

# 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

## 差校正

章鶴群<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>

氣象衛星中心<sup>1</sup> 氣象科技中心<sup>2</sup> 氣象資訊中心<sup>3</sup>

中央氣象局

國立中央大學 大氣科學系<sup>4</sup>

### 摘要

本研究使用類神經網路 (Artificial Neural Networks, ANN) 針對中央氣象局 WRF 系集預報系統 (WRF Ensemble Prediction System, WEPS) 的颱風降雨機率預報進行偏差修正。ANN 利用雷達降雨估計當做學習對象, 分析過去大量的降雨預報資料(訓練資料), 得到實際降雨機率與預報降雨機率之間的非線性關係, 再使用所得到的經驗係數來校正另一組降雨機率預報 (驗證資料), 使修正後的機率值接近降雨事件實際發生的機率。

本研究選取 2013 到 2015 年共 12 個颱風個案進行實驗, 採用交叉驗證 (cross validation) 方式進行校正, 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。可信度 (Reliability) 與區辨能力 (Discrimination ability) 等校驗分析顯示 ANN 方法可有效改善預報偏差。

關鍵字：類神經網路、系集預報、降雨機率

## 一、前言

降雨機率的預報是否有足夠的可信度當作防救災決策的參考是重要的研究課題, 亦是本研究最終的目標。中央氣象局發展一套以 WRF 區域模式為基礎的系集預報系統 (WRF Ensemble Prediction System, WEPS) 以改善決定性預報的不確定性, 透過不同的參數方法及隨機擾動, 產生 20 個系集成員, 並期望成員的預報結果能涵蓋大部分的預報不確定性, 以提供具有參考價值的機率預報。本研究是以 WEPS 20 個成員的降雨預報產生系集降雨機率預報。

降雨機率的計算方法需先決定降雨門檻, 再檢視 20 個成員所預報的降雨量, 計算超過所選定的門檻值所佔的比例。舉例來說, 若 24 小時降雨量超過 50 毫米的成員為 2 個, 則降雨機率為 0.1 (2/20)。WEPS 系集降雨機率預報的可信度分析顯示預報機率有偏高的情況 (吳等, 2016), 表示 WEPS 預報降雨有偏濕 (圖 1) 的特性。

類神經網路 (Artificial Neural Networks, ANN) 的基本概念是模仿人類的神經網路系統具有誤差容忍度高 (fault tolerance), 平行運算 (parallel computation) 以及強大的學習能力等特性, 地下水水流模式即應用

此種特性, 在不知道現地各種參數的情況下, 利用觀測到的輸入與輸出之成對資料訓練類神經網路, 取代大量精準參數與觀測資料 (劉等, 2009)。ANN 亦可以應用在系集預報的降雨機率修正, 利用長時間或大量的觀測數據, 透過 ANN 的學習, 能夠增加降雨機率預報的技術得分 (Skill Scores) (Yuan et al. 2008)。

颱風降雨對於台灣地區的影響十分重要, 除了是本島夏季主要降水的來源之外, 其強降雨所造成的災害, 更是造成人民生命、財產、安全及社會的經濟損失。因此本研究的重點在於使用類神經網路的學習特性, 針對颱風的降雨機率預報進行偏差修正, 以提高其預報可信度, 增加其使用價值。

## 二、ANN 校正方法

方程式(1)為一價值函數, 其中  $P^0$  為觀測的降雨機率, 利用實際降雨資料所得到的降雨機率僅有 0 和 1 兩種情況, 分別代表沒發生和有發生超過門檻值的降雨事件,  $P$  則是修正的 WEPS 降雨預報機率, 亦即使用透過 ANN 運算所得到的輸出值:

$$J = \min \sum_{t=1}^N \frac{1}{2} \{P_t^0(x, t) - P_t(x, t)\}^2 \quad \text{---(1)}$$

ANN的架構分為三部份（圖2），分別是輸入層（Input Layer）、隱藏層（Hidden Layer）和輸出層（Output Layer），本研究的設計是輸入層有7個神經元，分別輸入7個降雨門檻的機率預報值。進行校正之前，我們先利用三年所有颱風個案，統計個案的24小時累積觀測降水分布，訂定一組降水門檻（25, 30, 40, 50, 60, 70, 80, 100, 130, 160, 200, 250, 300, 350, 400, 450, 500 mm/day），目的在於使不同降水門檻間之訓練樣本數相當。要特別說明的是：校正是分別針對每個選定的降水門檻來進行（本研究為50, 80, 130, 200 mm/day 共4個降水門檻）。以校正50 mm/day的降雨機率為例，輸入給ANN進行學習的數值包含50 mm/day以及其前後3個門檻的機率值（25, 30, 40, 50, 60, 70, 80mm/day）。隱藏層有4個神經元，輸出層有1個神經元，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。令 $w_{ij}$ 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。輸出層的輸出值可以用公式(2)來表示：

$$P = Z_k(x, t) = S(a^{hidden} + \sum_{j=1}^n u_{j,k} y_j) = S(a^{hidden} + \sum_{j=1}^n u_{j,k} S(a^{input} + \sum_{i=1}^n w_{i,j} x_i)) \quad \text{---(2)}$$

其中定義了S函數（公式3）來計算各神經元所接收到的資訊的總量：

$$S(a) = \frac{1}{1+e^{-a}} \begin{cases} -\infty < a < \infty \\ 0 < s < 1 \end{cases} \quad \text{---(3)}$$

將公式(2)代入公式(1)，價值函數改寫成：

$$J(u_{j,k}, w_{i,j}) = \min \sum_{i=1}^n \frac{1}{2} \{ P_i^o(x, t) - S[\sum_{j=1}^{n+1} u_{j,k} S(\sum_{i=1}^{n+1} w_{i,j} x_i)] \}^2 \quad \text{---(4)}$$

若能夠令公式(4)的價值函數有極小值，則表示得到了最佳的降雨機率，此ANN所使用的數值解方為Linear Least Square SIMplex（LLSSIM），此方法同時使用了線性的最小平方方法（Linear Least Square）及Multi-start Simplex Search Algorithm去求得最佳的 $w_{ij}$ 和 $u_{jk}$ ，使價值函數有最小值（Hsu, 1996）。利用訓練的資料取得最佳的 $w_{ij}$ 和 $u_{jk}$ ，再將需校正的資料輸入公式(2)並使用訓練所得的 $w_{ij}$ 和 $u_{jk}$ ，可得到校正後的降雨機率。

### 三、實驗設計

本研究主要採用ANN方法以QPESUMS的雷達觀測估計之降雨量做為ANN的學習資料，來修正WEPS的預報降雨機率。所採用的個案資料包括2013年至2015年侵台的12個颱風（表1）之0-24小時累積降雨預報。

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最後的校正結果中。

## 四、校驗方法

本研究使用了reliability diagram、relative operating characteristic (ROC)和Brier skill score (BrSS)來評估ANN校正前後的WEPS降雨機率預報表現，以了解ANN對於颱風降雨機率預報的校正成效。

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，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）。ROC可用以評估降雨事件的發生或不發生是否具有區辨能力，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。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。

## 五、校驗結果

### （一）信賴程度分析

圖3為WEPS在降雨門檻 $\geq 50$ mm下校正前後的信賴度分析圖（reliability diagram），藍色點虛線是ANN修正前的情況，發現降雨機率預報較實際觀測頻率高，顯示WEPS的雨量預報具有濕偏差的特性，但是經過修正後有明顯改善，紅色實線為經過ANN修正後的結果，可以看出明顯接近圖中對角線，表示降雨機率預報接近實際觀測頻率。圖4、圖5和圖6為80、130和200mm/day的分析情況，同樣的在修正前預報的降雨機率頻率偏高，但經過ANN修正後均有改善。

### （二）Briere Skill Score分析

圖7藍色實線為WEPS降雨機率預報在ANN校正前之BrSS值，可以發現BrSS隨著降雨門檻提高而降低，其中130和200mm/day兩門檻值的BrSS<0，表示WEPS預報相對於氣候預報不具有預報能力。校正後的BrSS值明顯提高，且都大於0，表示校正後的WEPS預報相較於氣候預報比校正前的預報更具有預報技術，特別是在降雨門檻值大（例如：130和200mm/day）的時候更是明顯。

### （三）ROC分析

從ROC的分析結果可知經過修正後WEPS的降雨機率預報具有更好的區辨能力。雖然降雨門檻值低的時候，從ROC的分析沒有辦法看明顯看出區辨能力的改善，例如修正的門檻值為50mm/day時（圖8），修正前的藍色實線與修正後的紅色實線幾乎重疊在一起，但是80、130和200mm/day可以看出紅色實線在曲線下的面積稍大於藍色實線（圖9、圖10和圖11）。圖12則

是將不同門檻值的ROC曲線下面積標示在同一張圖上，以折線圖的方式表現，在80、130和200mm/day的ROC曲線下面積，修正後比修正前大，且數值大於0.85，表示區辨能力在透過ANN調整後有提升。

## 六、 結論

由於數值模式大多存在系統性偏差，在提供預報資訊給使用者前，若能先透過校正程序修正預報偏差，將可提高WEPS 預報的可信度，讓使用者做出較佳的決策，減少防災行動上的成本浪費。因此針對WEPS 的預報偏差進行校正，是必要的手段。

由於WEPS對於颱風降雨有過度預報（濕偏差）的狀況，透過ANN校正能夠改善其偏差狀況，增加其信賴程度，提升對颱風降雨事件的區辨能力及增加其預報能力。本研究僅針對颱風降雨預報進行討論，但颱風預報還有路徑、強度、綜觀環境場等不確定因素，未來可嘗試對颱風路徑、強度、綜觀環境場等不同特性進行校正，進一步評估ANN之校正成效。

## 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。天氣分析研討會。

劉宏仁、徐年盛、李天浩、林聖鈞，2009：應用 AIC 指標優選地下水水位模擬之類神經網路結構。中國土木水利工學刊，第二十一卷，第二期，183-192。

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.

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

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, 477–491.

## 附表及附圖

	Typhoon	Start - end
2013	Souluk (TY 01)	0000 UTC 11 Jul – 0000 UTC 14 Jul
	Trami (TY 02)	0000 UTC 20 Aug – 0000 UTC 22 Aug
	Kong-Rey (TY 03)	0000 UTC 27 Aug – 0000 UTC 30 Aug
	Usage (TY 04)	1800 UTC 19 Sep – 1200 UTC 22 Sep
	Fitow (TY 05)	1800 UTC 04 Oct – 0600 UTC 07 Oct
2014	Matmo (TY 06)	0000 UTC 21 Jul – 1800 UTC 23 Jul
	Fung-Wong (TY 07)	1800 UTC 18 Sep – 1200 UTC 22 Sep
2015	Linfa (TY 08)	1800 UTC 05 Jul – 0600 UTC 08 Jul
	Chan-Hom (TY 09)	1200 UTC 08 Jul – 0600 UTC 11 Jul
	Soudelor (TY 10)	1800 UTC 05 Aug – 0600 UTC 09 Aug
	Goni (TY 11)	0600 UTC 20 Aug – 1800 UTC 23 Aug
	Dujan (TY 12)	1800 UTC 26 Sep – 1800 UTC 29 Sep

表1、2013 年至2015 年颱風個案列表。

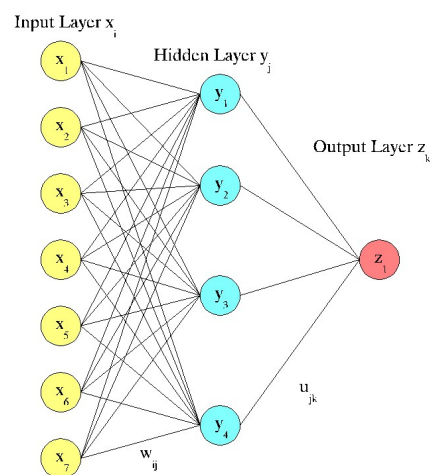


圖1、類神經網路架構圖，輸入層神經元共七個（黃色），隱藏層有四個（藍色），輸出層有一個（紅色）。

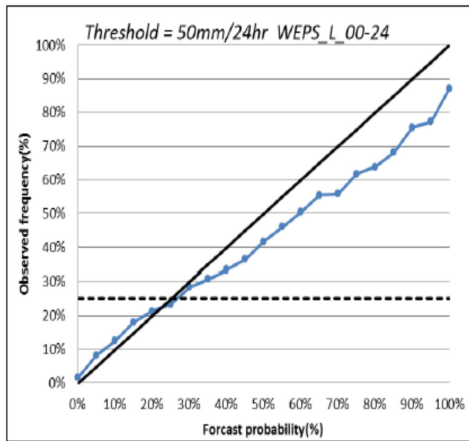


圖2、reliability diagram 示意圖，縱軸為觀測頻率，橫軸為預報機率，藍色線為觀測頻率與預報機率間對應結果；而黑色粗線為完美預報的結果；黑色水平虛線為樣本的氣候發生頻率（吳等，2016）

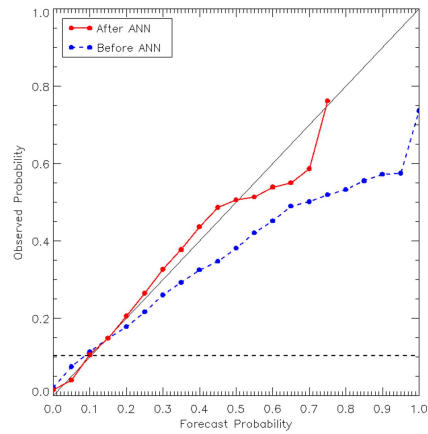


圖3、門檻值50mm/day的WEPS降雨預報，經過ANN修正前後的reliability diagram 藍虛線為修正前，紅實線為修正後，其餘說明同圖2。

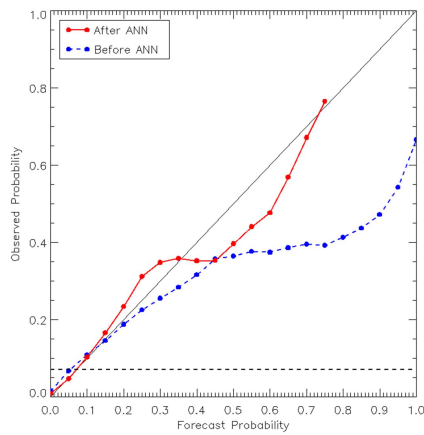


圖4、同圖3，為80mm/day的reliability diagram。

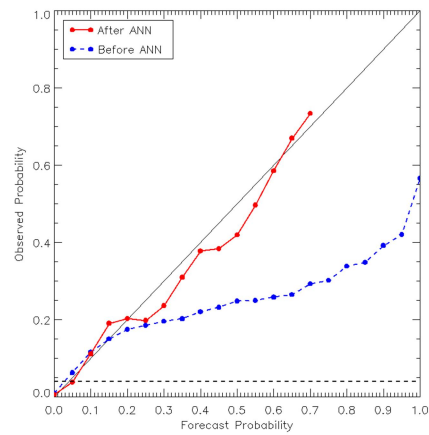


圖5、同圖3，為130mm/day的reliability diagram。

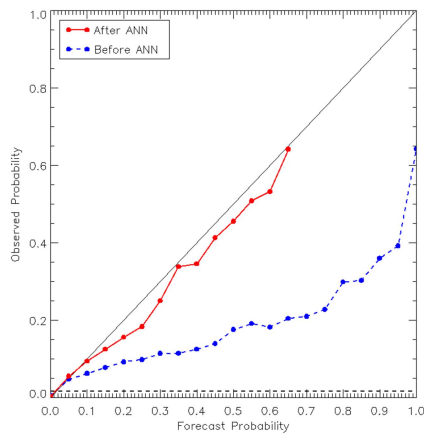


圖6、同圖3，為130mm/day的reliability diagram。

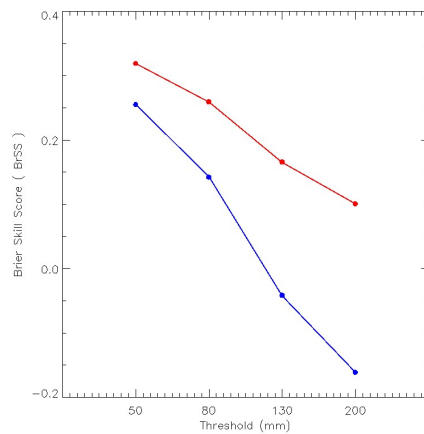


圖7、BrSS分析圖，橫軸是降雨量的門檻值，縱軸是Brier Skill Score，藍線為ANN修正前，紅線為修正後。

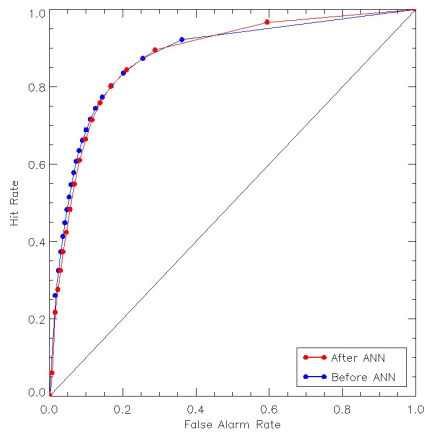


圖8，ROC曲線分析圖，縱軸是命中率 (hitrate)、橫軸是錯誤預報率 (false alarm rate)，藍線是ANN修正前，紅線是修正後。此圖的降雨門檻值為50mm/day。

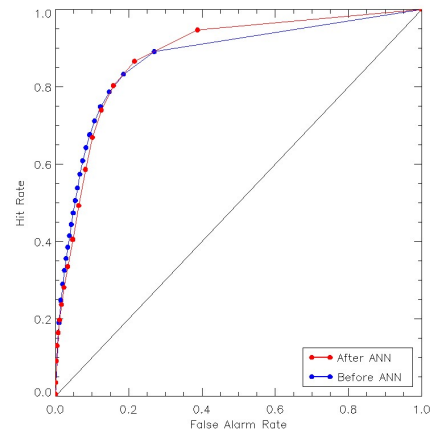


圖9，同圖8，降雨門檻值為80mm/day。

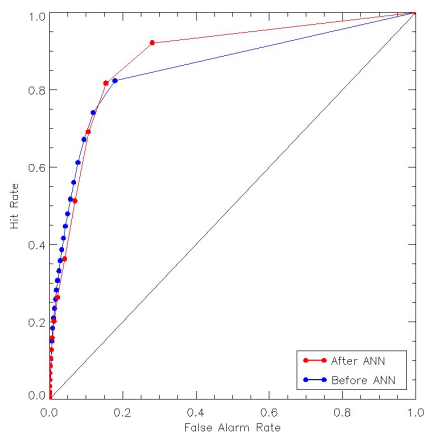


圖10，同圖8，降雨門檻值為130mm/day。

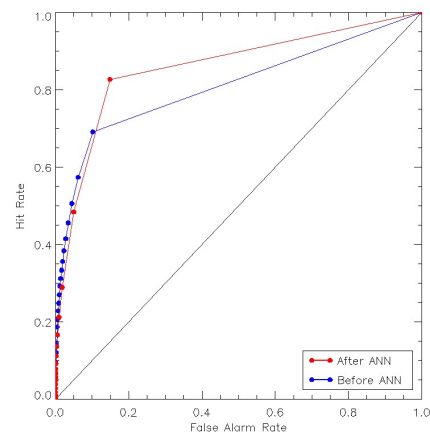


圖11，同圖8，降雨門檻值為200mm/day。

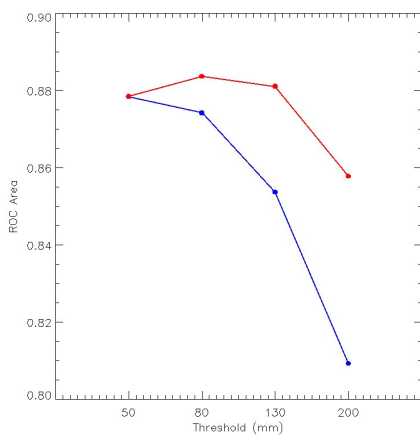


圖12，各降雨門檻值所得到ROC曲線下面積，縱軸為面積，橫軸為降雨車檻值，藍線是ANN修正前，紅線是修正後。