

LAPS短時預報系統之系集降水預報發展

張惠玲 黃葳凡 王溫和 陳嘉榮 丘台光

中央氣象局氣象衛星中心

摘要

本研究使用中央氣象局LAPS(Local Analysis and Prediction System)短時預報系統進行系集預報(ensemble forecast)，目的在於降低預報偏差並將預報的不確定性傳達給使用者，以期得到更具參考價值的短時(0-12小時)定量降水預報。

目前發展的LAPS系集降水預報是時間延遲之多模式系集降水機率預報(probabilistic quantitative precipitation forecast (PQPF) of time-lagged multimodel ensemble)。PQPF的優點在於它是利用系集成員實際的資料分布狀況來決定機率，用以描述降水超過某一門檻值的機率，因而可將數值天氣預報過程中的不確定性傳達給預報員或使用者；時間延遲之系集預報(time-lagged ensemble)的優點則在於可利用過去幾報的預報結果來增加系集成員而不需耗費額外計算資源；而多模式系集預報(multimodel ensemble)則可藉著增加系集散度(spread)以攫取預報過程中更多的不確定性來源。除系集成員之組合設計外，對於有偏差(bias)的系集預報系統而言，後端校正(calibration)是一個相當重要的程序，因而未來也將分別採用線性迴歸(linear regression, LR)和類神經網路(artificial neural network, ANN)兩種方法，對系集降水機率預報進行校正以降低預報偏差。此校正程序相當於尋求較佳的系集成員權重，因而藉此校正程序可解決系集成員之權重無法恰當給定的困擾。

關鍵字：系集降水機率、預報線性迴歸、類神經網路

一、前言

氣象局自2002年起與美國海洋暨大氣總署之預報系統實驗室(NOAA/FSL)技術合作發展台灣地區中尺度模式短時預報系統LAPS，目的即在於提升對劇烈天氣的短時(0-12小時)降水預報能力。LAPS短時預報系統主要包括觀測資料導入、非絕熱資料同化及模式預報三部份，其作業流程示意圖參見圖一。LAPS短時預報系統的特色在於大氣分析及初始化過程中已包含雲、雨等非絕熱作用，故可有效解決傳統模式在積分初期的調整問題(spin-up problem)，因而於模式預報初期即可獲得較準確的雨量預報資料。目前進入LAPS分析的資料包括模式預報場、地面觀測資料、探空資料、飛機觀測資料、都卜勒雷達回波場及徑向風場資料、地球同步衛星紅外線及可見光資料與衛星雲導風資料。在觀測資料引進系統後，LAPS隨即進行風場分析、地面場分析、溫度場分析、雲分析及水汽場分析，分析完後再進到一個動力平衡模組，目的在於使動量場、質量場與雲分析的垂直運動場三者保持動力上的平衡，降低模式調整時間及減低重力波雜訊，提供熱啟動(hot-start)模式有較好的初始場。最後再利用中尺度數值模式MM5(fifth-generation

Pennsylvania State University-National Center for Atmospheric Research Mesoscale Model)或WRF(Weather Research and Forecasting model)進行數值預報，輸出即時預報產品。

傳統的數值天氣預報植基於決定論觀點，企圖由最佳的單一初始狀態得到無偏差的單一預報結果，但整個數值天氣預報系統包括觀測資料收集、資料同化、模式預報等過程，每個過程都存在著無法避免的不確定性。渾沌理論(chaos theory)認為初始條件極微小的改變，經過不斷放大，對其未來狀態將造成巨大的差別。系集預報的發展即由此觀點出發，認為天氣預報是一個隨機過程(stochastic process)，而初始狀態應該是一個擾動的機率密度函數(probability density function, PDF)分佈，並且企圖去量化預報過程中的不確定性，以估計未來的機率分布狀況，這是數值天氣預報發展過程中極為重要的轉變。

簡等(2003)利用MM5進行系集降水預報，並以技術得分ETS和BIAS進行校驗，藉以評估系集降水預報系統的預報能力，結果顯示系集平均方式產生的降水預報並非是最佳預報，其技術得分在所有預報(共六個系集成員)中約排名第二或第三。此外，研究中亦嘗試以權重方式來進行系集降水預報，包括(1)由模式在A時段(0-12小時預報)的預報成績決

定 B 時段(12~24 小時預報)及 C 時段(24 ~36 小時預報)的權重；(2)由模式前一年的預報表現決定當年的權重；從技術得分之校驗結果，顯示權重方式的確能提升降水預報能力。簡等(2005)變動初始場、積雲參數法及微物理過程以組成各系集成員，並嘗試各種不同的系集成員組合，以評估各系集預報之降水預報能力，結果發現系集平均之降水預報，不論成員多寡，相對於單一模式成員皆有絕對的優勢；在初始場、積雲參數法及微物理過程三個變因中，考慮所有變因所組成的系集預報可得到最佳的降水預報成績，而在三個變因中，以變動初始場最有利於系集降水預報，變動積雲參數法次之，變動微物理過程最差。這兩篇研究為中尺度模式系集預報提供了重要的參考。

LAPS 短時預報系統之系集預報發展，也是先針對不同的微物理參數化、不同的初始場以及不同模式進行敏感度測試(模式水平分辨率為 9 公里，未使用積雲參數化)。結果發現採用不同的微物理參數化，對於 LAPS 0~6 或 6~12 小時的定量降水預報並沒有明顯的差別；而採用不同的初始場以及不同模式則有明顯的差別，亦即，不同初始場以及不同模式是比較顯著的不確定性因子。因此最後是保留 5 個由不同初始場和不同模式所構成的成員做為系集預報發展所用，分別是(1)以氣象局區域模式 NFS 為背景場的 LAPS-MM5；(2)以 NCEP 全球模式 GFS 為背景場的 LAPS-MM5；(3)以 NFS 為背景場的 LAPS-WRF(ARW)；(4)以 GFS 為背景場的 LAPS-WRF(ARW) 和 (5) 以 GFS 為背景場的 LAPS-WRF(NM)。之後，將分別代稱為(1)LAPS-MM5 : NFS；(2) LAPS-MM5 : GFS；(3)LAPS-WRF/ARW : NFS；(4)LAPS-WRF/ARW : GFS 以及 LAPS-WRF/ NM : GFS。

LAPS 初期的系集預報發展是以 5 個成員平均的方式來進行降水預報，經由長期的校驗結果分析，顯示此系集預報並非技術得分最高的預報，同簡等(2003)。原因可能源自(1)各成員預報能力良窳不齊，某些偏差較大的成員導致整個平均結果較差；(2)系集預報系統無法考慮預報過程中所有的不確定性，因此可能未納入影響降水預報的重要因子。針對第一個原因，我們考慮給予表現較好的模式有較高的權重，表現較差的模式權重較低，但究竟該用什麼標準來界定模式表現的好壞呢？且又該如何由模式表現優劣來給予合適的權重呢？Lu et al.(2007)利用線性迴歸(LR)方法有效地校正一個有偏差的系集預報系統，此外，類神經網路(ANN)技術亦證實可以有效改進 POPF(Mullen and Buizza 2004)。本研究將分別採用 LR 和 ANN 兩種方法，對系集降水機率預報進行校正，以降低預報偏差。此校正程序相當於尋求較佳的系集成員權重，因而藉此校正程序，可解決系集成員之權重無法恰當給定的困擾。

針對第二個原因，限於電腦計算資源及作業時

間之考量，我們無法如前述簡等(2005)在研究中多方考慮各種不確定性因子(初始場、積雲參數法及微物理過程)，組成多達 17 個成員的系集預報，以得到更好的預報結果。因而在研究中擬發展 time-lagged multi-model ensemble，以提供更有參考價值的短時降水預報；關於系集成員的組合設計以及 LR 和 ANN 的校正程序，將在以下的研究方法中詳述。

二、研究方法及初步結果

(一)所採用的資料

1. 模式資料

目前發展的期間延遲多模式系集降水機率預報是採用 LAPS 短時預報系統中的 2 個模式，分別是 LAPS-MM5、LAPS-WRF/ARW (advanced research WRF dynamic core)。關於模式採用的微物理參數化方法，LAPS-MM5 使用 Schults scheme，而 LAPS-WRF/ARW 使用 WSM 5-class scheme。至於模式使用的背景場，LAPS-MM5 和 LAPS-WRF/ARW 兩模式均有採用 NCEP 全球模式 GFS 及氣象局區域模式 NFS 兩種背景場，因此多模式系集成員包括 4 個模式，分別是(1)LAPS-MM5 : NFS；(2) LAPS-MM5 : GFS；(3)LAPS-WRF/ARW : NFS 以及(4)LAPS-WRF/ARW : GFS，4 個模式涵蓋的範圍如圖二所示。

2. 用以校驗之觀測資料

本研究將以 QPESUMS (Quantitative Precipitation Estimation and Segregation Using Multiple Sensors) 經過自動雨量站雨量校正之估量降雨為觀測資料，進行系集預報的後端校正。

(二) time-lagged multimodel ensemble 之設計

time-lagged ensemble 的優點在於：可利用前幾報的預報結果來增加系集成員，而不需耗費額外的計算資源，而 multimodel ensemble 則可藉著增加系集數目，以獲取預報過程中更多的不確定性來源，因此本研究將採用 time-lagged multimodel ensemble 來進行定量降水機率預報 (POPF)，time-lagged multimodel ensemble 的示意圖如圖三。

前述的 4 個模式(LAPS-MM5 : NFS、LAPS-MM5 : GFS、LAPS-WRF/ARW : NFS、及 LAPS-WRF/ARW : GFS) 每 3 小時預報(初始化)一次，預報長度為 24 小時，只取其較佳的前 12 小時預報來進行 time-lagged multimodel ensemble 定量降水機率預報；舉例說明，針對 0-6 小時累積降水機率預報，每個模式的 time-lagged 成員將包括 0-6 小時、3-9 小時及 6-12 小時等三種 POPF，因此 4 個模式共可組成 12(4×3=12) 個成員的 time-lagged multimodel ensemble。

(三)初步結果：

圖四是時間延遲多模式系集降水機率預報之個案(2008 年 6 月 5 日 00Z)結果，此個案受滯留鋒影響，在台灣

西南外海有明顯的中尺度渦旋和中尺度對流系統，中尺度對流系統之移入為台灣南部及東南部帶來豪大雨。圖四左欄是 LAPS 分別在 1, 5, 10, 15, 20, 25, 40mm/6hrs 等 7 個降水門檻時之預報 0-6 小時的降水機率分布，右欄是 QPESUMS 觀測的實際降水機率分布，陰影區(機率為 1)代表估計的降水超過圖上方所示的降水門檻。

(四)校正方法：

除系集成員之組合設計外，對於有偏差的系集預報系統而言，後端校正是一個相當重要的程序(Mass 2003)。本研究未來也將採用 LR 和 ANN 兩種方法，作為系集預報偏差修正(bias correction)的校正程序；此處之校正指的是可信度(reliability)，亦即觀測和預報之分佈在統計上的一致性。做法是使用 7 個依序的機率值做為 LR 和 ANN 模式的輸入資料，以校正 QPFs。先針對每一格點依序計算一組降水門檻(0.25, 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50mm/6h)的預報降水機率，然後分別針對每個選定的降水門檻進行校正，也就是對以此門檻為中心的 7 個相鄰門檻之降水機率進行訓練(training)，以最小化系集預報和觀測之降水機率間的誤差，簡言之，校正程序就是在進行偏差修正；而利用某個降水門檻為中心的 7 個相鄰門檻之降水機率來進行訓練，目的在於擷取以此門檻為中心的機率密度函數(PDF)分布，舉例說明，針對 15mm/6h 此降水門檻進行校正時，是以 1, 5, 10, 15, 20, 25, 30mm/6h 等 7 個降水門檻之降水機率做為校正模式的輸入資料；而在 5mm/6h 此降水門檻時，是以 0.25, 1, 5, 10, 15, 20, 25mm/6h 等 7 個降水門檻之降水機率做為校正模式的輸入資料。至於觀測的降水機率是採用二分法的機率值；若觀測降水達到選定的降水門檻，其值為 1，否則為 0。在不同時間不同格點的一組(7 個)預報降水機率即為一組訓練資料，整個校正過程即是在求出輸入資料(模式預報)與輸出資料(觀測)間的穩定關係，亦即求出權重，此權重可藉著加入新的校正資料而輕易更新。求出權重後，即可將權重應用到一組驗證資料(validation data)中，得到新的 QPF 值(每一格點有一 QPF 值，即校正後的 QPF 值)；此 QPF 值若為負值，則強制設定為 0，若大於 1，則強制設定為 1。

上述的驗證部分是採用交叉驗證(cross-validation)：倘若有 10 個個案，是以其中 9 個個案去校正剩餘的第 10 個個案；亦即，利用 9 個個案校正後得到的權重作為第 10 個個案(驗證資料)的權重，求出驗證個案偏差修正後的 QPF 值。針對所有選定的降水門檻進行偏差修正後，最後還需要一個檢驗步驟以確定每個格點的機率分布呈現單調分佈(monotonic distribution)，因為降水達到較大門檻的事件一定也達到較小門檻，所以較小門檻的降水機率理當高於或等於較大門檻的降水機率。經過單調分佈檢驗後的降水機率即為最後的 QPF 產品。以下再分別針對 LR 和 ANN 方法詳細說明。

1. 線性迴歸(linear regression, LR)方法

LR 方程如下：

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

其中 $M=7$ ， $f_i(x, t)$ 是前述 7 個依序的機率輸入值(也就是系集預報在 7 個降水門檻時的平均降水機率)， $P(x, t)$ 是觀測的降水機率，而 a 是一個常數，代表誤差乘餘項。藉著最小化觀測的降水機率 $P(x, t)$ 和模式導出的降水機率 $(a + \sum_{i=1}^M b_i f_i(x, t))$ 之間的誤差，即可求出 a 和 b_i 各項係數，簡言之，就是利用最小平方方法求出各個迴歸係數：

$$\min \sum_{j=1}^n \left[P(x, t) - \left(a + \sum_{i=1}^M b_i f_i(x, t) \right) \right]^2 \quad (2)$$

求出各迴歸係數後，將驗證資料的系集預報降水機率代入方程式(1)，即可得到修正的降水機率，也就是偏差修正後的降水機率。此機率若為負值(或 >1)，則重設為 0(或 1)。

2. 類神經網路(artificial neural network, ANN)方法

本研究亦將採用一個三層前饋式(feed-forward)的類神經網路模式以進行偏差校正，ANN 的基本架構如圖五所示。本研究之輸入層(input layer)有 7 個節點(node)，分別接收前述 7 個依序的機率值；輸出層(output layer)只有 1 個節點，輸出 1 個介於 0 和 1 之間的機率值，亦即預期的無偏差機率值；隱藏層(hidden layer)位於輸入層和輸出層之間，是中間處理的黑盒子，透過一組權重將輸入資料轉換為輸出資料；而權重則是透過一個雙曲線函數(sigmoid function)藉著最小化輸出機率和實際觀測降水機率的均根差(root-mean-square error, RMSE)來求取。

簡言之，整個校正過程需要一組用以訓練的輸入資料，和一組無偏差的目標資料(即用以校驗的觀測資料)，並在訓練過程中定義一個客觀函數(objective function)，用以測量在某一降水門檻時，類神經網路模式的輸出機率和實際觀測降水機率(目標機率)間的差距，本研究採用的客觀函數是 RMSE：

$$\sqrt{\frac{1}{n} \sum_{j=1}^n (p_j - o_j)^2} \quad (3)$$

其中 p_j 是類神經網路模式由一組(n個)訓練資料求得的輸出機率， o_j 是與訓練資料相對應的觀測降水機率(若觀測降水率達到降水門檻，值為 1，否則為 0)，而 n 是用以訓練的樣本數。類神經網路模式藉著最小化客觀函數 RMSE，以搜尋最佳的權重，並決定出輸入與輸出資料間的關係式。

三、結論及展望

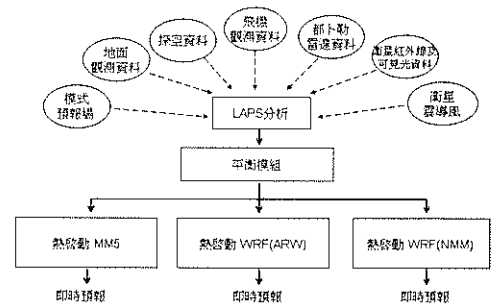
在中尺度數值模式之短時預報方面，國內尚未有系集降水機率預報之發展，本研究嘗試發展時間較短之多模式系集降水機率預報的目的，除了期望得到較準確的定量降水預報外，更重要的是將預報過程中的不確定性傳遞給預報員或使用者；系集預報最終的目的不是要得到一個最佳的預報結果，而是希望提供更多的可能性資訊。一些出乎預報之外的

狀別時有所見，當多個模式顯示颱風將往某特定方向移動，惟獨一家模式大相逕庭時，這種小的機率代表一種可能性，不應該完全忽略；而當各家模式預報的差異較大時，代表系統較不穩定，不確定性很大。系集降水機率預報的優點在於它是利用系集成員實際的資料分布狀況來決定機率，用以描述降水超過某一門檻值的可幾率，因而可將數值天氣預報過程中的不確定性傳達給預報員或使用者。這將是未來數值天氣預報的發展趨勢。

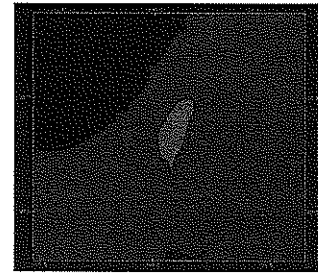
如何設計系集預報的組成成員以及如何決定系集成員的權重是系集預報發展過程會遇到的兩個主要問題，本研究針對此二問題提出時間延遲之多模式系集預報的構想，並利用後端校正程序來解決權重給定的困擾。不管是LR或是ANN模式，不同的訓練樣本將會產生不同的權重，其中非線性的ANN模式比LR需要更多的訓練樣本以得到較好的結果；如何選定訓練樣本以及該納入多少訓練樣本才能達到最大的效益將是一個重要的研究課題。

參考文獻：

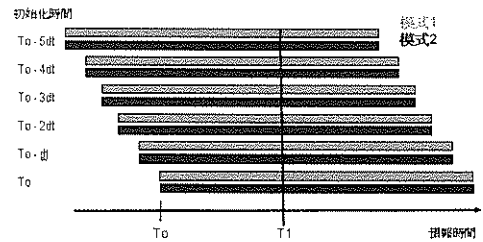
1. 簡國基與鄧仁星，2005：中央氣象局LAPS/MMS 系統在短期(0-12 小時)定量降水預報之應用—甄高(Vamco, 2003)颱風個案研究，大氣科學，33，77 - 99。
2. 簡芳菁、蕭育琪、周仲島、林沛練、楊明仁、洪景山、鄧仁星、林慧娟，2003：MMS 系集降水預報之校驗，大氣科學，31，77 - 94。
3. 簡芳菁、柳蕊秦、周仲島、林沛練、洪景山、蕭冷鳳，2005：2003 年梅雨季MMS 系集降水預報，33，255 - 275。
4. Lu, C., H. Yuan, B. Schwartz, and S. Benjamin, 2007: Short-range forecast using time-lagged ensembles. *Wea. Forecasting*, 22, 580-595.
5. Mass, C. F., 2003: IFPS and the future of the National Weather Service. *Wea. Forecasting*, 18, 75-79.
6. Mullen, S. L., and R. Buizza, 2004: Calibration of probabilistic precipitation forecasts from the ECMWF EPS by an artificial neural network. Preprints, *17th Conf. on Probability and Statistics in the Atmospheric Sciences*, Seattle, WA, Amer. Meteor. Soc., J5.6.



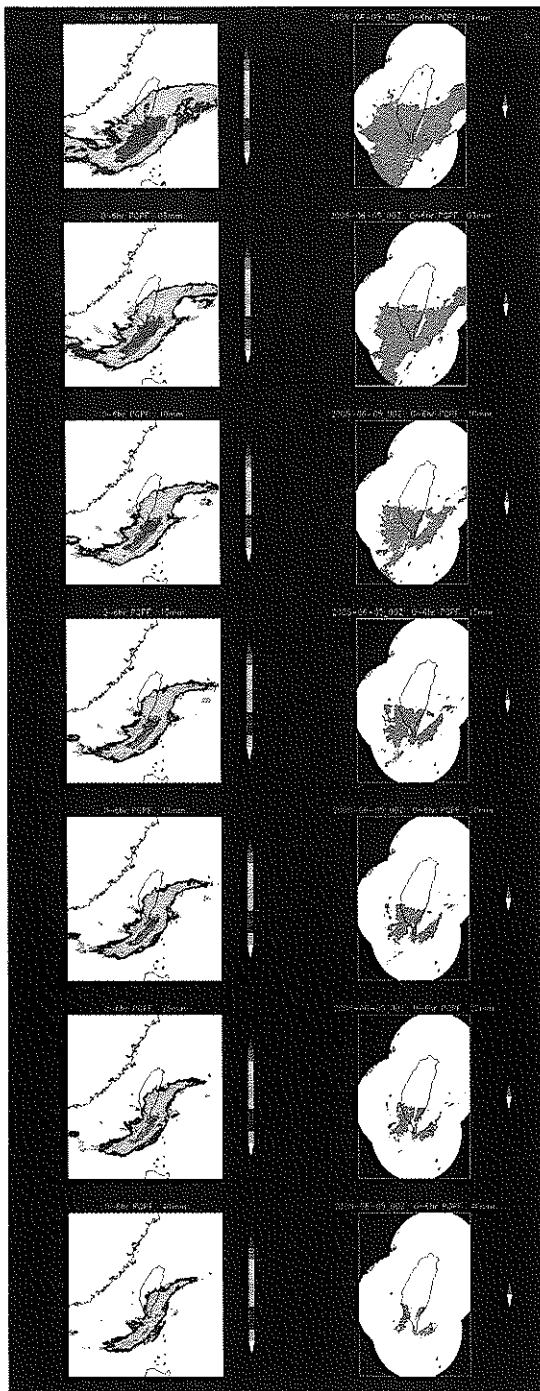
圖一、LAPS 短時預報系統之作業流程示意圖。



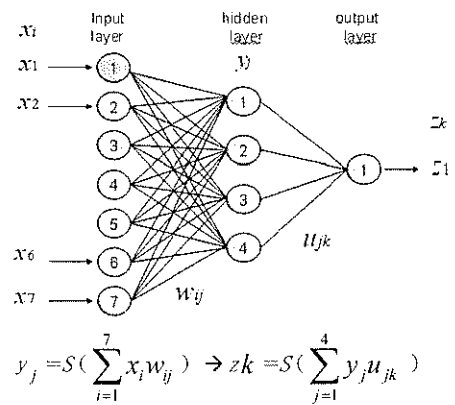
圖二、LAPS 短時預報系統之模式範圍。



圖三、time-lagged multi-model ensemble 的示意圖。



圖四、左欄是APS預報0~6小時的降水機率分佈，右欄是QPESUMS觀測的實際降水機率分佈，陰影區代表右欄中的降水超過圖上方所示的降水門檻。個案時間是2008年6月5日00Z。



圖五、類神經網路架構圖。