

# 應用類神經網路進行台灣地區颱風系集降雨機率預報校正

莊美誼<sup>1,2</sup> 王彥雯<sup>2</sup> 黃煜鈞<sup>2</sup> 李美賢<sup>3</sup> 張惠玲<sup>1</sup> 洪景山<sup>1</sup> 蕭朱杏<sup>2\*</sup>

<sup>1</sup>交通部中央氣象局

<sup>2</sup>國立臺灣大學流行病學與預防醫學研究所

<sup>3</sup>臺北市立大學數學系

(中華民國 111 年 3 月 1 日收稿；中華民國 111 年 8 月 24 日定稿)

## 摘要

颱風定量降雨預報是氣象防災的重中之重，建立在系集預報所提供的大數據基礎上，利用類神經網路技術建立颱風定量降雨預報校正模型，將是一個值得期待的趨勢。然而，由於雨量特殊的統計特性以及颱風個案稀少，因而使得應用類神經網路技術建立預報校正模型時將面臨相當的挑戰。本研究旨在探討類神經網路方法應用在台灣地區颱風雨量預報校正的可行性，以及在不同颱風雨量下的預報校正成效；特別是在雨量極大時，類神經網路校正模型是否能改善原始系集降雨機率預報。

以 2015 年蘇迪勒颱風為例，本研究建立 3 個淺層類神經網路與 1 個深層類神經網路，進行颱風未來 24 小時累積降雨預報校正，評比結果顯示淺層類神經網路有較佳的校正成效。本研究同時考量颱風快速變化的特性，使用非均等權重方法來建構模型，結果顯示非均等權重方法比均等權重方法所建立之淺層類神經網路模型有較佳的校正成效。本研究也將結果應用到 2013-2015 年間侵台的其他 11 個颱風，同樣發現非均等權重方法的淺層類神經網路能校正颱風原始系集降雨機率預報。

**關鍵字：**校正、系集預報、類神經網路、定量降雨機率預報、颱風雨量

---

\*通訊作者地址：蕭朱杏教授，10055 台北市徐州路 17 號 5 樓 523 室，

Email: ckhsiao@ntu.edu.tw，Tel: +886-2-3366-8032

## 一、前言

數值天氣預報 (Numerical Weather Prediction, NWP) 可以分為兩種形式，一是決定性預報，以單一初始場進行預報；二是系集預報 (Ensemble prediction)，考慮了模式初始、邊界以及模式積分過程擾動，產製模式預報的不確定性資訊 (Li et al. 2020)。建立在系集預報提供之大數據的基礎上，使用統計之後處理技術所開發的加值產品，往往可以更進一步提升系集預報的應用層面，類神經網路技術進行颱風定量降雨預報校正是一個值得期待的趨勢。本研究利用系集預報輸出，運用現有的類神經網路方法，進行颱風雨量預報校正，並加以優化類神經網路方法使其更適用於颱風雨量的資料型態，同時評比這些方法的校正成效及其預報效能。

常見的統計處理方法大致上可以分為以下四類：1. 迴歸模型：例如羅吉斯迴歸 (Logistic Regression, Hamill et al. 2004)、非同質高斯迴歸 (Non-homogeneous Gaussian Regression, Gneiting et al. 2005)、分位數迴歸 (Quantile Regression, Massidda and Marrocu 2018)；2. EMOS 統計模型 (Ensemble Model Output Statistics, Thrarinsdottir and Gneiting 2010)；3. BMA 貝氏統計模型 (Bayesian Model

Averaging, Raftery et al. 2005) 以及 4. 機器學習預測模型，例如類神經網路模型 (Neural Network, Lauret et al. 2014, Rasp and Lerch 2018, Bremnes 2020, Grönquist et al. 2020)。其中，過去許多研究嘗試使用機器學習中類神經網路模型並運用系集預報資料進行加值產品之研發，這也是本研究所關注的重點。提到類神經網路，由於缺乏預測結果的可解釋性，人們經常以「黑盒子」來形容；然而，相較於結果的解釋，預測的表現往往是類神經網路方法的主要目標。在系集預報的後續加值產品處理中，因為預測的表現也是主要目標之一，所以本研究利用類神經網路模型，並進一步探討這樣的模型是否適用於颱風雨量的預測。

過去曾有一些研究使用類神經網路於系集預報後續加值產品處理，如：Rasp and Lerch (2018) 針對氣候溫度預測提出的類神經網路處理 (Neural Networks for Postprocessing Ensemble Weather Forecasts)，其概念是將欲預測之氣候因子的系集預報成員的統計量，即平均值與標準差，視為預測因子，同時採用多個會影響主要觀測氣候之影響因子的系集預報成員統計量做為共變數，再使用類神經網路中嵌入的技術納入地點資訊，來建立類神經網路模型；他們在常態分布的假設下，預測出經過校正後的

氣候因子之機率分佈參數值，也就是常態分佈之期望值與標準差。Bremnes (2020) 的研究為結合迴歸統計模型中的分位數迴歸方法(quantile regression)與機器預測模型中的類神經網路方法，再利用伯斯坦多項式(Bernstein Polynomials)及類神經網路概念作為基礎之模型，此模型以系集預報成員作為預測因子，經由類神經網路預測伯斯坦係數，建立出一個分位數迴歸預測模型(Ensemble Postprocessing Using Quantile Function Regression Based on Neural Networks and Bernstein Polynomials)。該模型只需建立單一模型，就能同時預測不同分位數的機率預報分佈，優於過去方法需分別建立不同分位數的模型與預測。由於分位數迴歸的特性，無須考慮預設氣候因子所服從的分佈，並且可以特別關注使用者有興趣的分位數分佈，有利於預測極端氣候。同時 Grönquist et al. (2020) 也在研究中提出深度學習模型(Deep Learning for Post-Processing Ensemble Weather Forecasts)，運用多層 CNN 卷積神經網路來進行預測，屬於深度的類神經網路，其預測結果為預測機率預報分佈的參數值，該研究應用其技術預測全地球表面溫度的平均數與標準差。

以上研究多以常態分佈之期望值與標準差作為代表。若被預測的變數分佈本來就近似於常態分佈，例如溫度，則

這樣的設定是合理的。然而，若變數的分佈不是近似於常態分佈，例如本研究所關心的雨量，通常會有兩種處理方式：一是將模型的輸出值轉換成已知分佈的參數值，例如當預測因子分佈服從伽瑪分佈時，可以針對伽瑪分佈的參數進行預測；第二種方式是不假設雨量為常態分佈並將預測結果作為定值預報，模型輸出的兩個節點可視為定值的預測，以及其定值預測之不確定性。在本研究中，將先假設雨量為常態分配並進行分析，同時在文末討論所使用的資料是否符合常態分佈，此外，也使用上述兩種方法，探討採用非常態之假設對結果的影響。

本研究使用 Rasp and Lerch (2018) 的概念，並修改其類神經網路處理的部分過程，使模型能作為颱風雨量預報的校正方法。颱風未來 24 小時累積雨量為本研究的目標，相較於其他研究皆使用多種影響因子的預測資料，本研究單純以系集颱風雨量預報資料進行分析，沒有納入其他影響因素的資訊（如風速、溫度或壓力等）；所以相對而言，本研究嘗試藉由使用類神經網路模型技術，提升系集預報對颱風定量降水預報的能力，同時比較不同形式的類神經網路技術對颱風雨量預報的校正成效，以及探討本研究所使用的非均等權重資料方法是否有助於提升校正成效。此外，由於

颱風的型態多樣，不同的路徑、不同的雨量強度、不同的速度或是暴風圈的遠近，皆可能形成不同的情境，並且各個颱風個案可提供之預報次數也相對稀少，因此，本研究也探討，類神經網路是否適合使用少量的樣本數來建立模型、進行預測；是否應將不同時序的資料視為相等重要的訓練資料；並進一步探討若以颱風雨量系集預報成員的平均值與標準差作為預測因子，以常態分佈的參數值作為預測結果是否合理。研究方法將會在第二節呈現，第三節則以侵襲台灣的蘇迪勒颱風進行案例分析，包含颱風個案資料介紹，以及颱風雨量預報校正表現。最後，第四節則是針對研究中的相關議題進行討論。

## 二、研究目的與方法

本研究目的為建立類神經網路模型，對台灣颱風未來 24 小時累積雨量進行校正預報，特別是利用颱風初期幾次系集成員的雨量預報、及相對的實際觀測雨量，進行類神經網路模型之訓練，再針對該颱風接下來的系集成員的雨量預報來得到校正後的預報。本研究考慮四種類神經網路模型，本節中（一）至（三）的模型屬於淺層類神經網路，而（四）的模型屬於深層類神經網路模型。另外，模型（五）則會使用非均等權重資料方法於模型（一）至（四）中

最佳的模式，以不同形式的訓練集資料進行預測，再藉由不同的指標，檢查前述類神經網路預測颱風雨量的表現，尤其關注於強降雨時，模型的預測能力能否提升，並探討在類神經網路處理中，隱藏層、節點與層數對於颱風雨量預測表現的影響，以及比較經調整之訓練集資料（非均等權重資料方法）進行建模與未調整之訓練集資料進行建模的差異。

本研究中類神經網路模型的輸入層及輸出層與 Rasp and Lerch (2018) 的設計概念相同，亦即，使用系集成員的雨量預報樣本平均值 ( $M_{s,t}$ ) 與樣本標準差 ( $S_{s,t}$ ) 作為預測因子，將數據做為模型輸入層 (Input layer) 的節點 (Node) 資料；而預測結果則為預報的常態機率分佈之期望值 ( $\mu_{s,t}$ ) 與標準差 ( $\sigma_{s,t}$ )，也就是模型中輸出層 (Output layer) 的節點資訊。每個模型皆包含輸入層及輸出層，部分模型包含隱藏層 (Hidden layer)。每個模型總共有  $I$  層，以  $i = 1, \dots, I$  表示第幾層。每一層層內節點總數為  $N_i$ ，例如  $N_1$  代表第一層輸入層層內的總節點數。第  $i$  層層內的節點資料組成的向量表示為  $X^{(i)}$ ， $i = 1, \dots, I$ ，而  $X^{(i)}$  的維度為  $N_i$ ，

$$X^{(i)} = [x_1^{(i)}, \dots, x_j^{(i)}, \dots, x_{N_i}^{(i)}]$$

各層中的節點 ( $j$ ) 數值均為上一層各節點加權 ( $W^{(i)}$ ) 之總和加上誤差值 ( $b^{(i)}$ )

後，再經過激勵函數(Activation function, 以  $A$  表示)運算所得之結果，例如：

$$x_j^{(i+1)} = A \left( \left[ x_1^{(i)}, \dots, x_k^{(i)}, \dots, x_{N_i}^{(i)} \right] \times \left[ w_{j,1}^{(i)} \dots w_{j,N_i}^{(i)} \right]^T + \left[ b_{j,1}^{(i)} \dots b_{j,N_i}^{(i)} \right]^T \right)$$

也可簡單以向量表示為

$$X^{(i+1)} = A(X^{(i)}W^{(i)} + b^{(i)})$$

其中，

$$W^{(i)} = \begin{bmatrix} w_{1,1}^{(i)} & \dots & w_{N_i,1}^{(i)} \\ \vdots & \ddots & \vdots \\ w_{1,N_i}^{(i)} & \dots & w_{N_i,N_i}^{(i)} \end{bmatrix}_{N_i \times N_i},$$

$$b^{(i)} = \left[ b_1^{(i)}, \dots, b_{N_i}^{(i)} \right]_{N_i \times 1}。$$

以上模型將逐一於每個網格點，針對第四報開始的未來 24 小時累積降雨預報、結合訓練集資料，逐報分別進行建模與預測，訓練集資料則以滾動方式進行更新（流程如圖 1(e)，此流程之設計將於本文第四節進一步說明）。而模型中使用的激勵函數為 Softplus 函數，如以下式子(1)，(Zheng et al. (2015), Nwankpa et al. (2018))，並使用梯度下降法 Gradient decent (Chong and Zak (2013)) 來幫助模型獲得最佳的參數解。損失函數則使用氣象中經常作為機率預報校驗的連續分級概率評分(Continuous Ranked Probability Score)，簡稱為 CRPS (式子(2))，此評分方式不僅考慮預報誤差同時也考慮預報的不確定性，這也是其他相關研究經常使用的指標，如 Hersbach

(2000)與 Ferro (2013)。式子(2)中的 $F_Y(t)$ 為預測之累積機率， $y$ 為真實觀測值，CRPS 數值小表示預測機率分布與真實值接近。

$$\text{Softplus}(u) = \ln(1 + \exp(u)) \quad (1)$$

$$\text{CRPS} = \int_{-\infty}^{\infty} [(F_Y(t) - 1(y \leq t))]^2 dt \quad (2)$$

### (一) 淺層類神經網路 Fully Connected Neural Network (FCN)模型

第一個模型稱為 Fully Connected Neural Network (後續簡稱為 FCN)，屬淺層類神經網路，是全部模型當中最簡單的模型，僅包含輸入層與輸出層，中間沒有隱藏層，因此  $I = 2$ ，結構為  $N_{\{1,2\}} = \{2,2\}$ ，如圖 1(a)。例如在颱風雨量的應用時，第一層輸入層  $N_1 = 2$ ，包含了系集資料的平均值(mean)與標準差(sd)。

### (二) 淺層類神經網路 Neural Network with 2 Nodes (NN2N)模型

第二個模型則稱為 Neural Network with Two Nodes (後續簡稱為 NN2N)，同樣屬於淺層類神經網路，除了輸入層與輸出層外，中間包含一層兩個節點的隱藏層，因此  $I = 3$ ，結構為  $N_{\{1,2,3\}} = \{2,2,2\}$ ，如圖 1(b)。這個模型可與第一個 FCN 模型比較，以區辨隱藏層是否有助於提升預測的表現。

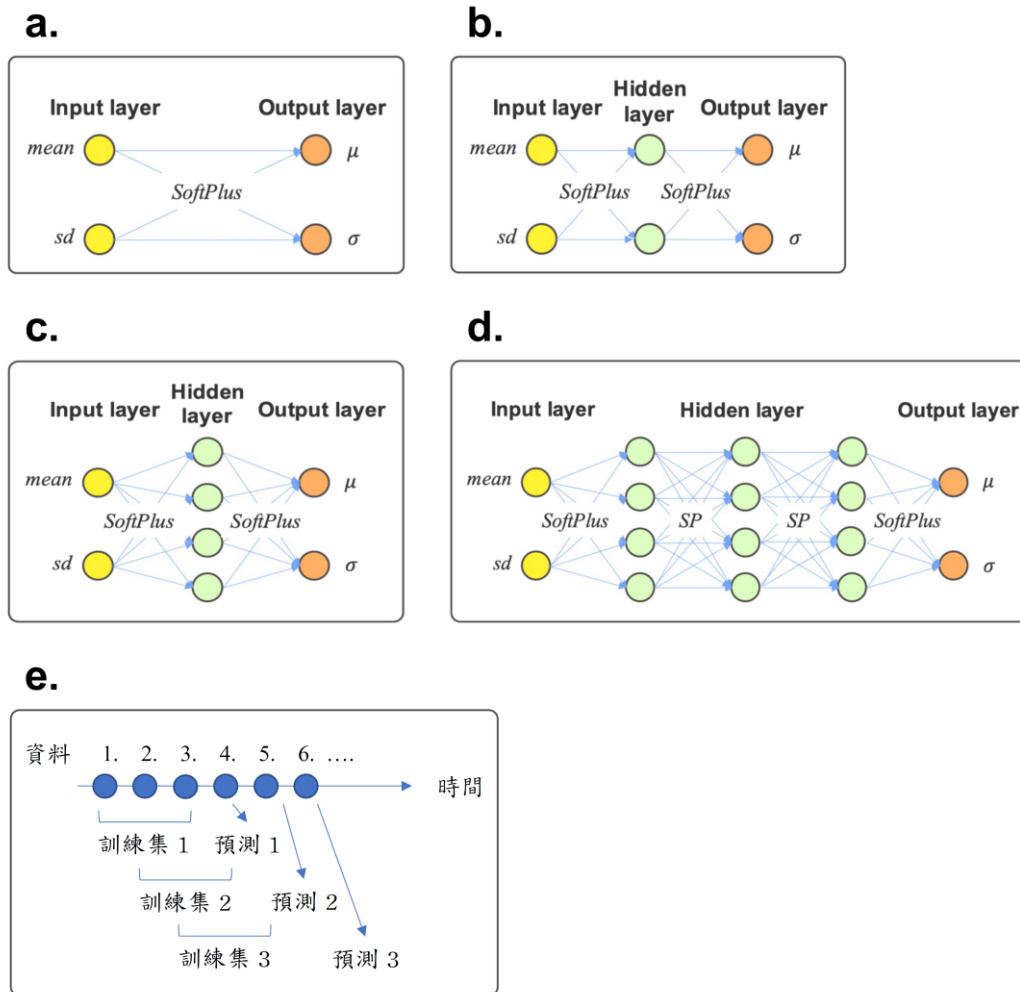


圖 1：(a) FCN 類神經網路模型結構圖。最左邊的 Input layer 表示模型中輸入系集成員預報訓練集中每一報的平均值(mean)與標準差(sd)，訓練集有幾報就輸入幾對。最右邊的 Output layer 代表的是要被估計出來的累積雨量的常態分佈的母體期望值( $\mu$ )與標準差( $\sigma$ )。(b) NN2N 類神經網路模型結構圖。最左邊的 Input layer 與最右邊的 Output layer 的說明與圖(a)相同，中間則增加一個有兩個節點的隱藏層。(c) NN4N 類神經網路模型結構圖。最左邊的 Input layer 與最右邊的 Output layer 的說明與圖(a)相同，中間則增加一個有四個節點的隱藏層。(d) NN3H4N 類神經網路模型結構圖。最左邊的 Input layer 與最右邊的 Output layer 的說明與圖(a)相同，中間則增加三個有四個節點的隱藏層。(e) 滾動式挑選訓練集資料，例如第 1-3 報為訓練集時，第 4 報為模型預測標的；第 2-4 報為訓練集時，第 5 報為模型預測標的；以此類推。

### (三) 淺層類神經網路 Neural Network with 4 Nodes (NN4N)模型

第三個模型稱為 Neural Network with Four Nodes (後續簡稱為 NN4N)，

為本研究中淺層類神經網路的最後一種，與 NN2N 結構上大致相同，差別在於中間的隱藏層共包含四個節點，因此  $I = 3$ ，結構為  $N_{\{1,2,3\}} = \{2,4,2\}$ 。這個模型可與第二個模型 NN2N 進行比較，

區辨隱藏層節點數的差異是否會影響預測的表現，如圖 1(c)。

#### (四) 深度類神經網路 Neural Network with 3 Hidden Layers and 4 Nodes for each Hidden Layer (NN3H4N)模型

最後一個模型稱為 Neural Network with Three Hidden Layers and Four node for each Hidden Layers (後續簡稱為 NN3H4N)，屬於深度類神經網路，也是本研究中較為複雜的模型，除了以上每個模型皆有的輸入層與輸出層外，中間包含了三層隱藏層，因此  $I = 5$ 。而且，每個隱藏層皆包含了四個節點，結構為  $N_{\{1,2,3,4,5\}} = \{2,4,4,4,2\}$ ，如圖 1(d)。這個模型可與其他三個模型比較，觀察使用複雜的深層類神經網路模型，是否與使用淺層類神經網路模型的表現有所差異。

#### (五) 非均等權重的訓練集 (Unequal Weight Training Data)

由於考量到颱風雨量變化速度很快，加上每次系集成員資料間隔時間通常為 6 小時，若合併太久遠的資料建立模型，則較久遠的資料將會干擾模型的預測表現，因此除了使用圖 1(e)的方式滾動式挑選訓練集資料外，本研究也嘗試以非均等權重的訓練集資料建立模型：給予較新的訓練集資料較大的權重，較舊的資料則給予較小的權重，亦

即給予新舊資料不均等的重要性。由於模型使用的損失函數是以積分方式計算 CRPS，因此我們可以令在迭代運算中所計算的損失函數，即所有訓練集 (共  $N$  筆) 的 CRPS 數值加總為 Total CRPS (簡稱 TC)，式子(3)的 TC 為未經調整的訓練集資料；而式子(4)的 TC 則是進行權重 ( $p^1 \dots p^N$ ) 調整的訓練集資料。

$$TC = CRPS^1 + CRPS^2 + \dots + CRPS^N \quad (3)$$

$$TC = p^1 \times CRPS^1 + p^2 \times CRPS^2 + \dots + p^N \times CRPS^N, \quad (4)$$

$$p^1 + p^2 + \dots + p^N = N$$

在第三節的應用中，將選擇上述 (一) 至 (四) 四種模型中表現最好的模型，也就是淺層類神經網路，加入非均等權重的方法進行雨量預測，將測試以下兩種權重比 (從較久遠的資料至最新的資料)：1:1:2 (簡稱 UE112) 與 1:2:4 (簡稱 UE124)，若預測表現相較於未調整之均等權重佳，可以推測對於颱風雨量預測來說，較為近期的資料應較為重要。針對以上模型分析所使用的 R 程式碼皆置於 GitHub ([https://github.com/joycejuang/Typhoon-Precipitation\\_JJuang.git](https://github.com/joycejuang/Typhoon-Precipitation_JJuang.git))。

### 三、研究資料與分析

中央氣象局自 2011 年起建置以 WRF 區域模式為基礎之系集預報系統

(WRF Ensemble Prediction System, WEPS, Li et al. 2020; Chang et al. 2018), 本研究使用之 WEPS 有 20 個系集成員, 擁有三層巢狀網格, 水平解析度分別為 45、15 及 5 公里, 模式範圍如同 Chen and Hong (2017) 文內之圖 3 所示, 本研究採用解析度最高的第三層網格降雨預報資料進行 0-24 小時系集颱風累積降雨機率預報校正。

分析使用的資料則採用中央氣象局劇烈天氣監測系統(Quantitative Precipitation Estimation and Segregation Using Multiple Sensors, QPESUMS, Chang et al. 2021)的雷達估計降雨(Quantitative Precipitation Estimation, QPE)做為校正與校驗的觀測資料, 水平解析度是  $0.0125^\circ \times 0.0125^\circ$ , 範圍範圍自經度  $118^\circ$  至  $123.5^\circ$ , 緯度自  $20^\circ$  至  $27^\circ$ 。由於 QPESUMS QPE 的範圍小於 WEPS 第三層網格範圍, 且 WEPS 的水平解析度是 QPESUMS QPE 的 4 倍, 因此本研究是以 QPESUMS QPE 範圍內的模式格點做為共同格點, 進行預報校正與校驗。其中, 海洋區域只在雷達偵測範圍內有降雨估計值, 因此每一報能進行校正與校驗的共有 12,170 個格點, 其中屬於台灣陸地的只有 1,341 個格點。至於每一預報的更新頻率為 6 小時, 這是自 2013 年天兔(Usage)颱風之後, 由原本 12 小時的更新頻率改變而來。

觀察 2013-2015 年間侵台的 12 個颱風中(表 1), 蘇利颱風、康妮颱風、天兔颱風、蘇迪勒颱風(以下表示為 TY10)以及杜鵑颱風, 在部分地區出現較極端的強降雨, 而康妮颱風、麥德姆颱風以及蘇迪勒颱風的觀測降雨量相對也較高。由於颱風強降雨經常造成嚴重災損, 因此本研究聚焦在較大降雨預報之校正成效, 選擇雨量較大的蘇迪勒颱風(TY10)進行類神經網路應用的案例分析, 比較淺層與深層的類神經網路架構以及非均等權重資料集方法的校正成效。

蘇迪勒颱風(Soudelor)於 2015 年 8 月侵襲台灣, 本研究使用的降雨預報共有 15 報, 由於需要共 3 報資料作為訓練資料集, 例如採用每一個準備被預報校正的前 3 報(第 1 報時間為 1200 UTC 06 Aug 2015、第 2 報時間為 1800 UTC 06 Aug 2015、第 3 報時間為 0000 UTC 07 Aug 2015)作為訓練資料, 再於當報(即此時期的第 4 報, 時間為 0600 UTC 07 Aug 2015)進行預報校正; 以此類推, 共可校正 12 報(第 4 報至第 15 報)。圖 2(a)為蘇迪勒颱風過境台灣陸地的路徑及暴風圈半徑範圍, 圖 2(b)為進行預報校正之 12 報的觀測降雨分佈, 其中第 2 至第 9 報(TY10\_02~TY10\_09)有較強的降雨量, 因此在校正成效評估部分會特別針對這八個雨量較大的預報進行討論。



表 1：本研究所用 2013 年至 2015 年侵襲台灣之颱風個案資料集，列表資訊包含颱風編號、颱風名稱、颱風過境時間、資料集總報數以及颱風近臺強度。

年份	颱風編號(名稱)	颱風過境時間(起-迄)	資料集 總報數	近臺 強度
2013	TY01 (Soulik 蘇力)	0000 UTC 11 Jul – 0000 UTC 14 Jul	7	強烈
	TY02 (Trami 潭美)	0000 UTC 20 Aug – 0000 UTC 22 Aug	4	輕度
	TY03 (Kong-Rey 康妮)	0000 UTC 27 Aug – 0000 UTC 30 Aug	7	輕度
	TY04 (Usage 天兔)	1800 UTC 19 Sep – 1200 UTC 22 Sep	12	強烈
	TY05 (Fitow 菲特)	1800 UTC 04 Oct – 0600 UTC 07 Oct	11	中度
2014	TY06 (Matmo 麥德姆)	0000 UTC 21 Jul – 1800 UTC 23 Jul	12	中度
	TY07 (Fung-Wong 鳳凰)	1800 UTC 18 Sep – 1200 UTC 22 Sep	15	輕度
2015	TY08 (Linfa 蓮花)	1800 UTC 05 Jul – 0600 UTC 08 Jul	11	輕度
	TY09 (Chan-Hom 昌鴻)	1200 UTC 08 Jul – 0600 UTC 11 Jul	12	中度
	TY10 (Soudelor 蘇迪勒)	1800 UTC 05 Aug – 0600 UTC 09 Aug	15	中度
	TY11 (Goni 天鵝)	0600 UTC 20 Aug – 1800 UTC 23 Aug	15	強烈
	TY12 (Dujan 杜鵑)	1800 UTC 26 Sep – 1800 UTC 29 Sep	13	強烈

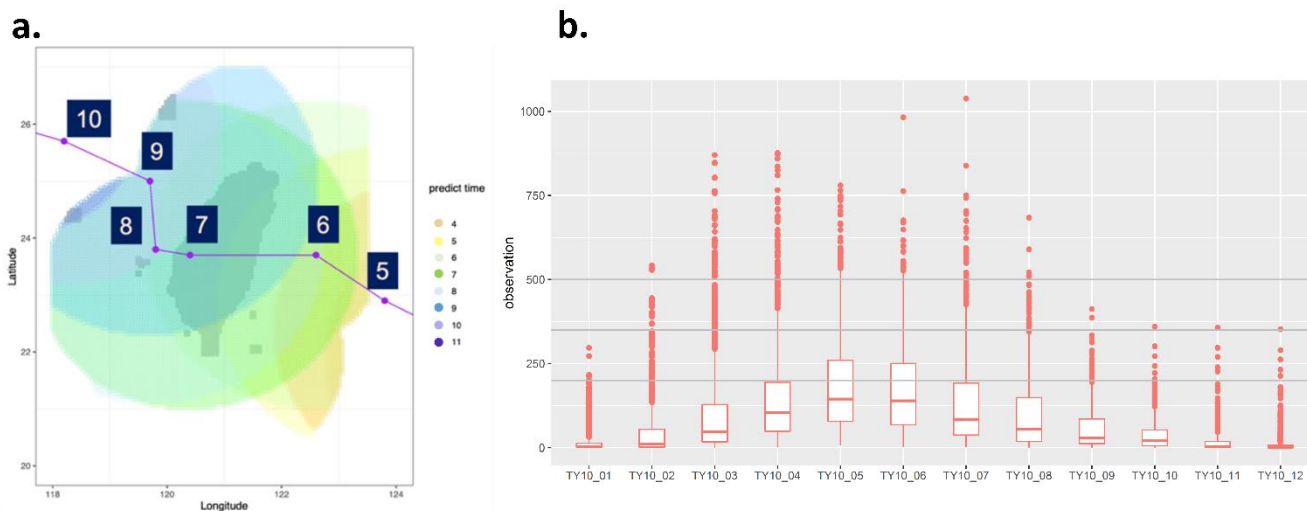


圖 2：(a) 蘇迪勒颱風過境路徑與半徑圖（標號為系集預測的報數，圖中包含第 5 報到第 10 報）。(b)蘇迪勒颱風(TY10)預測 12 報真實觀測值（標號為預測報數，從第 1 報 TY10\_01 到第 12 報 TY10\_12）。

預報校正結果評估部分，使用 CRPSS 數值盒鬚圖將各個模型的預報校正結果與原始系集預報(raw ensemble

members)互相比較，若 CRPSS（式子 (5)，數值範圍為  $(-\infty, 1]$ ）數值越高，則顯示校正成效越好。本研究使用的參

考模型是直接使用系集預報成員的平均值，再探索經類神經網路校正後的 12 報是否有較高的技術得分 CRPSS。

$$CRPSS = \frac{CRPS_r - CRPS}{CRPS_r},$$

$CRPS_r$ : reference model (5)

本研究將以最佳模型(機率預報誤差最小)占比圖呈現不同類神經網路模型的評比結果，亦即針對陸地格點統計最佳模型(CRPS 最小)的比例，繪製成不同觀

測降雨區間的最佳模型占比圖(圖 3)。至於非均等權重方法(第二節的模型五)將套用在 FCN 模型，因為此模型相較於其它模型有相對較佳的預報校正結果；採用的權重比包括(從較久遠的資料至最新的資料) 1:1:2 (簡稱 FCNUE112) 與 1:2:4 (簡稱 FCNUE124) 兩種。

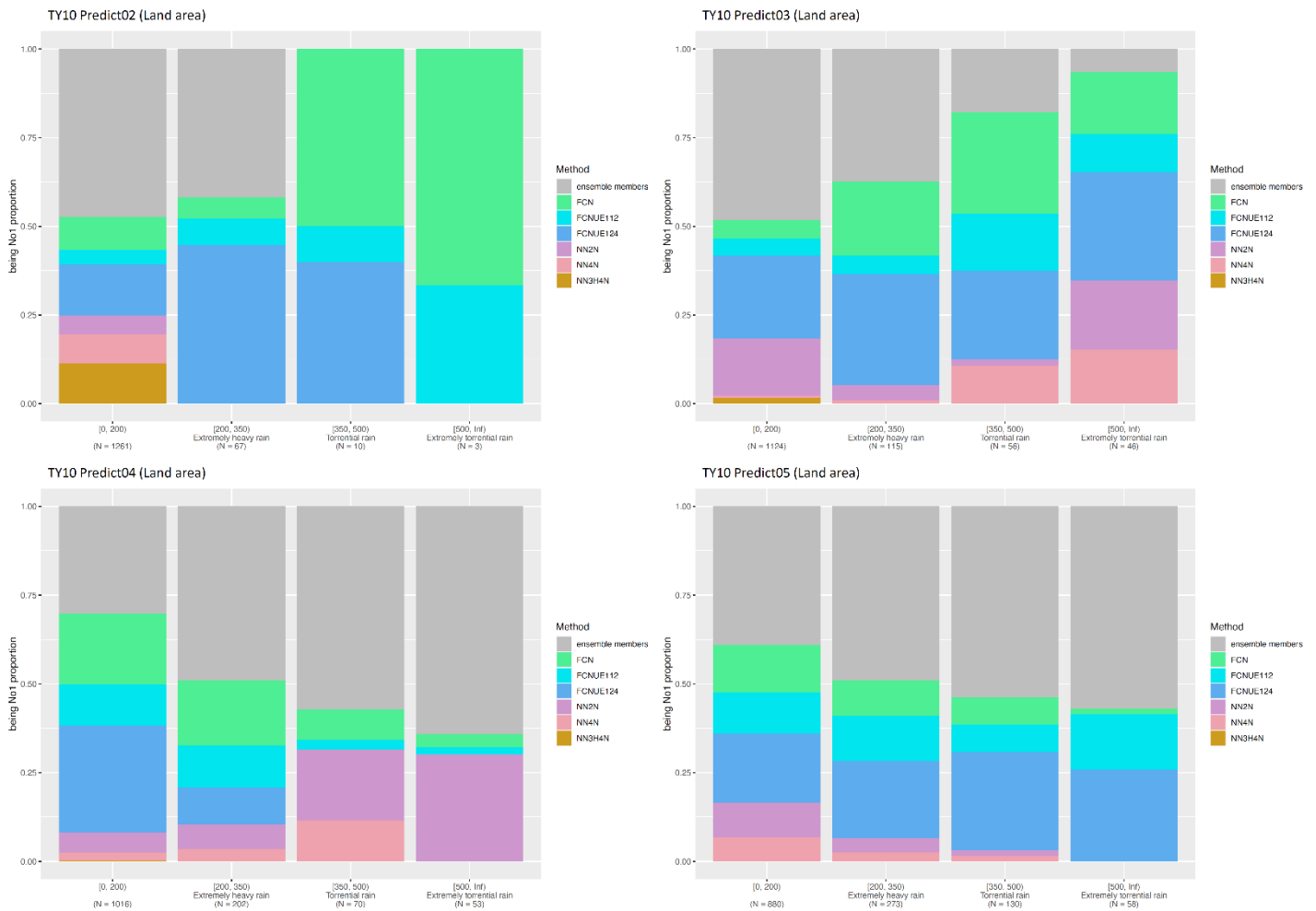


圖 3：只針對陸地(Land area)格點之後處理最佳模型占比圖，分別為從第 2 報(predict02)到第 9 報(predict09)，時間為第 2 報(1800 UTC 06 Aug 2015)至第 9 報(時間為 1200 UTC 08 Aug 2015)。X 軸為雨量真實觀測值範圍的四個類別，分別為[0,200), [200,350), [350,500), [500,Inf)，單位為毫米(mm)，N 代表該類別下的總格點數。Y 軸為各模型屬於最佳模型的格點數佔該類別的比例(以 CRPS 進行比較)。顏色則代表不同之模型。

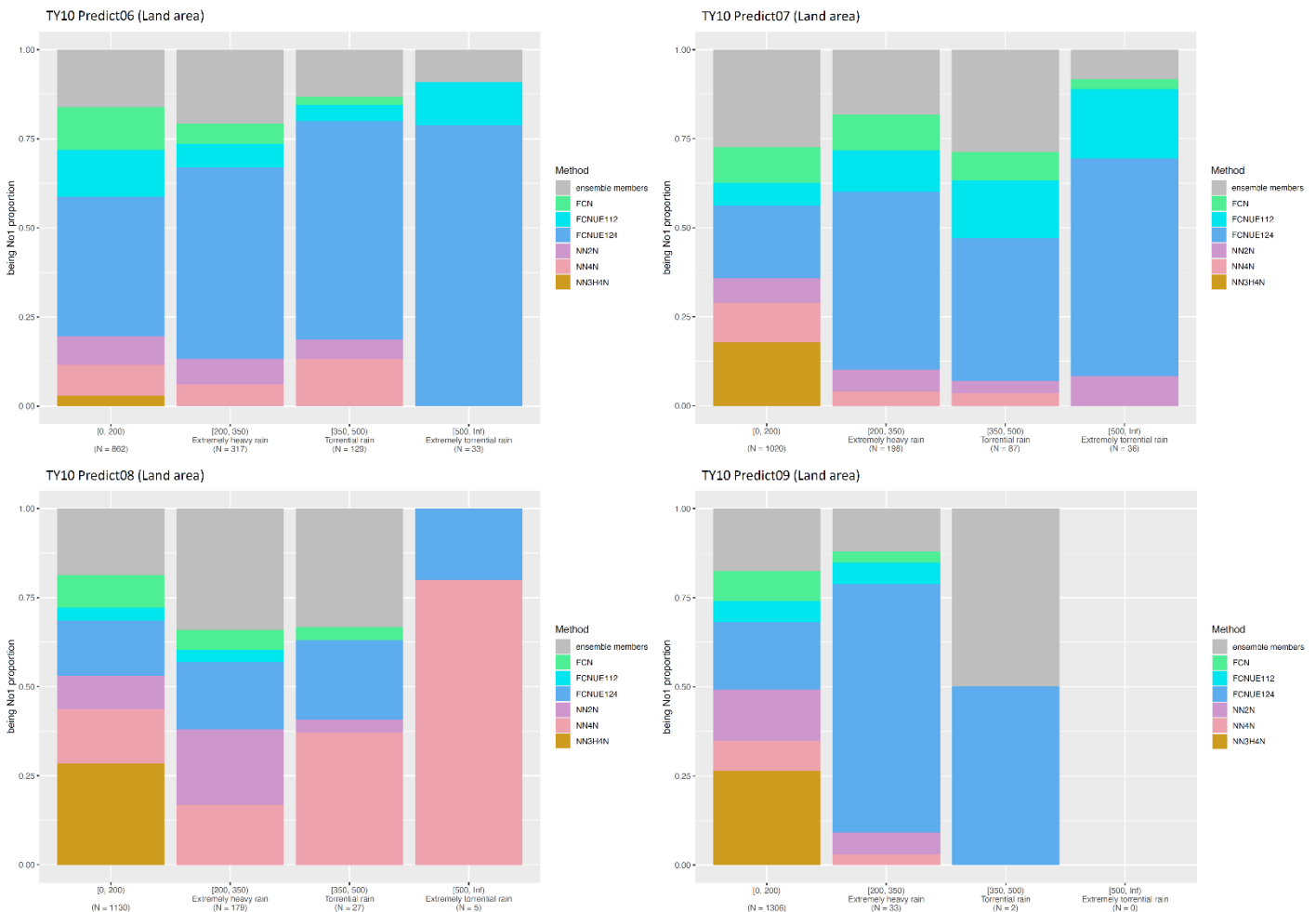


圖 3 續：只針對陸地(Land area)格點之後處理最佳模型占比圖，分別為從第 2 報(predict02)到第 9 報(predict09)，時間為第 2 報(1800 UTC 06 Aug 2015)至第 9 報(時間為 1200 UTC 08 Aug 2015)。X 軸為雨量真實觀測值範圍的四個類別，分別為[0,200), [200,350), [350,500), [500,Inf)，單位為毫米(mm)，N 代表該類別下的總格點數。Y 軸為各模型屬於最佳模型的格點數佔該類別的比例（以 CRPS 進行比較）。顏色則代表不同之模型。

表 2 的 CRPSS 結果顯示: 經由非均等權重方法處理後，校正後的預報技術得分相對提升，圖 3 中並可明顯看出 FCNUE124 相較於其他模型有較高的最佳模型占比。由 CRPSS 的分析數據（仍以原始系集機率預報作為對照模型）可看出，不管真實觀測值是介於 200 至 350 毫米、介於 350 至 500 毫米、或是高於 500 毫米，FCN、FCNUE112 及

FCNUE124 皆可提升 30%左右的預測校正表現。然而，NN2N、NN4N 及 NN3H4N 的預測校正表現則沒那麼高或甚至可能降低。

綜合以上蘇迪勒(TY10)的預測結果，獲得以下四點觀察：(1) FCN 不包含隱藏層、與 NN2N 及 NN4N 包含一層隱藏層，三者皆屬於淺層類神經網路模型，在此個案中，不包含隱藏層的模型

表 2：不同類神經網路模型於不同強降雨的情況(真實觀測值)下所對應的 CRPSS 數值之第一分位(Q1)、中位數(黑體粗字)以及第三分位數(Q3)，並皆以原始系集成員預報作為參考模型。

真實觀測值	統計量	FCN	CNUE112	FCNUE124	NN2N	NN4N	NN3H4N
200-350mm	Q3	0.6319	0.6736	0.6950	0.2713	0.3861	-0.6238
	<b>中位數</b>	<b>0.2853</b>	<b>0.3194</b>	<b>0.3260</b>	<b>-0.2873</b>	<b>-0.1253</b>	<b>-1.7832</b>
	Q1	-0.2540	-0.2943	-0.3514	-1.1948	-1.0532	-6.2003
350-500mm	Q3	0.6173	0.6543	0.6732	0.4887	0.5382	-0.9310
	<b>中位數</b>	<b>0.3030</b>	<b>0.3281</b>	<b>0.3276</b>	<b>0.0869</b>	<b>0.1808</b>	<b>-2.3209</b>
	Q1	-0.3470	-0.2017	-0.2118	-0.7701	-0.6233	-6.6703
≥500mm	Q3	0.6799	0.7189	0.7466	0.5947	0.6205	-1.7600
	<b>中位數</b>	<b>0.2314</b>	<b>0.3385</b>	<b>0.3693</b>	<b>0.0848</b>	<b>0.1436</b>	<b>-4.1029</b>
	Q1	-0.8497	-0.7159	-0.6469	-0.8944	-0.8021	-8.8996
p-value (Kruskal-Wallis):		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

優於包含單層隱藏層的模型。(2) NN2N 及 NN4N 兩者的差別為節點數的不同，在強降雨的情況下 NN4N 稍微優於 NN2N，不過整體來說預測表現差異並不大。(3) 在此個案中，淺層類神經網路模型(FCN、NN2N 與 NN4N)明顯比深層類神經網路模型(NN3H4N)的表現好。

(4) 未經權重調整的 FCN 與經過調整的 FCNUE112 與 FCNUE124 比較時，發現經過調整的預測結果是較佳的，甚至給予較近期資料較大的權重時，會有最佳的預測表現，即模型 FCNUE124 優於

FCNUE112。對於颱風資料的特色，如過境台灣時間短導致模式預報報數不多，以及氣象狀態變化極快，使得納入太過久遠的資料未必能提升預測表現，本研究發現非均等權重方法可能適用於颱風雨量這類型的資料，並有助於提升預報校正成效。並且若越近期的資訊貢獻量越大時，未來研究或許也可以考慮不一樣的比重，例如 1:1:4 或是 1:2:8 等，或是找出最優化的比重。

另外，雖然在個案蘇迪勒(TY10)中 FCN、FCNUE112 與 FCNUE124 整體而

言預測表現良好，但在第四與第五報的預報結果仍然可以發現校正後的預報表現有瞬間低下的狀況，這兩報的時機點為蘇迪勒颱風(TY10)登陸台灣的前後兩個時間點，推測可能與台灣降雨分佈轉換的時間點有關。再加上本研究使用每個格點分別建模，因此若原先雨量較低甚至沒有降雨的格點，在降雨分佈轉換後雨量如果突然增強，將使得訓練資料無法給予較有利且充分的資訊，而導致預報校正成效不佳。

除以 CRPSS 評估預報校正成效外，本研究也利用機率預報圖觀察模型校正後的機率預報在三個雨量門檻（分別為 200mm, 350mm 與 500mm）下的「降雨機率」與實際的「觀測降雨指標」是否有一致的降雨空間分布。比較表現較好的 FCN、FCNUE112 以及 FCNUE124 三個模型校正後的機率預報與原始系集機率預報的空間分布。結果發現（圖 4 至圖 6），經過類神經網路校正後，對於在真實累積降雨量有達雨量門檻（以下稱事件發生）的網格點，這些點的預測事件發生的機率較高，而對於事件未發生的那些網格點，預測事件發生的機率則較低，由此可得經校正後機率預報表現提升並掌握降雨空間分布，其中 FCNUE124 則呈現較佳的表現（圖 4 至圖 6）。這些圖是本研究校正方法有效的另一項支持。

藉由以上結果，本研究進一步使用上述研究結果中校正成效較佳的模型架構 FCNUE124 模型，於其他 12 個颱風（表 1）進行預報校正，並進行所有預報校正的綜合校驗分析。結果顯示，原始系集成員經 FCNUE124 模型校正後可提升機率預報表現，例如圖 7 顯示經校正後 CRPS 數值皆有降低。此外，圖 8 顯示預報區辨能力數值如經校正後之準確率 (Precision)、召回率 (Recall rate)、及 F1 Score 等皆有提升，同時也能掌握降雨空間分布。其中準確率、召回率、F1 Score 等數值之定義為：

Precision(準確率)

$$= \frac{\text{no. (True Positive)}}{\text{no. (True Positive) + no. (False Positive)}}$$

準確率是顯示所有預測為正的樣本中(即預報認為會達到累積雨量)，有多少為真實正樣本(實際觀察值達到了累積雨量)。召回率則顯示所有真實正樣本當中，能夠被預測為正樣本的比例：

Recall rate (召回率)

$$= \frac{\text{no. (True Positive)}}{\text{no. (True Positive) + no. (False Negative)}}$$

而這兩個數值可以計算 F1 Score：

$$\text{F1 score} = \frac{2 \times \text{Precision} \times \text{Recall rate}}{\text{Precision} + \text{Recall rate}}$$

圖 8 中 Y 軸數值為經 Precision 與 Recall 數值運算所得的 F1 Score、以及對應的 Precision 與 Recall，數值越高顯示預報區辨能力越佳。X 軸為作為判斷是

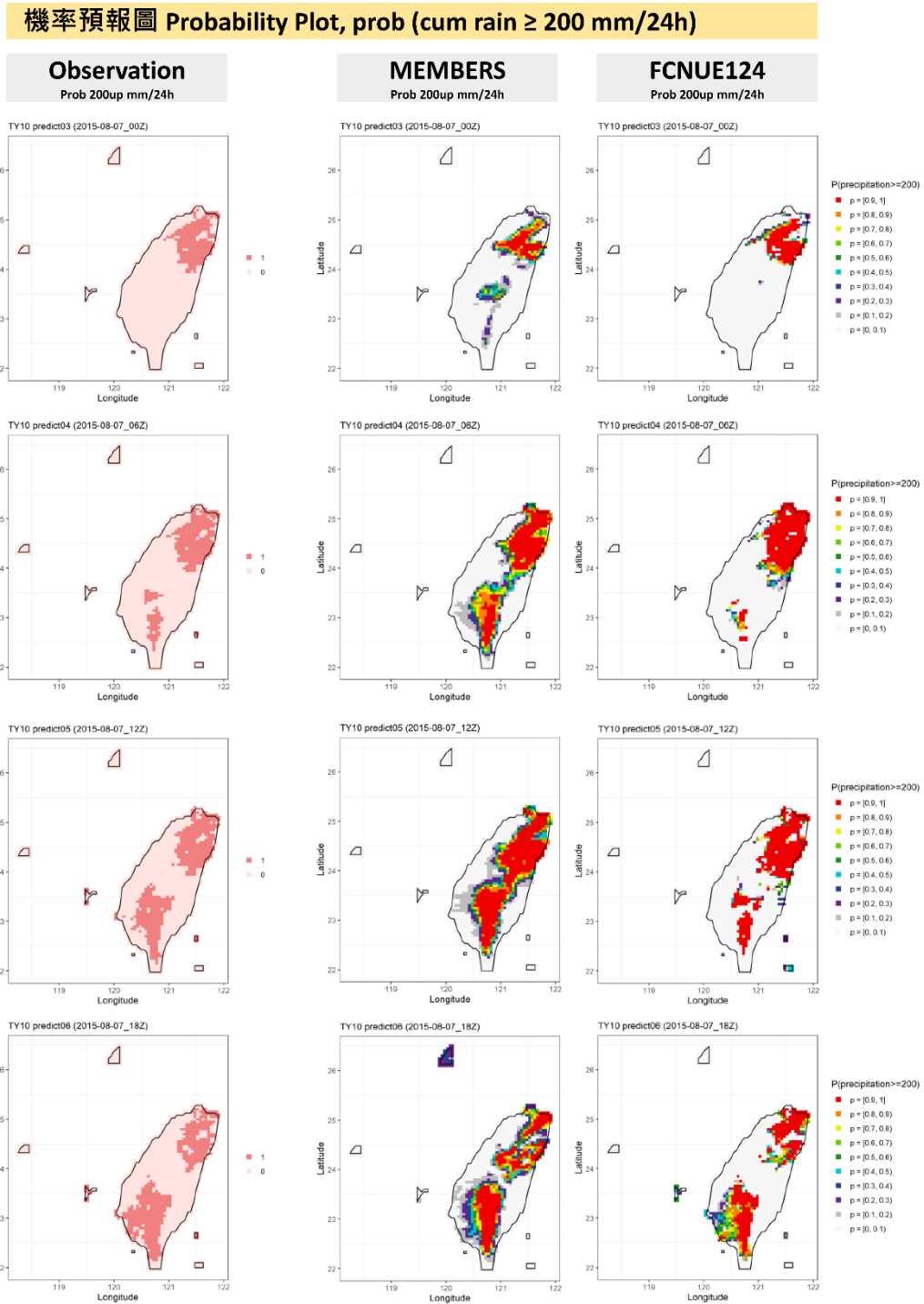


圖 4：蘇迪勒颱風第 3 至 6 報機率預報圖，累積降雨門檻為 200mm/24h(豪雨, Extremely heavy rain)，亦即利用模型所得之分佈來預測累積雨量超過(≥)200 毫米的機率， $p = \text{Pr}(\text{累積雨量超過} \geq 200)$ 。第一欄為真實觀測值降雨指標，若實際雨量達所設門檻，則呈現深紅色( $p=1$ )；第二、三欄分別為原始的系集模式機率預報圖以及 FCNUE124 模型校正後機率預報圖，不同顏色表示模型預報雨量達所設門檻之機率( $p$ )值不同，若該機率值大則顏色接近深紅色。由上至下列依序為預報時間: 第 3 報(0000 UTC 07 Aug 2015)、第 4 報(0600 UTC 07 Aug 2015)、第 5 報(1200 UTC 07 Aug 2015)以及第 6 報(1800 UTC 07 Aug 2015)。

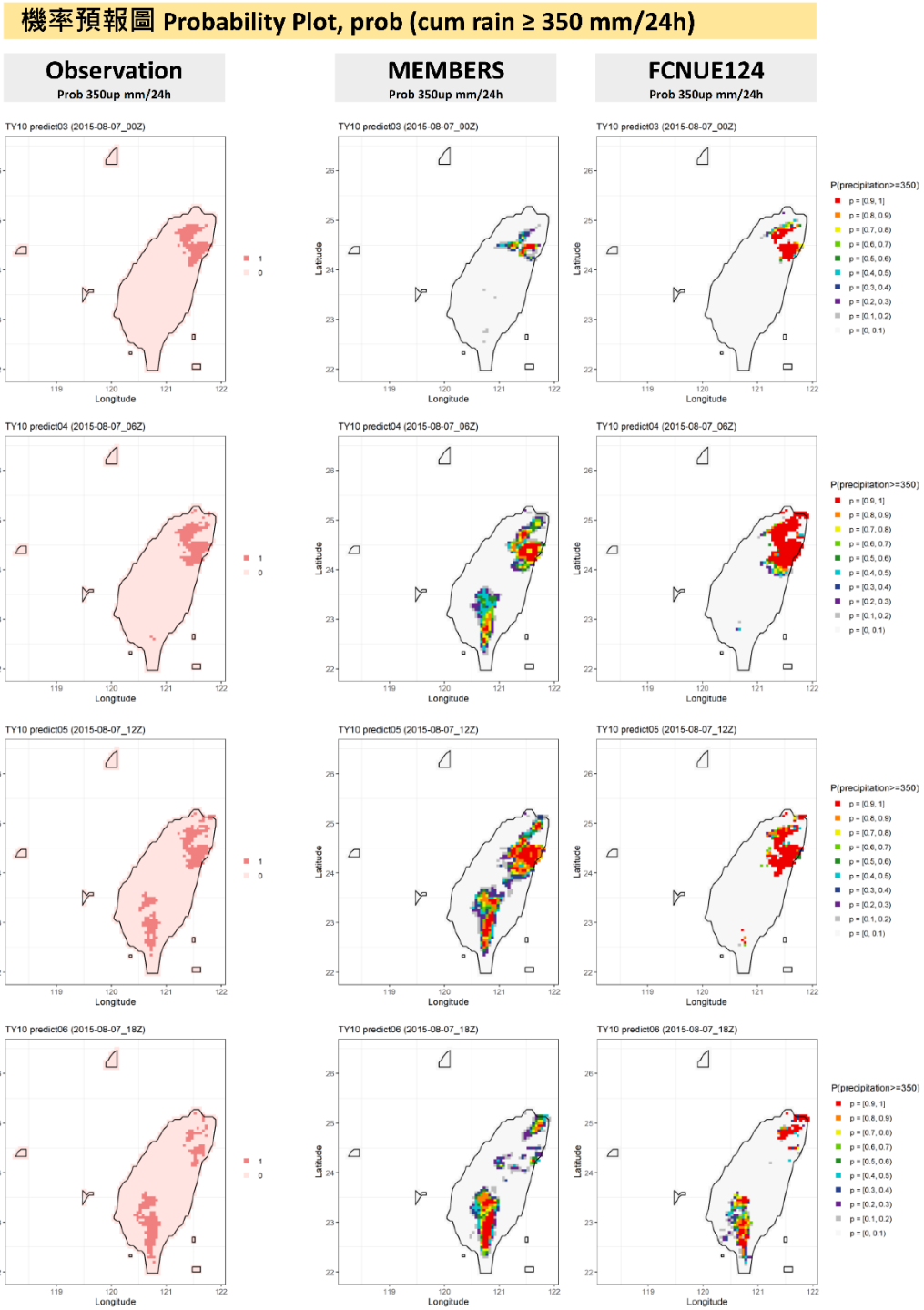


圖 5：蘇迪勒颱風第 3 至 6 報機率預報圖，累積降雨門檻為 350mm/24h(大豪雨, Torrential rain)，亦即利用模型所得之分佈來預測累積雨量超過(≥)350 毫米的機率， $p = \text{Pr}(\text{累積雨量超過} \geq 350)$ 。第一欄為真實觀測值降雨指標，若實際雨量達所設門檻，則呈現深紅色( $p=1$ )；第二、三欄分別為原始的系集模式機率預報圖以及 FCNUE124 模型校正後機率預報圖，不同顏色表示模型預報雨量達所設門檻之機率( $p$ )值不同，若該機率值大則顏色接近深紅色。由上至下列依序為預報時間: 第 3 報(0000 UTC 07 Aug 2015)、第 4 報(0600 UTC 07 Aug 2015)、第 5 報(1200 UTC 07 Aug 2015)以及第 6 報(1800 UTC 07 Aug 2015)。



機率預報圖 Probability Plot, prob (cum rain ≥ 500 mm/24h)

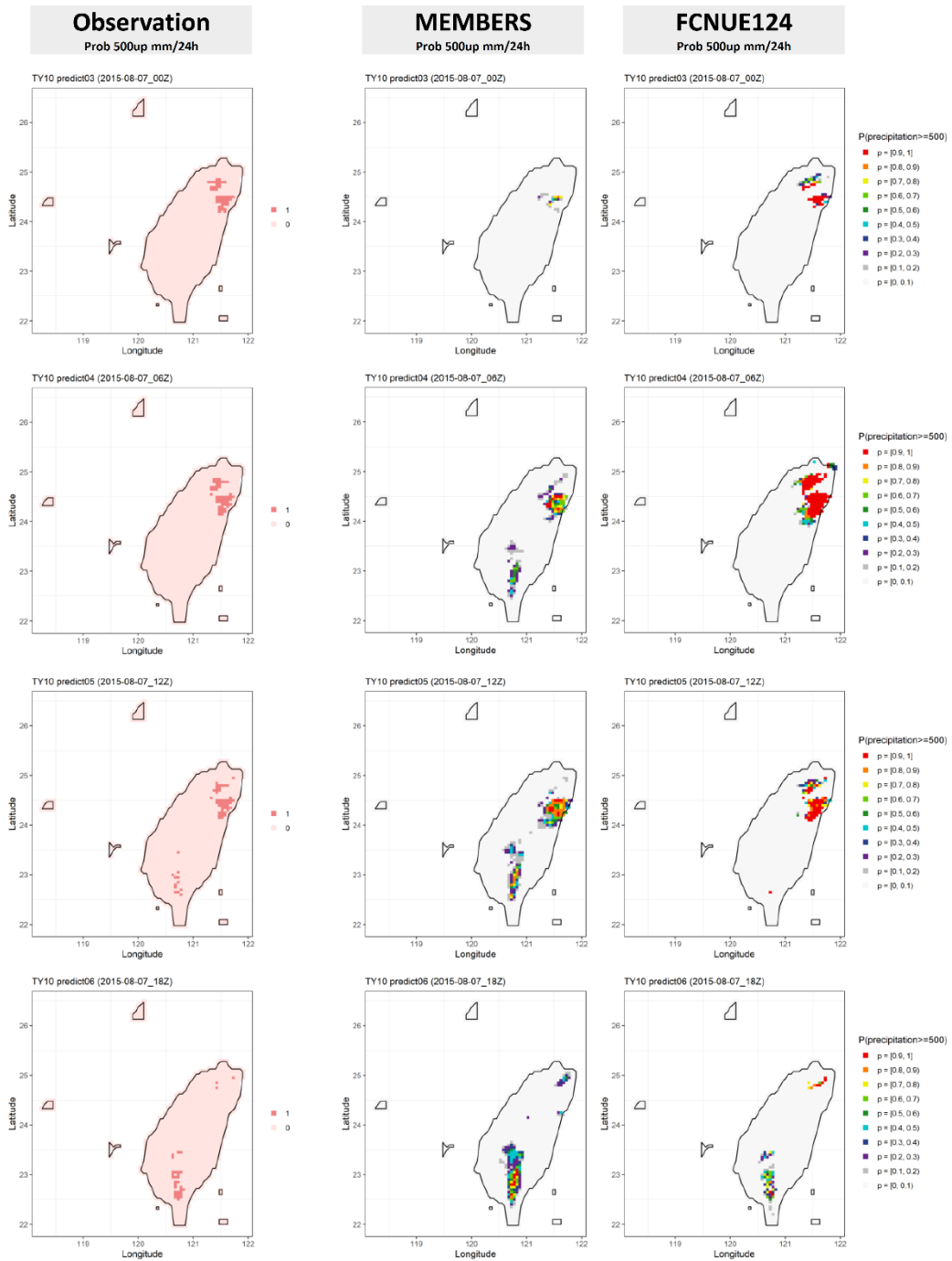


圖 6：蘇迪勒颱風第 3 至 6 報機率預報圖，累積降雨門檻為 500mm/24h(超大豪雨, Extremely torrential rain)，亦即利用模型所得之分佈來預測累積雨量超過(≥)500 毫米的機率， $p = \text{Pr}(\text{累積雨量超過} \geq 500)$ 。第一欄為真實觀測值降雨指標，若實際雨量達所設門檻，則呈現深紅色( $p=1$ )；第二、三欄分別為原始的系集模式機率預報圖以及 FCNUE124 模型校正後機率預報圖，不同顏色表示模型預報雨量達所設門檻之機率( $p$ )值不同，若該機率值大則顏色接近深紅色。由上至下列依序為預報時間: 第 3 報(0000 UTC 07 Aug 2015)、第 4 報(0600 UTC 07 Aug 2015)、第 5 報(1200 UTC 07 Aug 2015)以及第 6 報(1800 UTC 07 Aug 2015)。



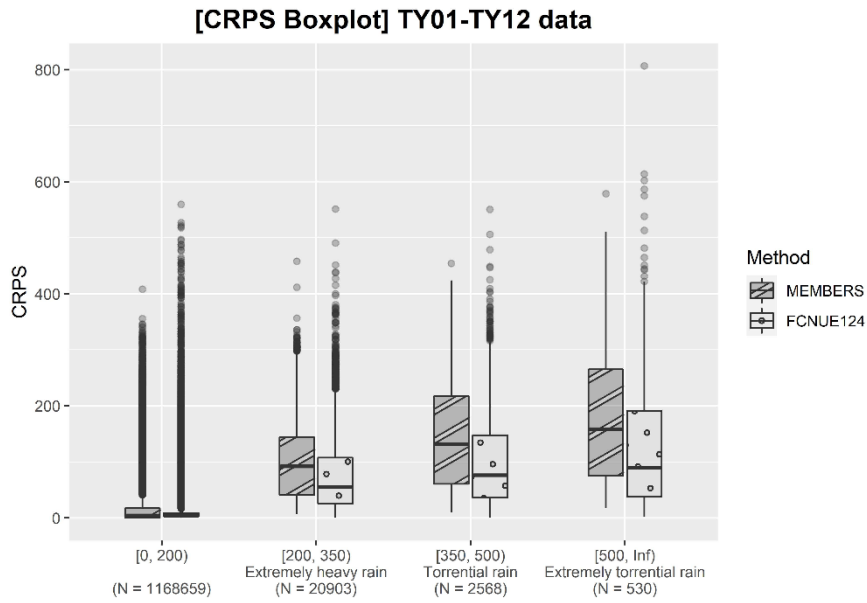


圖 7：校正 12 個颱風之結果 CRPS 數值盒鬚圖。Y 軸為 CRPS 數值，X 軸為真實觀測雨量大小的四種類別，每一種類別中，左邊為經過 FCNUE124 校正後的 CRPS 數值，右邊為原始預報系集成員(Member)預報的 CRPS 數值。CRPS 數值越小則顯示機率預報表現越佳。

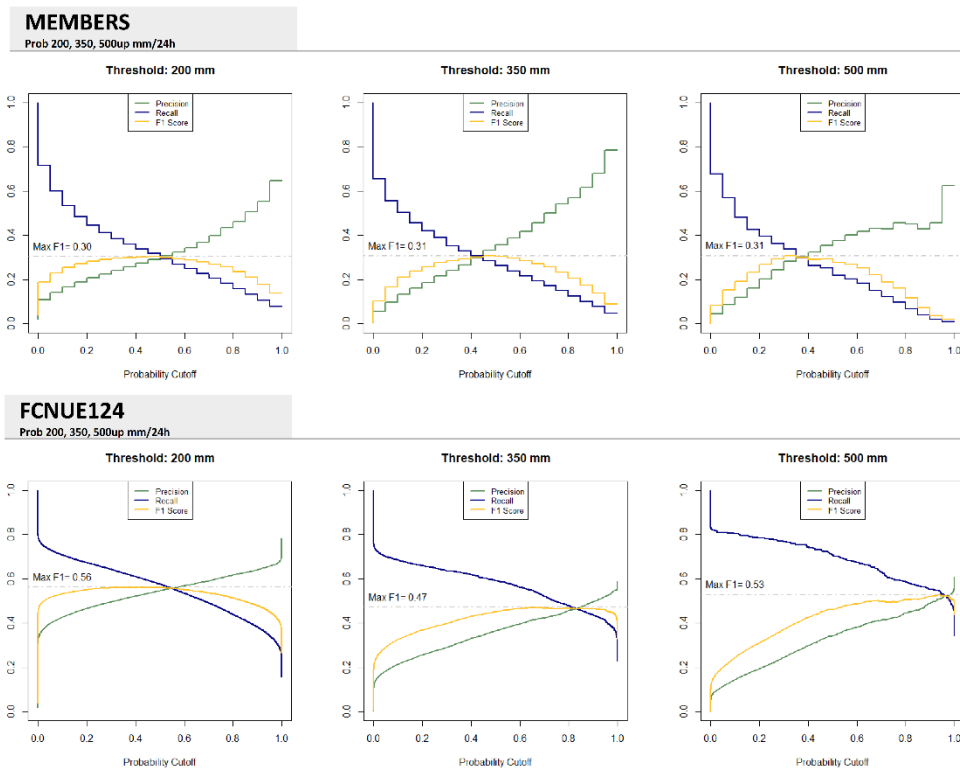


圖 8：評估準則，包含 Y 軸的 F1 Score、Precision、Recall 等數值。X 軸為作為判斷是否為降雨的預報機率門檻。上排為原始系集成員預報(RAW)所得結果，下排為 FCNUE124 所得結果。由左至右排分別是累積降雨量門檻 200mm/24h, 350mm/24h 以及 500mm/24h。

否為降雨的預報機率門檻，例如若校正後的預報機率超過 X 軸上的門檻值，則視為預報認為會下雨達累積雨量(positive)，再以此與真實情況是否達累積雨量進行比較，以決定預報正確(true positive)或預報不正確(false negative)，來計算各個判斷數值；所有預報正確的格點數為下面準確率的分子。最大 F1 Score 數值所在的位置則為最佳預報機率門檻(切點)。

#### 四、結論與討論

本研究使用四種類神經網路模型以及非均等資料權重方法來進行台灣颱風雨量預報校正，探討何種類型的模型較適合颱風雨量預報校正，觀察類神經網路的隱藏層存在重要性、隱藏層節點數以及隱藏層層數對於颱風雨量預報校正的影響。對於非均等資料權重方法，則探討在颱風雨量訓練樣本數稀少且變化快速的情形之下，是否相較於均等資料權重，能有較良好的預測表現。本研究發現，使用非均等資料權重方法於淺層類神經網路模型的 FCN 模式，在本個案有較好的預報校正結果。以下分別討論本研究使用模型的細節，並討論本研究可能的限制及未來可以研究的方向。

##### (一) 淺層與深層類神經網路及節點數之影響

本研究期望藉由建立的四種模型來觀察類神經網路的架構如何影響颱風雨量預報校正成效，首先是將 FCN 與其他三種模型進行比較以了解隱藏層的存在是否能夠提升預測能力，除了本文的蘇迪勒(TY10)颱風，另外也分析了杜鵑颱風(TY12)，結果同樣顯示增加隱藏層未必有較佳的預報校正結果。此外，也發現隱藏層的節點數的增加僅能些微的預報校正成效。本研究針對這兩個颱風，將 FCN、NN2N 與 NN4N 三種淺層類神經模型與 NN3H4N 深層類神經網路模型進行預報校正結果評比，結果都顯示較複雜的深層類神經網路模型的校正成效明顯差於淺層類神經網路模型。因此，本研究認為導致深層類神經網路 NN3H4N 表現不佳的原因，一是颱風降雨機率預報校正不適合太過複雜的模型，這樣的現象同樣出現在 Rasp and Lerch (2018)研究中，過多的隱藏層(深層類神經網路)並沒有改善預報校正結果，反而使得結果較差；二是因為在訓練樣本數過小的情況下使用太過複雜的模型導致模型過度擬合(overfitting)而造成預報校正不佳的狀況。例如在 FCN 模式中，輸入三報的平均值與標準差共六筆資料、加上實際雨量觀察值三筆共九個資料點，而參數個數有四個權重(weight)與兩個偏差(bias)共六個參數，如圖 1b，輸入資料的個數比參數個數

多，參數可以被最佳化；反觀深層類神經網路的參數個數就比輸入資料的個數要多得多，的確容易有過擬合的問題。

## (二) 滾動式更新訓練集資料

本研究是逐格點逐預報時間分別進行建模與預報，而建模時所需的訓練資料集則是以滾動方式（圖 1(e)）進行挑選，理由為颱風多半行進快速，時序過遠的資料所擁有的資訊會與當下的情境有所差異，藉由選擇時間較接近的資訊作為訓練集資料可降低因上述情形而導致的預測誤差。而本研究中所有颱風個案資料集分別僅包含 4 至 15 報原始預報資料（詳細資訊於表 1 第 4 欄位），若挑選較大的訓練集樣本數(報數)，則單一颱風能進行校正的預報報數將會減少，且期望校正之預報可包含陸地雨量的時段，因此本研究選擇使用三報資料作為訓練集資料。然而，最近的系集預報應該與想要校正的標的有最類似的情況，所以權重可能可以更大，未來可以嘗試 1:1:10，或是更極端但前二者非 0 的組合。

## (三) 不同時間點資料使用非均等權重

由於所有颱風個案皆顯示結構較單純的模型有較佳的預報校正成效，本研究進一步利用類神經網路中損失函數 CRPS 的計算方式以及概念，給予建立模型時使用的訓練樣本有不同權重的彈

性，當有利於預報校正的訓練樣本稀少且重要程度不同的情形下，能更加善用訓練樣本中的有利資訊，提升預報校正成效；蘇迪勒颱風(TY10)個案分析顯示此方法的確能提升預報校正成效（如表 2）。而杜鵑颱風(TY12)個案的最佳模型占比分析也同樣顯示非均等資料權重模型 FCNUE124 有比較大的最佳模型比例，給予時間較接近之樣本資料擁有較大的權重時，模型會有更好的預測表現。

## (四) 類神經網路模型內激勵函數 (Activation function)的選擇

在類神經網路中的激勵函數，原先一般多使用線性整流函數(Rectified Linear Unit, ReLU)來進行，但由於颱風期間的雨量是相當極端的，在有強降雨發生的同時，也有許多地區是未降雨或是極小雨的情形，因此接近零或等於零的數值不算少，再加上建模的訓練樣本數相當稀少，這樣的情形會造成 Dead ReLU Problem 的狀況發生(Lu et al. (2019), Arnekvist et al. (2020))，Dead ReLU Problem 指的是參數迭代運算中數值停滯於零無法更新的狀態，本研究也發生同樣情況，為了避免這樣的情形，本研究選擇使用 Softplus 來進行運算。

## (五) 系集預報資料分佈的常態假設

Rasp and Lerch (2018) 提出的方法，假設資料需符合常態分佈，接下來這兩位學者所使用的損失函數則為屬於常態分佈的 CRPS 函數，這個限制對於颱風雨量預報校正的影響有以下幾點。首先，本研究所使用的預測因子為 WEPS20 個雨量系集預報，並藉由類神經網路模型來校正服從常態的颱風雨量預報機率分佈的參數，在預報校正過程中需要考慮兩個面向：第一個面向是本研究使用的 20 個系集預報成員是否符合常態分配；第二個面向是此 20 個成員的原始分配對於預報校正是否有影響。

對於第一個面向，雖然一般普遍認為同一時間台灣地區(全部網格)雨量的分佈不是常態分佈，或是同一地區不同時間雨量的分佈不是常態，可能比較接近 Gamma 分佈，但要特別說明的是系集成員的分佈與實際雨量的分佈不同，同一時間同一地點的系集成員雨量的分佈應該是比較接近常態分佈。例如，本研究以個案蘇迪勒(TY10)為例，將個案資料中出現有強降雨或極端強降雨的五報（第三至第七報），每一報中每一個格點（皆包含 20 筆數值）分別進行 Shapiro-Wilk 常態分佈檢定。結果發現，全部格點中約一半(52%)未達統計顯著，無法推翻系集預報成員符合常態分佈的假設；再進一步將格點範圍限縮至台灣陸地進行 Shapiro-Wilk 常態分佈

檢定，發現只有 25%的格點之系集預報成員才達統計顯著推翻常態分佈。如果在台灣陸地的格點中挑選出雨量大於 200mm 的格點再進行一次 Shapiro-Wilk 常態分佈檢定，則發現僅有 14%的格點之系集預報成員會推翻常態分佈的假設。以上檢定結果顯示，根據系集成員資料，沒有強烈證據推翻常態分佈的假設。雖然這並不等於支持常態分佈，但未來可嘗試以其他分配進行分析，以檢測結論是否不同。

第二個面向則是探討此模型對於預報因子的分佈是否敏感，以釐清模型的預報校正成效是否會因為預報資料的分佈而有明顯差異。在研究初期，本研究也嘗試假設同一時間同一地點的系集成員的預報結果屬於右偏的分佈，所以在建立模型與預報前，利用 log 函數將資料進行轉換，使資料更為對稱更偏向常態，再進行評比此預報校正結果與假設系集成員預報為常態分佈時的預報校正結果。結果顯示，假設系集成員的預報為常態分佈的預報校正結果是較佳的；因此本研究後續都是以未經轉換的預報雨量資料，假設常態分布，來進行後續的預報校正與校正後的機率預報評估。

綜合第一面向與第二面向的討論，將系集分布視為常態分佈以進行建模與預報校正，是合理且不影響校正結果的做法。因此本研究以未經轉換的系集降

兩預報作為預測因子，並使用以常態分佈為假設的 CRPS 作為目標函數。另外，提到影響 CRPS 的分佈假設，若使用梯度下降法來進行求解的計算，在模型迭代過程中需進行兩次的微分，此時若使用非常態分佈假設的 CRPS，在計算上可能會遇到一些困難。而未來研究的天氣/氣候預測因子若是不符合常態分佈時，或可考慮更換損失函數的方式、或是改變求解之計算方法來解決，如最大似估計法等。

#### (六) 將路徑資料作為預測因子及擴大樣本

颱風雨量與颱風路徑有很大關係，因此，若能將路徑放入模型作為預測因子之一，可能可以進一步提升模型的表現。未來希望能累積較多個案，再嘗試將相同路徑的颱風納入建模樣本，再搭配非均等權重方法來進行預報校正，應該能更提升校正成效，這應是未來值得進一步探討的研究方向。此外，本研究雖能提升原始系集的降雨機率預報數值，但並無法調整原始系集預報的降雨數值；未來如能有更多颱風雨量數據時，應可進一步評估本研究的模式是否能進一步提升原始系集預報的可靠性；此外，也能彌補本研究沒有足夠資料作為驗證(validation)資料集的限制。

#### (七) 無法即時校正颱風雨量預報及可能方向

本研究提出的預報校正方法並無法作為即時的校正方法，這是因為本研究利用前三報的 24 小時系級成員預報及對應的真實降雨資料作為訓練資料集，進行模式的訓練，再針對第四報進行校正；然而，當颱風正在發生之時，這個前三報所對應的真實降雨資料尚未被收集，所以此時並無法執行本研究的方法。這是本研究的一大限制，因此本研究提出的預報校正方法並不能作為即時的校正方法。如果要進行即時的預報校正，可能需要不同的資料及作法，例如第一種作法是將時間區間變短，例如使用未來六小時降雨量預報而非未來二十四小時降雨量預報，那麼仍然可以利用本文提出的方法，使用前三報等三個時間點的六小時降雨量預報及對應的真實降雨量數值作為訓練集資料，一起放入模式中訓練，在第四報的時間點時，進行預報校正。然而，這個做法有兩個困難，一個是必須假設預報或真實雨量的紀錄都沒有延遲的情形，如果有延遲，那可能就得再往前使用更早的預報。另一個困難則是這個想法目前還沒有資料可以進行檢測，因此無法評估成效。第二種進行即時校正的作法則是利用相同路徑的颱風資料。相同路徑的颱風帶給各地的雨量比較相似，所以可能可以作

為訓練集資料；然而，這種做法的困難在於現有資料庫中，相同路徑颱風且有同樣時間區間雨量預報的紀錄並不多，所以也需要等待資料擴充之後，才有辦法進行分析及評估成效。第三種作法則是改變機器學習的模型，例如使用卷積神經網路(convolution neural network)，善用鄰近格點的資料而不是每個格點的模式只使用同一個格點的資料。這就不同於本文採用的模式，也是本研究團隊接下來預計嘗試的研究方向。

## 參考文獻

- Arnekvist, I., Carvalho, J. F., Kragic, D., and Stork, J. A. (2020). The effect of target normalization and momentum on dying ReLU. arXiv:2005.06195.
- Bremnes, J. B. (2020). Ensemble postprocessing using quantile function regression based on neural networks and Bernstein polynomials. *Mon. Wea. Rev.*, **148(1)**, 403-414.
- Chang, H. L., Chen, K. J., Wu, C. J., Wang, T., Hong, J. S., and Yang, S. C. (2018). Evaluation, Calibration and Economic Value Analysis of the Probabilistic Precipitation Forecasts from WRF Ensemble Prediction System in Taiwan Area-Part 1: Forecast Evaluation. *Atmospheric Sciences (Meteorological Society of the ROC)*, **46(1)**, 69-104.
- Chang, P. L., Zhang, J., Tang, Y. S., Tang, Lin, P. F., L., Langston, C., Kaney, B., Chen, C. R., and Howard, K. (2021). An Operational Multi-Radar Multi-Sensor QPE System in Taiwan. *Bull Am Meteorol Soc*, **102**, E555–E577, <https://doi.org/10.1175/bams-d-20-0043.1>.
- Chen, Y. R., Hong, J. S. (2017). Bias Correction of Surface Temperature Prediction by Using the Decaying Average Algorithm. *Atmospheric Sciences (Meteorological Society of the ROC)*, **45(1)**, 25-42.
- Chong, E. K. P., and Zak, S. H. (2013). An Introduction to Optimization (Fourth ed.), "Gradient Methods". Hoboken: Wiley, 131-160. ISBN 978-1-1182-7901-4.
- Ferro, C. A. T. (2013). Fair scores for ensemble forecasts. *Q. J. R. Meteorol. Soc.*, **140(683)**, 1917-1923.
- Gneiting, T., Raftery, A. E., Westveld III, A. H., and Goldman T. (2005). Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS

- estimation. *Mon. Wea. Rev.*, **133(5)**, 1098-1118.
- Gneiting, T., Raftery, A. E., Westveld III, A. H., and Goldman T., 2005: Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Wea. Rev.*, **133(5)**, 1098-1118.
- Grönquist, P., Yao, C., Ben-Nun, T., Dryden, N., Dueben, P., Li, S., and Hoefler, T. (2020). Deep learning for post-processing ensemble weather forecasts. arXiv preprint arXiv:2005.08748.
- Hamill, T. M., Whitaker, J. S., and Wei, X. (2004). Ensemble reforecasting: Improving medium-range forecast skill using retrospective forecasts. *Mon. Wea. Rev.*, **132(6)**, 1434-1447.
- Hersbach, H. (2000). Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems. *Weather and Forecasting*, **15(5)**, 559-570.
- Höhlein, K., Kern, M., Hewson, T., and Westermann, R. (2020). A comparative study of convolutional neural network models for wind field downscaling. *Meteor. Appl.*, **27(6)**, e1961.
- Lauret, P., Diagne, M., and David, M. (2014). A neural network post-processing approach to improving NWP solar radiation forecasts. *Energy Procedia*, **57**, 1044-1052.
- Li, C. H., J. Berner, J. S. Hong, C. T. Fong, and Y. H. Kuo, 2020: The Taiwan WRF Ensemble Prediction System: Scientific Description, Model-Error Representation and Performance Results. *Asia-Pacific J Atmos Sci*, **56**, 1–15.
- Lu, L., Shin, Y., Su, Y., and Karniadakis, G. E. (2019). Dying ReLU and Initialization: Theory and Numerical Examples. arXiv preprint arXiv:1903.06733.
- Massidda, L., and Marrocu, M. (2018). Quantile regression post-processing of weather forecast for short-term solar power probabilistic forecasting. *Energies*, **11(7)**, 1763, 1-20.
- Nwankpa, C., Ijomah, W., Gachagan, A., and Marshall, S. (2018). Activation functions: Comparison of trends in practice and research for deep learning. arXiv preprint arXiv:1811.03378.
- Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M. (2005). Using Bayesian model averaging to

- calibrate forecast ensembles. *Mon. Wea. Rev.*, **133(5)**, 1155-1174.
- Rasp, S., and Lerch, S. (2018). Neural networks for postprocessing ensemble weather forecasts. *Mon. Wea. Rev.*, **146(11)**, 3885-3900.
- Taillardat, M., Mestre, O., Zamo, M., and Naveau, P. (2016). Calibrated ensemble forecasts using quantile regression forests and ensemble model output statistics. *Mon. Wea. Rev.*, **144(6)**, 2375-2393.
- Thorarinsdottir, T. L., and Gneiting, T. (2010). Probabilistic forecasts of wind speed: ensemble model output statistics by using heteroscedastic censored regression. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **173(2)**, 371-388.
- Vannitsem, S., Bremnes, J. B., Demaeyer, J., Evans, G. R., Flowerdew, J., Hemri, S., Lerch, S., Roberts, N., Theis, S., Atencia, A., Bouallègue, Z. B., Bhend, J., Dabernig, M., De Cruz, L., Hieta, L., Mestre, O., Moret, L., Plenković, I. O., Schmeits, M., Taillardat, M., Van den Bergh, J., Van Schaeybroeck, B., Whan, K., and Ylhaisi, J. (2020). Statistical postprocessing for weather forecasts—review, challenges and avenues in a big data world. arXiv preprint arXiv:2004.06582.
- Vannitsem, S., Wilks, D., and Messner, J. (2018). Statistical Postprocessing of Ensemble Forecasts. *Elsevier*. eBook ISBN 978-012-812-248-8.
- Williams, R. M., Ferro, C. A. T., and Kwasniok, F. (2013). A comparison of ensemble post-processing methods for extreme events. *Q. J. R. Meteorol. Soc.*, **140(680)**, 1112-1120.
- Zheng, H., Yang, Z., Liu, W., Liang, J., and Li, Y. (2015). Improving deep neural networks using softplus units. *In Proc. IEEE Int. Joint Conf. Neural Networks (IJCNN)*, 1-4. doi: 10.1109/IJCNN.2015.7280459.



# Calibration of Ensemble Probabilistic Forecasts (PQPF) of Typhoon Precipitation Over Taiwan Using Neural Network Models

Joyce Juang<sup>1,2</sup>, Charlotte Wang<sup>2</sup>, Yu-Jyun Huang<sup>2</sup>, Mei-Hsien Lee<sup>3</sup>, Hui-Ling Chang<sup>1</sup>,  
Jing-Shan Hong<sup>1</sup>, and Chuhsing Kate Hsiao<sup>2\*</sup>

1. Central Weather Bureau, Taiwan

2. Institute of Epidemiology and Preventive Medicine, National Taiwan University

3. Department of Mathematics, University of Taipei

(manuscript received 1 March 2022 ; in final form 24 August 2022)

## ABSTRACT

Focusing on typhoon precipitation, this study adopts the neural network (NN) technique to calibrate the bias of 0-24-h probabilistic quantitative precipitation forecasts (PQPF) over Taiwan produced by the WRF ensemble prediction system (WEPS) developed by the Central Weather Bureau. The goals are to explore whether the NN post-processing technique is suitable for the calibration of typhoon precipitation forecasts and to compare the calibration effects among different structures of NN models.

Four NN models (three are shallow NNs and one is a deep NN) are developed and applied to calibrate the precipitation forecasts of Typhoons Soudelor, which brought torrential rain during the period of its passing through Taiwan. Evaluation results indicate that the post-processing procedures improve the performance of calibration of ensemble precipitation forecasts, especially for heavy precipitation. The simple shallow NNs perform better than the deep learning NN model. In addition, we compare the

forecast performance of the NN models when equal temporal weights (unadjusted) or unequal weights are used for the training data. The results with unequal temporal weights have lower forecast errors than the model with equal weights. This model is then applied to other typhoons that attacked Taiwan during 1013-2015, and the finding remains that the unequal weights of simple shallow NN can improve the calibration. However, it is noted that the NN models in this study can correct only the values of forecast probability, not the precipitation pattern.

**Keywords: calibration, ensemble forecast, neural network, probabilistic quantitative precipitation forecasts, typhoon precipitation. doi: 10.53106/025400022022125003001**