

東亞地區二週系集降雨預報評估分析

周柿均¹ 張惠玲²

1 中央氣象局科技中心

2 中央氣象局資訊中心

(中華民國 109 年 1 月 16 日 收稿；中華民國 109 年 4 月 27 日定稿)

摘 要

本研究主要是評估中央氣象局(Central Weather Bureau, CWB)與美國國家環境預報中心(National Centers for Environmental Prediction, NCEP)之全球系集預報系統在東亞地區的二週降雨預報表現，包含系集離散度之合理性、系集預報與系集機率預報之評估，同時透過模式後處理由系集導出三種單一預報，並評估其在臺灣鄰近區域之強降雨預報能力，目的在於評比兩系集系統的預報表現，並提供使用者較佳的降雨預報指引。

評估結果顯示：(1)兩模式之系集降雨預報有離散度不足的現象，以夏半年最明顯，且 CWB 模式較 NCEP 嚴重。(2)第二週降雨機率預報的可信度不高，但在冬季有良好的潛在可用性。(3)夏半年和梅雨季可採用機率擬合平均 (Probability-Matched Mean, PM)，而冬半年採用第 80 百分位(Quantitative Precipitation Forecast Percentile, QPF Percentile)作為降雨預報指引。

關鍵字：系集預測系統、系集機率預報、可信度、潛在可用性、系集離散度

一、前言與研究背景

國際上現行氣象預報的範圍，根據世界氣象組織(World Meteorological Organization, WMO)之定義可分為 0 至 2 小時的即時預報、2 至 12 小時的極短期天氣預報、12 至 72 小時的短期天氣預報、72 至 240 小時的中長期天氣預報、10 至 30 天的展期天氣預報、30 天至 2 年的長期預報，以及 2 年以上的氣候預報 (<https://www.wmo.int/pages/prog/www/DPS/GDPS-Supplement5-AppI-4.html>)。其中公認較具有可預報度為 7 天以內的天氣預報與季預報，7 天以內的預報主要參考天氣預報模式，而季預報則是參考氣候預報模式(陳，2010)。

隨著電腦運算效能的提高以及模式資料同化技術與參數化方法的進步，數值模式的預報能力明顯提升，然而從 10 天內的短、中期延長到展期(extended-range)天氣預報的過程中，由於模式不可避免地存在各種數值運算誤差、背景與觀測誤差，且誤差會隨積分時間增加而放大，進而嚴重影響預報結果，因此二週以上的預報不確定性(forecast uncertainty)相當大；且時間尺度介於 10-30 天的季內至季節預報(Sub-seasonal to Seasonal prediction, S2S prediction)同時受大氣初始條件與緩慢

變化之邊界條件所影響，在不同時間與空間尺度的大氣、陸地與海洋系統的交互作用下，S2S 預報是發展無縫隙預報(seamless forecast)中極具挑戰的時間尺度，長期以來被視為「可預報度的沙漠(predictability desert)」，因此各國氣象單位無不致力於提升 10-30 天的展期預報技術，期望提供短期天氣至較長期氣候的完整預報，聯接起天氣預報和短期氣候預測的橋樑，使在預報服務上沒有缺口或接縫，進而達到無縫隙預報的願景(White et al. 2017)，以因應氣候變遷下對預警的需求。

傳統數值天氣預報為單一決定性預報(deterministic forecast)，期望由最佳的初始場獲得最佳的預報結果。然而，預報過程中的種種不確定性，包括觀測資料、模式初始場、資料同化和數值模式的不確定性等，使得決定性預報有時會全然偏離實際的觀測；故完整的預報除了提供未來大氣最可能的狀態外，還需定量地描述預報的不確定性，因此發展出系集預報(ensemble forecast)(Leith, 1974)。

系集預報的概念源自大氣系統的混沌特性(Lorenz, 1963)，而系集預報主要是透過各種不同的擾動方法得到一組能夠合理描述預報過程中種種不確定性的系集成員，以得到未來天氣狀況可能的分布狀況。近年來系集預報技

術發展日趨成熟，加上計算資源充足，已可輕易產生大量的系集成員，但如何產生具有合理離散度的系集成員以反映預報的不確性，以及如何從大量的預報資料中萃取出有價值的資訊，是系集預報應用上的重要關鍵。

本研究透過系集機率預報將不確定性量化，藉以傳達預報的不確定性給使用者，幫助使用者了解天氣系統演變之可能性。然而一般民眾不知道如何以機率預報做決策，對於機率預報的接受度普遍較低，而下游非氣象相關單位的使用者需要決定性降雨預報進行後端應用，例如水利單位進行洪水預報、土石流預報或都市淹水預報等模型時，只能以決定性降雨預報做為其水文模式的初始場，因此機率預報在推廣上相對不易。為此，本研究透過模式後處理方法將系集預報轉為包含不確定訊息的單一預報，包括系集平均(ensemble mean)、機率擬合平均(Probability-Matched Mean, PM)與系集降雨預報百分位(Quantitative Precipitation Forecast Percentile, QPF Percentile)等三種預報，藉由評估此三種系集導出之單一預報在不同季節與不同降雨強度下的預報表現，幫助使用者了解預報產品特性，同時回饋給預報產品發展者，作為產品改進之重要依據。

由於臺灣地區的展期天氣系統與大尺度波動(低頻的熱帶波動)息息相關，因此目前許多研究聚焦在評估數值模式對這些大尺度波動的預報能力。然而，模式掌握到這些具有可預報度的大尺度波動時，並不表示對於臺灣地區的展期降雨具有預報能力。因此，本研究主要是直接評估中央氣象局(Central Weather Bureau, CWB)全球系集預報系統(Global Ensemble Prediction System, GEPS)與美國國家環境預報中心(National Centers for Environmental Prediction, NCEP)全球系集預報系統(Global Ensemble Forecast System, GEFS)在東亞地區的二週降雨預報表現，包含系集離散度之合理性、系集預報與系集機率預報評估，同時透過模式後處理發展系集導出之單一預報，並評估其在臺灣鄰近區域之強降雨預報能力，目的在於評比兩系集系統的預報表現，並提供使用者在不同天氣系統下之最佳預報指引。此論文第二部份主要說明系集模式與觀測資料來源，第三部分介紹校驗方法與各種系集導出之單一預報，第四部份為「系集預報品質」、「系集機率預報評估」與「系集預報導出之單一預報校驗」的結果分析與討論，最後一部分是結論與展望。

二、使用資料說明

(一) 系集模式

1. 中央氣象局(CWB)全球模式系集預報系統(GEPS)

CWB 目前作業中的 GEPS 共有 20 個系集成員，每日啟動兩次預報(00 UTC 與 12 UTC)，預報長度為 45 天，12 UTC 預報時間相同，但僅執行控制組。降雨為 12 小時累積雨量，逐 12 小時輸出；空間解析度為 $0.5^{\circ} \times 0.5^{\circ}$ 。本研究僅針對 00 UTC 預報結果進行分析。

2. 美國國家環境預報中心(NCEP)全球系集預報系統(GEFS)

NCEP 所發展的 GEFS 全球系集預報系統可提供長達 16 天之預報，由 20 組系集成員組成，每日啟動 00 UTC 與 12 UTC 兩次預報，降雨為 6 小時累積雨量，逐 6 小時輸出；空間解析度為 $1^{\circ} \times 1^{\circ}$ 。本研究僅針對 00 UTC 預報進行分析。

(二) 觀測資料

雨量資料採用美國國家海洋暨大氣總署(National Oceanic and Atmospheric Administration, NOAA)提供美國氣候預測中心變形技術(Climate Prediction Center MORPHing, CMORPH)整合之衛星反演降水產品，利用低軌道衛星觀測之微波估算衛

星降水率，若微波觀測資料不足時則使用地球同步軌道衛星觀測之紅外線資料，最後將衛星反演降水率資料整合成一筆完整且連續的全球資料(Joyce et al., 2004)。時間解析度為 24 小時，空間解析度為 $0.25^{\circ} \times 0.25^{\circ}$ 。

本研究僅採用一年的校驗期，是因為氣象局於 2018 年 6 月完成 GEPS 機率預報系統之建置，並透過統計後處理方法，作業化產製系集導出之單一預報，因此本研究之校驗起始時間為 2018 年 7 月。受限於可採用的資料時間不長，加上考量模式在不同季節之預報表現與特性有明顯的差異，為使不同季節的樣本數差異不要過大，故本文使用 2018 年 7 月至 2019 年 6 月完整一年之兩組系集預報與觀測資料進行校驗，並將 2018 年 7 月至 11 月定義為夏半年，2018 年 12 月至 2019 年 4 月定義為冬半年，另將 2019 年 5 月至 6 月獨立定義為梅雨季。校驗空間範圍分為兩個部分：東亞地區範圍取東經 105 至 150 度、北緯 10 至 40 度；臺灣鄰近範圍取東經 115 至 130 度、北緯 15 至 30 度。

由於本研究採用格點對格點進行校驗，由於預報與觀測的空間解析度不同，且模式網格點與觀測格點剛好完全錯開，因此以較低解析度之預報

格點為校驗格點：我們分別針對兩系集模式將預報格點相鄰四點之觀測格點資料內插至解析度較低的預報格點，後續在共同格點上進行校驗。

(三) 個案資料

為評估系集模式對臺灣鄰近區域較強降水系統的預報表現，本研究在「系集預報導出之單一預報校驗」僅挑選較強降雨事件作為校驗樣本；因校驗期不長，因此只要累積雨量期間包含強降雨個案均納入校驗樣本，目的在於盡可能避免統計樣本不足造成取樣誤差(Sampling error)，以提高校驗結果的統計代表性。所採用的校驗樣本在夏半年為 31 報，冬半年為 27 報，梅雨季為 21 報。

三、方法說明

(一) 主要校驗方法

(1) 序位直方圖(Rank Histogram)

逐格點將預報值與觀測由小自大進行排序，紀錄觀測值落在哪個序位，在該序位累計一次，待統計完所有時間與網格點的資料後，即可得到 Rank Histogram。Rank histogram 主要是用以評估系集預報的離散度能否合理地涵蓋到真實觀測，達到所有成員可以等機率預報到真實觀測的一致性，因此當圖形平坦時，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)。若圖形的峰值偏向左(右)側，表示真值大部分落在較小(大)值序位，預報模式有高報(低報)的情形 (Statistical Methods in the Atmospheric Sciences, figure 8.26; Wilks 2011)。因此透過 Rank Histogram 能夠了解系集模式的預報偏差，以及系集離散度 (ensemble spread) 是否合理。

(2) 系集離散度 - 預報能力關係 (Spread-Skill Relationship)

Spread-Skill Relationship 主要用以評估系集離散度能否合理地反應預報不確定性，亦即評估預報系統在系集離散度較小/大的情況下，是否具有較小/大的預報誤差或較好/差的預報技術 (forecast skill; Kalnay and Dalcher 1987)。因為系集平均可用來表示系集成員的中間趨勢 (central tendency)，因此系集平均的方均根誤差 (Root Mean Square Error, RMSE) 較小(大)，表示系集模式對某個天氣系統的預報誤差較小(大)，系集模式能夠(無法)較準確地預報此天氣系統，亦即，此天氣系統的可預報度較大(小)。如果預報系統具有合理的系集離散度，那麼當天氣系統

的可預報度越大(小)時，系集離散度理當反映出預報的不確定性，而隨之變小(大)。我們將系集平均之 RMSE 與系集離散度以散布圖(scatter plot)的形式呈現，若統計樣本多數散布於對角線上方，代表系集離散度不足(under-dispersive)，反之，表示系集離散程度太大(over-dispersive)。如果系集離散度與預報誤差有良好的線性關係，且量值相當，則我們可以透過系集離散度的大小來預測預報誤差的大小；換言之，若該次預報的系集離散度小，代表所預報之天氣系統的不確定性比較小，即可預測系集預報系統在該次的預報誤差較小，預報表現較佳。

(3) 可信度圖(Reliability Diagram)

若給定某一發生門檻，統計樣本中達到該門檻之系集個數與總系集個數之比例，可計算出預報機率(forecast probability)；接著對應各預報機率之滿足該門檻之預報樣本中，實際觀測亦達到該門檻個數之比例，即為觀測發生頻率(observed relative frequency)。透過各預報機率與對應之觀測發生頻率連線得到信賴曲線分布，可進一步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)。

若各預報機率皆等於觀測頻率，

信賴曲線會落在可信度圖對角線上，表示此預報系統完美可信。若信賴曲線普遍落在對角線上下方，表示預報機率通常大(小)於觀測發生頻率，說明模式機率預報為過度預報(預報不足)；若信賴曲線在高機率預報高(低)報、低機率預報低(高)報，顯示機率預報具有過度自信(自信不足)的情形(Statistical Methods in the Atmospheric Sciences, figure 8.8a; Wilks 2011)。

(4) 相對操作特徵曲線(Relative Operating Characteristics, ROC)

不同於 Reliability Diagram 以預報為條件機率之條件判斷，ROC 技術得分是以觀測作為條件機率的條件判斷。給定一組遞增的降雨門檻(表示當預報機率大於等於此降雨門檻時，視為預報有事件發生)，將機率預報轉至單一預報(即 Yes/No 預報)，並對應觀測實際發生情形建立二維列聯表(Contingency Table; Finley 1884)，進一步得到命中率(Hit rate)與錯誤預警率(False Alarm rate)，計算方式詳見附錄。將各預報機率的 Hit rate 與 False Alarm rate 打點連線，即可繪製 ROC 曲線。曲線下方的面積(ROC area)代表預報模式區辨事件發生與否的能力(Discrimination Ability)，是判定預報系統能力的一個重要指

標。若 ROC area 等於 1 表示完美預報；若曲線落在對角線上，ROC area 等於 0.5，常被視為隨機預報；若 ROC area 大(小)於 0.5，表示此預報系統有(無)區辨能力。Buizza et al. (1999)認為 ROC area 大於 0.7 表示此預報模式具有良好的區辨能力，故後續校驗分析依據 ROC area 大於 0.7 作為判斷模式是否具有預報能力的標準。

(5) 性能圖(Performance Diagram)

Performance Diagram(Robber 2009)可同時顯示偵測率(Probability of Detection, POD)、成功預報率(Success Ratio, SR)、預兆得分(Threat Score, TS)、偏倚得分(Bias Score, BS)等四種校驗分數(計算方式詳見附錄)，使校驗者可以整合性地去評估模式的預報表現。曲線由左下至右上表示 TS 得分逐漸增加，越接近 1 表示預報結果越好(Wang 2014)，因此性能圖的右上角代表完美預報，左下角代表無預報技術。校驗結果落在對角線上表示 BS 得分為 1，代表無偏差之預報；落在對角線上(下)方表示 BS 得分大(小)於 1，代表過度預報(低報)(Wilks 1995)。

(二) 由系集導出之單一預報

(1) 系集平均

系集平均具有濾除非線性誤差的特性，有效減少模式預報的誤差，

通常可以提供更準確的預報結果(Buizza and Palmer 1998)，因此常被使用在系集預報導出之單一預報。雖然系集平均可以大致反應出接近實際的降雨空間分布(rainfall pattern)，但平均過程會將預報結果平滑，容易造成大雨低估、無雨過度預報的問題(Ebert et al. 2001)。

(2) 機率擬合平均

為了改善系集平均造成極值被平滑掉的缺點，Ebert (2001)發展機率擬合平均(PM)方法：假設系集平均之降雨排序分布與真實情況一致，以系集平均大小排序之空間分布為基礎，將所有格點之系集成員打散後進行大小排序，依序以系集成員個數分組，並取各組其中位數作為該排序之預報值，依序重新填回系集平均排序之空間分布，達到重新分配降雨頻率之作用，使得降雨極值能被維持，這在防災應用上相當有用。由於 PM 的過程中不會改變系集平均降雨排序之空間分布，因此 PM 與系集平均的降雨空間分布會完全相同，僅有降雨量值上的差異，具有改善系集平均平滑掉極值與改善無雨過度預報之成效，進而得到更接近真實的降雨率。

(3) 系集降雨預報百分位

QPF percentile 的做法為逐格點

將系集成員預報值由小至大排序，以某個百分位作為雨量預報。因為系集百分位為逐格點導出降雨預報，因此降雨分布與系集平均不相同，高系集百分位的降雨容易攫取到極端強降雨發生，過去研究(黃等, 2016)指出第 80 百分位的 QPF 在臺灣地區有較好的降雨預報表現，可作為良好之預報指引，故本研究採用第 80 百分位作為降雨預報指引之重要參考，後續簡稱 PR80。

四、結果與討論

(一)系集預報品質

(1) 序位直方圖

不論在夏半年、冬半年或梅雨季，GEPS 與 GEFS 在最小序位的發生頻率明顯高出許多(圖 1)，表示兩預報模式明顯高報，且以冬半年最嚴重。然而，我們發現兩模式容易在觀測無雨的情形預報出大範圍的毛毛雨，這些毛毛雨不是我們關切的降雨事件，但佔有很大比例，因此我們移除雨量觀測未達 3.5 mm/week 的統計樣本，重新進行序位直方圖分析，結果兩模式在最小序位的比例明顯降低(圖 2)，表示圖 1 中的嚴重過度預報主要是模式大範圍的毛毛雨所造成。整體而言，

兩模式降雨預報之系集離散度明顯不足。

(2) 系集離散度-預報能力關係

由於數值模式預報大範圍毛毛雨並非預報不確定性來源，而是模式無法合適處理所導致，因此該部分與 Rank Histogram 做法相同，僅採納週累積雨量達到 3.5mm/week 的樣本，避免像小雨等非我們關注之降雨型態主導結果。結果顯示在夏半年(圖 3a)與梅雨季(圖 3c)以 GEFS 的系集散度與 RMSE 具有高度線性關係，各點相關係數(Correlation Coefficient)高達 0.9，決定係數(Correlation of Determination)在 0.8 以上(系集散度、RMSE、相關係數與決定係數計算方法詳見附錄)，表示 GEFS 之系集散度能夠合理地反應第二週累積雨量預報的不確定性；在冬半年(圖 3b)兩系集模式的系集散度與 RMSE 相關係數均高於 0.8，且決定係數在 0.65 以上。然而，GEPS 與 GEFS 大部分的統計樣本落在圖 3 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)，模式系集散度均呈現偏小的現象，顯示兩模式系集離散度普遍不足，與前述 Rank Histogram 結果

一致。

另外，我們也評估不同預報前導時間(lead time)的系集平均 RMSE 與 ensemble spread 變化。相同地，為了移除大部分的毛毛雨事件，我們僅採納日雨量大於 0.5mm/day 之

樣本，結果(圖 4)顯示兩系集模式在二週內日雨量預報皆有系集離散度不足的情形：(1)在預報模式方面，以 GEPS 較 GEFS 嚴重；(2)在季節方面，以夏半年最明顯，冬半年最輕微。

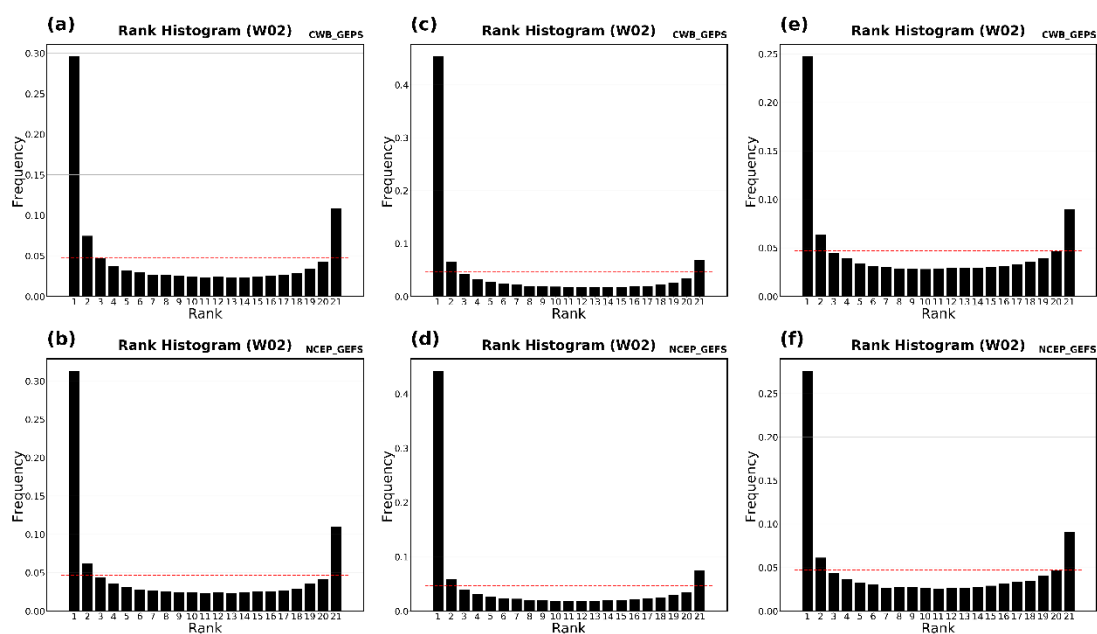


圖 1、第二週累積雨量原始預報的序位直方圖，其中(a)、(b)為夏半年，(c)、(d)為冬半年，(e)、(f)為梅雨季；上排為 GEPS 的結果，下排為 GEFS 的結果。黑色直條表示觀測於各順位樣本數占總樣本之比例，紅色水平虛線代表觀測降雨順位均勻分布在所有順序狀況下的頻率。

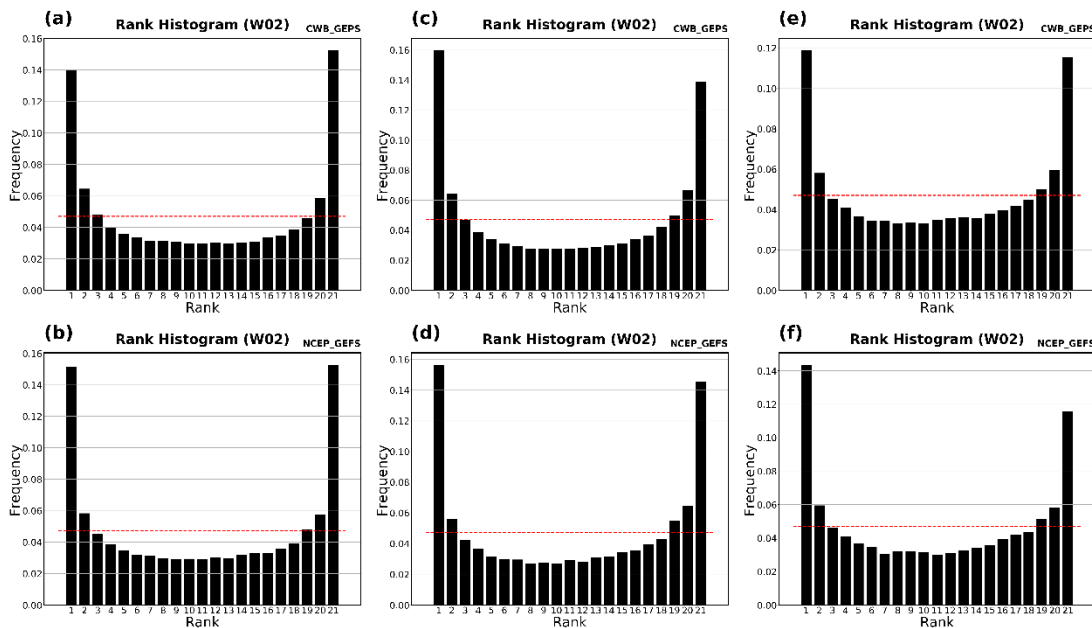


圖 2、同圖 1，但僅納入降雨觀測值達到 3.5mm/week 的樣本。

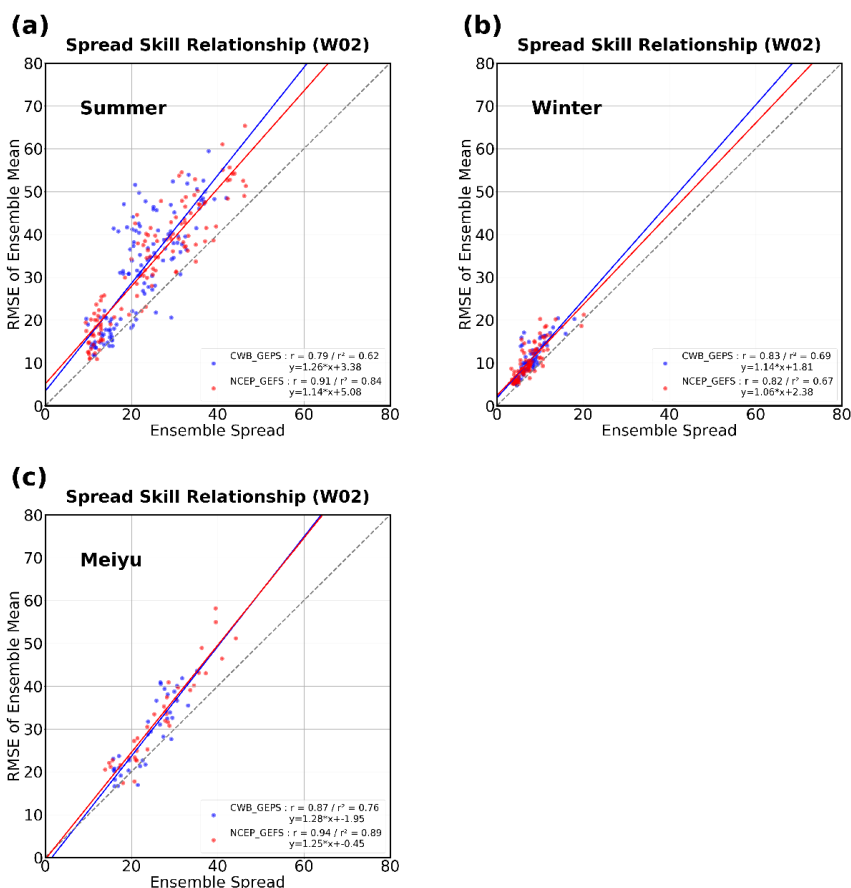


圖 3、系集離散度與系集平均之方均根誤差的散佈圖，其中(a)為夏半年，(b)為冬半年，(c)為梅雨季。圖例標示出線性迴歸線(實線)、相關係數(r)、決定係數(r^2)與線性迴歸方程。藍色為 GEPS 的結果，紅色為 GEFS 的結果。

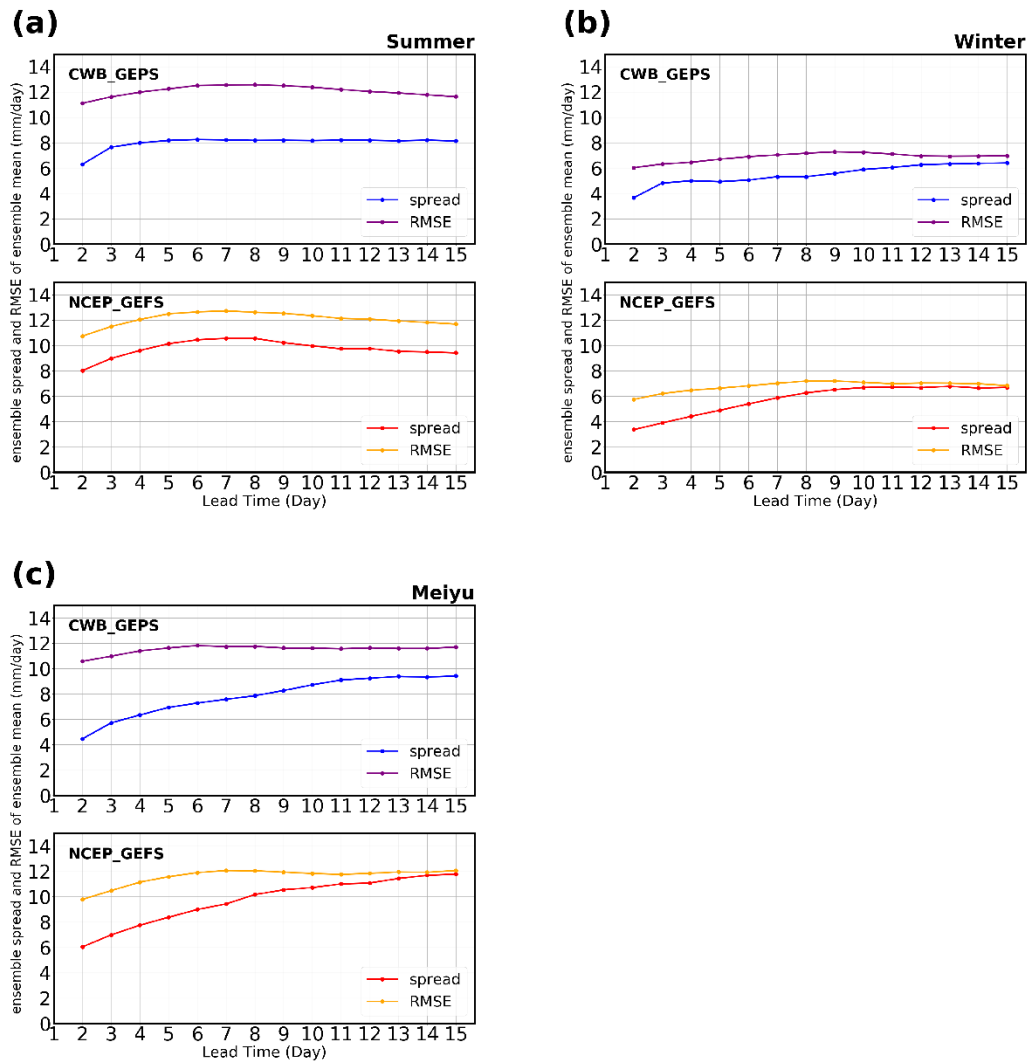


圖 4、前導時間第 2 天至第 15 天日累積雨量系集平均之 RMSE 與 ensemble spread 變化圖，其中(a)為夏半年，(b)為冬半年，(c)為梅雨季；上(下)圖為 GEPS(GEFS)的結果，藍色(紅色)為 ensemble spread，紫色(橘色)為系集平均的 RMSE。

(二)系集機率預報評估

(1) 可信度圖

評估 GEPS 與 GEFS 的機率預報可信度時，若信賴曲線兩端越偏離對角線且越接近水平(垂直)，表示該系統具有過度自信(自信不足)的特性，對於系集模式而言，即表

示模式具有系集離散度不足(過大)的情形；若信賴曲線大多數落在著色區域，則模式的整體預報能力越好，Brier skill score(BrSS)越高(BrSS 計算方式詳見附錄)。結果顯示(1)模式在夏半年(圖 5a)與梅雨季(圖 5c)顯示系集離散度不足，且以低於

觀測發生頻率之預報機率具有較佳的預報能力(BrSS 大於 0)，但整體而言機率預報可信度不高；(2)因圖 5b 的 Sample Size 顯示統計樣本大部分落在低預報機率，且樣本之觀測發生頻率也明顯較其他兩個季節低，表示在冬半年模式不容易預報週累積雨量達到 30mm 降雨事件，可信度圖的信賴曲線容易落在對角線下方，表示系集機率預報有高報的情形。另外，比較兩模式的信賴曲線分布，以 GEPS 高報的情況普遍較 GEFS 嚴重，同樣以低預報機率具有較佳的預報能力。

(2) 相對操作特徵曲線

圖 6 顯示(1)在模式方面，雖然夏半年 GEFS 的 ROC 面積略大於 GEPS，整體而言 GEPS 與 GEFS 在相同季節的 ROC 面積差異不大，表示針對第二週累積雨量兩模式的潛在可用性相近；(2)在季節方面，雖然 GEPS 與 GEFS 的區辨能力在夏半年表現較差(圖 6a, b)，但 ROC 面積仍有 0.6 以上，表示系集模式的第二週累積降雨預報優於隨機預報，有一定程度的可用性，但偏差校正後的預報能力有限。兩模式在冬半年的 ROC 面積皆大於夏半

年與梅雨季，表示冬半年的模式潛在可用性明顯較高，此與圖 4 顯示東亞冬半年天氣可預報度較高的結果一致；(3)在降雨門檻方面，GEPS 與 GEFS 在冬半年與梅雨季的潛可用性呈現逐降雨門檻提高而增加(圖 6c-f)，表示兩模式在冬半年與梅雨季期間降雨強度較大的降雨事件可能有較優的區辨能力。

另外，我們針對夏半年、冬半年與梅雨季分別選取第 50 與 80 百分位的觀測降雨強度作為門檻，分析不同降雨門檻之 ROC 面積隨預報前導時間之變化(圖 7)，進一步探究系集機率預報具有潛在可用性的預報天數。結果顯示：(1)GEFS 之 ROC 面積高於 0.7 的預報天數比 GEPS 長，換言之，GEFS 的潛在可用性優於 GEPS；(2)不同季節中以冬半年之潛在可用性最高，兩系集模式之可預報天數皆可長達 10 日以上，再次驗證冬半年可預報度較高；(3)在冬半年與梅雨季，當預報前導時間較長時，較強降雨門檻之 ROC 面積略大於較小降雨門檻，與前述週雨量預報之 ROC 結果一致，此現象在夏半年則不明顯。

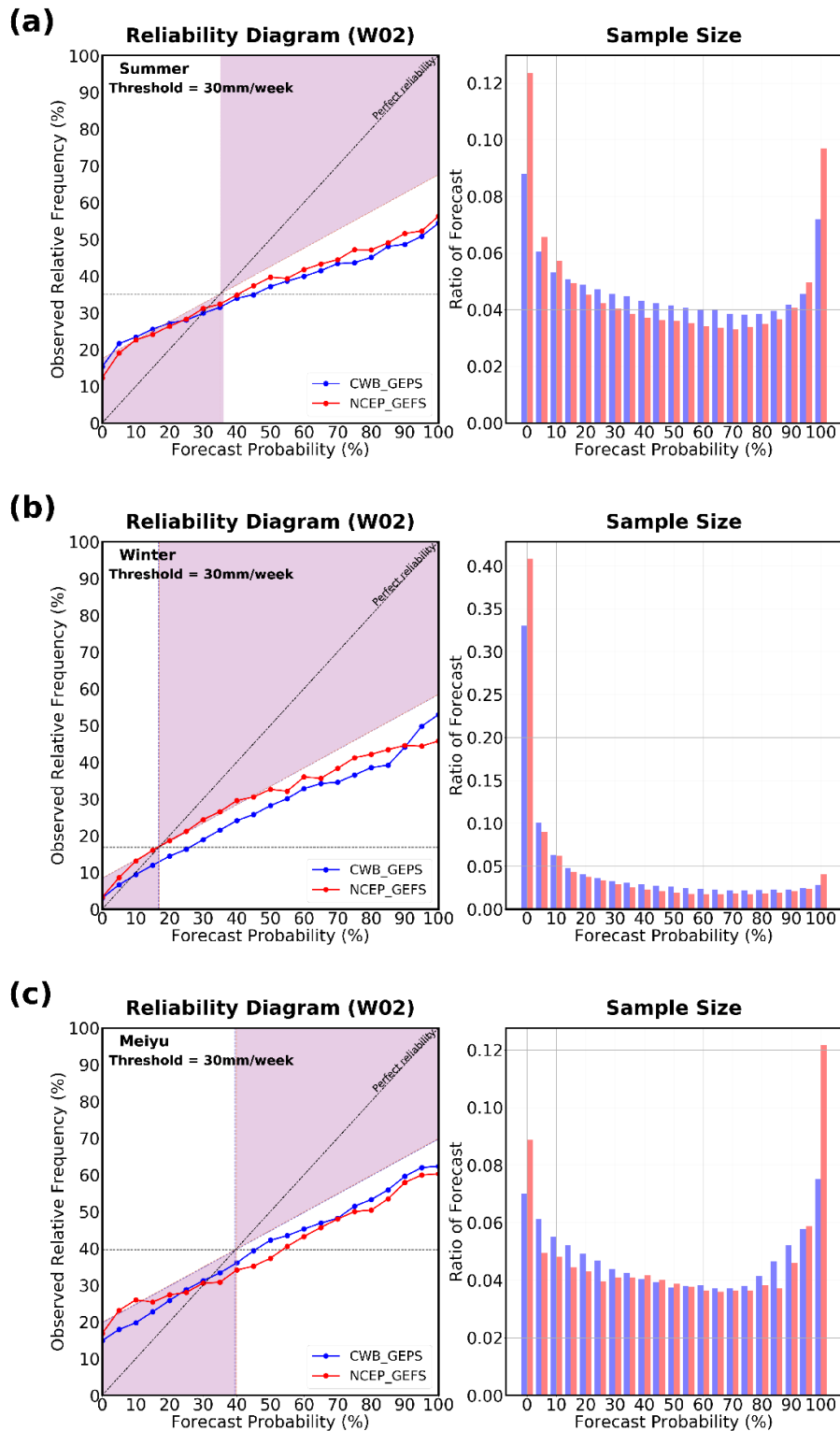


圖 5、第二週累積雨量達到降雨門檻 30 mm/week 的可信度圖，其中(a)為夏半年，(b)為冬半年，(c)為梅雨季。圖中水平黑色虛線代表機率預報相對於氣候不具有解析能力(no resolution)；黑色對角線虛線代表完美預報之信賴曲線分布。著色區域內表示相對於氣候預報具有誤差改善的預報能力(Brier Skill Score > 0)，著色區域外不具有預報技術(no skill)。

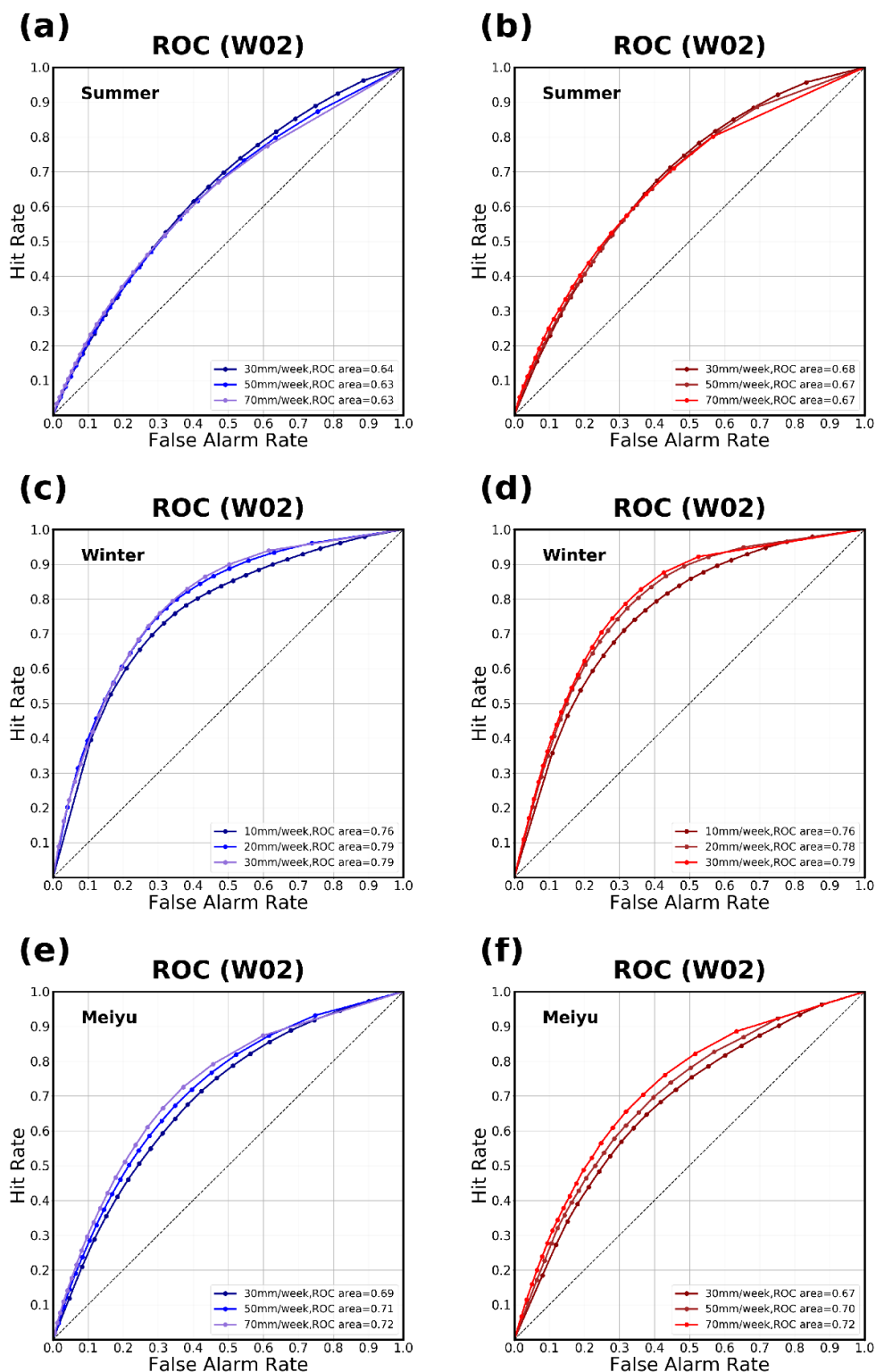


圖 6、第二週累積雨量在不同降雨門檻下的 ROC 曲線。圖(a)、(b)為夏半年，降雨門檻包含 30、50、70 mm/week；圖(c)、(d)為冬半年，降雨門檻包含 10、20、30 mm/week；圖(e)、(f)為梅雨季，降雨門檻包含 30、50、70 mm/week。圖中黑色對角虛線表示隨機預報。左欄(a)、(c)、(e)為 GEPS，右欄(b)、(d)、(f)為 GEFS。

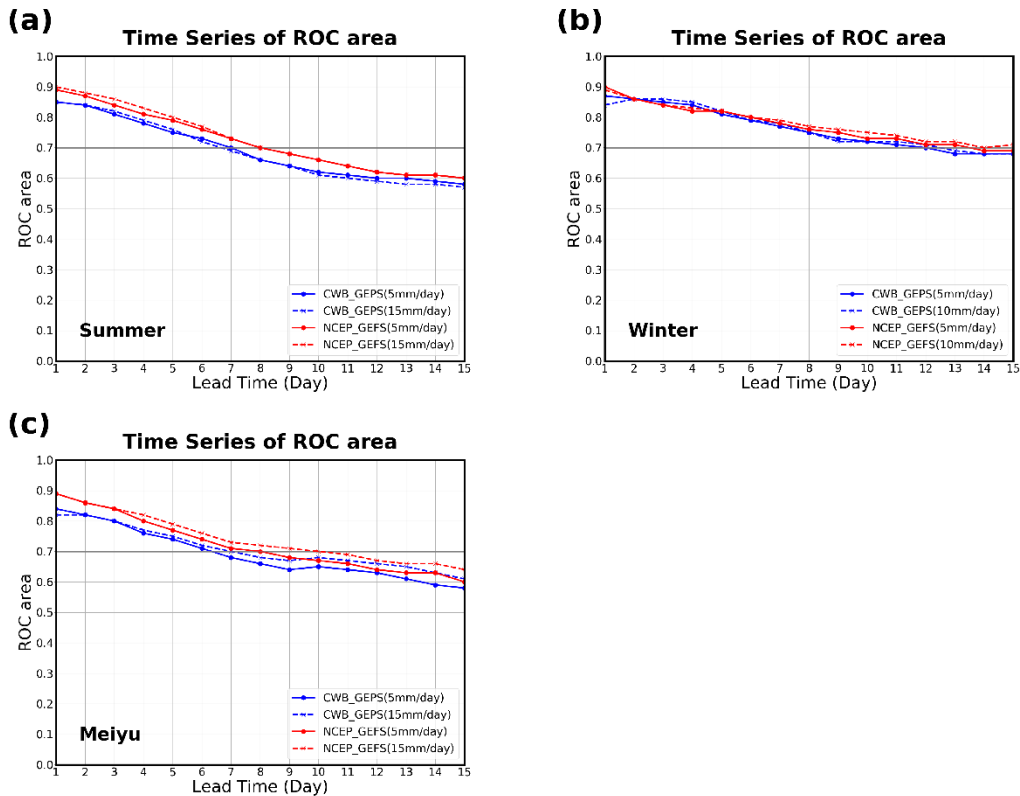


圖 7、預報時間第 1 天至第 15 天日累積雨量在約第 50、80 百分位降雨門檻下之 ROC 面積時間序列，其中(a)為夏半年，(b)為冬半年，(c)為梅雨季；紅色為 GEPS 的結果，藍色為 GEFS 的結果；實線為約第 50 百分位降雨門檻(夏半年、冬半年與梅雨季均為 5mm/day)，虛線為約第 80 百分位降雨門檻(夏半年與梅雨季為 15mm/day，冬半年為 10mm/day)。

(三)由系集預報導出之單一降雨預報校驗

(1) 性能圖

透過圖 8，我們進行不同模式與不同模式後處理方法(mean、PM 與 PR80)在不同季節以及不同降雨強度(i.e., 不同百分位之降雨門檻，是依據各季節之觀測樣本求得)的評比。整體而言：(1)雖 PR80 在夏季與梅雨季有明顯降雨過度預報的情形，但冬季對應相同降雨門檻時，PR80 普遍比 PM 有較高的 TS、

POD 與 SR，表示此過度預報主要反映在增加預報命中個數(h)，而不是增大錯誤預報個數(f)；(2) PM 在小雨與大雨的情況下均可使 BS 較接近 1(圖 8a-d)，表示 PM 可以修正降雨頻率偏差；(3)針對降雨強度較大的事件，GEFS 比 GEPS 有較好的預報表現，即 TS、POD 和 SR 較高，但在梅雨季(圖 8e, f)有嚴重的過度預報。整體評估建議梅雨季和夏半年採用 PM，而冬半年採用 PR80 作為預報指引。

(2) 箱形圖(Boxplot)

為了評估模式對強度較大降雨事件的預報能力，我們選用各季節第 80 百分位之週累積觀測降雨量作為降雨門檻，針對不同季節不同模式後處理方法的預報校驗分數 Tukey boxplot 進行分析。結果顯示：(1)在模式後處理方法方面，雖然 FAR 差異不大，但 PR80 有相對較高的 POD 與 BS，表示 PR80 藉由過度預報使得預報命中個數(h)提高，但不會過度增加錯誤預報(f)發生的情形，因此有較佳的預報表現。(2)在夏半年(圖 9a)GEPS 與 GEFS 的校驗分數 IQR 較其他兩季節大，表示夏半年兩模式各報間的預報表現相對不穩定。(3)冬半年(圖 9b)顯示 GEPS 的 POD 較 GEFS 高，且 FAR 亦較 GEFS 小，表示 GEPS 在冬半年有較良好的預報表現。(4)在模式方面，在兩模式 FAR 差異不大的情況下，梅雨季(圖 9c)以 GEFS 有較高的 BS 與 POD 校驗得分，表示 GEFS 透過顯著的過度預報提升校驗得分。整體而言，針對強度大的降雨門檻，模式校驗分數容易因過度預報而提高，故 PR80 有較佳的預報表現。在夏半年與梅雨季主要以 GEFS 有較佳的預報結果，在冬半年則以 GEPS 表現稍好。

(3) 相對均方根誤差(Relative Root-Mean-Square Error, RRMSE)之箱型圖

為了避免隨降雨強度增加而成長之降雨誤差，採用更公平客觀之 RRMSE(計算方式參見附錄)，評析不同後處理方法預報各降兩個案之誤差表現(圖 10)：(1)在季節方面，預報誤差在冬半年最小(圖 10b)，主要因為冬半年天氣系統可預報度較高，對於降雨空間分布、強度之掌握度相對於其他季節較佳；(2)在後處理方法方面，以 mean 預報誤差最小，源自於系集平均過程中會濾掉非線性系統性誤差。另外，圖 10c 顯示梅雨季 GEPS 的預報誤差比 GEFS 小，與前述 GEFS 降雨過度預報較嚴重有關。

(4) 空間相關係數之箱型圖與時序圖(time series)

0823 熱帶低壓影響期間(圖 11a 底圖粉紅色區域)預報與觀測有較高的相關性，且預報降雨空間分布大致上與觀測相似(圖 12a)，表示模式對天氣系統具有一定的掌握能力；山竹颱風靠近臺灣地區期間(圖 11a 底圖黃色區域)空間相關係數較低，原因為颱風路徑預報偏差所導致(圖 12b)。(2)冬半年降雨可預報度較夏半年高，故預報與觀

測相關性相對容易維持在相關係數 0.4 以上(圖 11b 底圖粉紅色區域), 例如圖 12c; 但四月份受到可預報度較低之中尺度天氣系統影響, 導致預報與觀測相關性降低(圖 11b 底圖黃色區域)。(3)梅雨季期間容易受到北方冷高壓與南方暖溼氣團交界位置與勢力發展所影響, 因此模式在梅雨季的可預報度比較低; 雖校驗期間有兩個降雨事件其相關係數大於 0.4(圖 11c、圖 12d), 但整體而言預報與觀測相關性普遍較低。綜合以上敘述, 臺灣附近區域預報與觀測較高之空間相關性容易出現在顯著降雨事件發生期間, 然而關鍵仍為數值模式對天氣系統強度與位置有一定的掌握能力。

(5) 時間相關係數空間分布 (Temporal Correlation Map)

GEPS 在臺灣鄰近降雨趨勢(時間相關性)之空間分布分析顯示: (1) 在夏半年(圖 13a)普遍東亞地區相關程度不高, 表示預報無法良好掌握降雨趨勢。(2)在冬半年(圖 13b)

部分區域之相關程度明顯比夏半年高, 且高相關性主要落在臺灣南側海域, 表示此區域的降雨趨勢表現相對穩定, 故時間相關係數較高; 在臺灣北部海域可能受到模式對鋒面位置或降雨強度之變異度較大, 使得預報趨勢掌握度較差, 故時間相關性不高。(3)梅雨季統計樣本僅兩個月間的降雨事件, 再加上天氣系統變異度較大, 可能導致時間相關係數高低分布較零散(圖 13c), 且通過顯著性檢定之格點較少, 未來應集整更多樣本提升校驗結果之統計代表性。

在 RRMSE 部分, 夏半年在臺灣東側太平洋區域有誤差表現較大之區域, 此與常發生在太平洋之天氣系統相關, 例如颱風。另外, 在中國大陸之預報誤差較大, 可能原因包含模式對陸地降雨特性描述能力與觀測資料的準確性等, 本研究在此不深入討論。綜合相關係數與 RRMSE 分析, 皆指出系集模式針對第二週累積降雨存在相當大的預報不確定性。

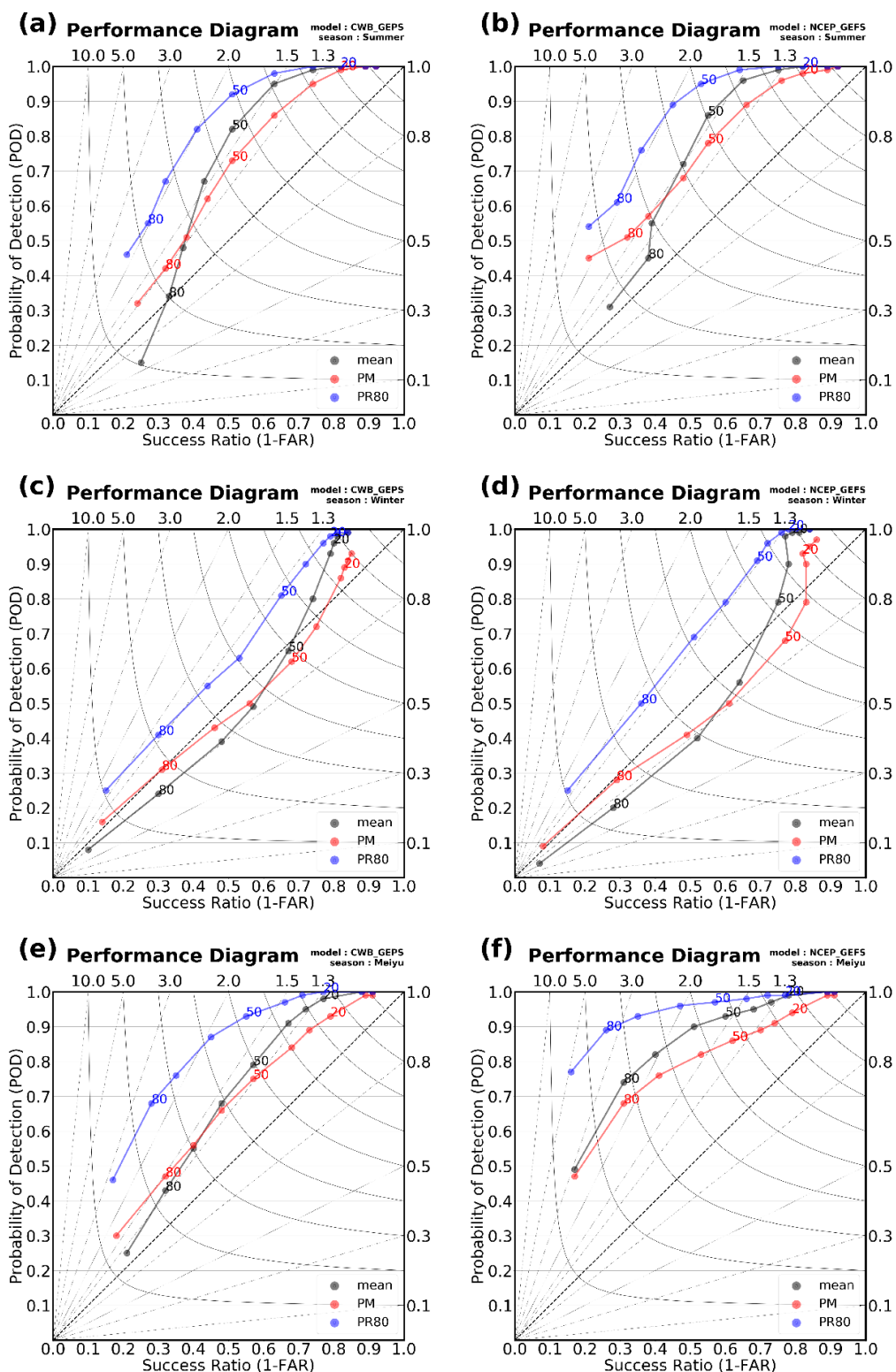


圖 8、各季節三種模式後處理方法導出第二週累積雨量預報在十個不同百分位降雨門檻(亦即不同降雨強度)下之性能圖。圖(a)、(b)為夏半年，圖(c)、(d)為冬半年，圖(e)、(f)為梅雨季。圖中黑色對角虛線表示隨機預報。左欄為 GEPS，右欄為 GEFS。不同數字表示不同百分位降雨門檻。模式後處理方法包含 mean、PM、PR80 三種，顏色分別以黑色、紅色與藍色表示。

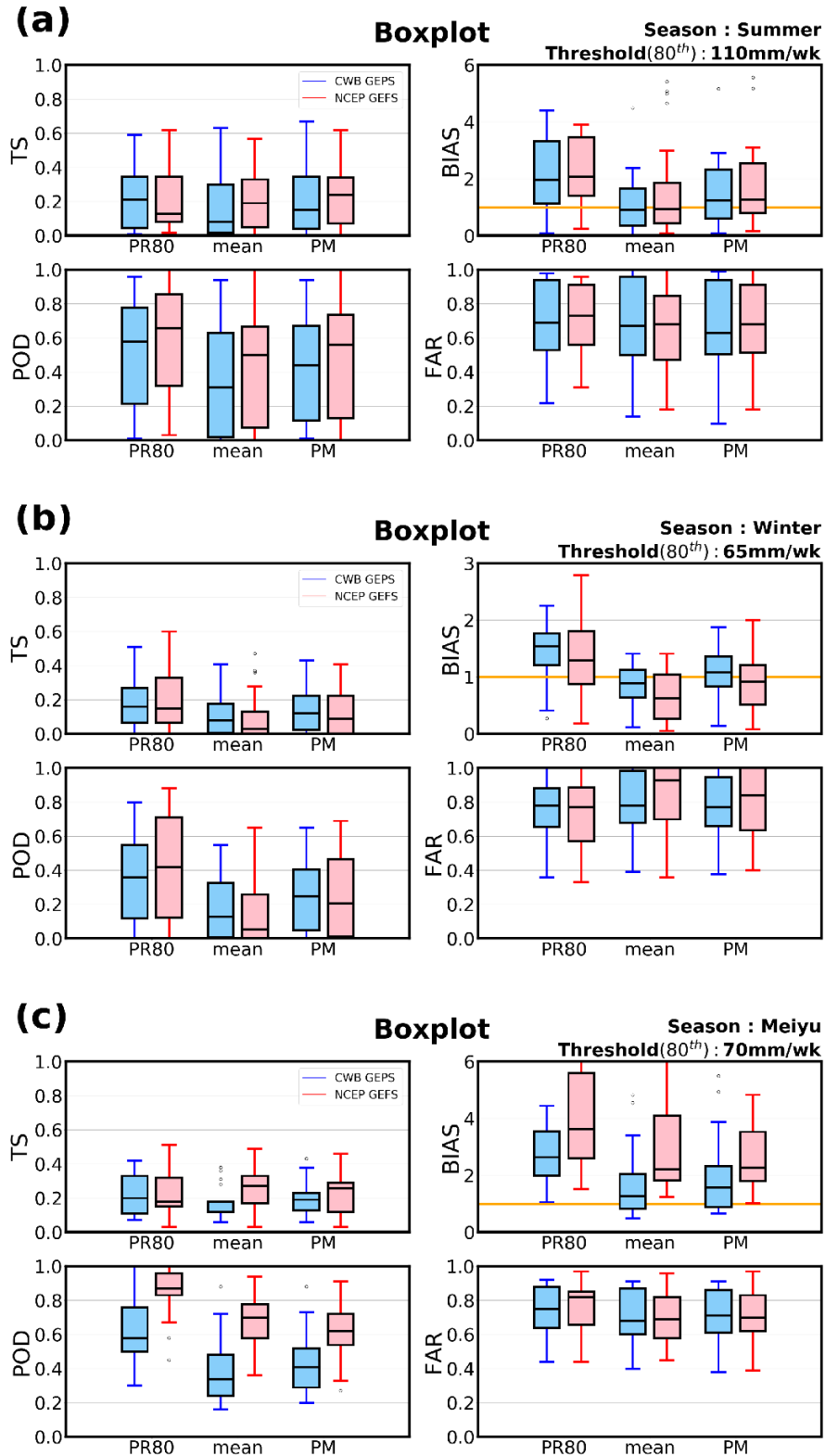


圖 9、在臺灣鄰近範圍 GEPS 與 GEFS 針對三種模式後處理方法預報第二週累積雨量之四種校驗分數箱型圖，其中(a)為夏半年，(b)為冬半年，(c)為梅雨季，第 80 百分位降雨門檻分別為 110、65、70 mm/week；藍色為 GEPS 的結果，紅色為 GEFS 的結果。

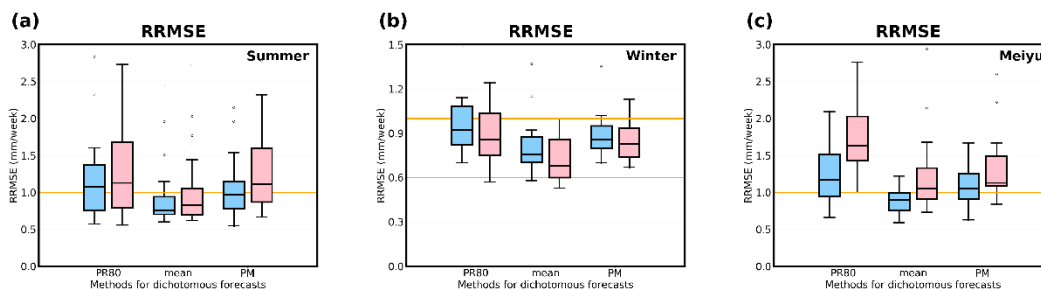


圖 10、GEPS 與 GEFS 三種模式後處理方法針對臺灣鄰近第二週累積雨量之正規化均方根誤差(RRMSE)，其中(a)為夏半年，(b)為冬半年，(c)為梅雨季；藍色為 GEPS 的結果，紅色為 GEFS 的結果。

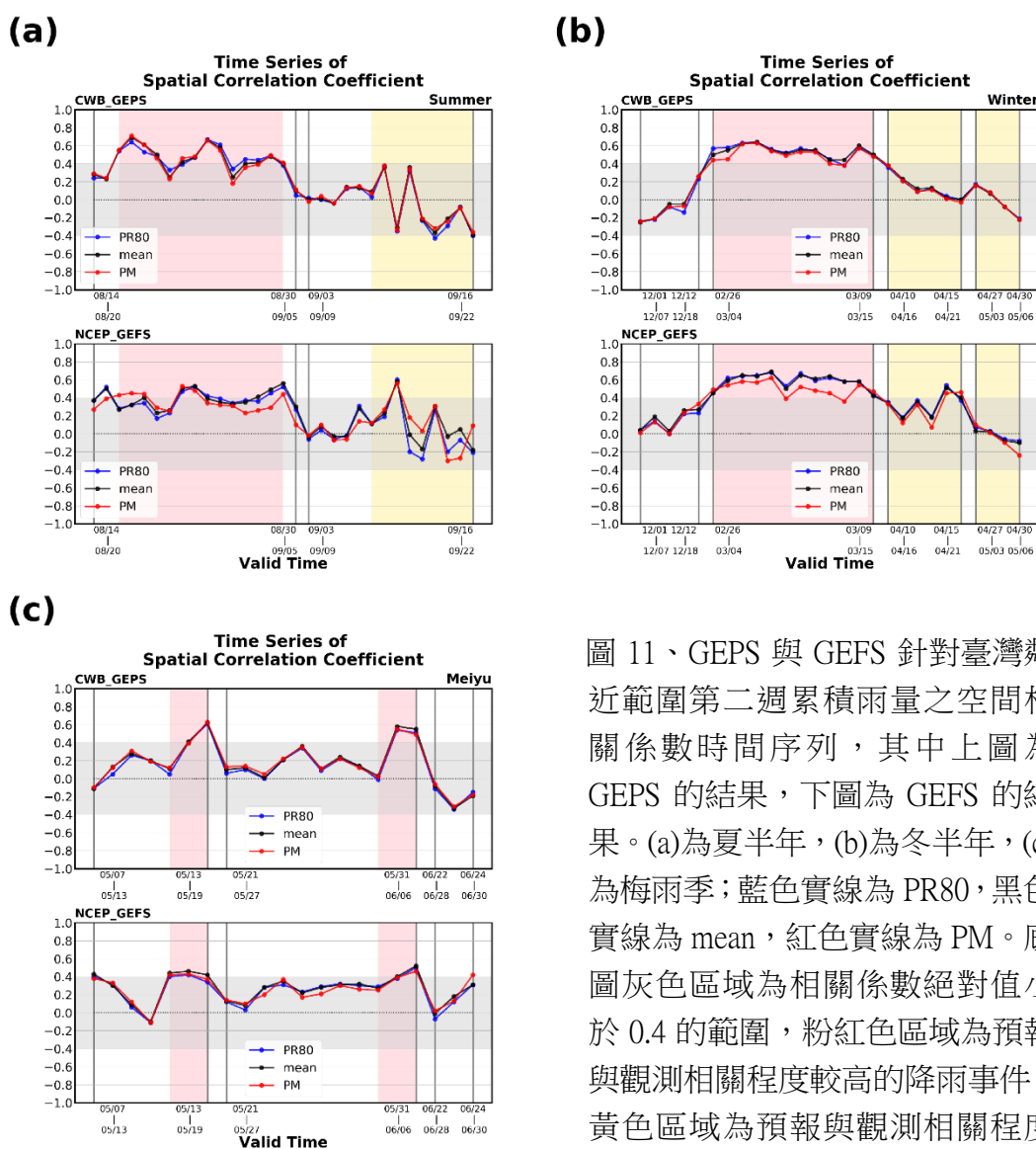


圖 11、GEPS 與 GEFS 針對臺灣鄰近範圍第二週累積雨量之空間相關係數時間序列，其中上圖為 GEPS 的結果，下圖為 GEFS 的結果。(a)為夏半年，(b)為冬半年，(c)為梅雨季；藍色實線為 PR80，黑色實線為 mean，紅色實線為 PM。底圖灰色區域為相關係數絕對值小於 0.4 的範圍，粉紅色區域為預報與觀測相關程度較高的降雨事件，黃色區域為預報與觀測相關程度較低的降雨事件。

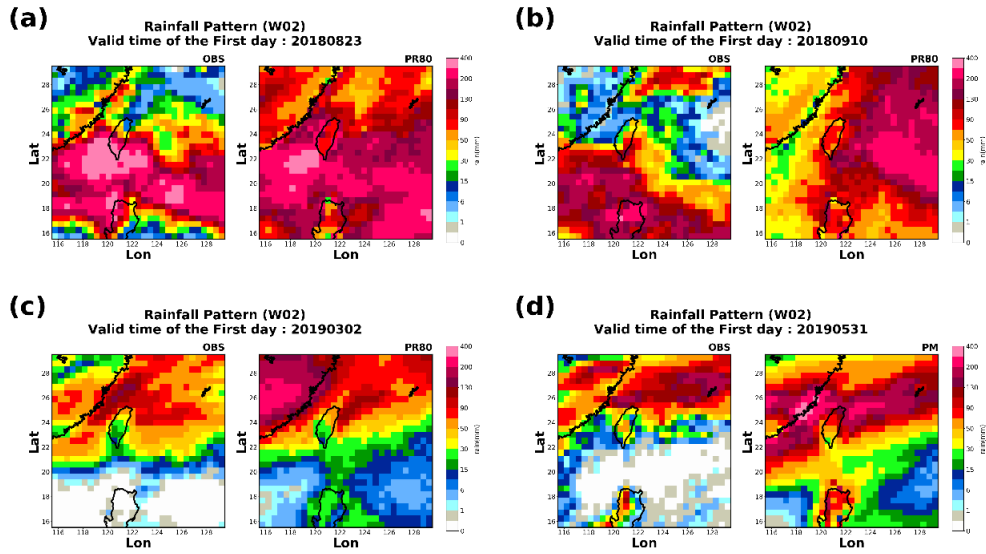


圖 12、CMORPH 觀測與 GEPS 預報於共同格點之第二週累積雨量空間分布。四個個案有效時間分別為(a)2018/08/23 至 2018/08/29、(b)2018/09/10 至 2018/09/16、(c)2019/03/02 至 2019/03/08、(d)2019/05/31 至 2019/06/06。顏色對照長條圖為累積降雨量值色階，單位為 mm/week。

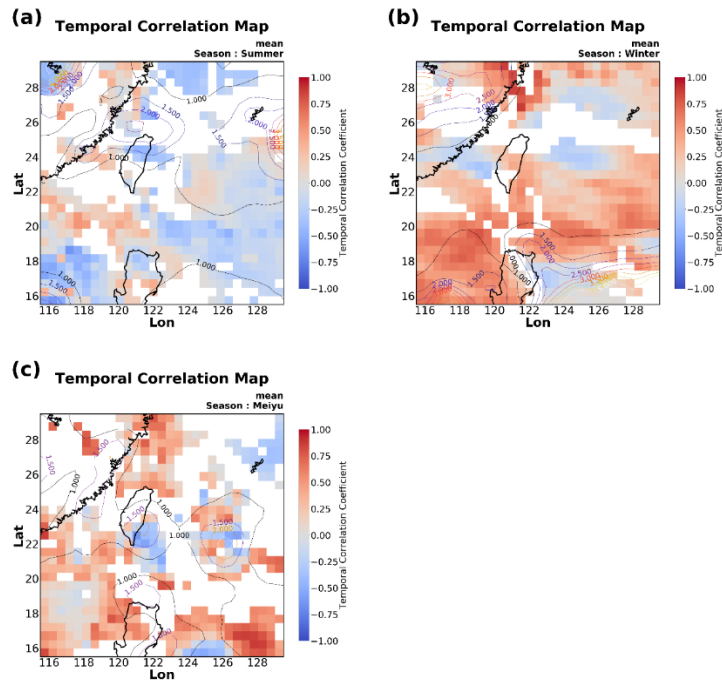


圖 13、針對臺灣鄰近範圍 GEPS 系集平均與 CMORPH 觀測之時間相關係數與 RRMSE 空間分布圖，X 軸為經度，Y 軸為緯度，其中(a)為夏半年，(b)為冬半年，(c)為梅雨季。顏色對照長條圖為時間相關係數色階，暖色系表示正相關，冷色系表示負相關，顏色越深表示觀測與預報相關程度越高，反之顏色越淺相關程度越小。等值線為各格點之 RRMSE，單位為 mm/week。空白區域表示未通過統計顯著性檢定(90%信心水準)。

五、結論與未來展望

隨著數值天氣預報的發展，近年來國際上致力於發展無縫隙預報，欲提供短期天氣至長期氣候完整的預報產品，然而展期預報因各種不確定因素的長時間誤差累積，導致預報不確定性相當大。為了涵蓋預報過程中之不確定性，各國積極發展系集預報以定量描述預報的不確定性。系集預報可透過給定某一降雨門檻計算得到相應的機率型產品，但民眾與政府防災單位普遍有不知如何使用機率預報的困難，因此氣象局會透過各種統計後處理方法由系集預報導出單一預報產品，此產品不同於單一模式的決定性預報，可視為包含不確定性訊息的單一預報。本研究目的在於評估 CWB GEPS 與 NCEP GEFS 在東亞地區的二週降雨預報表現，採用完整一年的資料以各種定量、定性校驗方式進行三大部分評估：

(一) 系集預報品質評估

不論是第二週累積雨量或日雨量預報，Rank Histogram 與 Spread-Skill Relationship 分析均顯示兩系集預報系統在東亞地區的離散度不足：(1)在季節方面，夏半年最明顯，冬半年最輕微。(2)在模式方面，以 GEPS 較 GEFS 嚴重。

(二) 系集機率預報校驗

Reliability Diagram 的分析顯示 GEPS 與 GEFS 在東亞地區第二週累積雨量機率預報的可信度不大，而 ROC 分析顯示兩模式在冬半年的潛在可用性明顯高於其他季節，有預報能力的天數均可長達 10 天以上，且 GEFS 有預報能力的天數較 GEPS 長。

(三) 由系集預報導出之單一預報校驗

三種不同模式後處理方法在臺灣附近的預報表現顯示：PR80 在冬半年可透過輕微的過度預報而提高 TS 和 POD 校驗分數；PM 具有修正降雨頻率偏差的特性；mean 因過程中濾除非線性誤差使其具有最小的預報誤差。整體評估建議兩模式在夏半年和梅雨季採用 PM，而冬半年採用 PR80 作為預報指引。針對降雨強度較大的事件，在夏半年與梅雨季以 GEFS 的預報表現較佳，冬半年則以 GEPS 稍好。校驗結果同時顯示：在顯著降雨事件期間，預報與觀測之空間相關性相對較高。

雖就機率型或單一預報校驗方法而言，樣本是足夠的，但因為本研究僅採用一年的資料，校驗結果無法反應模式預報能力的年際變化，例如梅雨季降雨預報能力之年際變化，故不完全符合系集預報藉由大樣本推估預報產品的預報水準與可信程度的基本精神。後續累積兩年資料後，我們將重新評估預報表現，並透過統計推論得到

不同信心水準下的信賴區間，以使預報評估更具統計代表性。

綜合以上結論，系集預報目的在於攫取展期降雨預報過程中的不確定性訊息，進一步得到機率型或透過後處理方法導出單一預報產品，不過一般民眾與下游單位不易應用預報資料

採取各式行動。故未來將進行系集預報產品之經濟價值(Economic Value, EV)分析，評估使用者參考系集預報能夠獲得多少效益，並在作業上提供各使用者採取防災行動的最佳機率門檻，以利後續決策判斷與防災預警。

致謝

感謝中央氣象局提供 CWB GEPS 降雨資料與電腦計算資源，以及審查委員與中央氣象局科技中心羅存文、陳昫靖等多位同仁寶貴的建議，使得

本論文得以更加嚴謹與完整。同時感謝科技部自然司之研究經費補助，本研究是在專題研究計畫(MOST 109-2111-M-052-001)補助下完成。

附錄

(一) 主要計算公式：

(1) 系集平均的方均根誤差(E)：

$$E(\lambda, \phi) = \left[\left(\psi_a(\lambda, \phi) - \bar{\psi}(\lambda, \phi) \right)^2 \right]^{\frac{1}{2}} = |\psi_a(\lambda, \phi) - \bar{\psi}(\lambda, \phi)| \quad (\text{式 1})$$

其中 λ, ϕ 分別為 x 格點與 y 格點； $\psi_a(\lambda, \phi)$ 為格點 (λ, ϕ) 之變量， $\bar{\psi}(\lambda, \phi)$ 為格點 (λ, ϕ) 之系集平均

(2) 系集離散度(S)：

$$S(\lambda, \phi) = \left\{ \frac{1}{N} \sum_{j=1}^N [\psi_j(\lambda, \phi) - \bar{\psi}(\lambda, \phi)]^2 \right\}^{\frac{1}{2}} \quad (\text{式 2})$$

$$\bar{\psi}(\lambda, \phi) = \frac{1}{N} \sum_{j=1}^N \psi_j$$

其中 $\psi_j(\lambda, \phi)$ 為該格點 (λ, ϕ) 某一系集成員 j 之變量， $\bar{\psi}(\lambda, \phi)$ 為該格點 (λ, ϕ) 之系集平均； N 為系集成員個數

(3) Brier skill score(*BrSS*)

BrSS 可評估系集機率預報相對於參考預報(一般是採用樣本氣候預報)的誤差改善程度，即相對於使用氣候資訊進行預報時，系集機率預報是否具有預報技術。計算公式如下：

$$BrSS = \frac{BrS - BrS_{ref}}{BrS_{perf} - BrS_{ref}} = 1 - \frac{BrS}{BrS_{ref}} \quad (式 3)$$

其中 *Brier score*(*BrS*)是預報系統的機率預報誤差大小，下標 *ref*與 *perf*分別代表參考預報與完美預報，*BrS* 定義如下：

$$BrS = \frac{1}{N} \sum_{j=1}^N (P_j - O_j)^2 \quad (式 4)$$

其中 *N*是所有校驗樣本數，*P_j*是預報機率，*O_j*是觀測頻率

BrSS 的值域從-∞到 1，BrSS=0 表示機率預報相對於氣候預報沒有預報能力，BrSS=1 表示完美預報，而 BrSS>0 代表機率預報相對於氣候預報是有預報能力的。

(4)相對均方根誤差(*RRMSE*)：

$$RRMSE = \frac{\sqrt{\frac{1}{N} \sum_{k=1}^N (R_f(k) - R_o(k))^2}}{\sqrt{\frac{1}{N} \sum_{k=1}^N (R_o(k))^2}} \quad (式 5)$$

其中 $R_f(k)$ 為該格點雨量預報值， $R_o(k)$ 為該格點雨量觀測值；*N*為總網格點數

(5)相關係數：

$$r = \frac{\sum_{k=1}^N (R_f(k) - \bar{R}_f)(R_o(k) - \bar{R}_o)}{\sqrt{\sum_{k=1}^N (R_f(k) - \bar{R}_f)^2 \sum_{k=1}^N (R_o(k) - \bar{R}_o)^2}} \quad (式 6)$$

其中 \bar{R}_f 為雨量預報平均值， \bar{R}_o 為雨量觀測平均值

(二) 二維列聯表：

	Observations		Forecasts	
		Yes	No	
Yes		<i>h</i> (hits)	<i>m</i> (misses)	

No	f (false alarms)	c (correct negatives)
----	-----------------------	----------------------------

表 1：二維列聯表。 h 為命中格點數，表示觀測與模式都達到降雨 \geq 門檻值的樣本數； f 為誤報格點數，表示觀測未達到，但模式有達到降雨 \geq 門檻值的樣本數； m 為漏報格點數，表示觀測有達到，但模式沒有達到降水 \geq 門檻樣本數； c 為正確識別樣本數，表示觀測與模式皆未達到降雨 \geq 門檻值的樣本數。

命中率與錯誤預警率的計算方式分別為：

$$\text{Hit Rate} = \frac{h}{h + m}, \quad (\text{式 7})$$

$$\text{False Alarm Rate} = \frac{f}{f + c} \quad (\text{式 8})$$

五種校驗分數計算方式分別為：

$$\text{POD} = \frac{h}{h + m}, \quad (\text{式 9})$$

$$\text{FAR} = \frac{f}{h + f}, \quad (\text{式 10})$$

$$\text{SR} = 1 - \text{FAR} = \frac{h}{h + f}, \quad (\text{式 11})$$

$$\text{TS} = \frac{h}{h + m + f}, \quad (\text{式 12})$$

$$\text{BS} = \frac{h + f}{h + m} \quad (\text{式 13})$$

(三) 箱形圖

箱型圖是一種描述性統計的圖示，主要利用最小值、第一四分位數或下四分位數(Q1，即由小至大第 25 百分位的數據)、中位數、第三四分位或上四分位數(Q3，即由小至大第 75 百分位的數據)與最大值五種統計量，透過視覺化方式描述樣本的分布特性。常見的分析方式包含四分位距(Interquartile

Range, IQR)，藉由上、下四分位數之差值大小分析數據的離散分布，若 IQR 越小(大)說明樣本越集中(離散)；觀察中位數偏向上四分位數或下四分位數可分析統計樣本之偏向等。另外，箱型圖有多種變形，Tukey 箱型圖是其中一種，可直觀地檢視資料之離群值(Outlier)，這些少數的離群值可能導致樣本平均不具代表性，或是使用者對於樣本分

布的誤判。最簡單的離群值(Outlier)定義為超出下限 $Q1-1.5\times IQR$ 與上限 $Q3-$

$1.5\times IQR$ 範圍的樣本。

參考文獻

李志昕、洪景山，2011：區域系集預報系統研究：物理參數化擾動。《大氣科學》，39，95-116。

___、洪景山，2014：區域系集預報系統研究：系集成員產生方式之評估。《大氣科學》，42，153-179。

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。《105 年天氣分析與預報研討會》。中央氣象局，臺北，臺灣。

周柿均、張惠玲、陳昀靖、羅存文、陳韋廷，2019：東亞地區二週系集降雨預報之校驗分析。《108 年天氣分析與預報研討會》，中央氣象局，臺北，臺灣。

卓盈旻、林昀靜、胡志文，2018：運用氣象局短期氣候預報系統發展臺灣梅雨季季內預報。《107 年天氣分析與預報研討會》，中央氣象局，臺北，臺灣。

陳孟詩，2010：中央氣象局月季長期天氣展望之預報校驗。《99 年天氣分析與預報研討會》，中央氣象局，臺北，臺灣。

陳思穎、黃婉如，2017：TRMM、CMORPH、PERSIANN 三組衛星資料對臺灣降雨日變化特色的表現能力評估。《大氣科學》，45，167-191。

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。

___、陳昀靖、周柿均、羅存文、馮智勇，2019：中央氣象局二週溫度與降雨預報產品之介紹。《108 年天氣分析與預報研討會》，中央氣象局，臺北，臺灣。

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。

葉世瑄、林沛練、洪景山、黃椿喜，2016：機率擬合之系集定量降水預報後處理方法。《大氣科學》，44，83-112。

鄭明典，2001：數值天氣預報近期的發展。《物理雙月刊》，廿三卷三期。

蘇奕叡、洪景山、李志昕，2016：系集機率擬合平均定量降水預報產品

- 之分析: 以 2014 年梅雨季為例。
大氣科學, **44**, 113-133。
- Brunet, G., and Coauthors, 2010: Collaboration of the weather and climate communities to advance subseasonal-to-seasonal prediction. *Bull. Amer. Meteor. Soc.*, **91**, 1397 – 1406.
- Buizza, R., 1997: Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system. *Mon. Wea. Rev.*, **125**, 99 – 119.
- ___, M. Miller, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Q. J. R. Meteorol. Soc.*, **125**, 2887 – 2908.
- ___, M. Miller, P. L. Houtekamer, Z. Toth, G. Pellerin, M. Wei, and Y. Zhu, 2005: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon. Wea. Rev.*, **133**, 1076 – 1097.
- ___, and T. N. Palmer, 1995: The singular vector structure of the atmospheric general circulation. *J. Atmos. Sci.*, **52**, 1434-1456.
- ___, and T. N. Palmer, 1998: Impact of ensemble size on ensemble prediction. *Mon. Wea. Rev.*, **126**, 2503-2518.
- Chang, H. L., H. Yuan, P. L. Lin, 2012: Short-Range (0-12h) PQPFs from Time-Lagged Multimodel Ensembles Using LAPS. *Mon. Wea. Rev.*, **140**, 1496 – 1516.
- ___, B.G. Brown, P. S. Chu, Y. C. Liou, and W. H. Wang, 2017: Nowcast Guidance of Afternoon Convection Initiation for Taiwan. *Weather and forecasting*, **32**, 1801-1817.
- Du, J., S. L. Mullen, and F. Sanders, 1997: Short-range ensemble forecasting of quantitative precipitation. *Mon. Wea. Rev.*, **125**, 2427-2459.
- Ebert, E. E., Ability of a poor man's ensemble to predict the probability and distribution of precipitation, *Mon. Wea. Rev.*, **129**, 2461 – 2480, 2001.
- Finley, J.P., 1884: Tornado predictions. *Amer. Meteor. J.*, **1**, 85-88.
- Gascón, E., Hewson, T., and Haiden, T., 2018: Improving predictions of precipitation type at the surface: Description and verification of two new products from the ECMWF ensemble. *Weather and Forecasting*, **33**(1), 89 – 108.
- Hamill, 2001: Interpretation of rank

- histograms for verifying ensemble forecasts. *Mon. Wea. Rev.*, **129**, 550 – 560.
- ___, and J. Juras, 2006: Measuring forecast skill: Is it real skill or is it the varying climatology? *Quart. J. Roy. Meteor. Soc.*, **132**, 2905 – 2923.
- Jolliffe, I. T., and D. B. Stephenson, 2011: *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. Wiley, 254 pp.
- Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, **5**, 487-503.
- Kalnay, E., and A. Dalcher, 1987: Forecasting forecast skill. *Mon. Wea. Rev.*, **115**, 349 – 356.
- Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**, 409-418.
- Lewis, W. R., W. J. Steenburgh, T. I. Alcott, and J. J. Rutz, 2017: GEFS precipitation forecasts and the implications of statistical downscaling over the western United States. *Wea. Forecasting*, **32**, 1007 – 1028.
- Mason, S. J., and N. E. Graham, 1999: Conditional probabilities, relative operating characteristics, and relative operating levels. *Wea. Forecasting*, **14**, 713 – 725.
- Robber, P.-J., 2009: Notes and Correspondence - Visualizing Multiple Measures of Forecast Quality. *Wea. Forecasting*, **24**, 601-608.
- Sharma, S., R. Siddique, N. Balderas, J.D. Fuentes, S. Reed, P. Ahnert, R. Shedd, B. Astifan, R. Cabrera, A. Laing, M. Klein, and A. Mejia, 2017: Eastern U.S. Verification of Ensemble Precipitation Forecasts. *Weather Forecast.* **32**, 117 – 139.
- Tukey, J., 1977: *Exploratory data analysis*. Reading, MA: Addison-Wesley Publishing Company.
- Vigaud, N., A. W. Robertson, and M. K. Tippett, 2017a: Multimodel ensembling of subseasonal precipitation forecasts over North America. *Mon. Wea. Rev.*, **145**, 3913 – 3928.
- ___, A. W. Robertson, M. K. Tippett, and N. Acharya, 2017b: Subseasonal

- predictability of boreal summer monsoon rainfall from ensemble forecasts. *Front. Environ. Sci.*, **5**, 2197 – 19.
- Wang, C. C., 2014: On the calculation and correction of equitable threat score for model quantitative precipitation forecasts for small verification areas: The example of Taiwan. *Wea. Forecasting*, **29**, 788 – 798.
- Whitaker, J. S., and A. F. Lough, 1998 : The relationship between ensemble spread and ensemble mean skill. *Mon. Wea. Rev.*, **126**, 3292-3302.
- White, C. J. et al., 2017: Potential applications of subseasonal-to-seasonal (S2S) predictions. *Meteor. Appl.*, **24**, 315-325.
- Wilks, D. S., 2000b: *Diagnostic verification of the Climate Prediction Center long-lead outlooks*, 1995-1998.
- ___, 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. Academic Press, 704 pp.
- Wobus, R. L, and E. Kalnay, 1995: Three years of operational prediction of forecast skill at NMC. *Mon. Wea. Rev.*, **123**, 2132-2148.
- Zhou, X., Y. Zhu, B. Fu, D. Hou, J. Peng, Y. Luo, and W. Li, 2018: The Development of Next NCEP Global Ensemble Forecast System. *43rd NOAA Annual Climate Diagnostics and Prediction Workshop*, 167-171.
- Zhu, Y., 2005: Ensemble forecast: A new approach to uncertainty and predictability. *Adv. Atmos. Sci.*, **22** (6), 781-788.
- ___, and Coauthors, 2018: Towards the improvement of sub-seasonal prediction in the NCEP Global Ensemble Forecast System (GEFS). *J. Geophys. Res. Atmos.*, **123**, 6732 – 6745.

Evaluation of Week-2 Ensemble Precipitation Forecasts over East Asia

Shih-Chun Chou¹ and Hui-Ling Chang²

1 Research and Development Center, Central Weather Bureau, Taipei, Taiwan

2 Meteorological Information Center, Central Weather Bureau, Taipei, Taiwan

(manuscript received 16 January 2020 ; in final form 27 April 2020)

Abstract

Focusing on week-2 precipitation forecasts of ensemble forecast systems of Central Weather Bureau (CWB) and National Centers for Environmental Prediction (NCEP), this study evaluates the spread-skill relationship, ensemble probabilistic forecasts, and dichotomous forecasts derived from ensemble forecasts. The aim is to compare the performance of both forecast systems and provides more suitable forecast guidance for users.

Results indicate that 1) both ensemble prediction systems are under-dispersive especially in summer, and the CWB forecast system appears to have more under-dispersion relative to the NCEP; 2) week-2 probabilistic precipitation forecasts do not have high reliability, but have good potential usefulness in winter; 3) the better post-processing method for precipitation forecast guidance is the PM (Probability-Matched Mean) during summer and Mei-Yu season, and the 80th percentile of ensemble forecasts in winter.

Key Words: ensemble forecast system, ensemble probabilistic forecast, reliability, potential usefulness, ensemble spread,

doi: 10.3966/025400022019064702003