

臺灣地區系集預報系統導出決定性預報第二週累積降雨校驗

周柿均 張惠玲 陳昀靖 羅存文 朱心宇

中央氣象局

摘 要

現今各國極力發展無縫隙預報，為涵蓋預報過程中種種的預報不確定因素，系集預報技術日趨成熟，並可輕易產生大量的系集成員進而產製機率型的預報產品。然而，一般民眾不知道如何使用機率預報進行決策，且下游非氣象單位需要決定型的預報資訊進行後端應用，因此機率預報在推廣上相對不易。故本研究針對臺灣地區第二週累積雨量，進一步評估中央氣象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)透過統計後處理方法導出決定性預報的預報表現，採用的後處理方法包括系集平均(ensemble mean)、機率擬合平均(Probability-Matched Mean, PM)與系集第 80 百分位數(Quantitative Precipitation Forecast Percentile 80th, QFPF80)。研究目的為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。

結果顯示：GEPS 與 GEFS 導出決定性預報的預報表現相近，其中以 GEFS 比 GEPS 具有較嚴重的過度預報，季節方面則以冬半年的預報表現穩定性最佳。除與各系集模式的降雨預報能力相關外，預報表現同時反映各統計後處理方法自身的特性：QFPF80 普遍具有最高的 Bias 與 POD，PM 的 Bias 較 ensemble mean 高，顯示透過重新分配降雨頻率以掌握降雨極值的特性等。整體而言，本研究建議 GEPS/GEFS 導出第二週累積雨量決定性預報的後處理方法：在夏半年採用 QFPF80/PM，冬半年使用 QFPF80/QFPF80，梅雨季參考 PM/ensemble mean。

關鍵字：系集預測系統、統計後處理

一、前言

近年來各國致力於發展無縫隙預報，期望能夠提供短期天氣至氣候預測完整的預報產品。然而，數值預報系統中存在著非線性過程，許多不確定性皆可能造成模式預報結果相當大的差異。為涵蓋預報過程中種種不確定因素，因此發展出系集預報：透過不同擾動方式得到一組能夠合理描述預報不確定性的系集成員，並依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。隨著系集預報技術日趨成熟，已可輕易產生大量的系集成員進而產製機率型的預報產品，然而，

一般民眾不知道如何使用機率預報進行決策，且下游非氣象單位的使用者常需使用決定型的明確資訊進行後端應用，導致機率預報在推廣上相對不易，尤其是降雨這類變數在時間與空間上具有高度不連續特性，氣象單位依據各方需求，研發統計後處理的策略以應用系集預報系統，進而產製合適的定量降水預報產品。

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)在第二週的降雨預報表現，將各 20 組系集成員採用三種統計後處理方法，包括系集平均(ensemble mean)、機率擬合平均(PM)與系集第 80 百分位數(QFPF80)，導出決定性第二週累積降雨預

報，並以性能圖(performance diagram)、箱型圖(box plot)、正規化均方根誤差(relative root mean square error, RRMSE)與空間相關係數(spatial correlation coefficient)等各種校驗分數等進行評估，以助於使用者了解定量降水預報產品特性，同時回饋給模式研發者，作為模式改進之重要依據。

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。初始時間範圍從 2018 年 7 月至 2020 年 6 月，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；水平解析度分別為 $0.5^\circ \times 0.5^\circ$ 與 $1^\circ \times 1^\circ$ ，校驗範圍取東經 115 至 130 度、北緯 15 至 30 度。

2.2 觀測資料

雨量資料採用歐洲中期天氣預報中心(European Center for Medium-Range Weather Forecasts, ECMWF)所推出的 ERA5 再分析資料，以作為取代 ERA-interim 的再分析產品，時間解析度為 1 小時，水平解析度為 $0.25^\circ \times 0.25^\circ$ ，資料範圍涵蓋全球，資料時間從 1950 年至迄今。

三、研究方法

3.1 決定性預報校驗方法

3.1.1 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，可透過二維列聯表(Contingency Table; Finley 1884)計算得到：

$$POD = \frac{h}{h+m}, \quad SR = 1 - FAR = \frac{h}{h+f}$$

$$TS = \frac{h}{h+m+f}, \quad BS = \frac{h+f}{h+m}$$

(其中，h 為命中格點數，f 為誤報格點數，m 為漏報格點數，c 為正確識別樣本數)

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，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(下)方表示 BS 得分大(小)於 1，代表過度預報(低報)。

3.1.2 箱型圖(box plot)

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)，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。

3.1.3 正規化均方根誤差(RRMSE)

一般常使用均方根誤差(RMSE)作為預報誤差的一種度量方式。然而因預報系統的降雨預報誤差表現會受到降雨強度影響，為更公平地進行評估，將 RMSE 除以觀測標準差以正規化預報誤差，計算方式如下：

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

(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)

3.1.4 空間相關係數(spatial correlation coefficient)

逐格點計算觀測與預報之間的相關係數，計算得到空間相關係數，值域介於 -1 至 1 之間。若 r 等於 0，表示觀測與預報無相關程度，r 越接近 1(-1)，觀測與預報間的正(負)相關程度越高。計算方式如下：

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

(其中 R_f 為雨量預報平均值， R_o 為雨量觀測平均值)

3.2 統計後處理方法

3.2.1 系集平均(ensemble mean)

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，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)。

3.2.2 機率擬合平均(PM)

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。

3.2.3 系集降雨預報百分位數

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 QFPF80。

四、結果與討論

由於數值模式預報至二週以上的不確定性相對提高許多，因此本篇以第二週累積雨量的形式以涵蓋時間不確定性進行後續分析，而非日累積雨量。考量不同季節造成強降水的主要天氣型態不同，因此將校驗期分為三個季節：夏半年 7 月至 11 月，冬半年 12 月至 4 月，梅雨季則為 5 月與 6 月，並依季節挑選出具

有一定降雨強度的個案，樣本數分別為 68、62、59 報。

Performance diagram(圖 1)顯示，在統計後處理方法：(1)系集平均過程使得整體量值平滑，小雨高估大雨低估的現象會反映在 performance diagram 的 Bias：高(低)降雨閾值 Bias 較小(大)；PM 因重新分配降雨頻率進而改善系集平均小雨高估大雨低估的情形則反映在 Bias 的變化：PM 在高(低)降雨閾值其 Bias 相對 ensemble mean 較大(小)，且因夏半年與梅雨季容易發生局部的強降雨極值，PM 在夏半年與梅雨季調整降雨頻率與掌握極值的特性較冬半年顯著；為因應防災對強降雨之迫切需求，採用 QFPF80 進一步了解系集模式預報降雨極值的能力，故在三種後處理方法中，QFPF80 達到某一降雨閾值的預報格點數普遍最多，即反映在 performance diagram 的 Bias：具有最高的 Bias，表示相對於 ensemble mean 與 PM，QFPF80 過度預報最顯著。另觀察 POD 與 SR：在高降雨閾值下(例如降雨強度第 80, 90 百分位數閾值)，三種後處理方法的 SR 差異不大，其中以 QFPF80 具有最高的 POD，表示 QFPF80 的過度預報並非過度增加錯報格點(列聯表的 f)，且因命中格點數(列聯表的 h)增加，使得 QFPF80 普遍具有最高的 TS。

若比較 GEPS 與 GEFS 的預報表現，圖 1 顯示兩系集模式在各季節的預報表現相近，不過在夏半年與梅雨季高降雨閾值下，GEFS 比 GEPS 具有較高的 Bias，表示 GEFS 在夏半年與梅雨季預報強降雨量值的能力較 GEPS 高，在冬半年兩系集模式預報表現相當。整體而言，雖考量防災預警的需求，通常不會採用有低報現象的模式預報，然而太過嚴重的過度預報對於使用者而言無法獲得有效的訊息，綜合以上考量，良好的降雨預報其 Bias 應落於 1.0 至 1.5 間，故本研究建議系集導出第二週累積雨量決定性預報的後處理方法在夏半年 GEPS/GEFS 採用 QFPF80/PM，冬半年 GEPS/GEFS 採用 QFPF80/QFPF80，梅雨季 GEPS/GEFS 使用 PM/ensemble mean。

除透過 performance diagram 進行系集模式整體降雨預報表現外，我們進一步評比模式在各樣本間對於高降雨閾值的預報能力穩定性，選用各季節觀測第二週累積雨量第 80 百分位數作為降雨閾值，進行 TS、POD、FAR、Bias 等校驗分數的 Tukey 箱形圖分析。

圖 2 顯示：(1)在模式後處理方法方面， QPFP80 相對於 ensemble mean 與 PM 具有較高的 POD 與 BS，不過三種後處理方法的 FAR 差異不大，表示 QPFP80 雖過度預報最顯著，但不會過度增加錯誤預報個數(列聯表的 f)，同時因預報命中個數(列聯表的 h)提高，且 TS 校驗分數也比 ensemble mean 與 PM 稍高，與 performance diagram 三種後處理方法在第 80 百分位閾值的結果一致。然而觀察 QPFP80 的 Bias，除冬半年外，梅雨季兩模式 Bias 的 Q1 高於 1.5，顯示 QPFP80 在梅雨季的過度預報嚴重，對於後端應用參考性不高，類似情形同樣發生在夏半年 GEFS 導出的 QPFP80；(2)在季節方面，兩模式皆以冬(夏)半年具有最佳(差)的預報表現，包含 TS、POD 校驗分數高(低)、FAR 校驗分數低(高)、Bias 接近(不接近)1，且在冬(夏)半年 GEFS 與 GEFS 的校驗分數 IQR 最小(大)，表示兩模式在冬(夏)半年各報間的預報表現相對穩定(不穩定)。

為了解系集模式預報表現的空間分布，以時間相關係數評估模式預報降雨趨勢的能力，並以 RRMSE 作為預報誤差之度量。由於各季節透過不同統計後處理方法導出的預報表現差異不大，因此後續以系集平均作為代表進行分析。圖 3 結果顯示，以系集模式而言：兩系集模式在各季節的預報表現空間分布差異不大，其中臺灣本島區域在夏半年與梅雨季以 GEFS 具有較 GEPS 良好的預報趨勢。以季節而言：(1)系集模式在夏半年臺灣區域具有正的預報趨勢，其中以臺灣東側的預報誤差較臺灣西側小；在臺灣西南側至華南沿岸海域的預報趨勢為負值，且 RRMSE 非常大，說明系集模式在該範圍的降雨預報能力不佳；而太平洋海域受限於模式對颱風路徑、強度的預報能力，使得預報與觀測幾乎無相關程度。(2)冬半年在包含臺灣以南區域的預報趨勢不佳，原因為該區域在冬季為弱綜觀天氣型態，相對於東北鋒面自北方南下之強綜觀天氣型態，較無法準確掌握偶發性降雨，且本篇校驗期冬半年涵蓋 3 至 4 月春季天氣系統影響，此期間降雨型態包含由華南地區引進的強降雨系統，相對冬季東北鋒面降雨系統，其可預報度較低；在預報誤差方面，以中國北方陸域的預報誤差最大，且向東遞減，但具有正的預報趨勢，原因為兩模式容易在中國北方陸域過度預報降雨量值，同時與模式掌握鋒面東北-西南兩

帶位置的預報能力相關。(3)梅雨季期間以臺灣東南側海域的預報趨勢較臺灣北部區域差；除中國大陸外，在預報誤差空間分布呈現西南-東北走向，可能與季節性西南暖溼季風或華南雲雨帶向東移入臺灣地區相關。

五、結論與展望

為達到無縫隙預報的願景，各國無不致力於發展系集預報技術，系集預報除了能夠產生機率型預報以提供預報過程中不確定資訊外，透過合適的統計後處理方法亦可產製決定型的預報產品，以因應一般民眾與下游非氣象單位需求。延續先前系集機率預報校驗校驗結果(周, 2019)，本研究針對 CWB GEPS 與 NCEP GEFS 透過統計後處理方法導出決定性第二週累積降雨之預報表現進行評估。結果顯示：(1)GEPS 與 GEFS 導出決定性預報的預報表現相當，其中以 GEFS 相較於 GEPS 具有較嚴重的過度預報；(2)在高降雨閾值下，QPFP80 普遍具有最高的 Bias 與 POD，PM 的 Bias 較 ensemble mean 高，顯示透過重新分配降雨頻率以掌握降雨極值的特性；(3)GEPS 與 GEFS 在冬半年的預報表現穩定性最佳。整體而言，兩系集模式導出決定性預報的校驗結果，除了與系集模式的預報能力有關外，同時反映各統計後處理方法自身的特性。本研究建議 GEPS/GEFS 導出第二週累積雨量決定性預報的後處理方法：在夏半年採用 QPFP80/PM，冬半年使用 QPFP80/QPFP80，梅雨季參考 PM/ensemble mean。

系集預報的目的在於攫取預報過程中的不確定性訊息，一般而言，評估系集預報系統可從系集預報品質、系集機率預報校驗，以及由統計後處理方法導出決定型預報進行該模式預報能力全方位的評估。若進一步考量使用者依據系集預報判斷是否採取行動，能夠更深入了解各類型預報產品對於使用者提供多少應用價值，即系集預報的經濟價值分析(economic value analysis)，以利於後續防災預警的決策。

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。105 年天氣分析與預報研討會。中央氣象局。

周栴均、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。

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。

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。

Buizza, R., A. Hollingsworth, F. Lalurette, 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) PPDFs from time-lagged multimodel
ensembles using LAPS. Mon. Wea. Rev., 140,
1496–1516

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.

Leith, C. E., 1974: Theoretical skill of Monte Carlo
forecasts. Mon. Wea. Rev., 102, 409-418.

Zhu, Y., 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.

附圖

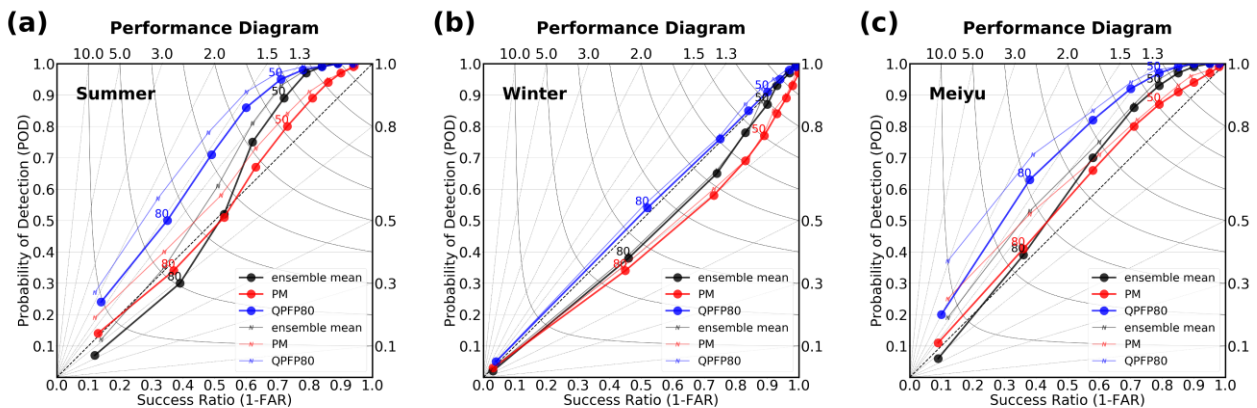


圖 1，GEPS 與 GEFS 由三種統計後處理方法預報第二週累積雨量在不同百分位降雨閾值下之性能圖。(a)為夏半年，(b)冬半年與(c)梅雨季。統計後處理方法包含 ensemble mean、PM、QFPF80，顏色分別以黑色、紅色與藍色表示；實心圓標記代表 GEPS 的結果，「N」標記代表 GEFS 的結果，同一條線上的標記由右上至左下表示由小至大共 9 個百分位數降雨閾值，並特別以數字標示 GEPS 第 50 與 80 百分位數降雨閾值位置。圖中黑色對角虛線表示隨機預報。

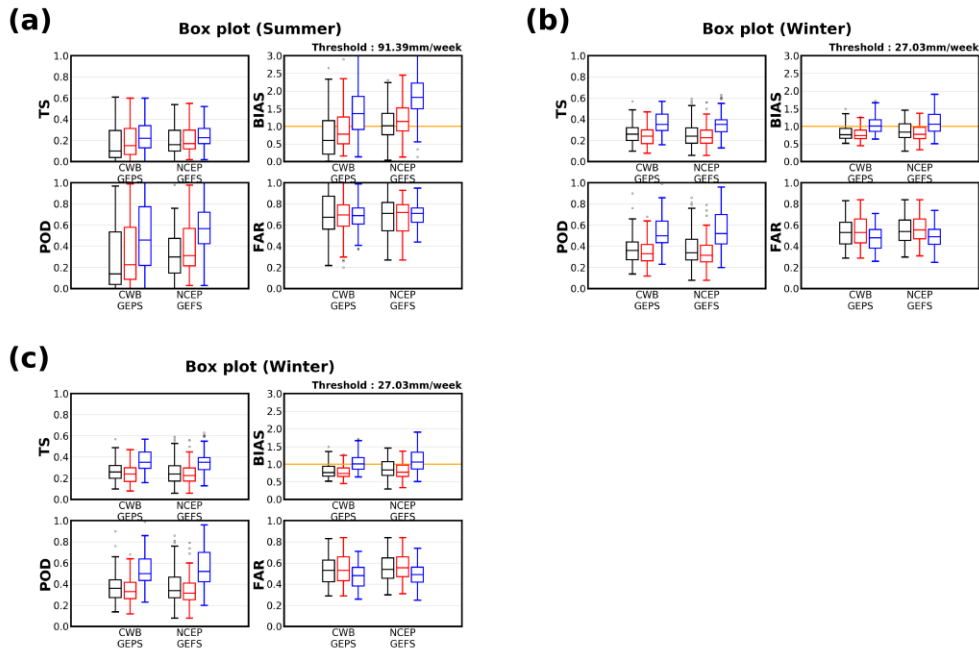


圖 2，GEPS 與 GEFS 透過三種模式後處理方法導出第二週累積雨量預報之四種校驗分數箱型圖，(a)為夏半年，(b)冬半年，(c)梅雨季。橫軸為系集模式，縱軸為校驗分數。統計後處理方法包含 ensemble mean、PM、QPFP80，顏色分別以黑色、紅色與藍色表示。圖右上角標示第 80 百分位降雨閾值。

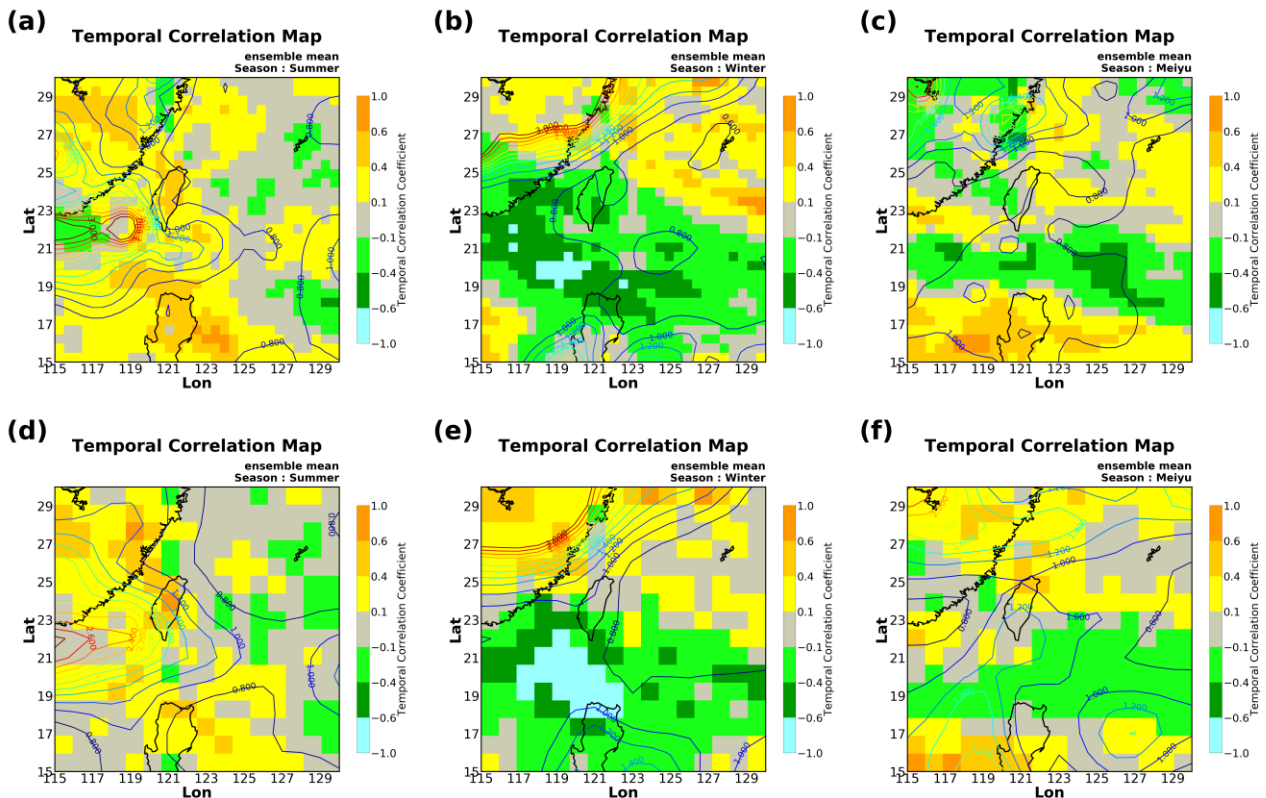


圖 3，系集平均與 ERA5 reanalysis 之時間相關係數與 RRMSE 空間分布圖，橫軸為經度，縱軸為緯度。上排為 GEPS 的結果，下排為 GEFS 的結果。(a)、(d)為夏半年，(b)、(e)冬半年，(c)、(f)梅雨季。顏色對照長條圖為時間相關係數色階。等值線為 RRMSE，單位為 mm/week。