

# 本土化極短期強對流機率預報系統

陳威志<sup>1</sup> 林允才<sup>1</sup> 王溫和<sup>1</sup> 吳佳蓉<sup>1</sup> 劉郁青<sup>1</sup> 馮智勇<sup>2</sup> 陳奕翰<sup>2</sup> 林涵芳<sup>2</sup>

<sup>1</sup>中央氣象局衛星中心 <sup>2</sup>多采科技有限公司

## 摘要

目前中央氣象局臺灣自動即時預報系統 (TANC) 之首要目標, 為針對弱綜觀環境下、未來 1 小時之午後對流 (定義為雷達回波 $\geq 35$  dBZ) 發生的可能性。最初使用回波發生頻率、回波變化趨勢、對流可用位能、對流抑制能、層化平均濕度、輻合輻散量、對流胞發生起始位置、對流起始位置參數等 8 個因子, 藉由模糊邏輯並基於美國國家大氣研究中心 (NCAR) 設定之權重與隸屬函數, 計算格點發生強對流之可能性 (likelihood)。

在數值模式的空間解析度不斷提升並發展逐時資料同化的模擬策略, 以及觀測方面的衛星、雷達及地面真實場也都持續精進下, 勢必需要彈性納入更具解釋能力的因子, 並依據臺灣的地理特性分區提供機率預報。本研究透過空間分群並結合數值模式、衛星、雷達及地面真實場等本土化因子建置複邏輯斯迴歸模型, 相較於現行 TANC 系統, 顯示該模型除有較佳之預報表現外, 也具有因應各地天氣特性來增減或更換預報因子的彈性。

**關鍵字:** 極短期、弱綜觀、機率預報、午後對流、複邏輯斯迴歸

## 一、前言

近年來除了全球暖化的影響之外, 隨著都市化的發展, 都市熱島效應也對臺灣的降雨造成影響, 極端降雨事件發生的頻率提高、延時縮短且強度加劇。相較於梅雨鋒面、颱風等天氣系統, 這種強對流系統具有隨著時空尺度變化迅速之特性, 可能在短時間內生成、發展及產生劇烈降雨等; 然而不完善的模式物理與對流參數化, 以及觀測資料密度不足以解析中尺度邊界層特性, 皆使預報更具挑戰性, 因此如何提升即時預報 (nowcast) 的表現日益重要。

目前中央氣象局臺灣自動即時預報系統 (Taiwan Auto-NowCaster, TANC) 之首要目標, 為針對弱綜觀環境下、0.01 度 (約 1 公里) 空間格點未來 1 小時之午後對流 (定義為雷達回波 $\geq 35$  dBZ) 發生的可能性 (黃等 2016)。最初使用回波發生頻率、回波變化趨勢、對流可用位能、對流抑制能、層化平均濕度、輻合輻散量、對流胞發生起始位置、對流起始位置參數等 8 個因子, 藉由模糊邏輯演算法 (fuzzy logic algorithm) 並基於美國國家大氣研究中心 (NCAR) 設定之權重與隸屬函數, 計算格點發生強對流之可能性 (likelihood)。

然而可能性與氣象慣用的機率在定義並不相同, 難以將可能性預報轉換成機率預報, 應用上受到限制。馮等 (2020) 以 TANC 預報因子透過複邏輯斯迴歸 (multiple logistic regression) 方法建置模型, 並有

相當的預報表現。在數值模式的空間解析度不斷提升, 且已開始發展逐時資料同化的模擬策略, 而在觀測方面的衛星、雷達與地面真實場亦持續精進下, 極短期強對流機率預報系統勢必將需彈性納入更具解釋能力的本土化因子, 並依據臺灣的地理特性分區提供機率預報。

第二段簡述本研究使用之預報因子資料以及個案挑選標準, 並且簡介複邏輯斯迴歸、前進選擇法、K-means 集群分析以及機率預報校驗方法, 第三段敘述利用本土化因子建置之複邏輯斯迴歸對流機率預報系統發展概況與成效, 第四段以個案的方式校驗 TANC 系統之可能性預報、以及個別利用 TANC 因子與本土化因子建置之複邏輯斯迴歸對流機率預報系統, 最後則為總結與未來展望。

## 二、使用資料與研究方法

(一) 使用資料:

### 1. TANC 系統 8 個預報因子:

TANC 系統原始使用之 8 個預報因子, 包含回波發生頻率、回波變化趨勢、對流可用位能、對流抑制能、層化平均濕度、輻合輻散量、對流胞發生起始位置及對流起始位置參數。其中, 回波發生頻率、回波變化趨勢為氣候因子, 來自 Lin et al. (2011) 統計 2005-2008 年間暖季 (5-10 月) 弱綜觀環境午後 (12-20L) 雷達回波 $\geq 40$  dBZ 的發生頻率與變化趨勢。

對流可用位能、對流抑制能及層化平均濕度來自氣象局CWB-RWRF模式輸出；輻合輻散量來自地面觀測演算得出；而對流胞發生起始位置及對流起始位置參數則是以雷達觀測為基礎後製之產品。

## 2. 本土化預報因子：

建置複邏輯斯迴歸模型所使用之本土化預報因子，係從120個因子中挑選出各類型資料較有解釋能力之79個因子（圖1），包含了氣象局鄉鎮尺度網格分析值（Ground Truth, GT）、模式（MODEL）、雷達（RADAR）及衛星資料（SAT）。

- (1) 鄉鎮尺度網格分析值資料：包含本身參數，並再計算出溫度與露點溫度差（T-Td），及這些參數隨時間變化率（delta XX），以及分8個方位以數學角呈現之風向。
- (2) 模式資料：使用CWB-RWRF，涵蓋5層（1000、925、850、700及500 hPa）參數。TS為地面溫度，PS為地面氣壓，UMF為上升水氣通量，BLH為邊界層高度，TPW為可用降水量，GHT為重力位勢高度，CV為雷達整合回波。
- (3) 雷達資料：包含氣候因子（同原TANC系統）及反演參數。當中反演參數RadarCu為雷達回波外延，而mosaicPOS為計算出的對流位置函數，兩者皆來自QPESUMS資料。
- (4) 衛星資料：包含各波段亮溫（Brightness Temperature,  $T_B$ ）及過衝雲頂（Overshooting Cloud Top, OCT,  $T_{B6.7-T_{B11.0}} > 2.5 K$ ; Ackerman 1996），皆來自CLAVRx資料。各波段亮溫中， $T_{B13.3}$ 是利用波長13.3  $\mu m$ 可以吸收二氧化碳的特性來突顯高層雲（Mecikalski and Bedka 2006; Lee et al. 2017），常用來評估可能的對流性降雨位置。OCT則利用水氣頻道與紅外線頻道的亮溫差來看平流層中是否有水氣，表示對流是否伴隨有過衝雲頂。

以上這些因子會分區依照前進選擇法逐項挑選出15個具解釋能力的因子來建置模型。

## 3. 弱綜觀個案挑選標準：

使用2018-2019年79個建模個案（表1）及2021年23個校驗個案（表2）皆為暖季弱綜觀環境下之午後對流個案，挑選條件如下：

- (1) 1.5度以內無綜觀系統。
- (2) 1000 hPa~850 hPa風力小於15 kts。
- (3) 天氣概況中本島無華南雲雨帶、南方雲系、低壓帶等字眼。

- (4) 雷達合成回波圖距本島海岸線1.5度內無大片移入的回波。

## (二) 複邏輯斯迴歸（Multiple Logistic Regression）：

原理是利用訓練資料進行模型建置，透過如前進選擇法（forward selection）的篩選程序得到具解釋能力的預報因子並建立複迴歸模型（multiple regression），再將各預報因子的最新輸出代入迴歸模型得到新的機率預報值。複邏輯斯迴歸方程式如下式：

$$p = \frac{e^{(X\beta)}}{1 + e^{(X\beta)}}, X\beta = \beta_0 + \sum_{i=1}^{N_X} \beta_i X_i \quad (1)$$

$p$ 為二分（dichotomous）Y資料發生機率值 $p(\beta)$ ， $\beta_i$ 為待求係數， $N_X$ 則為X變數個數。

本研究採用最大似法（max. likelihood method）推求式中各項係數。假設共有 $N_{rec}$ 筆X、Y配對資料（ $N_{rec} > N_X$ ）供迴歸建模，則發生機率為 $p(\beta)$ 的情況下，概似函數值 $l(\beta)$ 可表示為：

$$l(\beta) = \prod_{i=1}^{N_{rec}} p(X_i)^{Y_i} \cdot (1 - p(X_i))^{1-Y_i} \quad (2)$$

其中，當事件發生時， $Y_i = 1$ ；若事件未發生， $Y_i = 0$ 。為了方便計算，採用對數概似函數值 $\ln[l(\beta)]$ 為：

$$\ln[l(\beta)] = \sum_{i=1}^{N_{rec}} [Y_i \cdot (\beta X_i) - \ln(1 + e^{\beta X_i})] \quad (3)$$

式(3)最大值處一階微分值为0，一階微分如下：

$$F(\beta) = \frac{\partial \ln[l(\beta)]}{\partial \beta_j} = \sum_{i=1}^{N_{rec}} X_{ij} \left[ Y_i - \frac{e^{\beta X_i}}{1 + e^{\beta X_i}} \right] \quad (4)$$

故令式(4)等於0，得一非線性聯立方程組，則可利用疊代求解係數矩陣 $\hat{\beta}$ 。

## (三) 前進選擇法（Forward Selection）：

前進選擇法逐項增加迴歸方程式中的X預報因子至 $N_X$ 個，挑選出有解釋能力之因子，其執行步驟說明如下：

1. 建立Y資料與各單一X變數的邏輯斯迴歸方程式。
2. 計算步驟1中各迴歸方程式係數 $\beta_1$ 的卡方（Chi-Square）統計值，若其對應的發生機率超過使用者設定的門檻值（p-value）則剔除該變數進入前進選擇流程；若發生機率小於門檻值，則保留為候選變數。
3. 於步驟2中的候選變數中以赤池訊息準則（Akaike's Information Criterion, AIC）為指標，挑選解釋變異量最大的X變數為式(1)中的 $X_1$ 變數，得到迴歸方程式。赤池訊息準則的計算方式為： $AIC = 2n - 2 \ln(l(\beta))$ ， $n$ 為迴歸式使用的變數個數。
4. 於前個步驟得到的方程式增加一個X變數，同樣以AIC為指標，保留模式解釋能力最大

者以及差異小於使用者設定的寬容值 (tolerance) 為候選名單。

5. 利用已納入模式的X變數線性估計步驟4候選名單中的各個變數，而相關係數最低的X變數代表與模式中既有的變數共線程度最低，納入迴歸模式。
6. 重複步驟4與步驟5直到挑選出 $N_x$ 個變數為止。

#### (四) K-means 集群分析 (K-means Clustering) :

K-means為非監督式學習。舉例來說，給1組身高體重資料，但不說明這組資料哪些是男生或是女生，利用非監督式學習可以嘗試著區分出男生、女生。本研究將2018年9個午後對流個案共196個時間，8個因子作為特徵，32422個格點作為樣本，試著利用以上特徵分出午後對流的空間分布特徵，運作步驟如下：

1. 設定群數。
2. 評估各個樣本到各群心的距離，如果樣本距離第n個群心更近，則該樣本屬於第n群。
3. 計算每個群中樣本的平均位置，將群心移動至該位置。
4. 重複以上步驟直到各群心位置不再發生變化。

#### (五) 機率預報校驗方法：

使用性能圖 (performance diagram, 如圖4) 校驗TANC可能性預報與複雜邏輯迴歸機率預報。橫軸為成功率SR (Success Ratio), 縱軸為偵測率POD (Probability Of Detection), 其式如下：

$$POD = \frac{Hit}{Hit + Miss} \quad (5)$$

$$SR = 1 - \frac{FalseAlarm}{Hit + FalseAlarm} \quad (6)$$

POD愈高代表預報能掌握住大部分觀測有發生的事件，而SR愈高代表預報的假警報愈少。而圖中的曲線為預兆得分TS (Threat Score)、虛線為偏倚得分BIAS (Bias score) 的等值線。若當預報結果出現在圖的左上角時，代表雖然能掌握住大多數觀測有發生的事件，但也有過多的假警報，整體來說有過度預報的情形。反之，若預報結果出現在右下角時，代表雖然出現假警報的次數較少，但卻漏掉許多觀測有發生的事件，因此有保守預報的情形。當預報結果出現在右上角時，代表此預報結果能掌握多數觀測有發生的事件且鮮少有假警報的情況，因此當預報結果越靠近右上角代表預報表現越佳。

### 三、本土化因子機率預報建置

在建置本土化因子複雜邏輯迴歸模型前，同樣依循過往研究(馮等2020)，使用K-means集群分析對

午後對流個案空間分類的5個分區：北區、西北、沿海、西南、山區(圖2)，再針對各分區建置模型。

圖3可以看到各分區模型的表現，在較低機率的預報有接近對角線的可靠度，代表預報接近實際觀測的頻率，隨著預報的機率愈高而稍微呈現過度預報，在預報較高機率的資料量相較去年的TANC因子複雜邏輯迴歸模型有明顯增加；從ROC圖來看，當ROC曲線下面積 (Area Under Curve, AUC)  $\geq 0.7$ 時，代表模式具有好的區辨能力 (Buizza et al. 1999)，可以看到各分區的AUC皆在0.9左右，顯示對於 $\geq 35$  dBZ對流發生事件都有良好的區辨能力，尤其當中北區的AUC甚至高達0.94。

再從各分區挑選具解釋能力的因子(表3)來看，除了各分區皆有RadarCu的雷達因子之外，也有OCT及TB133的衛星因子被選入，可能是來自於衛星資料的時空解析度相較以往均有所提升。不過衛星因子當中的OCT是利用 $T_{B6.7}$ 及 $T_{B11.0}$ 兩波段的亮溫差 $>2.5K$ 來判斷過衝雲頂，當發展到平流層的高雲時才容易符合此條件，對於初生的對流反而不易掌握。從因子排序中可見OCT的貢獻度僅次於RadarCu，推測是我們選擇建模大部分都是對流較強的個案所致。另外，模式因子不同於TANC系統是以選取中低層平均之熱力因子為主，除了有選入中低層的熱力因子之外，也有選入部分的動力因子。地面觀測資料在TANC系統僅採用以動力為主的輻合輻散場，本土化因子將各觀測參數並考慮其時間變化率個別納入挑選，可以看到熱力因子亦有其貢獻度，且不亞於動力因子。

藉由這些地面、雷達、衛星觀測及模式各層原始資料的納入挑選建置模型，可以看出預報因子的貢獻度與TANC系統的權重略有差異，且在各分區表現有所不同。

### 四、與現行TANC系統校驗表現

本研究收集2021年5到8月間的22個弱綜觀午後對流個案(表2)，比較機率預報系統與現行TANC系統的預報表現情況。為了更清楚了解加入不同預報因子模型之間的表現差異，並再分為TANC因子與本土化因子複雜邏輯迴歸模型，與TANC可能性預報的預報表現相互比較。

根據過去TANC可能性預報校驗顯示0.4為可能性預報之最佳門檻，表示可能性 $\geq 0.4$ 可視為發生 $\geq 35$  dBZ對流；而複雜邏輯迴歸機率預報則是在建模擬合後，利用ROC曲線尋找各機率門檻下命中率(Hit Rate)減去假警報率(False Alarm Rate)的最大值作為此機率預報模型的最佳機率門檻。本研究將TANC因子與本土化因子的2個複雜邏輯迴歸模型，與TANC可能性預報結果比較；當中TANC因子模型最佳機率門檻按分區(北區、西北、沿海、西南、山區)為0.26、

0.23、0.29、0.27、0.21，本土化因子模型為0.77、0.75、0.73、0.67、0.69。

再依據上述最佳機率門檻，將2021年個案的可能性預報、機率預報轉為決定性預報，並進一步計算校驗得分繪製性能圖（圖4）。結果顯示：TANC可能性預報呈過度預報，TANC因子與本土化因子機率預報則有些微的保守預報。BIAS顯示在TANC可能性預報多介於3~5、TANC因子機率預報多介於0.5~0.8、本土化因子機率預報多介於0.5~1之間，代表複雜邏輯斯迴歸模型預報事件發生的頻率較接近於實際觀測的頻率。TANC可能性預報與TANC因子機率預報的TS大部分落在0.1~0.2之間，然而本土化因子機率預報則在0.1左右。本土化因子的建模表現與個案校驗預期結果的落差可能源自於建模資料量的差異，由於在同一時間如缺少部分因子，便無法納入建模，使得本土化因子模型的資料量少於TANC因子模型，可以掌握到的對流特徵較少，而導致在預報表現上稍有落差。

## 五、總結與未來展望

本研究依據各地不同分區的氣候特性，使用對當地具解釋能力的不同本土化因子來建置預報未來1小時回波 $\geq 35$  dBZ對流發生機率的複雜邏輯斯迴歸模型。相較於TANC系統的8個預報因子，本土化因子涵蓋氣象局所產出之各類原始資料，包含衛星遙測、雷達觀測、地面網格化真實場與CWB-RWRF模式各層參數，不受限於固定時段(12-20L)的回波發生頻率、回波變化趨勢等氣候統計因子，並具備後續加入更具解釋能力預報因子的可擴充性。

從複雜邏輯斯迴歸依照具解釋能力選入的TANC預報因子排序中，可知與現行TANC可能性預報各預報因子依權重大小的排序有一致性；本土化因子的部分則是納入各項參數，可以更完整考慮個別熱力及動力因子的貢獻度，因此可見兩模型預報因子的貢獻度略有差異，且在各分區表現有所不同。

利用2021年5至8月間的22個弱綜觀午後對流個案進行TANC可能性預報、TANC因子與本土化因子複雜邏輯斯迴歸機率預報的分區校驗結果比較，可知複雜邏輯斯迴歸機率預報比現行之TANC可能性預報有較佳的預報表現；本土化因子相對於TANC因子之複雜邏輯斯迴歸機率預報在建模時有較佳表現，然而在個案校驗的預報表現則稍有落差。推測是建模時資料量的差異所導致，後續可以透過篩選同一時間各自模型都擁有完整的因子才進入建模，以確保在比較時能有較一致的基準。

目前本土化因子複雜邏輯斯迴歸模型中，雷達預報因子僅包含回波相關參數及氣候統計因子，尚未納入風場資料，後續建議利用氣象局與中央大學合作發展之多都卜勒氣象雷達三維風場合成技術

(WInd Synthesis System using Doppler Measurement, WISSDOM, Liou and Chang 2009)，反演沿地形上之三維風場資料，並透過公式求得輻合輻散場及渦度場，再經由熱動力反演技術亦可反演溫度場及壓力場之三維熱動力場資料，藉由水平及垂直上高解析度的動力及熱力資料，以期可以更加掌握對流的發展；另外也有研究指出從陸基式GPS反演之可降水量，可以作為對流生成與否的參考依據（Benevides et al. 2015），評估以上新加入之本土化因子是否將再進一步提升複雜邏輯斯迴歸模型的預報表現。長遠目標將擴大預報對象至混合型（mixed regime）天氣乃至強綜觀影響之對流系統生成及發展，以完備各種環境下對流系統之即時預報。

## 六、參考文獻

1. 馮智勇、陳奕翰、林涵芳、林允才、王溫和、吳佳蓉與劉郁青，2020：極短期強對流機率預報技術發展。109年天氣分析與預報研討會，中央氣象局。
2. 黃葳芃、張惠玲、林允才、吳佳蓉、唐玉霜、汪琮、邱健倫與陳嘉榮，2016：臺灣自動即時預報系統(TANC)之發展現況。105年天氣分析與預報研討會，中央氣象局。
3. Ackerman, S. A., 1996: Global satellite observations of negative brightness temperature differences between 11 and 6.7 microns. *J. Atmos. Sci.*, **53**, 2803–2812.
4. Barron, J., D. Fleet, and S. Beauchemin, 1994: Performance of optical flow techniques. *Int. J. Comput. Vision*, **12**, 43–77.
5. Benevides, P., Catalao, J., and Miranda, P. M. A., 2015: On the inclusion of GPS precipitable water vapour in the nowcasting of rainfall. *Nat. Hazards Earth Syst. Sci.*, **15**, 2605–2616.
6. Buizza, R., M. Milleer, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteorol. Soc.*, **125**, 2887–2908.
7. Chang, H. L., H. Yuan, and P. L. Lin, 2012: Short-range (0-12h) PQPFs from time-lagged multimodel ensembles using LAPS. *Mon. Wea. Rev.*, **140**, 1496–1516.
8. Chang, H. L., 1997: The study on the afternoon convection in Taiwan area. M.S. thesis. Department of Atmospheric Sciences, National Central University, Taiwan. 200 pp.
9. S.-C. Yang, H. Yuan, P. L. Lin and Y. C. Liou, 2015: Analysis of relative operating characteristic and economic value using the LAPS ensemble

- prediction system in Taiwan area. *Mon. Wea. Rev.*, **143**, 1833–1848.
10. Chen, Y. L., and J. Li, 1995: Characteristics of surface airflow and pressure patterns over the island of Taiwan during TAMEX. *Mon. Wea. Rev.*, **123**, 695–716.
  11. Jou, B. J.-D., 1994: Mountain-originated mesoscale precipitation system in northern Taiwan: A case study 21 June 1991. *Terr. Atmos. Ocean. Sci.*, **5**, 169–197.
  12. Lee, S., Han, H., Im, J., Jang, E, 2017: Detection of deterministic and probabilistic convective initiation using Himawari-8 Advanced Himawari Imager data. *Atmos. Meas. Tech.*, **10**, 1859–1874.
  13. Lin, P. F., P. L. Chang, Ben J. D. Jou, J. W. Wilson and R. D. Roberts, 2011: Warm season afternoon thunderstorm characteristics under weak synoptic-scale forcing over Taiwan Island. *Wea. Forecasting*, **26**, 44–60.
  14. Liou, Y.-C., Y.-J. Chang., 2009: A variational multiple-Doppler radar three-dimensional wind synthesis method and its impacts on thermodynamic retrieval. *Mon. Wea. Rev.*, **137**, 3992–4010.
  15. Mecikalski, J. R., and K. M. Bedka, 2006: Forecasting convective initiation by monitoring the evolution of moving cumulus in daytime GOES imagery. *Mon. Wea. Rev.*, **134**, 49–78.
  16. P. L. Chang, Ben J. D. Jou, J. W. Wilson and R. D. Roberts, 2012: Objective prediction of warm season afternoon thunderstorms in northern Taiwan using a fuzzy logic approach, *Wea. Forecasting*, **27**, 1178–1197.
  17. Roebber, P. J., 2009: Visualizing multiple measures of forecast quality. *Wea. Forecasting*, **24**, 601–608.

## 七、附圖、表

表 1 弱綜觀環境午後對流建模個案列表

2018 年											
05/01	05/11	05/12	05/13	05/15	05/16	05/17	05/18	05/19	05/20	05/24	05/28
05/29	05/30	06/04	06/05	06/06	06/09	06/25	06/26	06/27	07/07	07/08	07/13
07/19	07/20	07/24	07/25	07/26	07/27	07/28	07/29	07/30	08/04	08/05	08/06
08/07	08/08	08/10	08/11	09/01	09/02	09/05	09/06	09/07	09/12	09/13	09/17
09/18	09/19	09/20	09/21	09/22	09/30	10/22	10/23				
2019 年											
05/13	05/22	05/24	05/25	06/05	06/17	07/16	07/24	07/26	07/27	07/29	07/30
08/02	08/03	08/13	08/22	08/27	08/29	09/02	09/09	09/15	10/03	10/04	

表 2 弱綜觀環境午後對流校驗個案列表

2021 年											
05/10	05/13	05/19	05/23	06/10	06/11	06/12	07/03	07/08	07/09	07/11	07/18
07/29	08/11	08/14	08/15	08/20	08/21	08/24	08/25	08/26	08/27		

表 3 本土化預報因子各分區選擇順序，與 TANC 因子可能性預報及機率預報依比重大小排序比較。

分區 順序	本土化因子機率預報					順 序	TANC 因子	
	北區	西北	沿海	西南	山區		可能性預報	機率預報
1	RadarCu	RadarCu	RadarCu	mosaicPOS	RadarCu	1	RadarCu	RadarCu
2	OCT	OCT	OCT	OCT	OCT	2	ClimoFreq	ClimoFreq
3	1000_TT	Clim_freq	1000_Q	1000_TT	1000_Theta	3	RHavg	ClimoTrend
4	700_Q	mosaicPOS	850_Q	700_Theta	TB133	4	StormInitLoc	RHavg
5	PS_MD	850_Q	Td	TB133	TPW_MD	5	ClimoTrend	StormInitLoc
6	delta Td	1000_Q	1000_UU	delta T	BLH_MD	6	SurfDiv	CAPE
7	850_TT	850_VV	CIN_MD	Clim_freq	PS_MD	7	CIN	CIN
8	TT	delta Td	mosaicPOS	RadarCu	delta Td	8	CAPE	SurfDiv
9	TB133	925_Q	TB133	850_Q	mosaicPOS			
10	WD	T-Td	500_Theta	500_Q	Td			
11	CIN_MD	PS_MD	850_UU	850_VV	Clim_freq			
12	1000_VV	TPW_MD	BLH_MD	VV	850_UU			
13	Clim_freq	CV_MD	925_Q	delta U	500_dBZ			
14	850_GHT	700_VV	TS_MD	delta V	925_GHT			
15	500_TT	TB133	500_dBZ	RH_MD	delta V			

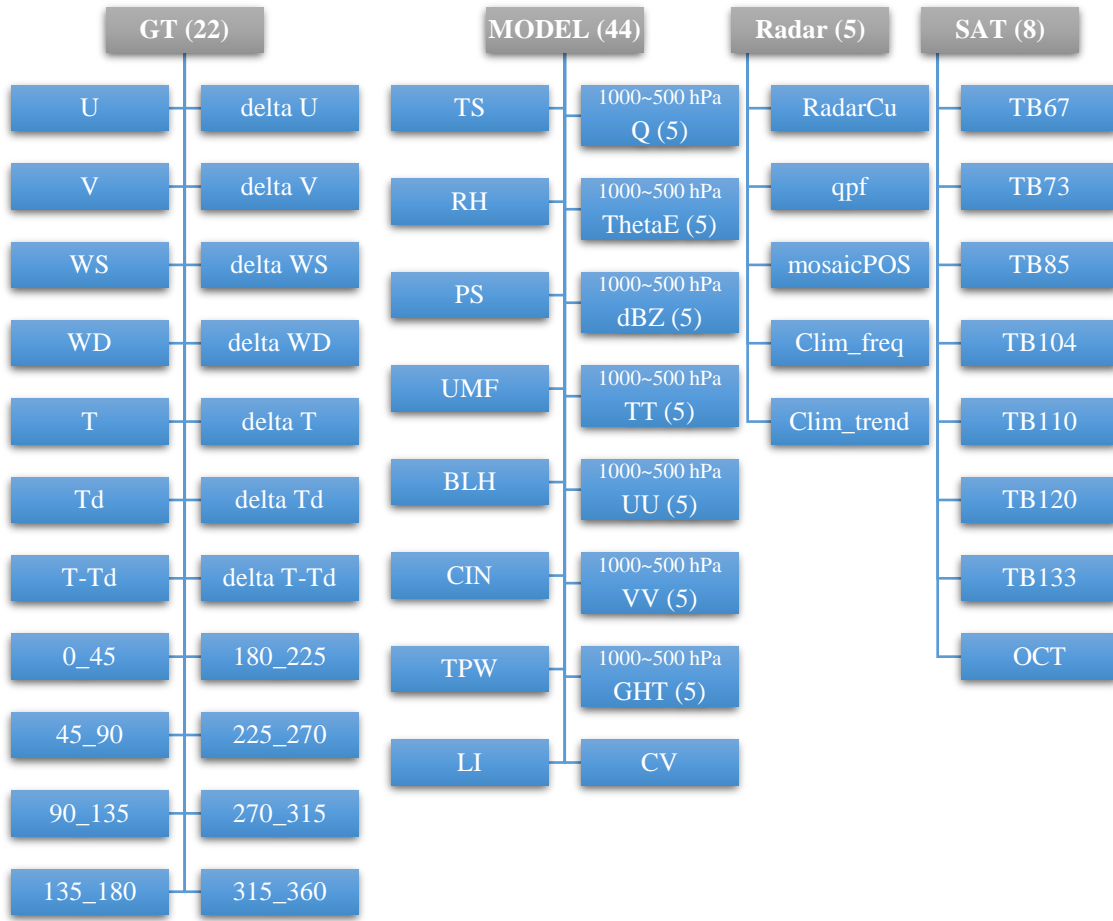


圖 1 複邏輯斯迴歸機率預報的 79 個本土化預報因子列表（括號內為數量）。

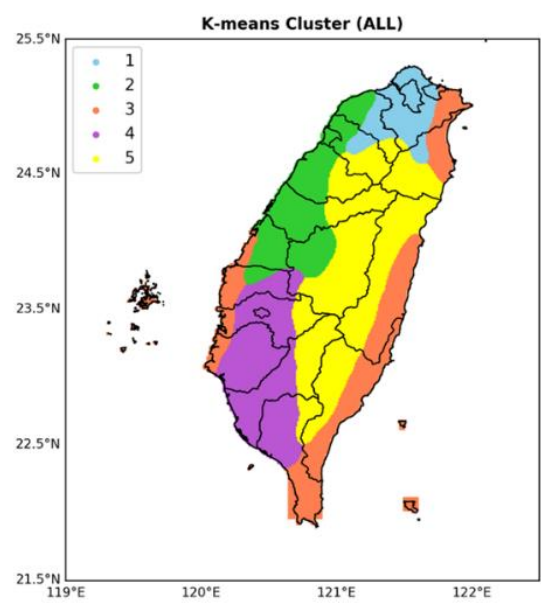


圖 2 K-means 空間分群圖（淺藍色為北區、綠色為西北、橘色為沿海、紫色為西南、黃色為山區）。

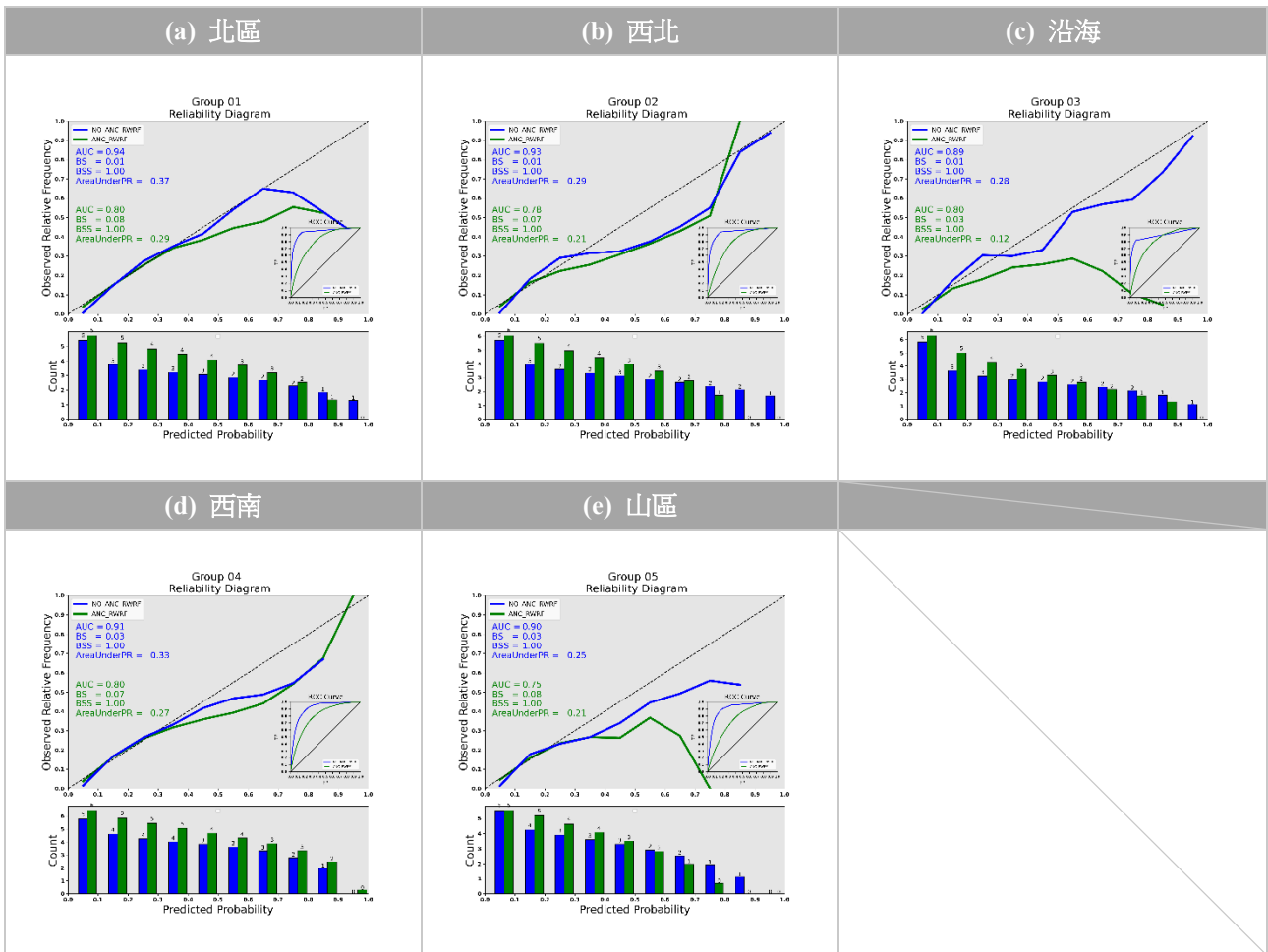


圖 3 2018-2019 年個案本土化因子（藍線）、TANC 因子（綠線）分群建模可靠度圖及 ROC 曲線比較（(a)北區、(b)西北、(c)沿海、(d)西南、(e)山區）。

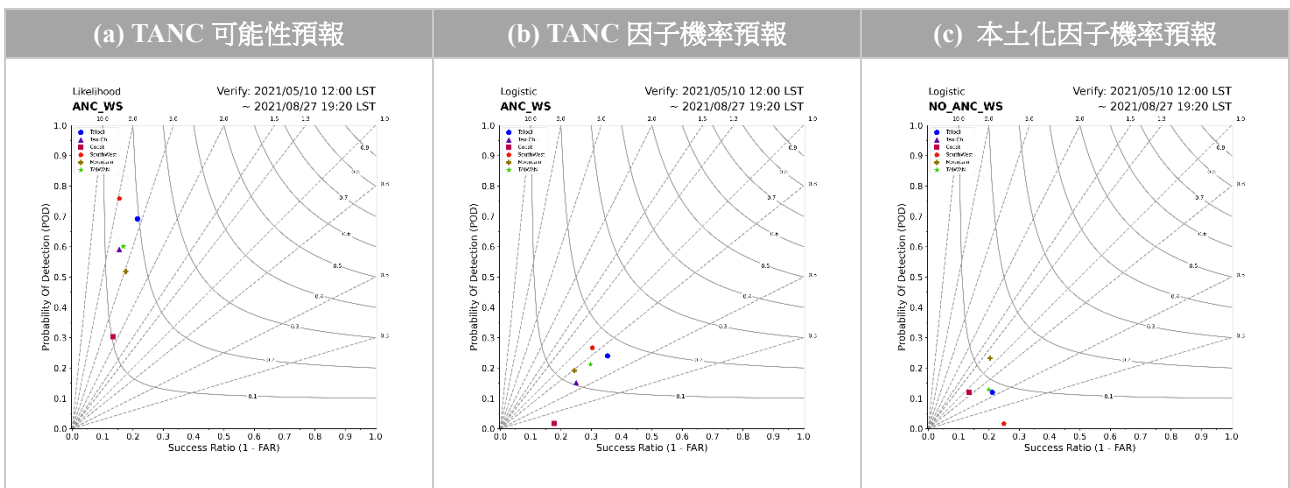


圖 4 2021 年個案分群校驗性能圖（(a)TANC 可能性預報、(b)TANC 因子機率預報、(c)本土化因子機率預報；藍、紫、深紅、紅、褐、綠點各為北區、西北、沿海、西南、山區、不分群平均）。



# Probability Nowcasts of Convective Storm Initiation

Wei-Jhih Chen<sup>1</sup>, Yun-Tsai Lin<sup>1</sup>, Wen-Ho Wang<sup>1</sup>, Chia-Jung Wu<sup>1</sup>, Yu-Ching Liu<sup>1</sup>, Chih-Yung Feng<sup>2</sup>, Yi-Han Chen<sup>2</sup>, Han-Fang Lin<sup>2</sup>

<sup>1</sup>Meteorological Satellite Center, Central Weather Bureau

<sup>2</sup>Manysplendid Infotech, Ltd.

## Abstract

To estimate the afternoon thunderstorm (radar echo  $\geq 35$  dBZ) initiation in 1 hour under weak synoptic-scale forcing, Central Weather Bureau Taiwan Auto-NowCaster (TANC) applied a fuzzy logic approach by using 8 predictors, including radar echo change trend, radar echo frequency, convective available potential energy, convective inhibition, layer average humidity, surface divergence, radar echo extrapolation, and the starting position of convection, to produce 1-hr likelihood nowcasts. The weight and the membership function of each predictor was based on the setting of NCAR. To include more interpretable predictors according to different area, this study classified the Taiwan land grids into 5 categories by k-means clustering method and established multiple logistic regression for each category with predictors, including numerical model outputs, satellite, radar, and ground truth fields. The logistic model shows that it is not only comparable to TANC, but also can further improve the forecast performance.

**Keyword:** nowcasts, weak synoptic-scale forcing, probability forecast, afternoon thunderstorm, multivariate logistic regression