

東亞地區二週系集降雨預報之校驗分析

周柿均¹ 張惠玲² 陳昀靖¹ 羅存文¹ 陳韋廷¹

中央氣象局科技中心¹

中央氣象局衛星中心²

摘要

現今對於短預報時效的天氣預報可信度大幅提高，然而模式預報誤差會隨著預報時距的增長而放大，為落實無縫隙預報願景，中央氣象局正持續發展 2 至 4 週預報技術，期望能達到幫助長期預報的資訊分析與強化災害預警的應用。

本研究主要針對本局全球系集預測系統(CWB GEPS)與美國國家環境預報中心全球系集預報系統(NCEP GEFS)，進行夏季與冬季二週降雨預報評估，以了解兩模式系集散度的合理性、系集機率預報的可信度與潛在可用性。結果顯示兩模式系集散度夏季較冬季不足；系集機率預報的可信度會隨著降雨門檻提高與預報時間增長而降低；第二週累積雨量系集機率預報的潛在可用性隨著降雨門檻適當地增加而提高，且冬季對東亞地區降雨事件的區辨能力較夏季高。此評估結果除了助於使用者了解模式現階段的預報能力外，同時可提供模式發展者作為模式調整之依據。

關鍵字：系集預測系統、系集離散度、可信度、潛在可用性

一、前言

數值預報系統中存在著非線性過程，且對微小初始條件變化有相當程度的敏感，因此許多不確定性皆可能造成模式預報結果相當大的差異。傳統數值天氣預報為單一決定性預報，期望由最佳的單一初始資料獲得最佳模式預報結果。這種決定性預報無法掌握預報過程中的不確定性，對於所有可能的天氣變化或極端天氣預測有其困難度。不同於單一決定性的論點，系集預報能夠根據多個不同系集成員預報結果對天氣狀態產生近似的機率分布，將不確定性量化，並透過預報機率的方式傳達不確定性給使用者，彌補單一模式預報的不足。然而，系集預報系統的組成，需考量到系集預報成員的預報品質與預報成效，因此必須透過校驗分析了解系集預報是否能有效攫取預報誤差相關的不確定性來源。

近年來各國致力於發展無縫隙預報，希望能夠提供短期天氣至氣候預測完整的預報產品。然而，

預報能力會受限於隨預報時距增加而放大的預報預差，因此 2 至 4 週預報準確性一直是艱鉅挑戰的課題，對模式開發者而言了解數值模式的預報極限相當重要。

本研究欲了解本局全球系集預測系統(CWB GEPS)與美國國家環境預報中心全球系集預報系統(NCEP GEFS)在二週內降雨預報表現，將各 20 組系集成員以各種定量校驗方式進行評估，常見的校驗方法包含 spread-skill relationship、Rank Histogram、Reliability Diagram、Relative Operating Characteristic(ROC)、Brier Skill Score(BrSS)等。數值模式偏差資訊可提供給系統開發者或使用者，針對預報進行調整與預報結果校正，進而提升預報產品的價值。

二、模式與觀測資料

2.1 模式資料

本研究採用本局全球系集預測系統(CWB GEPS)與美國國家環境預報中心全球系集預報系統(NCEP

GEFS)之降雨系集預報場，皆包含 20 組系集成員。初始時間範圍為 2018/07/01 至 2019/02/28，預報時段為第 1 至 14 天。時間解析度分別為 12 小時與 6 小時，自行計算日累積雨量；空間解析度分別為 $0.5^\circ \times 0.5^\circ$ 與 $1^\circ \times 1^\circ$ ，校驗範圍取東經 105 至 150 度、北緯 10 至 40 度。

2.2 觀測資料

雨量資料採用美國國家海洋暨大氣總署(NOAA)提供美國氣候預測中心變形技術(Climature Prediction Center MORPHing, CMORPH)整合之高解析度衛星反演日降水產品，空間解析度為 $0.25^\circ \times 0.25^\circ$ 。

由於本研究採用格點對格點進行校驗。因模式網格點與觀測格點剛好完全錯開，因此共同格點處理方式為將最近四點高解析度觀測格點資料內插至解析度較低的兩模式網格點上，再進行後續的校驗分析。

三、研究方法

3.1 評估預報系統之系集離散度與預報偏差

3.1.1 Rank Histogram

將由小自大排序的預報值與觀測進行比較，紀錄觀測值落在哪個區間，在該區間累計一次，待統計完所有時間與網格點的資料後，即可得到 Rank Histogram。若圖形呈現 U 型分布，表示此系集模式離散程度不足(under-dispersive)；若呈現 A 型分布，表示離散程度太大(over-dispersive)；當圖形平坦時，表示離散程度合理，不確定性與真實情況相符。若圖形的峰值偏向左側，表示真值大部分落在較小值的順位，預報模式有高報(over-forecasting)的情形；反之，或圖形峰值偏向右側，預報普遍有低報(under-forecasting)的現象。因此透過 Rank Histogram 能夠了解系集模式的預報偏差，以及評估系集預報相對於真值的離散程度。

3.1.2 spread-skill relationship

系集離散度-預報能力關係可評估系集成員分歧是否足夠，能否適當地反應預報的不確定性。我們以系集平均之方均根誤差(RMSE)作為預報誤差之量測，以系集散度(ensemble spread)作為散度之量測。

若預報誤差與系集散度有良好的線性關係，表示系集離散程度合理，可利用系集散度大小來預測系集預報系統針對不同可預報度天氣事件的預報能力。

3.2 評估系集機率預報的預報能力

3.2.1 Reliability Diagram

給定某一發生門檻，將系集預報轉換成機率預報，進而了解預報機率與觀測頻率的一致性，即預報系統的可信度(reliability)。若預報機率等於觀測頻率(即信賴曲線在對角線上)，表示完美可信。水平虛線為氣候發生頻率，當觀測頻率接近氣候發生頻率時，表示機率預報不具有解析能力(no resolution)；著色區域與 Brier Skill Score(BrSS)直接相關，用來評估系集機率預報降雨事件時，相對於氣候預報的改善程度。若信賴曲線落在著色區域內，表示 BrSS 大於 0，即機率預報相對於氣候預報具有預報能力；著色區域邊界表示 BrSS 等於 0，不具有預報能力(no skill)。

3.2.2 Relative Operating Characteristic (ROC)

給定一組遞增的預警門檻(表示當預報機率大於等於此預警門檻時，預報有事件發生)，將機率預報轉至二分類預報(即 Yes/No 預報)，並對應觀測實際發生情形建立二維列聯表，得到命中率(Hit rate)與錯誤預警率(False Alarm rate)。不同的預警門檻會得到不同的 Hit rate 與 False Alarm rate，即可繪製 ROC 曲線，曲線下方的面積(Area Under the Curve, AUC)代表預報模式區辨事件發生與否的能力。若 AUC 等於 1 表示完美預報，AUC 大於 0.7 表示此預報模式具有良好的區辨能力。另外，本研究將不同預警門檻之 AUC 繪製折線圖，進一步比較逐預報時距，系集機率預報在不同天氣事件的區辨能力。

四、結果與討論

數值模式預報時距增長至第二週時，預報降雨的不確定性相對地提高許多，因此我們主要以第二週累積雨量，而非日累積雨量，來進行校驗分析。由於東亞夏季與冬季的降雨特徵差異較大，因此將 2018 年 07 月至 10 月定義為夏季，2018 年 11 月至 2019 年 02 月定義為冬季。

定性上，將第二週累積雨量觀測值與所有系集成員的預報值進行排序，並統計觀測落在每一個順位(rank)的發生頻率，即 Rank Histogram。若使用預報模式原始預報值，如圖 1，不論在夏季或冬季，CWB GEPS 與 NCEP GEFS 結果顯示觀測多數落在預報值

最小與最大順位，且最小順位發生頻率高出許多，表示兩個預報模式皆有明顯高報(over-forecasting)以及系集離散度不足(under-dispersive)的問題。多數研究指出數值預報模式在觀測小雨時易有過度預報的系統性偏差。為了驗證預報模式的濕偏差來自於觀測極小雨的樣本，我們僅納入降雨觀測值達到週累積雨量 3.5mm 的統計樣本，再與系集成員預報值進行排序，結果如圖 2。相較於原始預報結果(圖 1)，圖 2 中兩預報模式雖然仍存在系集離散度不足的問題，但在最小順位的比例明顯降低，大幅改善預報系統的濕偏差，表示兩個系集預報系統在觀測極小雨(<3.5mm/week)時，確實有過度預報的情形。

定量上，我們採用系集離散度-預報能力關係 (spread-skill relationship)，如圖 3 所示，圖中每一點系集離散度與 RMSE 為每個初始時間逐網格點預報第二週累積雨量之系集離散度與 RMSE，再進行空間範圍平均。結果顯示，在夏季以 NCEP GEFS 的系集離散度與 RMSE 具有高度線性關係，各點相關係數高達 0.9，決定係數 (coefficient of determination, r^2) 在 0.8 以上，表示 NCEP GEFS 之系集離散度能夠合理地反應第二週累積雨量預報的不確定性；在冬季以 CWB GEPS 的系集離散度與 RMSE 有較良好的線性關係，相關係數高於 0.84，且決定係數高於 0.7。然而，不論是夏季或冬季，CWB GEPS 與 NCEP GEFS 大部分的點落在圖 3 散布圖對角線的上方，表示不論系集平均的 RMSE 較小(天氣系統可預報度較大)或較大(天氣系統可預報度較小)，模式系集離散度容易出現偏小的現象，顯示兩預報模式在第二週累積雨量預報有明顯系集離散度不足(under-dispersive)的問題，與前述 Rank Histogram 結果一致，且夏季較冬季嚴重。

另外，為了瞭解各模式雨量預報之系集離散度與 RMSE 逐預報時間(日)的發展，我們以日累積雨量預報觀察從預報第 1 天至第 14 天系集離散度與 RMSE 變化，如圖 4。結果顯示在夏季，兩預報模式的 RMSE 至第二週有達到飽和的現象，而 NCEP GEFS 在接近初始時間有相對較大的系集離散度，且隨預報時間增加，系集離散度在第二週與 RMSE 相當；在冬季，兩預報模式之系集離散度與 RMSE 發展過程相似，且 RMSE 比夏季小，表示冬季東亞整體的天氣系統可預報度較大。

上述透過定性與定量校驗分析指出 CWB GEPS 與 NCEP GEFS 在第二週累積雨量預報有嚴重不足的系集離散度。接下來，將系集預報轉換至機率預報，透過 Reliability Diagram 了解預報模式針對不同天氣事件發生的可信度表現。圖 5 與圖 6 顯示兩預報模式分別在夏季與冬季降雨發生門檻為 10mm/week 與 50mm/week 的信賴曲線分布。若信賴曲線越接近圖中的對角線，表示機率預報的偏差越小，可信度越高。不論是夏季或冬季，在第二週累積雨量達到 10mm 的機率預報結果(如圖 5)呈現低預報機率低報且高預報機率高報，可信度仍有很大的進步空間。我們發現預報機率为 0%時，對應到觀測發生频率大約為 30%，表示在所有系集成員沒有達到 10mm/week 門檻時，有部分觀測有達到此發生門檻，在 Rank Histogram 中，這些觀測達到發生門檻的統計樣本會排在最大順位，反之，在預報機率为 100%的預報樣本中，觀測發生频率約為 70%，表示系集成員都達到此發生門檻時，有部分觀測沒有達到此門檻，這些統計樣本在 Rank Histogram 會落在最小排序。因此，信賴曲線兩端越偏離對角線且越接近水平，可能隱含系集離散度不足的現象。另外，因 10mm/week 的統計樣本多數落在 sample size 的最右側，顯示觀測小雨時兩預報模式主要呈現高報(over-forecasting)的現象，若降雨門檻提高至 50mm/week 時(如圖 6)，因多數統計樣本落在 sample size 的最左側，表示兩預報模式在週累積雨量較大時會傾向低報(under-forecasting)。

Reliability Diagram 主要提供模式機率預報可信度的資訊，了解預報天氣事件發生時實際觀測發生的情況。相反地，若觀測有天氣事件發生時，我們採用 ROC 曲線下面積(AUC)判斷模式對事件的區辨能力 (discrimination ability)。當 AUC 大於 0.7 時，代表預報系統具有良好的區辨能力(Buizza et al. 1999)。圖 7 顯示 CWB GEPS 與 NCEP GEFS 在不同預警門檻下 ROC 曲線的分布，曲線上每個點代表在預報有事件發生時各預報機率的命中率(Hit Rate)與錯誤預報率(False Alarm Rate)。夏季與冬季的降雨預警門檻不同是依各季降東亞降雨特性而設定，若以相同的預警門檻 10、30、50mm/week 而言，圖 7 顯示在夏季以 NCEP GEFS 的潛在可用性較 CWB GEPS 高，在冬季兩者差

異不大。然而各降雨預警門檻之 AUC 冬季整體皆優於夏季，表示冬季對第二週累積雨量預報的潛在可用性較夏季高，可能與前述提到東亞冬季天氣可預報度較夏季高有關。雖然兩預報模式的區辨能力在夏季表現不是相當良好，但 AUC 仍有 0.6 以上，表示模式對第二週累積降雨仍有一定的預報能力。因此，預報指引開發者需針對預報結果進行調整與校正，使系集模式達到真正的潛在可用性，進而提升機率預報的預報能力。

另外，我們針對夏季與冬季分別選取兩個降雨預警門檻，逐預報時距觀察 AUC 變化，如圖 8 所示。結果顯示夏季整體以 NCEP GEFS 的潛在可用性較高，AUC 維持在 0.7 以上的預報時距比 CWB GEPS 多一天；冬季兩預報模式在有極小雨(1mm/day)的區辨能力相似，當預警門檻提高至 20mm/day 時，兩預報模式 AUC 在 0.7 以上的天數可延長至 10 天以上。因不同的預警門檻代表不同程度的降雨事件，可藉由找出 AUC 最大值訂出最佳的預警降雨門檻。本文主要討論預報模式的潛在可用性，因此關於最佳預警門檻的訂定暫不討論。

五、結論與展望

本研究採用 CWB GEPS 與 NCEP GEFS 針對東亞地區第二週累積雨量預報進行校驗分析。不論在冬季或夏季，CWB GEPS 與 NCEP GEFS 有明顯不足的系集離散度，且在觀測極小雨時，兩預報模式皆有過度預報的系統性偏差。在機率預報方面，雖然兩預報模式可信度不高，但因兩預報模式在本研究測試的預警門檻之 ROC 曲線下面積(AUC)大於 0.6，顯示針對一般中小雨的情況，若能夠透過校正程序修正預報偏差，CWB GEPS 與 NCEP GEFS 第二週累積雨量預報仍有一定的潛在可用性。

本研究採用 2018 年中至 2019 年初之東亞範圍系集預報與機率預報進行校驗評估，但結果可能會受到統計樣本數量與特性影響。未來將針對台灣附近特定降雨個案進行決定性預報校驗，進一步了解數值模式對局地特定天氣事件的預報能力，找出原始系集模式的預報極限，提供降雨預報指引開發者更有價值的參考依據。

六、參考文獻

- 吳佳蓉、汪琮、陳冠儒、張惠玲、洪景山、楊舒芝，2016：WRF 系集預報系統 (WEPS) 之颱風降水與風速預報評估。2016 年天氣分析研討會。中央氣象局。
- 張惠玲、陳冠儒、吳佳蓉、汪琮、洪景山、楊舒芝，2018：臺灣地區 WRF 颱風系集降雨機率預報之評估、校正與經濟價值分析-第一部分：預報評估。大氣科學，46，71-106。
- 黃椿喜、葉世瑄、呂國臣、洪景山，2016：系集定量降水預報方法之探討與分析—系集平均、機率擬合平均與超越之定量降水預報。大氣科學，44，173-196。
- Buizza, R., A. Hollingsworth, F. Lalaurette, A. Ghelli, 1999: Probabilistic Predictions of Precipitation Using the ECMWF Ensemble Prediction System. Wea. Forecasting, 14, 168-189.
- Chang, H. L., H. Yuan, and P. L. Lin, 2012: Short-range (0–12h) PQPFs from time-lagged multimodel ensembles using LAPS. Mon. Wea. Rev., 140, 1496–1516

附圖

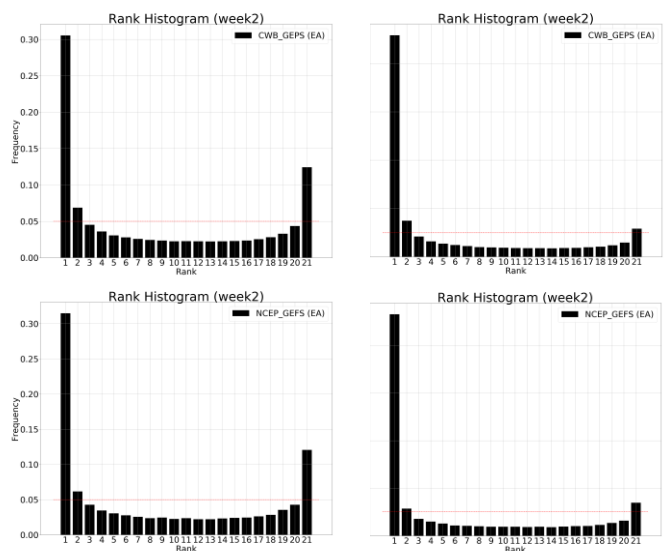


圖 1，第二週累積雨量原始預報的 Rank Histogram。左側為夏季，右側為冬季；上排為 CWB GEPS 的結

果，下排為 NCEP GEFS 的結果。紅色水平虛線代表觀測降雨順位均勻分布在所有順序狀況下的頻率。

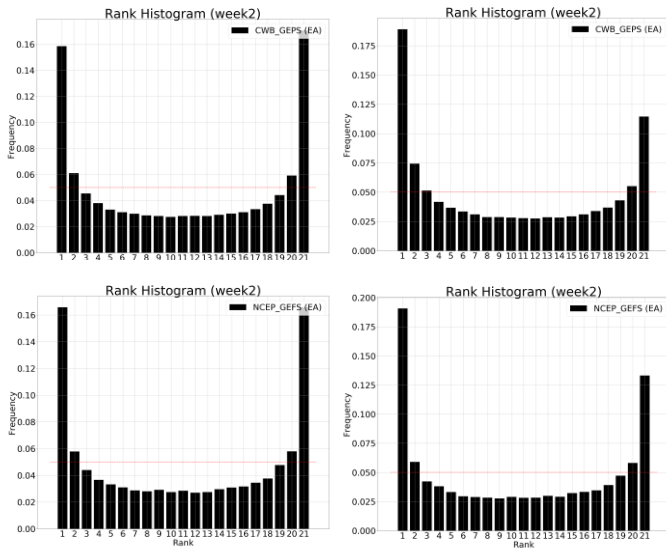


圖 2，類似圖 1，僅納入降雨觀測值達到週累積雨量 3.5mm 的第二週累積雨量預報的 Rank Histogram。

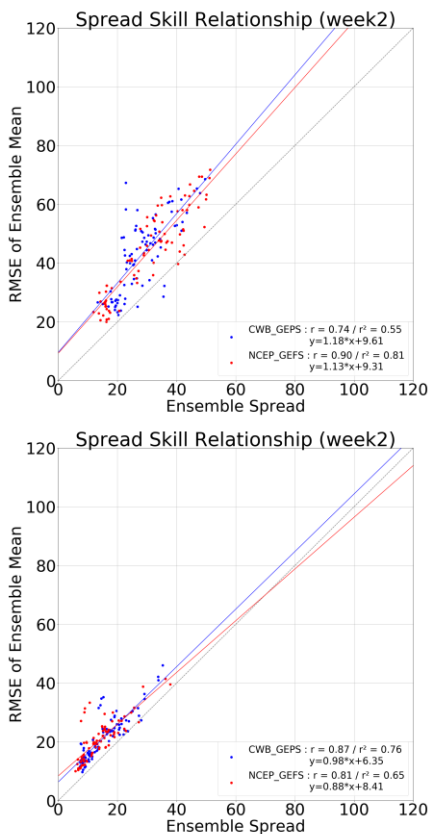


圖 3，系集散度與系集平均之方均根誤差的散佈圖。上圖為夏季，下圖為冬季。圖例標示出線性迴歸線(實線)、相關係數(r)、決定係數(r^2)與線性迴歸方程。藍色為 CWB GEFS，紅色為 NCEP GEFS。

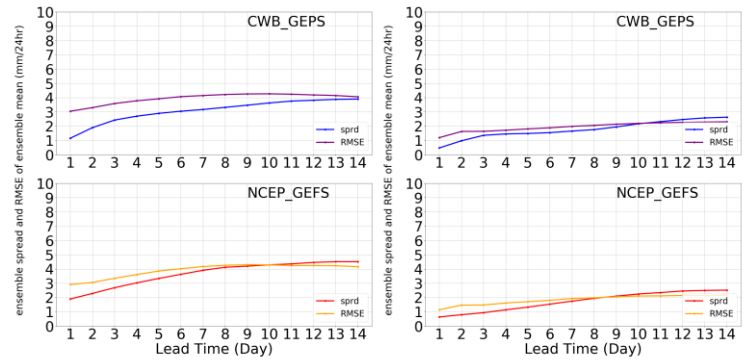


圖 4，逐預報時間第 1 天至第 14 天日累積雨量系集平均之 ensemble spread 與 RMSE 變化圖。上圖為夏季，下圖為冬季。藍色為 CWB GEFS 的結果，紅色為 NCEP GEFS 的結果。

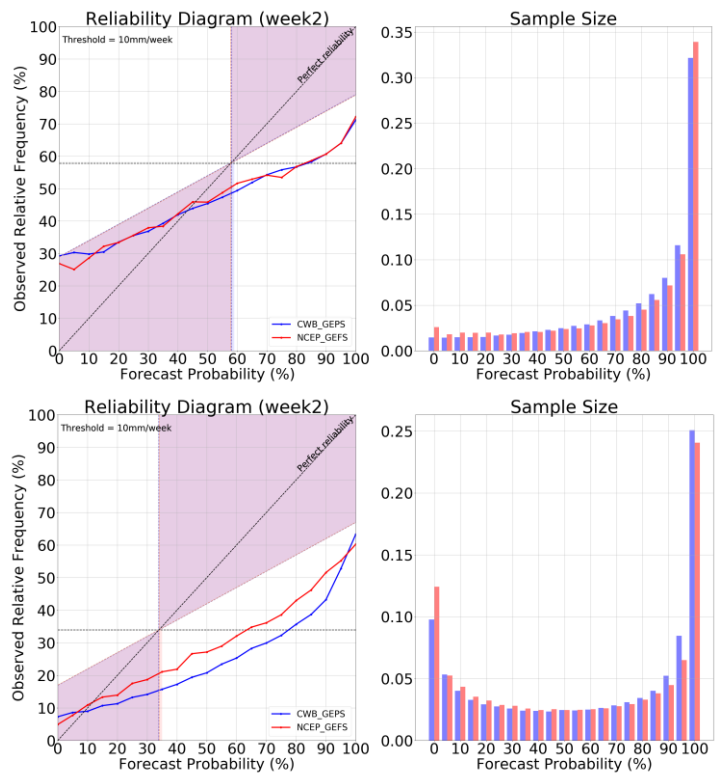


圖 5，第二週累積雨量達到降雨發生門檻 10mm/week 的 Reliability Diagram。上圖為夏季，下圖為冬季。圖中水平黑色水平虛線代表機率預報相對於氣候不具有區辨能力(no resolution)；黑色對角線虛線代表完美預報之信賴曲線分布。著色區域內表示相對於氣候預報具有誤差改善的預報能力(Brier Skill Score > 0)，著色區域邊界不具有預報技術(no skill)

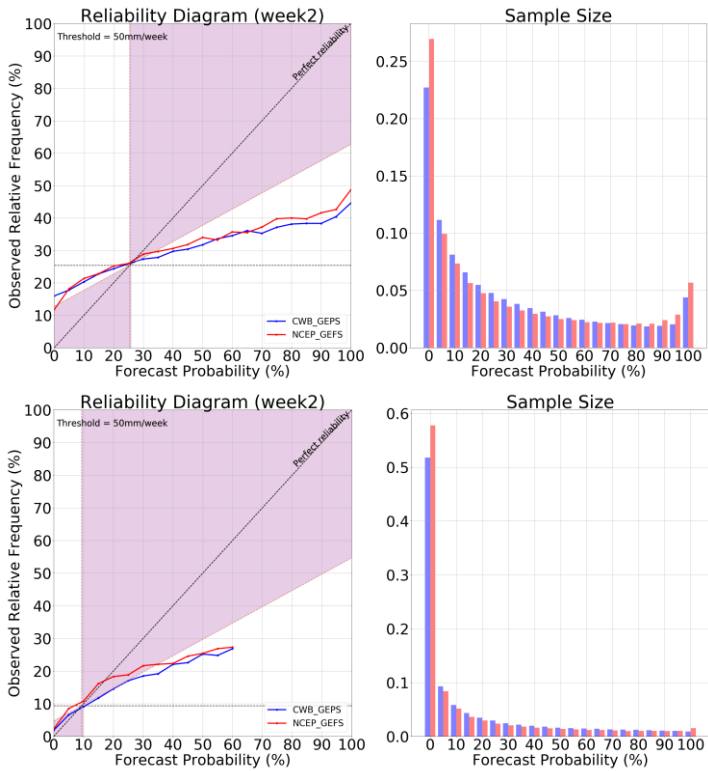


圖 6，第二週累積雨量達到降雨發生門檻 50mm/week 的 Reliability Diagram。其他說明同圖 5。

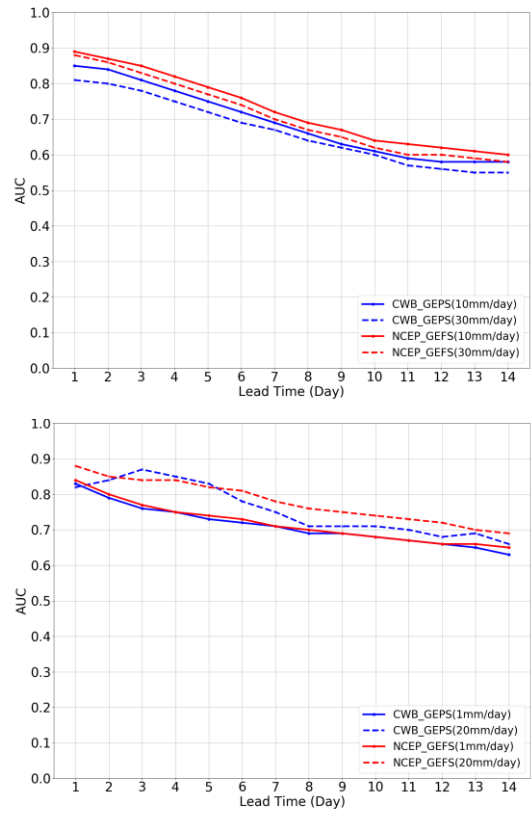


圖 8，逐預報時間第 1 天至第 14 天日累積雨量在不同降雨門檻下 ROC 曲線下面積(AUC)變化圖。上圖為夏季，降雨門檻包含 10、30mm/week；下圖為冬季，降雨門檻包含 1、20mm/week。

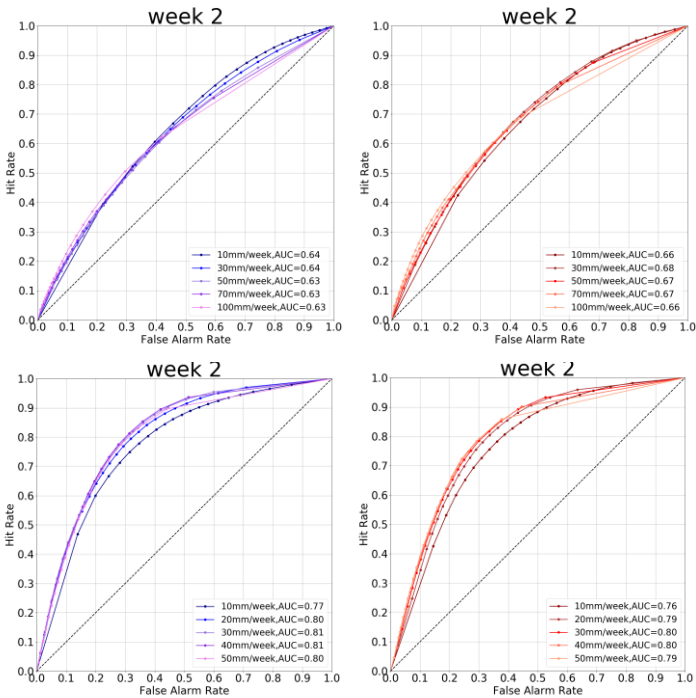


圖 7，第二週累積雨量在不同降雨門檻下的 ROC 曲線。上圖為夏季，降雨門檻包含 10、30、50、70、100mm/week；下圖為冬季，降雨門檻包含 10、20、30、40、50mm/week。圖中黑色對角虛線表示隨機預報。左邊為 CWB GEPS，右邊為 NCEP GEFS。