

NAR Labs 國家實驗研究院

台灣颱風洪水研究中心

應用類神經網路於整合定量降雨系集 預報實驗產品之初步研究

102年天氣分析與預報研討會

吳明璋、陳嬾竹、江宙君、黃麗蓉、李清勝

1. 動機與目的
2. 系集預報表現分析
3. 整合方法與初步結果
4. 結論與未來方向

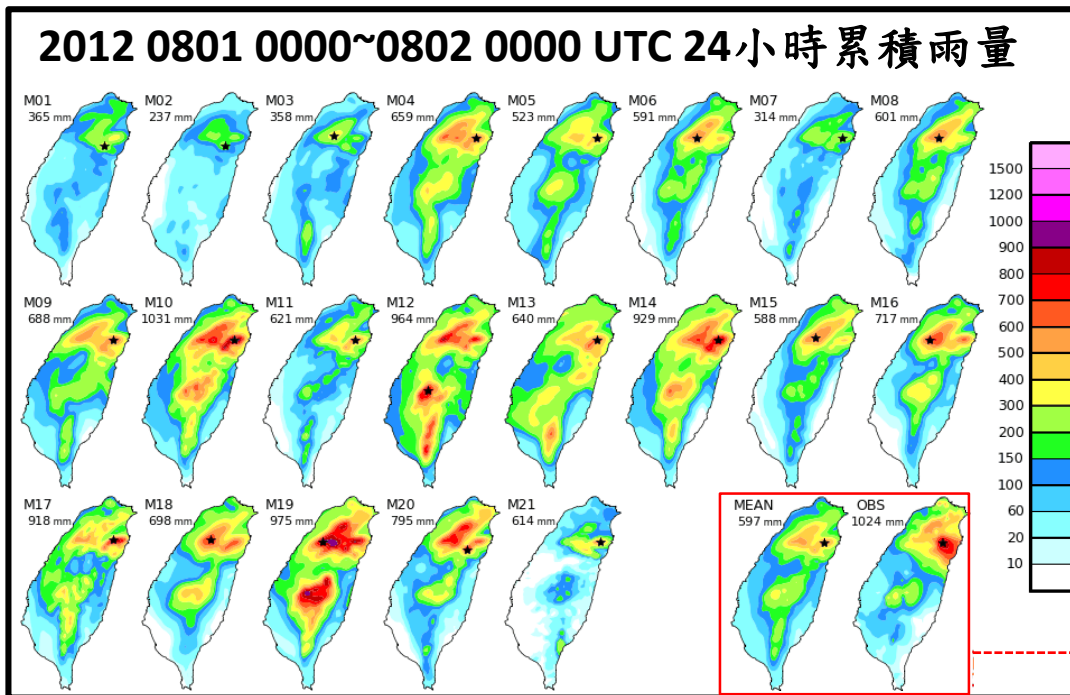
定量降雨系集預報實驗

- 國研院颱洪中心結合學術界與相關作業單位的研發能量與資源，執行定量降雨系集預報實驗 (Taiwan Cooperative Precipitation Ensemble Forecast Experiment, TAPEX)。
- 目的為提升颱風侵台期間定量降雨預報能力。
- 實驗產品可供災防單位於應變作業時之參考，協助提升防救災預警與應變效能。
- TAPEX每日提供4次預報，每次約20組實驗結果。



1. 動機與目的

- 但因系集實驗的成員個數眾多，導致後續水文應用端在流量或淹水預報工作時使用上的困難。



- 以系集成員的平均值，提供使用者參考。
- 僅用平均值可能會低估極大值，對防災應用上較為不利。

□ 實測值為1025mm
□ 系集平均值約近600mm

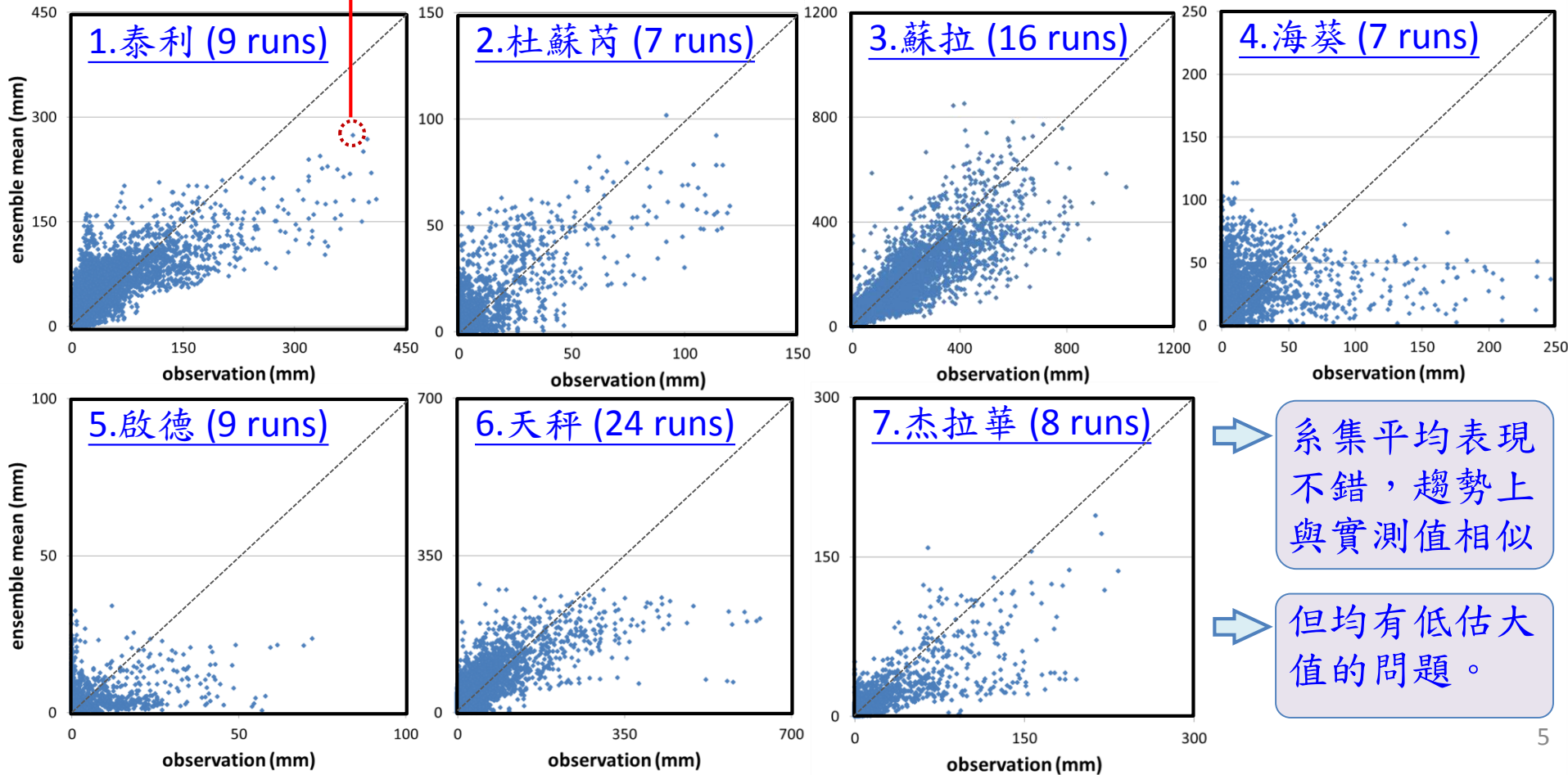
因此

為解決低估極大值的問題，需要發展更佳的整合方式，達到最佳化系集實驗預報產品之目的。

2. 系集預報表現分析

- 2012年7場颱風事件中系集預報平均值的表現

某一測站，某一時刻，未來24小時累積雨量的預報表現



2. 系集預報表現分析

- 系集預報平均值與實測值的落點分析

命中(Hit)

Forecast_ensemble mean (mm)

區間	0~100	100~200	200~300	300~400	400~500	500~600	600~700	700~800	800~900	900~1000	1000~1100	總和
0~100	33621	593	17	4	0	1	0	0	0	0	0	34236
100~200	991	915	166	19	2	0	0	0	0	0	0	2093
200~300	84	388	289	60	14	0	1	0	0	0	0	836
300~400	9	101	179	106	30	8	1	0	1	0	0	435
400~500	1	28	83	91	32	19	6	1	1	0	0	262
500~600	3	10	29	48	32	13	10	3	0	0	0	148
600~700	0	2	8	15	10	12	5	3	0	0	0	55
700~800	0	0	2	9	5	0	1	2	0	0	0	19
800~900	0	0	0	4	2	0	1	0	0	0	0	7
900~1000	0	0	0	0	1	1	0	0	0	0	0	2
1000~1100	0	0	0	0	0	1	0	0	0	0	0	1
總和	34709	2037	773	356	128	55	25	9	2	0	0	38094

嚴重高估

嚴重低估

高達9成

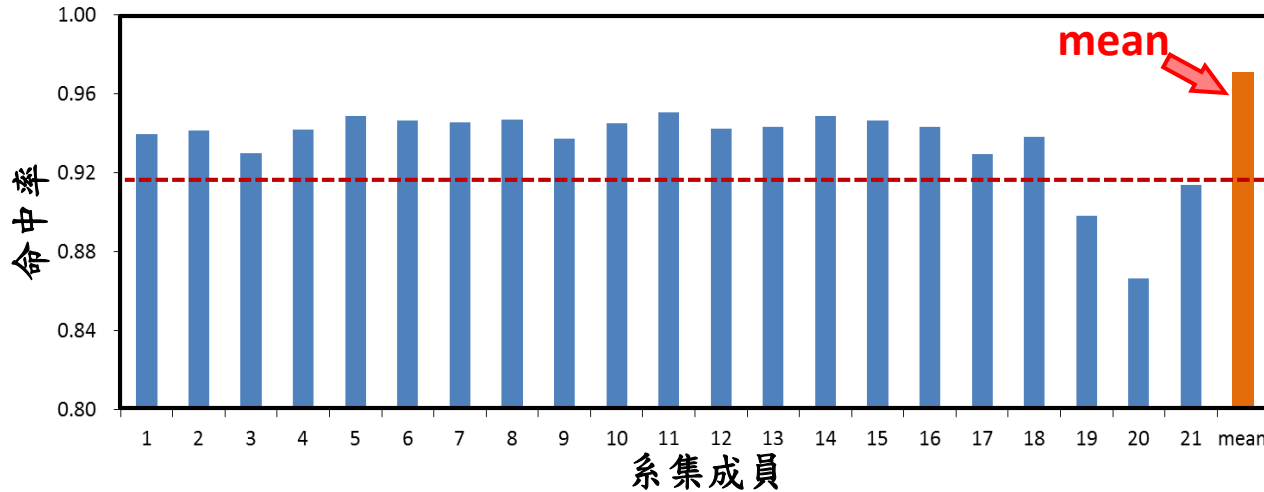
約7成

約僅剩6成

系集平均值在300mm以下時有很高的命中率，但隨預報值的增大，命中率會隨著下降。

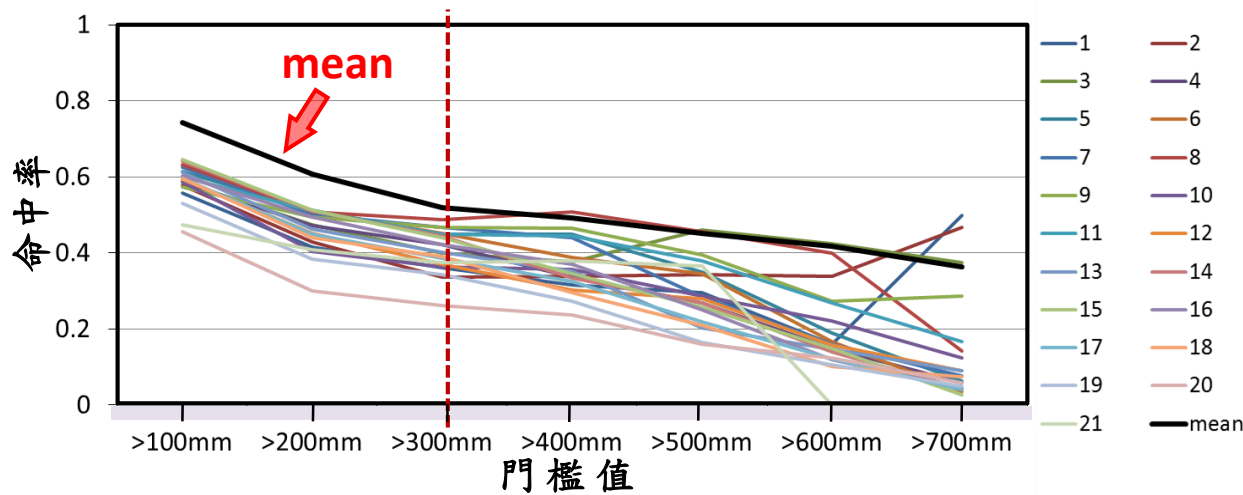
2. 系集預報表現分析

- 各系集成員預報值準確命中的比例：



- Member_19~21 較低，
- 其餘Member表現差不多，約有0.92以上。
- 系集平均表現最好。

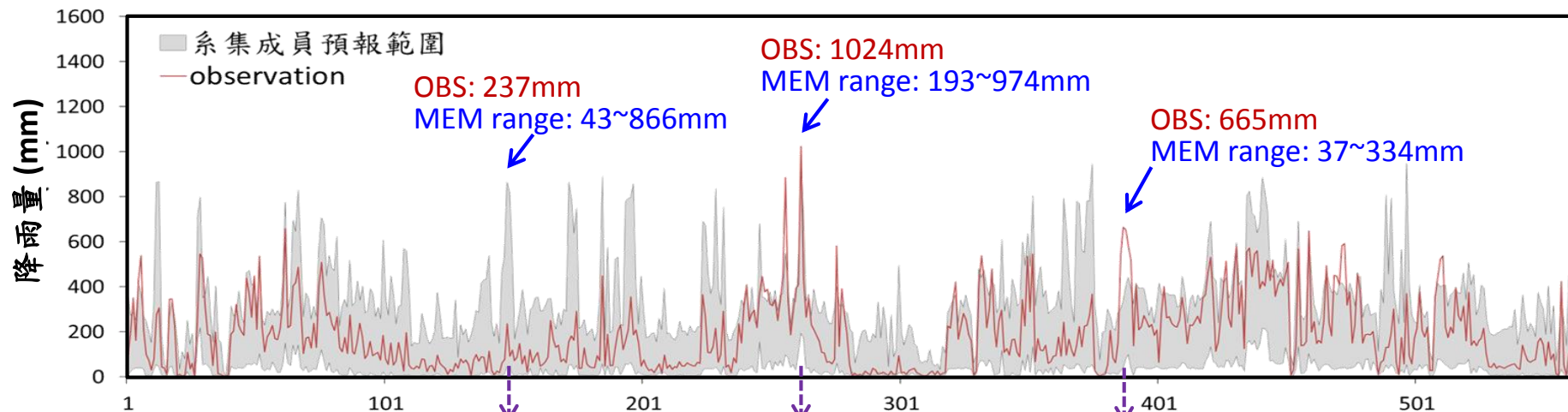
- 不同門檻值條件下，各系集成員預報值的命中率：



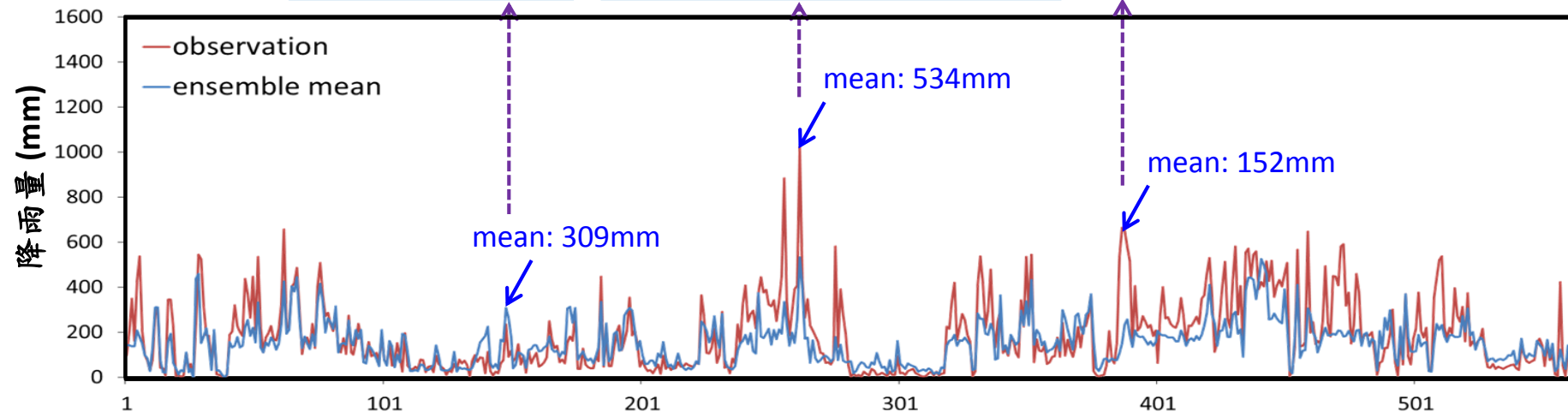
- 隨門檻值提高：命中率漸下降。
- 系集成員表現有時會優於系集平均。

2. 系集預報表現分析

- 以UTC 2012_0801_00~0802_00時段展示系集成員(MEM)的表現：



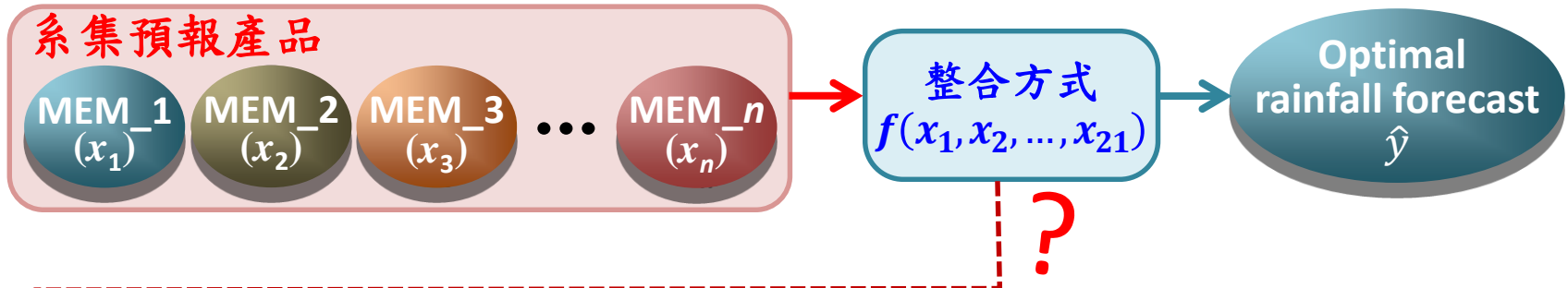
系集平均表現優 系集成員表現優於系集平均 系集成員與平均表現均不佳



測站序號

3. 整合策略

- 整合系集預報產品的目的：



策略

依據各成員的過往表現，以統計方法建置合適的整合方式。

1. 要依據哪部分的表現？

➤ 依據過去颱風事件期間的模式表現

2. 利用什麼統計方法建置整合方式？

➤ 類神經網路(artificial neural network, ANN)

3. 整合方法_類神經網路

- 類神經網路是一種模仿生物大腦神經系統之資訊處理系統，能對外界輸入訊號進行學習、儲存及回想等動作，可視為一種特殊之統計模式。
- 優點是不需事先做過多假設，透過反覆學習訓練即能處理大量的複雜資料。
- 應用類神經網路技術進行整合

a. 倒傳遞神經網路(back propagation neural network , BPN)

b. 支援向量機(support vector machine, SVM)

c. 自組織線性輸出網路(self organizing linear output, SOLO)

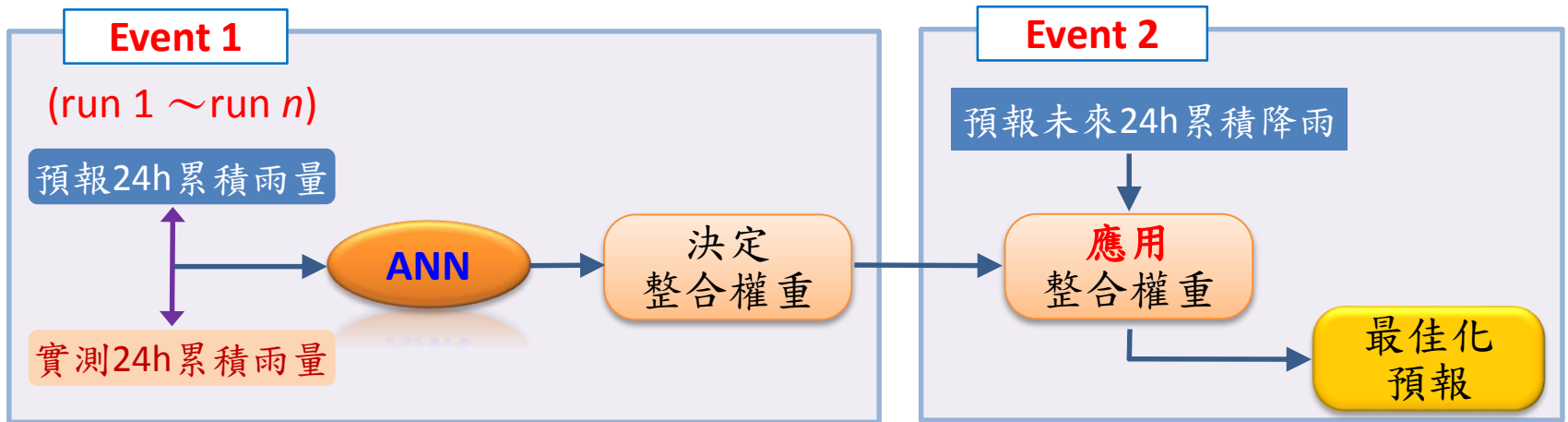
BPN與SVM均可視為非線性的回歸分析技術，但兩者採用不同機制決定回歸式的形式與參數。

SOLO為線性的回歸分析技術，採用片段線性的概念，描述非線性問題。

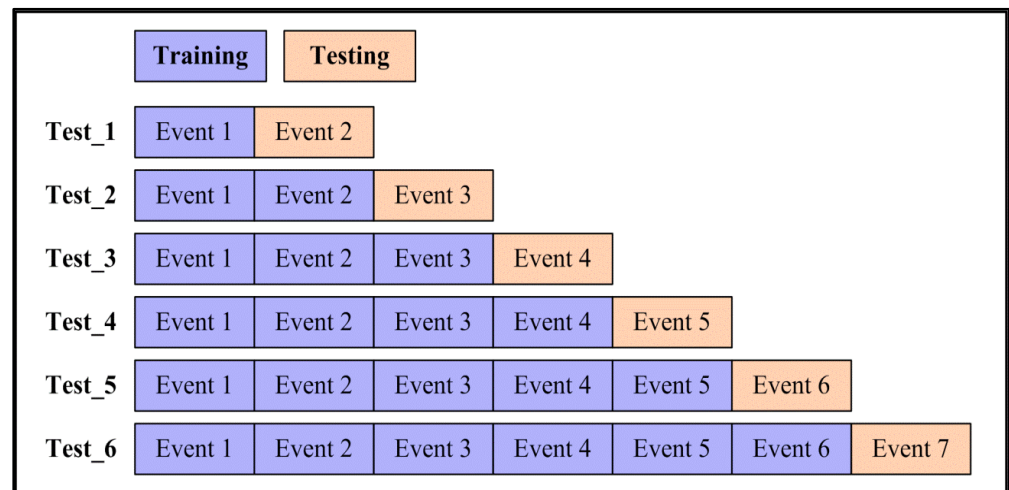


3. 試驗設計

- 「應用類神經網路(ANN)並依據過去颱風事件期間的模式表現」進行整合的試驗設計如下：

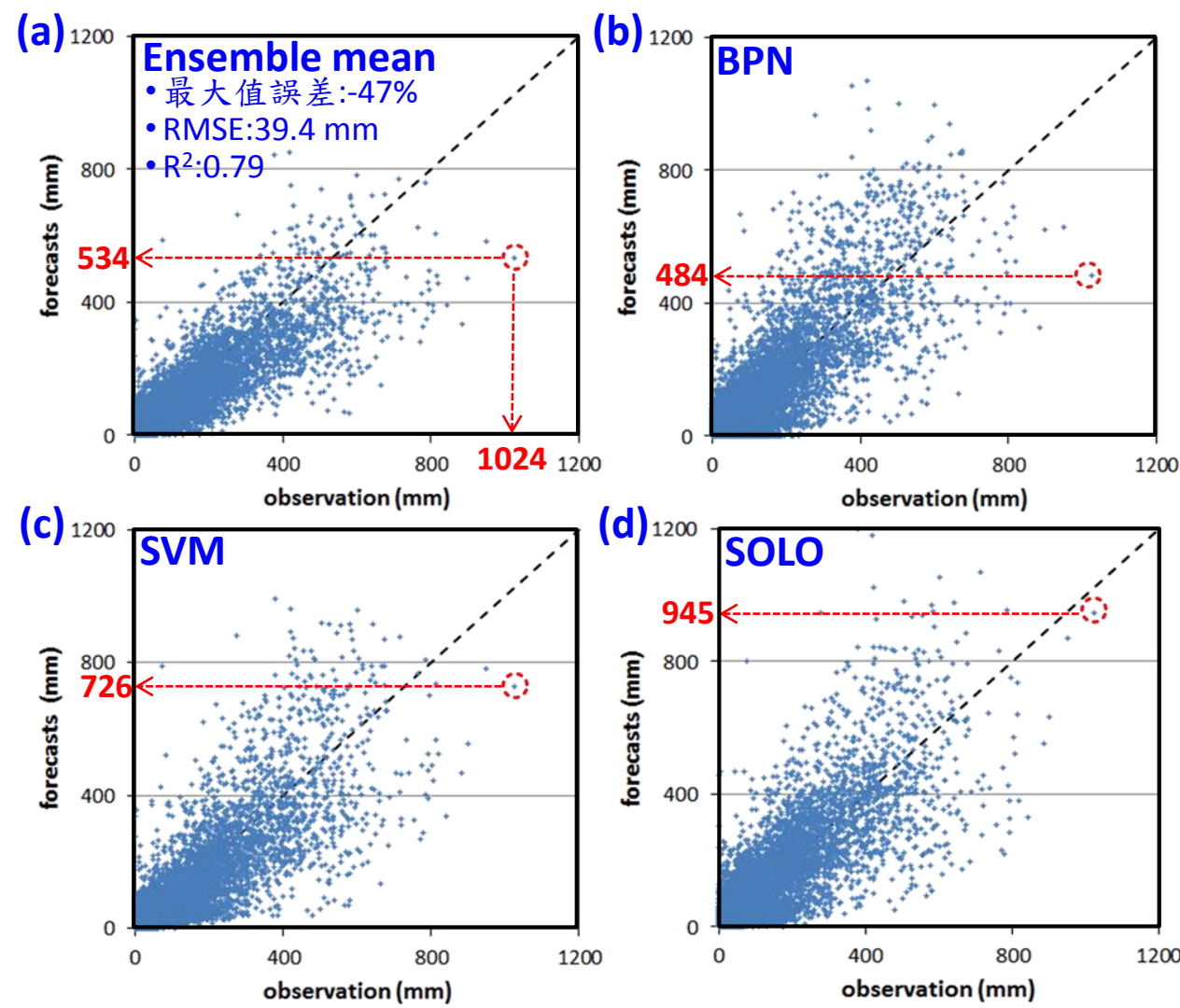


- 採用的測試方式：
 - 將過去歷史事件作為訓練，以最新一場事件作為測試。
 - 2012年7場颱風事件，共可有6個測試結果！



3. 初步試驗結果

6場颱風事件整體測試結果



除BPN外，SVM與SOLO均可降低極大值的預報誤差。

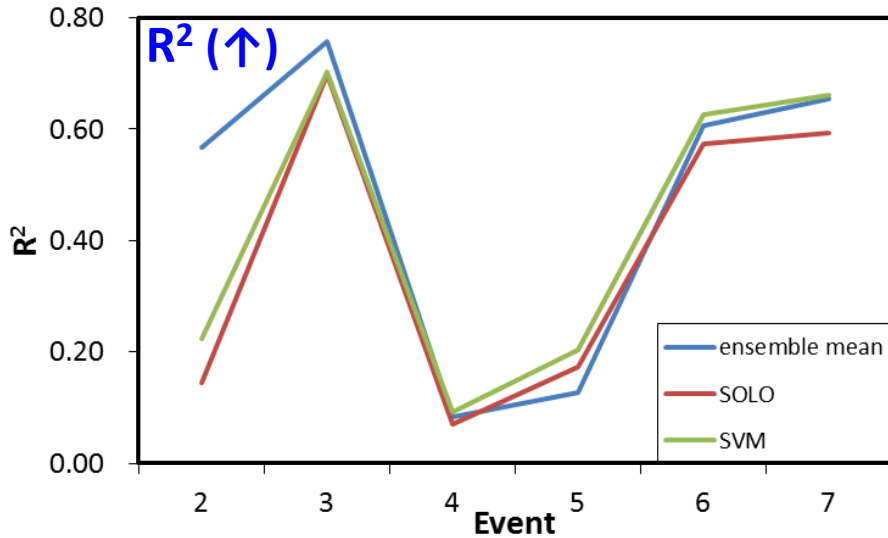
- SVM: -47% → -29%
- SOLO: -47% → -8%

結果亦顯示，整體上，整合後的預報值有較多的高估情況。

- SVM
 R^2 : 0.74
RMSE: 44.1 mm
- SOLO
 R^2 : 0.74
RMSE: 45.4 mm

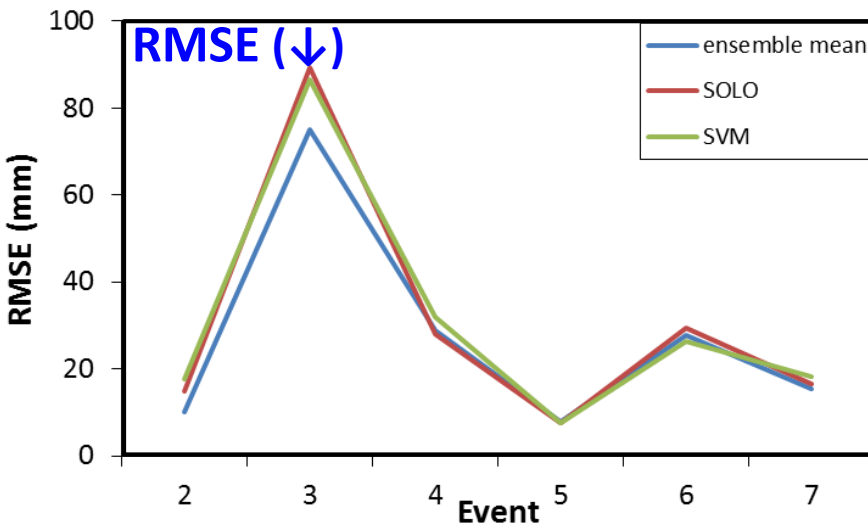
3. 初步試驗結果

6場颱風事件個別測試結果



Coefficient of determination (R^2)

	Ensemble mean	SOLO	SVM
2	0.568	0.145	0.223
3	0.758	0.697	0.702
4	0.087	0.071	0.092
5	0.127	0.173	0.205
6	0.607	0.573	0.626
7	0.654	0.594	0.662

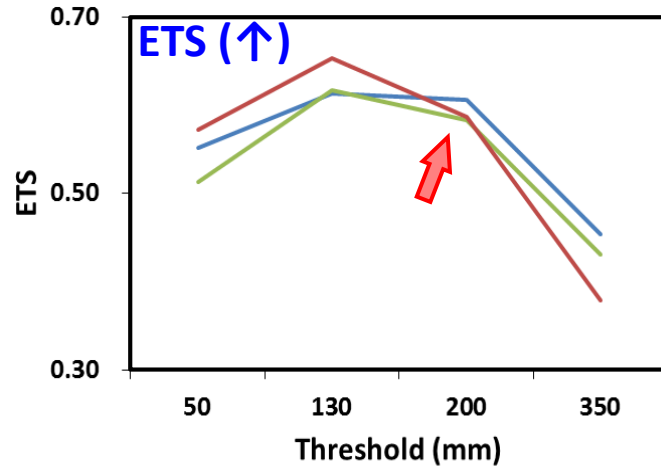
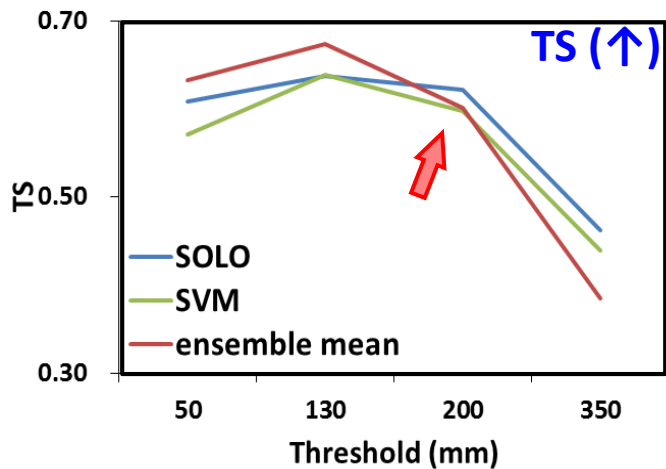


Root mean square error (RMSE)

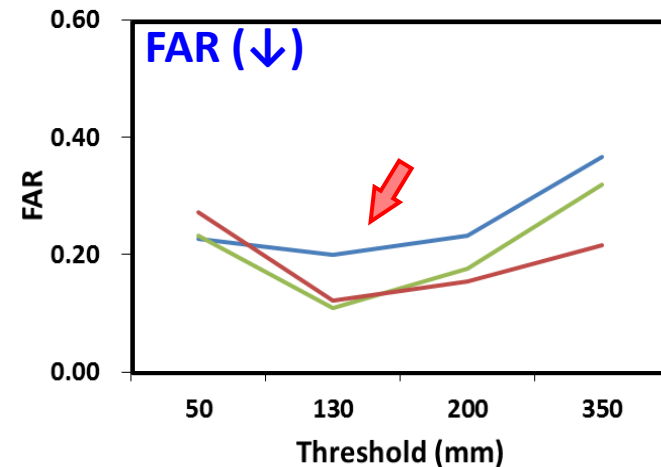
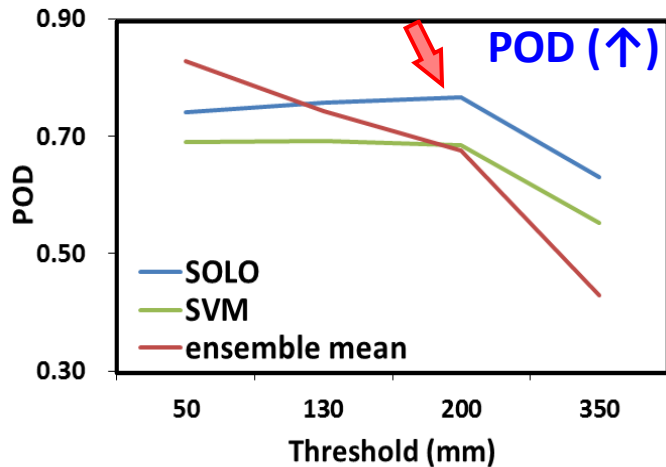
	Ensemble mean	SOLO	SVM
2	10.081	14.979	17.681
3	74.935	89.240	86.525
4	28.780	28.049	31.981
5	7.870	7.462	7.512
6	27.850	29.491	26.299
7	15.510	16.512	18.167

3. 初步試驗結果

- 以預兆得分(threat score, TS)、公正預兆得分(equitable threat score, ETS)、可偵測機率(probability of detection, POD)、誤報率(false alarm ratio, FAR)等4個技術指標量化試驗結果。



整合後，可提高大於200 mm部分之TS與ETS值。

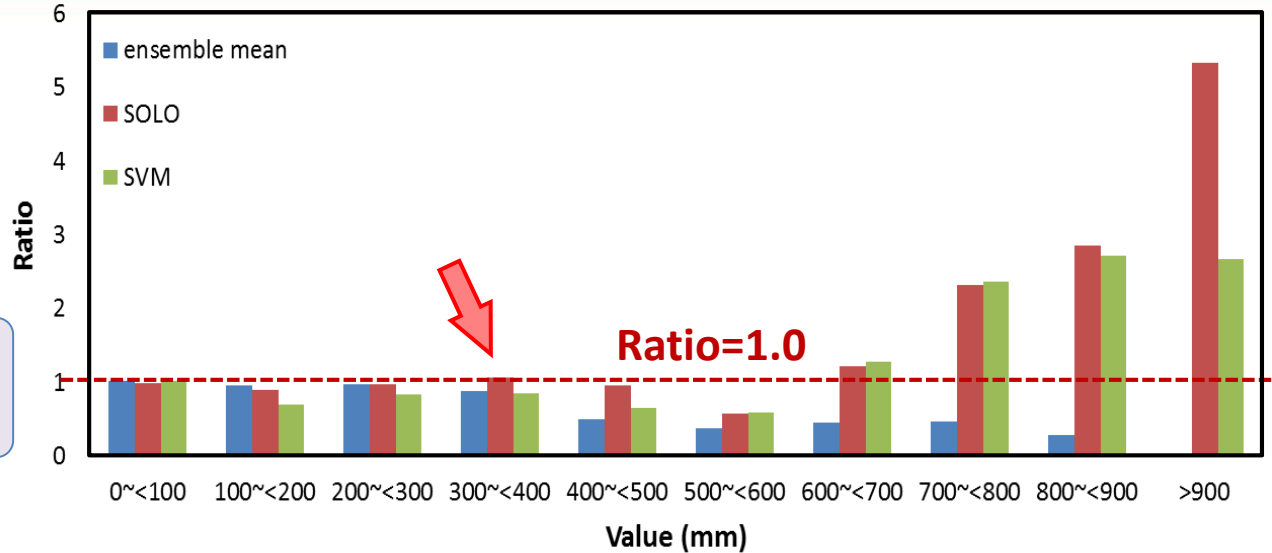


整合後，雖可提高POD值，但FAR值卻也提高了。

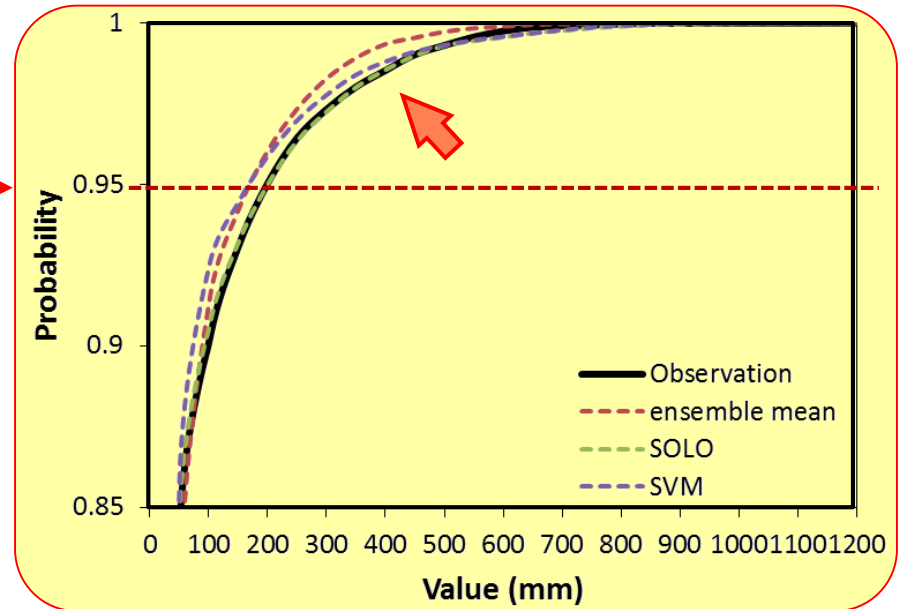
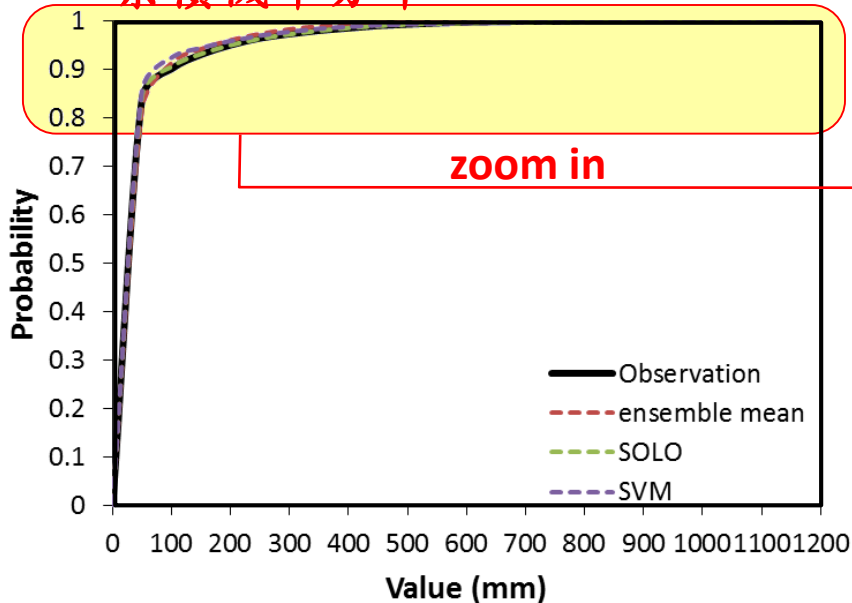
3. 初步試驗結果

- 進一步，分析整合前後預報值與實測值位於不同區間的資料數比率。

$$\text{Ratio} = \frac{\text{模式資料個數}}{\text{實測資料個數}}$$



累積機率分布



4. 結論與未來方向

• 結論

- 初步測試結果顯示，整合後確可改善極大值低估問題。
- 整合後預報值的累積機率表現更符合真實。
- 但整合後部分區域則有略微高估的現象。
- 應用類神經網路進行整合需要較大量且較具代表性的歷史事件資料，才能得到較佳的整合結果。

• 未來進行方向

- 可針對類神經網路的學習策略進行調整與改進
- 亦可配合使用氣候模式法的預估資料作為類神經網路的輸入項。期望可提升全台灣24hr累積降雨的空間分布與最大量值的預報表現。

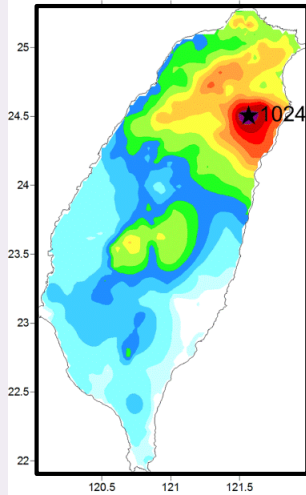
簡報結束 感謝聆聽



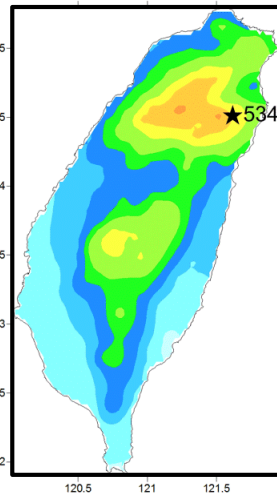
3. 初步試驗結果

2012 0801 00~0802 00 UTC 24小時累積雨量

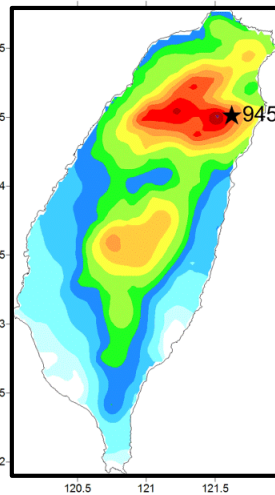
Observation



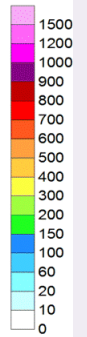
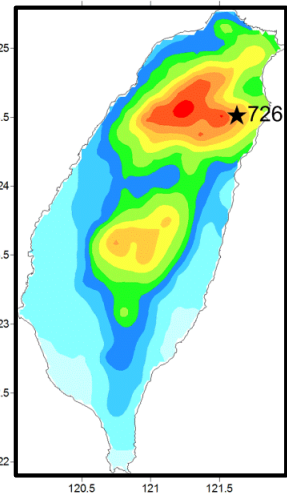
Ensemble mean SOLO



SOLO

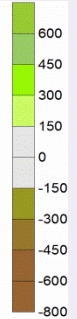
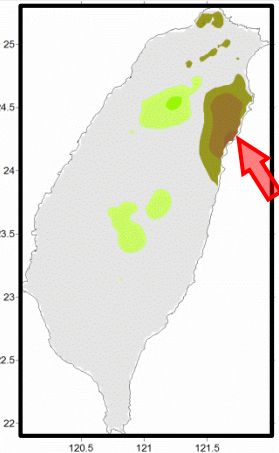
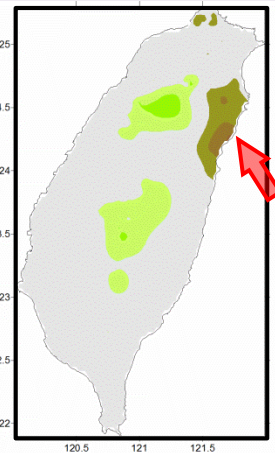
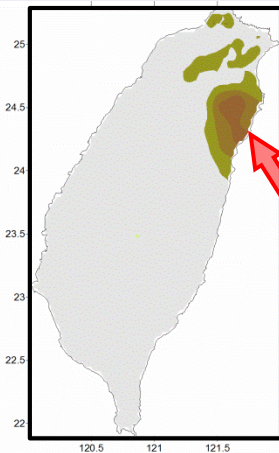


SVM



誤差空間分布

正：高估，負：低估



類神經網路(ANN)

- 應用類神經網路整合21個系集成員雨量預估結果之概念，可用數學式： $\hat{y} = f_{ANN}(x_1, x_2, \dots, x_{21})$ 表示。
- 其中， x_1, x_2, \dots, x_{21} 分別為21個系集成員各自對某一雨量測站未來24小時累積雨量的預估資料， \hat{y} 為經類神經網路整合後的預報。

◆ 倒傳遞神經網路(back propagation neural network, BPN) (Rumelhart et al., 1986)

$$\hat{y} = f\left(\sum_{j=1}^J w_j^O f\left(\sum_{i=1}^I w_{ij}^H x_i\right)\right) = \left(1 + \exp\left(-\sum_{j=1}^J w_j^O \left(1 + \exp\left(-\sum_{i=1}^I w_{ij}^H x_i\right)\right)^{-1}\right)\right)^{-1}$$

其中 x 為輸入項， w^H 與 w^O 為權重， \hat{y} 為輸出。

◆ 支援向量機(support vector machine, SVM) (Vapnik, 1995)

$$\hat{y} = \sum_{k=1}^{N_{SV}} w_k^* f(\mathbf{x}, \mathbf{x}_k) = \sum_{k=1}^{N_{SV}} w_k^* \exp\left(-\left(I\right)^{-1} \sum_{i=1}^I (x_i - x_{ki})^2\right)$$

其中 x_k 為支援向量， w^* 為權重。

◆ 自組織線性輸出網路(self organizing linear output, SOLO) (Hsu et al., 2002)

$$\hat{y} = \mathbf{w}\mathbf{x} = \begin{cases} \sum_{i=1}^I w_i^1 x_i & \text{if } \mathbf{x} \in \text{cluster 1} \\ \sum_{i=1}^I w_i^2 x_i & \text{if } \mathbf{x} \in \text{cluster 2} \\ \vdots & \\ \sum_{i=1}^I w_i^n x_i & \text{if } \mathbf{x} \in \text{cluster } n \end{cases}$$

其中 w^n 為第 n 個群集的權重。

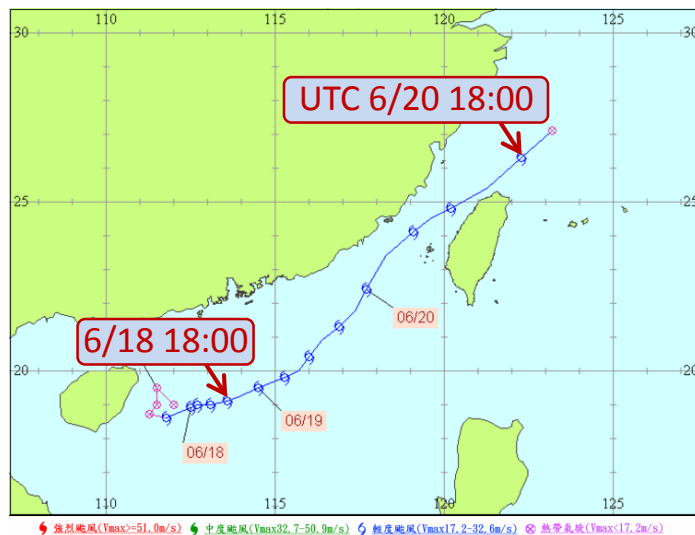


附錄資料

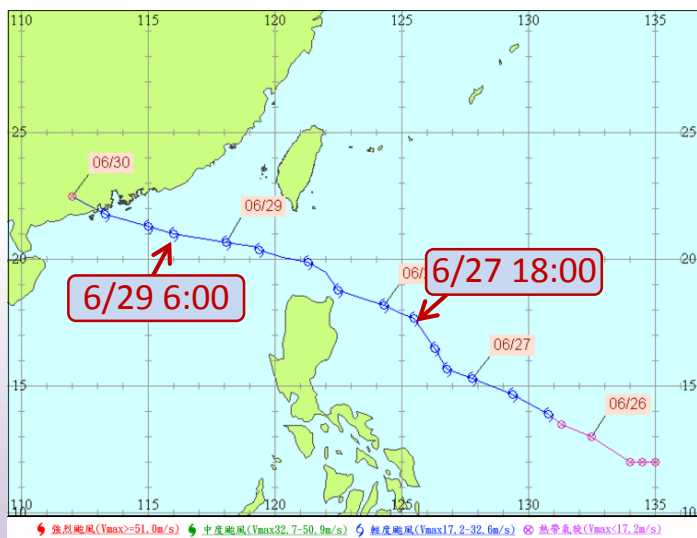
附錄：2012年7場颱風事件

中文名稱	英文名稱	警報期間 (LST)	近台強度	近台近中心最低氣壓 (hPa)	近台近中心最大風速 (m/s)	七級風暴風半徑 (km)
泰利	TALIM	06/19~06/21	輕度	985.0	25.0	150
杜蘇芮	DOKSURI	06/28~06/29	輕度	995.0	23.0	120
蘇拉	SAOLA	07/30~08/03	中度	960.0	38.0	220
海葵	HAIKUI	08/06~08/07	中度	960.0	35.0	180
啟德	KAI-TAK	08/14~08/15	輕度	995.0	20.0	150
天秤	TEMBIN	08/26~08/28	中度	965.0	35.0	180
	TEMBIN	08/21~08/25	中度	945.0	45.0	180
杰拉華	JELAWAT	09/27~09/28	強烈	910.0	55.0	250

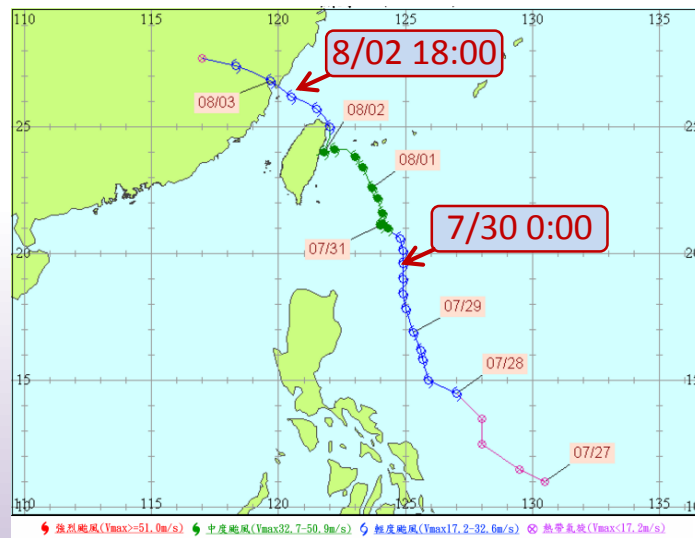
1. 泰利 (Talim 0618-18~0620-18, 共9個DTG)



2. 杜蘇芮 (Doksuri 0627-18~0629-06, 共7個DTG)

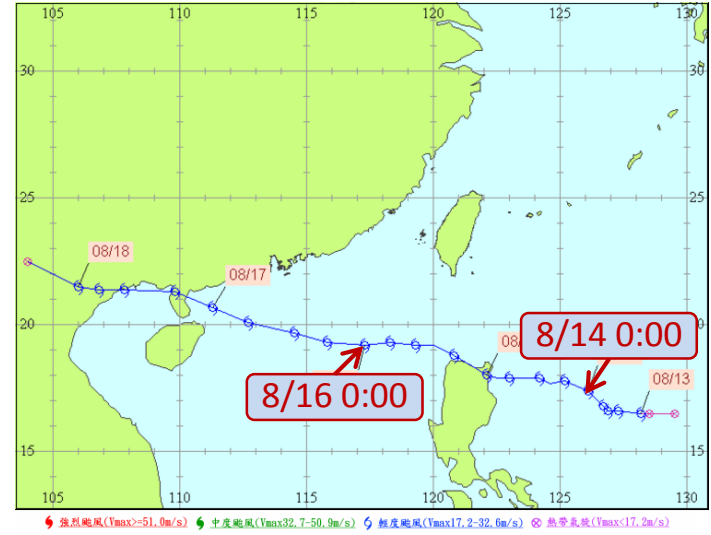
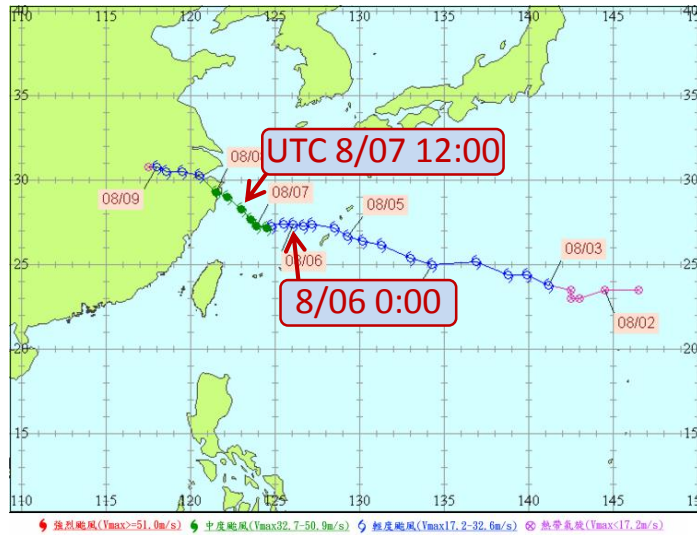


3. 蘇拉 (Saola 0730-00~0802-18 共16個DTG)

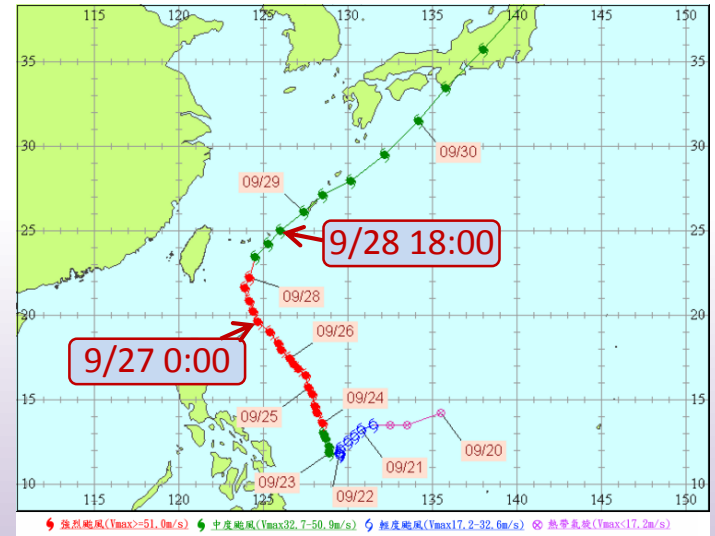
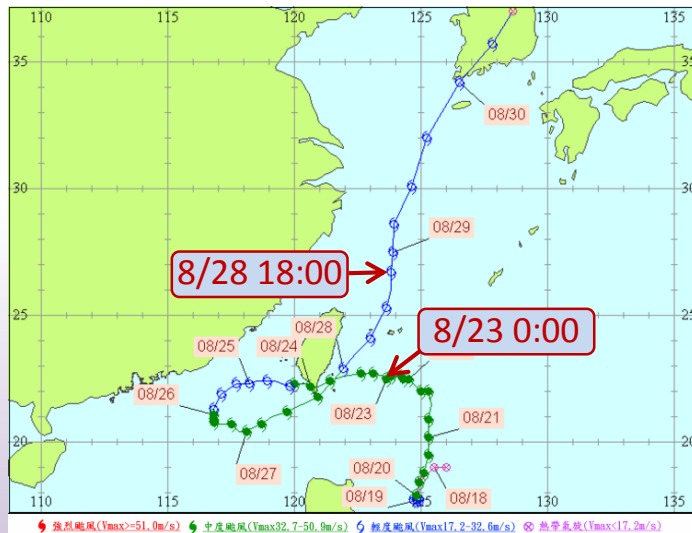


附錄：2012年度7場颱風事件

4.海葵(Haikui 0806-00~0807-12, 共7個DTG) 5.啟德(Kai-Tak 0814-00~0816-00, 共9個DTG)



6.天秤(Tembin 0823-00~0828-18, 共24個DTG) 7.杰拉華(Jelawat 0927-00~0928-18, 共8個DTG)



• 類神經網路設定

a. 倒傳遞神經網路(back propagation neural network, BPN)



- 使用3層架構的BPN(1 input layer, 1 hidden layer, and 1 output layer)，10個隱藏層神經元(hidden layer neuron)，學習速率(learning rate)為0.8，最大學習次數(learning epoch)為5000。

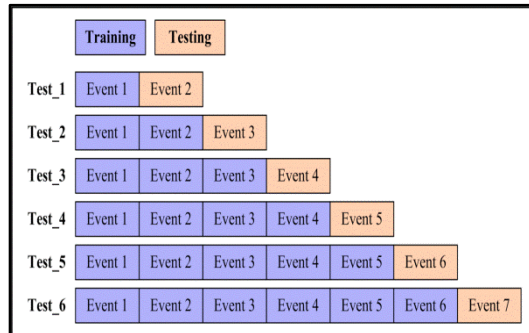
b. 支援向量機(support vector machine, SVM)

- 使用的成本係數(cost function)為1，誤差容忍度(tolerance)為1%。

c. 自組織線性輸出網路(self organizing linear output, SOLO)

- 採用2*2的拓樸(topology)結構。

- 使用類神經網路進行整合需仰賴大量的系集實驗預報結果，藉由調整各系集成員的組合權重將雨量預報穩定化，而非改進個別成員之預報表現。



名稱	警報期間 (LST)	資料使用時段 (UTC)
泰利 (TALIM)	06/19~06/21	6/18 18:00~6/20 18:00
杜蘇芮 (DOKSURI)	06/28~06/29	6/18 18:00~6/20 18:00
蘇拉 (SAOLA)	07/30~08/03	6/18 18:00~6/20 18:00
海葵 (HAIKUI)	08/06~08/07	6/18 18:00~6/20 18:00
啟德 (KAI-TAK)	08/14~08/15	6/18 18:00~6/20 18:00
天秤 (TEMBIN)	08/21~08/28	6/18 18:00~6/20 18:00
杰拉華 (JELAWAT)	09/27~09/28	6/18 18:00~6/20 18:00

中文名稱	英文名稱	警報期間 (LST)	近台強度
泰利	TALIM	06/19~06/21	輕度
杜蘇芮	DOKSURI	06/28~06/29	輕度
蘇拉	SAOLA	07/30~08/03	中度
海葵	HAIKUI	08/06~08/07	中度
啟德	KAI-TAK	08/14~08/15	輕度
天秤	TEMBIN	08/21~08/28	中度
杰拉華	JELAWAT	09/27~09/28	強烈