

利用長短期記憶模型提升嘉義 地區能見度預測之準確性

**Enhancing the Accuracy of Visibility Forecasts in the Chiayi Area
Using Long Short-Term Memory (LSTM) Networks**



國防大學理工學院 環境資訊工程學系
共同作者：蔡世樵、陳致穎
報告人：翁語晨



≡ 研究動機與目的

研究動機：

- 嘉義地區冬季易因輻射冷卻效應而出現輻射霧，造成低能見度情況。
- 低能見度事件影響區域航空與交通安全，準確預測成為重要任務。
EX: 2000年4月24日，一架班機在嘉義機場降落時因低能見度而偏出跑道，顯示低能見度對飛航安全的威脅。
- 現行經驗公式準確度有限，需引入更強的資料驅動預測技術（如AI模型）。

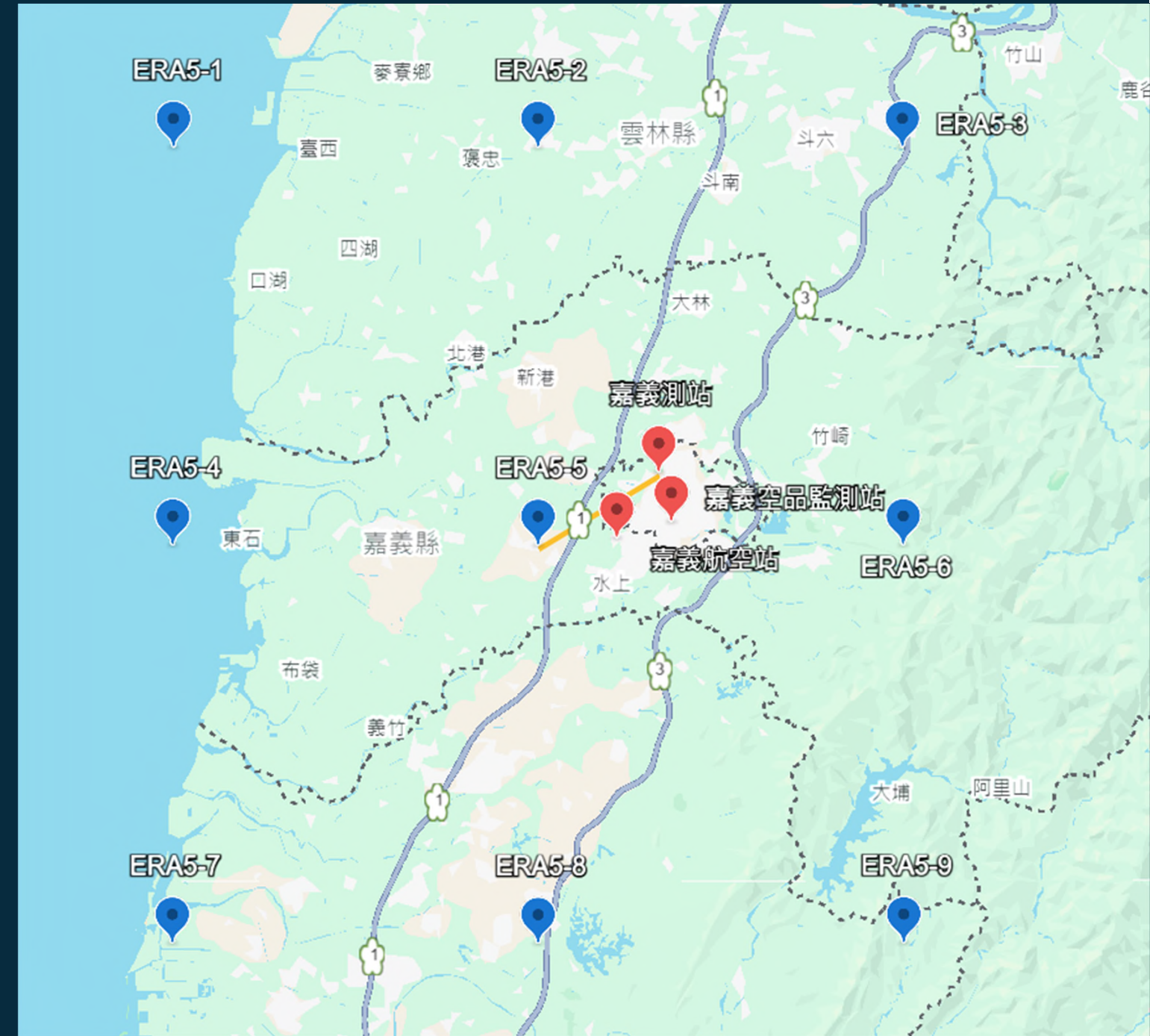
研究目的：

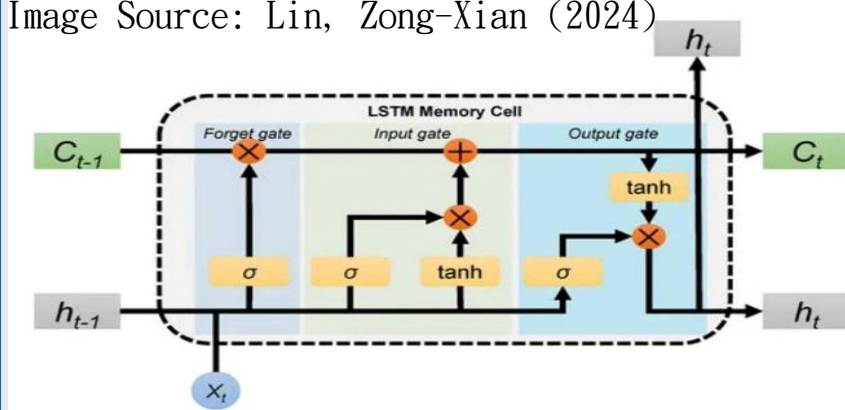
- 建構以氣象與空氣品質為基礎的能見度預測模型。
- 比較傳統公式（FSL、SW99）、多元線性回歸與LSTM模型之預測表現。
- 尋求提升能見度預測之準確度，以作為區域風險預報的輔助工具。

1. 中央氣象署逐時觀測資料（2013 - 2023）：
嘉義測站：氣溫、相對濕度、風速、降水量、能見度等觀測變數。
2. 環保署空氣品質逐時監測資料（2013 - 2023）：
包含PM2.5與PM10，作為能見度預測之重要污染指標。
3. 大氣科學研究與應用資料庫：
作為低能見度個案分析中大尺度天氣系統背景（如高壓出海、高壓迴流、東北季風等）判讀依據。

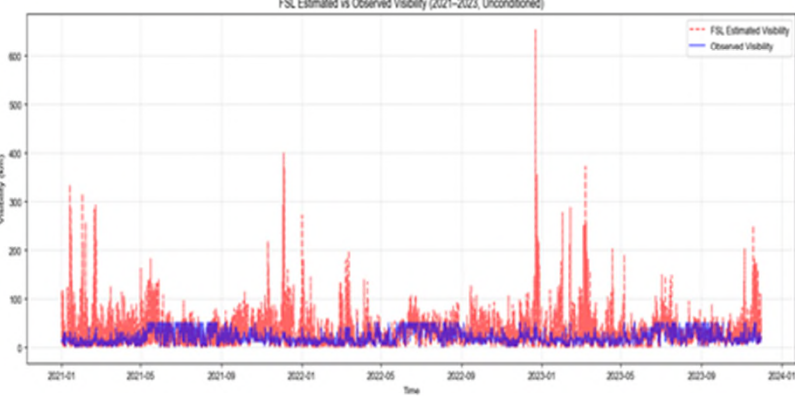
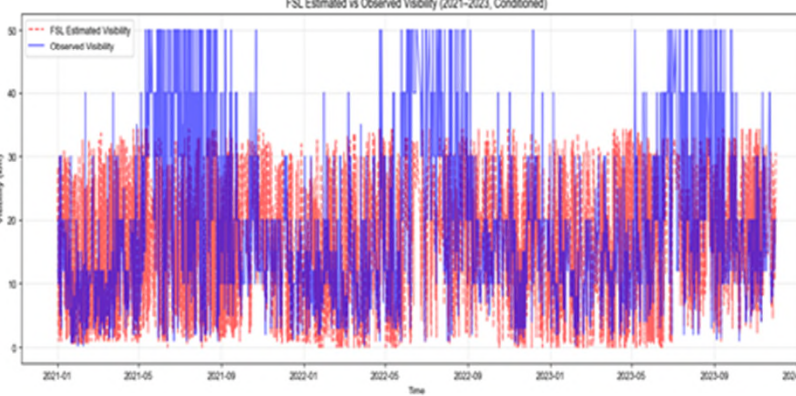
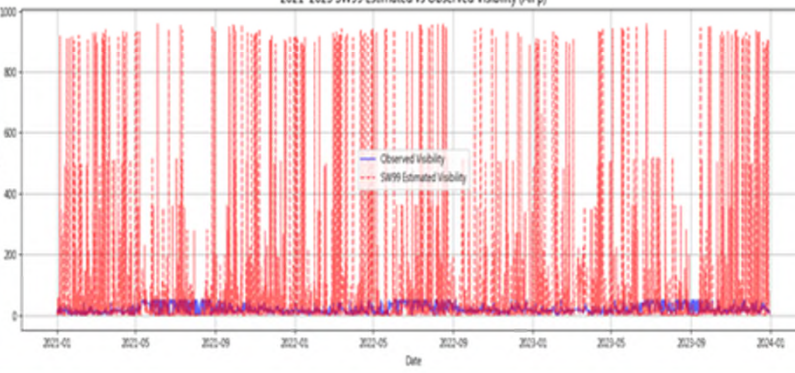
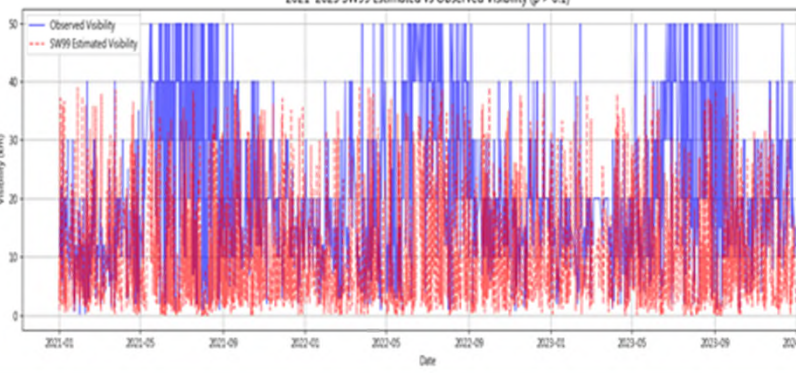
4. ERA5再分析資料 (2021 - 2023) :

- 空間範圍：
北緯 $23.2^{\circ} - 23.8^{\circ}$;
東經 $120.1^{\circ} - 120.7^{\circ}$ (涵蓋嘉義地區)
- 時間解析度：逐小時
- 使用變數：
(1) Total Column Liquid Water(TCLW):
雲中液態水含量 (SW99使用)
(2) Fraction of Cloud Cover:
霧日個案雲量分析



模型類型	模型名稱	原理	備註
能見度公式	FSL	由 NOAA 所提，根據溫度與露點溫度差 (T - Td) 與相對濕度 (RH) 估算大氣能見度	$x_{vis} = 1.61 * 6000 * \frac{T - T_d}{(RH)^{1.75}}$
	SW99	根據光學消光理論 (Beer's Law)，以雲中液態水含量估算大氣中的光學厚度與消光係數 β ，進而推算能見度	$x_{vis} = \frac{-\ln(0.02)}{\beta}$ $\beta = 144.7C^{0.88}$ $C=LWC = Qcloud \times \rho$
統計回歸方法	多元線性回歸(MLR)	假設能見度與各氣象與污染變數呈線性關係，適合初步探討影響因子與能見度的線性關聯性，並可快速建立預測模型	$Y = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots \varepsilon_i$ 自變數X(ex:氣溫、PM2.5...) 因變數Y(本研究為能見度)
機器學習模型	LSTM	時序神經網路，可以用來解決長序列訓練過程中梯度消失和梯度爆炸問題，具記憶門控可學時間依賴性	Image Source: Lin, Zong-Xian (2024) 

≡ FSL、SW99估計誤差調整(2021-2023年)

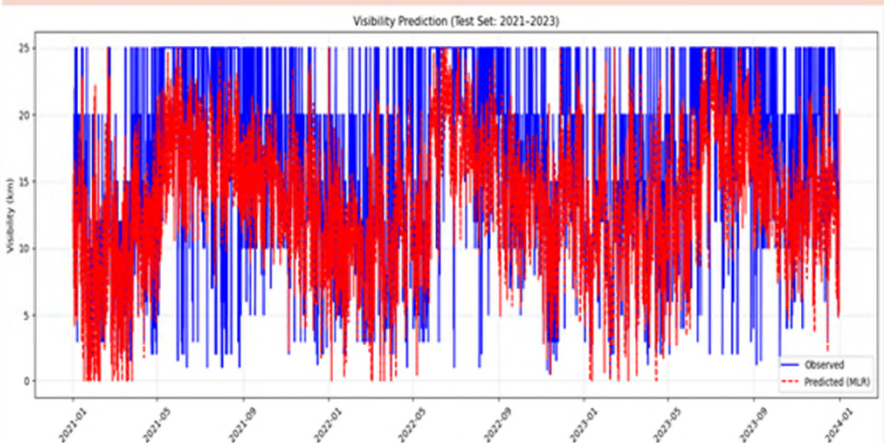
傳統能見度公式	條件篩選前	條件篩選後
FSL	RMSE: 36.03 R^2 : -5.95	$T - T_d \leq 6$ 且 $RH \geq 70\%$ RMSE: 9.21 R^2 : 0.12
		
SW99	RMSE: 198.45 R^2 : -214.49	$\beta > 0.1$ RMSE: 21.29 R^2 : -1.43
		

多元線性回歸模型表現

模型類型

模型測試集表現(2021-2023)

RMSE: 5.56
R²: 0.41



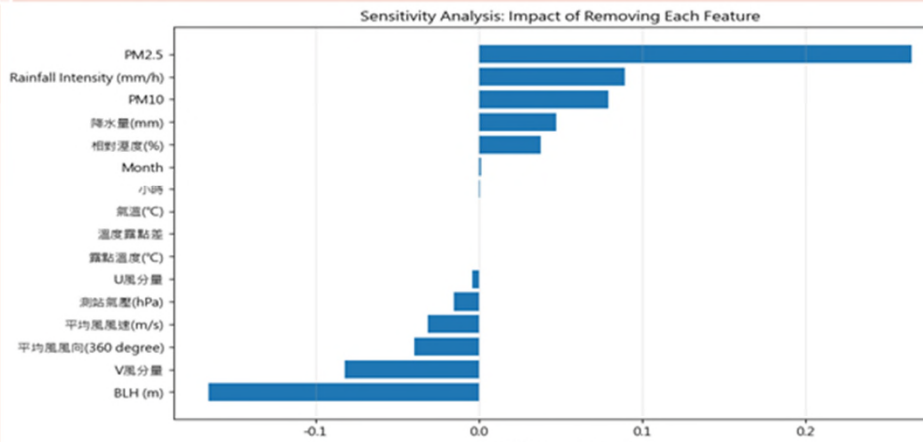
全部參數

主要變數/公式

$$\text{Visibility} = 116.092 + (-0.035 \times \text{Month}) + (0.020 \times \text{Hour}) + (-0.074 \times \text{測站氣壓 (hPa)}) + (-0.016 \times \text{氣溫 (°C)}) + (0.068 \times \text{露點溫度 (°C)}) + (-0.297 \times \text{相對溼度 (\%)}) + (0.420 \times \text{平均風風速 (m/s)}) + (-0.003 \times \text{平均風風向 (360 degree)}) + (-0.314 \times \text{U風分量}) + (0.743 \times \text{V風分量}) + (0.00028 \times \text{降水量 (mm)}) + (-0.000086 \times \text{Rainfall Intensity (mm/h)}) + (-0.005 \times \text{BLH (m)}) + (-0.154 \times \text{PM2.5}) + (-0.020 \times \text{PM10})$$

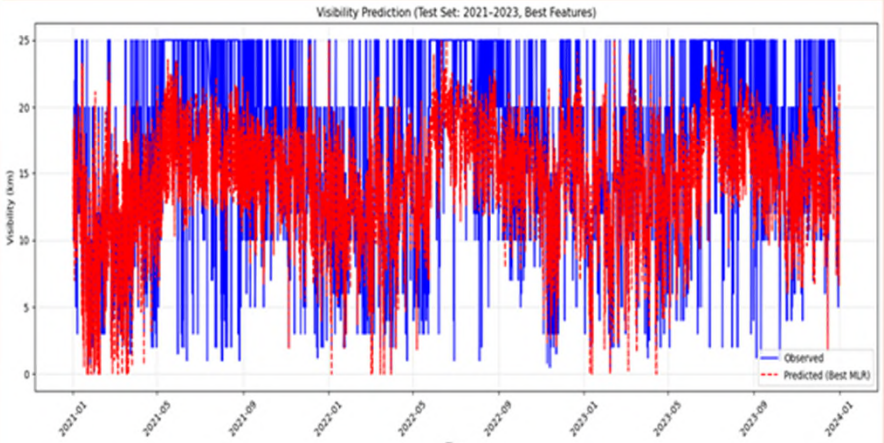
敏感度測試

ΔRMSE 為正 → 移除該特徵使模型表現變差 → 該特徵重要。
ΔRMSE 為負 → 移除該特徵反而略為改善模型 → 該特徵可能干擾模型。



經敏感度分析最佳變數組合

RMSE: 5.22
R²: 0.48



主要變數/公式

$$\text{Visibility} = 32.160 + (-0.005 \times \text{Month}) + (0.046 \times \text{Hour}) + (-0.197 \times \text{相對溼度 (\%)}) + (0.00028 \times \text{降水量 (mm)}) + (-0.000086 \times \text{Rainfall Intensity (mm/h)}) + (-0.171 \times \text{PM2.5}) + (-0.030 \times \text{PM10})$$

敏感度測試: 透過移除單一特徵的方式, 來檢視該特徵對模型表現的影響

Rank	Removed Feature	ΔRMSE (Removed - Baseline)	New RMSE
1	PM2.5	+0.264818	5.852284
2	Rainfall Intensity (mm/h)	+0.089276	5.676741
3	PM10	+0.079319	5.666784
4	降水量(mm)	+0.047696	5.635162
5	相對溼度(%)	+0.038005	5.625470
6	Month	+0.001390	5.588855
7	小時	+0.000605	5.588070
8	氣溫(°C)	~0.000000	5.587465
9	溫度露點差	~0.000000	5.587465
10	露點溫度(°C)	~0.000000	5.587465
11	U風分量	-0.003945	5.583520
12	測站氣壓(hPa)	-0.015095	5.572370
13	平均風風速(m/s)	-0.031159	5.556307
14	平均風風向(360 degree)	-0.039686	5.547780
15	V風分量	-0.082391	5.505074
16	BLH (m)	-0.165514	5.421951

≡ LSTM預測效能與比較

Preface

Methods

Analyze

Conclusion

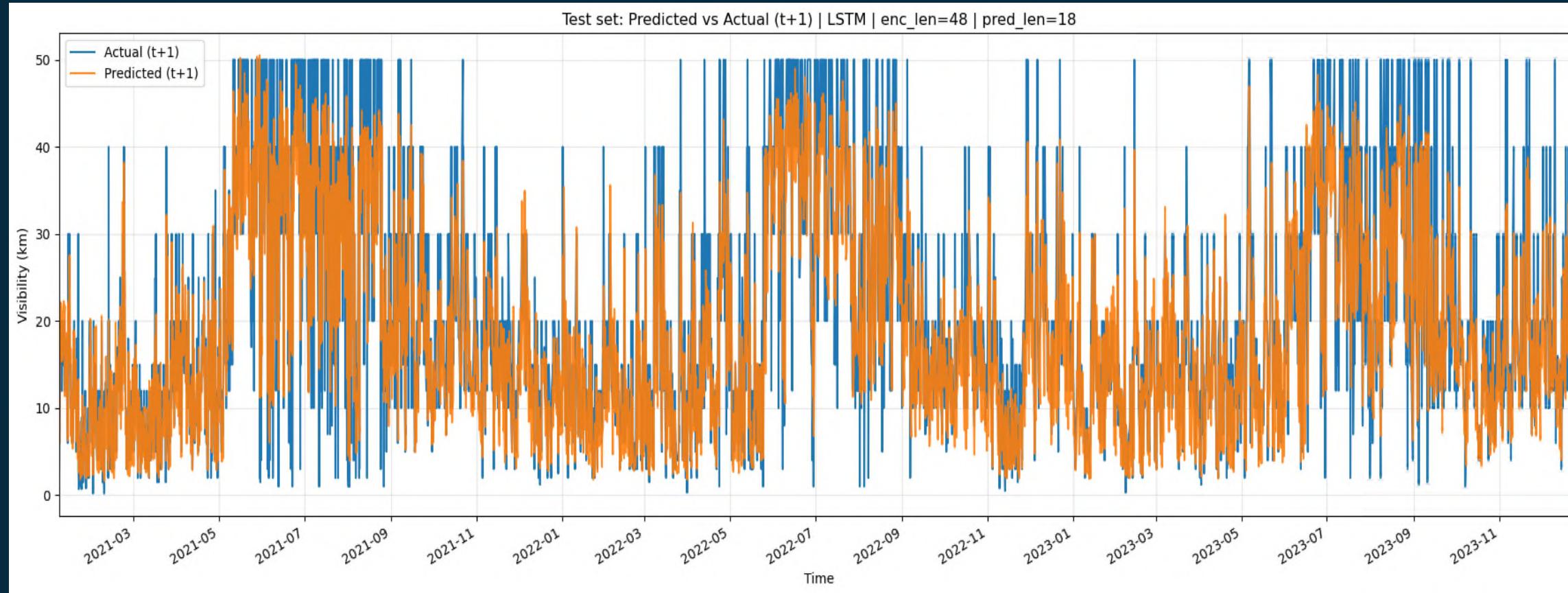
1. 參數設定及預測表現
2. 參數設定及預測表現(防止過擬合)

輸入之特徵參數根據測站能見度與氣象參數之相關係數

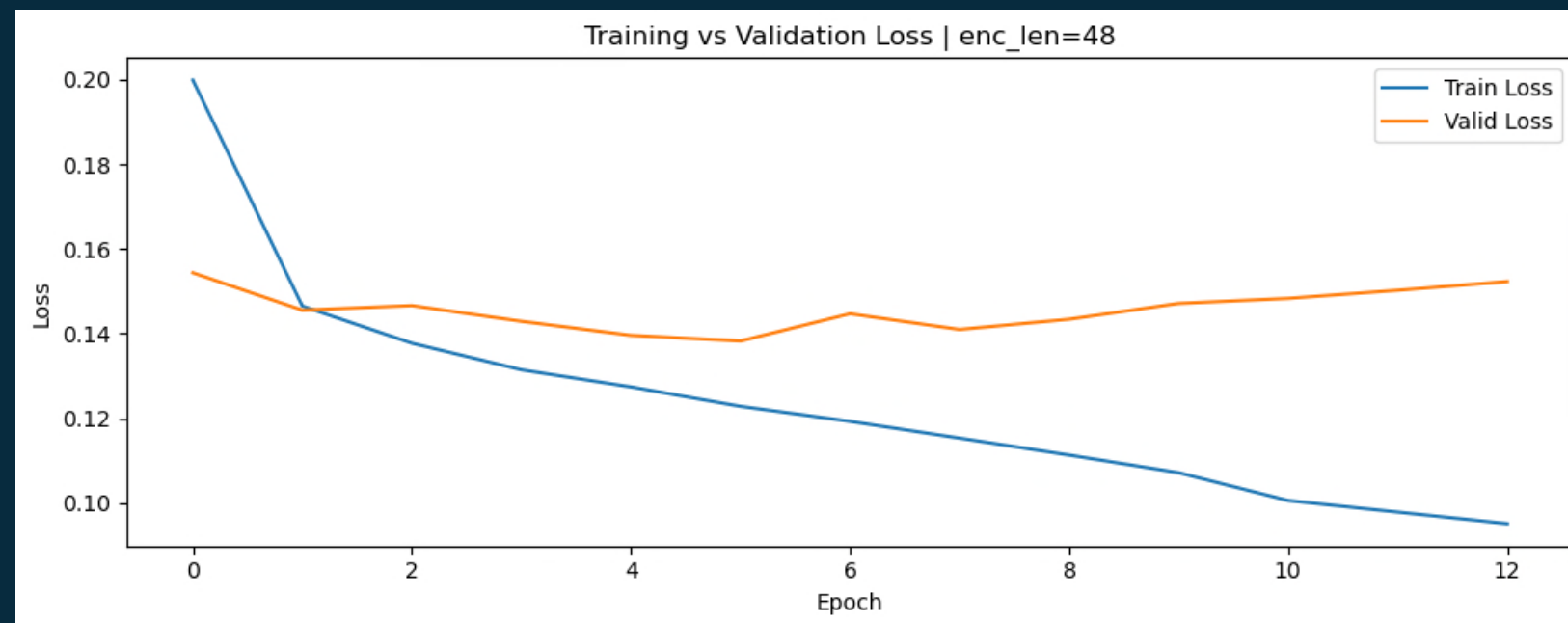
	溫度	露點溫度	氣壓	相對溼度	降水量	PM2.5	PM10	邊界層高度	降水強度
全時段	0.45	0.31	-0.33	-0.33	-0.08	-0.58	-0.25	0.24	-0.17

- 根據皮爾森相關係數，低度相關為正負0.1至0.3(黑)，中度相關為正負0.3至0.5(黃)，高度相關為正負0.5至1(紅)。
- 依相關性高低依序為PM2.5、溫度、氣壓、相對濕度，作為後續LSTM模型的特徵參數。

≡ LSTM預測表現



測試集 t+1 (2021-2023年):
RMSE(km)=7.1936
MAE(km)=4.8194
 $R^2=0.7214$



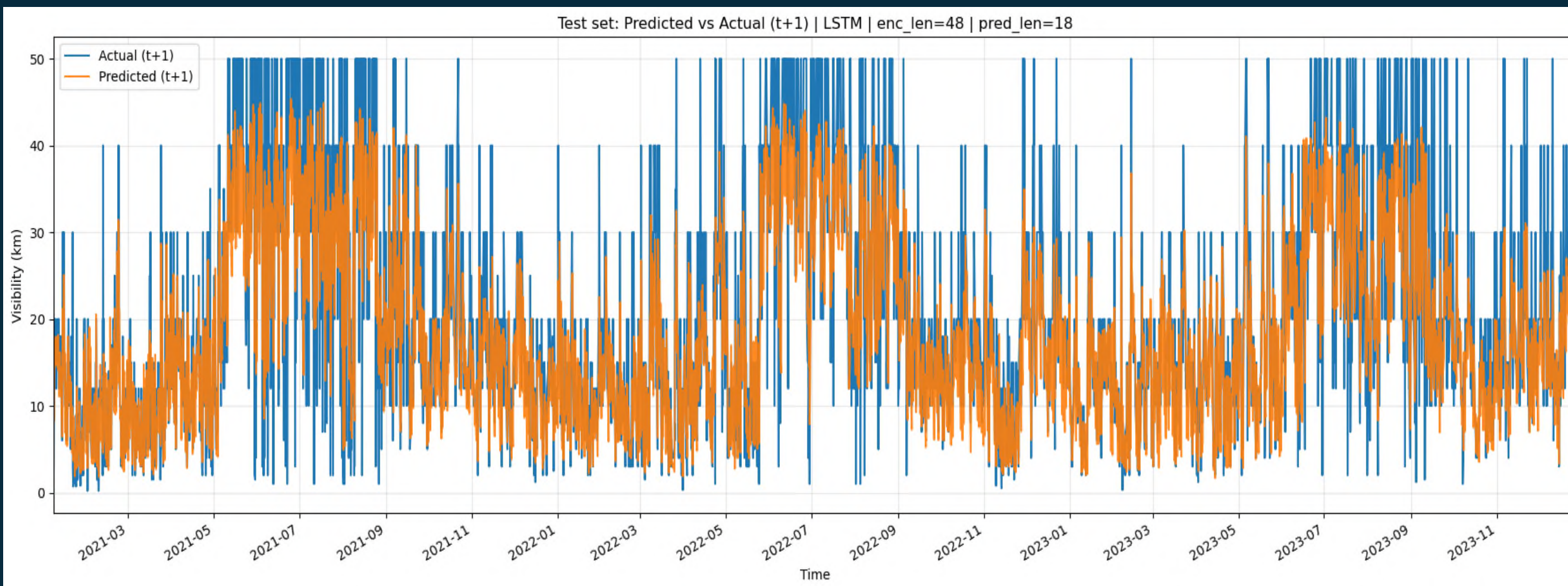
前期: 訓練集和驗證集都在下降。
中期: 訓練集繼續下降, 但驗證集開始震盪。
後期: 訓練集loss繼續下降, 但驗證集loss趨勢往上。
=>過擬合現象

≡ LSTM模式(防止過擬合)參數設定

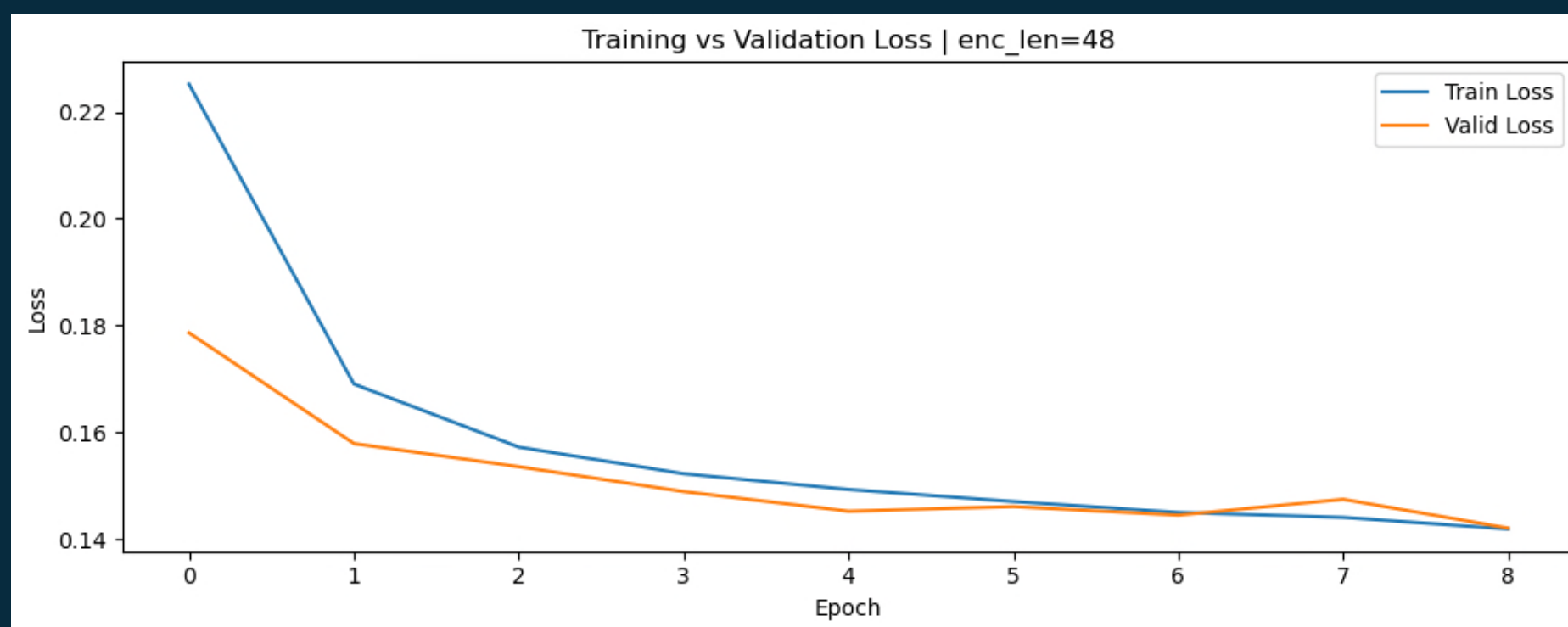
原版、防止過擬合版：差異對照

參數	原版	防過擬合版	差異/說明
LSTM 規模	enc/dec hidden=128、layers=2	enc/dec hidden=64、layers=1	縮小模型容量，降低過擬合風險
LSTM Dropout	0.25	0.40	增加Dropout，訓練時更強的隨機失活
FC Dropout	0.25	0.50	輸出前更強正則
權重衰減	weight_decay=5e-5	5e-4	更強L2
EarlyStopping	patience=8	patience=4	提早停止，避免過訓
LR調整	Plateau：patience=4	Plateau：patience=2	加速LR調整，更快降學習率
LayerNorm	無	Decoder輸出加LayerNorm	穩定訓練、抑制過擬合
資料增強	無	Gaussian noise (0.02) + time-step masking (p=0.10)	只在訓練時對 Encoder/Decoder 輸入套用

≡ LSTM(防止過擬合)預測表現



測試集 t+1 (2021-2023 年):
RMSE(km)=7.3671,
MAE(km)=5.0555,
 $R^2=0.7078$



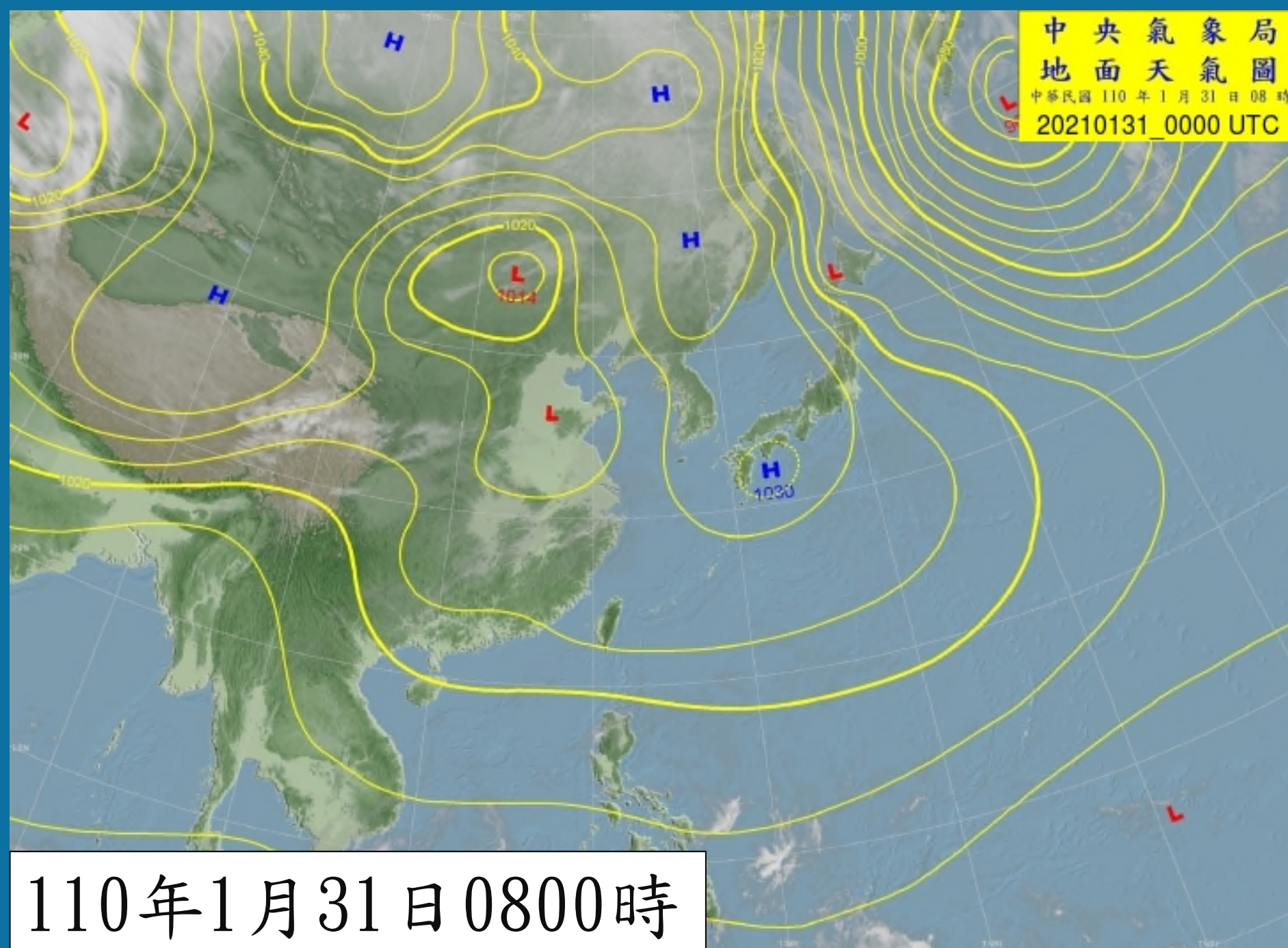
訓練集 vs 驗證集 差距很小
驗證 Loss 沒有明顯回升
=> 無過擬合現象

≡ 低能見度個案分析

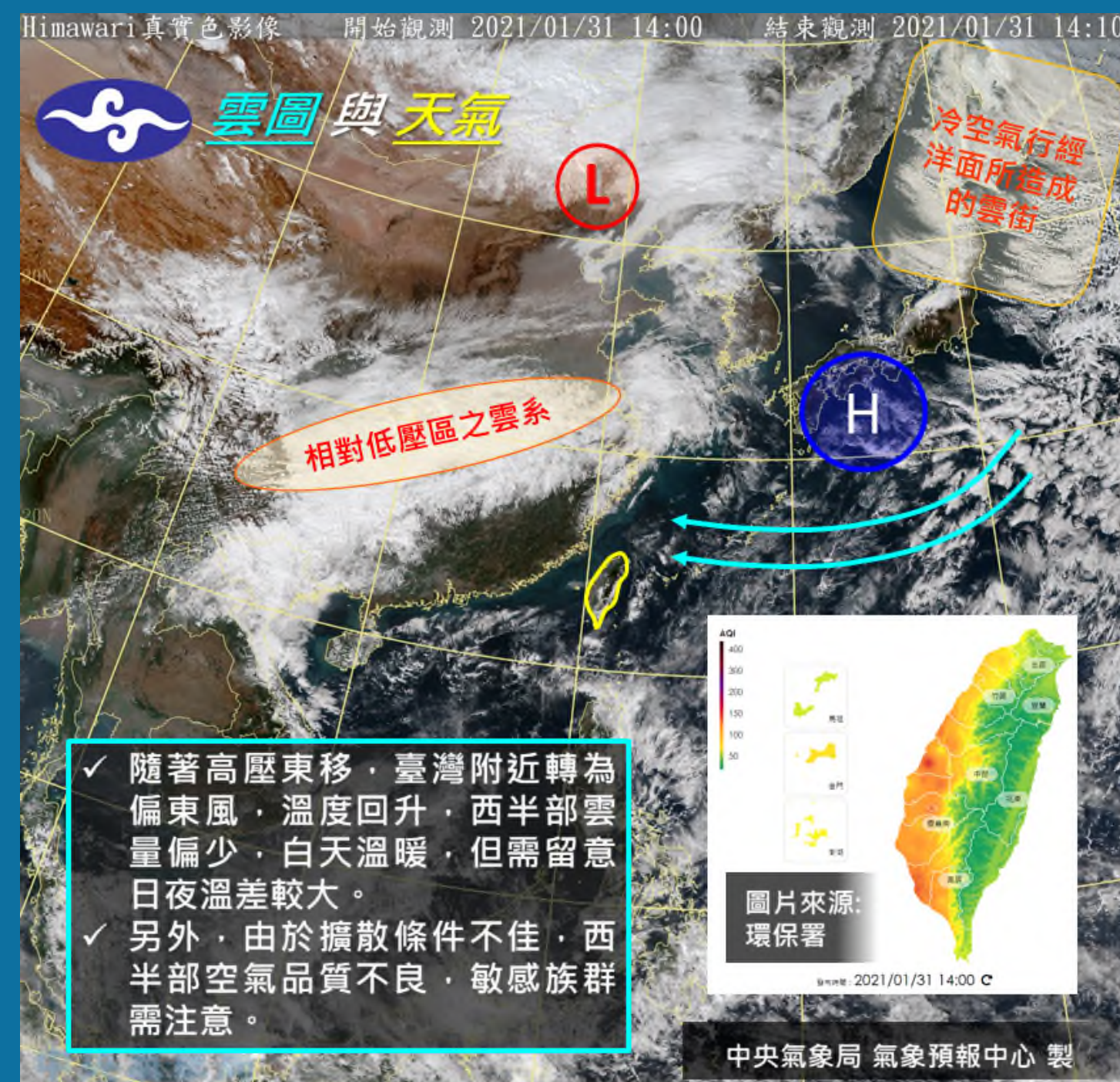
1. 個案一：2021/1/31
2. 個案二：2021/2/8

三個案一 綜觀天氣分析

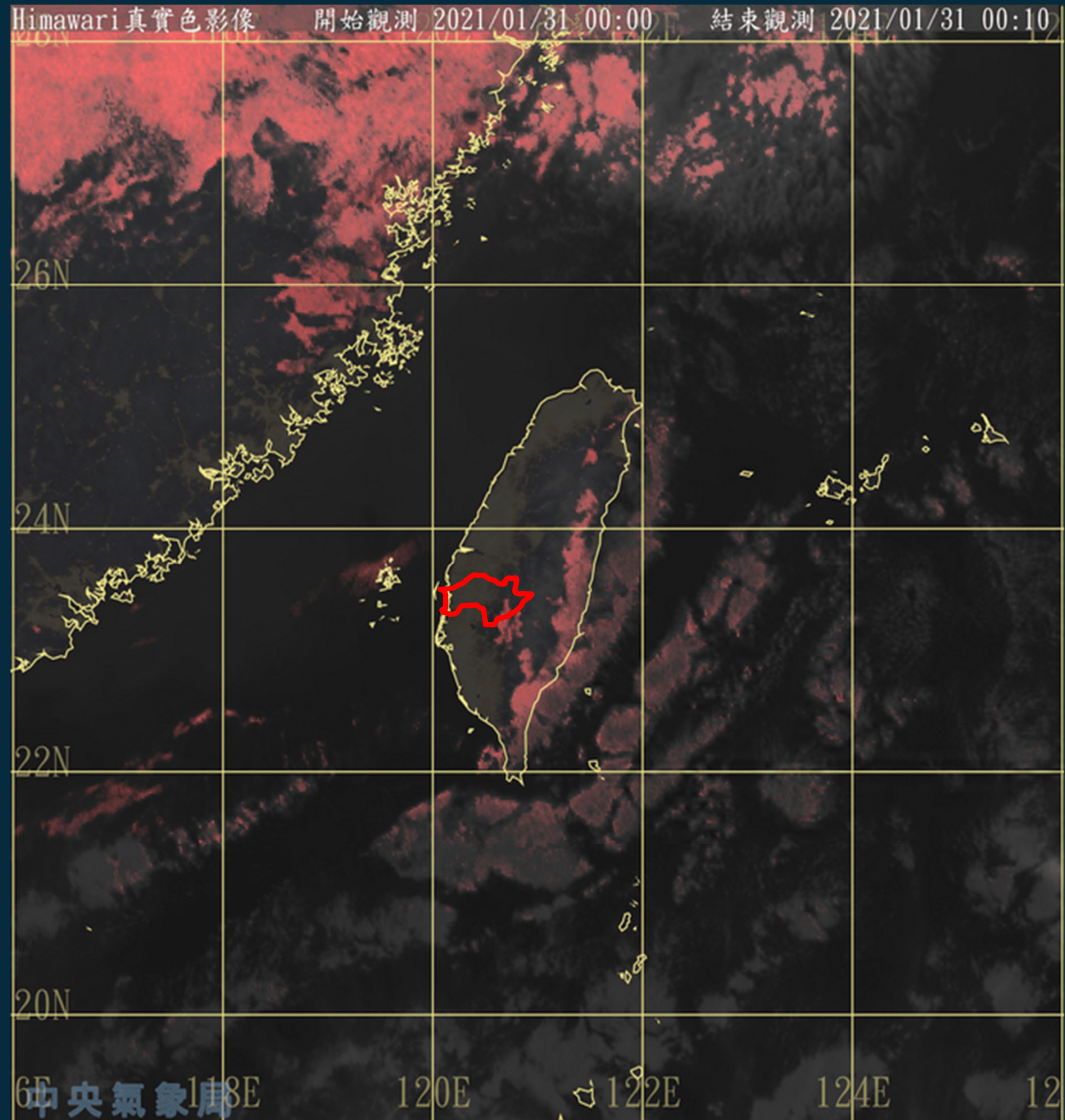
2021年1月31日08時，天氣系統為大陸冷氣團尾端，高壓出海，臺灣地區風場為偏東風。西半部位於背風側，擴散條件不佳，雲量較少、天氣穩定。



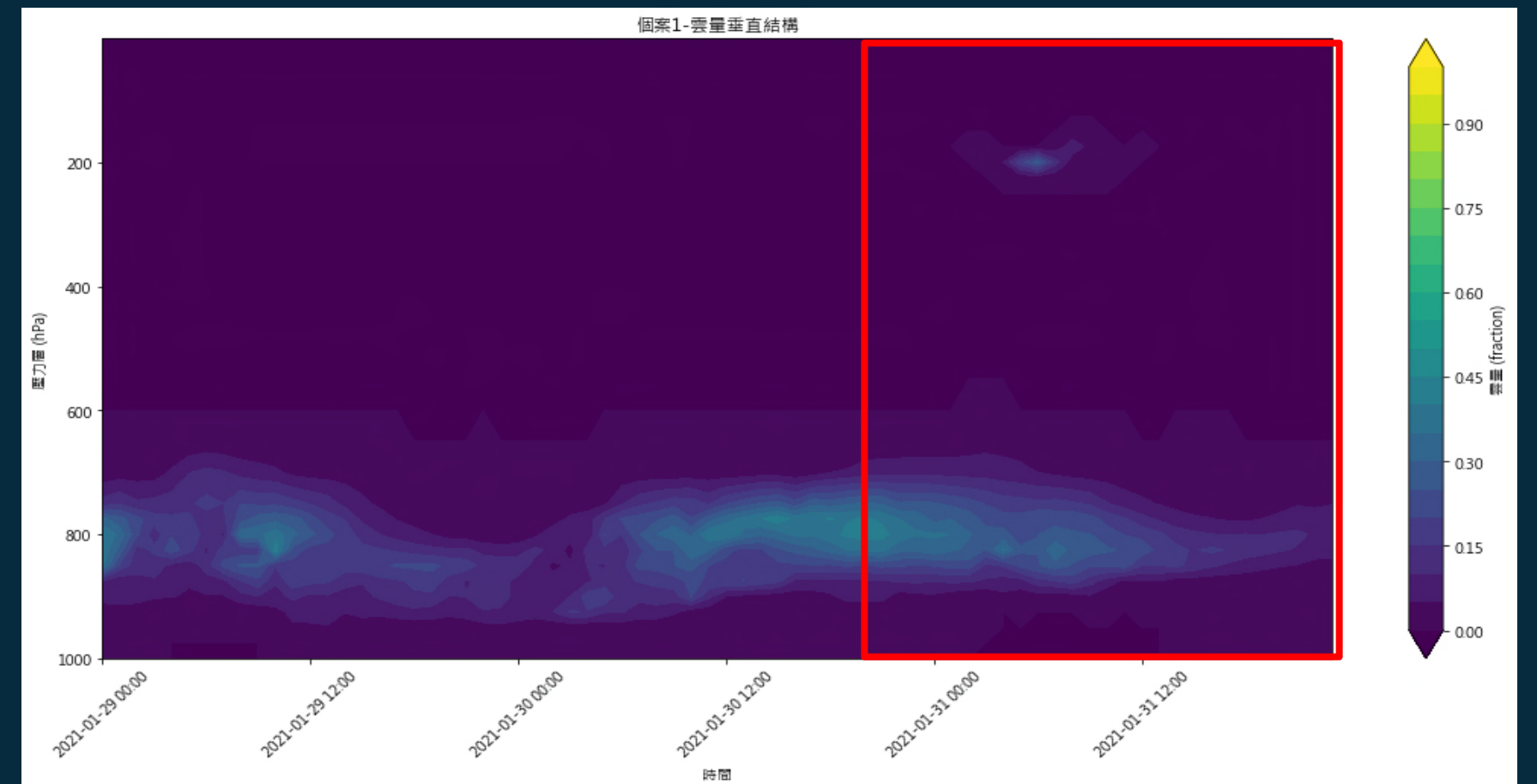
110年1月31日0800時



≡ 個案一 衛星雲圖及ERA5雲量

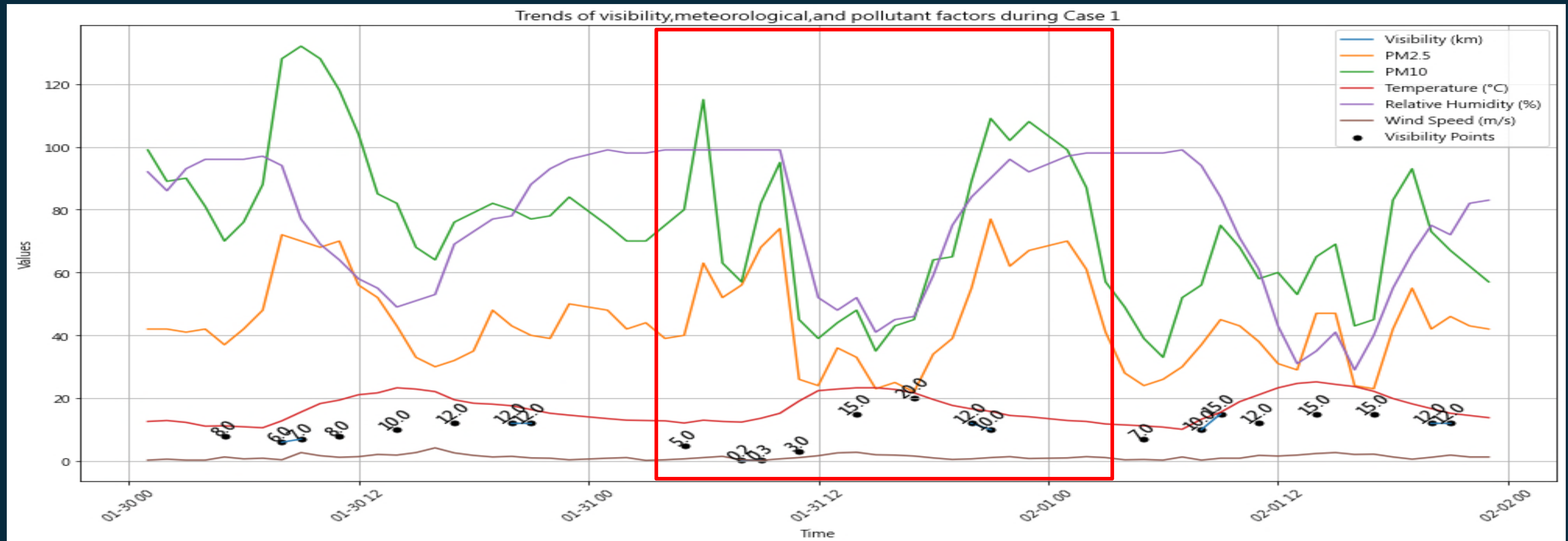


- 雲圖顯示西部沿海一帶有雲層堆積。
- ERA5雲量稀少，雲主要集中在於 875–925hPa。



個案一 能見度與氣象污染因子變化趨勢

個案一與相對濕度高、風速低及清晨氣溫下降密切相關，顯示為輻射冷卻主導的輻射霧事件；能見度下降期間PM2.5與PM10有上升，表污染粒子有參與凝結作用。

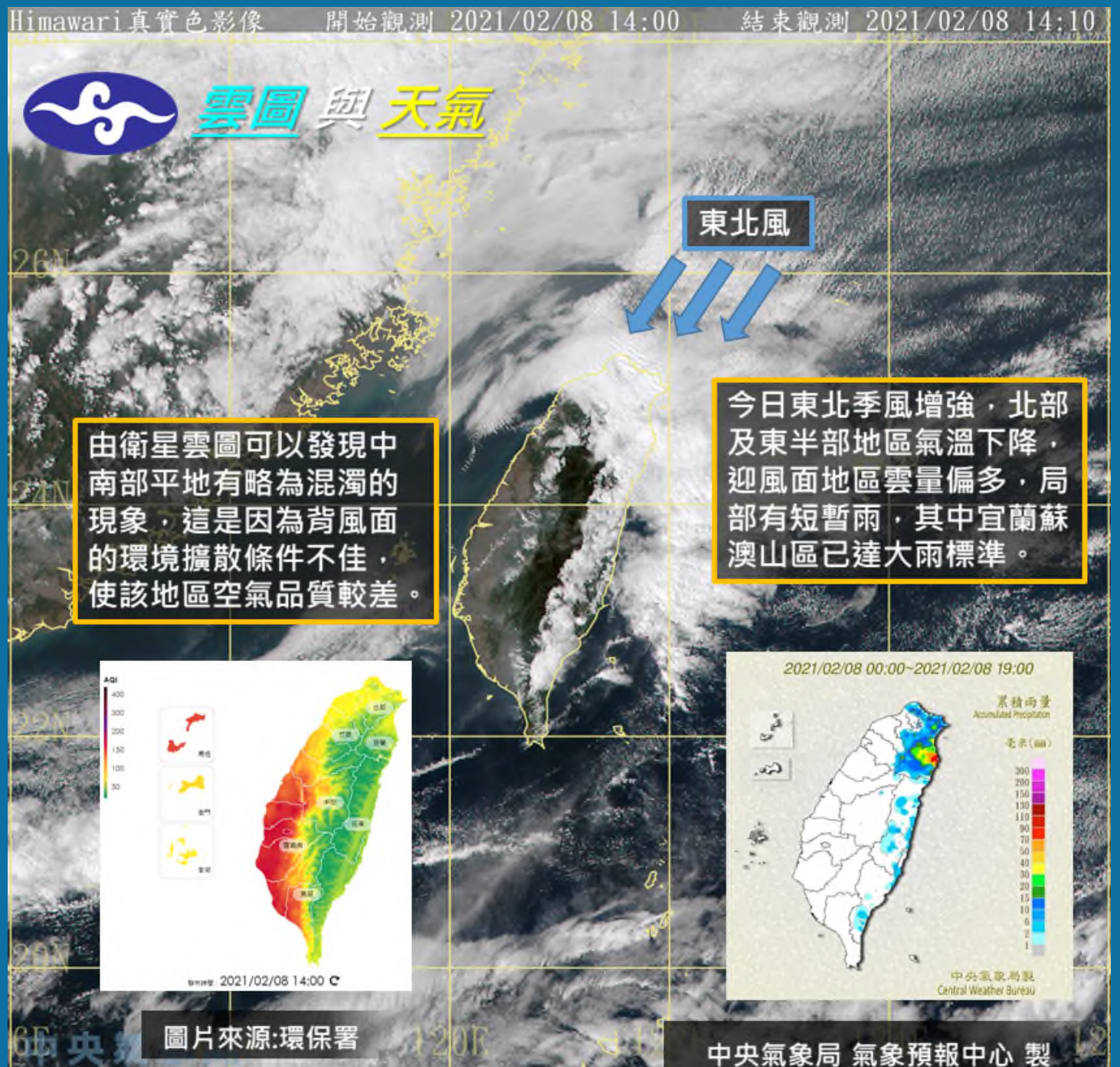
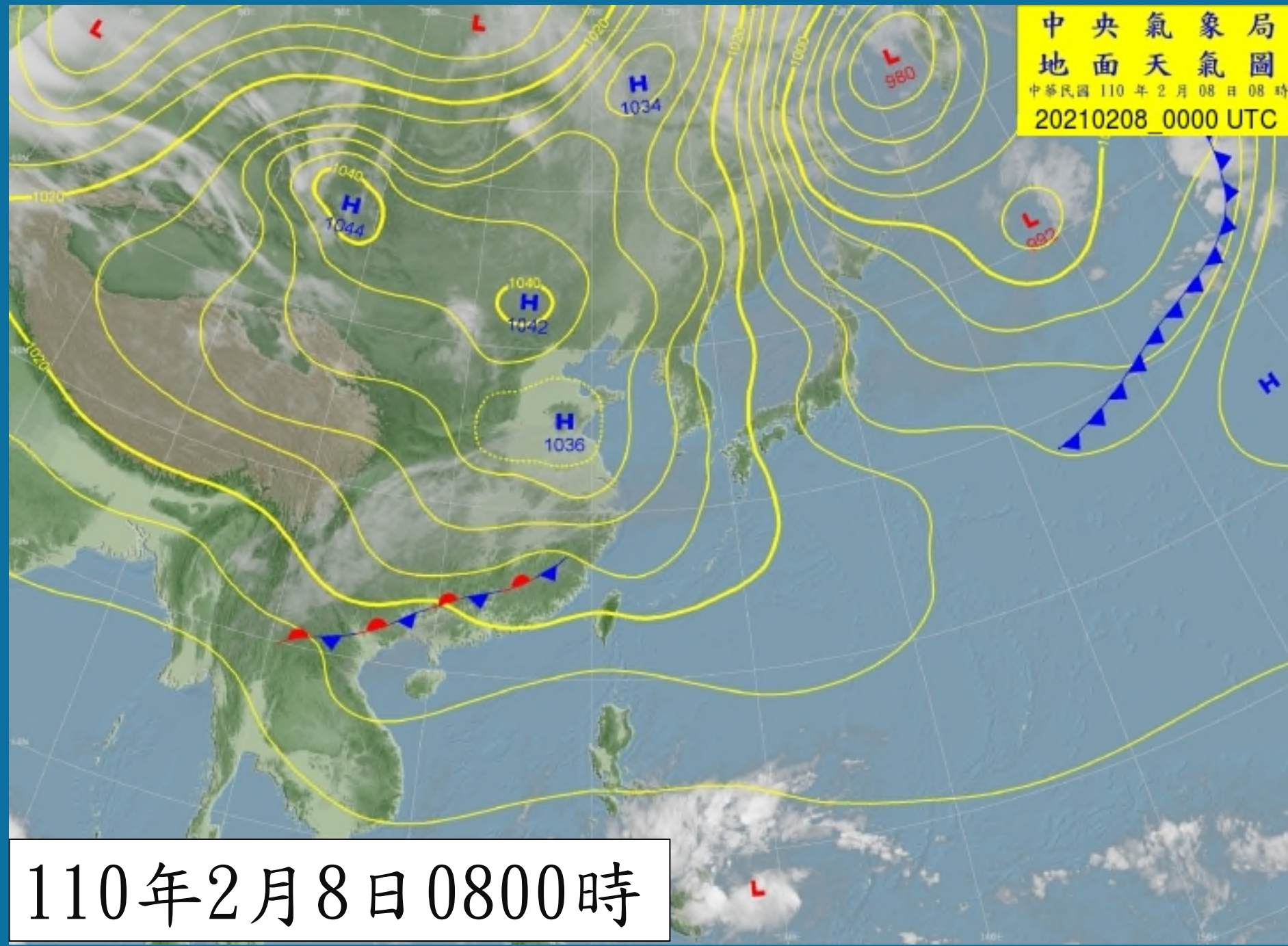


三個案一 各模型預測結果

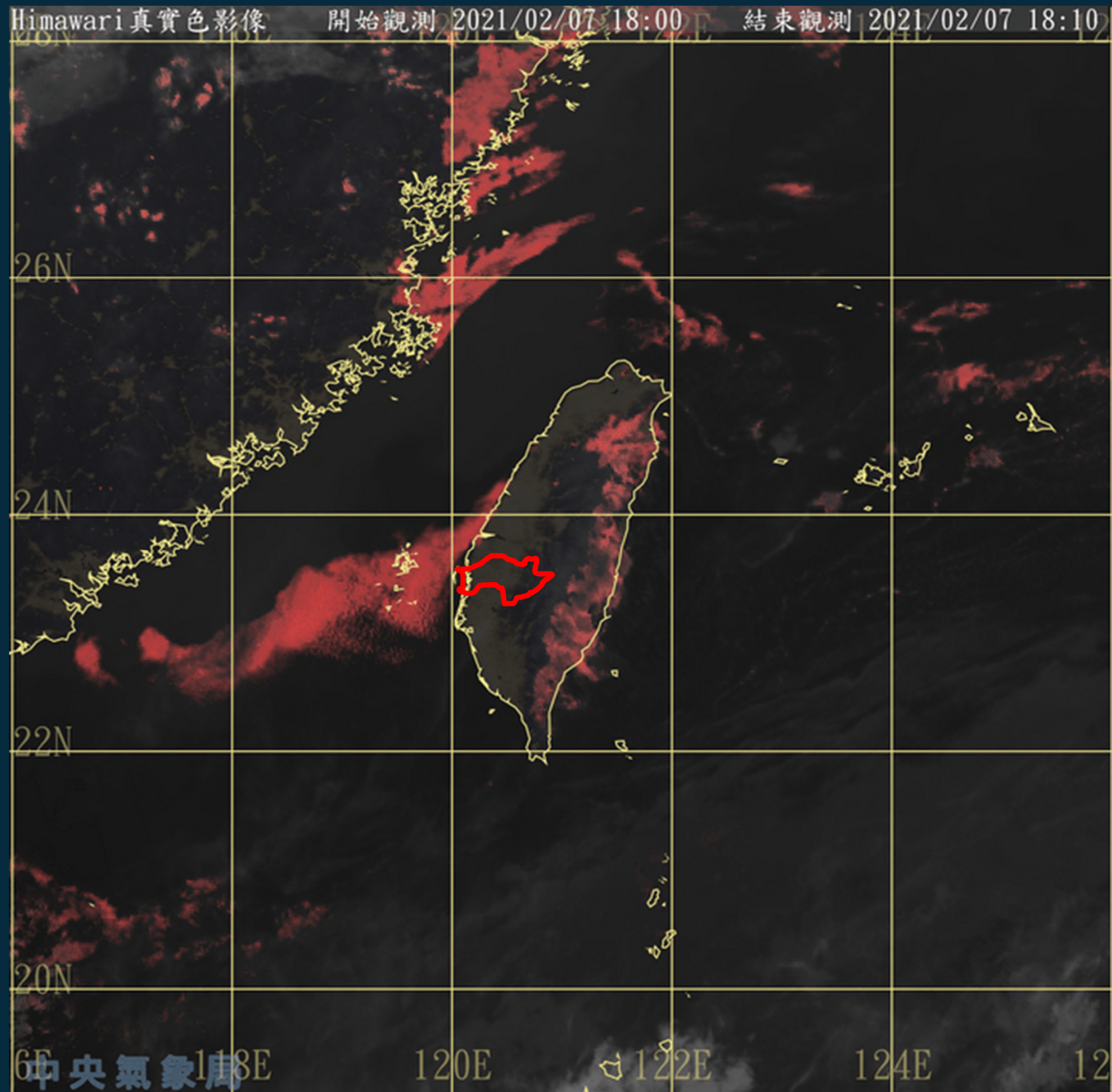
Model	Case 1(2021/1/31)		圖
	RMSE	R ²	
FSL	2.72	0.69	
SW99	8.62	-1.59	
MLR	4.47	0.23	
LSTM(18h)	2.46	0.74	
LSTM-防過擬合(18h)	2.72	0.68	

三個案二 綜觀天氣分析

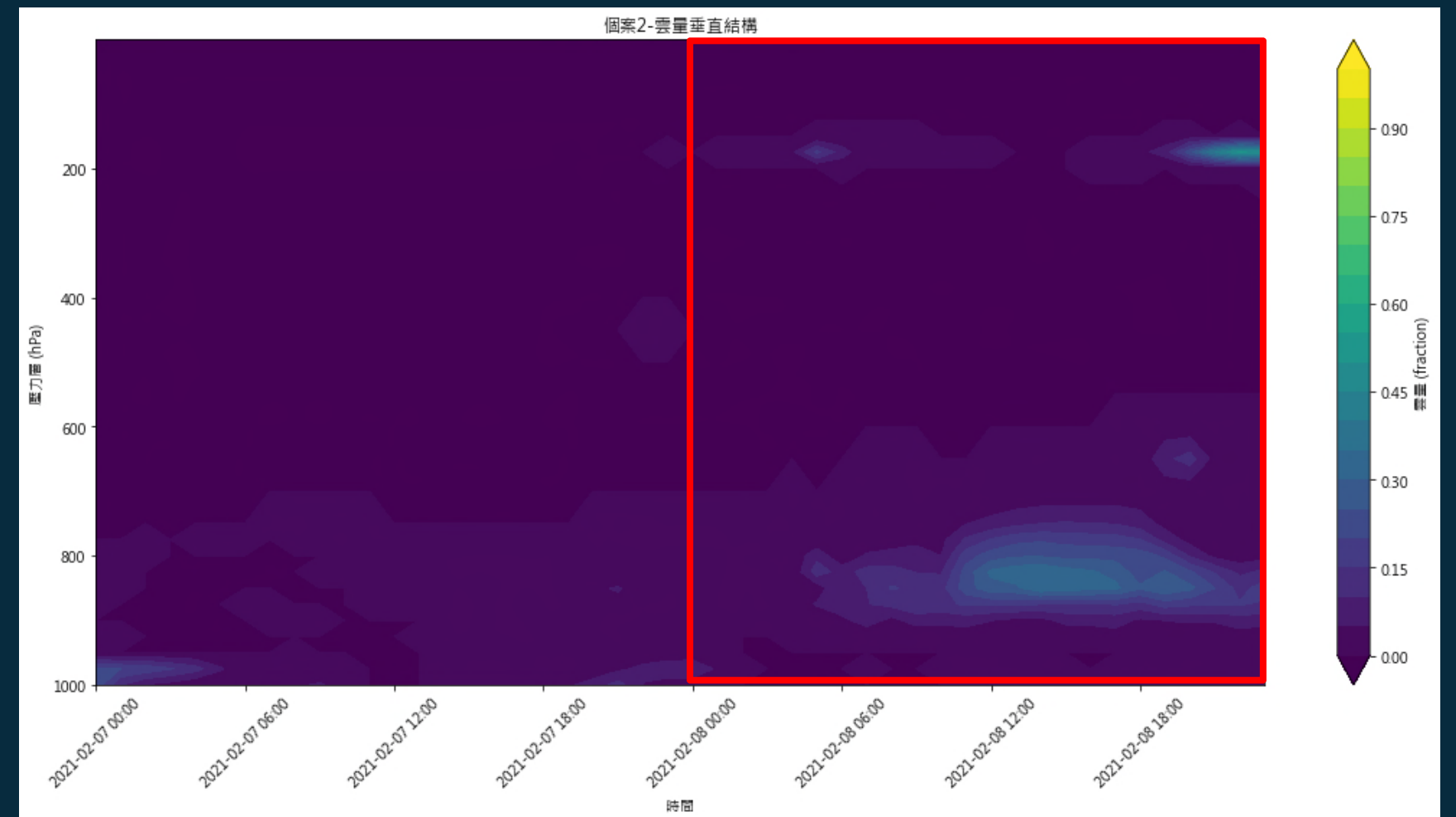
2021年2月8日08時，天氣系統為東北季風增強，臺灣地區風場為東北風。中南部為背風面，擴散條件不佳，汙染物累積，空氣品質差（橘紅）。



個案二 衛星雲圖及ERA5雲量

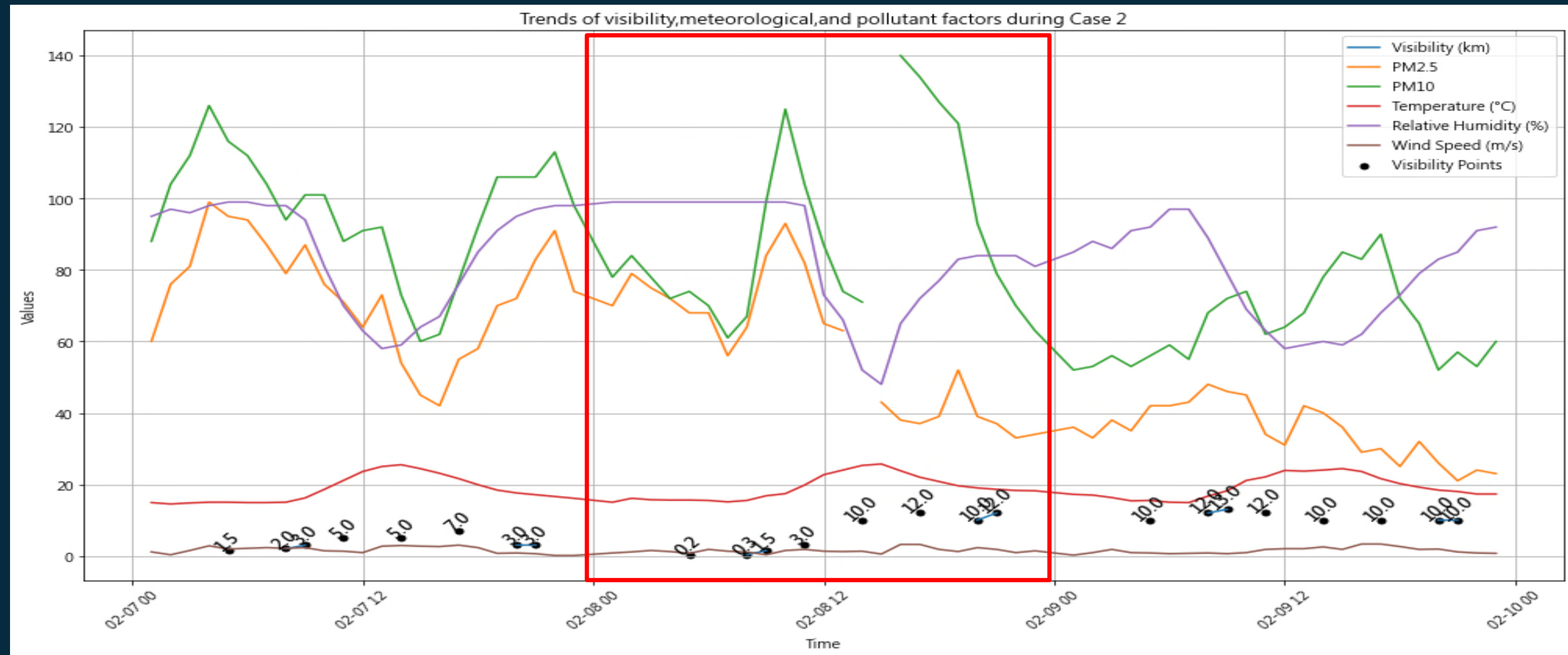


- 雲圖可見雲系僅覆蓋臺灣西部沿海一帶
- ERA5雲量稀少，雲層集中在900–800hPa



≡ 個案二 能見度與氣象污染因子變化趨勢

能見度急劇下降至1公里以下，PM2.5與PM10顯著上升，風速低於2m/s、相對濕度接近100%，氣溫穩定下降。顯示該時段為輻射冷卻、污染物滯留、濕度接近飽和造成能見度下降。



三個案二 各模型預測結果

Model	Case 2(2021/2/8)		圖
	RMSE	R ²	
FSL	1.19	0.94	
SW99	5.03	-0.63	
MLR	2.55	0.56	
LSTM(18h)	2.96	0.50	
LSTM-防過擬合(18h)	3.12	0.44	

- 傳統公式 (FSL、SW99) 在嘉義地區的預測準確度有限，尤其SW99表現極差，顯示其假設條件與本地氣候特性不符。
- 多元線性回歸能部分改善，但仍受限於線性假設。
- LSTM模型在測試集與個案中整體表現優於傳統方法，能更好捕捉能見度與氣象/污染因子之間的非線性關係。

Model	Case 1(2021/1/31)		Case 2(2021/2/8)	
	高壓出海		東北季風增強	
	輻射冷卻+空汙		輻射冷卻+空汙	
	RMSE	R ²	RMSE	R ²
FSL	2.72	0.69	1.19	0.94
SW99	8.62	-1.59	5.03	-0.63
MLR	4.47	0.23	2.55	0.56
LSTM(18h)	2.46	0.74	2.96	0.50
LSTM-防過擬合(18h)	2.72	0.68	3.12	0.44

1. 擴展應用至其他易霧區域

後續預計導入金門地區，驗證模型在不同地形與氣候條件下之適用性。

2. 導入更多低能見度事件

納入不同天氣系統下之低能見度事件，增加樣本多樣性，以提升模型在不同天氣條件下的泛化能力。

3. 引用其他AI模型提升預測表現

導入其他深度學習架構如Transformer或RNN，並與LSTM模型進行比較，探討其在捕捉長期依賴性與多特徵交互作用下的效果，以進一步提升能見度預測準確度。

1. Liang, C. W., Chang, C. C., Hsiao, C. Y., & Liang, C. J. (2023). Prediction and Analysis of Atmospheric Visibility in Five Terrain Types with Artificial Intelligence.
2. 梁佳齡, 廖杞昌, & 汪建良. (2017). 臺灣及其週邊海域低空層能見度預報與分析.
3. 黃淑倫, & 陳玟諭. (2020). 民俗活動期間社區大氣懸浮微粒濃度變化 - 以嘉義縣朴子市為例.
4. 梁延剛, 胡文志, & 楊敬基. (2007). 香港能見度、大氣懸浮粒子濃度與氣象條件的關係.
5. 林冠伶, 江宙君, 林欣弘, 朱容練, 徐理寰, 黃紹欽, & 于宜強. (2024). 離島機場能見度AI分析系統介紹 - 以金、馬地區機場為例. 災害防救電子報.
6. 許義宏, 林達遠, 曹勝傑, 陳子健, & 李俊穎. (2024). 馬祖港域能見度探討之研究(2/2)能見度預測模式之概念驗證. 國際標準書號.
7. 環境部. (2024). 影響空氣品質的氣象要素. 環境部空氣品質監測網.
<https://airtw.moenv.gov.tw/CHT/Encyclopedia/pedia07/pedia7.aspx>

報告完畢 恭請指導

Thank you for your listening.