓呈現精確度的同時也可以看出對降雨狀況的掌握度。即計算限制必須連續兩日以上，都是在列聯表 (表 1) 中的 A (命中) 以及 D (正確識別) 的表現之於總次數，稱為「預報穩定準確率」(Accuracy of Data Stability Score, ADSS)。

列聯表		
(事件次數)	觀測降雨	觀測無降雨
預測降雨	A 命中	B 誤報
預測無降雨	C 漏報	D 正確識別
預報準確率 (Accuracy)	Accuracy = $\frac{A + D}{A + B + C + D}$	
預報穩定準確率 (ADSS)	ADSS = $\frac{A + D - E}{A + B + C + D} \leq \text{Accuracy}$	
E: 非連續命中/正確識別的次數		

表 1，列聯表。將原先預報準確率計算式的分子減去非連續命中或是正確識別的次數，即為預報穩定準確率。而這樣的預報穩定準確率透過下圖中 (圖 14) 可以發現，

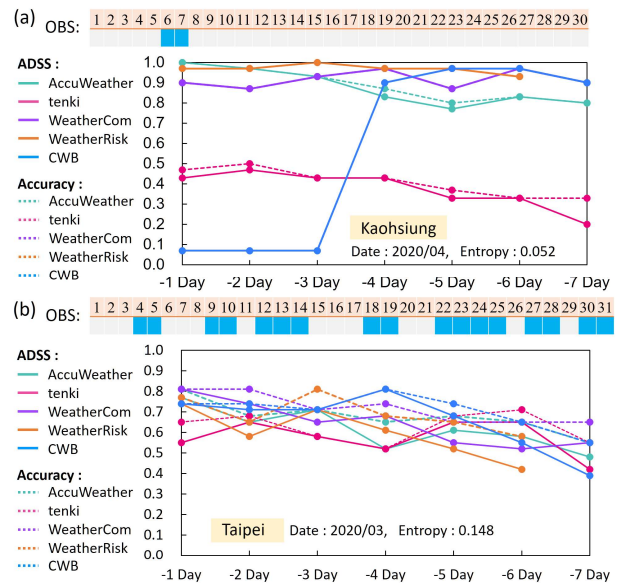


圖 14，(a)在整月的連續波動性平均之偏差值為較低的 0.052 (bits)。(b)整月的連續波動性平均之偏差值為偏高的 0.148 (bits)。在 (a)(b) 中，原先 Accuracy 就比較高的值，在 ADSS 時之差異不大；但是在較低的狀況，則會明顯表現出差異。

Accuracy 的線條(虛線)與 ADSS(實線)之間的差值對於 Accuracy 越高的預報來說普遍越小;而在原先 Accuracy 的值越低或是預報日期越後面時,差值則越大。也就是說,透過 ADSS 我們可以更容易地去看該預報所呈現的 Accuracy 以及在連續天數的掌握度,較不會受到連續漏報或是誤報的事件所干擾。

五、 結論

在資訊爆炸時代,民眾對於氣象預報的需求也不僅只在於三天的預報,要求更即時性(十分鐘、一小時、三小時等…)、同時也更長期性(五天、七天、十四天等…)。而各國氣象市場越來越蓬勃之際,除了逐步提升自己的預報能力以外,也需要建立對應的適當校驗機制與平台。因此,本研究針對現行四家世界知名的氣象公司與中央氣象局本身一同進行七天預報的校驗,並選取台灣四個地區(台北、台中、高雄、花蓮)來做為校驗目標。在高溫預報與低溫預報方面,採用 MAE、誤差小於等於兩度、以及偏差值分析這三種工具來校驗和解析資料。結果顯示五家氣象預報公司在高溫預報的部分,有著普遍低估的現象;在低溫預報的部分,則較高溫預報時來的好。AccuWeather、WeatherCom、WeatherRisk、CWB 在不同地區上都有小於 2 度 MAE 的表現。

在降雨預報校驗的部分,除了常用的得分指標 TS(CSI)、BS 以外,我們引入了連續性波動平均散度(Mean-Fluctuation-Continuation divergence)方法來協助我們分析降雨天候複雜度,以及預報穩定準確率(Accuracy of Data Stability Score)來更進一步地解析各家降雨預報的精確度技術掌握度。在資料結果顯示,部分氣象預報公司在預報前三天與後四天的預報有較明顯地起伏,推測為揉合了至少兩種差

異較大的模式所造成。另外,AccuWeather 在 Accuracy 與 ADSS 的差值相對較小,表示對連續天氣掌握度佳。

在各家氣象預報公司所開發的預報模式,其時間解析度和空間解析度正不斷提升的情況下,區域模式與全球模式的差距可能會越來越小。如何利用校驗資料來精進自己的預報產品與技術,也將會是個重要的議題。透過本研究所建立的分析與校驗機制,讓預報員更能了解自身預報產品的優劣,以求更進一步地改進預報技術與作業流程。也冀望能與世界更多氣象預報公司合作,一同砥礪前進。提供更準確的預報,協助預報使用者創造更多價值。

六、 參考文獻

- [1] Finley, J.P., 1884: Tornado predictions. *Amer. Meteor. J.*, 1, 85-88
- [2] C. E. Shannon, 1948: "A mathematical theory of communication," in *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379-423
- [3] Peter Dayan, 1995: *Neural Computation* 7, pp. 889-904
- [4] Aurélien Géron, 1995: "Praxis Einstieg Machine Learning mit Scikit-Learn und TensorFlow: Konzepte, Tools und Techniken für intelligente Systeme", pp. 173
- [5] Kullback, S.; Leibler, R. A., 1951: "On Information and Sufficiency", *Ann. Math. Statist.* 22, no. 1, 79-86.
- [6] Jianhua Lin, 1991: "Divergence Measures Based on the Shannon Entropy", *IEEE Transactions of information theory*, vol. 37, no. 1
- [7] Godfrey, Charles, Siddons, Arthur Warry, 1908: *Modern Geometry*, University Press, pp. 20

Verification of the Seven-Day Weather Forecast from Weather Companies Worldwide

Chen-Han Huang, Chung-Kung Lin, Ping-Yu Lin, Treng-Shi Huang, Kuo-Chen Lu

Weather Forecast Center,

Central Weather Bureau

Abstract

People presently who can obtain weather forecast easily from hundreds of webs and apps that given by some Non-governmental organization or companies, no longer only from the national weather department. Here we study the properties and the relationship of the seven-day forecast of four weather companies worldwide that are AccuWeather, tenki, Weather.com, WetherRisk, and the central weather bureau of Taiwan in four cities, Taipei, Taichung, Kaohsiung, and Hualien city. By diagnosing the high-temperature, low-temperature, and rain phase forecast separately, we conclude that the forecasting performance is in the different regional conditions and periods.

The result shows significant different performance on daily and monthly diagnostics after analyzing the temperature and rain phase forecast data from Feb. to July 2020. At first, in the verification of temperature forecast, all companies have a much smaller bias in low-temperature than at high-temperature, and the four companies (AccuWeather, Weather.com, WetherRisk, and CWB) have less than 2 degree Celsius in MAE in all months and cities.

In the verification of the rain phase, the performance of some companies had considerable disparity between the first three days and after the fourth day that blend with two or more models. Additionally, they all show over-forecast in the months with numerous rainy days, such as monsoon season or about March to May. Besides, we can measure the steady performance of the accuracy metric by utilizing the Accuracy of Data Stability Score (ADSS) in the unstable weather situation of alternate sunshine and rain. By using the entropy that is derived from Mean-Fluctuation-Continuation divergence (MFC-divergence), it shows not only how stable the monthly weather but the change with different seasons and years.

This paper proposes to set up a verification system for weather companies worldwide, and to develop more precisely and stably forecast models on a seven-day weather forecast.

Key Words: Weather Forecast Company, Forecast, Temperature, Rain, Verification, Stability, Entropy