

基於集成學習改善卷積神經網路之 系集定量降水預報後處理方法

Improving Quantitative Precipitation Forecast Post-processing Method of Convolutional Neural Networks Based on Ensemble Learning

作者：葉致源、林冠成、張保亮

報告人：賴昱丞

日期：2024/09/04



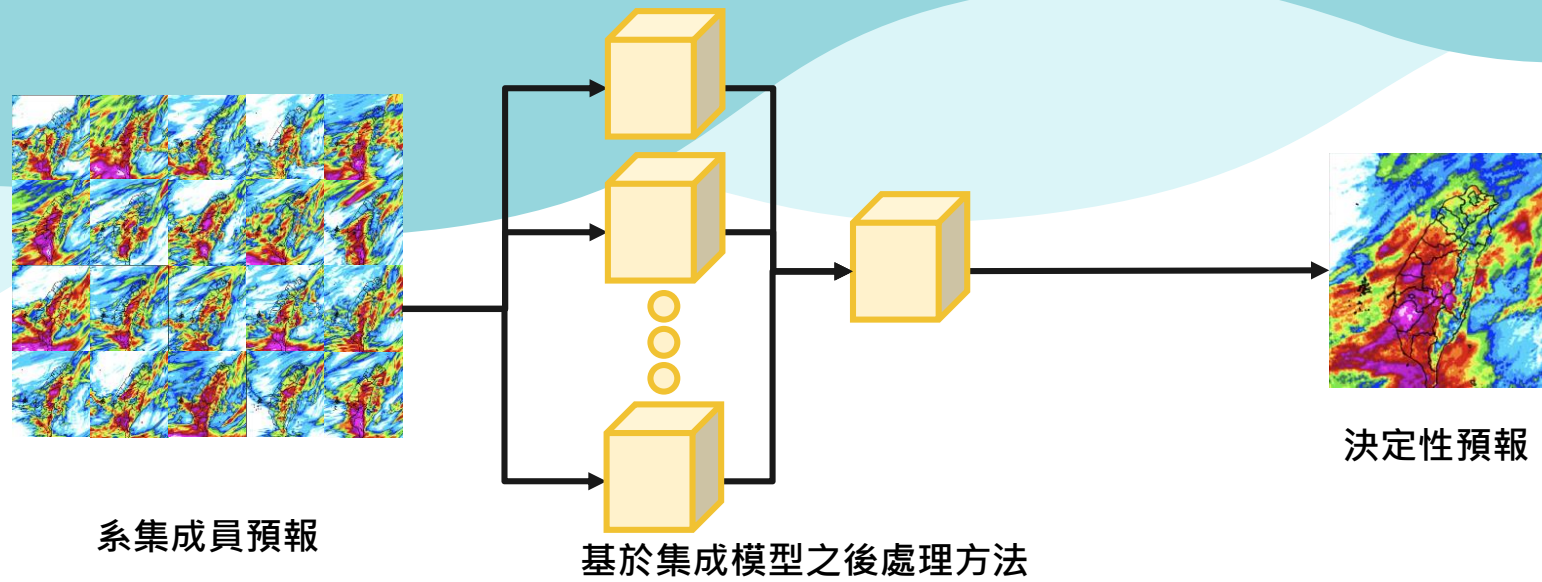
基於集成學習改善卷積神經網路之 系集定量降水預報後處理方法

Improving Quantitative Precipitation Forecast Post-processing Method of
Convolutional Neural Networks Based on Ensemble Learning

作者：葉致源、林冠成、張保亮

報告人：賴昱丞

日期：2024/09/04



基於集成學習改善卷積神經網路之 系集定量降水預報後處理方法

Improving Quantitative Precipitation Forecast Post-processing Method of
Convolutional Neural Networks Based on Ensemble Learning

作者：葉致源、林冠成、張保亮

報告人：賴昱丞

日期：2024/09/04

大綱

- 1 背景、動機與目的
- 2 研究方法
- 3 實驗結果與討論
- 4 結論與未來方向

1 背景、動機與目的

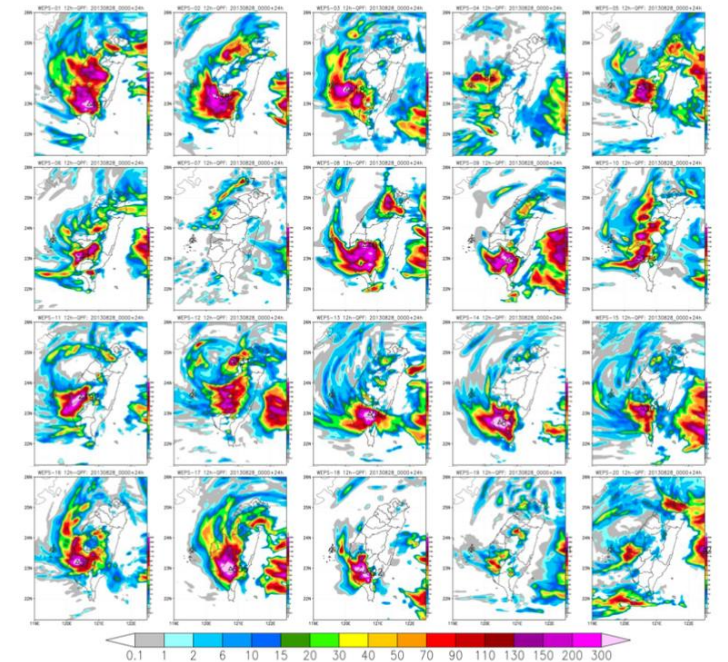
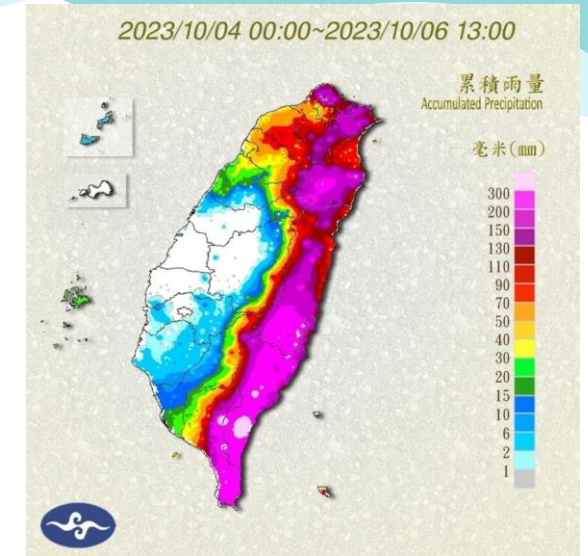
背景介紹-定量降水預報 (QPF)

• 定義

- 在一段時間內，於特定地點的累積總雨量
- 採用數值天氣預報 (NWP) 模式

• 問題

- 初始場不完美、大氣變化速度快 → 具有很高的不確定性
- 無法在有限的時間內提升預報能力

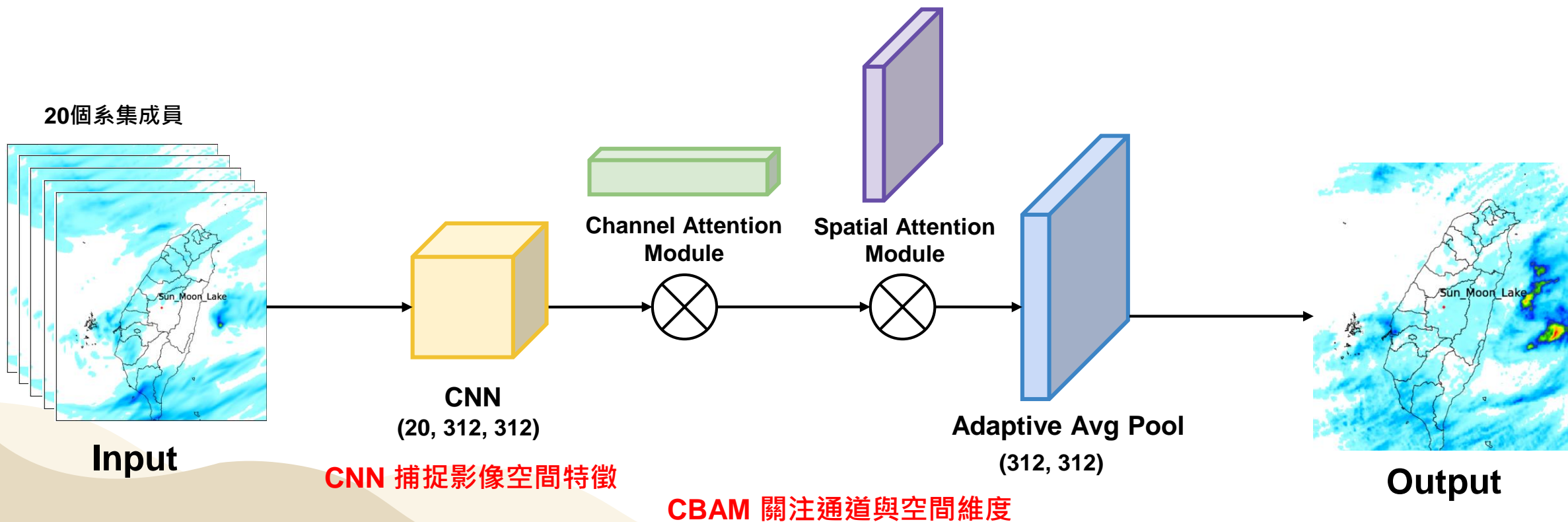


背景介紹-系集降水預報後處理方法

類型	非固定參數法			深度學習
方法	平均法 (Mean)	機率擬合平均 (PM)	新機率擬合平均 (NPM)	卷積神經網路 CNN
優點	計算簡單 較佳的空間相對分布	改善平均法大雨低報	改善大雨高報問題	改善降雨空間分佈 、累積降雨量
缺點	大雨或極端降雨低報	未考慮到成員間的差異性 大雨存在過度預報的問題	由於降雨空間分佈 取決於平均法 導致降雨分佈不準確	降雨空間分佈仍有待修正 降雨量取決於系集成員

背景介紹-運用深度學習於系集後處理方法

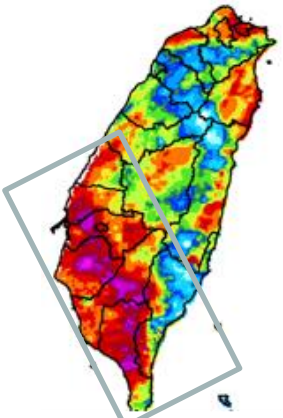
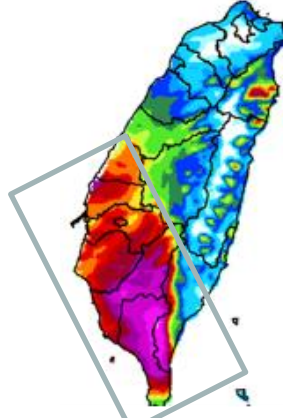
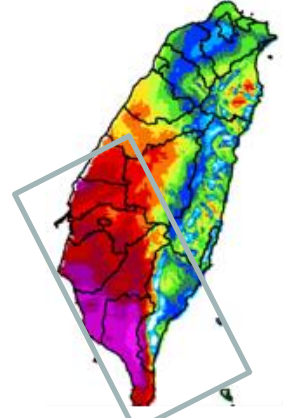
- 陳宛婷, 2022, 卷積神經網路結合空間注意力機制
 - 加入前 6 小時觀測值作為偏差修正

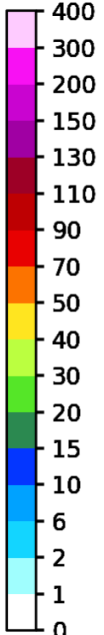


研究動機-預測結果仍有誤差

降雨空間分佈與降雨量預報結果仍有修正空間

2021年6月21日 梅雨個案

	OBS (實際觀測值)	NPM (新機率擬合平均法)	CNN+CBAM 遷移學習
RMSE	-	44.3930	39.6838
SSIM	-	0.3615	0.4101
可視化圖			



A vertical color scale legend for rainfall intensity, ranging from 0 to 400. The scale uses a rainbow color gradient: 0 (white), 1 (light blue), 2 (blue), 6 (cyan), 10 (green), 15 (light green), 20 (green), 30 (yellow-green), 40 (yellow), 50 (orange), 70 (red-orange), 90 (red), 110 (dark red), 130 (magenta), 150 (purple), 200 (dark purple), 300 (black), and 400 (white).

研究動機-使用平均絕對損失作為損失函數

- Mingheng Jiang 等人提出

- 以Overall損失修正整體降雨強度

$$QWMSE\ Loss = Loss_{overall} + Loss_{torrential}$$

- 類似於平均絕對損失 (MAE)

- 於不同閾值下計算Torrential損失

- 類似於平均平方損失 (MSE) 提供模型更多修正資訊

- 實驗顯示引入Overall後可有效**提高預報準確度**

研究動機-以集成學習改善預報準確度

- Jonathan A. Weyn等人提出
 - 結合多種大型天氣模型結果進行平均集成
 - Pangu-Weather、GraphCast、MS-Nowcasting、FourCastNet 等
- 實驗顯示透過多模型集成方式可有效提升整體預報的多樣性
 - 透過集成平均**減少預報誤差**

研究目的

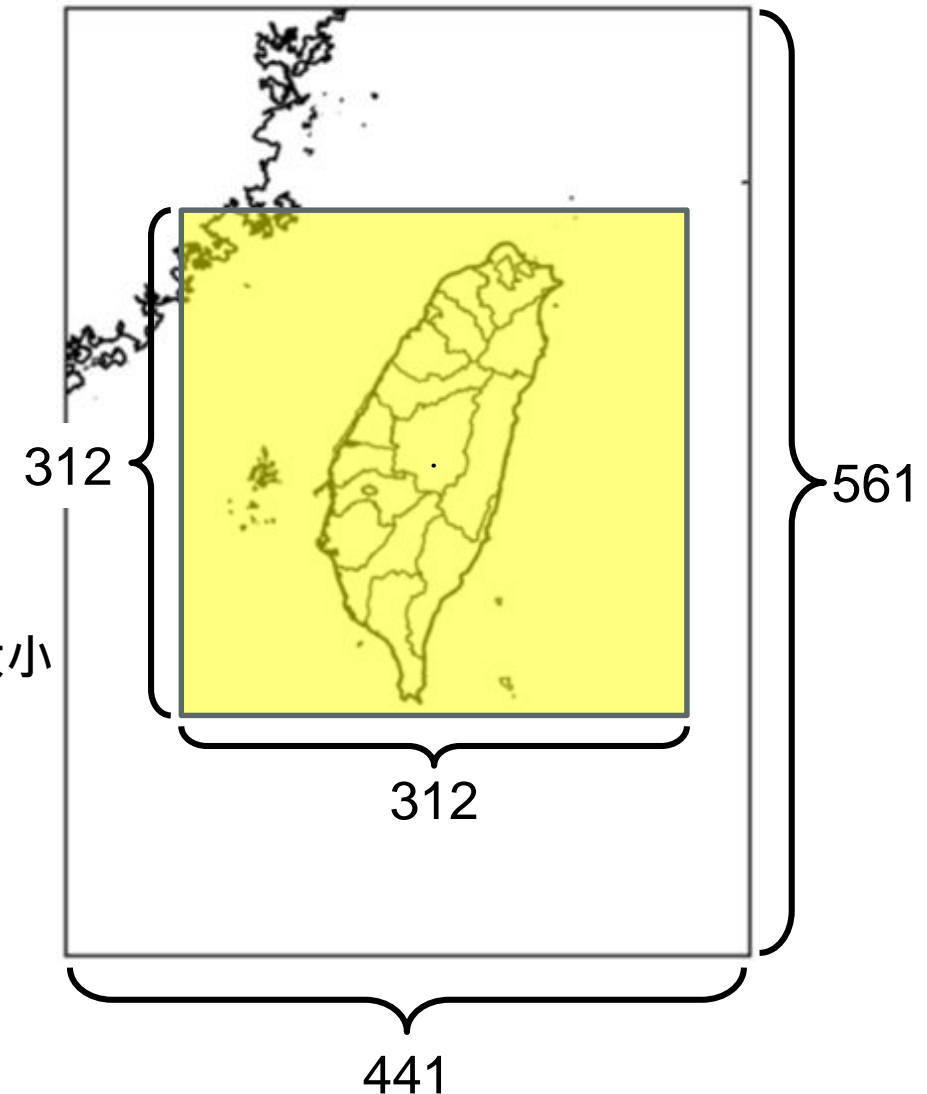
- 改進模型訓練品質
 - 不同加權損失函數優化模型訓練
 - 使用集成學習整合多模型優勢
- 改善降雨空間分佈
 - 加入MAE作為損失函數，改善模型整體降雨量高報問題
 - 驗證前6小時觀測值作為校正，是否有效改善預報表現



2 研究方法

實驗資料

- 來源：中央氣象署-系集預報系統 (WEPS)
- 資料集
 - 訓練：2018 ~ 2021年5~8月、2023年5、6月
 - 驗證：2023年7月
 - 測試：2023年8月
- 網格大小：561 x 441
- 以日月潭測站為中心點，往南移動 0.18° 劃設 312×312 的大小
- 網格解析度：1.3 公里 x 1.3 公里
 - 約等於 $0.0125^\circ \times 0.0125^\circ$
- 預報時段：第 6-30 小時



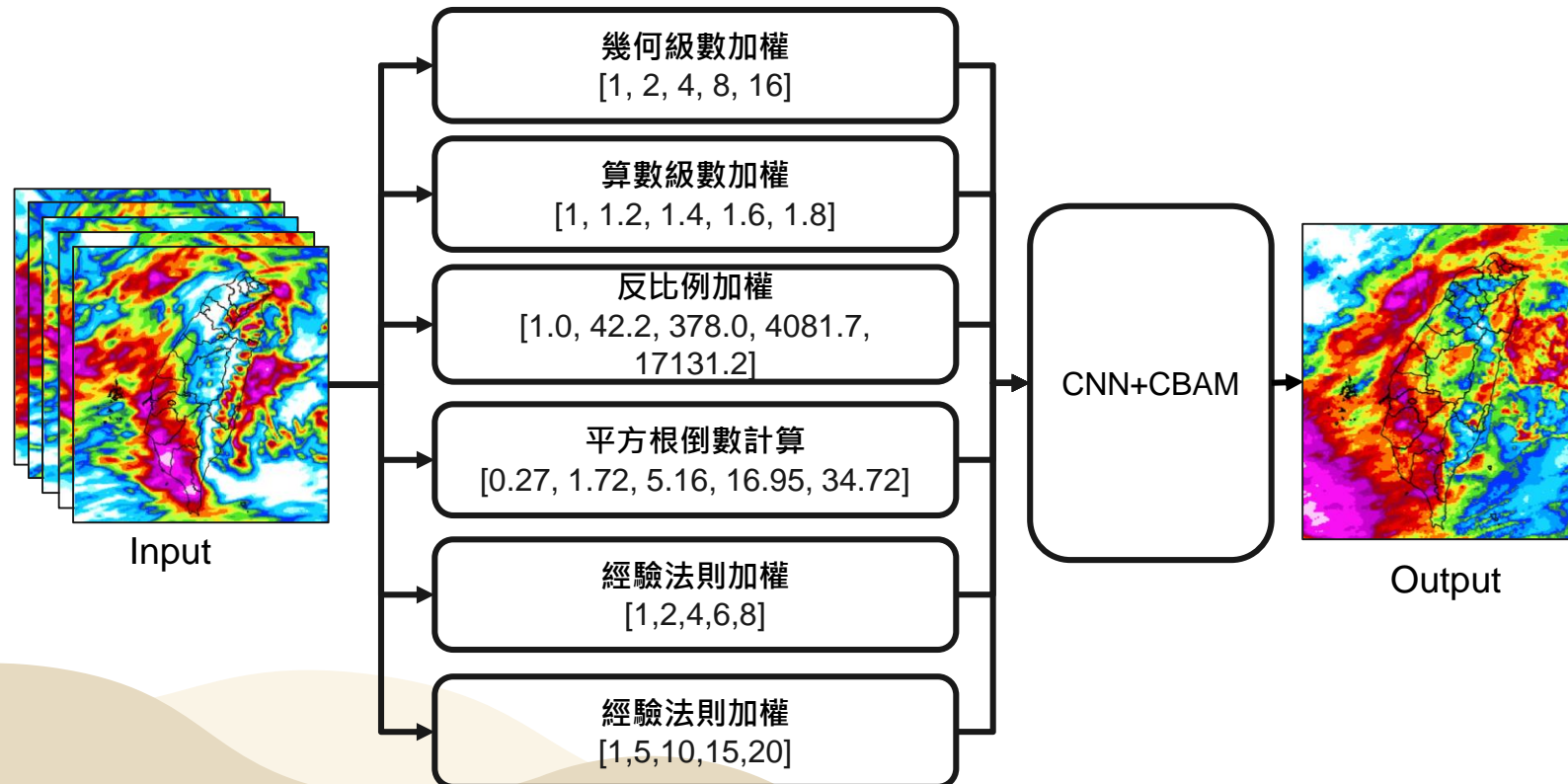
實驗一、比較加權損失函數權重

- 依照中央氣象署定義的累積降雨區間
- 加權方式
 1. 幾何級數加權
 2. 算數級數加權
 3. 反比例加權
 4. 平方根倒數加權
 5. 經驗法則加權

累積降雨區間	2018至2021年5至8月、 2023年5、6月 實際觀測值所有格點	資料佔比
$0 \leq x < 80$	1.98311922e8	97.34%
$80 \leq x < 200$	4.828193e6	2.37%
$200 \leq x < 350$	5.39068e5	0.26%
$350 \leq x < 500$	4.9916e4	0.02%
$500 \leq x$	1.1893e4	0.01%

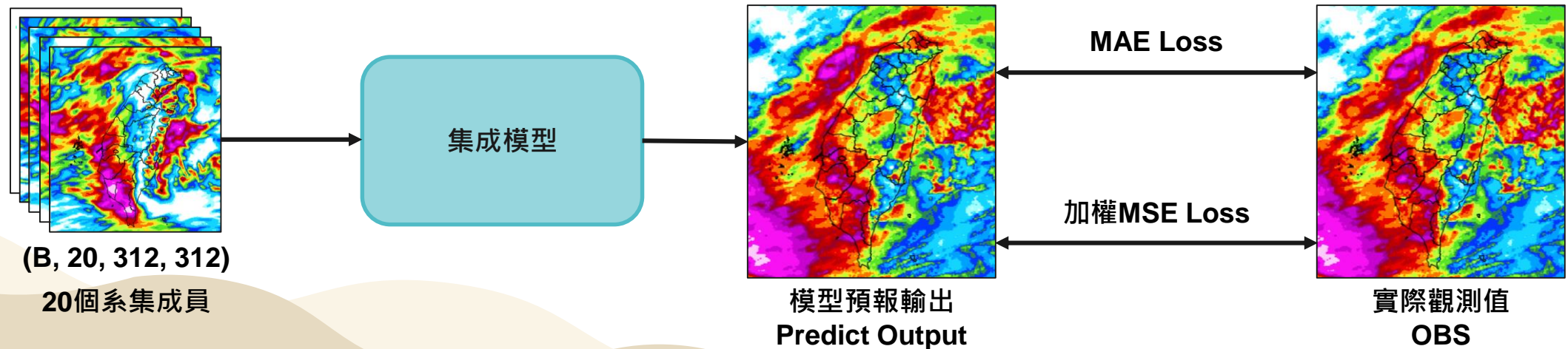
實驗二、結合集成學習改善降雨空間分佈預測

- 不同的權重值可能適用於不同個案
 - 結合各模型之間的預報結果進行集成學習預測
 - 分別取實驗一之預訓練模型



實驗三、結合MAE作為損失函數

- 降低整體累積降雨量的預報誤差
 - 避免小雨高報導致預報準確度下降
 - 參考Mingheng Jiang (2024) 等人的研究
 - $Total Loss = MAE Loss + Weighted MSE Loss$



評估指標

1. 均方根誤差 (RMSE)
 - 回歸演算法的評估指標之一
 - 越小表示預報準確度越高
2. 結構相似性指標 (SSIM)
 - 值介於0~1之間
 - 越大表示，兩張圖像越相近
3. 連續排名機率係數 (Continuous Ranked Probability Score, CRPS)
 - 氣象上常用於評估預報結果與真實值的分佈差異
 - 值越小預報越精準



3 實驗結果與分析

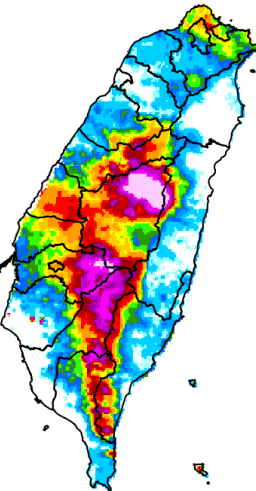
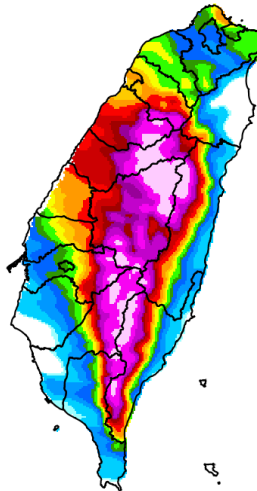
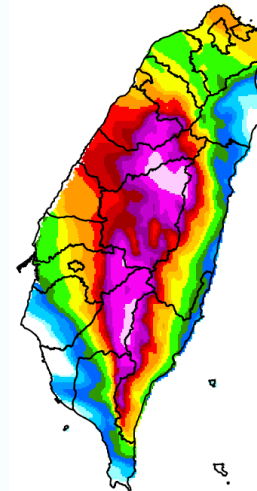
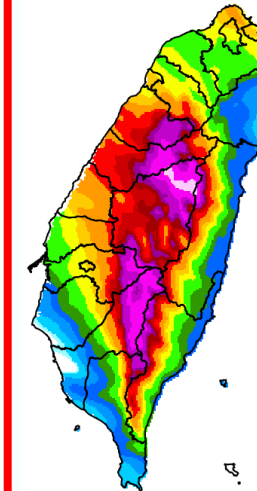
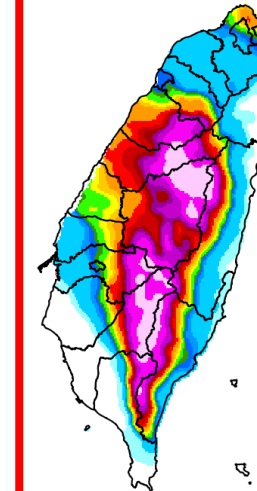
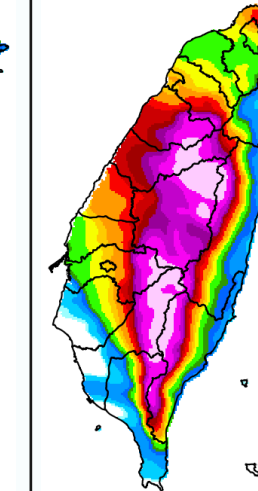
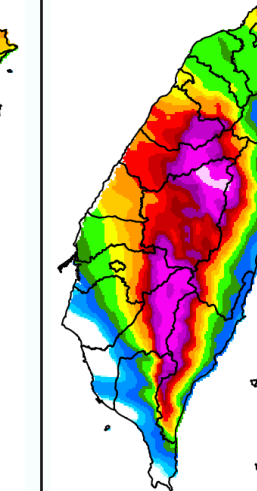
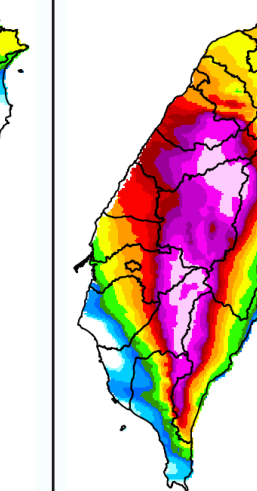
實驗一、比較加權損失函數權重 – 未加入偏差修正 – 2023年整月平均

不加入偏差修正時，算數級數加權的整體平均結果最佳
其餘加權方式誤差太大，導致平均結果不佳

	統計方法	幾何級數加權	算數級數加權	反比例加權	平方根倒數計算	經驗加權法	
OBS	NPM	[1, 2, 4, 8, 16]	[1, 1.2, 1.4, 1.6, 1.8]	[1.0, 42.2, 378.0, 4081.7, 17131.2]	[0.27, 1.72, 5.16, 16.95, 34.72]	[1,2,4,6,8]	[1,5,10,15,20]
RMSE	27.4596	22.3221	21.4319	23.8447	24.1602	21.6936	24.5580
SSIM	0.6019	0.6192	0.6025	0.6378	0.6197	0.6280	0.6040
CRPS	9.8525	8.9011	8.8991	8.5590	9.2541	8.4590	9.9740

實驗一、比較加權損失函數權重 – 未加入偏差修正 – 卡努颱風

算數級數加權法整體降雨量最接近OBS，同時避免了大雨過報的問題

	統計方法	幾何級數加權	算數級數加權	反比例加權	平方根倒數	經驗加權法	
OBS	NPM	[1, 2, 4, 8, 16]	[1, 1.2, 1.4, 1.6, 1.8]	[1.0, 42.2, 378.0, 4081.7, 17131.2]	[0.27, 1.72, 5.16, 16.95, 34.72]	[1,2,4,6,8]	[1,5,10,15,20]
							
RMSE	69.5745	48.7548	39.9516	56.1688	66.4801	43.1156	66.0689
SSIM	0.4136	0.4504	0.4616	0.4986	0.4482	0.4462	0.4166
CRPS	22.9805	19.9652	18.1827	20.1719	23.2212	18.7298	24.6276

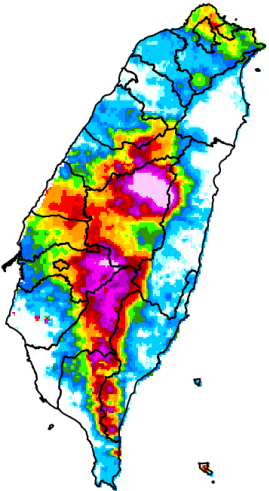
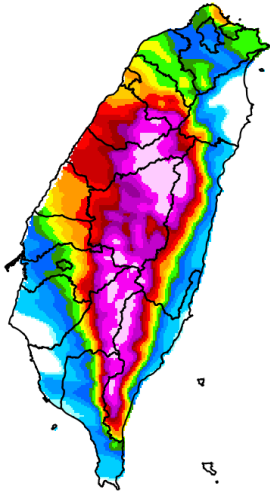
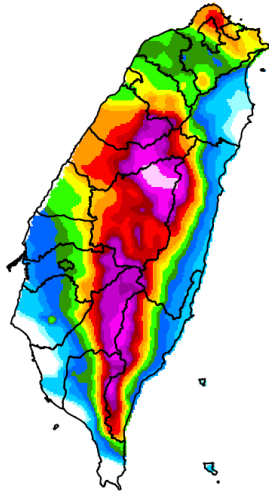
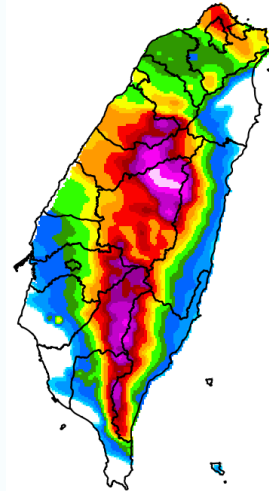
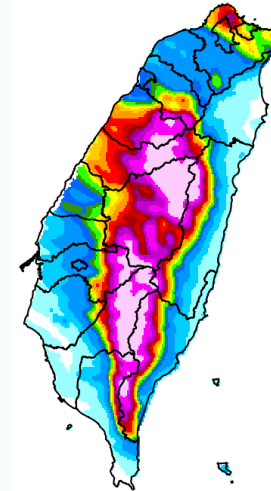
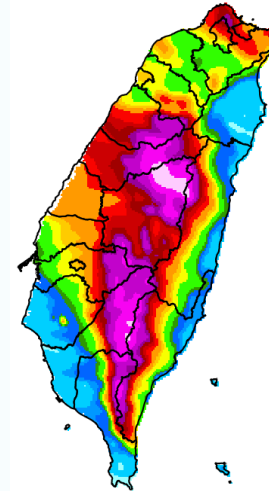
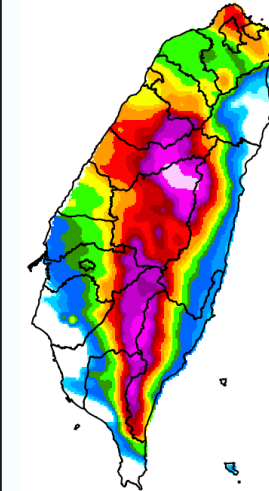
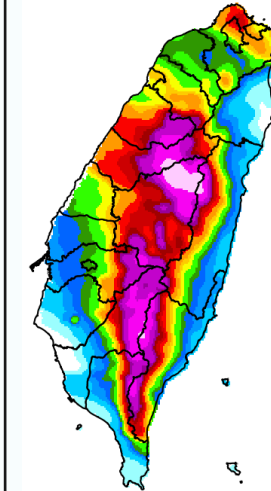
實驗一、比較加權損失函數權重 – 加入偏差修正 – 2023年整月平均整

在加入偏差修正進行偏差修正後，算數級數仍有綜合最低的誤差表現

	統計方法	幾何級數加權	算數級數加權	反比例加權	平方根倒數計算	經驗加權法	
OBS	NPM	[1, 2, 4, 8, 16]	[1, 1.2, 1.4, 1.6, 1.8]	[1.0, 42.2, 378.0, 4081.7, 17131.2]	[0.27, 1.72, 5.16, 16.95, 34.72]	[1,2,4,6,8]	[1,5,10,15,20]
RMSE	27.4596	18.7316	18.1630	23.0001	20.9353	19.1637	19.3707
SSIM	0.6019	0.6836	0.6843	0.6816	0.5391	0.6722	0.6755
CRPS	9.8525	6.9568	6.9713	7.8948	9.4621	7.3553	7.2439

實驗一、比較加權損失函數權重 – 加入偏差修正 – 卡努颱風

加入偏差修正，更能使算數級數加權正確的反應降雨空間分佈

	統計方法	幾何級數加權	算數級數加權	反比例加權	平方根倒數	經驗加權法	
OBS	NPM	[1, 2, 4, 8, 16]	[1, 1.2, 1.4, 1.6, 1.8]	[1.0, 42.2, 378.0, 4081.7, 17131.2]	[0.27, 1.72, 5.16, 16.95, 34.72]	[1,2,4,6,8]	[1,5,10,15,20]
							
RMSE	69.5745	39.2836	35.2615	73.9281	44.2268	40.6855	43.0483
SSIM	0.4136	0.4946	0.4819	0.5212	0.4499	0.4580	0.5024
CRPS	22.9805	16.4414	15.7062	21.8135	18.7312	17.0183	17.0937

實驗一、比較加權損失函數權重 - 小結

- 算數級數加權整體誤差最低
 - 不論加入或未加入偏差修正
 - 整體降雨空間分佈最接近OBS
- 不同個案下可能需要使用不同的權重進行預報

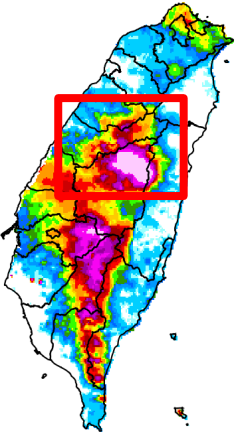
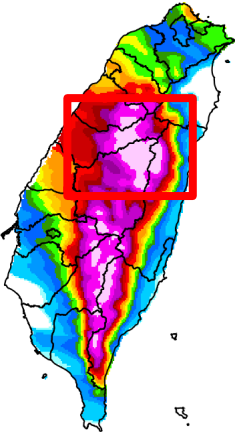
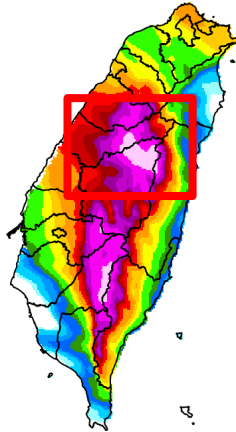
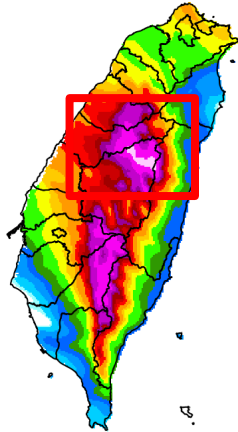
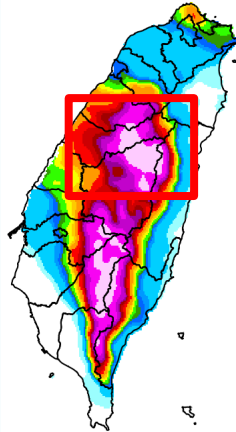
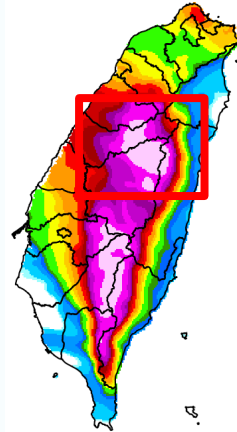
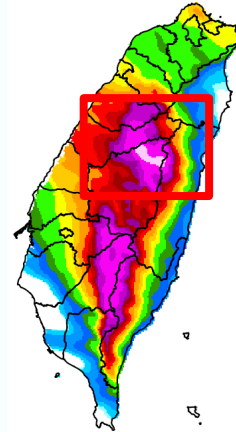
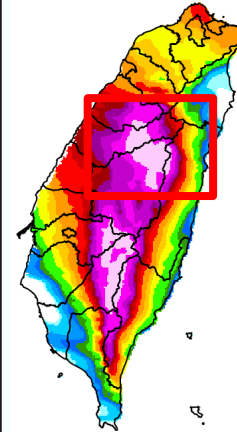
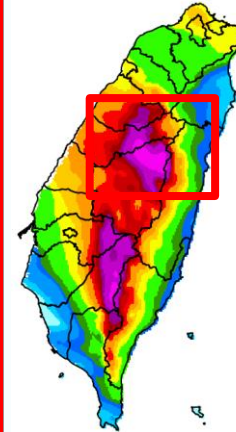
實驗二、集成不同權重 – 不加入偏差修正 – 2023年整月平均

集成不同權重後，整月平均RMSE略優於單一權重值
SSIM與CRPS則不如算數級數加權

	統計方法	幾何級數加權	算數級數加權	反比例加權	平方根倒數	經驗加權法		集成模型
OBS	NPM	[1, 2, 4, 8, 16]	[1, 1.2, 1.4, 1.6, 1.8]	[1.0, 42.2, 378.0, 4081.7, 17131.2]	[0.27, 1.72, 5.16, 16.95, 34.72]	[1,2,4,6,8]	[1,5,10,15,20]	
RMSE	27.4596	22.3221	21.4319	23.8447	24.1602	21.6936	24.5580	21.3653
SSIM	0.6019	0.6192	0.6025	0.6378	0.6197	0.6280	0.6040	0.5630
CRPS	9.8525	8.9011	8.8991	8.5590	9.2541	8.4590	9.9740	9.2802

實驗二、集成不同權重 – 不加入偏差修正 – 卡努颱風

集成預報誤差進一步降低，整體降雨空間更接近實際觀測值
但同時也帶來大雨些微低報問題

	統計方法	幾何級數加權	算數級數加權	反比例加權	平方根倒數	經驗加權法		集成模型	
OBS	NPM	[1, 2, 4, 8, 16]	[1, 1.2, 1.4, 1.6, 1.8]	[1.0, 42.2, 378.0, 4081.7, 17131.2]	[0.27, 1.72, 5.16, 16.95, 34.72]	[1,2,4,6,8]	[1,5,10,15,20]		
									
RMSE	69.5745	48.7548	39.9516	56.1688	66.4801	43.1156	66.0689	36.0905	
SSIM	0.4136	0.4504	0.4616	0.4986	0.4482	0.4462	0.4166	0.4742	
CRPS	22.9805	19.9652	18.1827	20.1719	23.2212	18.7298	24.6276	17.5362	

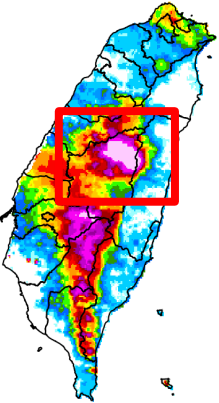
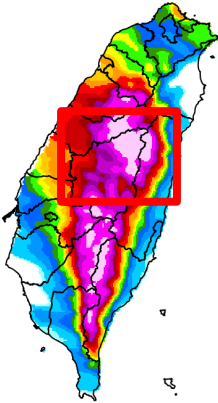
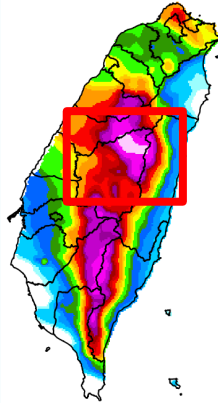
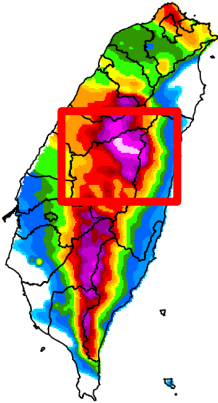
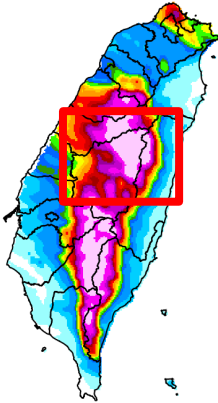
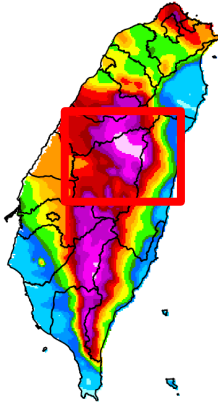
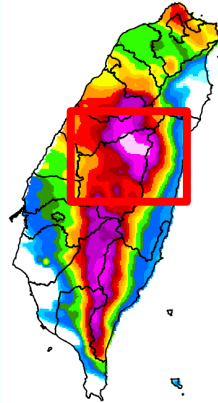
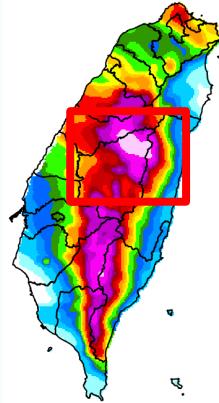
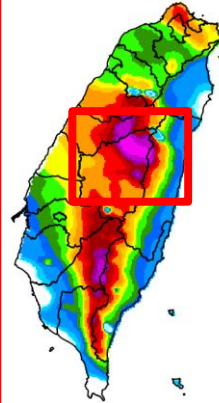
實驗二、集成不同權重 – 加入偏差修正 – 2023年整月平均

整體預報結果與未加入偏差修正相似
 相比算數級數加權，RMSE更低，但SSIM與CRPS略輸算數級數加權

	統計方法	幾何級數加權	算數級數加權	反比例加權	平方根倒數	經驗加權法		集成模型
OBS	NPM	[1, 2, 4, 8, 16]	[1, 1.2, 1.4, 1.6, 1.8]	[1.0, 42.2, 378.0, 4081.7, 17131.2]	[0.27, 1.72, 5.16, 16.95, 34.72]	[1,2,4,6,8]	[1,5,10,15,20]	
RMSE	27.4596	18.7316	18.1630	23.0001	20.9353	19.1637	19.3707	18.0587
SSIM	0.6019	0.6836	0.6843	0.6816	0.5391	0.6722	0.6755	0.6665
CRPS	9.8525	6.9568	6.9713	7.8948	9.4621	7.3553	7.2439	7.0928

實驗二、集成不同權重 – 加入偏差修正 – 卡努颱風

降低南投與中央山脈間的高報問題，整體降雨空間最接近實際觀測值

		統計方法	幾何級數加權	算數級數加權	反比例加權	平方根倒數	經驗加權法		集成模型
OBS	前6小時觀測值	NPM	[1, 2, 4, 8, 16]	[1, 1.2, 1.4, 1.6, 1.8]	[1.0, 42.2, 378.0, 4081.7, 17131.2]	[0.27, 1.72, 5.16, 16.95, 34.72]	[1,2,4,6,8]	[1,5,10,15,20]	
									
RMSE	-	69.5745	39.2836	35.2615	73.9281	44.2268	40.6855	43.0483	35.2931
SSIM	-	0.4136	0.4946	0.4819	0.5212	0.4499	0.4580	0.5024	0.5051
CRPS	-	22.9805	16.4414	15.7062	21.8135	18.7312	17.0183	17.0937	15.9274

實驗二、集成不同權重 – 小結

- 集成模型在預報極端大小雨時，出現高低報問題
 - 集成模型只接收6個模型的預報結果
 - ➔ 造成模型出現大雨低報，小雨高報
- 針對集成模型高低報問題進行修正
 - 加入MAE作為損失函數
 - 提升降雨極值準確度

實驗三、加入MAE修正預報誤差 – 不加入偏差修正

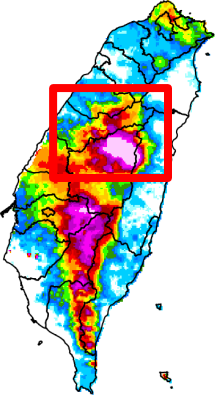
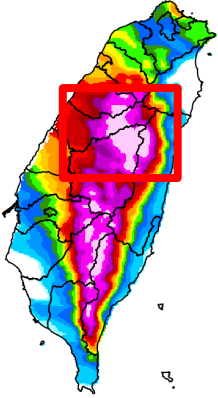
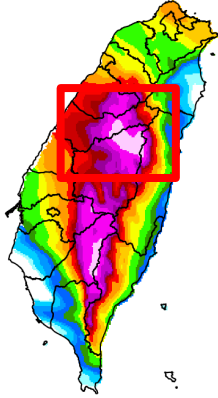
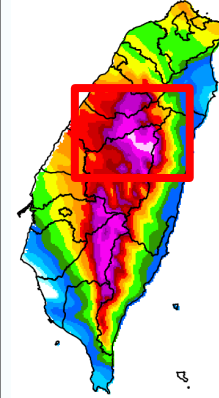
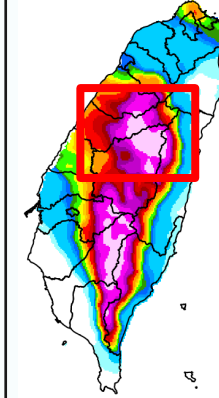
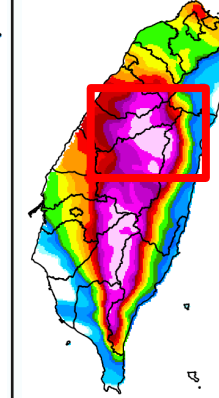
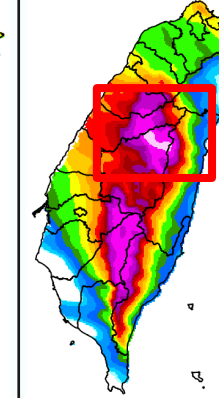
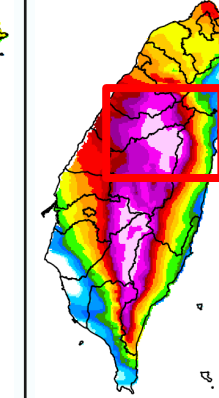
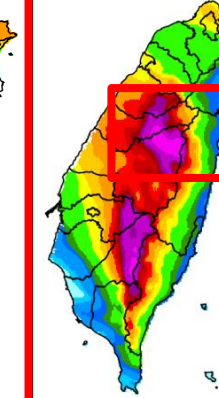
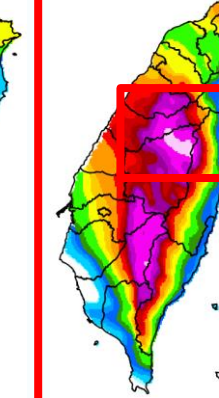
2023年整月平均

加入MAE後，RMSE誤差雖然稍高，但SSIM與CRPS獲得顯著提升

	統計方法	幾何級數加權	算數級數加權	反比例加權	平方根倒數計算	經驗加權法		集成模型	集成模型加入MAE
OBS	NPM	[1, 2, 4, 8, 16]	[1, 1.2, 1.4, 1.6, 1.8]	[1.0, 42.2, 378.0, 4081.7, 17131.2]	[0.27, 1.72, 5.16, 16.95, 34.72]	[1,2,4,6,8]	[1,5,10, 15,20]		
RMSE	27.4596	22.3221	21.4319	23.8447	24.1602	21.6936	24.5580	21.3653	21.6691
SSIM	0.6019	0.6192	0.6025	0.6378	0.6197	0.6280	0.6040	0.5630	0.6344
CRPS	9.8525	8.9011	8.8991	8.5590	9.2541	8.4590	9.9740	9.2802	8.5311

實驗三、加入MAE修正預報誤差 – 不加入偏差修正 – 卡努颱風

集成模型之預報表現較佳，但同時南投地區降雨極值低報

OBS	統計方法 NPM	幾何級數加權 [1, 2, 4, 8, 16]	算數級數加權 [1, 1.2, 1.4, 1.6, 1.8]	反比例加權 [1.0, 42.2, 378.0, 4081.7, 17131.2]	平方根倒數 [0.27, 1.72, 5.16, 16.95, 34.72]	經驗加權法		集成模型	集成模型 加入MAE
						[1,2,4,6,8]	[1,5,10,15,20]		
									
RMSE	69.5745	48.7548	39.9516	56.1688	66.4801	43.1156	66.0689	36.0905	43.1026
SSIM	0.4136	0.4504	0.4616	0.4986	0.4482	0.4462	0.4166	0.4742	0.4714
CRPS	22.9805	19.9652	18.1827	20.1719	23.2212	18.7298	24.6276	17.5362	18.5842

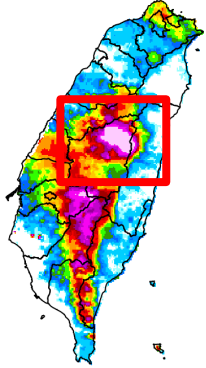
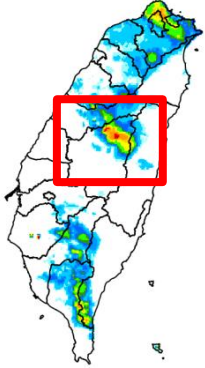
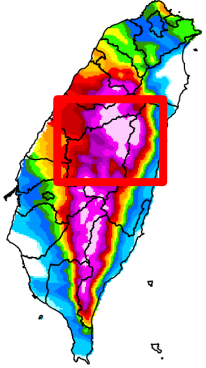
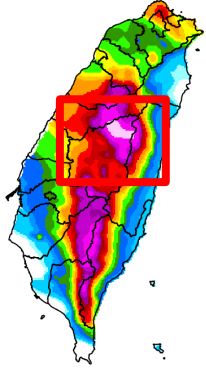
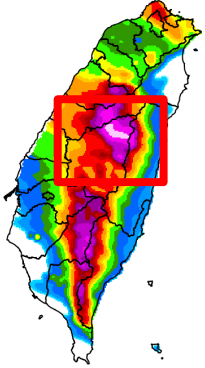
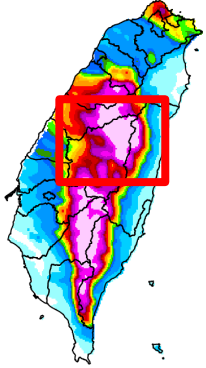
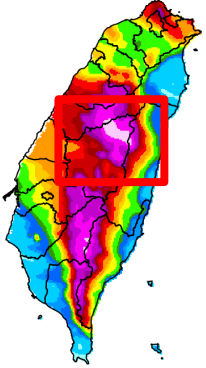
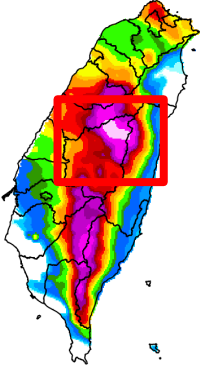
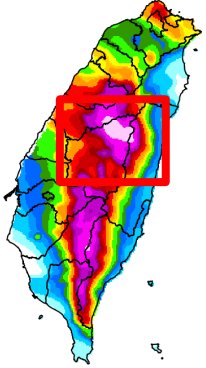
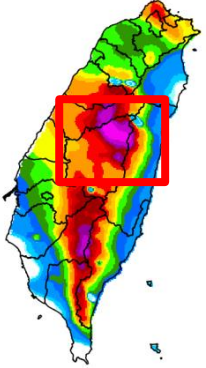
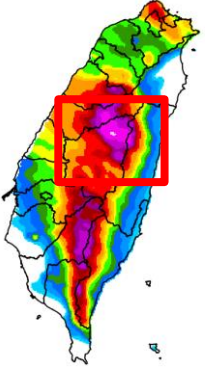
實驗三、加入MAE修正預報誤差 – 加入偏差修正 2023年整月平均

在加入偏差修正後，RMSE雖然些微高於未加入
但SSIM與CRPS顯著優於未加入MAE

	統計方法	幾何級數加權	算數級數加權	反比例加權	平方根 倒數計算	經驗加權法		集成模型	集成模型 加入MAE
OBS	NPM	[1, 2, 4, 8, 16]	[1, 1.2, 1.4, 1.6, 1.8]	[1.0, 42.2, 378.0, 4081.7, 17131.2]	[0.27, 1.72, 5.16, 16.95, 34.72]	[1,2,4,6,8]	[1,5,10,15,20]		
RMSE	27.4596	18.7316	18.1630	23.0001	20.9353	19.1637	19.3707	18.0587	18.0767
SSIM	0.6019	0.6836	0.6843	0.6816	0.5391	0.6722	0.6755	0.6665	0.6924
CRPS	9.8525	6.9568	6.9713	7.8948	9.4621	7.3553	7.2439	7.0928	6.8399

實驗三、加入MAE修正預報誤差 – 加入偏差修正 – 卡努颱風

加入偏差修正後，南投地區的大雨會有低報的問題

		統計方法	幾何級數加權	算數級數加權	反比例加權	平方根倒數計算	經驗加權法		集成模型	集成模型加入MAE
OBS	前6小時觀測值	NPM	[1, 2, 4, 8, 16]	[1, 1.2, 1.4, 1.6, 1.8]	[1.0, 42.2, 378.0, 4081.7, 17131.2]	[0.27, 1.72, 5.16, 16.95, 34.72]	[1,2,4,6,8]	[1,5,10,15,20]		
										
RMSE	-	69.5745	39.2836	35.2615	73.9281	44.2268	40.6855	43.0483	35.2931	35.5235
SSIM	-	0.4136	0.4946	0.4819	0.5212	0.4499	0.4580	0.5024	0.5051	0.4816
CRPS	-	22.9805	16.4414	15.7062	21.8135	18.7312	17.0183	17.0937	15.9274	15.7362

實驗三、加入MAE修正預報誤差 – 小結

- 加入MAE後，有效修正整體高低報問題
- 改善整體降雨空間分佈
 - 卡努颱風個案，未加入偏差修正下，成功還原大雨極值
- 加入偏差修正不一定會提升預報表現
 - 前6小時觀測值累積降雨量低，導致預報結果低報



4 結論與未來方向

結論

1. 在使用加權損失函數改善資料不平衡的情況下
 - 算術級數加權整體表現是最好
2. 使用集成學習整合各權重模型
 - 優點：整合多個模型的優勢，提升整體預報的準確度與穩定性
 - 缺點：預訓練模型特徵提取不佳 → 造成小雨高報、大雨低報
3. MAE 修正集成模型的降雨高低報問題
 - 改善降雨空間分佈、累積降雨量不準
4. 加入偏差修正不一定會提升預報表現

未來方向

1. 持續增加訓練樣本，豐富降雨型態
2. 判斷是否需要使用觀測值模型
 - 前 T 小時觀測值陸地區域總降雨量超過 N 毫米
 - 使用觀測值模型



感謝聆聽