

自我監督式深度學習方法應用於 溫度檢核機制評估

陳佳莉¹ 鄭軻安¹ 馮智勇¹ 劉坤波² 張博雄² 蔡立夫²

多采科技有限公司¹、中央氣象署²

114年天氣分析與預報研討會

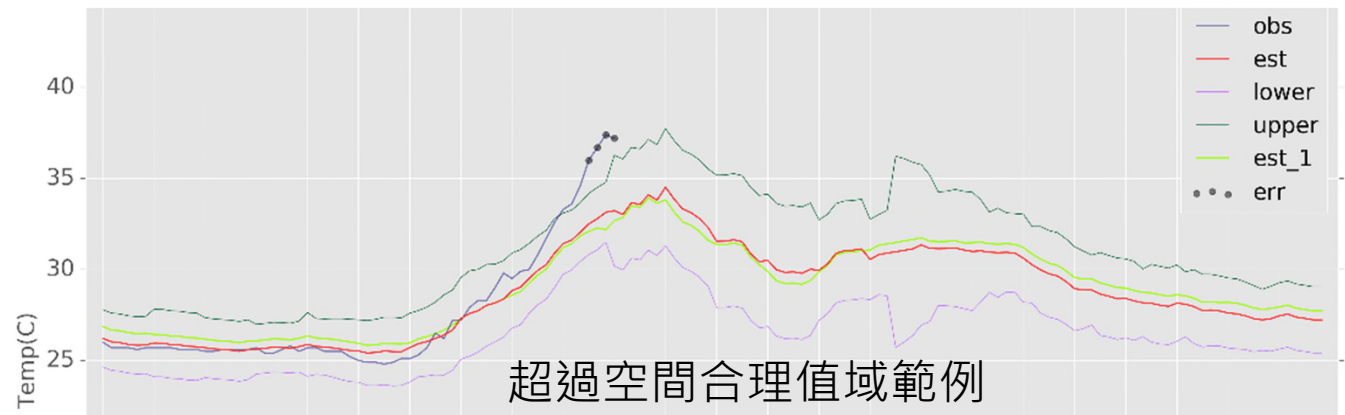
溫度檢核概況

■ 逐時與逐10分鐘溫度，歷史資料檢核

- ◆ 溫度的空間檢核利用克利金(Kriging)法推估溫度的合理值域，觀測值超出推估合理值域則視為異常。
- ◆ 空間檢核之誤判率過高，被提列個案接入時間檢核機制救回。
- ◆ 先前已開發各測站分別建立之AutoEncoder (AE)逐時溫度檢核模型，並上線運行中，可救回9成以上之誤提列。
- ◆ 但對部分降雨降溫或快速升溫等個案仍有誤判，故本次嘗試加入濕度資料。

■ 逐分鐘溫度，即時檢核

- ◆ 目前尚無自動檢核流程。

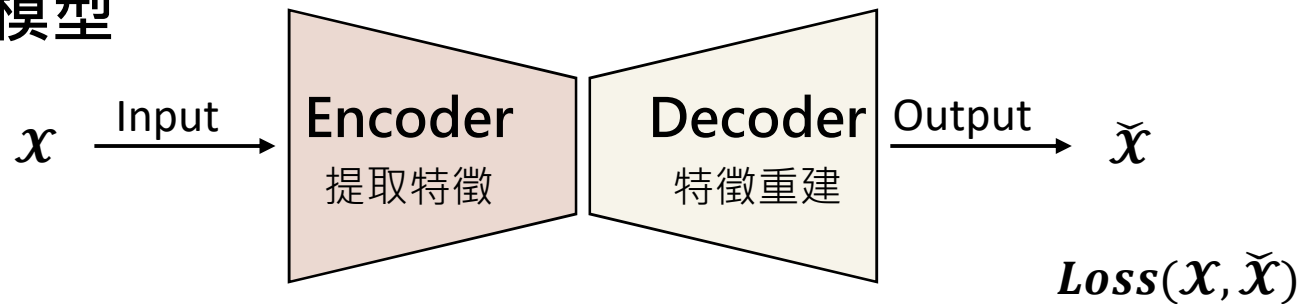


AutoEncoder異常檢核簡介

■ 溫度時間序列特性

- ◆ 正常資料量 >> 疑似異常資料量，且不易判斷是否真的異常。
- ◆ 故使用非監督式學習方法

■ 以正常資料訓練模型



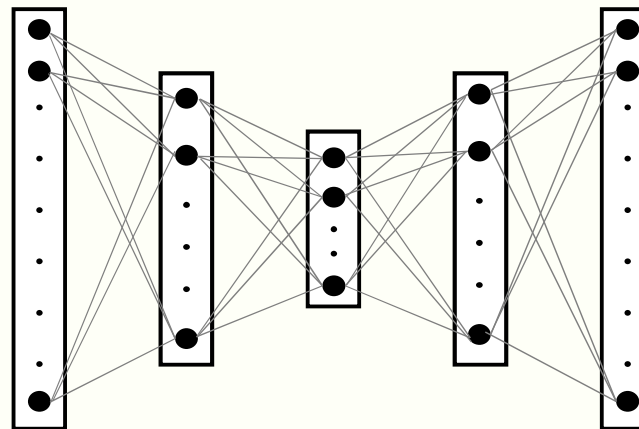
■ 應用

- ◆ 輸入正常資料AE可成功還原，輸入異常資料則輸出與輸入有明顯差異。

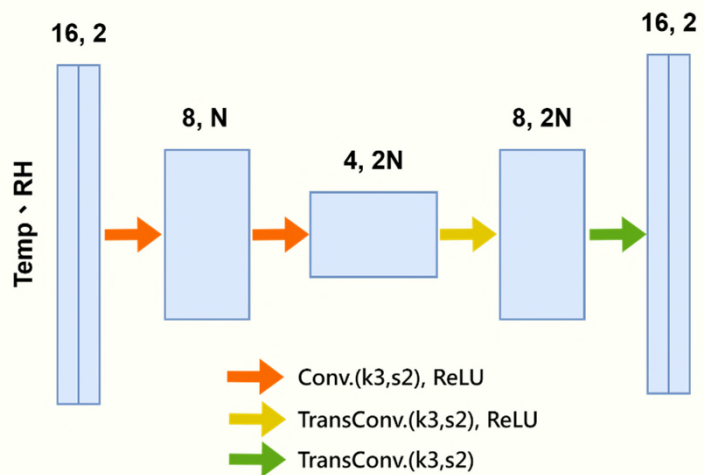
AutoEncoder 模型 架構測試

- 測試不同架構之AE，以 Convolution AE 擬合表現最佳

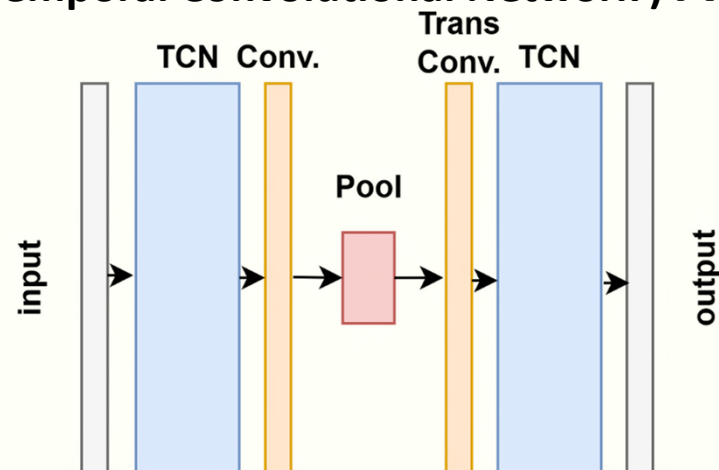
Fully connected AE



Convolution AE



TCN(Temporal Convolutional Network) AE



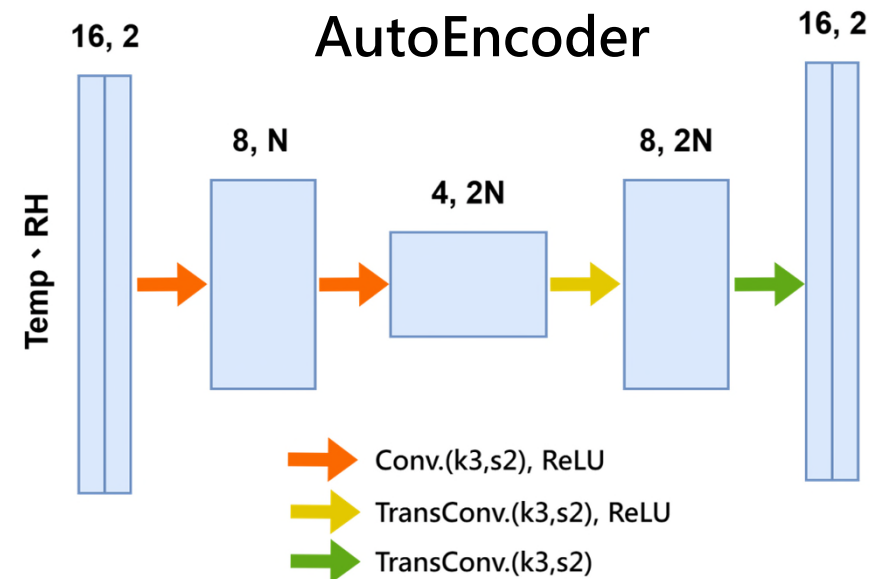
資料處理與訓練

■ 舊版溫度逐時、1分鐘AE模型

- ◆ 逐時AE，使用2018~2022年溫度經檢核之正常資料，各測站分別訓練模型。
- ◆ 1分鐘AE，使用2022年1分鐘溫度資料，各測站分別訓練模型。

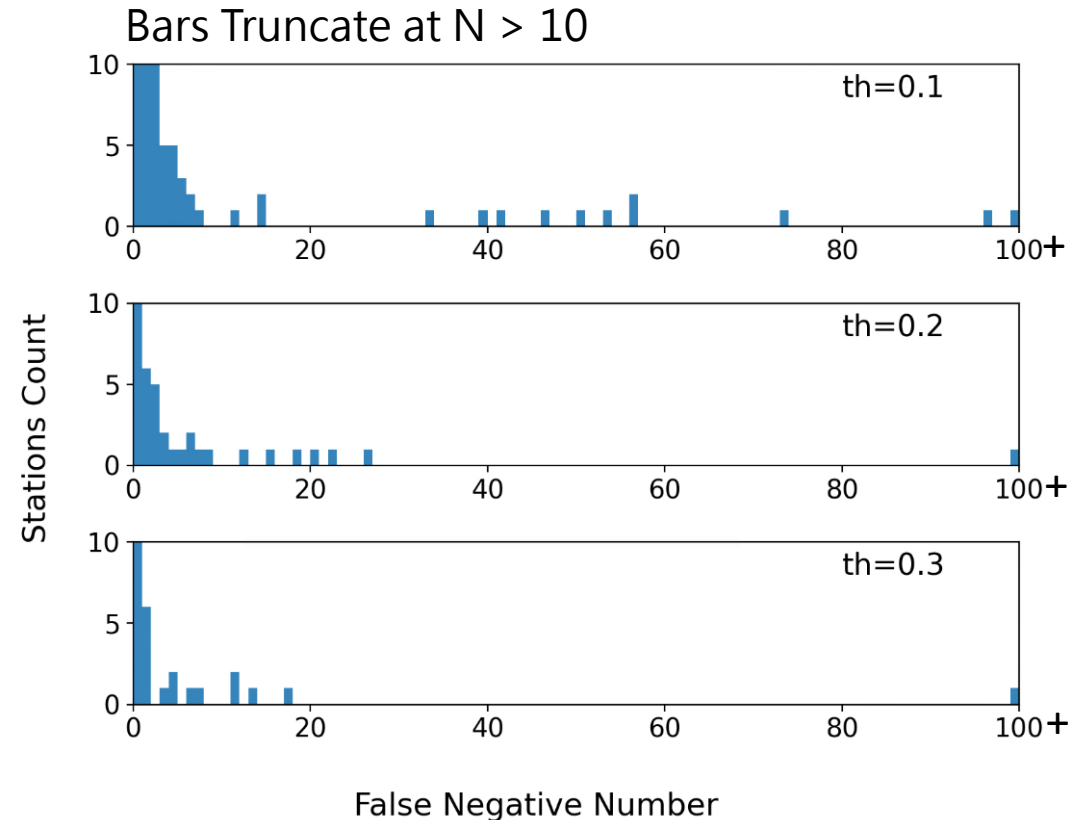
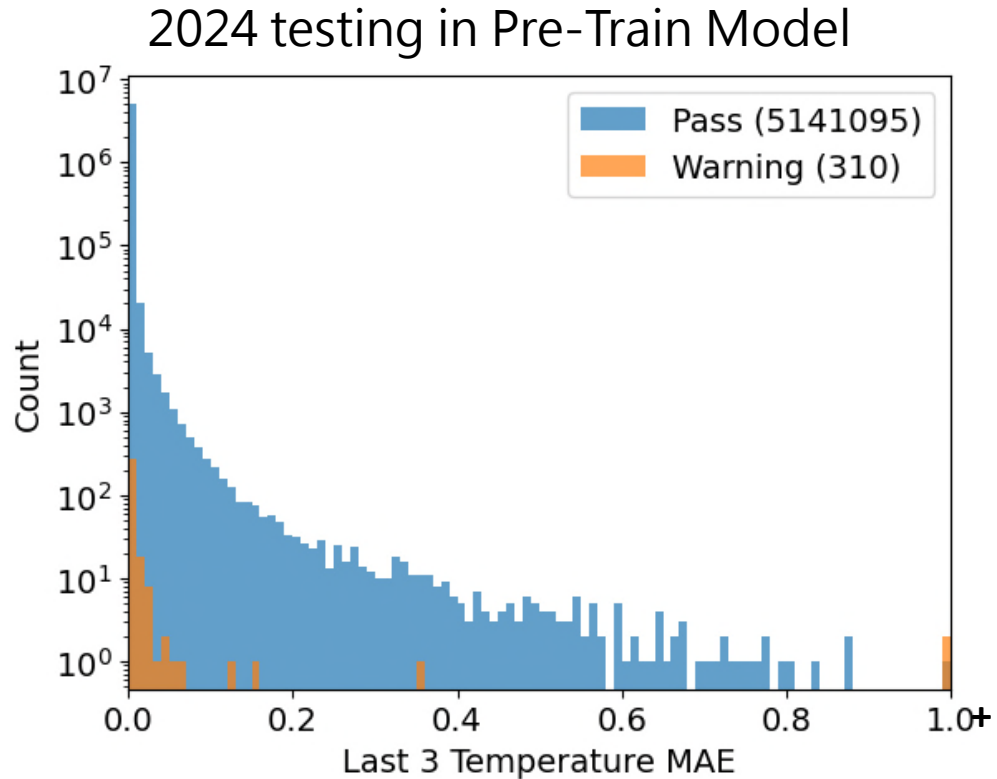
■ 新版加入濕度的逐時、10分鐘、1分鐘AE模型

- ◆ 皆先使用一整年所有測站的溫濕度資料建立pre-train模型。
- ◆ 輸入16筆序列資料，濕度標準化到0~1。
- ◆ 再依照測站特性遷移學習。



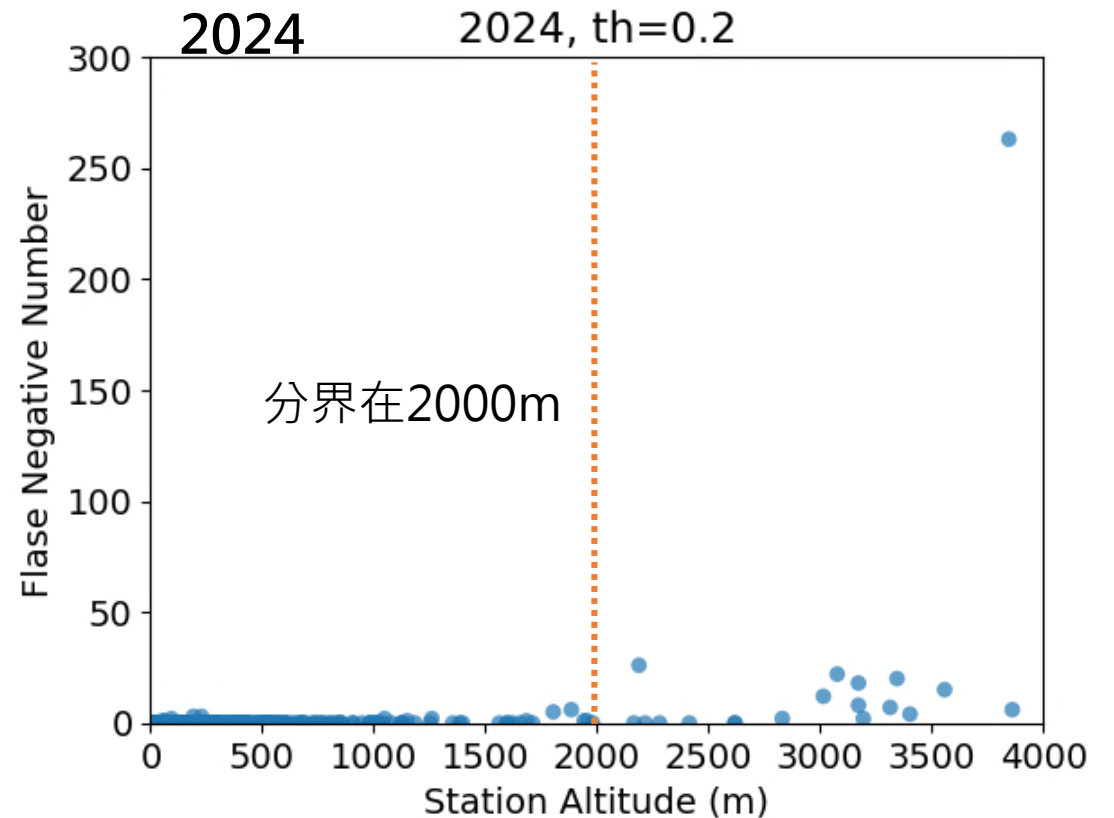
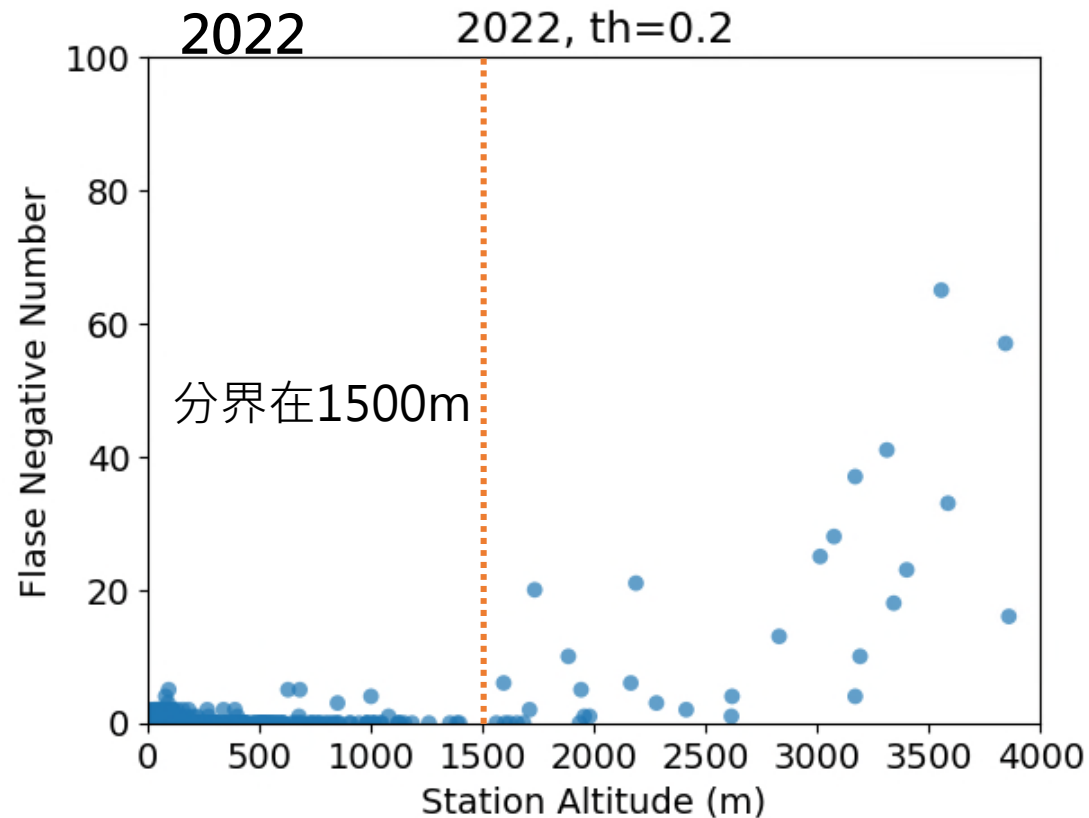
逐時與10分鐘AE檢核門檻設定

- 檢核最後3筆溫度之MAE(L3MAE)
- 檢核門檻以逐時AE為例，依測試年份誤判數測站分布定為0.2。



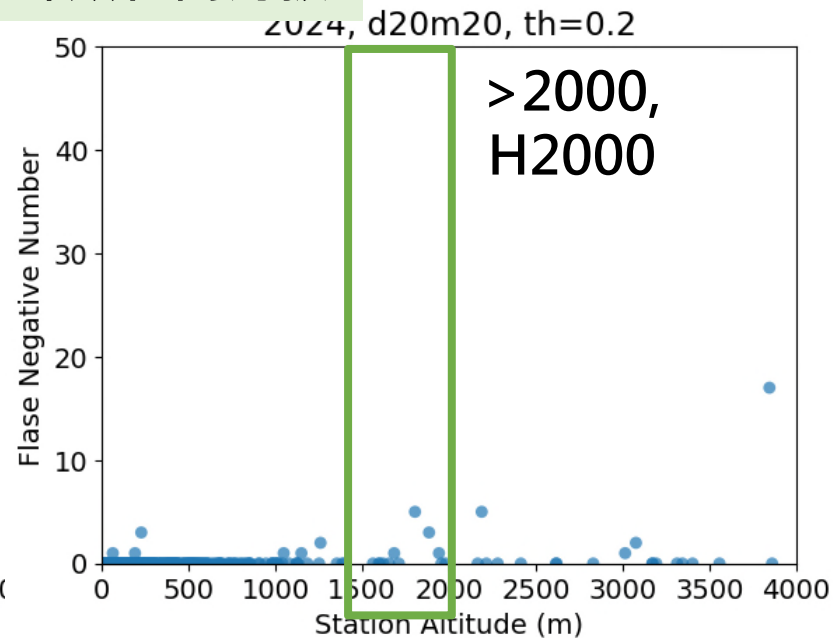
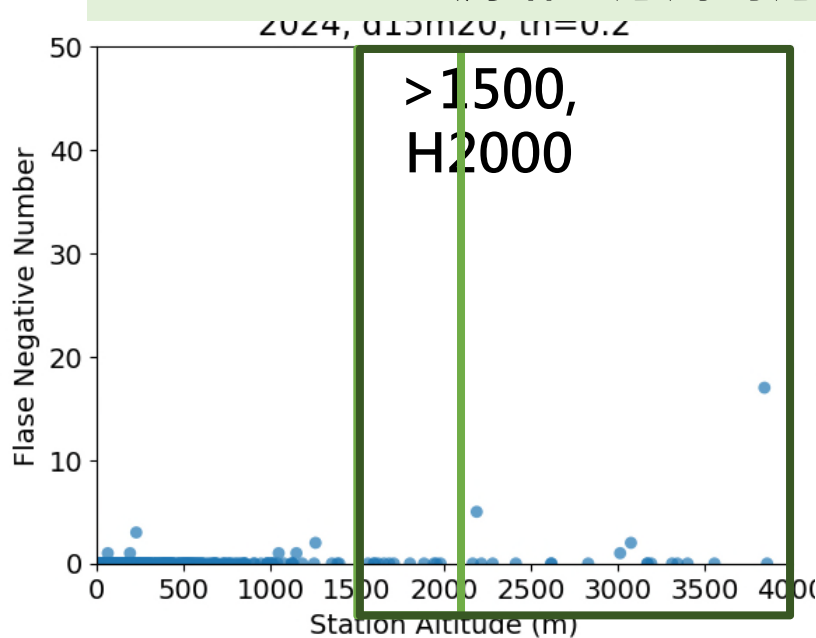
1HR Pre-Train AE測試

將2023年全部測站訓練之pre-train模型進行transfer learning，
以2000 m為分界→ H2000/L2000模型；
以1500 m為分界→ H1500/L1500模型。



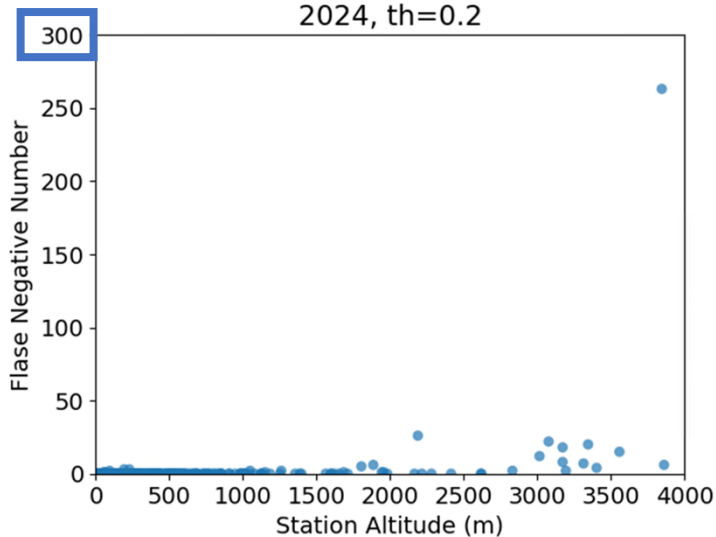
逐時AE依高度遷移學習後表現

1500~2000m測站，應用時應歸類在高海拔

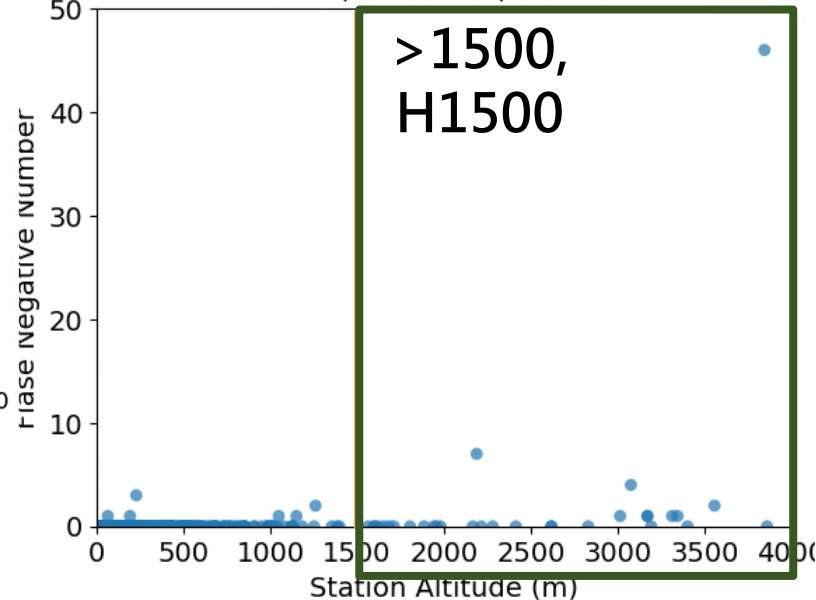


Pre-Train model

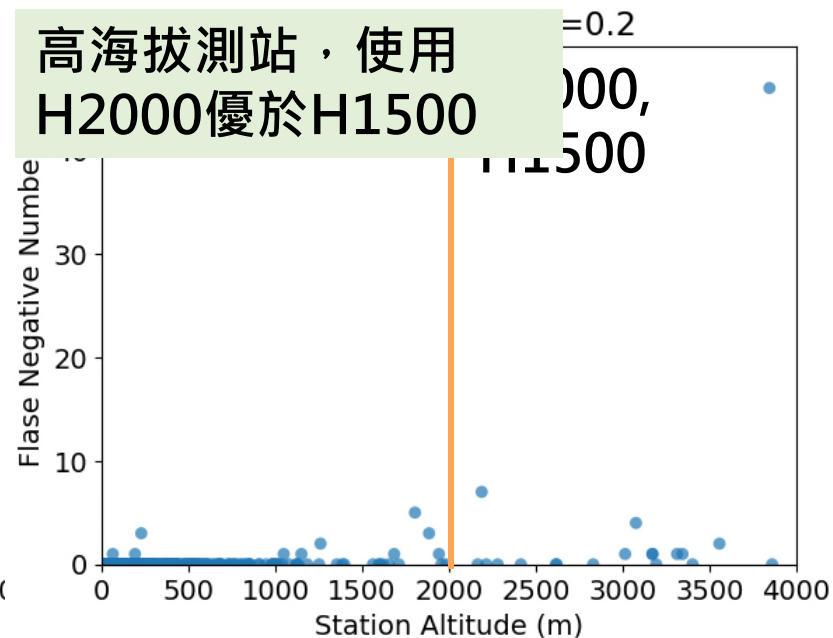
2024, th=0.2



2024, d15m15, th=0.2

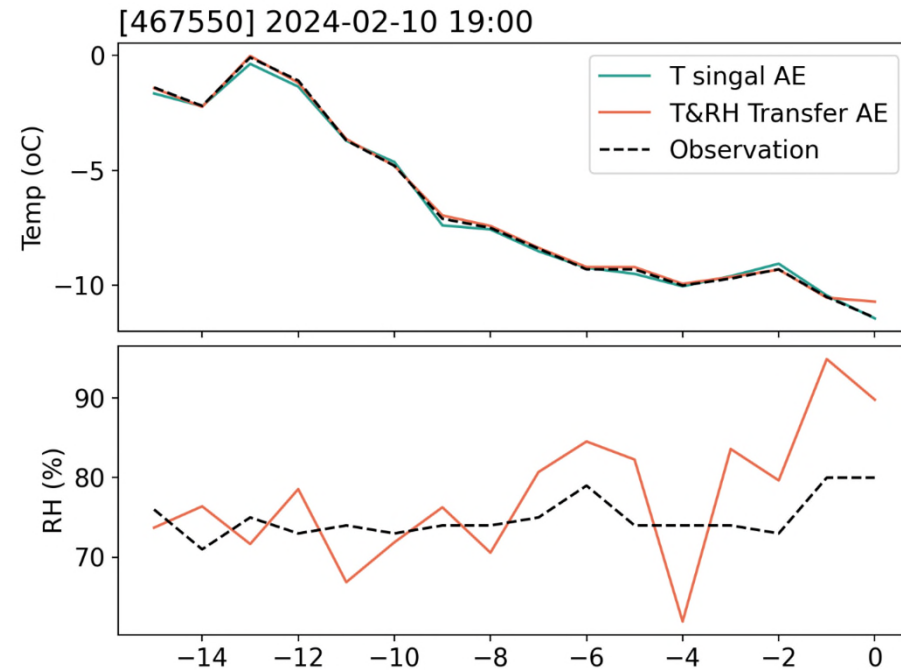
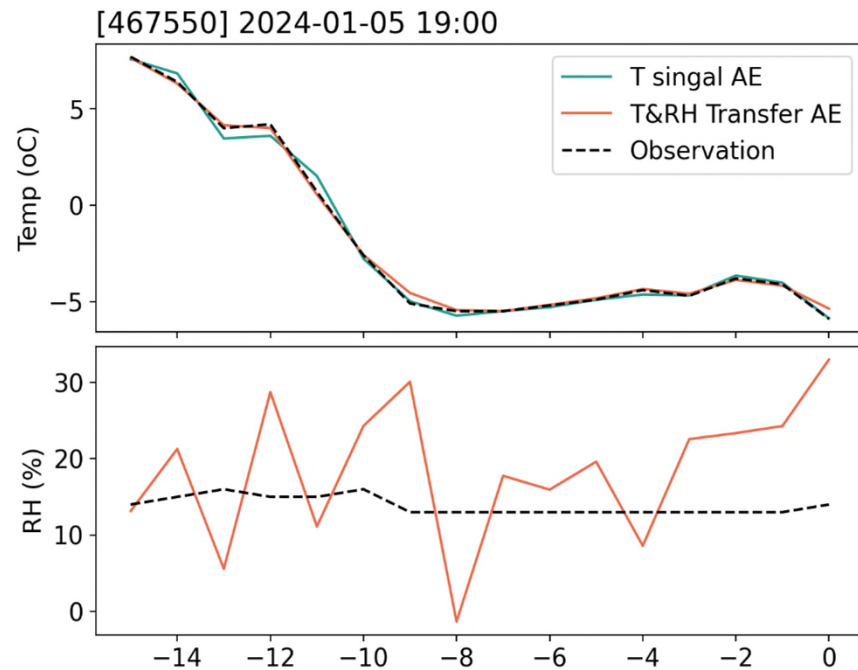


高海拔測站，使用 H2000優於H1500



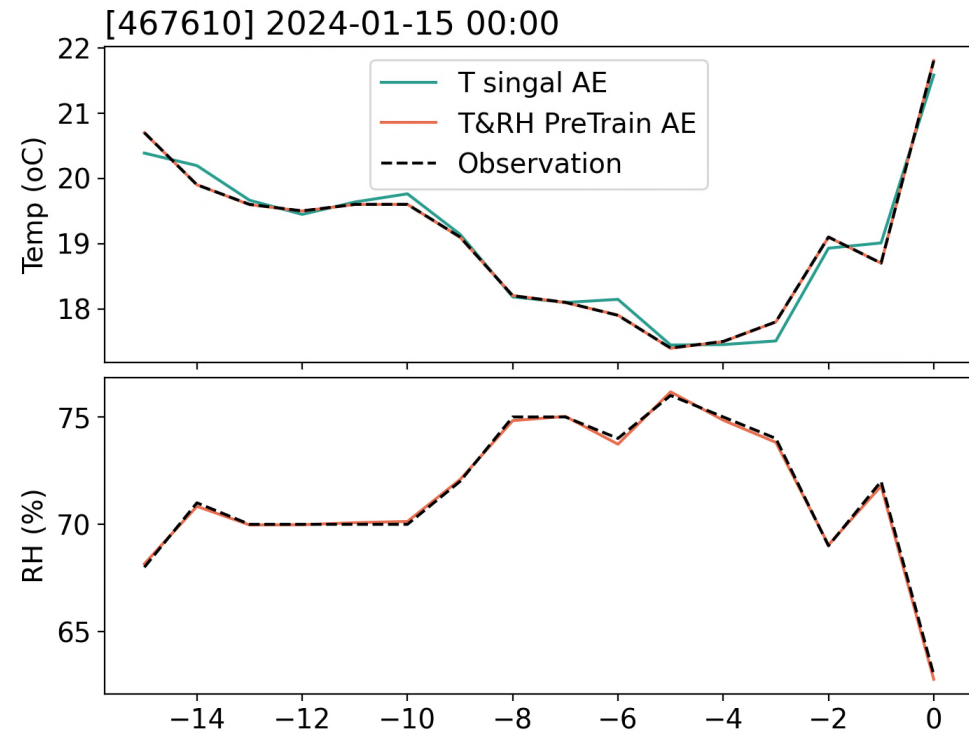
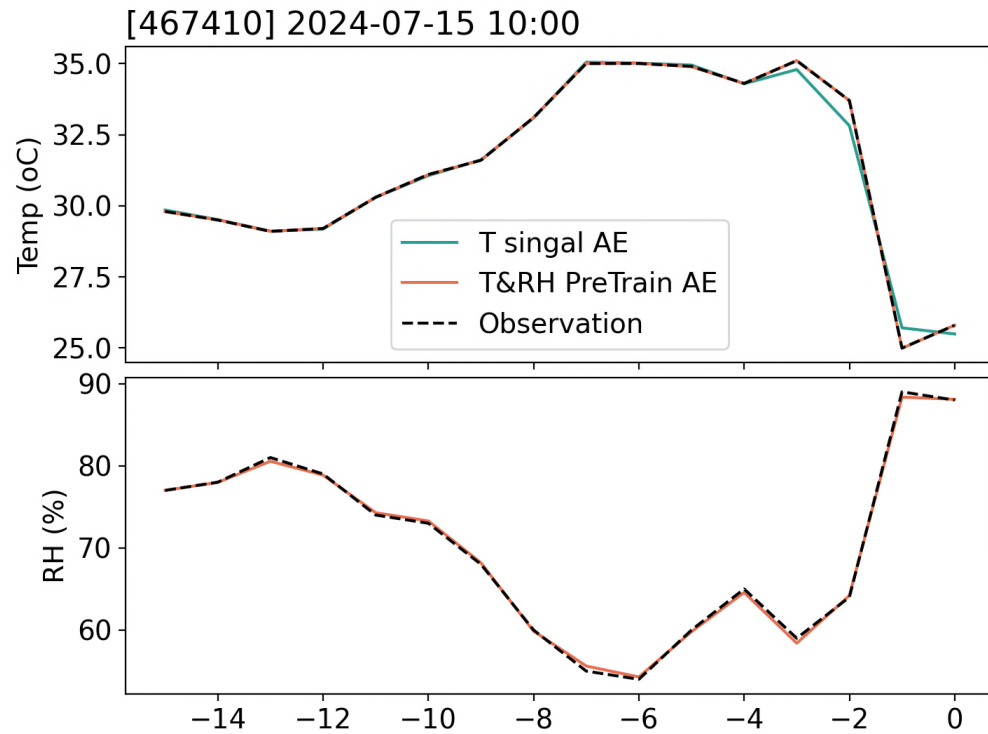
2024年逐時AE誤判個案數高之測站

- 2024年467550 (玉山)測站不通過H2000有多筆連續提列為同一個案。
- 玉山站疑似誤判個案在1~3月間，應為2023年冬季最低溫高於2024年之影響。
- 玉山考慮使用以前建置之單站溫度AE模型



Pre-Train模型與單測站模型比較

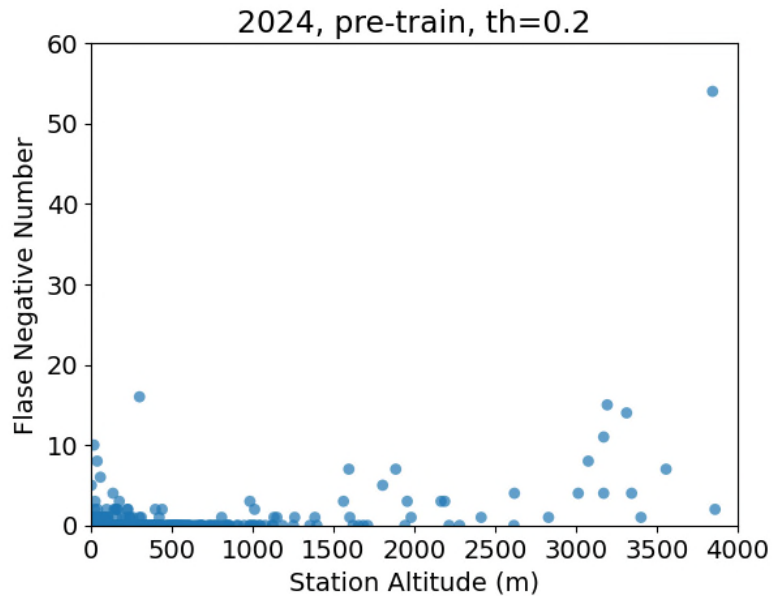
- 比較舊版單測站溫度AE與新版溫溼度AE通過門檻結果不相同的個數。
- 新版溫溼度AE擬合表現優於舊單測站溫度AE。



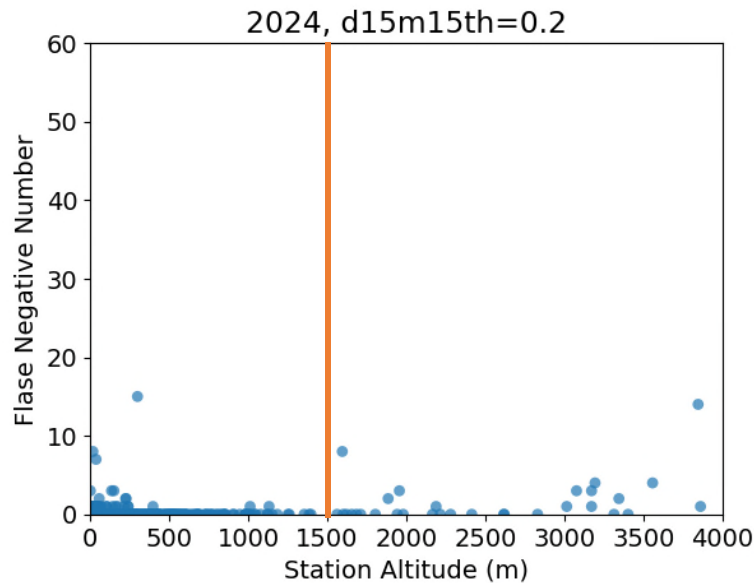
10分鐘AE依高度遷移學習後表現

- 10分鐘溫度序列測站分界也在1500 m。
- 高海拔使用H2000表現略優於H1500，但差異不明顯。
- 分高低海拔遷移後中高海拔測站FNN有明顯下降。

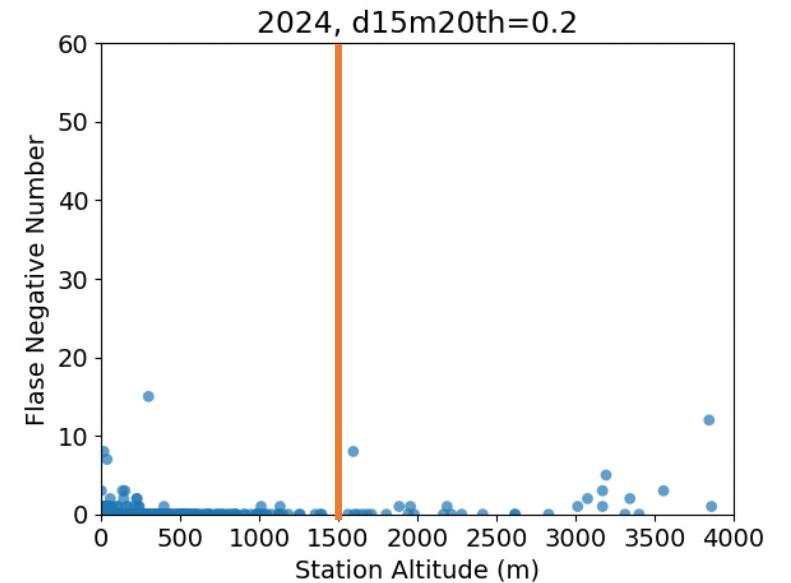
Pre-Train model



> 1500, H1500



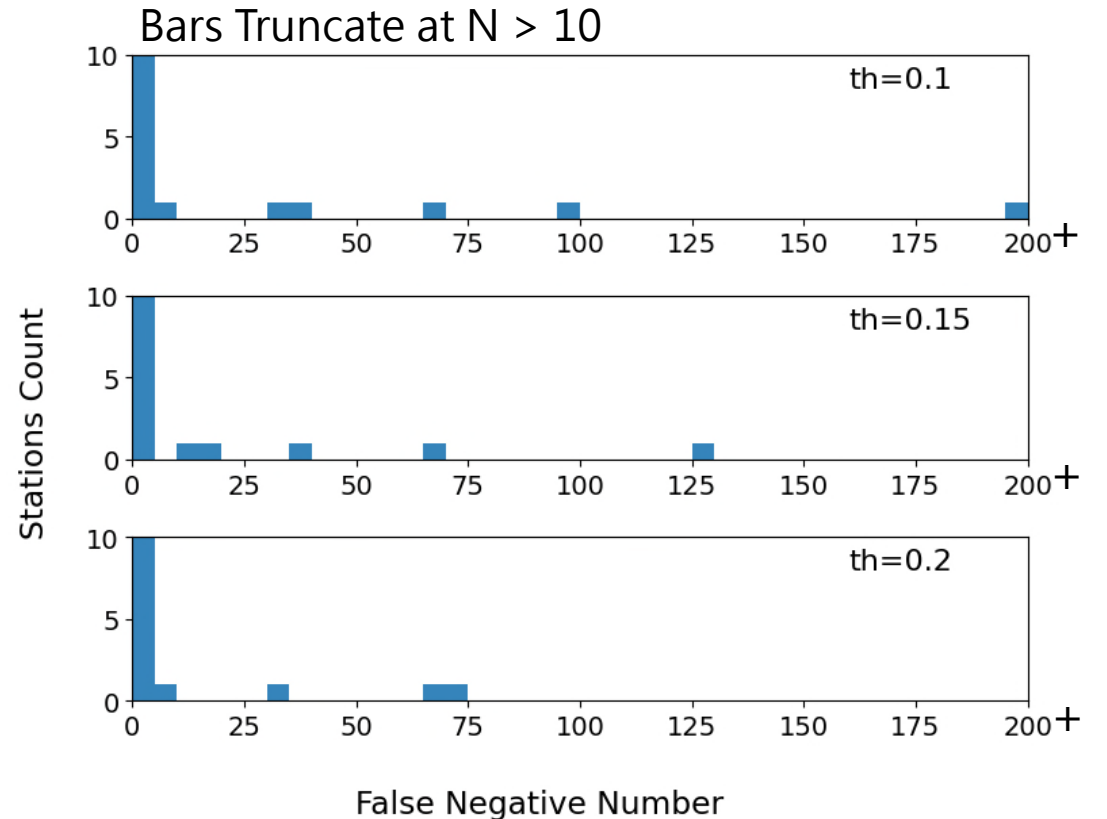
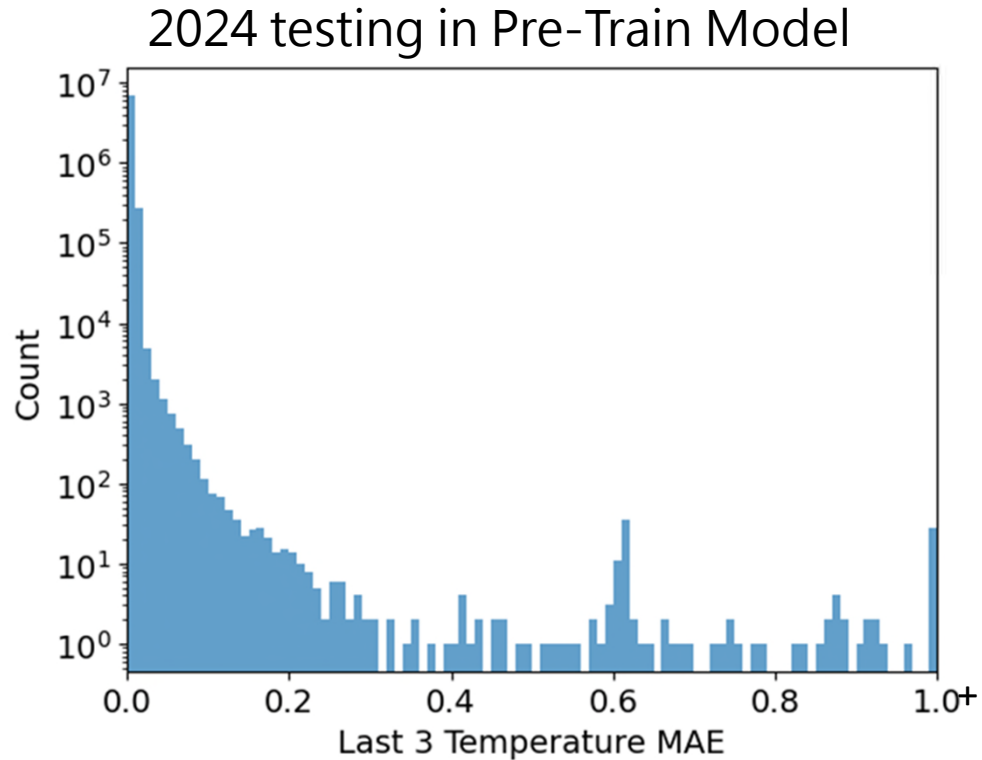
> 1500, H2000



2024

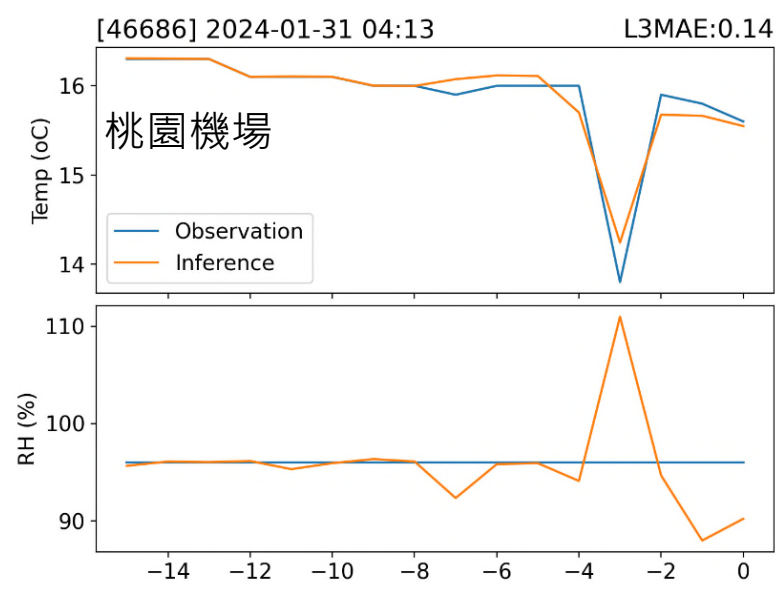
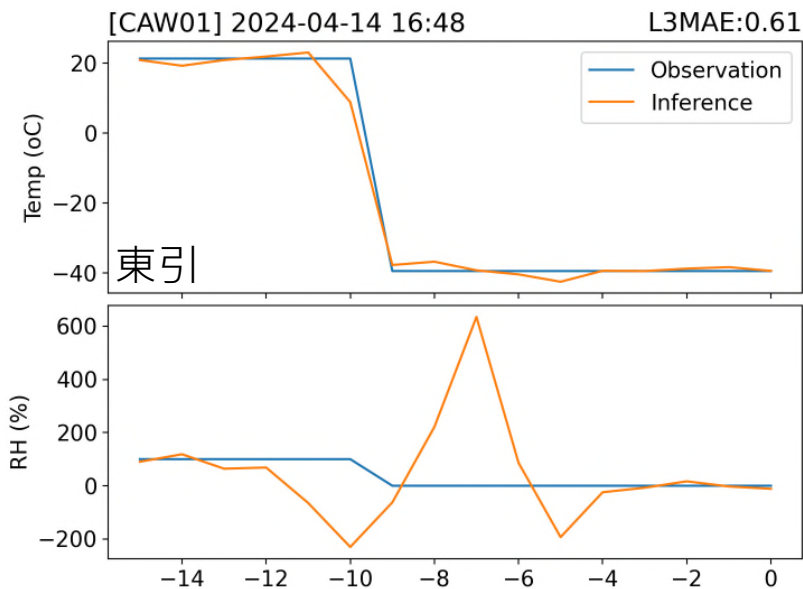
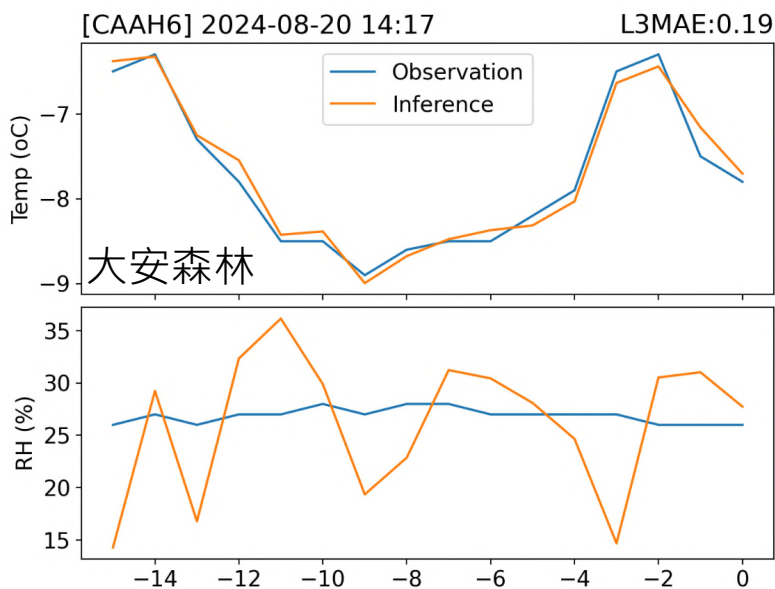
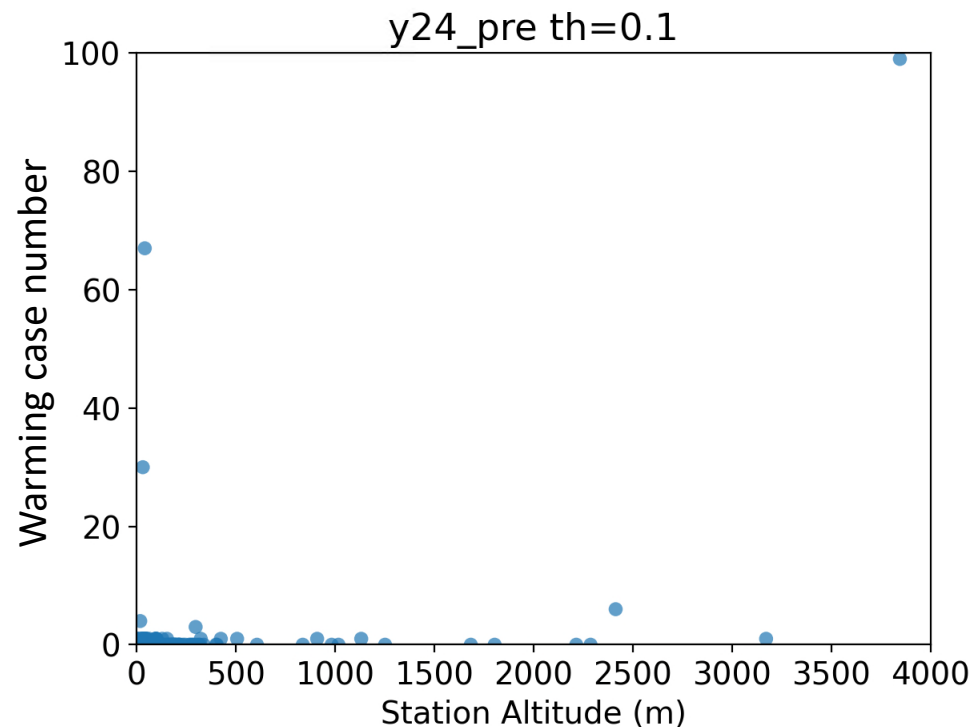
1分鐘AE檢核門檻設定

- 檢核最後3筆溫度之MAE(L3MAE)
- 依測試年份誤判數測站分布，定門檻為0.1。



1min Pre-Train AE之 2024年測試結果

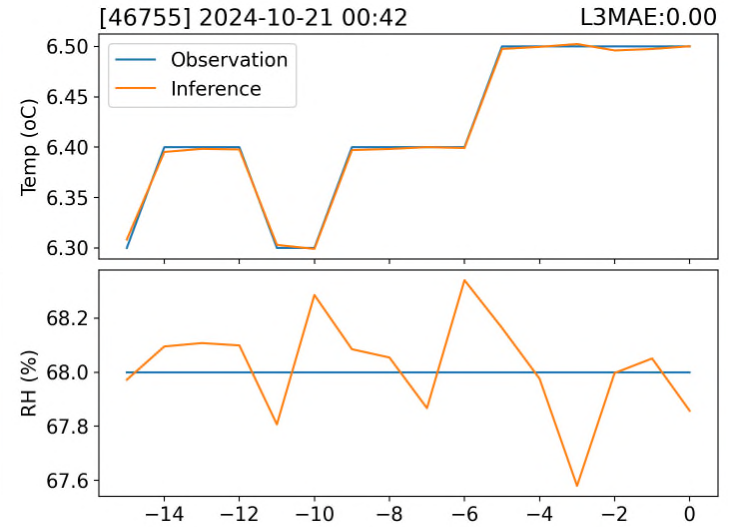
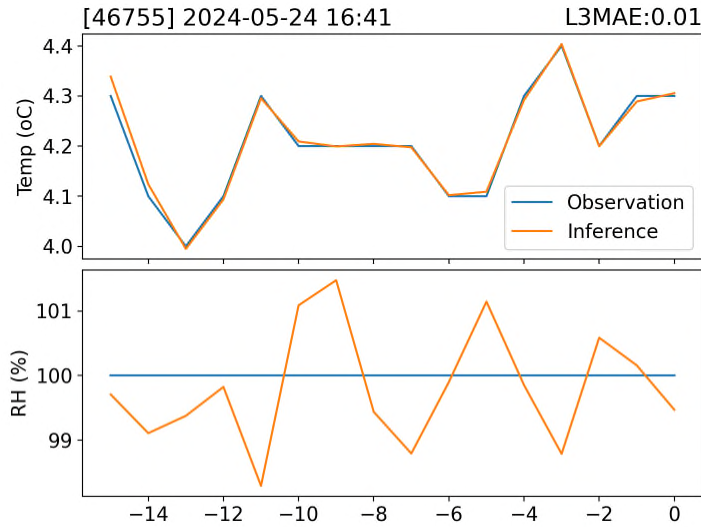
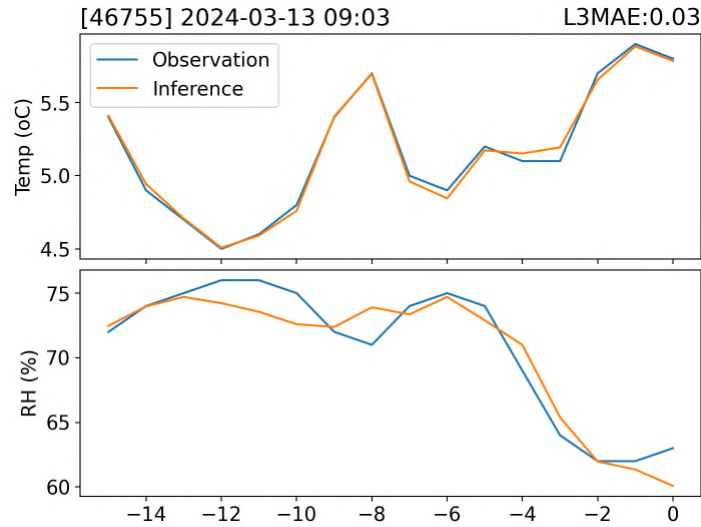
- ◆ Pre-Train模型擬合良好，不須再遷移學習
- ◆ 除玉山外，提列數量高之測站皆為異常個案



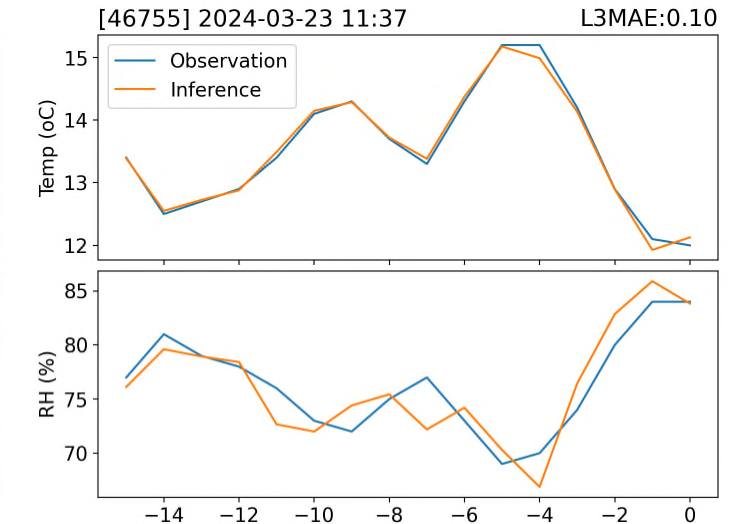
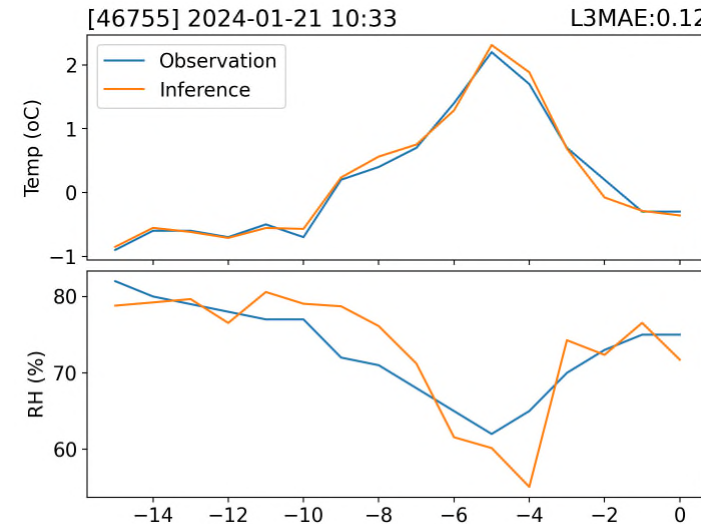
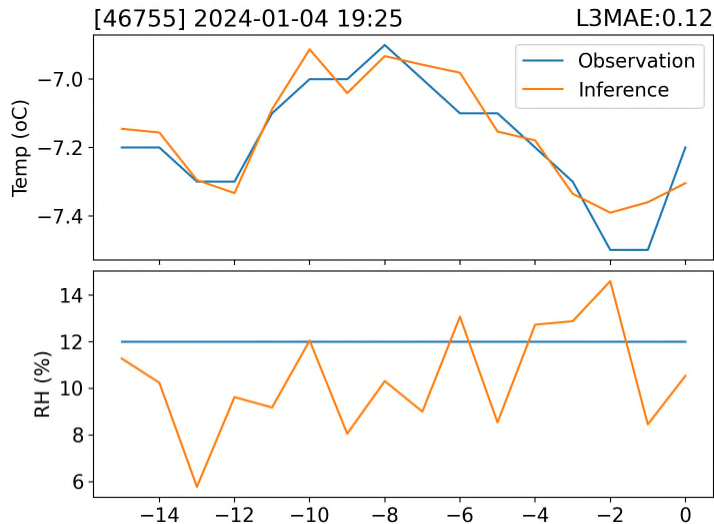
46755 玉山正常與提列個案

被提列個案為誤判可能性高，
考慮使用去年建立的單測站溫度AE。

正常



提列



結論

■ 逐時AE、10分鐘AE

- ◆ 溫溼度AE模型訓練時，以**2000**公尺為界分別訓練高/低海拔模型。
- ◆ 應用於溫度檢核時，則以**1500**公尺為界分別挑選高/低海拔模型使用。
- ◆ 玉山使用單測站溫度AE模型。
- ◆ 檢核門檻 $L3MAE > 0.2$

■ 逐分鐘AE

- ◆ 測站不分高度使用單一Pre-Train模型。
- ◆ 玉山使用單測站溫度AE模型。
- ◆ 檢核門檻 $L3MAE > 0.1$ 。