

# 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—第二部分：校正

章鶴群<sup>1</sup> 陳冠儒<sup>1</sup> 劉豫臻<sup>1</sup> 張惠玲<sup>2</sup> 洪景山<sup>3</sup> 楊舒芝<sup>4</sup>

<sup>1</sup> 中央氣象局第四組

<sup>2</sup> 中央氣象局科技中心

<sup>3</sup> 中央氣象局資訊中心

<sup>4</sup> 國立中央大學大氣物理研究所

(中華民國 107 年 4 月 19 日收稿，中華民國 107 年 6 月 27 日定稿)

## 摘 要

中央氣象局 WRF 系集預報系統 (WRF Ensemble Prediction System, WEPS) 所提供的颱風降雨機率預報 (probabilistic quantitative precipitation forecasts, PQPFs)，具有明顯的系統性偏差。本研究嘗試對 WEPS 所做的降雨機率預報進行校正以修正其偏差，令降雨預報的結果更具有實用價值，讓使用者參考該預報產品做出的決策能得到最大的經濟價值。

本研究使用劇烈天氣監測系統 (QPESUMS) 的雷達估計降雨量做為校正 WEPS 降雨機率的依據，校正的方法分別為線性迴歸法 (Linear Regression, LR) 及非線性的類神經網路法 (Artificial Neural Networks, ANN)，分析其校正後的可信度與區辨能力。在校正實驗中，將訓練樣本區分為全區和陸地，陸地再細分為平地 (地形高度低於 500 公尺) 和山區 (地形高度高於 500 公尺)，比較不同的訓練樣本對於校正後的結果有何影響。

研究結果發現，線性的 LR 和非線性的 ANN 方法都能夠成功地修正 WEPS 的降雨預報。校正後明顯增加了可信度，且對於降雨事件仍具有相當不錯的區辨能力，透過 Brier Skill Score (BrSS) 的分析則證明校正後提升了預報能力。而將訓練樣本侷限在陸地，能夠將預報結果校正到十分接近完美可信 (Perfectly reliable)，並對降雨事件的區辨能力良好，預報能力也提升。

關鍵詞：系集降雨機率預報、迴歸法、類神經網路法

## 一、前言

降雨機率預報 (probabilistic quantitative precipitation forecasts, PQPFs) 是整合系集預報系統各預報成員的降雨量預報, 以一個固定的降雨門檻去比較各預報成員的預報值是否達到該降雨門檻, 計算超過該降雨門檻成員數目佔總成員數的比例。舉例來說, 若有一個系集預報系統有 10 個預報成員, 其中 2 個成員的降雨預報超過降雨門檻值 50 毫米, 則可以說降雨超過 50 毫米的機率為  $2/10$ , 即 20%。PQPFs 是系集預報的一大特色, 有別於單一決定性預報只提供固定的降雨量, PQPFs 所提供的是降雨事件發生的機率。使用者根據自身的需求決定多大的降雨量是必須關注的事件, PQPFs 則提供該降雨事件發生的機率做為決策的參考。使用者可透過經濟價值 (Economic Value, EV) 分析來評估使用者參考了系集預報而採取決策後能夠獲得多少效益 (陳等, 2016)。

Chang et al. (2012) 分析局地分析與預報系統 (Local Analysis and Prediction System, LAPS) 預報 6 小時累積降雨的 PQPFs。LAPS 的系集預報成員共 12 個, 成員間的差異分別為不同的數值模式 (MM5 和 WRF), 不同的初始場 (NFS 和 AVN) 及不同時間區間 (0-6 小時、3-9 小時及 6-12 小時), 結果發現降雨預報機率具有系統性的偏差, PQPFs 大部份高於

實際降雨事件發生的頻率。而線性的迴歸法 (LR, Yuan et al. 2008) 可系統性修正 PQPFs 偏差, 讓系統性的偏差經過校正而減小, 使得校正後的 PQPFs 更接近實際降雨發生的頻率。Yuan et al. (2007) 將類神經網路法 (artificial neural network, ANN) 應用於 NCEP 的短期系集預報模式 (short-range ensemble forecasting, SREF) 並針對美西的區域預報進行校正實驗, 亦能夠有效的修正系集預報的系統性偏差。

中央氣象局所建置的 WRF 系集預報系統 (WRF Ensemble Prediction System, WEPS) 是透過不同的初始條件、邊界條件和模式參數化的設定 (李與洪 2011), 產生 20 個系集成員 (李與洪 2014), 比 Chang et al. (2012) 所用的 LAPS, 更能代表非線性動力誤差或模式誤差等不確定性。WEPS 的 PQPFs 經過校驗後發現也具有濕偏差的特性 (吳等人, 2016), 必須要經過校正, 預報的降雨機率才會接近實際降雨的發生頻率。

本研究為臺灣地區 WRF 颱風系集降雨機率預報之評估、校正與經濟價值分析的第二部份, 接續第一部份預報評估的研究 (張等 2018), 重點在於利用 LR 及 ANN 法校正 WEPS 的系統性偏差, 令後續的經濟價值分析能夠使用接近無偏差的分析樣本, 讓使用者能夠決定獲得最大利益的降雨機率門檻值。

表 1、區域系集預報系統之物理參數法設定（李與洪 2014）

	微物理參數法	邊界層參數	積雲參數
第 1 組	GCE	YSU	Grell
第 2 組	GCE	YSU	Tiedtke
第 3 組	GCE	MYJ	Betts-Miller
第 4 組	GCE	MYJ	K-F
第 5 組	GCE	MYJ	Tiedtke
第 6 組	GCE	MYJ	Old SAS
第 7 組	GCE	MYJ	New SAS
第 8 組	GCE	ACM2	Grell
第 9 組	GCE	ACM2	Tiedtke
第 10 組	GCE	ACM2	New SAS
第 11 組	WSMS	YSU	Tiedtke
第 12 組	WSMS	MYJ	Betts-Miller
第 13 組	WSMS	MYJ	K-F
第 14 組	WSMS	MYJ	Tiedtke
第 15 組	WSMS	MYJ	Old SAS
第 16 組	WSMS	MYJ	New SAS
第 17 組	WSMS	ACM2	Grell
第 18 組	WSMS	ACM2	Tiedtke
第 19 組	WSMS	ACM2	New SAS
第 20 組	WSMS	YSU	Grell

表 2、2013 到 2015 年颱風個案列表。

	颱風名稱（編號）	個案起始和結束時間
2013 年	蘇力 Soulik (TY01)	0000UTC 11 Jul - 0000UTC 14 Jul
	潭美 Trami (TY02)	0000UTC 20 Aug - 0000UTC 22 Aug
	康芮 Kong-Rey (TY03)	0000UTC 27 Aug - 0000UTC 30 Aug
	天兔 Usagi (TY04)	1800UTC 19 Sep - 1200UTC 22 Sep
	菲特 Fitow (TY05)	1800UTC 04 Oct - 0600UTC 07 Oct
2014 年	麥德姆 Matmo (TY06)	0000UTC 18 Sep - 1800UTC 23 Jul
	鳳凰 Fung-Wong (TY07)	1800UTC 18 Sep - 1200UTC 22 Sep
2015 年	蓮花 Linfa (TY08)	1800UTC 05 Jul - 0600UTC 08 Jul
	昌鴻 Chang-Hom (TY09)	1200UTC 08 Jul - 0600UTC 11 Jul
	蘇迪勒 Soudelor (TY10)	1800UTC 05 Aug - 0600UTC 09 Aug
	天鵝 Goni (TY11)	0600UTC 20 Aug - 1800UTC 23 Aug
	杜鵑 Dujan (TY12)	1800UTC 26 Sep - 1800UTC 29 Sep

本論文第二部分主要介紹 WEPS 及觀測資料。第三部分利用簡單的數學公式說明校正 PQPFs 的過程及方法。第四部分說明驗證的方法及敏感度實驗的設計。第五部份利用信賴程度分析、區辨能力分析和預報技術分析探討校正後的預報表現。最後一部分則針對前面所得到結果給出結論及未來的展望。

## 二、模式資料與觀測資料

WEPS 的系集成員所用的數值模式為 WRF/ARW，擁有三層槽狀網格，水平解析度分別為 45、15 及 5 公里。WEPS 對微物理參數法、邊界層參數法和積雲參數法進行不同參數法之組合，物理參數法設定如表 1 (李與洪 2014)，模式所需的邊界條件來自美國 NCEP GEFS 全球系集預報系統之預報場。預報作業每日執行 4 次，分別為每日的 00、06、12 和 18 UTC。本研究採用 5 公里解析度的網格資料進行系集颱風風雨機率預報。

降雨觀測資料使用中央氣象局、經濟部水利署與美國劇烈風暴實驗室 (National Severe Storm Laboratory, NSSL) 所合作發展之劇烈天氣監測系統 (Quantitative Precipitation Estimation and Segregation Using Multiple Sensors, QPESUMS) 的雷達估計降雨 (Quantitative Precipitation Estimation, QPE) 做為觀測資料 (Gourley et al. 2001)，解析度是  $0.0125^\circ \times 0.0125^\circ$ 。QPE 受到地形和雷達位置的影響，不同區域的準確性不同，在陸地區域的 QPE 經過雨量計的

校正，有較高的正確性，距離雷達站較遠的海洋地區，雷達波束離地面較高，QPE 低估的可能性相當高。

## 三、降雨機率的校正方法

為了能夠充分的了解 WEPS 的 PQPFs 在各不同降雨門檻值的表現，本研究選擇颱風侵襲台灣的強降水事件，透過 2013 到 2015 年的 12 個颱風個案建立 25 至 500mm/day 間共 17 個降雨門檻 (25, 30, 40, 50, 60, 70, 80, 100, 130, 160, 200, 250, 300, 350, 400, 450, 500 mm/day) 之 PQPF，再針對 50, 80, 130, 200 mm/day 的 PQPFs，做為校正門檻，使用 LR 和 ANN 兩種不同的校正方法進行校正。理論上可以針對每個校正門檻進行校正，但是並不是每個門檻都是我們所關心的，氣象局所訂的大雨及豪雨標準為每 24 小時超過 80 毫米及 200 毫米，所以選擇這兩個雨量值做為校正門檻，另再選雨量值較小的 50 毫米及 130 毫米做為比較。

PQPF 的校正過程分為訓練和驗證兩部份。在訓練的部份，需要一組訓練資料，訓練資料包括以校正門檻為中心的 7 個相鄰門檻之 PQPFs ( $f_i(x, t)$ ) 和 1 個觀測機率。所謂的觀測的降雨機率，是指觀測降雨量是否超過所設定的門檻值，分別以 1 代表降雨大於門檻值和 0 代表小於門檻值。

$$\varepsilon = \sum_{j=1}^N \{P_{obs}(x_j, t_j) - P_{ens}(x_j, t_j)\}^2 \dots \quad (1)$$

公式(1)表示觀測機率與 PQPFs 的平方差， $P_{obs}$  為觀測的降雨機率， $N$  為所訓練的

樣本數， $P_{ens} = P(f_1, f_2, \dots, f_7)$  為利用訓練資料中 7 個相鄰門檻之 PQPFs 透過函數關係所得到的校正後 PQPFs，此函數關係可以是線性，也可以是非線性。接下來探討的數學方法即是為了找尋一個最佳的關係，令公式(1)有最小值，表示訓練資料得到最佳的校正結果。完成訓練之後再行驗證的部份，另取一組與訓練資料完全不同的資料做為驗證，同樣取 7 個相鄰門檻之 PQPFs，代入訓練所使用的關係式中，即可得到校正後的機率。

### (一)、迴歸法(Linear Regression)

假設某降雨門檻校正後的 PQPFs 與 7 個相鄰降雨門檻的 PQPFs 是線性的關係[公式(2)] (Yuan et al. 2008, Chang et al. 2011)

$$P_{ens}(x, t) = a + \sum_{i=1}^M b_i f_i(x, t) \quad \dots(2)$$

$$J = \min \sum_{j=1}^N \{P_{obs}(x_j, t_j) - [a + \sum_{i=1}^M b_i f_i(x_j, t_j)]\}^2 \quad \dots(3)$$

其中  $M=7$ ， $f_i(x, t)$  是 7 個以校正門檻為中心的 7 個相鄰門檻之 PQPFs，常數  $a$  和  $b_1 \sim b_7$  為是待求的未知系數。公式(2)代入公式(1)可寫成公式(3)，令公式(3)中  $J$  對各迴歸係數進行一階微分為零，透過訓練資料便可建立線性方程組，並可解得  $a$  和  $b_1 \sim b_7$  使公式(3)有最小值。再將驗證樣本資料套入公式(2)中，則可得到校正後的 PQPFs。

### (二)、類神經網路法(Artificial Neural Network)

假設修正後的 PQPFs 與校正門檻為中心的 7

個相鄰門檻之 PQPFs 值為類神經網路

(ANN) 所建構的非線性關係，圖 1 為說明架構的示意圖。ANN 的架構分為三部份，分別是輸入層 (Input Layer)、隱藏層 (Hidden Layer) 和輸出層 (Output Layer)，分別有 7 個、4 個和 1 個神經元。輸入層的神經元分別是 7 個降雨門檻的 PQPFs。不同層的神經元彼此相連接，每個神經元所接收到的資料量為來自上一層傳送來的總合 (章等 2017)。

$$P_{ens}(x, t) = Z_k(x, t) = S(a^{hidden} + \sum_{j=1}^n u_{j,k} y_j) = S(a^{hidden} + \sum_{j=1}^n u_{j,k} S(a^{input} + \sum_{i=1}^m w_{i,j} x_i)) \quad \dots(4)$$

$$S(a) = \frac{1}{1+e^{-a}} \begin{cases} -\infty < a < \infty \\ 0 < s < 1 \end{cases} \quad \dots(5)$$

$$J(u_{j,k}, w_{i,j}) = \min \sum_{l=1}^N \frac{1}{2} \{P_{obs}(x_l, t_l) - S[\sum_{j=1}^{n+1} u_{j,k} S(\sum_{i=1}^{m+1} w_{i,j} f_i)]\}_l^2 \quad \dots(6)$$

公式(4)為輸出層與輸入層的非線性關係，等號左項是輸出層的數值，亦是調整後的 PQPFs， $w_{ij}$  為輸入層第  $i$  個神經元與隱藏層的第  $j$  個神經元的加權值， $u_{jk}$  則為隱藏層第  $j$  個神經元與輸出層第  $k$  個神經元的加權值。公式(4)中包含  $S$  函數[公式(5)]，功能為調整各神經元接收到的訊號總量在 0 到 1 之間， $N$  為樣本數目， $m=7$  (輸入層神經元個數)  $n=4$  (隱藏層神經元個數)。公式(4)代入公式(1)得到一個價值函數 Cost Function [公式(6)]，將訓練資料代入  $f_i$  並令其有極小值，可得到神經元之間的權重。利用 Linear Least Square SIMplex

(LLSSIM, Hsu 1996) 求得公式(6)極小值，此方法同時使用了線性的最小平方法 (Linear Least Square) 及 Multi-start Simplex Search

Algorithm 求得最佳的 $w_{ij}$ 和 $u_{ij}$ ，使價值函數有最小值。將取得的 $w_{ij}$ 和 $u_{ij}$ ，與驗證資料輸入公式(4)，可得到校正後的 PQPFs。

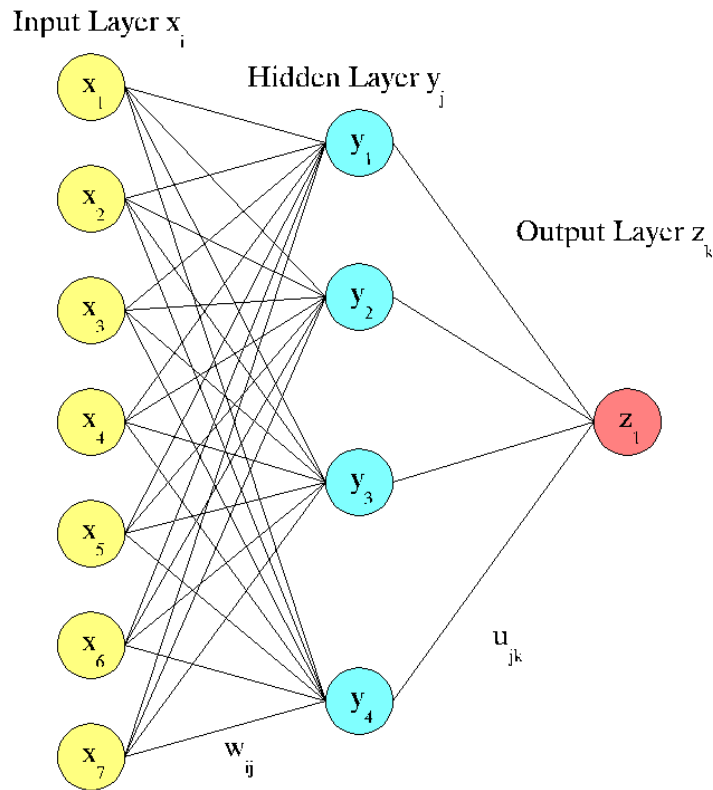


圖 1、類神經網路架構圖，輸入層神經元共七個（黃色），隱藏層有四個（藍色），輸出層有一個（紅色）。

#### 四、敏感度實驗設計

本研究使用 WEPS 的 0-24 小時累積降雨預報，2013-2015 共 12 個颱風個案如表 2。實驗的目的除了想探討不同的校正方法 (LR 和 ANN) 對校正結果有什麼不同之外，也想測試降雨的地區 (平地、山區或海洋) 不同，對於

校正的結果有什麼影響。陳等 (2016) 評估過 WEPS 的 PQPFs，不同的域預報偏差程度亦不同，可以猜想分開校正亦有不同的效果，因此在敏感度實驗中，模式範圍區分為全區、平地 (地形高度低於 500 公尺) 和山區 (地形高度高於 500 公尺)。

表 3、PQPF 校正實驗的設計與樣本說明。

實驗名稱	說明
EXP-T	統計樣本取自 QPESUMS 範圍內所有雷達覆蓋區域(包含海洋和陸地)；未經後端校正程序，本研究以此實驗做為參考實驗。
EXP-L	同 EXP-T，但統計樣本取自臺灣陸地區域。
EXP-P	同 EXP-T，但統計樣本取自臺灣平地(地形高度< 500 m)。
EXP-M	同 EXP-T，但統計樣本取自臺灣山區(地形高度≥ 500 m)。
EXP-T-LR	統計樣本同 EXP-T；採用 LR 方法進行校正；進行交叉驗證時，將統計樣本分為 12 組，分別是 12 個颱風個案(TY01~TY12)。
EXP-T-ANN	同 EXP-T-LR；但採用 ANN 方法進行校正。
EXP-L-LR	同 EXP-T-LR，但統計樣本取自臺灣陸地區域。
EXP-L-ANN	同 EXP-T-LR，但統計樣本取自臺灣陸地區域。
EXP-P-LR	同 EXP-T-LR，但統計樣本取自臺灣平地(地形高度< 500 m)。
EXP-M-LR	同 EXP-T-LR，但統計樣本取自臺灣山區(地形高度≥ 500 m)。

表 3 為 PQPF 校正實驗的設計與樣本說明，實驗名稱中 EXP-之後的字母代表樣本來源，T 代表全部的樣本 (total area)，L 代表實驗樣本僅來自台灣陸地區域 (land area)，M 和 P 代表來自陸地區域的山區 (mountain area) 和平地 (plain area)，除了參考實驗外，其他的實驗名稱再加所使用的校正方法是 LR (線性迴歸法) 或 ANN (類神經網路法) 以區別之間的不同。

為增加驗證樣本以提高統計校驗結果的代表性，我們是採用交叉驗證 (cross-validation) 的方式來進行校正；以其中 11 個颱風個案作為訓練樣本去校正剩餘的那 1 組 (驗證樣本)，亦即每個颱風個案輪流做為驗證樣本，最後再結合所有驗證樣本於校正結果中。

## 五、分析方法與結果討論

### (一) 信賴程度分析

Reliability diagram 可用以了解系集預報系統所預報的機率與實際觀測頻率之間的對應狀況。所謂的實際觀測頻率是統計所有訓練資料中 PQPFs 相同的格點資料，降雨量達到降雨門檻的比率。若兩者對應的情形一致，則預報的可信度越高，並且能夠進一步評估系集機率預報的偏差 (高報或低報)。從圖形上來解讀，完美可信 (Perfectly reliable) 的預報系統，Reliability diagram 中的信賴曲線 (Reliability curve) 將會落在對角線上，若預報機率高於觀測頻率，表示模式呈現過度預報 (Over-forecasting)，此時信賴曲線會落在對角線下方；反之，則落在對角線上方。水平虛線為氣候發生頻率，小於 10% 時屬於較罕見的極端事件，本研究選擇的雨量校正門檻 (50,

80, 130, 200 mm/day) 均達此標準 (圖 2), 表示所定訂的校正門檻是我們所需要關心的降雨事件。

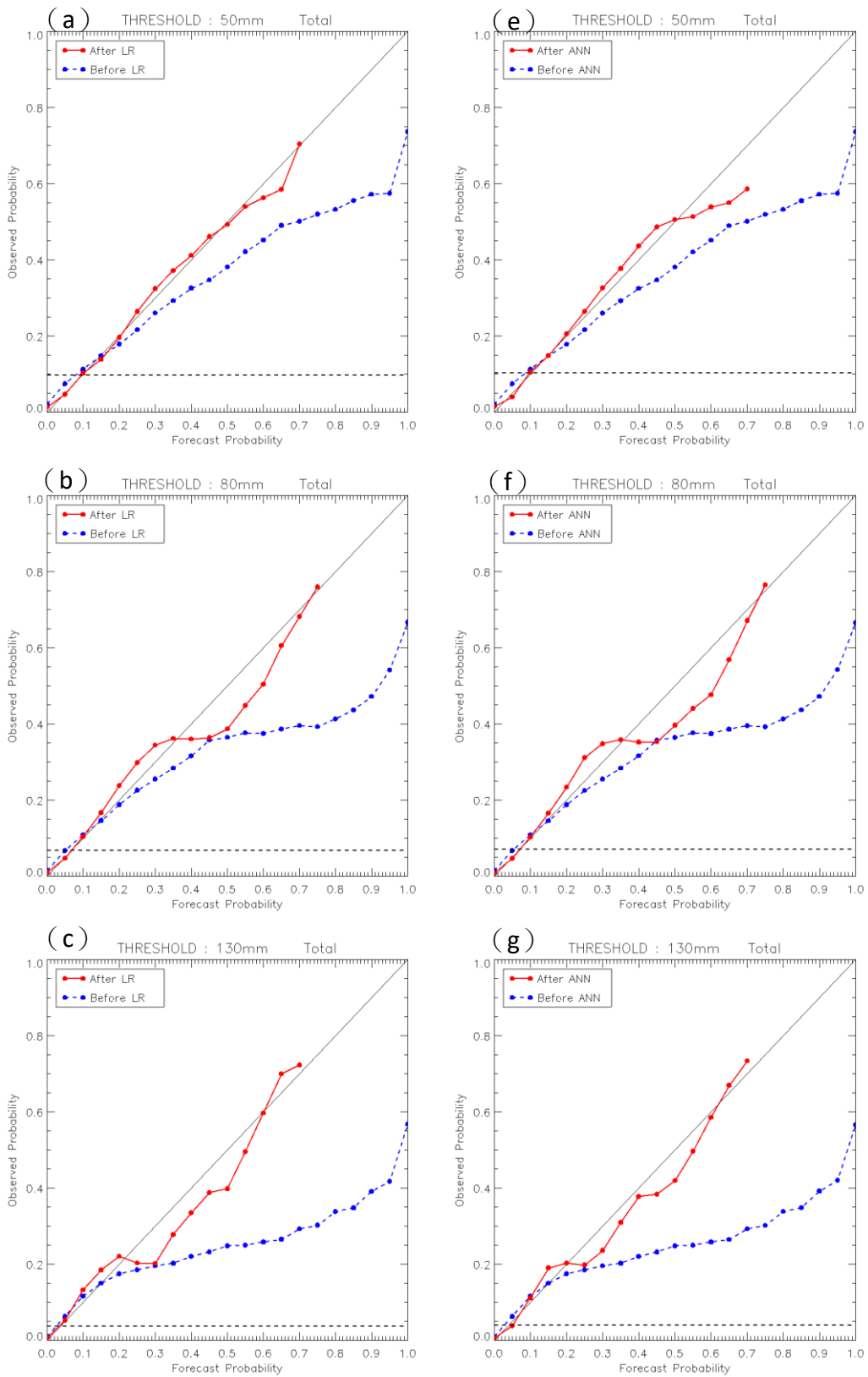
利用 Reliability diagram 比較校正前 (EXP-T) 與校正後 (EXP-T-LR、EXP-T-ANN) 的信賴程度 (圖 2), 藍色的虛線是 EXP-T 的分布情形, 大部份落在對角線的右下方, 表示 WEPS 的 PQPFs 總是有高估的情況, 降雨的門檻值愈高, 現象愈明顯。紅色實線是經過校正後的分布情況, 左欄是經過 LR 的校正 (EXP-T-LR), 右欄是 ANN 的校正 (EXP-T-ANN), 不論使用何種方式, 校正後的可信度提高, 比校正前 (EXP-T) 更接近對角線。比較 LR 和 ANN 兩種校正方式, 發現校正後的結果十分相似, 表示 WEPS 的系統性偏差對於校正方法不敏感, 不論線性或非線性的方法均能夠校正。

圖 2 校正後的結果與完美可信的情況尚有一段差距 (紅色實線仍沒有貼齊在對角線上), 並隨著預報機率高低而出現起伏的變化, 因為訓練樣本 (training sample) 和驗證樣本 (validation sample) 的預報偏差不一致所造成的。偏差不一致的原因可能是做為降雨量參考的 QPESUMS 資料在陸地和海洋上出現不一致的情況, 如同第二節中所提及, 降雨估計受雷達仰角的限制容易低估, 而陸地部份的估計則經過地面雨量站資料的校正而減少了

估計的偏差。將訓練和校正的資料侷限於陸地, 可以避開降雨估計海面 and 陸地上不一致的問題, 使訓練樣本和驗證樣本的預報偏差相對一致, 因此嘗試了 EXP-L-LR 和 EXP-L-ANN 兩組實驗, 與 EXP-L 進行比較, 圖 3 為只選取陸地部份樣本進行訓練和校正的結果, 校正前 (EXP-L) 的分布已經比圖 2 的 EXP-T 更接近對角線, 但機率小的部份有稍為低估 (落在對角線上方), 而機率高高的部份高估 (落在對角線下方) 隨著降雨門檻增加, 越為明顯, 經過校正後明顯十分貼近對角線, 使用 LR 和 ANN 校正均有相同的效果, 特別是降雨門檻低的部份, 由於達到降雨門檻的樣本數充足, 校正後幾乎達到完美 [圖 3(a)(b)(e)(f)]。

若再把陸地區分為山區和平地, 可以發現 PQPF 的偏差具有地域性, 在小雨量的門檻 (50mm 和 80mm), 山區 (EXP-M) 明顯呈現濕偏差 [圖 4(e)和(f)], 在平地 (EXP-P) 則沒有這麼明顯的濕偏差, 低 PQPF 低 (低於 50%) 還出現乾偏差 [圖 4(a)和(b)], 經過校正後可以完美的修正偏差。而高累積雨量降雨門檻 (130mm 和 200mm) 濕偏差明顯出現在系集預報中, 但再經過校正後使得達到這兩個門檻的樣本數很少, 只有低機率值有數值 [圖 4(c)(d)(g)(h)], 不利於後續經濟效益等分析評估。





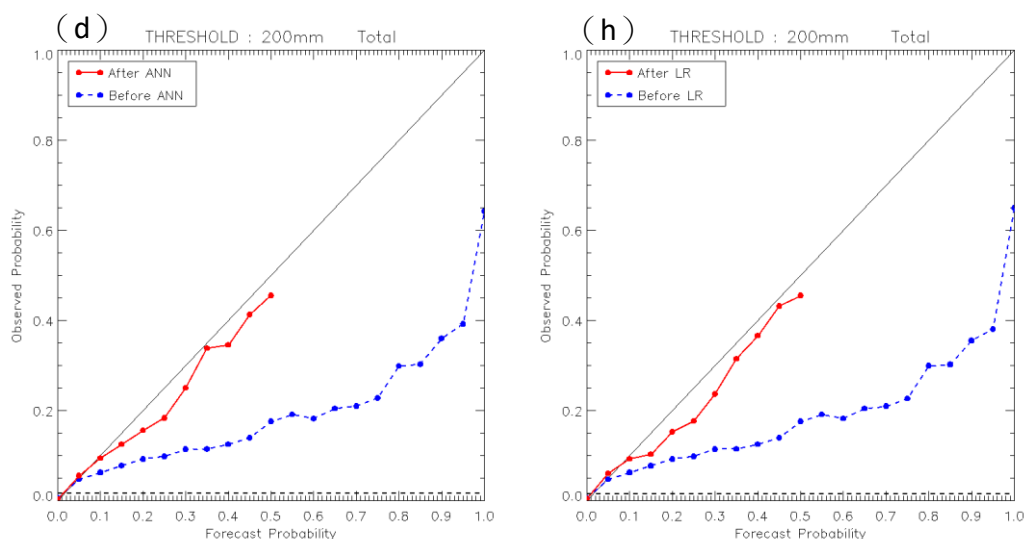
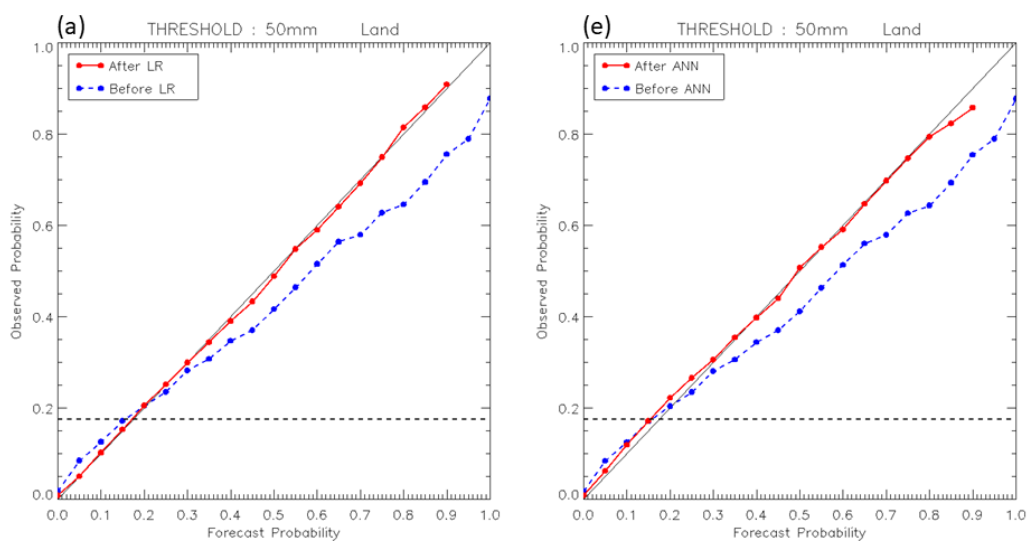


圖 2、WEPS 的 PQPF 在不同降水門檻(由上而下分別是 50、80、130、200 mm/day)的 reliability diagram，橫軸為預報的 PQPF，而縱軸是該 PQPF 的達到降雨門檻的發生頻率，右欄和左欄分別顯示兩種不同校正方法校正後的結果，左右兩欄的藍色虛線均為 EXP-T 的分布，而左欄紅色實線為 EXP-T-LR，右欄為 EXP-T-ANN 的結果。水平黑色虛線為該降雨門檻的氣候發生頻率。



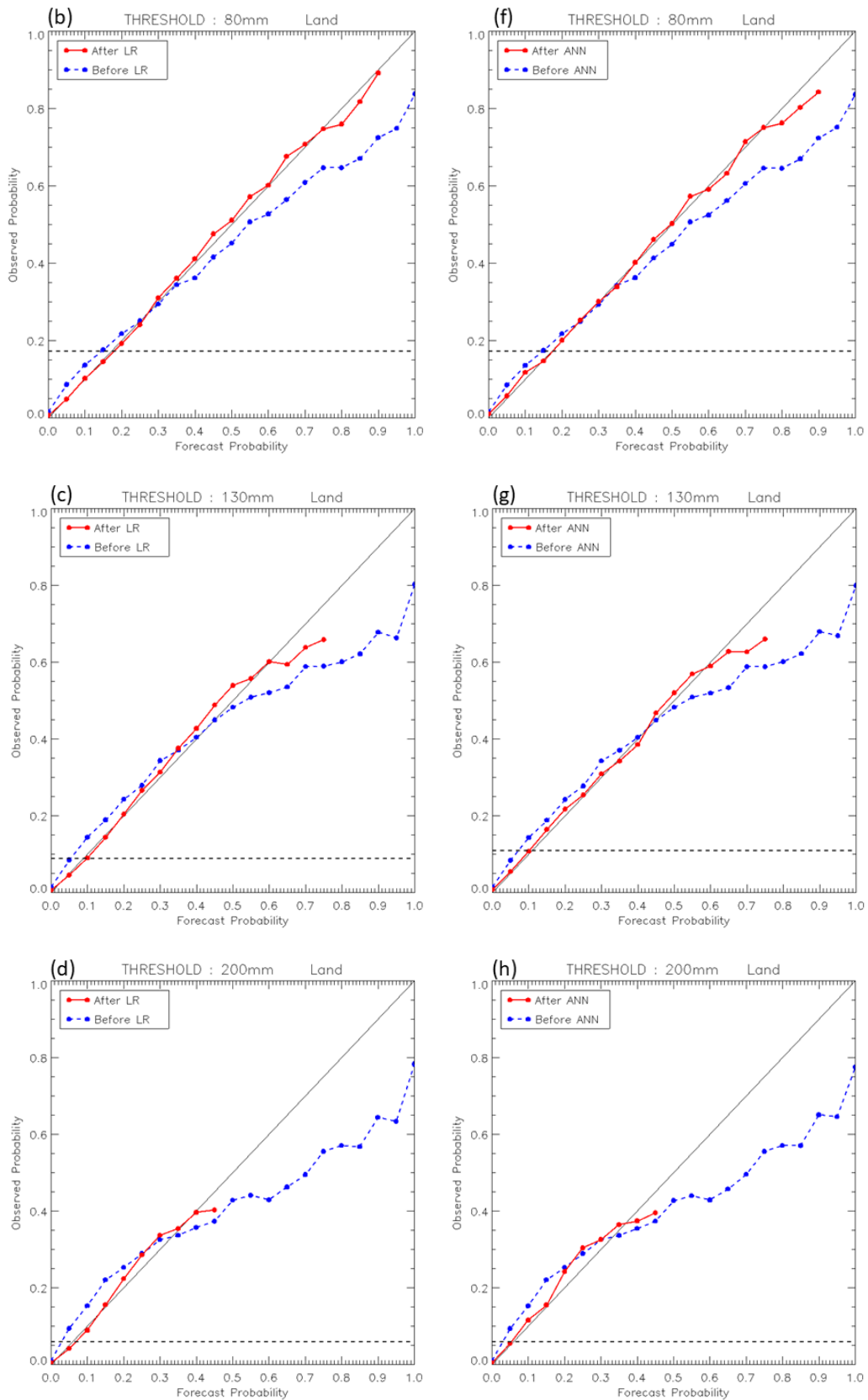
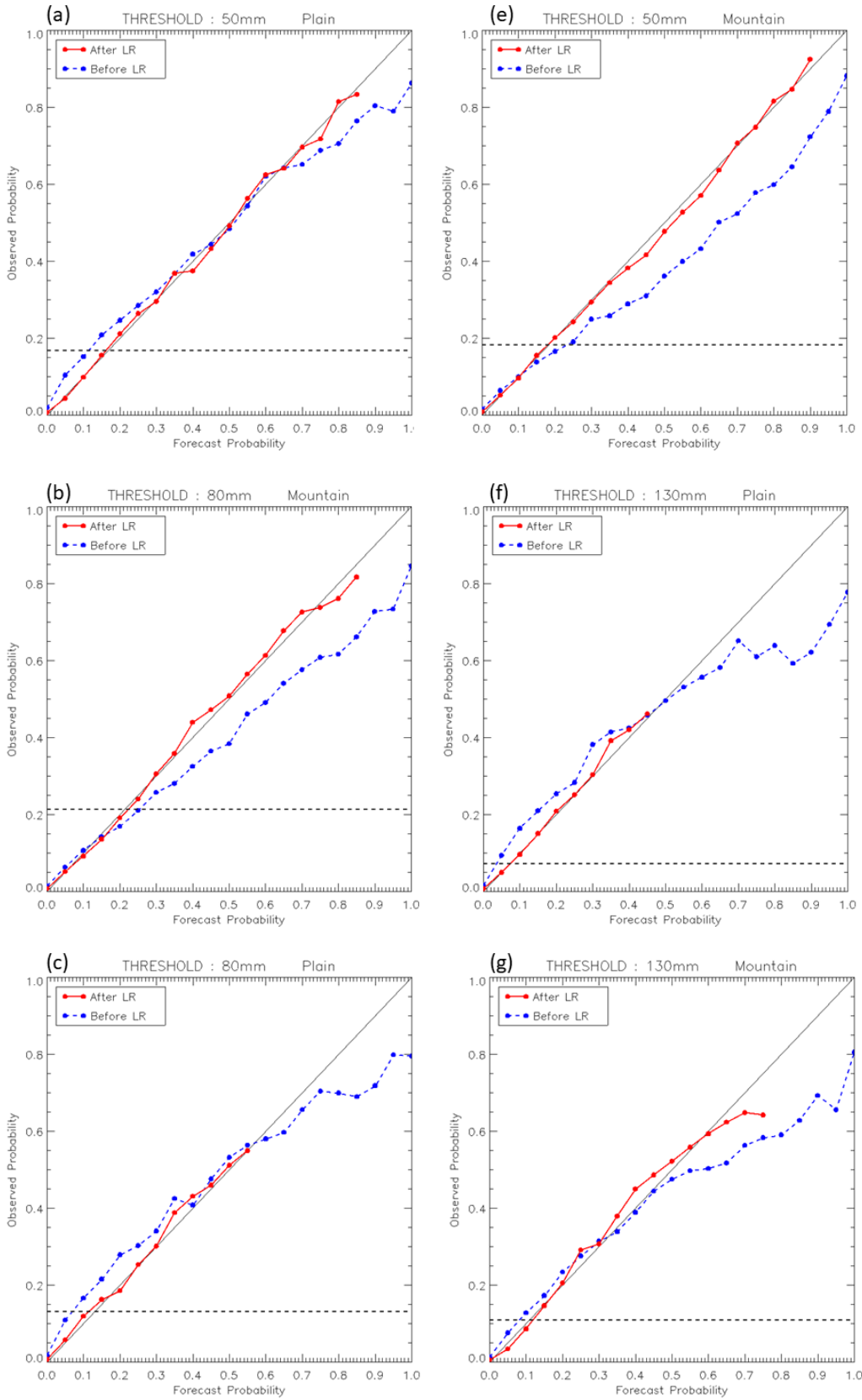


圖 3、同圖 2，不同在於訓練的資料樣本來自於陸地，分別是 EXP-L（藍色虛線）、EXP-L-LR（左欄紅色實線），EXP-L-ANN（右欄紅色實線）的結果。



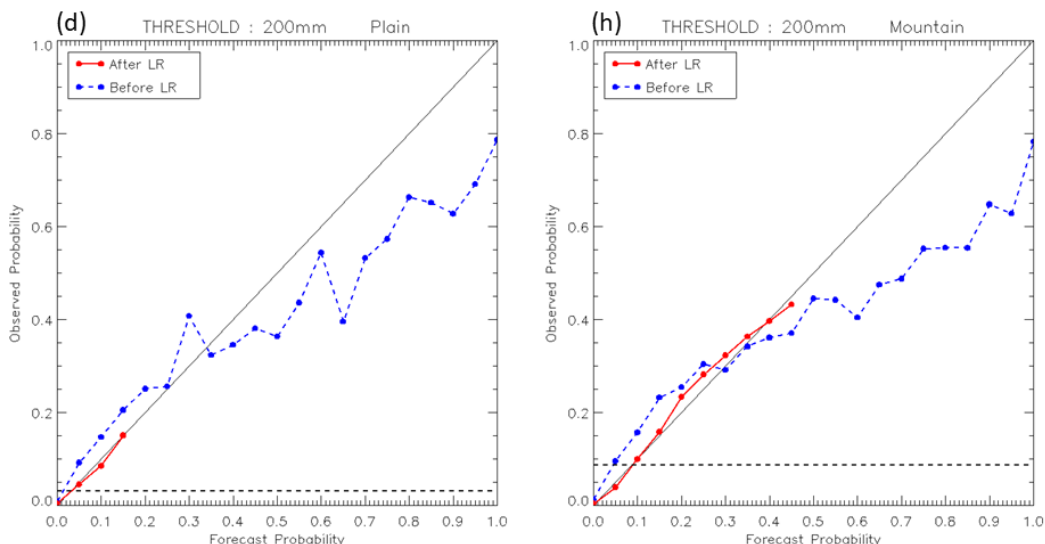


圖 4、訓練的資料樣本來自於平地（左欄）和山區（右欄），左欄的藍色虛線均為 EXP-P 的分布，紅色實線為 EXP-P-LR，右欄的藍色虛線均為 EXP-M 的分布，紅色實線為 EXP-M-LR。

## （二）區辨能力分析

### Relative operating characteristic

（ROC）可用以評估模式對於降雨事件的發生或不發生是否具有區辨能力。利用一組遞增的機率做為預警門檻，當預報機率大於或等於此門檻時，代表預報有事件發生。用圖形來表現則以 y 軸是命中率（hit rate）、x 軸是錯誤預報率（false alarm rate），繪出分布曲線。應用上主要以 ROC 曲線下的面積大小來評估系集預報系統的區辨能力；只要 ROC 曲線下面積大於 0.5（亦即對角線下方面積），即表示系集預報系統具有區辨該氣象事件之能力，若 ROC 曲線下面積大於 0.7 則表示具有良好的區辨能力（Buizza et al. 1999）。

將 EXP-T 和 EXP-T-LR 的校正前後的結果，依不同校正門檻的 ROC 面積大小畫成折線圖

（圖 5），不論校正前或校正後，均大於 0.8，表示原本的預報就具有良好的區辨能力，而校正後也有依然保有相同的區辨能力。圖 6 的統計樣本來自陸地的部份（EXP-L 和 EXP-L-LR），發現 EXP-L 比 EXP-T（圖 5）還要高，表示只取用陸地的資料進行分析，有較高的區辨能力。同樣的，修正前後（EXP-L 和 EXP-L-LR）均保有優異的區辨能力（圖 6）。ROC 的分析證明，校正改善信賴程度，而預報的區辨能力不受影響。

## （三）預報技術分析

Brier skill score (BrSS) 是用來評估系集預報在預報降雨事件時，相較於氣候預報的改善程度。BrSS 的定義如下（公式 7）：

$$BrSS = \frac{BrS - BrS_{ref}}{BrS_{pref} - BrS_{ref}} = 1 - \frac{BrS}{BrS_{ref}} \quad \dots(7)$$

$$BrS = \frac{1}{N} \sum_{j=1}^N (P_j - O_j)^2 \quad \dots(8)$$

$$BrS = \left[ \frac{1}{N} \sum_{i=1}^K n_i (P_i - \bar{O}_i)^2 \right] - \left[ \frac{1}{N} \sum_{i=1}^K n_i (\bar{O}_i - \bar{O})^2 \right]$$

$$O_{avg})^2] + [O_{avg}(1 - O_{avg})] \quad \dots(9)$$

$$BrSS = 1 - \frac{BrS}{BrS_{ref}}$$

$$= 1 - \frac{reliability - resolution + BrS_{ref}}{BrS_{ref}}$$

.....(10)

在 *BrSS* 中的 *BrS* 是預報系統的機率預報誤差大小[公式(8)]，*N* 是所有校驗樣本數，*P<sub>j</sub>* 是預報機率，*O<sub>j</sub>* 是觀測機率。下標 *ref* 為樣本氣候值。*BrSS* 的值域介於負無限大與 1 之間，若 *BrSS*=0 時，表示機率預報相對於氣候預報而言，不具有預報能力；當 *BrSS*>0 時，機率預報相對於氣候預報而言，具有預報能力；完美預報的 *BrSS*=1。

*BrS* 可分解成 *reliability*, *resolution* 以及 *uncertainty* 三項 (Murphy 1973)，依序是公式(9)的右側第一、第二及第三項，公式(9)中的 *N* 代表所有的校驗樣本數，*K-1* 是預報機率區間的總和（在本研究中 *K=21*），*n<sub>i</sub>*、*P<sub>i</sub>*、*O<sub>i</sub>* 分別代表預報機率區間 *i* 的校驗樣本數、預報機率之中間值及條件性觀測頻率，*O<sub>avg</sub>* 是樣本氣候頻率。*Reliability* 項代表 *PQPF* 的條件性偏差的大小，是負向指標，愈小代表結果愈好，*resolution* 項代表分辨極端事件的能力，是正向指標，愈大表示跟氣候預報比起來愈能夠區辨降水事件，*uncertainty* 為樣本的觀測變異，只和樣本的氣候頻率有關，並不會在校正過程中有所改變。

將公式(9)經過整理之後，可將 *BrSS* 改寫

成公式(10)，發現降低 *reliability* 項和提高 *resolution* 項均能提高 *BrSS*，而計算校正前後各項的變化量，就能夠知道 *PQPF* 經過校正後提升了 *BrSS*，主要的貢獻來自於何項。下列以 *LR* 的結果為例討論。

圖 7(a)顯示參考實驗 *EXP-T* 和 *EXP-T-LR* 依不同校正門檻 *BrSS* 數值的分布情況，校正前 (*EXP-T*，圖中的藍色虛線) *BrSS* 的數值隨門檻值增加而快速降低，大雨量 (降雨門檻大於 100mm) 的 *BrSS* 低於 0，表示相對於氣候預報而言，不具有預報能力。經過校正後可以改善此現象 (*EXP-T-LR*，圖中的紅色實線)，將 *BrSS* 提高到 0.35 與 0.1 之間，表示機率預報與氣候預報相比具有預報能力。圖 7(b)是 *BrSS* 中 *reliability* 項和 *resolution* 項的分布，表示 *BrSS* 的提高主要是透過校正程序降低 *reliability* 項來達成的，而 *resolution* 在校正前後並沒有明顯改變，在 *ROC* 的分析也証實了校正程序不會影響區辨能力，原因是公式(1)在數學上的意義與公式(9)的 *reliability* 項相近，因此極小化公式(1)的作用等同於降低公式(9)的 *reliability* 項。

若是訓練的樣本只採用陸地的部份，則是 *BrSS* 在修正前就十分具有預報能力，修正後僅稍為增加 *BrSS* 的數值[圖(8a)]，分項的結果也與 *EXP-T* 和 *EXP-T-LR* 相同，*BrSS* 的提升靠降低 *reliability* 來達成[圖 8(b)]。校正可大幅改善模式預報可信度，進而提升 *BrSS*，但對於改善區辨率並沒有明顯差異。

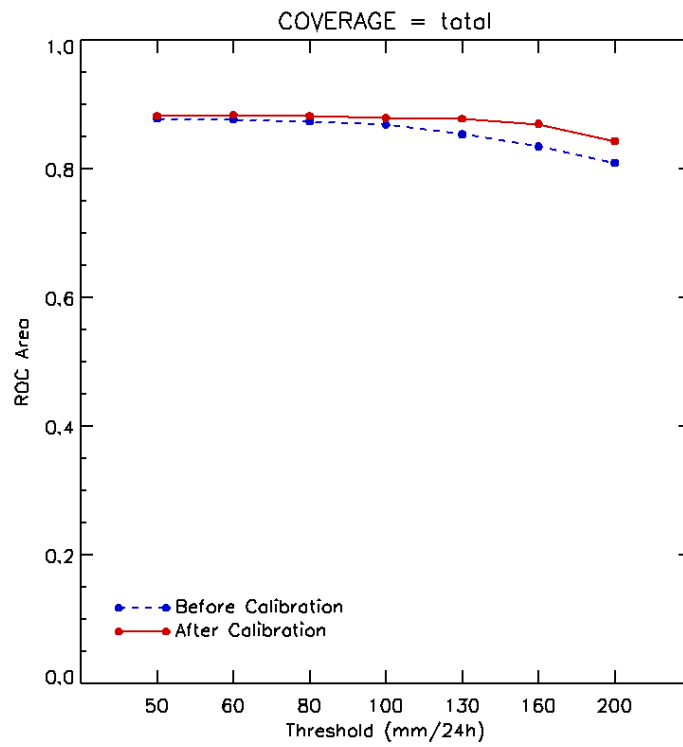


圖 5、 EXP-T (藍色虛線) 和 EXP-T-LR (紅色實線) 的 PQPF 在不同降水門檻下的 ROC 面積。

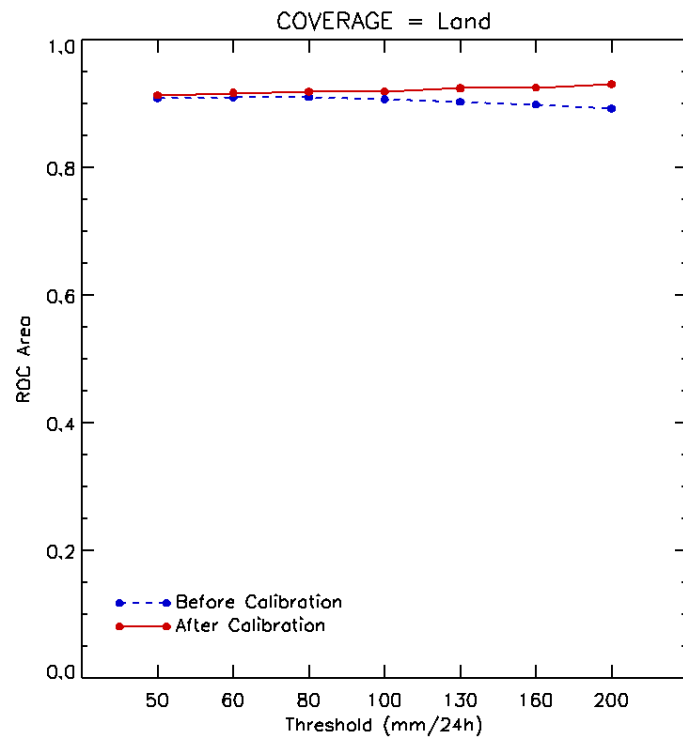


圖 6、同圖 5，為 EXP-L 和 EXP-L-LR 的 ROC 面積。

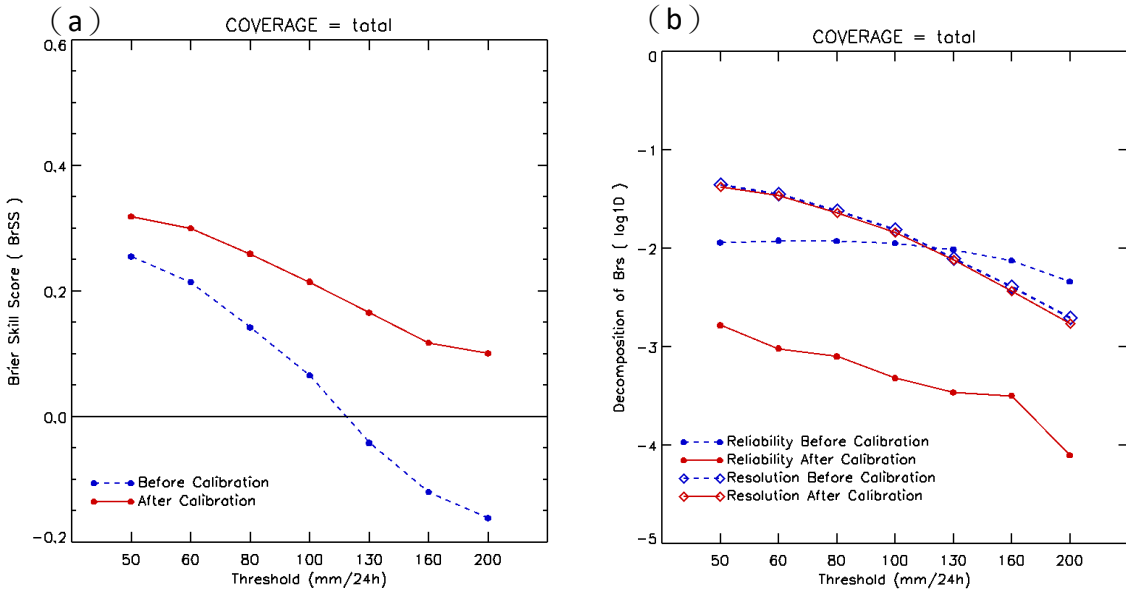


圖 7、EXP-T (藍色虛線) 和 EXP-T-LR (紅色實線) 的 PQPF 在不同降水門檻下的(a) Brier Skill Scores(BrSS), (b)Brier Score 分解，圖形實心和菱形空心標記分別是 Reliability 和 Resolution 項。

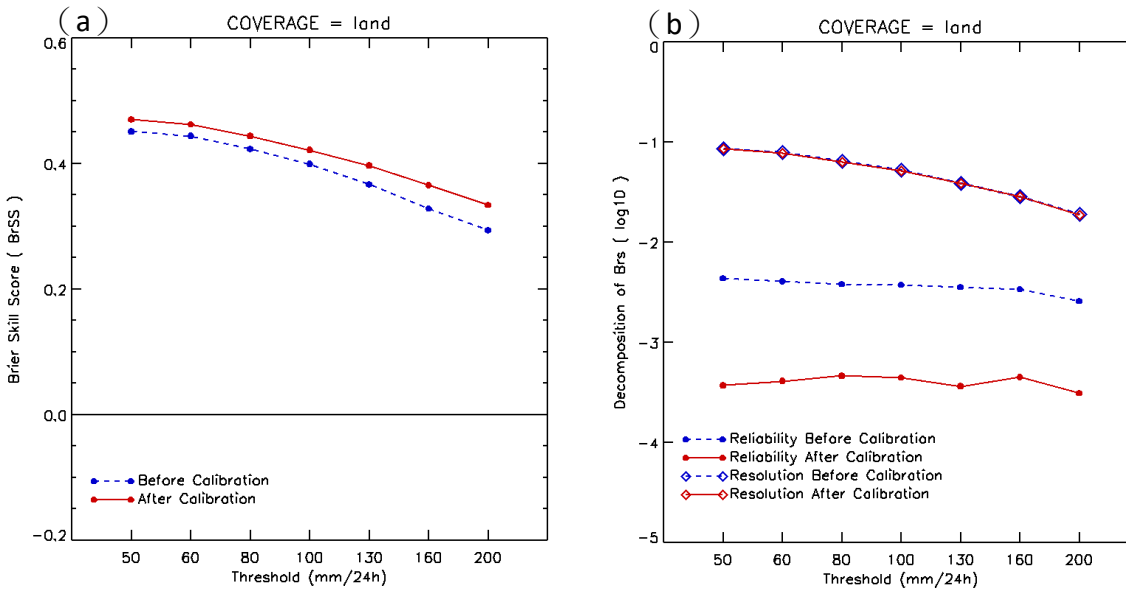


圖 8、同圖 7，訓練的樣本資料是取用陸地的部份。

## 六、結論與未來展望

本研究針對 WEPS 的 PQPFs 進行後端的預報偏差校正。欲得到良好的校正結果，需要具有相似氣候特徵的長期資料做為訓練樣本，依地形或氣候特性進行分區和大量的樣本數

目進行訓練 (Yaun et al. 2008)，因此收集了 2013 到 2015 年的 12 個颱風的 WEPS 預報資料進行實驗，以了解在不同降雨門檻值的表現。透過敏感度實驗來探討影響校正結果的因子，包括不同校正方法以及不同的地表分



區對校正結果的影響，不同分區包括全區和陸地、高山和平地。

測試結果顯示不管採用線性迴歸 (LR) 或非線性的類神經網路 (ANN)，校正結果非常相似，不論線性或非線性的方法均能有效修正預報偏差，但 LR 的計算簡單且快速，因此在 LR 和 ANN 方法校正結果不分軒輊的情況下，接下來的所有的討論都使用 LR 的校正結果。簡而言之，WEPS 的系統性誤差是可以經過校正而改善的，經過解讀 Reliability diagram 證明信賴程度經過校正明顯提升，ROC 的分析結果保證了校正程序不會讓預報對降雨事件失去區辨能力，而 BrSS 的分析證明校正後的預報優於校正前，表示預報能力有所提升。再進一步的分解 BrSS 為 reliability, resolution 以及 uncertainty 三個項目後發現，預報能力提升主要是靠降低 reliability 達成，代表校正程序降低機率預報的條件性偏差 (conditional bias)，而這明顯的改變在 Reliability diagram 也可以看到。Resolution 表示機率預報在區辨降水事件發生與否時，是否能有別於氣候預報的能力，在分析後發現影響不大，這與 ROC 的分析結果一致，校正並不會改變預報對降雨事件的區辨能力。

為了讓校驗的結果更為可信，將訓練的樣本侷限在陸地，以消除 QPESUMS 受限於雷達觀測及雨量站分布於陸地所造成的降雨估計可能不一致的情況，好處是能夠有較精確的降雨估計做為校正的參考值，但缺點是大

大減少了訓練用的樣本數 (QPESUMS 範圍內所有雷達覆蓋區域海洋和陸地的比例是 9 比 1)。經由分析 Reliability diagram，以及計算 ROC 的面積及 BrSS 的數值所得到的結果顯示，EXP-L 比 EXP-T 擁有更佳的信賴度、區辨能力和預報能力，這表示陸地上的降雨預報有較小的系統性偏差。經過校正 (EXP-L-LR) 後則信賴程度和預報能力均有提升，特別是在 Reliability diagram 上看到十分接近完美可信 (perfectly reliable) 的情況 (圖 3(a) 的紅色實線十分貼近對角線)，表示在陸地區域的訓練樣本和驗證樣本有一致性的預報偏差分布，這對於校正很有幫助。

若再將陸地區分為山區和平地，樣本數目比陸地再減少一半 (山區和平地的比例約 1 比 1)，在雨量門檻值較低時，透過 Reliability diagram 分析仍能會有不錯的結果，但是在雨量門檻值較高的情況，則是降水達到這門檻的樣本數很少，校正後的結果只有低機率值有數值，對於後續經濟效益的分析評估並沒有幫助。

本研究專注於校正 WEPS 對颱風降雨預報的系統性偏差，比較了不同校正方法，但是並沒有如預期想像，非線性的類神經網路法能夠優於線性迴歸法。為什麼複雜的數學計算沒有辦法得到更好的校正結果？是樣本數不足？還是颱風個案使用 LR 就足以修正系統性的偏差？原因可能是目前的 EPS configuration 並未抓到非線性的降水特徵，也

可能和 ANN 的參數設定有關，若要仔細探究 ANN 和 LR 之間的校正結果的相同或相異，可以從上游的 ESP 的降雨預報和 ANN 細部的參數設定來分析。未來可再進一步測試各種不同的天氣型態，例如梅雨或午後對流，分析 WEPS 的預報對颱風之外的強降水事件的預報能力，並用不同方法校正其預報結果，期望能找出使用 LR 無法適當校正但 ANN 能勝任的天氣型態，使校正後的 PQPFs 應用於各種強降水事件的決策分析中均能有所助益。

## 致謝

感謝中央氣象局資訊中心提供 WEPS 降雨預報資料，以及衛星中心提供 QPESUMS 雷達降水估計資料和電腦計算資源，使本文能順利完成。本研究是在中華民國科技部專題研究計畫補助下完成，張惠玲由 MOST-104-2625-M-052-004、MOST-105-2625-M-052-001 以及 MOST-106-2625-M-052-001 所支持；楊舒芝由 MOST 106-2625-M-008-015 所支持。

## 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。
- 李志昕、洪景山，2014：區域系集定量降水預報之應用與分析研究。2014 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。
- 章鶴群、劉豫臻、張惠玲、洪景山、楊舒芝，2017：利用類神經網路方法進行 WRF 系集預報系統之颱風降雨機率預報偏差校正。2017 年天氣分析研討會。中央氣象局。
- 陳冠儒、張惠玲、吳佳蓉、汪琮、洪景山、楊舒芝，2016：中央氣象局 WRF 系集預報系統之颱風風雨機率預報評估。2016 年天氣分析研討會。中央氣象局。
- 陳冠儒、張惠玲、吳佳蓉、汪琮、洪景山、楊舒芝，2017：中央氣象局 WRF 系集預報系統之經濟價值分析。2017 年天氣分析研討會。中央氣象局。
- 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。
- Buizza, R., 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.
- 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. 447-491.

Gourley, J. J., J. Zhang, R. A. Maddox, C. M. Calvert, and K. W. Howard, 2001: A real-time precipitation monitoring algorithm—Quantitative Precipitation Estimation Using Multiple Sensors (QPE-SUMS). Preprints, Symp. on Precipitation Extremes: Prediction, Impacts, and Responses, Albuquerque, NM, *Amer. Meteor. Soc.*, 57–60.

Hsu, K. L., H. V. Gupta, and S. Sorooshian, 1996: A Superior training strategy for three-layer feedforward artificial neural networks. Department of Hydrology and Water Resources, University of Arizona. No. 96-030.

Murphy, A. H., 1973: A new vector partition of the probability score. *J. Appl. Meteor.*, **12**, 595-600.

Yuan, H., X. Gao, S. L. Mullen, S. Sorooshian, J. Du, and H. H. Juang, 2007: Calibration of probabilistic quantitative precipitation forecasts with an artificial neural network. *Wea. Forecasting*, **22**, 1287–1303.

Yuan, H., J. A. McGinley, P. J. Schultz, C. J. Anderson, and C. Lu, 2008: Short-range precipitation forecasts from time-lagged multimodel ensembles during the HMT-West-2006 campaign. *J. Hydrometeor.*, **9**,

# **Evaluation, calibration and economic value analysis of the probabilistic forecasts from WRF ensemble prediction system in Taiwan area- part 2: forecast calibration**

**Hochin Chang<sup>1</sup>, Guan-Ru Chen<sup>1</sup>, Yu-Jhen Liou<sup>1</sup>, Hui-Ling Chang<sup>1</sup>,  
Jing-Shan Hong<sup>1</sup>, and Chu-Chih Yang<sup>2</sup>**

**1 Central Weather Bureau ,Taipei , Taiwan**

**2 Department of Atmospheric Sciences, National Central University , Jhong-Li,  
Taiwan**

(manuscript received 19 April 2018 ; in final form 27 June 2018)

## **Abstract**

It is known that the WRF Ensemble Prediction System (WEPS) of the Central Weather Bureau has large systematic bias in precipitation. This study attempts to correct the systematic forecasting bias in WEPS, and improve the results of rainfall forecast to have more practical value. Users can obtain maximum economic value by making decisions based on this forecast probability product.

In this study, the estimate rainfall rate of Quantitative Precipitation Estimation and Segregation Using Multiple Sensors (QPESUMS) was used as the true values for correcting the PQPFs of WEPS. The correction methods were Linear Regression (LR) and Artificial Neural Networks (ANN). Calibration results show that these two techniques successfully correct the wet bias and improve the forecast skill, and calibration effect of these two techniques are quite similar. Verification results from different areas show that there are better reliability and discrimination over land areas after calibration than all area.

Key word: probabilistic quantitative precipitation forecasts (PQPFs), Linear Regression (LR), Artificial Neural Networks (ANN). doi: 10.3966/025400022018034601004