

機器學習技術建立現地型地震預警系統相關研究(II)

子計畫一

以類神經網路技術開發現地型地震最大震度即時預測系統

金台齡 李世璽 吳祐甄 尤俊昇

國立臺灣科技大學

1 中文摘要

臺灣地震頻率高並經常伴隨破壞性地震，而造成嚴重的損失。為了有效降低這些地震造成的災害，地震預警系統成為一種重要的防災工具。地震預警系統能夠在強震波發生之前提前發出警報，讓人們有更多時間採取適當的措施以減少災害。

地震預警系統主要分為兩種類型：現地型預警和區域型預警。現地型預警系統專注於提供震央附近區域的預警，以解決近震央地區的需求。區域型預警系統則利用多個近震央測站的數據，以檢測到的強烈地震波為依據，向較遠的地區發送預警信息。這種區域型預警系統確保了重要地區能夠提前收到地震警報，進一步降低災害風險。

本研究採用基於神經網路的方法，通過分析震央附近測站所記錄的地震波形數據，來預測目標測站的峰值地面加速度(PGA)。這項研究旨在運用神經網路技術實現區域型地震預警系統，使我們能夠更準確預測地震的強度，並為相關區域提供更早的警報，有助於降低地震帶來的災害風險。此研究期望能為地震預警技術的發展和實施提供有力的支持。

關鍵詞：類神經網路、地震預警、加速度預測

2 Abstract

The frequency of earthquake occurred in Taiwan is high and it often accompanied by destructive seismic events resulting in significant losses. In order to mitigate the disasters caused by these earthquakes, earthquake early warning (EEW) systems have emerged as crucial tools for disaster preparedness. EEW systems can issue warnings before the arrival of strong seismic waves, providing people with valuable time to take appropriate actions to reduce the impact of these disasters.

Earthquake early warning systems are primarily categorized into two types: on-site and regional. On-site Earthquake Early Warning Systems (EEWS) are designed to

provide alerts for areas close to the earthquake epicenter, addressing the immediate needs of those in proximity to the epicenter. In contrast, Regional EEWs leverage data from multiple nearby seismic stations to detect strong seismic waves and broadcast warnings to more distant areas. Regional EEWs ensure that critical regions receive early earthquake alerts, further reducing the risk of disasters.

This research employs a neural network-based approach, analyzing seismic waveform data recorded by stations near the earthquake epicenter to predict the peak ground acceleration (PGA) at a target station. The aim of this study is to utilize neural network technology to realize a regional earthquake early warning system, enhancing our ability to accurately forecast earthquake intensity and provide earlier warnings to relevant regions. This research holds the potential to provide robust support for the development and implementation of earthquake early warning technology.

Keywords: Neural Network、Earthquake early warning、PGA prediction

3 前言

地震一直以來都是地球上最具挑戰性的自然災害之一，其突發性和破壞性經常對人們的生命和財產造成重大風險。特別是在臺灣這樣一個地震活躍的地區，地震災害的頻率和嚴重性對我們的社會和經濟穩定產生了長期的威脅。

這些地震造成的損失不僅對個人和家庭造成了嚴重的影響，也對整個社會和經濟系統帶來了巨大的衝擊。因此，如何有效地減少地震所造成的災害，一直是人們關注的焦點。地震預警系統就是一個非常有前途且可以有效的解決方案，能夠在地震發生前對各地區發出警告，將地震所帶來的傷害最小化。

為了應對這一挑戰，本研究專注於利用神經網路技術來提高地震預警系統的性能，尤其是區域型預警系統。我們將通過分析地震波形數據，透過神經網路預測地震的強度，以實現更精確的預警。這個研究的目標是為提高地震預警技術的效益，從而更好地保護我們的社區和資產免受地震帶來的損害。

這份研究報告將深入探討我們所採用的方法、數據和實驗結果，並探討其在地震預警領域的潛在應用。我們期望這個研究能夠為地震預警技術的不斷改進和發展提供有價值的信息，並為地震災害風險的降低做出貢獻。

4 類神經網路應用於區域型地震預警

4.1 簡介

本研究是利用 P 波前 5 秒到 P 波後數秒的三軸加速度波形資料，結合類神經網路的技術，來預測模型的目標測站，此研究目的是希望利用近震央的幾個測站，能在地震發生後第一報之前快速且準確地預測出其他縣市區域的 PGA，且希望模型能 real-time 預測，隨著地震事件時間的推移，模型能因為得到更多波形資訊而預測得更準確，以達到區域型預警的效果。

目前採用的模型是利用圖卷積神經網路，考慮測站的位置以及測站間的距離關係，地震信號在傳遞時，相鄰測站距離若越小，信號在傳遞到兩測站的距離和時間也會較接近，PGA 可能會越相近，距離越大則反之，以這樣的空間資訊再結合原本時間序的波形資訊，讓模型依據這些特徵來預測出目標測站的 PGA。且為了要讓模型更容易的學習地震發生時，不同地區之間的能量衰減關係，因此模型加入了分區的概念，將偵測波形的測站定義在花蓮地區，輸出的目標測站定義在臺灣其他地區，此目的是為了讓模型更有效的學習花蓮與分別其他地區的地震波能量衰減關係，模型設計概念圖如圖 1，我們利用近震央的花蓮測站所偵測到的波形，透過測站間的距離，如圖 1 橘色的邊，讓模型學習地震波能量和距離的關係，最後會輸出目標測站的 PGA。

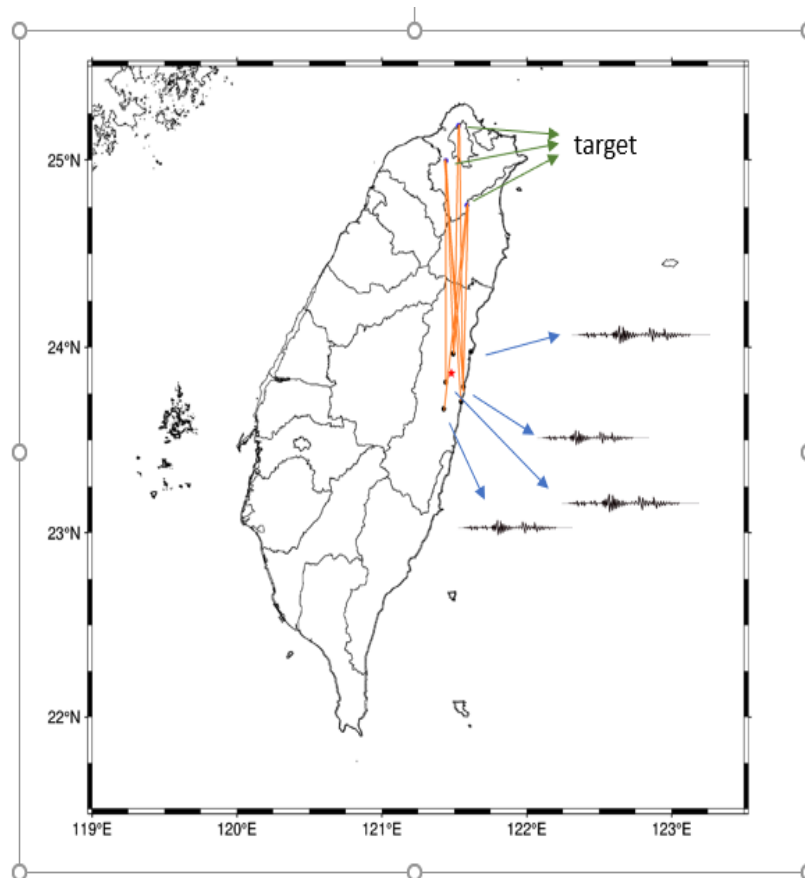


圖 1、模型概念圖

4.2 資料處理

本研究使用的資料集來自中央氣象署 CWA 觀測網中，測站分布如圖 2 所示，以綠色表示 21 個偵測測站及藍色的 44 個目標測站。資料集分為訓練集和測試集，訓練集時間從 2012 年到 2020 年且震央發生在花蓮地區的花蓮事件，測試集時間從 2021 年到 2023 年且震央發生在花蓮地區的花蓮事件。

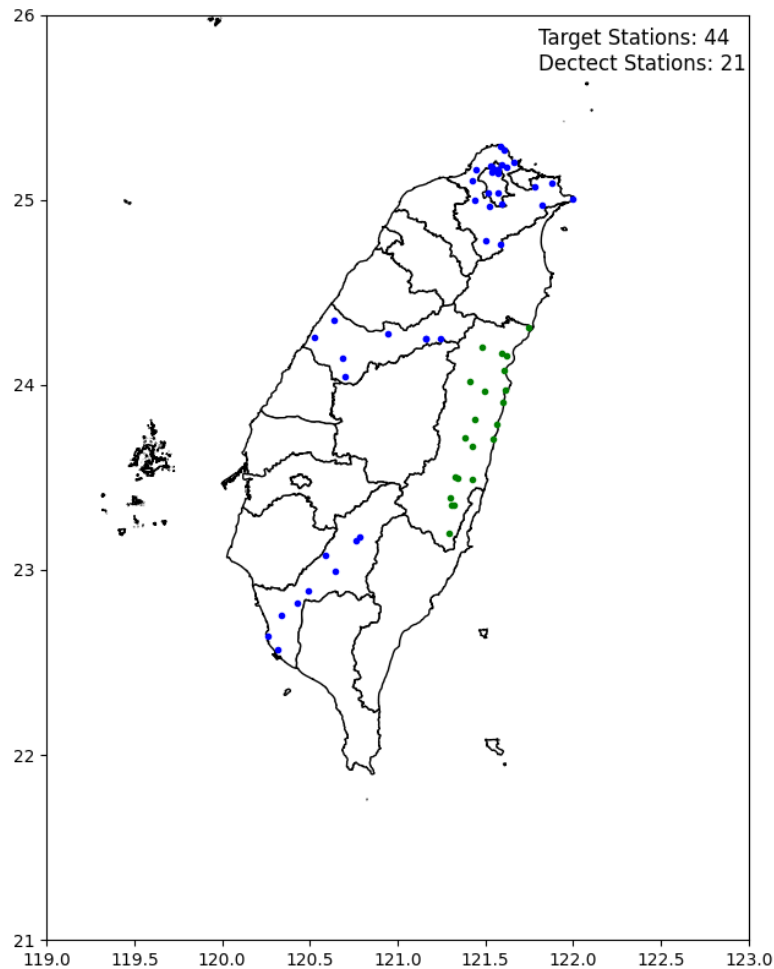


圖 2、CWA 測站分布圖

在本次研究中，我們所使用的地震事件在訓練集規模 1-2 有 7 筆，規模 2-3 有 601 筆，規模 3-4 有 1300 筆，規模 4-5 有 470 筆，規模 5 以上有 72 筆，總共 2450 筆地震事件，訓練集的規模和震源深度的範圍分佈及其規模和深度分布如圖 3 和圖 4 所示，測試集的規模 1-2 有 98 筆，規模 2-3 有 310 筆，規模 3-4 有 225 筆，規模 4-5 有 168 筆，規模 5 以上有 44 筆，總共 845 筆地震事件，規模和震源深度的範圍分佈如下圖 5 和圖 6，震央分佈圖如圖 7 和圖 8，紅點越大代表規模越大，為了讓模型能夠 real-time 的預測，我們會對資料集進行資料增強(data augmentation)

將時間點 T 之後的三軸加速度波形補 0，形成一個 30 秒時間窗的波形，以匹配我們模型的輸入。

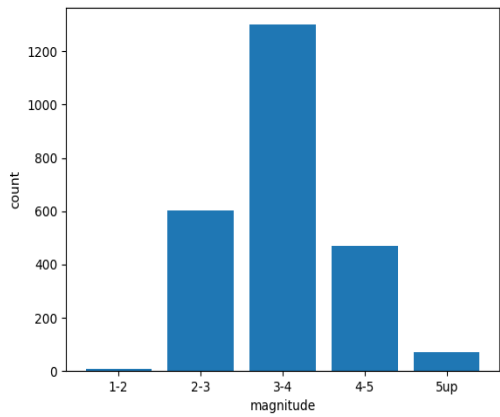


圖 3、訓練集規模分佈圖

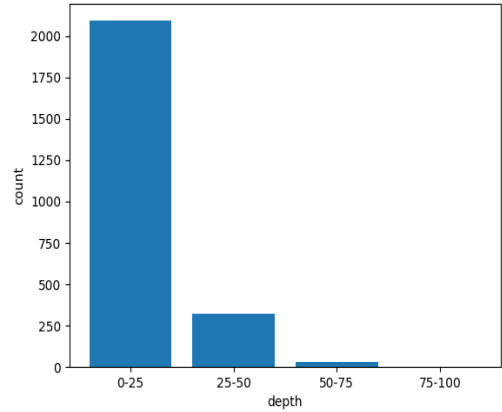


圖 4、訓練集深度分佈圖

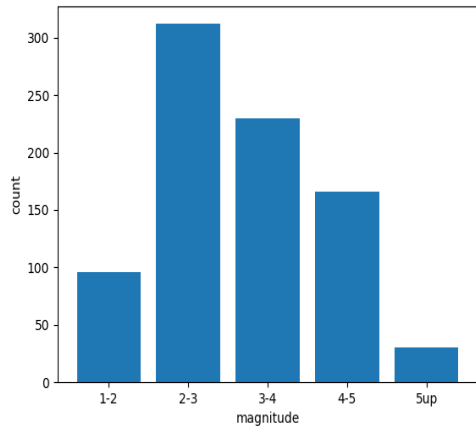


圖 5、測試集規模分佈圖

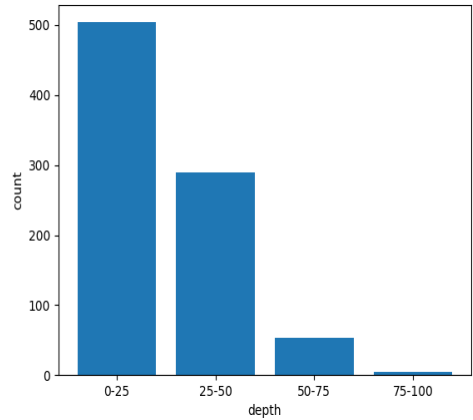


圖 6、測試集深度分佈圖

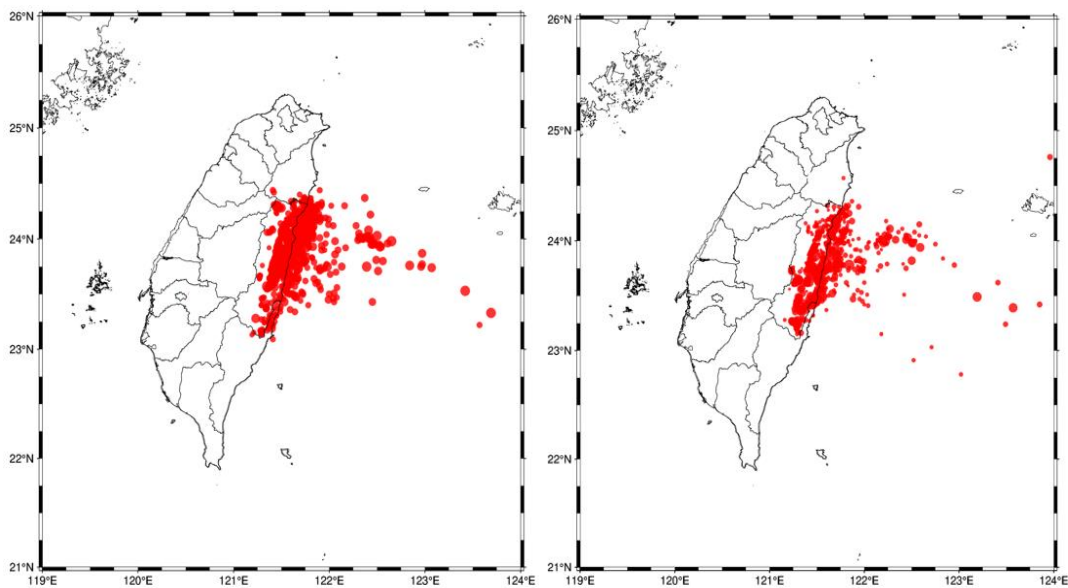


圖 7、訓練集震央分佈圖

圖 8、測試集震央分佈圖

4.3 模型架構

本研究的模型架構圖如圖 9，由輸入、卷積神經網路(Convolutional Neural Network)、Position embedding、圖卷積神經網路(Graph Convolutional Neural Network) [1][2]和輸出主要的五個部分組成。

首先，因為我們的模型有分區，所以模型的輸入會分成兩邊，detect 和 target 左右兩個部分，detect 部分的輸入有花蓮測站的 30 秒三軸加速度波形和花蓮測站經緯度，target 部分的輸入則只有目標測站的經緯度，左邊會先將花蓮測站的三軸加速度波形利用 2 層 2 維卷積神經網路(Convolutional neural network)和 4 層 1 維卷積神經網路(Convolutional neural network)進行特徵提取，kernel size 依序分別為 (1,1)、(1,3)、16、16、8、4，filter 分別為 8、32、64、128、32、16，第三和四層的卷積層後會接最大池化層(Max pooling 和加入機率為 0.1 的 Dropout，每層的卷積神經網路後都會接整流線性單位函數(Rectified Linear Unit, ReLU)，再來模型會通過 2 層全連接層(Fully Connected Layer)，units 分別為 1024 和 400，並在第一層全連接層後加入機率為 0.2 的 Dropout，最後形成具有各測站波形的特徵矩陣 M ， $M \in \mathbb{R}^{n \times d}$ ， n 為測站數， d 為特徵維度。

針對位置資訊，我們會將花蓮測站的經緯度做 Position embedding[3][4]，將偵測測站經緯度分別映射到一個 d 維的向量，形成一個矩陣 P ， $P \in \mathbb{R}^{n \times d}$ ， n 為測站數， d 為特徵維度，提取偵測測站位置的特徵後，再將加速度波形和測站位置的特徵相加，如下式 1，形成具有各測站的混合特徵矩陣 D ， $D \in \mathbb{R}^{n \times d}$ ， n 為測站數， d 為特徵維度。右邊則為我們目標輸出的測站，不會有波形資料輸入，只有測站位置的資訊，因此，我們將目標測站的經緯度做 Position embedding，將目標測站經緯度分別映射到一個 d 維的向量，形成一個矩陣 T ， $T \in \mathbb{R}^{k \times d}$ ， k 為測站數， d 為特徵維度，提取目標測站位置的特徵。

接著，我們會將花蓮和目標測站的特徵合併，如下式 2，形成一個具有各測站特徵的矩陣 X ， $X \in \mathbb{R}^{(n+k) \times d}$ ，再來，我們會考慮各測站間的距離關係，所以我們會計算花蓮和目標測站間的距離，並以 Adjacency matrix 的方式表示，計算方式為先利用經緯度計算兩兩測站間的距離，再對此矩陣做正規化，依照距離的大小比例，給予不同的權重，接著會將 Adjacency matrix 轉換成 Laplacian matrix，以利模型的訓練，轉換方式如下式 3， A 為 Adjacency matrix， I 為 Identity matrix， D 為 Degree matrix，Degree matrix 為一個對角矩陣，對角線為相鄰測站的個數[5]。接著會將特徵矩陣和權重矩陣透過四層圖卷積神經網路(Graph Convolutional Neural Network)學習測站之間的距離關係和地震波能量衰減與距離的關係，圖卷積神經網路的 Propagation rule 如下式 4， $H^{(l)}$ 為第 l 層的特徵矩陣， L 為由式 3 轉換成的拉普拉斯矩陣， W 為模型要學習的權重矩陣， θ 為激活函數(Activation Function)[6]，圖卷積神經網路(Graph Convolutional Neural Network)的核心操作是圖卷積層(Graph Convolutional Layer)，它通過將節點的特徵與鄰近節點的特徵進行加權總和，然後進行非線性變換來更新節點的特徵表示。這種加權總和的計算可以通過圖結構中的鄰接矩陣進行，通過多層的圖卷積層堆疊，模型可以學習到更複雜的空間上的特徵，每層的 filter 皆為 16，且每層的圖卷神經網路後皆會接整流線性單位函數(Rectified Linear Unit, ReLU)，最後將目標測站的特徵通過一個全連接層(Fully Connected Layer)進行 regression，輸出目標測站的 PGA。

$$D_{n \times d} = M + P \quad (\text{式 1})$$

$$X_{(n+k) \times d} = [D \ T] \quad (\text{式 2})$$

$$L = I - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \quad (\text{式 3})$$

$$H^{(l+1)} = f(H^{(l)}, L) \quad (\text{式 4})$$

$$= \partial(L \times H^{(l)} \times W^{(l)}) , H^{(0)} = X$$

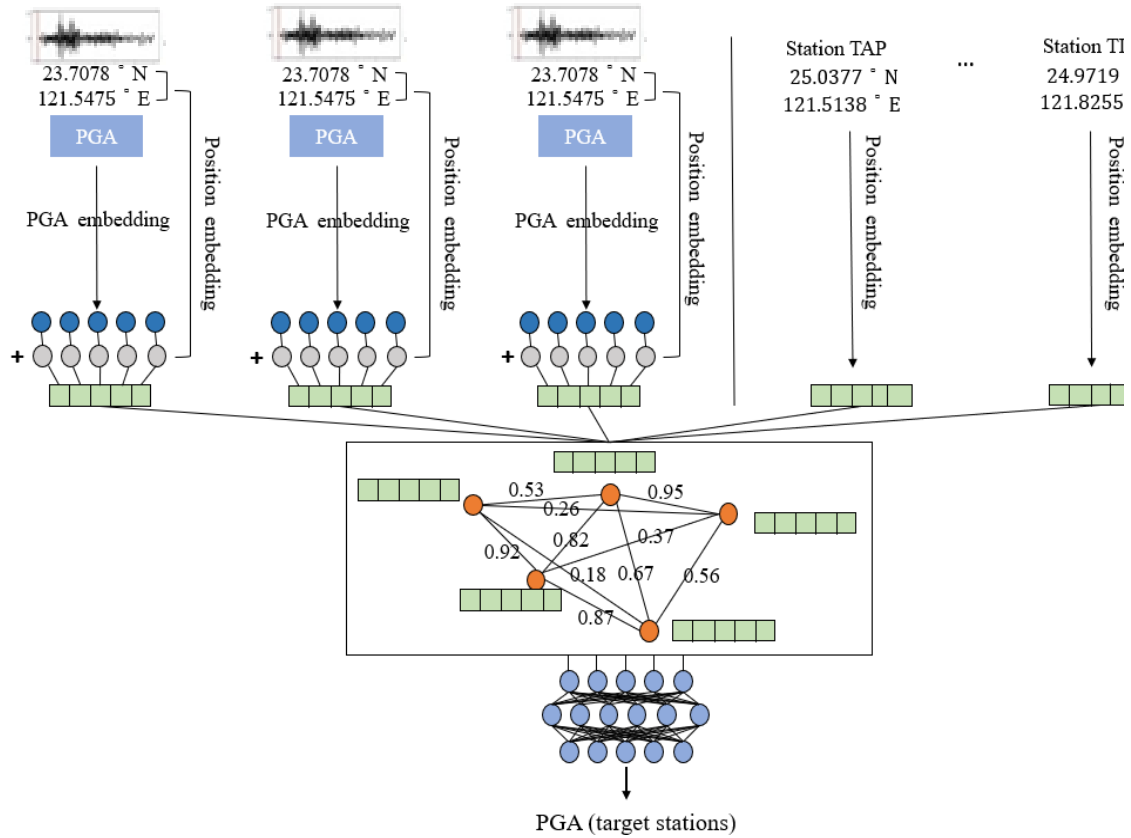


圖 9、模型架構圖

4.4 訓練過程

在本研究的訓練過程中所使用的優化器是 Adaptive Moment Estimation (Adam)，其在更新梯度時會考慮偏差校正，使每一次的學習率都有一個確定的範圍，讓參數的更新較為平穩。學習率是設定為 10^{-5} ，batch size 為 64，總訓練次數為 200 epoch，decay 為 10^{-7} ，EarlyStopping 的 patience 設為 20，避免模型過度擬合，Loss function 為均方誤差 (Mean Square Error, MSE)，其為實際值和預測值的平方差之平均。將偵

測到的三軸加速度波形輸入到模型後進行訓練，模型輸出目標測站的預測 PGA，依據 Loss function 對模型進行更新和調整，優化模型預測結果。

4.5 具體成果

4.5.1 第 026 號花蓮地震事件預警測試

在此實驗中，我們利用 P 波前 5 秒到當前時間點 T 的波形的當作模型輸入對此事件做評估，且我們訓練了三個分區模型，分別對臺北、臺中、高雄三個地區測試，如下圖 10，紅色框為偵測區域，橘色框為目標區域。選取 2021 年 4 月 18 日 22 時 14 分 37 秒，規模 6.2，深度 14.4 公里，震央位置位於北緯 23.86 度、東經 121.48 度的第 026 號花蓮地震。

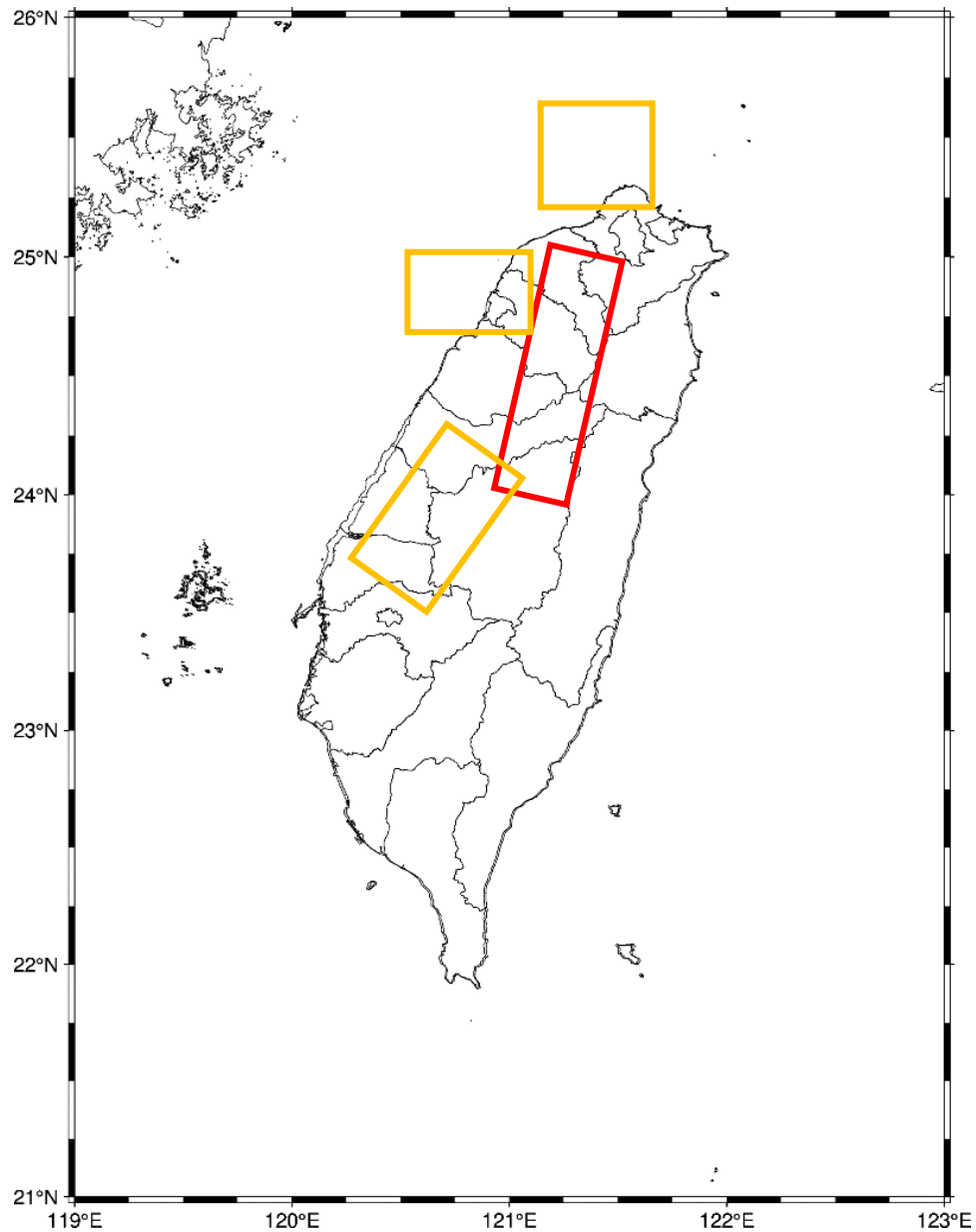


圖 10、模型分區圖

首先，我們先對臺北地區做測試，一開始模型使用第一個被 Pick 到 P 波的 ESL 測站的三軸加速度波形資料來做預測，預測結果如下圖 11，與圖 17 的真實 PGA 分布圖比較，可以看到還有明顯的低估，主要是因為地震波剛到，只使用了 1 個測站的波形資訊來做預測，模型所接收到的資訊太少，才会有此結果產生。再來，地震經過 3 秒後，花蓮地區有 6 個測站有接收到波形資料，因此，我們使用這 6 個測站的三軸加速度波形資料來做預測，預測結果如下圖 12，和圖 1 相比，模型在地震發生 3 秒後因為在接收到比較多的資訊，所以在預測的表現上也有提升。

在地震經過 5 秒後，有 11 個測站接收到地震波，因此，我們使用花蓮地區 11 個測站所接收到的三軸加速度波型當作模型的輸入來做預測，跟前一個 3 秒的結果相比，第 5 秒的預測結果離真實 PGA 值又更接近。地震發生 8 秒後，花蓮地區有 20 個測站接收到波形，我們使用這 20 個測站的資訊來預測，預測結果如下圖 14。再來圖 15、圖 16 皆有使用花蓮地區 21 個測站的波形資料預測出的結果，分別是地震發生後 10 秒以及 15 秒。事件發生大約 15 秒後新北地區的 ANP 測站為北部第一個 Pick 到地震波的測站，而我們的模型能在 15 秒內預測出與真實 PGA 相近的結果，由此可得知，此模型可以在地震波還沒到達北部區域時，預測出相近的 PGA 值。

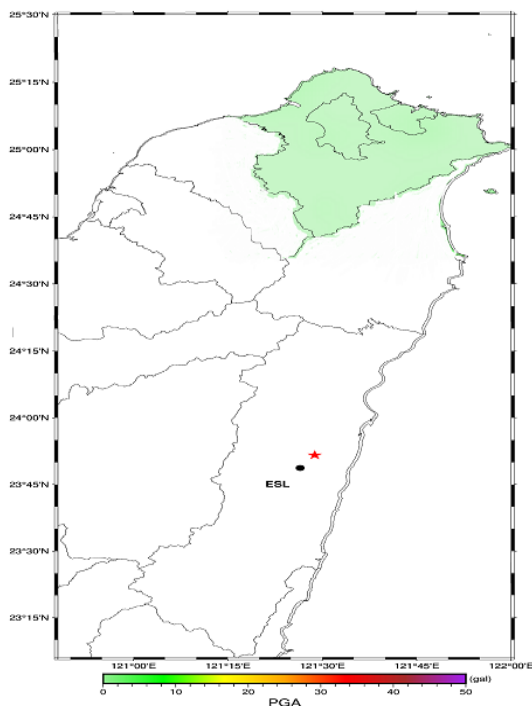


圖 11、P 波到後 1 秒 PGA 分佈圖

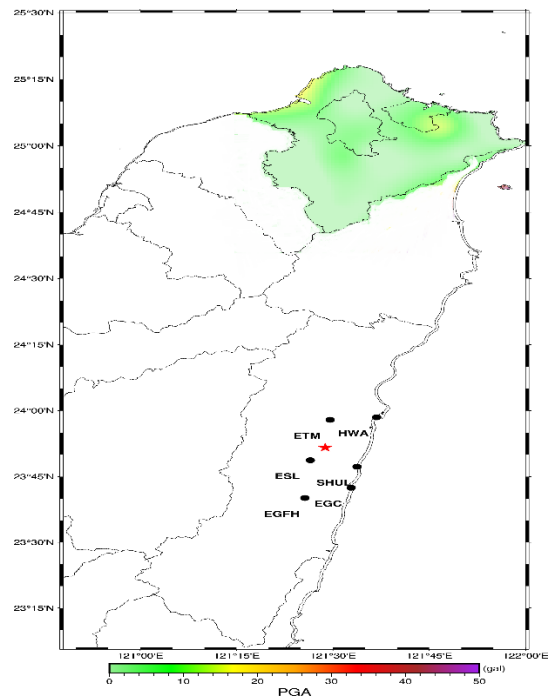


圖 12、P 波到後 3 秒 PGA 分佈圖

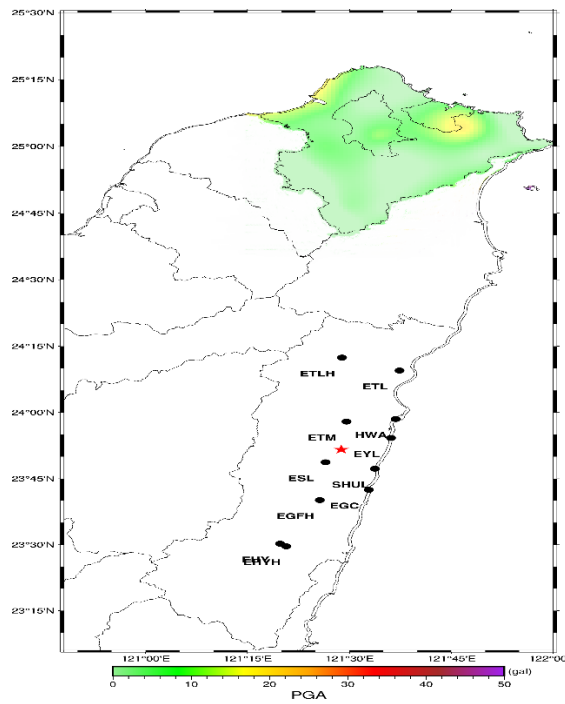


圖 13、P 波到後 5 秒 PGA 分佈圖

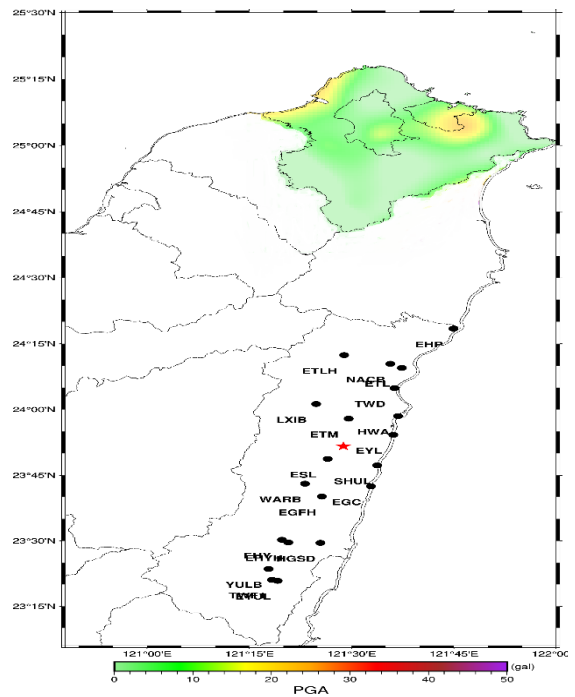


圖 14、P 波到後 8 秒 PGA 分佈圖

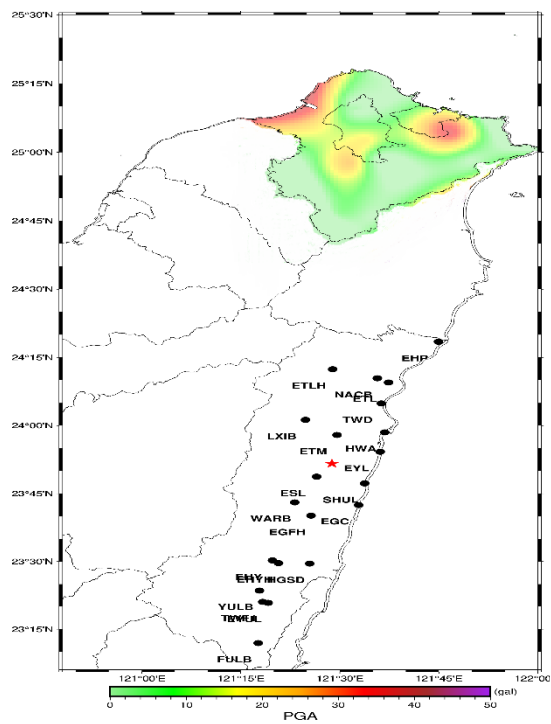


圖 15、P 波到後 10 秒 PGA 分佈圖

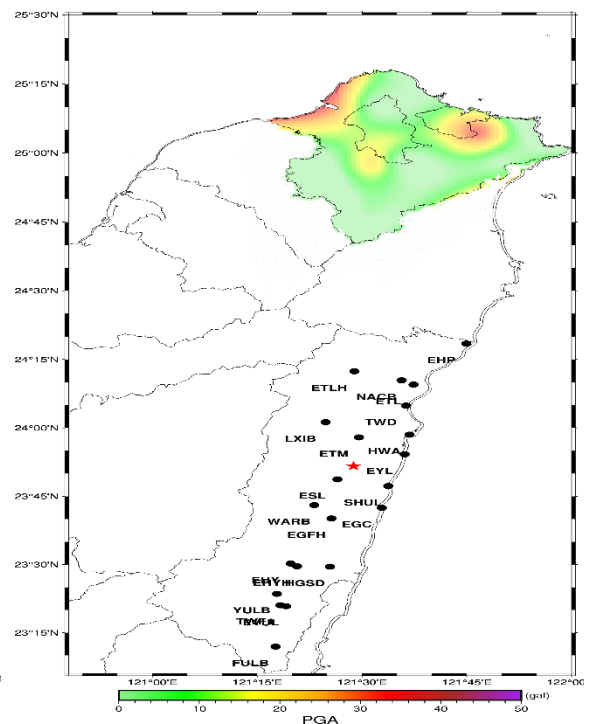


圖 16、P 波到後 15 秒 PGA 分佈圖

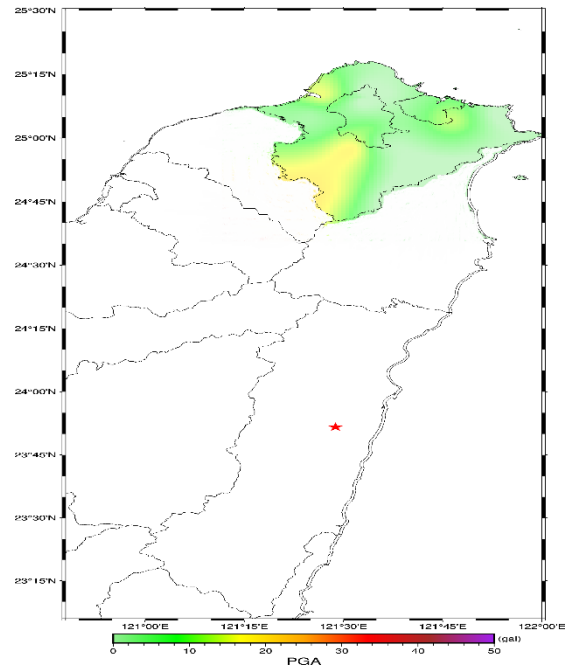


圖 17、真實 PGA 分佈圖

接著，我們對臺中地區做測試，一開始模型使用第一個被 Pick 到 P 波的 ESL 測站的三軸加速度波形資料來做預測預測結果如下圖 18，與圖 23 的真實 PGA 分佈圖比較，預測結果低估，主要是因為地震波剛到，該時間點只使用了 1 個測站的波形資訊來做預測，模型所接收到的資訊太少，才会有此結果產生。再來，地震經過 3 秒後，花蓮地區有 6 個測站有接收到波形資料，因此，我們使用這 6 個測站的三軸加速度波形資料來做預測，預測結果如下圖 19，預測結果與真實 PGA 已經非常相近，除了在靠近東邊的地區還有稍為低估。地震經 8 秒後，有 20 個測站接收到地震波，因此，我們使用花蓮地區 20 個測站所接收到的三軸加速度波型當作模型的輸入來做預測，此時間點東邊已經沒有低估的情況，預測結果如下圖 21。地震發生 10 秒後，此時間點的預測結果和 8 秒預測結果相似，預測結果如下圖 22。

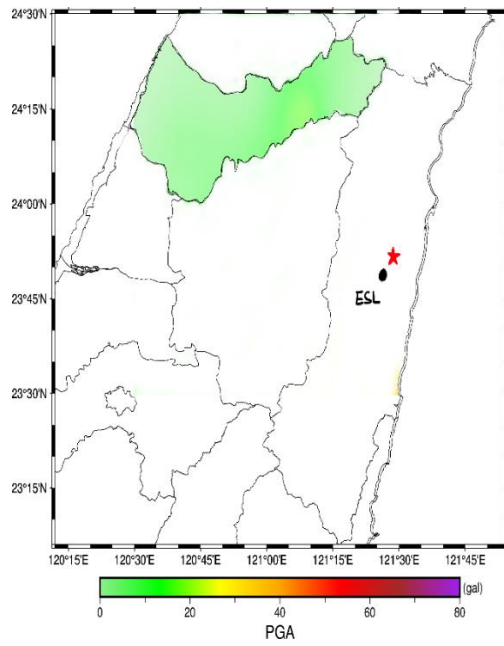


圖 18、P 波到後 1 秒 PGA 分佈圖

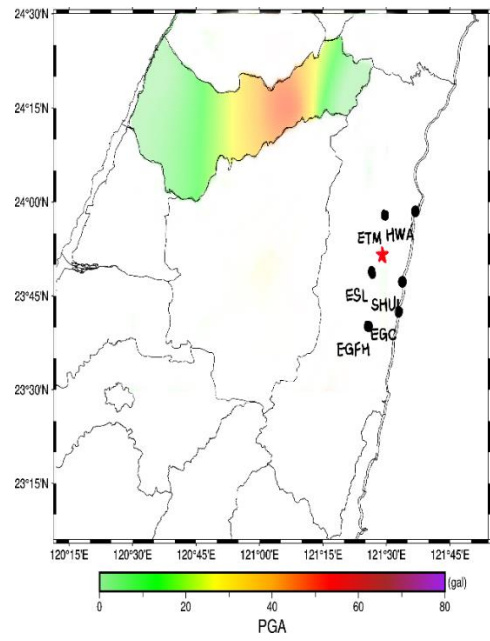


圖 19、P 波到後 3 秒 PGA 分佈圖

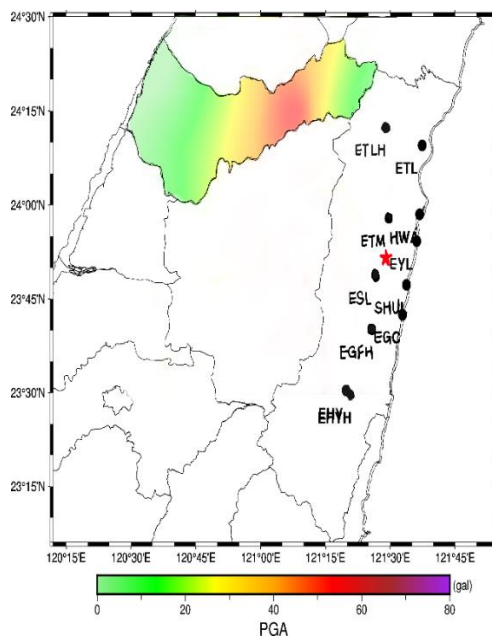


圖 20、P 波到後 5 秒 PGA 分佈圖

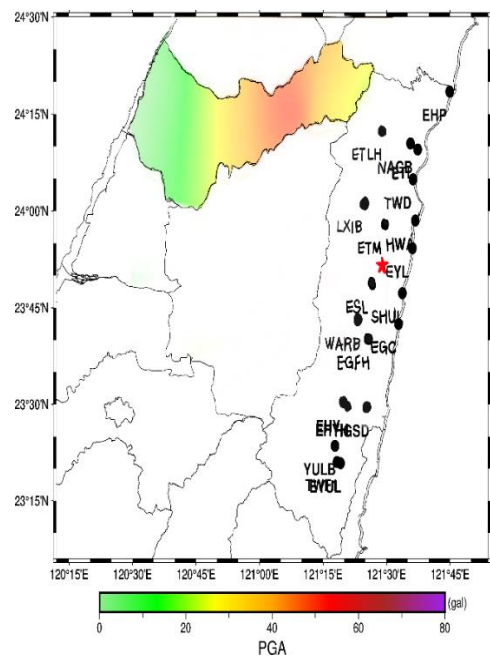


圖 21、P 波到後 8 秒 PGA 分佈圖

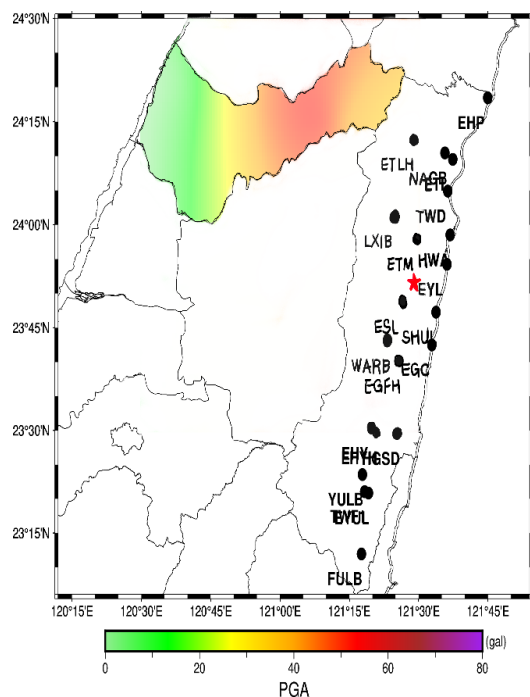


圖 22、P 波到後 10 秒 PGA 分佈圖

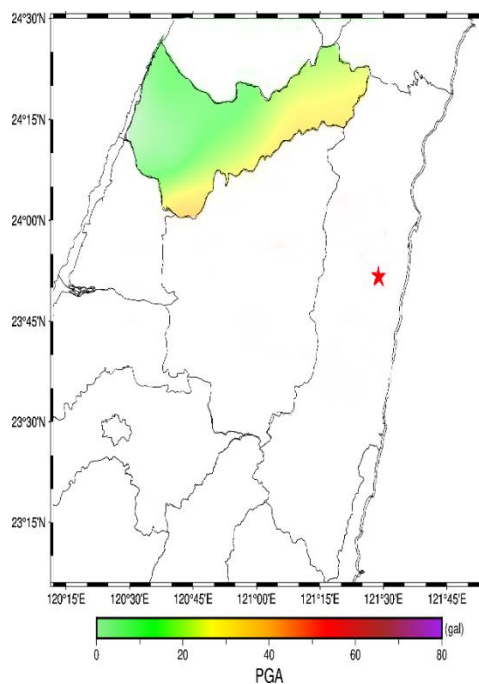


圖 23、真實 PGA 分佈圖

最後，我們對高雄地區做測試，一開始模型使用第一個被 Pick 到 P 波的 ESL 測站的三軸加速度波形資料來做預測預測結果如下圖 24，與圖 29 的真實 PGA 分布圖比較，預測結果低估，主要是因為地震波剛到，該時間點只使用了 1 個測站的波形資訊來做預測，模型所接收到的資訊太少，才会有此結果產生。再來，地震經過 3 秒後，花蓮地區有 6 個測站有接收到波形資料，因此，我們使用這 6 個測站的三軸加速度波形資料來做預測，預測結果如下圖 25，此時間點預測結果與真實 PGA 最相近，說明模型是有能力在地震波尚未到達該地區時，預測出相近的 PGA，5、8、10 的預測結果皆非常接近，預測結果如下圖 26、27、28。

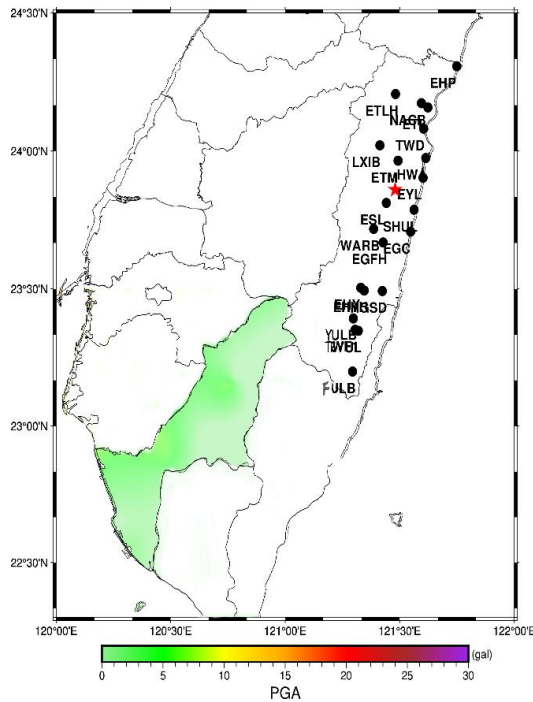


圖 28、P 波到後 10 秒 PGA 分佈圖

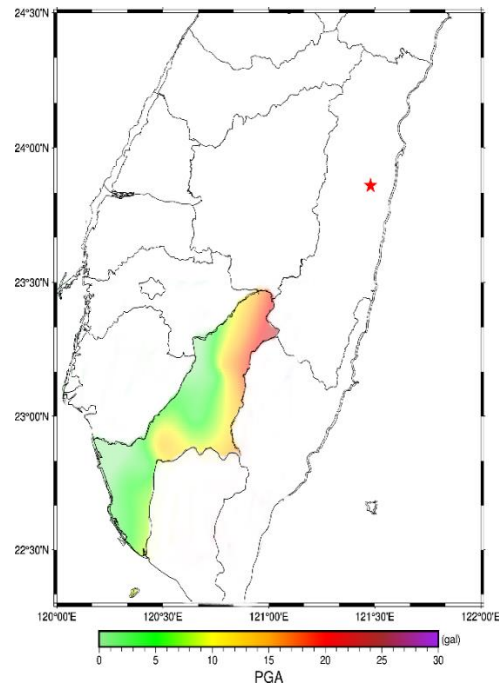


圖 29、真實 PGA 分佈圖

4.5.2 第 096 號花蓮地震事件預警測試

在此實驗中，我們利用 P 波前 5 秒到當前時間點 T 的波形的當作模型輸入對此事件做評估。選取 2021 年 9 月 26 日 6 時 21 分 19 秒，規模 5.7，深度 46.1 公里，震央位置位於北緯 24.32 度、東經 121.68 度的第 096 號花蓮地震。

首先，我們先對臺北地區做測試，一開始 P 波到時模型使用第一個被 Pick 到 P 波的 EHP 測站的三軸加速度波形資料來做預測預測結果如下圖 30，雖然有稍微低估，但此時間點僅用一個測站的波形資訊來預測，因此，誤差範圍是可接受的。再來，地震經過 3 秒後，花蓮地區有 7 個測站有接收到波形資料，因此，我們使用這 7 個測站的三軸加速度波形資料來做預測，預測結果如下圖 31，和圖 30 相比，模型在地震發生 3 秒後因為在接收到比較多的資訊，所以在預測的表現上也有提升，地震經過 5 秒後，有 9 個測站接收到地震波，因此，我們使用花蓮地區 9 個測站所接收到的三軸加速度波型當作模型的輸入來做預測。地震發生 8 秒後，花蓮地區有 13 個測站接收到波形，我們使用這 13 個測站的資訊來預測，預測結果如下圖 33，此事件模型低估的情形比較沒有這麼嚴重，誤差大約也都在 10gal 以內。再來，此模型可以在地震波還沒到達目標測站，提早預測出此事件目標測站真實的 PGA，事件發生大約 4 秒後新北地區的 NWL 測站為北部第一個 Pick 到地震波的測站，而我們的模型能在 4 秒內預測出與真實 PGA 相近的結果，由此可得知，此模型可以在地震波還沒到達北部區域時，預測出相近的 PGA 值。

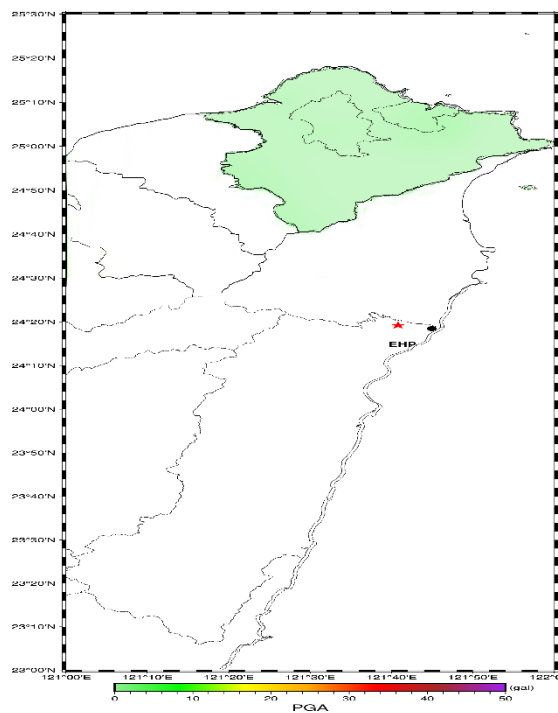


圖 30、P 波到後 1 秒 PGA 分佈圖

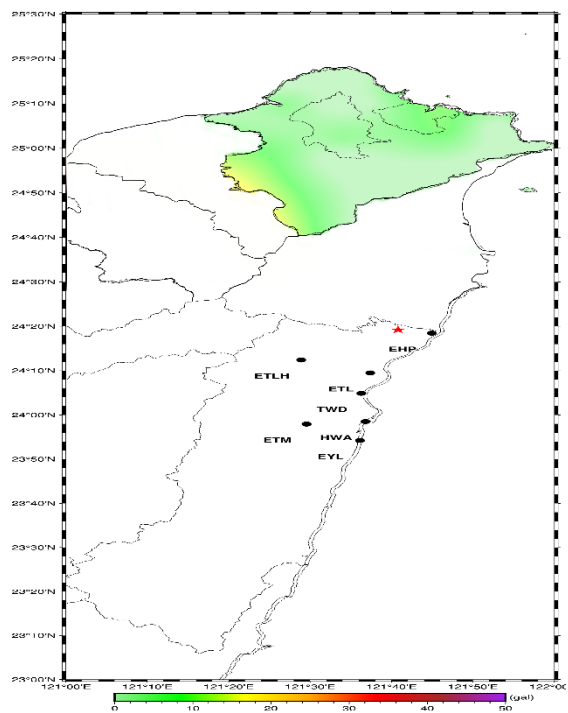


圖 31、P 波到後 3 秒 PGA 分佈圖

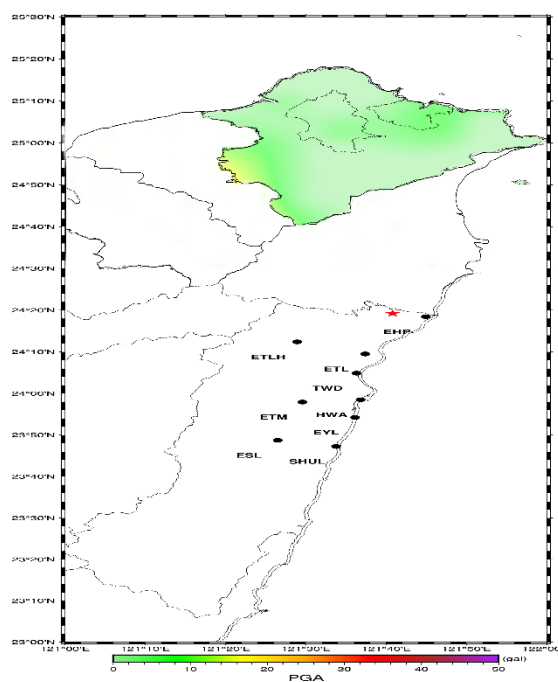


圖 32、P 波到後 5 秒 PGA 分佈圖

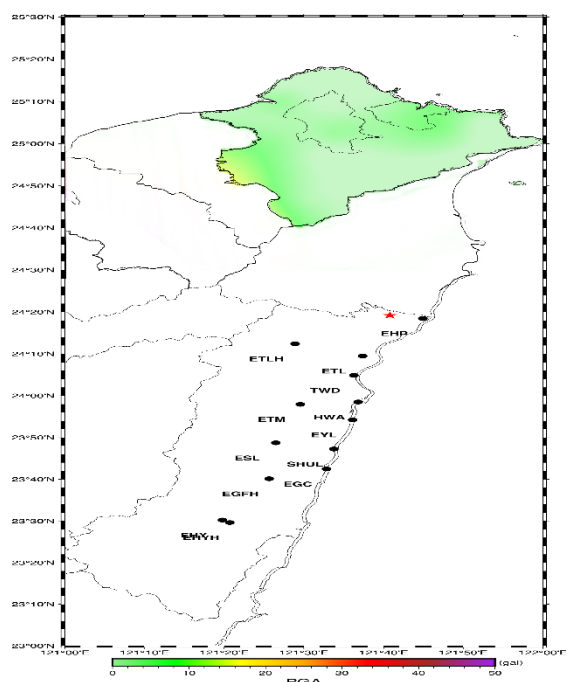


圖 33、P 波到後 8 秒 PGA 分佈圖

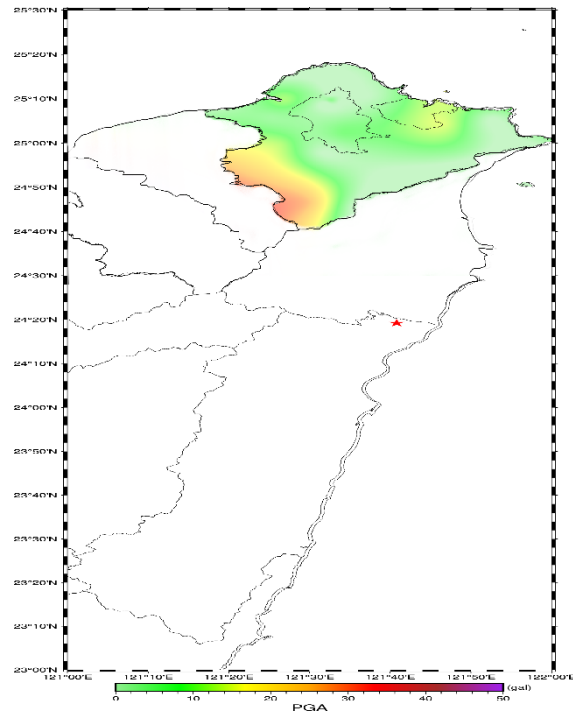


圖 34、真實 PGA 分佈圖

接著，我們對臺中地區做測試，P 波到時模型使用第一個被 Pick 到 P 波的 EHP 測站的三軸加速度波形資料來做預測，預測結果如下圖 35，地震發生 3 秒後，模型使用花蓮地區有 7 個測站的三軸加速度波形做預測，預測結果開始接近真實 PGA，真實 PGA 分布圖如下圖 39，除了在東半部稍為低估，但誤差皆在 10 gal 以內，說明了模型是有提前預測出相近 PGA 的能力，5 秒後，模型預測結果皆非常相似，5 秒預測結果如下圖 37，8 秒預測結果如下圖 38。

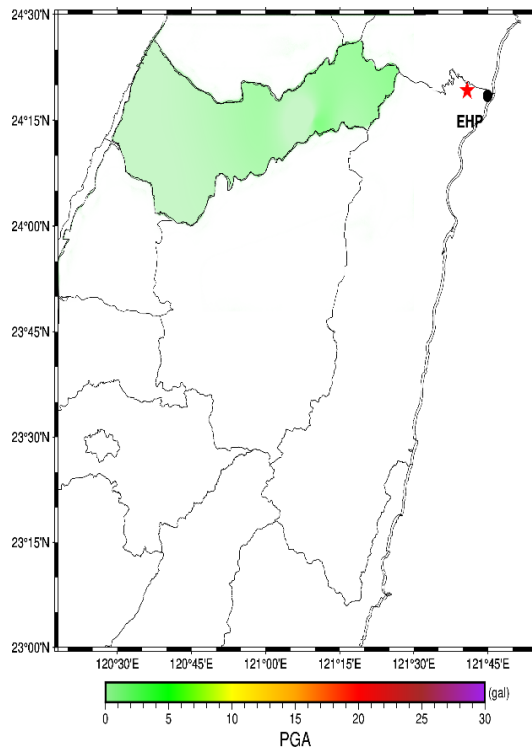


圖 35、P 波到後 1 秒 PGA 分佈圖

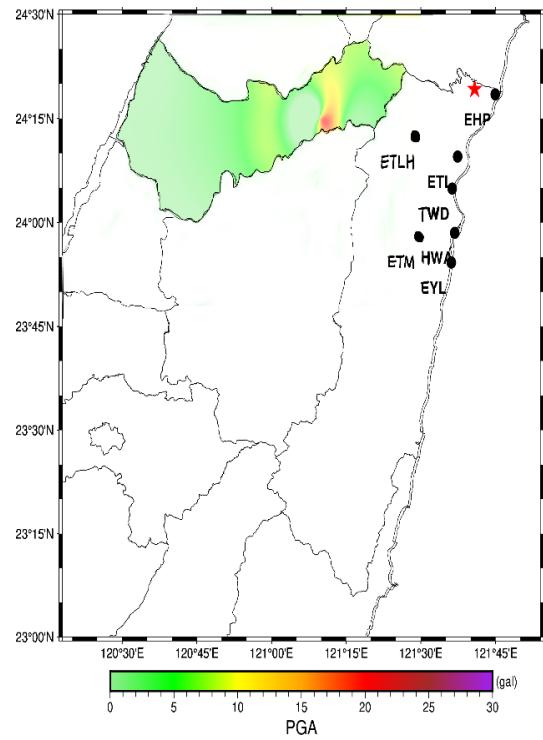


圖 36、P 波到後 3 秒 PGA 分佈圖

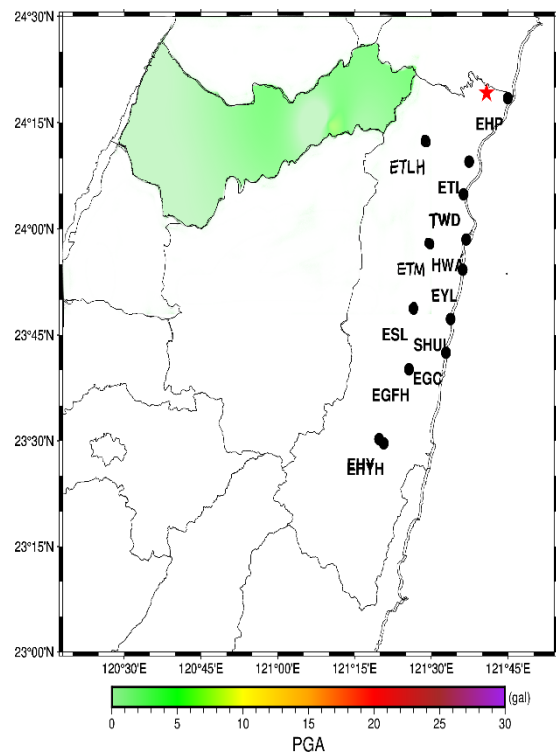
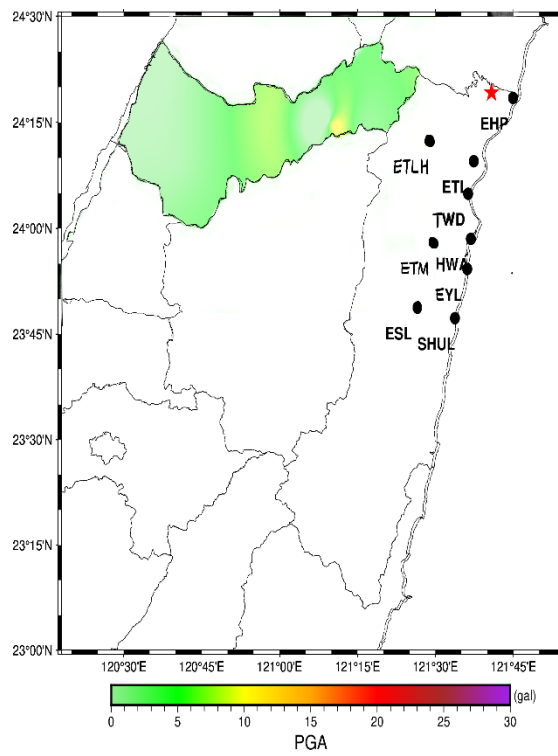


圖 37、P 波到後 5 秒 PGA 分佈圖

圖 38、P 波到後 8 秒 PGA 分佈圖

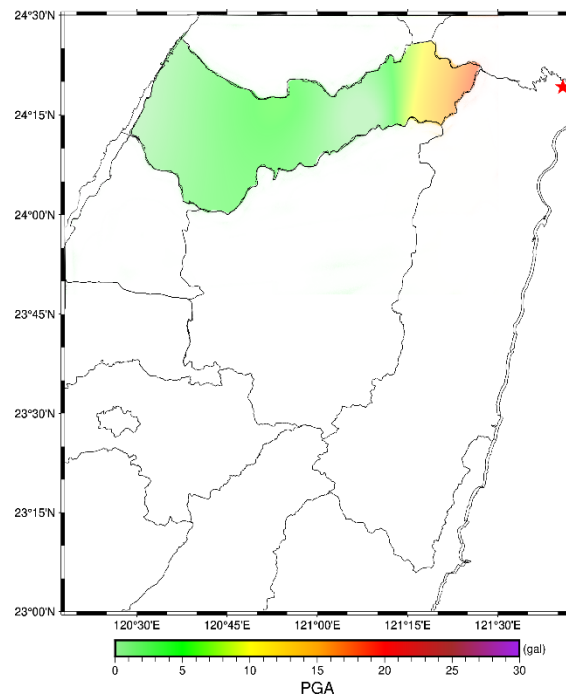


圖 39、真實 PGA 分佈圖

最後，我們對高雄地區進行測試，地震波剛到時，花蓮地區第一個被 Pick 到 P 波的 EHP 測站接收到地震波，我們使用此測站的三軸加速度波形資料來預測，此時間點的預測結果如下圖 40，地震發生 3 秒後，花蓮有 7 個測站接收到地震波，因此我們使用這 7 個測站的三軸加速度波形對高雄地區做預測，預測結果如下圖 41，此時間點的預測結果已經非常接近真實 PGA，真實 PGA 分布圖如下圖 45，說明了模型具有提前預測出相近 PGA 的能力，5 秒後的預測結果皆相似，下圖 42、43、44 分別為 5、10 秒以及 15 秒的預測結果。

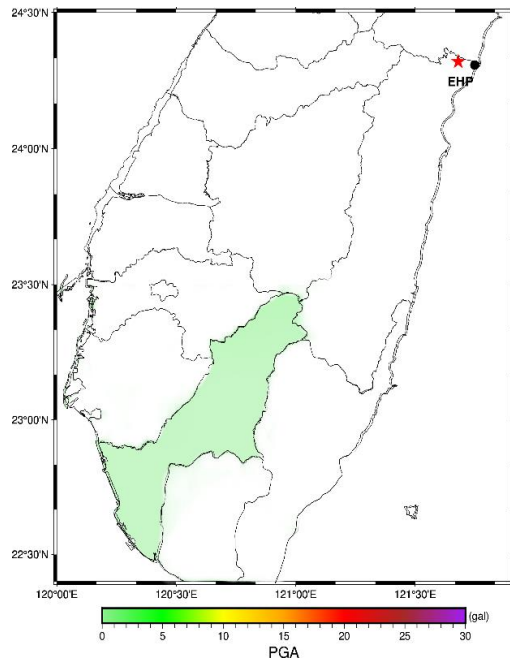


圖 40、P 波到後 1 秒 PGA 分佈圖

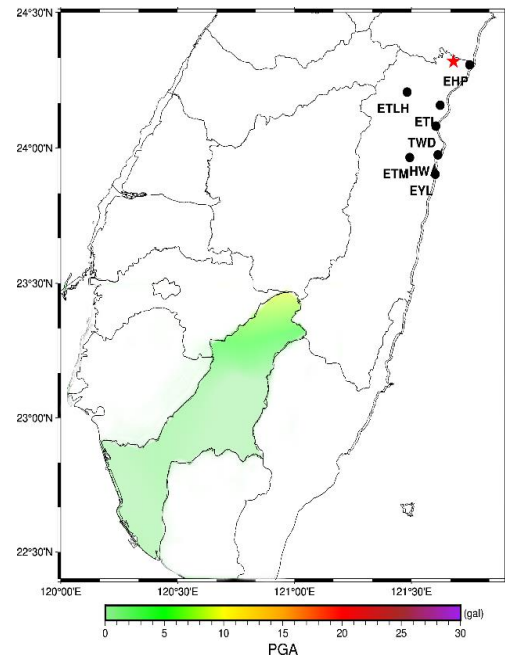


圖 41、P 波到後 3 秒 PGA 分佈圖

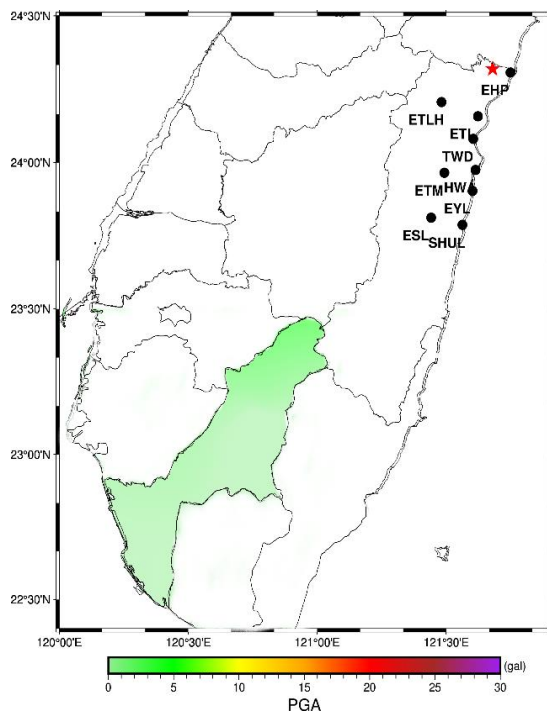


圖 42、P 波到後 5 秒 PGA 分佈圖

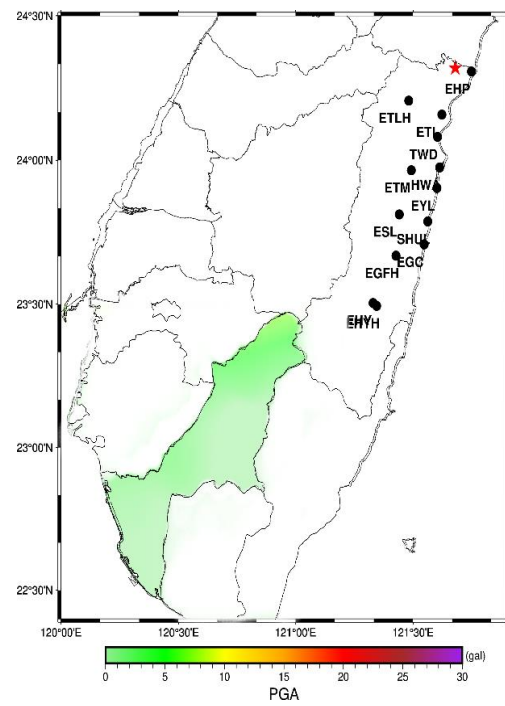


圖 43、P 波到後 10 秒 PGA 分佈圖

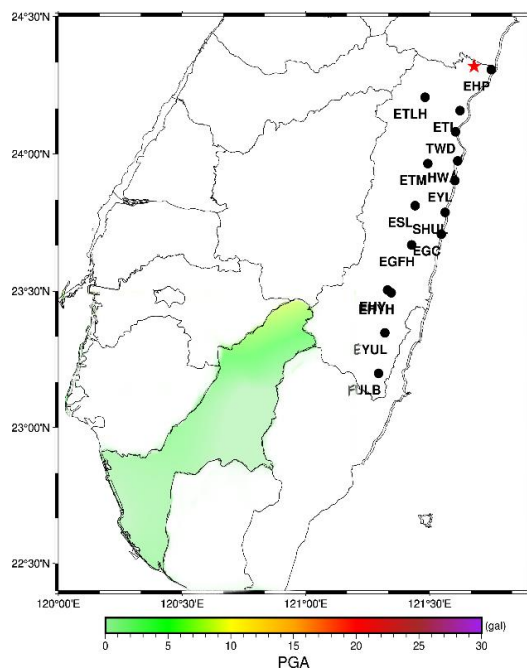


圖 44、P 波到後 15 秒 PGA 分佈圖

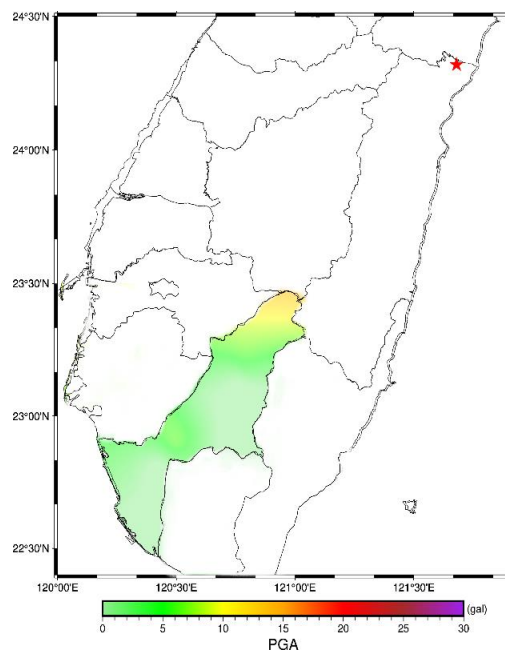


圖 45、真實 PGA 分佈圖

隨著地震時間的推移，我們的模型在得到更多的波形資訊後，模型的表現有越來越好的趨勢，預測的結果會越來越逼近真實的 PGA 結果，雖然在某些區域還是有低估的情況，但經過我們分區以及資料增強的處理，低估的情形也已經有明顯的改善，後續可以再加入一些地質的資訊，更進一步的降低低估的情況發生。再來，此模型可以在地震波還沒到達目標測站，提早預測出此事件目標測站真實的 PGA。

4.6 模型評估

為了瞭解模型是否能在地震波尚未達到目標區域而提早預測到 PGA 並發警報，因此我們將使用混淆矩陣對測試集進行分析，首先，我們要先定義閾值和發報條件，閾值設定: 2.5gal、8 gal、25 gal、80 gal，發報條件定義成觀測加速度比閾值加速度高則發報。再來是定義混淆矩陣的 True Positive (TP)、False Positive (FP)、True Negative (TN)、False Negative (FN)，定義如下表 1 所示。測試集共有 846 個花蓮地震事件，我們將以臺北地區的 NHDH 測站、臺中地區的 TWT 測站和高雄地區的 STYH 測站評估，使用 Precision、Recall、F1-score 評估，分別的計算方式如下式 5、6 和 7。

臺北模型選用 NHDH 測站評估，評估結果如下圖 46，25 和 80 gal 為 0 是因為測試集內此測站的真實 PGA 皆無超過 25 和 80 gal，而我們的模型也都沒有發報，因此皆屬於 True Negative (TN)。

臺中地區選用 TWT 測站評估，評估結果如下圖 47，80 gal 為 0 是因為測試集內此測站的真實 PGA 皆無超過 80 gal，而我們的模型也都沒有發報，因此皆屬於 True Negative (TN)。

高雄地區選用 STYH 測站評估，評估結果如下圖 48，因測試集內僅有 1 個事

件大於 25gal，而我們的模型漏報，因此 25gal 的各項評估值為 0。80 gal 為 0 是因為測試集內此測站的真實 PGA 皆無超過 80 gal，而我們的模型也都沒有發報，因此皆屬於 True Negative (TN)。

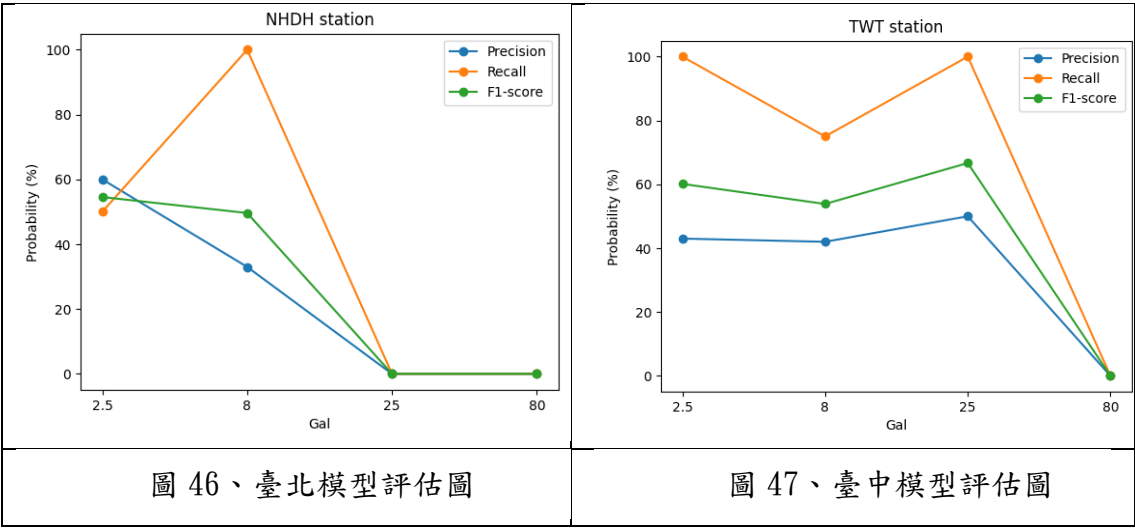
$$Precision = \frac{TP}{TP + FP} \quad (\text{式 5})$$

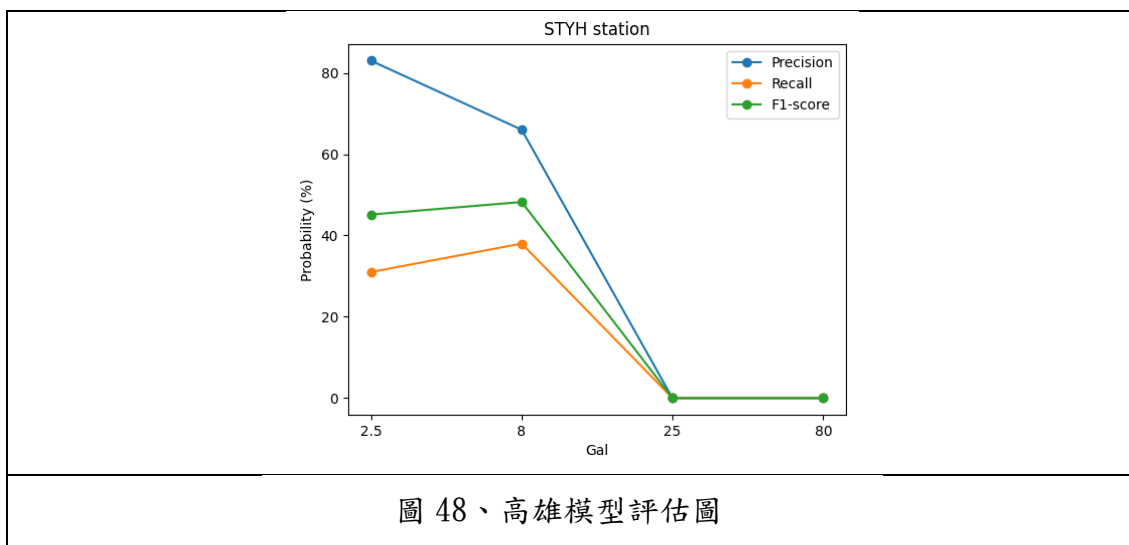
$$Recall = \frac{TP}{TP + FN} \quad (\text{式 6})$$

$$F1 = \frac{2PR}{P + R} = \frac{2TP}{2TP + FP + FN} \quad (\text{式 7})$$

表 1、評估方法

	Actual	
Predict	True Positive (TP): 預估有發報，實際有發報(發報時間比實際時間提前)。	False Positive (FP): 預估有發報，實際無發報。
	False Negative (FN): 1. 預估無發報，實際有發報。 2. 預估比實際發報時間晚。	True Negative (TN): 預估無發報，實際無發報。





4.7 結論

本研究利用卷積神經網路(CNN)，將三軸加速度波形做特徵提取，再利用測站的距離關係建立一個 weight matrix，然後利用圖卷積神經網路(GCN)，將其各測站的特徵依照 weight matrix 做 aggregation，此模型加入了分區的概念，讓模型更容易去學習不同地區之間的地震波能量衰減關係，進而改善低估的情況。隨著地震事件時間的推移，越來越多測站接收到波形，模型得到更多資訊後，在表現上也有更好的趨勢，因此模型具有 real-time 的能力，會再調整模型架構，使模型更為泛化，預測更加準確且快速，提升模型的預測能力。在資料方面，若能有更多可用且平衡的資料或是其他輸入資料類型和地震參數，可以提供給模型更多的判斷資料來提升預測結果。在未來，希望模型不要只侷限在花蓮地區的地震，使模型不管地震發生在臺灣任何地區，皆能快速針對目標區域做出準確的預測，並持續精進模型的性能，對大震度的地震事件深入做評估，讓模型達到能上線至地震預警系統之要求。

5 參考文獻

- [1] Bloemheuveld, S., van den Hoogen, J., Jozinović, D. et al. "Graph neural networks for multivariate time series regression with application to seismic data." *Int J Data Sci Anal* 16 (2023): 317–332
- [2] Zhang, Si, et al. "Graph convolutional networks: a comprehensive review." *Computational Social Networks* 6.1 (2019): 1-23.
- [3] Münchmeyer, Jannes, et al. "Earthquake magnitude and location estimation from real time seismic waveforms with a transformer network." *Geophysical Journal International* 226.2 (2021): 1086-1104.
- [4] Münchmeyer, Jannes, et al. "The transformer earthquake alerting model: A new versatile approach to earthquake early warning." *Geophysical Journal International* 225.1 (2021): 646-656.
- [5] Zhou, Cui, et al. "Graph neural networks: A review of methods and applications." *AI Open* 1 (2020): 57–81

- [6] Kipf, T.N., Welling, M. "Semi-Supervised Classification with Graph Convolutional Networks." *International Conference on Learning Representations* (2017).

機器學習技術建立現地型地震預警系統相關研究(II)

子計畫二

使用多測站資料與自我監督式學習於地震預警系統之研究

陳冠宇 林汶蔚 湯冠維 朱冠霖

國立臺灣科技大學

中文摘要

深度學習已在文字、影像與語音處理的許多任務上，取得大幅度的發展與進步。有鑑於此，本研究著眼於探究利用深度學習於提升地震預警的任務效能。本年度的研究分成三個方向。首先，延續過去的研究，我們發展出一套 GRADUATE 模型，透過結合時域與時頻域資訊，進行 P 波挑選任務，在實驗中進一步驗證其效能優於其他 state-of-the-art 模型，且在不同屬性資料下，亦保持穩定的效能。第二，有鑑於大型預訓練模型的成功，借鑑於語音處理領域中常見的技術，我們以 Wav2Vec2.0 模型為範本，利用 CWASN 與 TSMIP 資料作為訓練資料，初步嘗試大型預訓練模型在地震任務上的可行性。最後，我們延續前期成果，著眼於模型永續性問題，也就是探討利用多測站資料的深度學習地震預警模型，在面臨地震測站新增或移除時，如何處理的永續性作法，期望模型能持續提供穩定的預測結果，為後續實際上線運行時可能面對的問題做準備。

關鍵字：P 波挑選，訓練模型，TEAM

Abstract

Deep learning has made significant advances in many research subjects, including text, image, and speech processing. In light of this, this study focuses on exploring the use of deep learning to enhance the performance of earthquake early warning tasks. Our research results can be divided into three topics. First, based on our previous research results, we extended to derive a GRADUATE model that combines time-domain and time-frequency information for P-phase picking tasks. Through experiments, we further examined its performance superiority over other state-of-the-art models and its stability across different corpora. Second, inspired by the success of large pre-trained foundation models, we used the Wav2Vec2.0 model as an example. We explored the feasibility of large pre-trained models in earthquake early warning tasks by training on the CWASN and TSMIP datasets. Finally, continuing from previous achievements, we focused on the sustainability issue of the model, specifically addressing how a deep learning earthquake early warning model utilizing data from multiple stations handles the sustainability aspect when facing the addition or removal of seismometers. The goal is to ensure the model can not only provide stable predictions, but also prepare for facing potential challenges for real-world applications.

Keywords: Phase picking, Pretrained model, TEAM

第一章 介紹

本文提出了三個面向的研究，主要旨在運用深度學習技術發展具有穩定效能的模型，來處理地震領域的任務，例如 phase-picking、震度預測等。中央氣象署現行的地震預警系統多仰賴就有演算法，如 STA/LTA 負責處理 phase picking，這些演算法在多數情況下符合效能需求，但經研究發現，若遇到較少出現的波型屬性、模式，皆會使演算法效能有了非常大的衰退，影響著系統的可靠性，因此研究目的希望能將發展出的模型與系統做結合，以取代現有的演算法。

第一個研究延續過去一年的成果，發展出了專門針對 P 波挑選任務的 GRADUATE 模型。此模型結合時域與時頻域資訊作為輸入，讓模型同時考慮到震幅變化與頻率能量上的不同，多重考量下得到最適當的預測結果。在實驗中進一步驗證了其效能優於其他 state-of-the-art 模型，且在不同屬性資料下，亦保持穩定的效能。

第二個研究利用語音辨識領域中常見的技術，也就是用大量資料構建預訓練模型作為有效的特徵抽取模組，抽出富含資訊量的特徵向量後，再微調相對較小量的參數以適應各種不同的下游任務，以達到理想的效能表現。本研究利用 Wav2Vec2.0 模型，訓練在 CWASN 與 TSMIP 資料上，並使其能在 P 波挑選任務上得到一定的成果。此一成功也讓後續將預訓練模型套用至其他地震相關任務上的機率提高許多。

最後一個研究為多測站震度預測模型，參考 TEAM 模型的設計架構，以全臺灣測站同時作為輸入來預測未來震度的大小，並進行相對應的預警。為了更貼近真實情況，在訓練時加入「未使用測站」的向量，讓此向量在未來作為模型永續經營上有力的工具之一。

第二章 本研究提出之 P 波挑選模型

2.1 動機

P 波挑選模型的效能表現容易受到儀器影響，影響因素包括儀器屬性 (如 channel) 或是儀器建置環境 (如井下、地表地震儀)。這些因素都會使波形品質產生變化，讓模型的效能不穩。一般來說，我們會用 Signal-to-Noise Ratio (SNR) 衡量波形品質，經過相關研究發現，若能夠提升模型在低 SNR 值的波形上有穩定的預測能力，就能提升模型整體效能。

此外，在語音辨識領域中，使用到的音訊資料與 p 波挑選模型使用的波形資訊同屬於時間序列類別的資料，但常見的語音辨識模型會將音訊資料經過轉換後，產生具有頻率資訊的向量作為模型輸入 [2, 13, 14]，然而這樣的做法在 p 波挑選模型上卻不常見。因此我們認為除了原本就帶有時域資訊的波形資料外，若能將帶有頻率資訊的時頻域頻譜作為模型額外的輸入，對於提升模型效能上應有正面幫助。

2.2 模型設計

2.2.1 架構

根據研究動機提出了相關的 p 波挑選模型，名為 GRADUATE (Group Representation Across Dual-domain Understanding Against Turbulent Environments)。如圖 2.1 所示，GRADUATE 的架構由兩個分支組成：time domain branch 和 time-frequency domain branch。time-domain branch 以 Conformer [2] 為骨幹網絡，從三分量數據中提取特徵，並用不同感知領域的特徵來生成 time domain representation；而 time-frequency domain branch 以短時傅立葉變換 (Short-Time Fourier Transform, STFT) 頻譜圖作為輸入，當面臨嚴重的背景雜訊時，可解決模型性能下降的問題，在這方面，我們在模型的後部分中融合了來自兩個不同領域的資訊，以實現更穩定的預測，在圖中，GRADUATE 以 30 秒的地震波形作為輸入，維度為(批次大小, 3000, 3)，波形採樣為 100 Hz，產生 3000 個樣本點，而 3 表示地震波形的三分量。

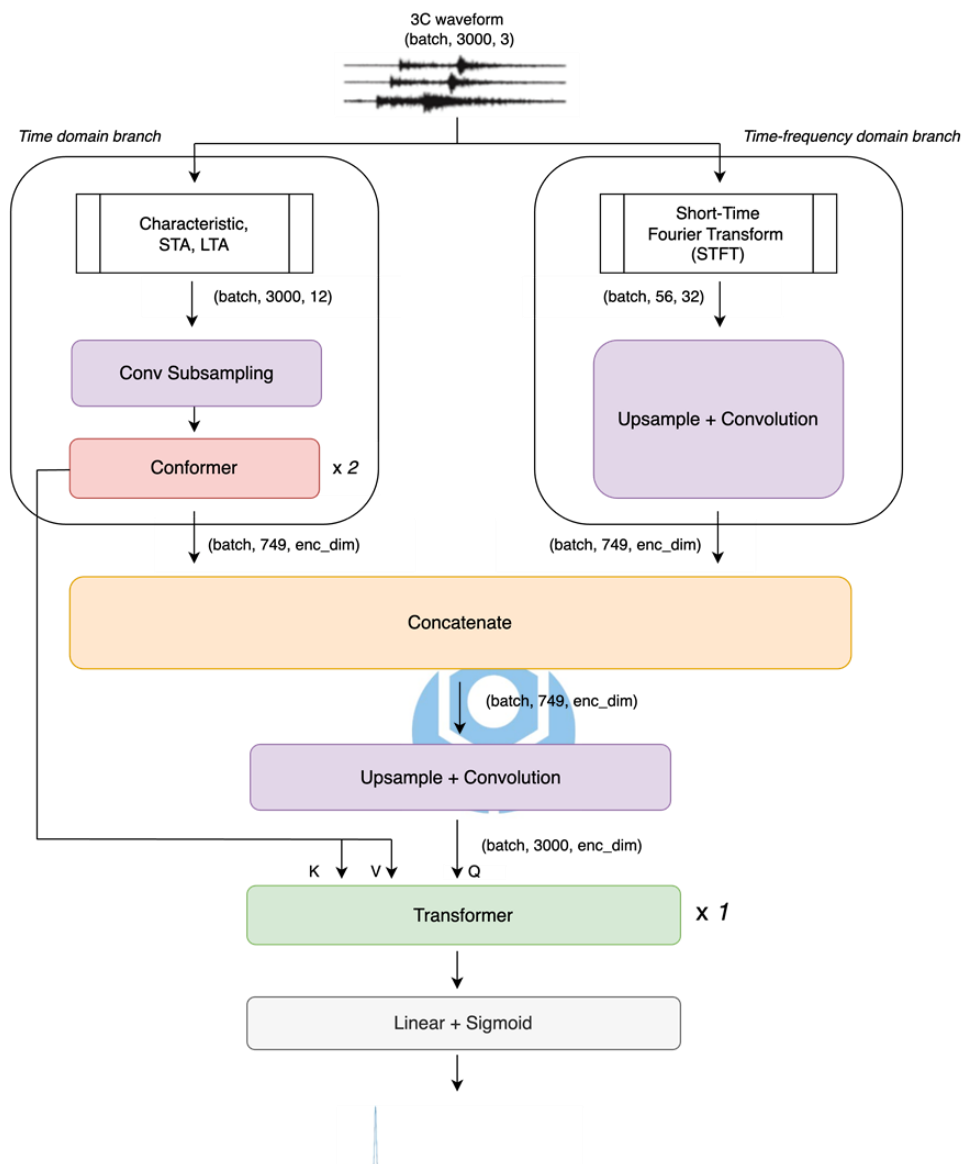


圖 2.1: GRADUATE 模型架構圖

2.2.2 Time Domain 分支

利用 Conformer 模型作為骨幹網路，他可以同時關注波形局部和全局的資訊。在過去許多 P 波挑選的模型研究中發現，大多數模型以 CNN 為主，因為地震相關波形數據擁有周期性，這使局部關係對任務來說至關重要；然而，全局關係也具有重要意義，不能被忽視，像規模預測任務需要更大的感知領域以提高預測的準確性。在這個分支中，除了原先的三分量外，為了提高模型的準確性，我們將幾個特徵添加到輸入向量中，分別是 characteristic, STA, 與 LTA。經過 Conformer 運算後，產生此分支的輸出，也就是具有時域資訊的向量表示法。

2.2.3 Time-Frequency Domain 分支

對於 P 波挑選模型而言，波形的品質在性能中發揮重要作用，背景雜訊的干擾程度直接影響了品質的好壞。以傳統方法為例，如 Short-Term Average / Long-Term Average (STA/LTA)，被雜訊嚴重干擾的波型容易產生誤報，使它們在許多情況下缺乏信心。基於深度學習的模型，在處理具有較重的背景雜訊資料方面取得了顯著的改進，但在處理較不明顯的地震訊號時，它們仍然會出現性能下降的現象。我們利用此分支來生成不同於時域的資訊，在取得經過正規化和濾波的波型後，通過取每個分量平方和的平方根來獲得合成震波，隨後應用短時傅立葉變換 (STFT) 後，得到波型相對應的頻譜圖。具體來說，取得頻譜圖的過程如下：

$$X_{\text{hybrid}} = \sqrt{Z^2 + N^2 + E^2},$$

$$X_{TF} = \text{STFT}(X_{\text{hybrid}}),$$

其中，Z、N 和 E 分別代表波形垂直、南北和東西分量，形成和城鎮波 H_{hybrid} ，接著通過 STFT 過程獲得頻譜圖 X_{TF} 。接著我們設計了一個模組，從頻譜圖中提取 time-frequency domain 特徵，鑒於波形中的大多數時間皆為背景雜訊，過大的模型感知領域可能導致模型過於關注背景雜訊，生成無用的特徵，所以我們選擇使用抽取局部特徵的方法從頻譜圖中提取 time-frequency 特徵。具體來說，我們將 Upsampling 與 1D CNN 堆疊，作為生成 time-frequency domain representation 的相關模組。

2.2.4 解碼器

經由兩個分支運算得到向量表示法後，接下來，我們需要考慮如何結合來自兩個不同領域的資訊。我們選擇直接將它們在維度上合併，然後經過一個全連接層，產生相對應的 dual-domain representation，最大程度地保留兩個領域上各自代表的資訊。一旦獲得了 dual-domain representation，為了使模型能夠對目標進行預測，利用可以關注全局資訊的 Transformer [1] 模組統合整個 representation 的資訊，因為模型前半段的兩個分支已經很好的提取了輸入資料的特徵，模型最後的部分可能需要更多的全局信息，以使模型能夠對整個隱藏狀態有全面的了解。模型有三個任務，分別是 p 波和 s 波的挑選，以及地震偵測。因此，我們將特徵向量映射到三維空間，分別代表每個時間點，對於三種任務各自的機率值。

2.3 實驗結果

2.3.1 資料集

我們使用自行整理，且在期中報告討論過的 CWA 資料集，包含來自 CWASN 以及 TSMIP 兩個地震網的資料。我們也從 CWASN 資料中額外抽取大量的雜訊來平衡使用資料中正負樣本的比例。資料數量如下表 2.1。

表 2.1：用於實驗之 CWA 資料集的波型數量

Subset	Training	Validation	Testing
CWASN	375,422	46,833	86,019
TSMIP	69,415	11,172	8,458
Noise in CWASN	599,248	71,746	135,920
Total	1,044,085	129,751	230,397

我們還在 STEAD [15] 上進行了實驗，STEAD 包含來自全球（不包括日本）的地震記錄，每筆資料皆為 60 秒，除了地震事件外，它還包括約 200,000 筆雜訊資料。資料數量如下表 2.2。

表 2.2：用於實驗之 STEAD 資料集的波型數量

Subset	Training	Validation	Testing
Seismogram	887,468	39,741	103,022
Noise	188,340	23,542	23,544
Total	1,075,808	63,283	126,566

進一步分析這兩個資料集的差異，波形品質會是比較的一項重點，而測站距離震央距離這項數據也是另外一個影響波形品質的關鍵，因此在表 2.3 中，我們統計了兩個資料集的 SNR 值分佈與震央距離分佈。我們發現 STEAD 相比 CWA 具有較高的平均 SNR 值，使模型更容易辨識 p 波到達的時間，然而，CWA 資料集的波形質量較差，這一結果主要與資料集內的平均震央距離有關，CWA 數據集的平均距離較大，影響了其 SNR 值分布，解釋了為什麼相對於 STEAD，其資料屬性較 noisy。

表 2.3：STEAD 與 CWA 資料集的比較，顯示了 Z 分量波形的 SNR 和震央距離的平均值和中位數（平均值/中位數）

	STEAD	CWA
Z-component SNR (dB)	22.06 / 19.92	11.09 / 10.52

Epicentral distance (km)	50.92 / 39.05	107.98 / 92.80
--------------------------	---------------	----------------

2.3.2 資料處理

對於每筆資料，我們選擇使用 Z-score normalization 與 1~10 Hz 的 bandpass 濾波器進行處理，以達到初步濾波的效果。在訓練期間，我們隨機將 p 波到時固定在預測時間窗的任意時間點，讓模型學習不同情境下的資料。在測試時，我們設計了的四個不同預測時間窗的長度，分別為 10、20 和 30 秒，讓我們能夠評估模型在不同長度的地震訊號下之效能表現。表 2.4 呈現出當預測時間窗不同之下，測試時固定 p 波到時的時間點。

表 2.4：不同輸入時間窗長度下，固定 p 波到達的時間點

Prediction window	The timesteps P-arrival fixed
10 s (1,000 samples)	250, 500, and 750-th sample
20 s (2,000 samples)	500, 1,000, 1,500, and 1,750-th sample
30 s (3,000 samples)	500, 1,500, 2,000, and 2,750-th sample

2.3.3 衡量指標

在測試中使用了三個指標：recall, precision, and F1-score。True-positive (TP) 是指當模型預測的 p 波到達時間與 ground-truth 之間的絕對距離小於 0.5 秒；false-positive (FP) 表示模型預測在上述的範圍之外；false-negative (FN) 表示模型在沒有存在 p 波的情況下被 pick；其餘情況被定義為 true-negative (TN)。Recall, precision, and F-score 的計算如下：

$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$F_{score} = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

此外，我們以秒為單位，計算誤差之平均值 μ 和標準差 σ 。

2.3.4 三十秒的預測時間窗

我們首先在較常見的情況下比較了 GRADUATE 與其他 baselines 的效能。此外，當輸入資料更長時，地震圖內訊號和雜訊之間的比例差異很大，這使我們可以觀察模型在這種情況下的表現。表 2.5 與 2.6 分別是在 STEAD 與 CWA 資料集上進行的 p-phase picking 的結果，在每個 p-arrival 測試情況下，GRADUATE 在性能上表現出色，超過了所有 baseline。就整體效能而言，與其他基於深度學習的方法相比，GRADUATE 在所有資料集上表現出更穩定的效能，我們將結果歸因於運用到時頻域的資訊，這有助於模型提高 recall 值的成績，即使用頻率的能量分佈來發現更多潛在的地震訊號，此外，利用不同粒度和結合不同長度感知領域的資訊，幫助模型更精確的挑選出 p-phase arrival time。

表 2.5：STEAD 資料集上的 *p*-wave phase picking 結果（預測時間窗= 30 秒）

P-arrival	Model	Recall	Precision	F1-score	$\mu(s)$	$\sigma(s)$
750	STA/LTA	0.9874	0.8569	0.9175	0.3361	1.2467
	EQTransformer	0.9988	0.9884	0.9936	0.0585	0.1633
	PhaseNet	0.9793	0.9630	0.9711	0.0916	0.5821
	RED-PAN	0.9847	0.9463	0.9651	0.1859	0.9843
	GRADUATE	0.9992	0.9881	0.9936	0.0358	0.1783
1500	STA/LTA	0.9999	0.8470	0.9172	0.3908	1.2313
	EQTransformer	0.9993	0.9878	0.9935	0.0560	0.1507
	PhaseNet	0.9817	0.9652	0.9734	0.0716	0.4172
	RED-PAN	0.9219	0.7838	0.8473	0.3192	2.2950
	GRADUATE	0.9999	0.9896	0.9947	0.0254	0.1440
2000	STA/LTA	0.9999	0.8464	0.9168	0.3916	1.2333
	EQTransformer	0.9992	0.9869	0.9930	0.0560	0.1636
	PhaseNet	0.9794	0.9655	0.9724	0.0704	0.3653
	RED-PAN	0.9106	0.8907	0.9005	0.1067	2.3559
	GRADUATE	0.9999	0.9890	0.9944	0.0249	0.1525
	STA/LTA	0.9999	0.8457	0.9164	0.4113	1.3174
	EQTransformer	0.9852	0.5014	0.6645	0.4533	0.4801

2750	PhaseNet	0.9825	0.9646	0.9735	0.0625	0.2774
	RED-PAN	0.7045	0.9379	0.8045	0.4253	1.7579
	GRADUATE	0.9997	0.9857	0.9926	0.0323	0.1986

表 2.6：CWA 資料集上的 *p*-wave phase picking 結果（預測時間窗= 30 秒）

P-arrival	Model	Recall	Precision	F1-score	$\mu(s)$	$\sigma(s)$
750	STA/LTA	0.9871	0.7757	0.8687	0.2294	0.6174
	EQTransformer	0.9961	0.9354	0.9648	0.1566	0.2867
	PhaseNet	0.9760	0.9364	0.9558	0.1642	0.4078
	RED-PAN	0.9470	0.9555	0.9512	0.1308	0.2307
	GRADUATE	0.9983	0.9446	0.9707	0.1469	0.3997
1500	STA/LTA	0.9870	0.7984	0.8827	0.2164	0.4677
	EQTransformer	0.9962	0.9360	0.9651	0.1631	0.3833
	PhaseNet	0.9780	0.9353	0.9562	0.1776	0.5772
	RED-PAN	0.9532	0.9553	0.9542	0.1313	0.3176
	GRADUATE	0.9985	0.9458	0.9714	0.1429	0.3107
2000	STA/LTA	0.9870	0.8093	0.8893	0.2068	0.4881
	EQTransformer	0.9962	0.9394	0.9670	0.1707	0.4643
	PhaseNet	0.9769	0.9302	0.9530	0.2004	0.8462
	RED-PAN	0.9487	0.9565	0.9526	0.1292	0.5057
	GRADUATE	0.9983	0.9459	0.9714	0.1394	0.3036
2750	STA/LTA	0.9848	0.8244	0.8975	0.1726	0.7652
	EQTransformer	0.9886	0.9006	0.9425	0.2715	1.2548
	PhaseNet	0.9715	0.9250	0.9477	0.3264	1.8239
	RED-PAN	0.7669	0.9623	0.8536	0.1139	1.4289
	GRADUATE	0.9687	0.8853	0.9251	0.2398	0.8352

2.3.5 十秒與二十秒的預測時間窗

三十秒的預測時間窗若要套用在即時系統上，在計算時間與資源上都無法有效減少，運用較短的時間窗便可以直接改善此一問題。在這一部分中，我們使用了長度為 10 秒與 20 秒的預測窗口，表 2.7 到 2.10 展示了實驗結果，我們發現，GRADUATE 在所有數據集上產生了比其他 baselines 更穩定的預測，包括 EQTransformer 和 STA/LTA，這表明長度為 10 秒的預測窗口更適合 GRADUATE 整合不同的粒度並充分利用來自 dual-domain 的資訊。

表 2.7：STEAD 資料集上 p-wave phase picking 的結果（預測窗口=10 秒）

P-arrival	Model	Recall	Precision	F1-score	$\mu(s)$	$\sigma(s)$
500	STA/LTA	0.9679	0.9015	0.9335	0.1628	0.4295
	EQTransformer	0.9934	0.9845	0.9890	0.0966	0.1595
	GRADUATE	0.9965	0.9894	0.9930	0.0864	0.1364
750	STA/LTA	0.9817	0.8895	0.9333	0.2352	0.7379
	EQTransformer	0.9947	0.9838	0.9892	0.0932	0.1632
	GRADUATE	0.9969	0.9899	0.9933	0.0882	0.1770

表 2.8：CWA 資料集上 p-wave phase picking 的結果（預測窗口=10 秒）

P-arrival	Model	Recall	Precision	F1-score	$\mu(s)$	$\sigma(s)$
500	STA/LTA	0.9201	0.8594	0.8887	0.2076	0.3383
	EQTransformer	0.9893	0.9395	0.9638	0.1489	0.2106
	GRADUATE	0.9960	0.9417	0.9681	0.1542	0.1995
750	STA/LTA	0.9187	0.8619	0.8894	0.1939	0.3338
	EQTransformer	0.9901	0.9391	0.9640	0.1558	0.2466
	GRADUATE	0.9954	0.9399	0.9669	0.1564	0.2265

表 2.9：STEAD 資料集上 p-wave phase picking 的結果（預測窗口=20 秒）

P-arrival	Model	Recall	Precision	F1-score	$\mu(s)$	$\sigma(s)$
500	STA/LTA	0.9852	0.8963	0.9386	0.2897	1.2820
	EQTransformer	0.9959	0.9869	0.9914	0.0715	0.1999
	GRADUATE	0.9983	0.9890	0.9936	0.0657	0.1433

1000	STA/LTA	0.9930	0.8183	0.8972	0.4531	1.1125
	EQTransformer	0.9969	0.9865	0.9917	0.0916	0.2304
	GRADUATE	0.9993	0.9902	0.9947	0.0576	0.1351
1500	STA/LTA	0.9999	0.8130	0.8968	0.5097	1.3008
	EQTransformer	0.9969	0.9851	0.9910	0.0993	0.2348
	GRADUATE	0.9995	0.9902	0.9949	0.0576	0.1222
1750	STA/LTA	0.9999	0.8132	0.8970	0.5318	1.3721
	EQTransformer	0.9963	0.9903	0.9933	0.1025	0.2644
	GRADUATE	0.9995	0.9904	0.9950	0.0631	0.1393

表 2.10：CWA 資料集上 p-wave phase picking 的結果（預測窗口=20 秒）

P-arrival	Model	Recall	Precision	F1-score	$\mu(s)$	$\sigma(s)$
500	STA/LTA	0.9691	0.8043	0.8790	0.2970	0.7141
	EQTransformer	0.9976	0.9413	0.9686	0.1450	0.2610
	GRADUATE	0.9981	0.9451	0.9709	0.1439	0.2884
1000	STA/LTA	0.9724	0.8184	0.8887	0.2234	0.5015
	EQTransformer	0.9978	0.9392	0.9676	0.1444	0.2452
	GRADUATE	0.9984	0.9450	0.9710	0.1419	0.2397
1500	STA/LTA	0.9715	0.8241	0.8918	0.2035	0.4700
	EQTransformer	0.9975	0.9381	0.9669	0.1507	0.3734
	GRADUATE	0.9982	0.9454	0.9711	0.1429	0.3327
1750	STA/LTA	0.9675	0.8290	0.8929	0.1811	0.5082
	EQTransformer	0.9961	0.9483	0.9716	0.1532	0.4591
	GRADUATE	0.9969	0.9592	0.9777	0.1384	0.3714

2.3.6 不同類型的地震儀

接著，我們分析了模型在不同類型儀器上的效能，EQTransformer、RED-PAN 和 GRADUATE 用來作為比較的模型，表 2.11 和 2.12 分別是針對不同 channel 和 location 儀器上的實驗結果，我們發現 GRADUATE 在不同的地震儀屬性下，成績依然優於其他 baseline，這表明 GRADUATE 生成更穩定的預測，不受地震儀特性的影響。

表 2.11：CWA 資料集上不同 channel 的效能

P-arrival	Channel	Model	Recall	Precision	F1-score	μ	σ
750	HL	EQTransformer	0.9968	0.9573	0.9767	0.1124	0.2265
		RED-PAN	0.9734	0.9701	0.9717	0.1040	0.2667
		GRADUATE	0.9991	0.9595	0.9789	0.1160	0.4684
	EH	EQTransformer	0.9906	0.9352	0.9621	0.1550	0.2615
		RED-PAN	0.9373	0.9549	0.9460	0.1267	0.1810
		GRADUATE	0.9981	0.9412	0.9688	0.1477	0.4184
	HH	EQTransformer	0.9920	0.9086	0.9484	0.2079	0.5553
		RED-PAN	0.9279	0.9417	0.9348	0.1616	0.2059
		GRADUATE	0.9978	0.9327	0.9641	0.1751	0.3050

表 2.12：CWA 資料集上不同 location 的效能

P-arrival	Location	Model	Recall	Precision	F1-score	μ	σ
750	Surface	EQTransformer	0.9929	0.9415	0.9666	0.1431	0.3331
		RED-PAN	0.9529	0.9613	0.9571	0.1208	0.1763
		GRADUATE	0.9986	0.9477	0.9725	0.1350	0.3140
	Borehole	EQTransformer	0.9873	0.9075	0.9457	0.2107	0.5535
		RED-PAN	0.9317	0.9393	0.9355	0.1691	0.2589
		GRADUATE	0.9977	0.9319	0.9637	0.1781	0.3370
	Seabed	EQTransformer	0.9977	0.9568	0.9768	0.1206	0.1908
		RED-PAN	0.9634	0.9480	0.9556	0.1084	0.2635
		GRADUATE	0.9981	0.9457	0.9712	0.2011	1.1803

2.3.7 不同 SNR 範圍的資料

波形的 SNR 值是影響模型效能的主要因素，其計算方式如下：

$$SNR = 10 \log_{10} \frac{|S_{95}|}{|N_{95}|}.$$

我們根據波形的 SNR 值，將整個 CWA 數據集分成多個區塊，並分別測試模型，結果顯示在表 2.13 中，GRADUATE 在所有區塊上的資料，其效能都優於 EQTransformer，表明我們的模型性能更穩定，我們還發現，當 SNR 不超過 0 dB 時，EQTransformer 和 GRADUATE 之間的性能差異變得更大，這證明了當波型受到強烈背景雜訊的干擾時，GRADUATE 比較更加穩定。

表 2.13：EQTransformer 和 GRADUATE 在 CWA 資料集不同 SNR 區間上效能比較

SNR interval	Model	Recall	Precision	F1-score	$\mu(s)$	$\sigma(s)$
<= 0 dB (3,879)	EQTransformer	0.9876	0.8892	0.9359	0.2601	0.3879
	GRADUATE	0.9884	0.9528	0.9703	0.2148	0.3733
0~5 dB (14,996)	EQTransformer	0.9886	0.9038	0.9443	0.2580	0.6055
	GRADUATE	0.9931	0.9461	0.9690	0.2117	0.5800
5~10 dB (20,756)	EQTransformer	0.9969	0.9271	0.9607	0.2164	0.6397
	GRADUATE	0.9984	0.9522	0.9748	0.1718	0.4501
10~15 dB (22,389)	EQTransformer	0.9993	0.9580	0.9782	0.1453	0.3359
	GRADUATE	0.9997	0.9692	0.9842	0.1224	0.2422
15~20 dB (16,793)	EQTransformer	0.9998	0.9787	0.9891	0.1066	0.3530
	GRADUATE	0.9998	0.9843	0.9920	0.0809	0.1733
20~25 dB (9,282)	EQTransformer	0.9997	0.9906	0.9951	0.0841	0.2137
	GRADUATE	0.9995	0.9934	0.9964	0.0553	0.1072
> 25 dB (5,022)	EQTransformer	0.9998	0.9966	0.9982	0.0737	0.1120
	GRADUATE	0.9998	0.9974	0.9986	0.0407	0.1029

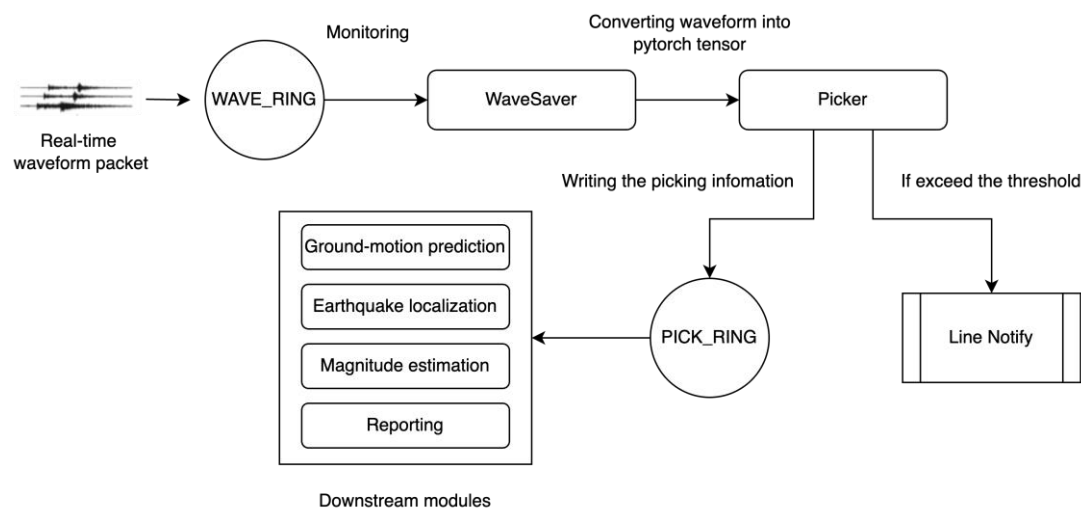


圖 2.2：EEW 系統與 GRADUATE 模型結合之架構流程

2.4 P 波挑選模型用於 EEW 系統之效能

Earthworm [8] 是一個專為地震預警系統 (EEW) 設計的即時系統，目前由中央氣象署作為早期警報系統的開發基礎，我們試圖將所提出的 p 波挑選模型 — GRADUATE 結合到系統中，以替代現有的模組。整個架構如圖 2.2 所示，系統由 WaveSaver 模塊監視，負責接收由測站發送的資料封包，當收到封包後，WaveSaver 會將封包資訊以及相對應的波型資料存至共享變數中，以供後續模組使用。

Picker 模組從共享變數中提取資訊，然後進行 GRADUATE 所需的前處理步驟，Picker 為每個測站預測出 picking 結果，然後將對應之 picking 資訊寫入共享記憶體，稱為 PICK_RING，供下游模組使用，例如規模預測模組。此外，我們還通過 Line Notify 向手機發送通知，讓我們可以即時接收到各測站的相關資訊，在運作一段時間後，我們觀察到在訓練集中的資料和真實世界資料之間存在著差異，也就是包含了模型在訓練期間未學習到的雜訊模式，這導致了高誤報率，因此，我們實施了一系列的防誤報機制：如果一個測站被 pick，並且在以其最近的 5 個測站中，有 2 個或更多測站也同樣被 pick，那麼該測站的 picking 結果被視為有效。這種方法減少了大多數誤報，儘管它也可能導致模型捨棄一些準確的預測時間上的延遲。此外，為了提供下游模組具有即時性的資訊，我們只考慮當預測的 phase arrival 發生在預測時間窗的最右側時，才將其視為有效，例如，當模型以 20 秒的時間窗作為輸入時，只有當預測的 phase arrival 發生在第 1,500 個樣本之後時，才被視為有效的 picking 結果。目前，我們已將輸入資料長度為 20 秒的模型部屬到系統中進行測試，該系統監測 TSMIP 地震網，該地震網包含總共 553 個即時測站。此外，我們還將 CWA 原本設計的 picker 之結果進行比較，也就是 STA/LTA 演算法。除了即時測試外，我們還利用 Earthworm 中的 tankplayer 模組回放了過去的重大事件，我們從 2022 年選擇了 3 月 23 日（事件編號 019）。

2.4.1 事件 1：2022 年 3 月 23 日

第一個事件發生在 2022 年 3 月 23 日（編號 019），事件的詳細資訊列在表 2.14 中，表中的 First Arrival 代表模型首次將 picking 資訊寫入共享記憶體 PICK_RING 之絕對時間。在圖 2.3 中，我們只展示了兩個 picker 首次偵測到 p 波時的情況，圖的左部分顯示了 AI picker，即 GRADUATE 的結果，而右側顯示了 CWA picker，即傳統的 STA/LTA 方法的結果。

表 2.14：有關 2022 年 3 月 23 日事件（編號 019）的詳細信息
(UTC+8)

Origin Time	2022/03/23 01:41:38
Epicenter	(121.61°E , 23.4°N)
Magnitude	6.7
Source Depth	25.7 km
Maximum Intensity	6 weak
First Arrival (AI picker)	01:41:50.41 (6 stations)
First Arrival (CWA picker)	01:41:49.17 (1 stations)

雖然結果顯示 CWA picker 比 AI picker 提前約一秒將 picking 資訊寫入 PICK_RING，但僅包含一個測站，然而，GRADUATE 在下一秒同時 picks 了六個測站。雖然從時間上看，在相似的前處理和後處理下，STA/LTA 僅使用演算法計算預測結果，GRADUATE 是基於深度學習的方法涉及額外的步驟，如圖形處理單元之間的數據傳輸和大量參數計算以獲得預測結果，因此 GRADUATE 比傳統方法慢，。然而，就預測結果而言，我們期望模型能夠實現更穩定的性能，這一期望可以從 GRADUATE 模型是等到更多鄰近測站被 pick 時，才寫入共享記憶體。

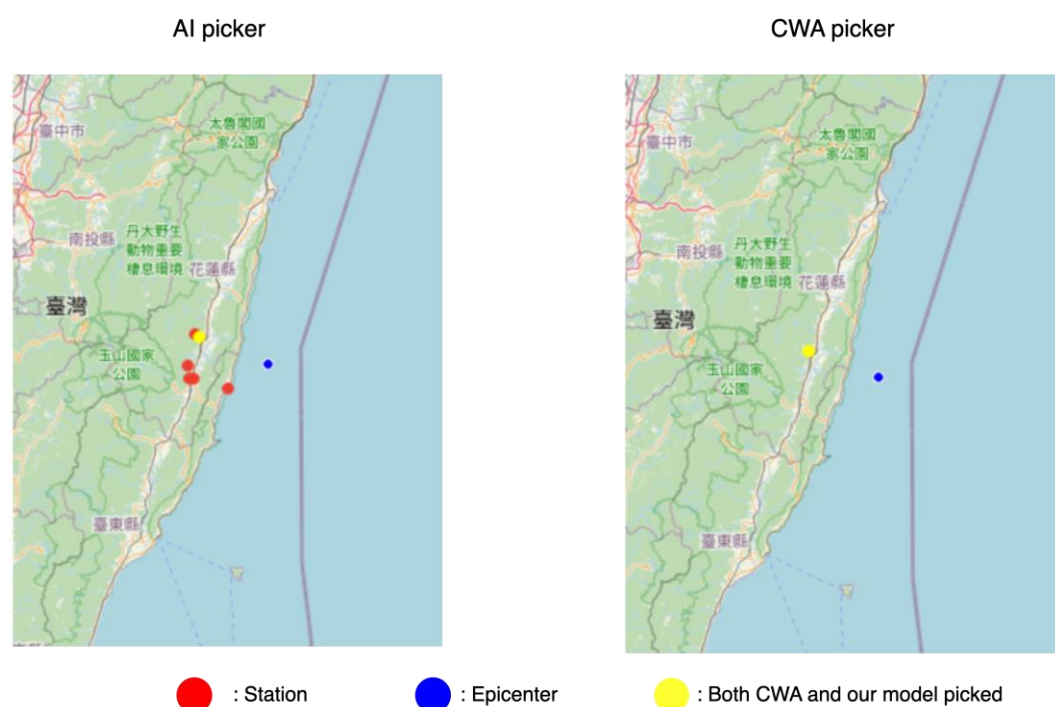


圖 2.3：AI picker（左）和 CWA picker（右）在 2022 年 3 月 23 日事件（No.019）中預測的震央和挑選站點

表 2.15：2022 年 3 月 23 日事件（No.019）中 AI picker 和 CWA 人工挑選之結果的 phase arrival 時間比較

Station	Manual	AI picker
EYUL (121.32°E , 23.35°N)	01:41:46.47	01:41:46.41
EHYH (121.35°E , 23.49°N)	01:41:46.04	01:41:46.14
ECB (121.45°E , 23.32°N)	01:41:45.91	01:41:45.92
TWFI (121.31°E , 23.35°N)	01:41:47.41	01:41:47.40
EHY (121.33°E , 23.50°N)	01:41:46.05	01:41:46.24
YULB (121.30°E , 23.39°N)	01:41:46.26	01:41:46.22

除了即時測試結果之外，我們進一步比較了 GRADUATE 預測的 phase arrival 時間與經 CWA 專家人工挑選之 phase arrival 時間，CWA 的結果來自 pfiles，表 2.15 顯示了結果，這 6 個測站的挑選誤差幾乎都不超過 0.1 秒，顯示 GRADUATE 的穩定度且能夠進行即時運算。

2.4.2 EEW 系統效能

在前部分我們提到，基於深度學習的 phase picker 未必能比傳統的 picker 更快產生 picking 結果，因此，穩定的效能已經成為我們期望的目標，在實際運作一段時間之後，我們發現每天有大量的地震事件，僅少數被視作有感地震，CWA 和 GRADUATE picker 都顯示出類似的預測趨勢，然而，不排除在這種情況下 picker 亦可能產生許多誤報。

表 2.16：CWA 和 AI picker 每日 picking 的測站數量比較

Date (UTC+0)	AI picker	CWA picker
9/29	69	293
9/30	124	375
10/01	100	343
10/02	210	418
10/03	133	429
10/04	130	467
10/05	133	298

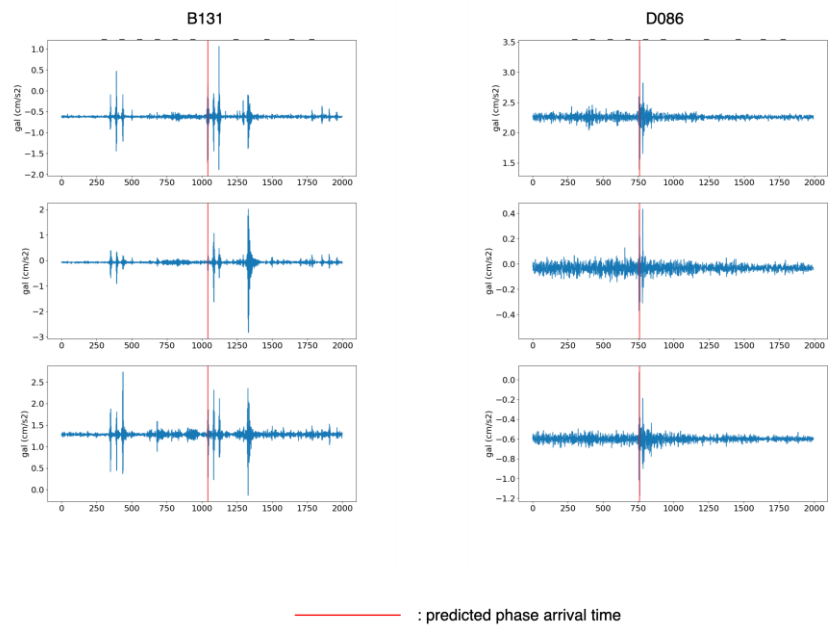


圖 2.4：使 CWA picker 容易誤報的常見波型

表 2.16 比較了 CWA 和 AI picker 每日 pick 的測站數量，AI picker picking 的測站數量遠低於 CWA picker，這可能跟防誤報機制套用有關，它過濾掉了許多只

有 picking 一個測站的情況，然而我們也發現之所以 CWA picker 每天會有大量的測站被 pick，原因與圖 2.4 的波型狀況有關，圖中顯示了 CWA picker 比較常見的誤報波型，這兩個波型面臨著瞬間強烈的震動，沒有明顯的 p 與 s 波，然而，我們採用的防誤報機制有效防止了這種誤報狀況。

圖 2.5 展示了地震預警系統中 Picker 模組發送的通知，picked stations 與對應之波型透過 Line Notify 應用程式發送到手機。

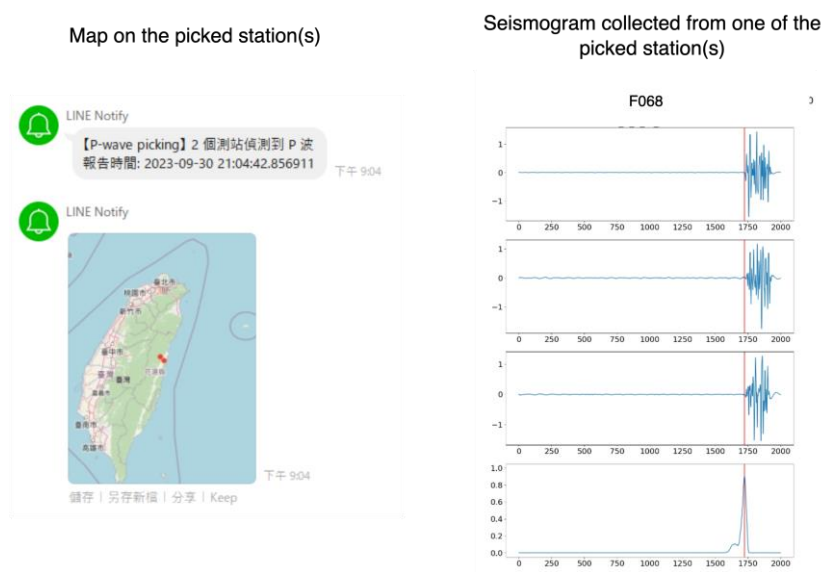


圖 2.5：Line Notify 上通知的範例

第三章 預訓練模型用於地震之研究

3.1 研究背景

傳統的地震監測方法往往受限於人力和時間成本，無法滿足大規模地震資料的高效處理需求。隨著深度學習技術的快速發展，我們可以利用其在圖像和序列資料處理方面的強大能力來改善地震監測和預測的準確性和效率。

而在前一次的計畫中分為兩個部分，分別為預測 p 波觸發時間來達地震預警之目的的 picking 模型 GRADUATE，以及透過多個測站的波形來預測測站未來的地表加速度值超過特定門檻值的機率的結合特定地點向量與閘門機制之全測站地震預警模型。預測 p 波的觸發時間的模型 GRADUATE 已經對波形中是否含有地震訊號的任務做到清楚辨識。而本研究的目的是利用深度學習技術開發一種預訓練模型，利用未標註數據進行非監督學習，訓練完成後，再通過少量標註數據進行微調。

地震數據集的獲取通常是一項複雜且耗時的任務。由於地震事件的罕見性和

對地震事件進行人工標註需要花費大量的時間以及人力，可用於深度學習模型訓練的地震數據集數量有限，並且有大量的地震數據並沒有經過人工標註或者審核標註正確性有疑慮。而這些未標註數據以及錯誤標註數據也隱含著寶貴的資訊，可以幫助提高地震波形 picking 和震度預測模型的性能。因此，這次計劃旨在開發一種能夠以非監督學習的方式利用未標註數據的預訓練模型，充分利用地震數據，提高模型的性能和適應性，並使其具備泛化能力，能夠適應在不同情況下的地震數據。因此，我們將借鑒 Wav2Vec [18]、Wav2Vec2.0 [19]、HuBert [20]，等先進的語音識別模型，並將他們對於語音波形進行的非監督學習以及自監督學習方法應用到地震資料中。

3.2 預訓練模型介紹

3.2.1 Wav2Vec

Wav2Vec 是一種自監督學習的語音表示學習模型，模型的特點是可以將原始的音頻訊號直接作為輸入，而無需進行語音訊號的前處理。他通過利用大量未標註的語音數據進行自我監督訓練，通過學習音頻中的上下文資訊來提取有意義的語音表示。具體來說，Wav2Vec 模型先將原始語音訊號 x 作為輸入，優化模型使其可以達到利用歷史和當前輸入的前後文語音訊號預測未來的樣本。本模型使用了兩個編碼器進行計算，分別是編碼器網路 (Encoder Network) 和前後文網路 (Context Network)。首先，編碼器網路 (f) 使用卷積層網路將語音訊號嵌入到特徵空間中，用類似於語言模型的方式，將每個 i 時間單位下的 x_i 映射為一個特徵向量 z_i ，再通過現有的特徵向量預測 z_{i+k} ，其中， k 為我們要預測的未來的時間點。然後，前後文網路 (g) 使用卷積層網路將編碼器網路輸出的特徵向量 $z_i, \dots, z_{i-v}$ 轉化為前後文表示 $c_i = g(z_i, \dots, z_{i-v})$ ，其中的 v 為接收的視野大小。之後，將以上兩個網路的輸出 Z, C 用於目標損失函數的計算。作者還利用了負採樣技術，從資料的概率分佈中抽取出負樣本 $\tilde{z}$ (干擾樣本)，最終模型的損失函數為區分正樣本和負樣本的對比損失函數 (Contrastive Loss)：

$$L_k = -\sum_{i=1}^{T-k} (\log \sigma(z_{i+k}^\top h_k(c_i)) + \lambda E[\log \sigma(-\tilde{z}^\top h_k(c_i))])$$

其中， σ 為 Sigmoid 函數， T 為序列的長度， λ 為負樣本的數量。 $h_k(c_i) = W_k c_i + b_k$ 是一個 step-specific 的線性層，即為對每個 k 都有一個對應的線性層。最後將所有 k 對應的 loss 相加就能得到最終的損失函數。在這個模型中，編碼器網路能夠讓模型從語音訊號中學習到語音的語義和特徵，前後文網路能夠讓模型捕捉到音訊的前後文資訊，提取更加準確和有意義的語音表示，而對比損失函數能夠通過最小化錯誤樣本的相似性來訓練模型，讓模型能夠學習到區分正確樣本和錯誤樣本的能力。通過以上的訓練目標和模型設計，Wav2Vec 能夠讓模型學到更準確的語音表示，而這些學習到的表示可以經過微調後用於語音識別，語音合成和其他語音相關的任務。

3.2.2 Wav2Vec2.0

Wav2Vec2.0 同樣通過利用大量未標註的語音數據進行自我監督訓練，通過學習音頻中的前後關係來提取有意義的語音表示。相對於 Wav2Vec，Wav2Vec2.0 從幾個方面進行了改進，包括引入了 Transformer 做為模型的編碼器，在預訓練過程中引入了乘積量化模組 (Product Quantization)，在模型訓練的過程中引入了遮罩 (Masking) 機制等。具體來說，Wav2Vec2.0 先將原始語音訊號 x 作為輸入，本模型同樣使用了兩個編碼器進行計算，分別是特徵編碼器 (Feature Encoder) 和 Transformer 編碼器 (由 Transformer 組成的前後文網路)。首先是由卷積層，歸一化層，以及啟動函數 GELU 組成的特徵編碼器 (f)，將輸入的原始音訊 X 映射為特徵向量表示 $z_1, \dots, z_T$ ，其中 T 代表了時間長度。特徵編碼器的輸出 Z 會傳至 Transformer 編碼器 (g)，Transformer 編碼器使用卷積層作為相對位置嵌入 (Relative Positional Embedding)。在這個部分中，Transformer 學習整個序列的前後關係並利用特徵編碼器的輸出來建構特徵表示 $c_1, \dots, c_T$ 。然後，特徵編碼器的輸出 Z 會經過乘積量化模組離散化為一組有限的語音表示 q_T 。乘積量化相當於從多個碼本 (Codebooks) 或組 (Groups) 中選擇“量化表示 (Quantized Representation)”並將它們拼接 (Concatenation) 起來。假設有 G 個碼本或組，每個碼本/組都具有 V 個項目 $e \in R^{V \times d/G}$ ，我們從每個碼本中選擇一個項目 (One Entry) 並將結果向量 $e_1, \dots, e_G$ 拼接起來，並應用線性變化 (Linear Transformation) $R^d \mapsto R^f$ 以獲得 $q \in R^f$ 。但是這個步驟無法直接進行反向傳播，所以在反向傳播中作者採用了 Gumbel Softmax 去輸出真實梯度以解決此問題。在進行模型訓練的時候，作者對模型中 Transformer 編碼器的一部分輸入進行了遮罩，並且使用在所有經過遮罩的時間點之間共用的且經過訓練的特徵向量去替換它們。模型的目標損失函數由一個對比學習損失函數 L_m 和一個多樣性損失函數 L_d 組成：

$$L = L_m + \alpha L_d$$

其中， α 為經過微調的超參數。在對比學習損失的部分，給定前後文網路輸出 c_T 以遮罩時間點 t 為中心，模型需要在 $K+1$ 個量化候選表示 $\tilde{q} \in Q_t$ 中識別正確的量化潛在語音表示，其中包括 q_T 和 K 個干擾樣本。並且還會計算前後文表示和量化潛在語音表示之間的餘弦相似度。對比學習損失函數如下：

$$L_m = -\log \frac{\exp(\text{sim}(c_T, q_T)/k)}{\sum_{\tilde{q} \sim Q_t} \exp(\text{sim}(c_T, \tilde{q})/k)}$$

而多樣性損失函數希望能夠增加量化碼本表示的使用，在這個函數中，平等地使用了每個碼本 (Code Book) G 中的 V 項目，利用在經過一批語音片段的碼本 $\underline{p}_g$ ，最大化每個平均 softmax 分佈 I 的熵。多樣性損失函數如下：

$$L_d = \frac{1}{GV} \sum_{g=1}^G -H(\underline{p}_g) = \frac{1}{GV} \sum_{g=1}^G \sum_{v=1}^V \underline{p}_{g,v} \log \underline{p}_{g,v}$$

在這個模型中的兩個編碼器能夠使模型學習到語音中的特徵表示以及前後關係，而額外引入的乘積量化模塊能夠生成離散的語音表示，能夠將無限的特徵表示空間限制成有限的離散空間，讓特徵的穩健性更強，減少受到干擾的影響。遮罩預測使模型需要借助已學習的特徵表示來填充遮罩位置，促使模型學習到有效的語音特徵。通過以上機制，Wav2Vec2.0 透過遮罩預測學習以及解決量化語音表示的對比任務，成功讓模型在對未標註數據進行語音處理預訓練上展現出巨大潛力。

3.2.3 HuBert

HuBert 是一種基於 BERT [22] 的自監督學習的語音辨識模型，它利用了離散化，遮罩預測以及集群分配的反覆運算細化等技術，實現了與 Wav2Vec2.0 類似或更好的表現。具體來說，在離散化的部分，給定 T 個 frame 單位的語音數據 $X = [x_1, \dots, x_T]$ ，假設模型能夠探測到的隱藏單位 (Hidden Units) 為 $h(X) = Z = [z_1, \dots, z_T]$ ，其中 $z_t \in [C]$ 是一個 C -class 的類別變量，而 h 則是一個非監督的聚類模型，例如 k-means。而在遮罩預測的部分，首先在如何遮罩的問題上，作者使用了和 Wav2Vec 類似的方法，首先隨機選擇 $p\%$ 的時間步驟作為開始的序列，然後從各個序列後面數 l 個 steps 作為要遮罩的區域。而在如何計算損失的問題，假設 $M \subset [T]$ 表示要遮罩的區域， $\tilde{X}$ 表示被遮罩後的遮罩預測模型 f 的輸入，我們定義在遮罩區域上的交叉熵損失函數 L_m (Cross-entropy Loss) 為：

$$L_m(f; X, M, Z) = \sum_{t \in M} \log p_f(z_t | \tilde{X}, t)$$

而相對應的我們可以定義在非遮罩區域上的交叉熵損失函數 L_u ，這個損失函數與上條式子類似，只是 $t \notin M$ ，最終的損失函數為：

$$L = \alpha L_m + (1 - \alpha) L_u$$

本模型中用到的 k-means 方法根據初始值和 k 值得不同，其表現也會大不相同。對於 Hubert 模型來說如果只是依賴一個 k-means 模型則其表現很可能會較大受到當前這個模型的影響，導致模型的魯棒性降低。為瞭解決這個問題作者引用了集群聚成 (Cluster Ensembles) 方法，使用多個不同參數的 k-means 模型來進行學習，這個時候假設我們用 $Z^{(k)}$ 表示第 k 個集群模型生成的序列，則 L_m 可以被定義為：

$$L_m(f; X, \{Z^{(k)}\}_k, M) = \sum_{t \in M} \sum_k \log p_f^{(k)}(z_t^{(k)} | \tilde{X}, t)$$

類似的損失函數也可以被定義為 L_u 。而除了使用集群聚成之外，模型裡改進的另一個方向是在整個學習過程中去改進集群分配。通過在學習到的潛在表示上訓練離散潛在模型使模型重新學習潛在表示來替代之前的集群模型，藉此來提高模型的精準度。通過以上方式，Hubert 可以使用大量的未標註語音數據進行自我訓練，學習到更加豐富和有意義的語音表示，從而改善語音識別等任務的準確性。

3.3 研究動機

而以上提到的這些自監督和非監督學習的模型中都有許多先進的技術與方法可以適用於預訓練模型的任務上，這些訓練方法都能成功在無人工標註的語音數據學習到隱含的特徵表示[18-21]。介於語音波形和地震波形同為時間序列下的波形資料，我們希望能夠將這些模型中的方法借用於地震研究中的波形 picking 模型，開發一個能夠在未標註的地震數據下學習地震波形中隱含的特徵表示的預訓練模型。我們可以將這些模型套用至地震資料上直接學習地震的隱含表示，將這些模型微調至下游的震波 picking 任務或是震度預測任務進行實驗，期望能夠達到更準確的預測成效。或是運用其中的預訓練任務的方法改進現有的 picking 或震度預測模型。

EQTransformer 是一個在 picking 任務中取得 SOTA 成績的深度學習模型。這個模型由一系列的輸入編碼器以及兩個作為地震震波 p 波和 s 波兩個分支情況下的解碼器，輸入編碼器由一系列串聯的卷積層，殘差網路卷積層(ResNet)，以及 Transformer 層組成。我們可以引用 Wav2Vec2.0 中提到的預訓練任務，將其套用至 EQTransformer 中的輸入編碼器。具體來說，模型將輸入的地震波形資料，經過卷積層和殘差網路卷積層提取後的特徵，在輸入 Transformer 前進行遮罩，並且引用乘積量化模組，將經過乘積量化後的卷積層編碼器的輸出表示與 Transformer 層的輸出表示進行對比學習，藉由以上方法學習原始地震波形的潛在特徵表示。並且在預訓練任務完成後，我們可以藉由微調解碼器的方式，將預訓練模型套用至下游的預測任務，例如 P 波和 S 波的觸發時間預測，震度預測等等。希望以此方式開發出一個泛化的地震預訓練模型，可以將其套用至未標註資料，減少數據預處理和人工標註所需的時間和人力成本，或是經過微調套用至各種地震預測任務上，達到更準確快速的地震災害檢測。

3.4 實驗流程與結果

目前，我們已經對地震波形的預訓練模型進行了初步的探索實驗，因為 Wav2vec2.0 其模型設計的泛用性以及效能都相對來說較好，我們選擇了 Wav2vec2.0 作為我們實驗中採用的模型。預訓練模型的主要訓練流程分為兩個關鍵步驟：

- 編碼器預訓練：在此階段，我們對編碼器進行預訓練，旨在使其從大量資料中學習到具有泛化能力的特徵。
- 解碼器微調：針對特定任務的需求，我們設計解碼器，使其接收預訓練編碼器輸出的特徵。隨後，我們對解碼器進行微調，以使其適應當前任務的具體要求。

這兩步的訓練流程有助於我們在地震波形分析中取得更為精確和有效的結果，通過預訓練模型的特徵學習和微調，使其能夠更好地適應特定地震分析任務。

預訓練模型的主要貢獻在於提供具有高度泛化能力的特徵。從理論上來說，若預訓練的編碼器能夠生成高品質的特徵，那麼在執行諸如分類、預測等任務時，僅需進行微調，並設計簡單的解碼器即可完成相應任務。因此，在預訓練模型的訓練過程中，我們特別注重編碼器的訓練效能。

在資料準備階段，我們選擇以 CWASN 和 TSMIP 的加速度波形作為訓練數據。對於原始波形，我們在輸入模型前進行了前處理。這包括使用濾波器從整個波形數據中去除無意義的低頻雜訊，隨後將資料切割為 60 秒的時間間隔，形成每筆資料。在此過程中，我們將地震波形隨機放置在 60 秒的時間內，確保每一筆資料都包含地震波形。在初步實驗中，我們以地震波形的 p 波檢測任務為研究對象，在同樣的實驗設置下，以 picking 任務效能作為衡量基準評估了在不同設定下編碼器訓練效能和質量的差異。對於 picking 任務的準確度效能，我們選擇使用 precision, recall, 與 F1-score 作為衡量指標。這些指標利用 True-Positive（真正例）、False-Positive（假正例）、True-Negative（真負例）、與 False-Negative（假負例）的數量，計算模型在預測 p 波的準確度。具體的衡量方式請參照表 3.1 和表 3.2。

表 3.1: TP, FP, TN, FN 之條件，其中模型預測值以“predicted”表示，經人工標注的值則以“ground-truth”表示。

TP	predicted 落在 ground-truth 前後 50 個 samples 範圍內
FP	predicted 落在 ground-truth 前 50 samples 範圍外
TN	Predicted 值沒有觸發，且 ground-truth 亦沒觸發
FN	Predicted 值沒有觸發，但 ground-truth 有觸發

表 3.2: Recall, precision, F1-score 計算方式以及其意義。

Recall	$\frac{TP}{TP + FN}$	在所有有包含 p 波的波形中，模型成功 picking 在正確時間點的比例
--------	----------------------	---------------------------------------

Precision	$\frac{TP}{TP + FP}$	在模型有 picking 的事件中，成功 picking 在正確時間點的比例
F1-score	$2 \times \frac{Precision \times Recall}{Precision + Recall}$	結合 recall 與 precision 的綜合分數

我們以 Wav2vec2.0 的初始配置為基礎，分別訓練了一個對音訊資料長度進行降維和另一個不進行降維的版本的預訓練模型。在預訓練完成後，我們對預訓練模型輸出的特徵向量進行微調，以適應解碼器在 Picking 任務上的需求，我們使用全連接層作為我們初步實驗的解碼器。通過 Picking 的準確度來初步評估地震波形在預訓練模型上的可行性，具體實驗結果詳見表 3.3。

表 3.3：原始設定下對長度進行降維和不進行降維的實驗結果比較

Method	Precision	Recall	F1-score
Downsample	0.5560	0.9413	0.6991
w/o Downsample	0.6668	0.8396	0.7433

根據表 3.3 的實驗結果，我們觀察到模型在 Picking 任務上已經展現出初步的辨識能力。這表明利用地震波形進行預訓練的模型具有相當的潛力。在接下來的實驗步驟中，我們著手嘗試將預訓練模型中編碼器的不同模組替換為地震領域中具有領先性能模型所使用的模組參數。我們引用了兩個地震 Picking 模型的領先模型，即 EQTransformer (EQT) 和 ViT Picking (ViT) [23] 的特徵提取器，以及時間序列預測模型中的領先模型 One Fits All (OFA) [24] 的特徵提取器。我們替換了 Wav2vec2.0 的特徵提取器，並根據這三個特徵提取器對編碼器進行了預訓練。在解碼器的部分，我們採用了全連接層（Linear）、卷積層（CNN）以及同時使用兩者的情況，作為解碼器進行微調。由於預訓練模型本身的訓練需要相當長的時間，並且考慮到在本次模組交換的消融實驗中存在多個模組需要調整，我們選擇僅使用資料集中占總數 20% 的子資料集作為訓練樣本，以提高實驗的效率。具體實驗結果詳見表 3.4。

表 3.4：不同特徵抽取器以及不同解碼器組合的實驗結果比較

Encoder	Decoder	Precision	Recall	F1-score
---------	---------	-----------	--------	----------

EQT	Linear	0.8034	0.9632	0.8761
	CNN	0.5522	0.8989	0.6841
	CNN + Linear	0.7774	0.8269	0.8014
ViT	Linear	0.4650	0.4196	0.4412
	CNN	0.4517	0.9574	0.6138
	CNN + Linear	0.6857	0.8357	0.7533
OFA	Linear	0.4793	0.5138	0.4959
	CNN	0.4214	0.6564	0.5133
	CNN + Linear	0.6938	0.7640	0.7272

根據實驗表 3.4 的結果，在不同特徵提取器和解碼器的組合下，EQTransformer (EQT) 的特徵提取器結合全連接層的解碼器，僅使用子資料集進行訓練時，相較於原始設置使用完整資料集的情況，性能得到了顯著提升。目前的實驗結果顯示，預訓練模型在 picking 任務上已經展現出一定的潛力，表明該模型具備基本的雜訊波形和地震波形的區分能力。在後續實驗中，由於預訓練模型的訓練過程中存在大量參數需要調整，並且在 picking 任務上與目前領先模型存在一定差距，因此在編碼器訓練的階段仍需進行參數的調整及優化。通過實驗，我們將進一步尋找適用於地震波形的預訓練模型參數組合。除此之外，為了驗證預訓練模型的泛化性，我們計畫引入其他與地震相關的任務進行實驗，例如規模預測、震度預測等。通過這些實驗，我們將確認預訓練模型是否正確學習到了地震波形中可泛化使用的特徵。這些探索將有助於進一步提升模型的性能和應用範圍。

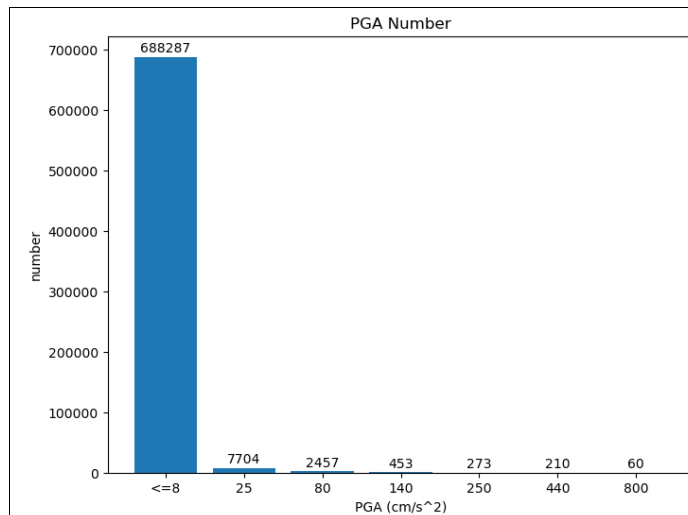
第四章 本研究提出之多測站地震預警模型

4.1 介紹

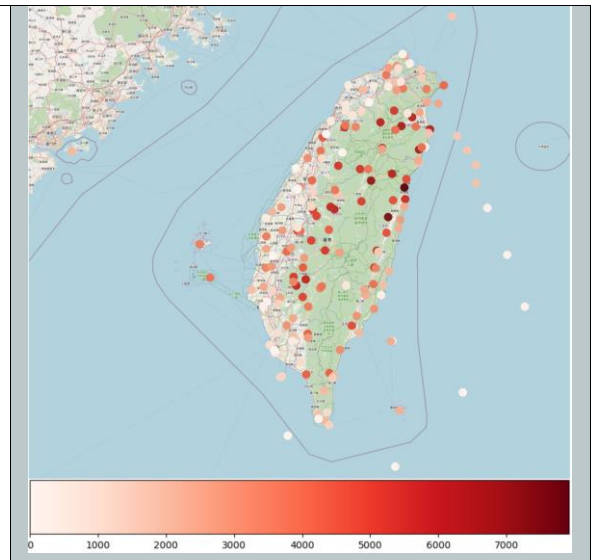
本研究著重在臺灣的區域型預警模型，利用 TEAM[3] 模型作為基底，訓練出提前預警時間 (leading time) 整體平均相較過去模型多 3 秒以上的表現。同時在最終運用模型時，分析藉由調整各測站機率閾值，以達全臺各區域皆有能力預警的目標。並且為了使訓練貼近實際運行，本實驗改善了資料集的一些部分，並在訓練時加入「未使用測站」的嵌入向量 (embedding vector)，此向量可在未來作為使模型永續經營的工具。

4.2 計畫背景與目的

地震預測任務本身存在著巨大的資料不平衡，統計 2012 至 2021 年中央氣象署地震觀測網 (CWASN) 中的有感地震事件，測站接收到的峰值地面加速度 (Peak Ground Acceleration) 中，如圖 4.1 所示，小於三級震度累計達六十餘萬次，相較大於三級震度只有一萬餘次。此外，事件測站的分佈也存在著區域不平衡，圖 4.2 為統計所有地震事件測站出現的次數圖，大多數事件集中在東臺灣以及中央山脈地區。為此，不僅在模型訓練時必須對資料做適當處理，實際應用時的調控也是緩解此種不平衡的方法之一。



[圖 4.1]



[圖 4.2]

目前的震度預估模型，主要以兩種形式作為呈現，一種是現地型預警模型，另一種是區域型預警模型。在現地型預警模型中，每個測站之間是彼此獨立的，一般做法是在接收到體波 (Primary wave) 訊號後，經過模型運算決定是否超過警報的閾值。此種作法的優點在於準確性較高，缺點是提前預警時間較短，並且在實際的

預警系統中，所有預警測站皆會受到前面的體波偵測模組影響。另一種作法是區域型預警模型，此種模型會以單測站或統合多個測站的資訊，判斷其他地區是否應該發布警報，優點是預警時間較長，且受到體波偵測模組的影響較小，缺點是準確性較差。

在地震模型的設計會有兩種主要的方向。分別是泛用性與針對性，TEAM 模型是以泛用性為主的模型方法，只需座標和波形的輸入。本實驗的目的在於訓練針對性的模型，專門預測全臺各地震度。因此模型的輸入是以資料集曾經出現過的所有測站，此種方法能夠提升許多的準確度，但在實際運用的永續性會是一大顧慮。當測站每隔一段時間進行更新，原本的輸入可能會無法使用，必須重新訓練才能再繼續使用。因此這個議題在本研究先以「未使用測站」作為測站標記，在模型訓練時標記出當下未使用的測站，一方面避免未使用測站讓模型學習混亂，另一方面提供模型一個區分測站的指標，在未來的研究可以嘗試將未學習過的測站套用在這些已知不再使用的測站中。

對於過去模型在實際系統的探討發現。資料集中存在一些可修改處。第一是有關加速度測站的選用，過去會包含一部分的速度型測站在其中，第二是有關測站漏值，訊號補零的現象，會因歸零校正波形而將零訊號變為突波，造成誤報。這些情況若有多加處理，可以使訓練更貼近實際需求。總結以上，本實驗想要使模型資料集更貼近實際系統運行情況，並且基於 TEAM 方法的區域型長預警時間基礎下，將準確性進一步提升，也朝模型的永續性展開初步實驗。

4.3 模型介紹

4.3.1 特徵提取

特徵提取是將已知現有知識，轉變為模型方法可以辨識、學習的特徵。本實驗會分為以下幾個方面進行：

- 加速度波形：主要以卷積神經網路 (convolutional neural network, CNN) 進行，首先將二維波形圖 (3000, 3)，其中 3000 為單筆波型資料之樣本點數，而 3 則為三分量的波型。透過二維卷積網路運算，學習三軸波形的特徵。再經過維度攤平變成一維波形，最後經過一維卷積網路將波形變為 500 維向量。
- 座標：將經度、緯度、深度，分別經過正弦餘弦處理，最終合併為 490 維度向量
- 未使用測站指標：將數值轉變為獨熱向量 (one-hot vector)，獲得 2 維度向量

在模型中，輸入測站與目標測站的特徵如下表 4.1。我們將輸入測站定為 25 個，目標測站定為資料集中所有的 249 個加速度測站。其中 25 個輸入測站可為任意測站。

表 4.1:輸入與目標測站之特徵

	特徵
輸入測站	加速度波型、座標
目標測站	座標、未使用測站指標

4.3.2 特徵結合

將特徵進行合併的方法，本實驗採取向量連接（concatenate）與向量相加的方式：

- 輸入測站座標、零填充連接
- 目標測站進行座標、未使用測站指標、零填充連接以上會得到各為（500,）維度的特徵向量
- 最後將輸入測站之波形向量和特徵向量進行向量相加

希望將以上之輸入測站、目標測站資訊結合，本實驗以變形模組（transformer）之多頭注意力機制為主體加以變化，以下為多頭注意力機制執行步驟：

- 建立三個權重矩陣，維度皆為（500, 500），特徵向量分別對其進行點積處理，獲得（274, 500）的查詢向量、鍵向量、值向量
- 將三個向量進行多頭注意力機制處理，本實驗設計以十個頭，將維度分別轉變為（10, 274, 50）、（10, 50, 274）和（10, 274, 50）
- 對查詢向量、鍵向量進行點積，獲得（10, 測站數, 測站數）的注意力矩陣，此矩陣代表分別在十個特徵資訊下，測站和測站之間的關係
- 將注意力矩陣對最後一軸進行歸一化指數函數（Softmax）啟動函式，轉換至 0 到 1 區間，代表在十個假定類別資訊間，每一個測站應該選取每一個自己與其他測站多大分量的特徵比例
- 將注意力矩陣和值向量進行點積，變換向量變為（274, 500），其中每個測站皆獲得十個隱藏類別的所有測站資訊

本實驗如同 TEAM，對於變形模組的變化為將目標測站、輸入測站未滿 25 個的測站之注意力矩陣減去很大的值，因此在歸一化指數函數後會是非常接近零的值，使得測站不會去學習目標測站和輸入測站未滿的部分，而是學習所有輸入測

站的特徵。

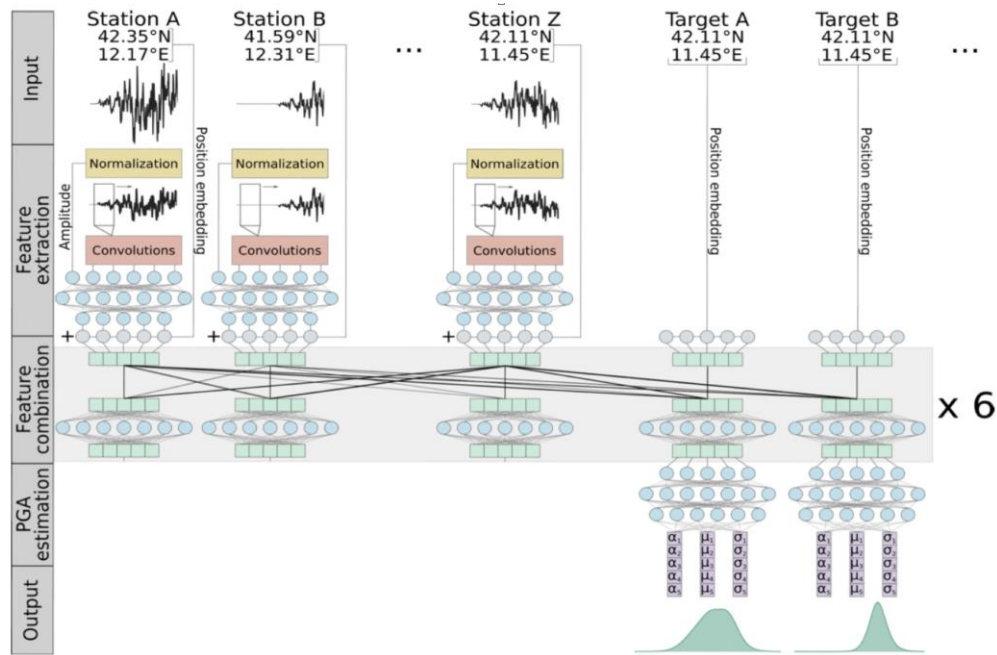


圖 4.3：模型架構

4.3.3 震度預測

經過以上變形模組，將會得到 $(274, 500)$ 的維度向量，我們選取目標測站的 $(249, 500)$ 維度向量，分別對各測站進行緻密層（Dense Layer）的點積處理，獲得各測站之混合高斯機率密度分佈模型參數：中位數、變異數以及混合權重比例，本實驗混合五組機率密度。因此最終獲得的維度向量為 $(249, 5, 3)$ 。每個測站皆有五組機率密度變數值，最終的震度預測階段將會計算各測站在給定加速度條件下的機率值，並人為給定每個測站在給定加速度下，超過多大的機率值應該發報。

本實驗將模型自第一個測站偵測到 p 波開始，每隔 0.8 秒輸出一個預測值，至第一個測站偵測到 p 波後 25 秒，因此在每個事件中皆有多個時間點下之預測值，最後再經由混淆指標的計算取得成績，混淆指標計算法如表 4.2 所示。實驗中有訓練多次模型，成績的挑選以四級之 F1 成績為優先，因此其餘震度會有所起伏。模型 F1 成績大約在 0.5，相較 TEAM 模型有平均 0.2 的進步幅度。在過去的 SENSE 模型方法中，由於是以 p 波偵測模組偵測到 p 波才發報的模型方法，因此在成績方面會優於本實驗模型。

表 4.2: 真陽性表示模型預測有發報，實際有發報，並且發報時間比實際時間更早；真陰性表示模型預測無法報，實際無發報；假陽性表示模型預測有發報，實際無法報；假陰性表示模型預測無發報，實際有發報，或是發報時間比實際時間晚

模型	真陽性 True Positive	真陰性 True Negative	假陽性 False Positive	假陰性 False Negative	假陰性 False Negative
標籤值 (label)	有值 (較大)	無值	無值	有值 (較小)	有值
預測值 (prediction)	有值 (較小)	無值	有值	有值 (較大)	無值

提前預警時間的部分，計算方式為發報時間點至 PGA 時間點的時間差，在 TEAM 和本實驗模型的表現是相近的，各震度約有 7 至 8 秒的預警時間；在 SENSE 模型中，預警時間則是 4 秒至 5 秒。以下成績大於四級之震度是採加速度值選取，以 gal 值 8、25、80、140、250、440、800 作為劃分，並以「類」為開頭作為命名。

表 4.3: 全測站以驗證集挑選之全測站 TEAM + 未使用測站指標模型成績

	三級	四級	類五弱	類五強	類六弱	類六強	類七級
準確度 (Accuracy)	0.996	0.998	0.999	0.999	0.999	0.999	0.999
精密度 (Precision)	0.691	0.603	0.485	0.537	0.589	0.404	0.310
召回率 (Recall)	0.416	0.495	0.479	0.457	0.412	0.548	0.562
F1-score	0.519	0.544	0.482	0.494	0.485	0.521	0.420

表 4.4: 全測站以驗證集挑選 alpha 之 F1 score 成績

模型	三級	四級	類五弱	類五強	類六弱	類六強	類七級
TEAM	0.363	0.373	0.361	0.362	0.327	0.444	0.372
全測站 TEAM + 未使用測站指標	0.519	0.544	0.482	0.494	0.485	0.521	0.420
SENSE	-	0.719	0.723	0.729	0.633	0.588	0.689

表 4.5: 全測站以驗證集挑選 alpha 之提前預警時間

模型	三級	四級	類五弱	類五強	類六弱	類六強	類七級
TEAM	10.35	7.67	8.57	7.97	6.94	6.35	7.53
全測站 TEAM + 未使用測站指標	9.99	8.47	8.23	7.59	8.13	7.74	8.70
SENSE	-	5.04	4.58	4.49	4.83	4.25	4.88

以上成績皆為所有測站在同一震度挑選相同的機率閾值，經實驗發現，在一些地震較不好發區域(舉例:本島西北部)，使用和地震好發區域相同之機率閾值，會造成無法發報情形。因此測站發報機率閾值的挑選，嘗試將全測站挑選相同的機率閾值，轉變成各測站挑選自己的機率閾值。實驗中可以發現，震度四級使用此挑選法可以改善臺北、桃園、宜蘭、花蓮一帶的不發報狀況;類五弱、類五強可以改善宜蘭、花蓮一帶的不發報狀況類;六弱以上的震度，則是表現類似的預測分佈;其中震度三級時使用此方法會造成反效果，若機率閾值設定過低，反而會導致過多測站發生誤報的情形。因此在挑選震度閾值時，震度四級、類五弱、類五強可視需求挑選是否要以各測站挑選自己的機率閾值，抉擇在於是否以較高的誤報率，換取不發報地區的發報機會。

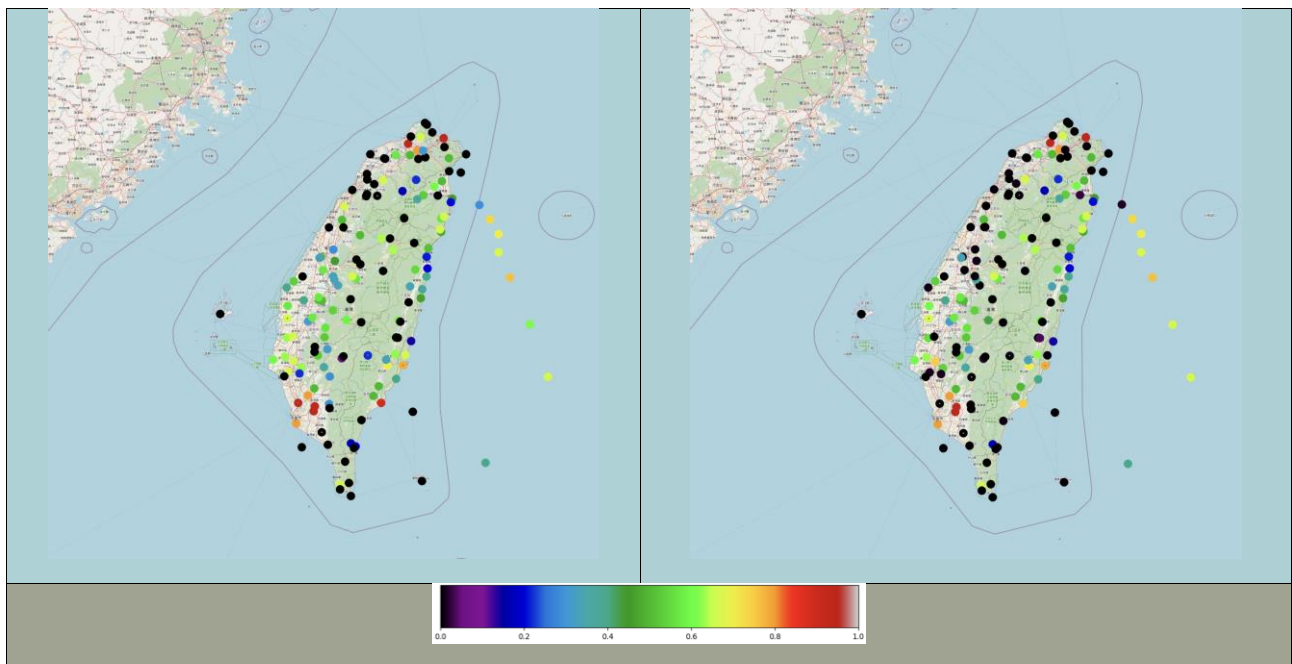


圖 4.4:三級全測站可視化 F1 score，左圖全測站挑選相同的機率閾值，右圖各測站挑選自己的機率閾值

表 4.6: 三級之混淆指標

模型	真陽性 True Positive	真陰性 True Negative	假陽性 False Positive	假陰性 False Negative	F1 score
全測站挑選相同的 機率閾值	1016	513531	453	1426	0.519
各測站挑選自己的 機率閾值	1200	439683	74301	1242	0.031

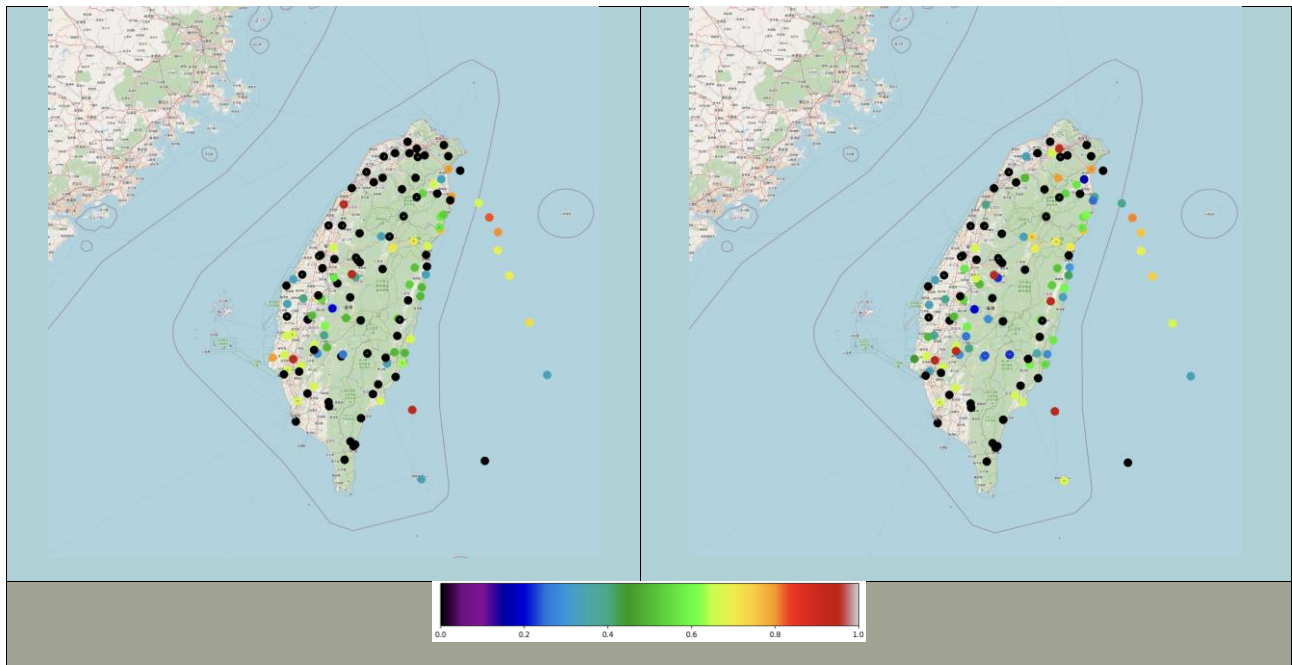


圖 4.5: 四級全測站可視化 F1 score，左圖全測站挑選相同的機率閾值，右圖各測站挑選自己的機率閾值

表 4.7: 四級之混淆指標

模型	真陽性 True Positive	真陰性 True Negative	假陽性 False Positive	假陰性 False Negative	F1 score
全測站挑選相同的 機率閾值	390	515383	256	397	0.544

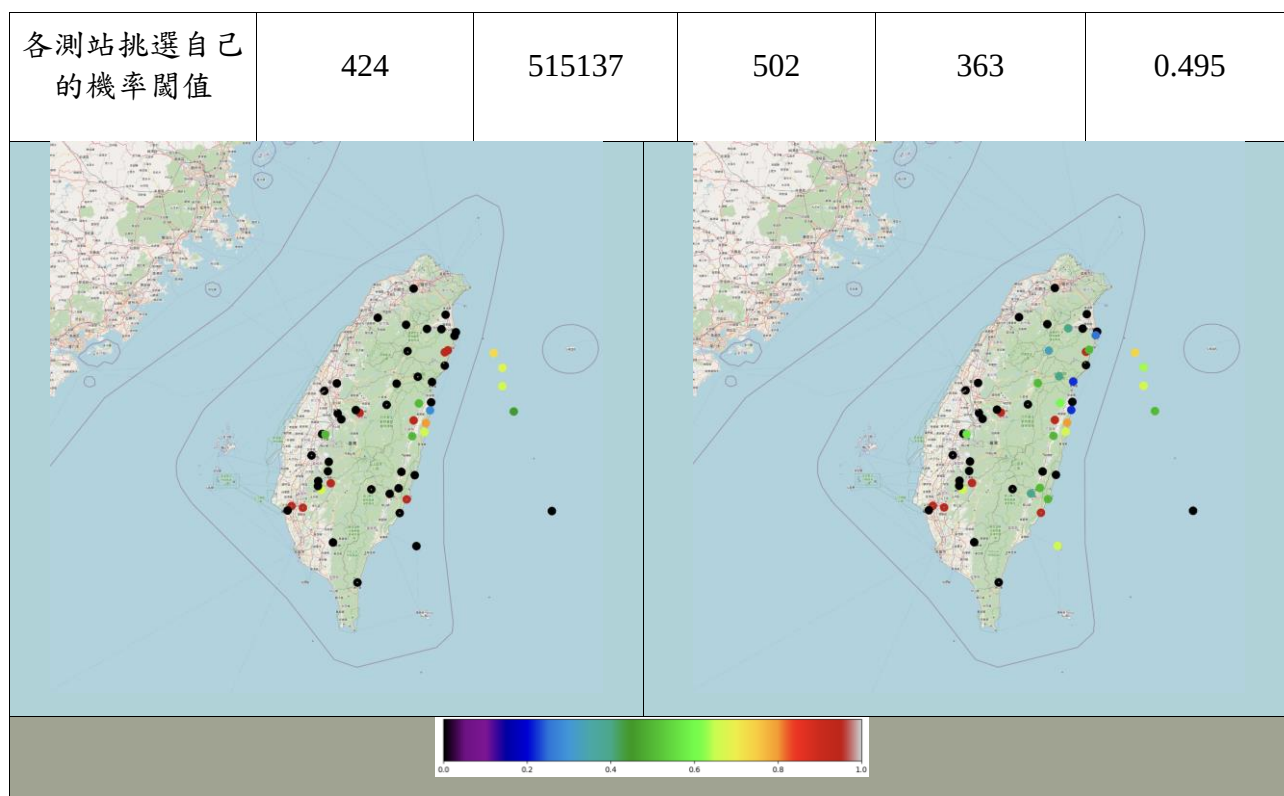


圖 4.6: 類五弱全測站可視化 F1 score，左圖全測站挑選相同的機率閾值，右圖各測站挑選自己的機率閾值

表 4.8: 類五弱之混淆指標

模型	真陽性 True Positive	真陰性 True Negative	假陽性 False Positive	假陰性 False Negative	F1 score
全測站挑選相同的 機率閾值	117	516058	124	127	0.482
各測站挑選自己的 機率閾值	125	515981	201	119	0.439

表 4.9: 類五強之混淆指標

模型	真陽性 True Positive	真陰性 True Negative	假陽性 False Positive	假陰性 False Negative	F1 score

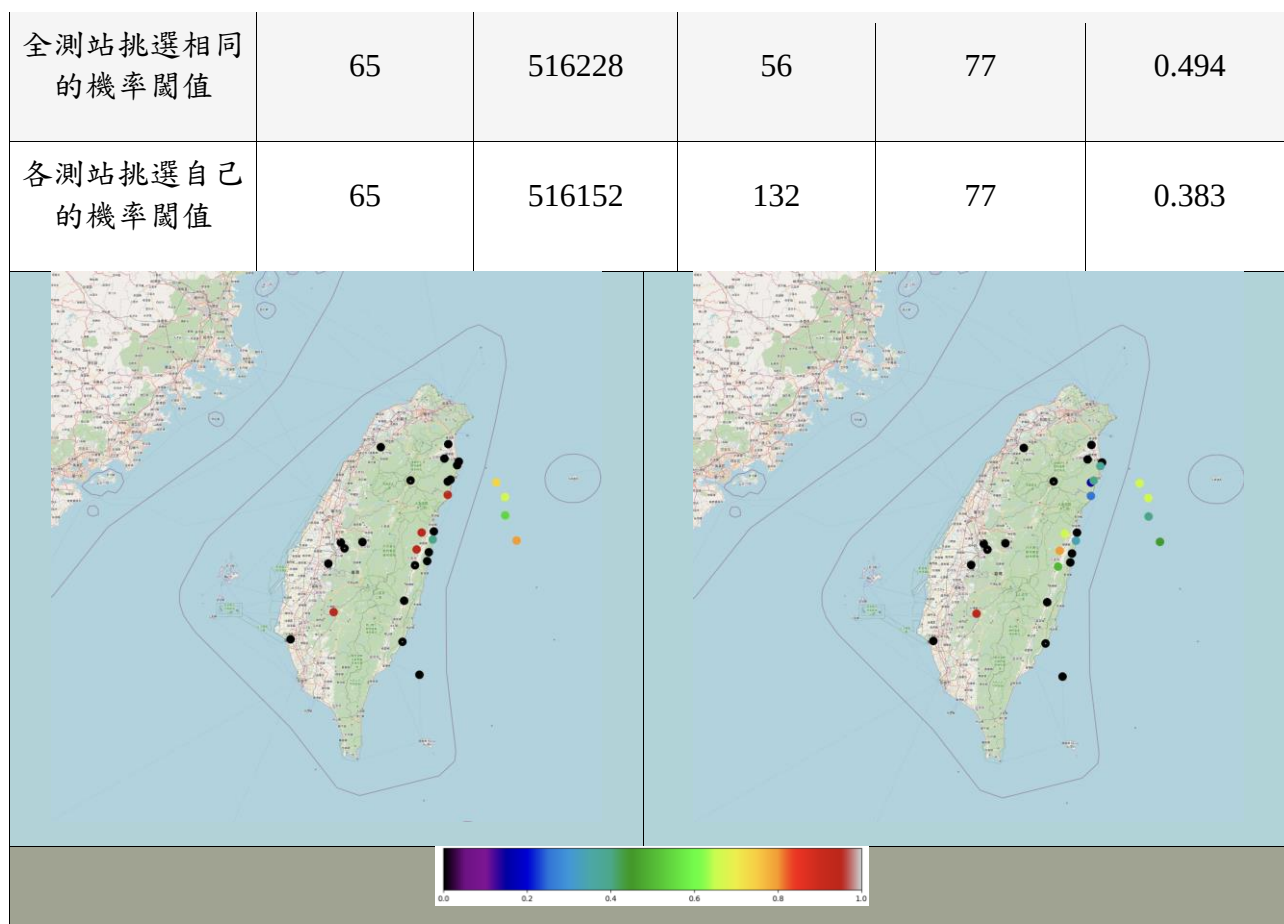


圖 4.7: 類五強全測站可視化 F1 score，左圖全測站挑選相同的機率閾值，右圖各測站挑選自己的機率閾值

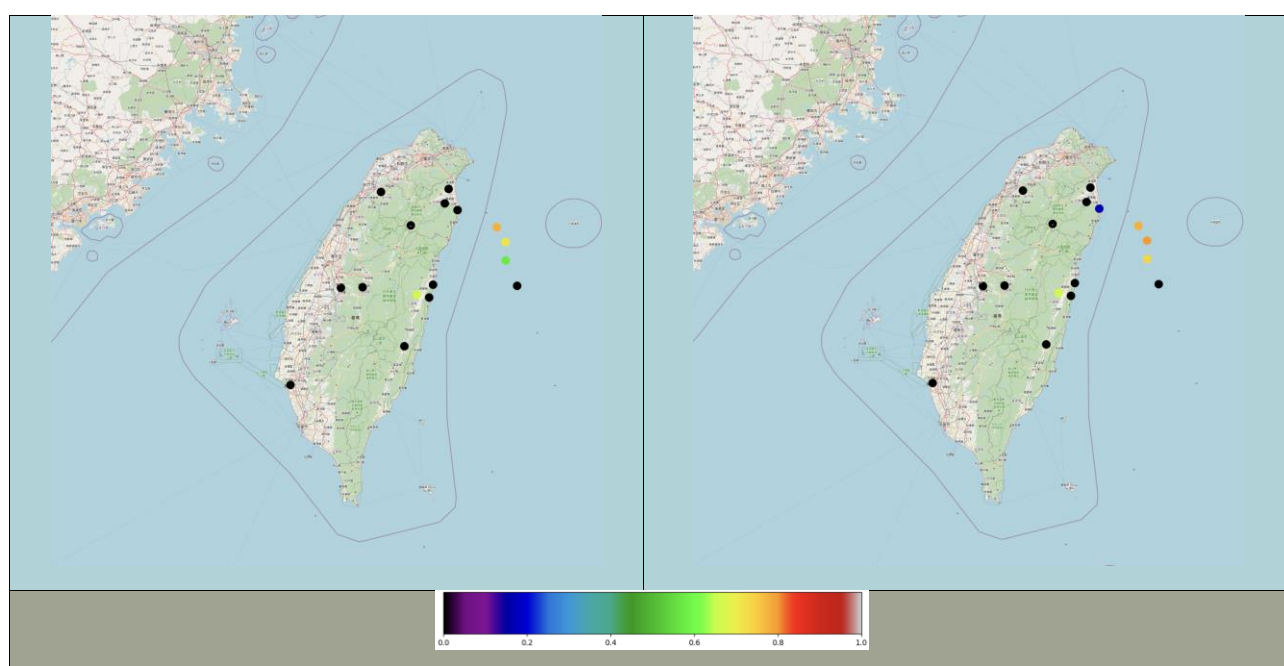


圖 4.8: 類六弱全測站可視化 F1 score，左圖全測站挑選相同的機率閾值，右圖各測站挑選自己的機率閾值

表 4.10: 類六弱之混淆指標

模型	真陽性 True Positive	真陰性 True Negative	假陽性 False Positive	假陰性 False Negative	F1 score
全測站挑選相同的 機率閾值	33	516323	23	47	0.485
各測站挑選自己的 機率閾值	34	516294	52	46	0.410

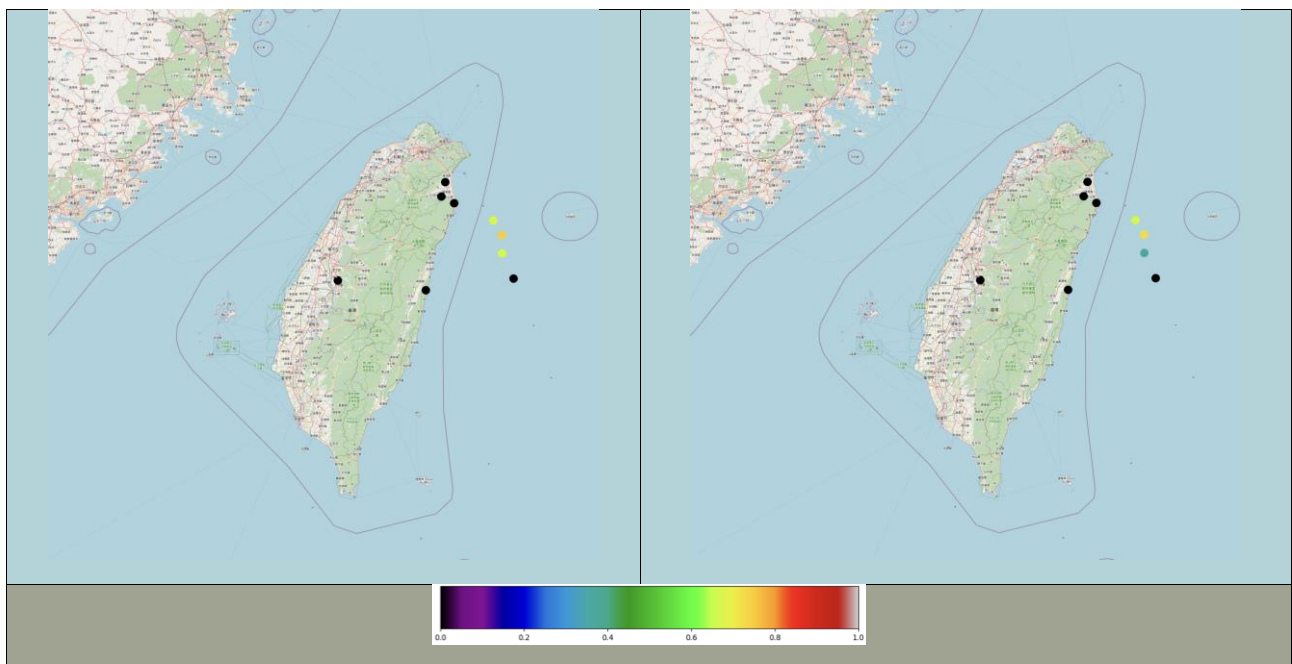


圖 4.9: 類六強全測站可視化 F1 score，左圖全測站挑選相同的機率閾值，右圖各測站挑選自己的機率閾值

表 4.11: 類六強之混淆指標

模型	真陽性 True Positive	真陰性 True Negative	假陽性 False Positive	假陰性 False Negative	F1 score
全測站挑選相同的 機率閾值	17	516370	25	14	0.521
各測站挑選自己的 機率閾值	16	516360	35	15	0.390

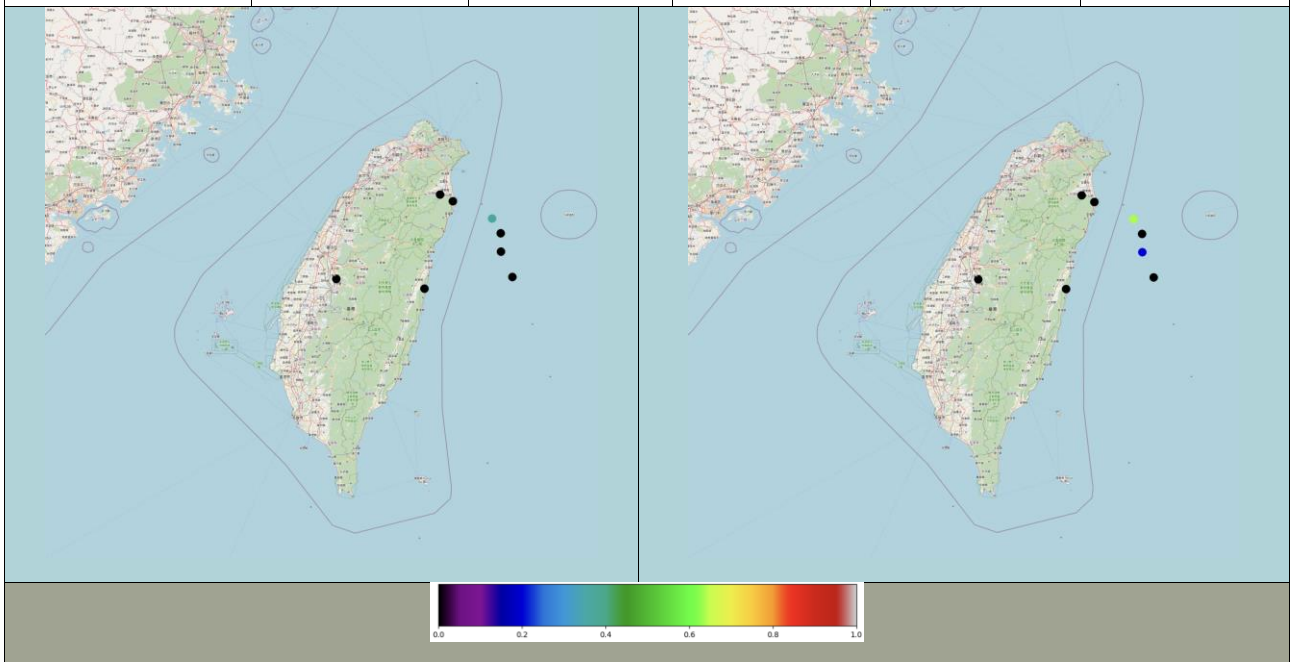


圖 4.10: 類七級全測站可視化 F1 score，左圖全測站挑選相同的機率閾值，右圖各測站挑選自己的機率閾值

表 4.12: 類七級之混淆指標

模型	真陽性 True Positive	真陰性 True Negative	假陽性 False Positive	假陰性 False Negative	F1 score
全測站挑選相同的 機率閾值	9	516390	20	7	0.420

各測站挑選自己的機率閾值	9	516381	29	7	0.333
--------------	---	--------	----	---	-------

4.3.4 事件分析

表 4.14 至表 4.18，挑選各地區的多測站事件分析，以四級地震為範例，黑色點為事件中大於四級震度的測站，白色點為事件中小於四級震度的測站，綠色點為提前發報之測站，紅色點為誤報之測站，黃色點為漏報之測站。秒數皆以第一個測站偵測到體波後開始計算，詳細的圖例解說如表 4.13 所示。此分析對四種模型方法做比較，分別是 TEAM 模型、未使用測站指標之全測站 TEAM、各測站機率閾值之未使用測站指標之全測站 TEAM、SENSE 模型。TEAM 為泛用性之模型方法。雖發報快速，但其誤報情形嚴重，並且誤報範圍廣。SENSE 模型預測的十分準確、測站誤報情況少，但在預測的時間上較慢。本實驗之兩種方法雖準確度較 SENSE 差，但完整的繼承了 TEAM 的快速發報，因此在絕大多數事件中，能於第一個測站偵測到體波後的幾秒鐘內預測大多數測站。在誤報的分佈上，大多位於應發報測站附近，較 TEAM 有顯著改善。

關於機率閾值（alpha）的挑選分析，挑選表 4.15 與表 4.16 此二影響橫跨東西之事件，此種類事件在模型的預測在模型的預測較為困難。在全測站挑選相同的機率閾值下，模型會如同絕大多數事件一樣預測半邊區域內的測站；但在各測站挑選自己的機率閾值下，模型有能力發報另一半邊的測站，雖然結果為誤報，但皆是位於鄰近的測站內。

總結，本實驗之模型在提前預警時間上表現傑出，雖在單點測站準確度上較為浮動，但在區域的預測上不會偏差太多。是否需要讓特定信心程度低之區域發報，則是一個可以依照策略作調控的選項。

表 4.13: 模型預測圖中點的呈現法

	初始狀態		提前		漏報	晚報	誤報
意義	測站 PGA 達四級	測站 PGA 未達四級	提前發報	提前發報，且四級時間已到	應發報，模型未發報	模型在四級時間後才發報	不應發報，模型發報
形態	黑	白	綠（大）	綠（小）	黃	紫	紅
時間軸始末			黑->綠（大）->綠（小）		黑->黃	黑->紫	白->紅

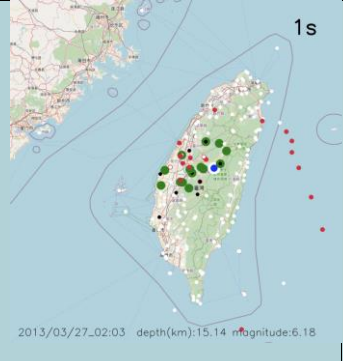
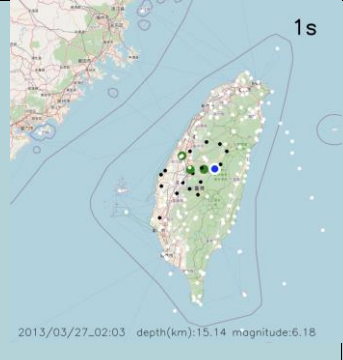
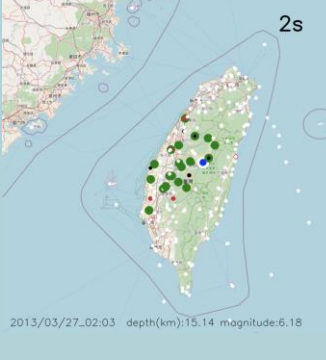
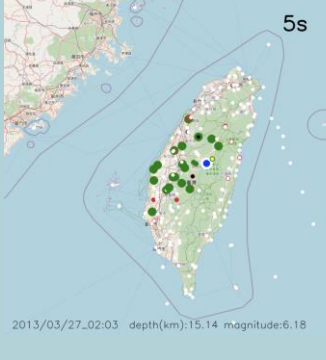
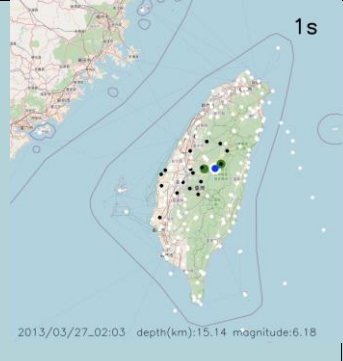
模型	1 秒	2 秒	5 秒
TEAM			
全測站 TEAM + 未使用測站指標			
全測站 TEAM + 未使用測站指標 測站 alpha			
SENSE			

表 4.14: 事件時間 2013 年 3 月 27 日 02 點 03 分，各模型自第一個測站偵測到體波後的模型預測狀況

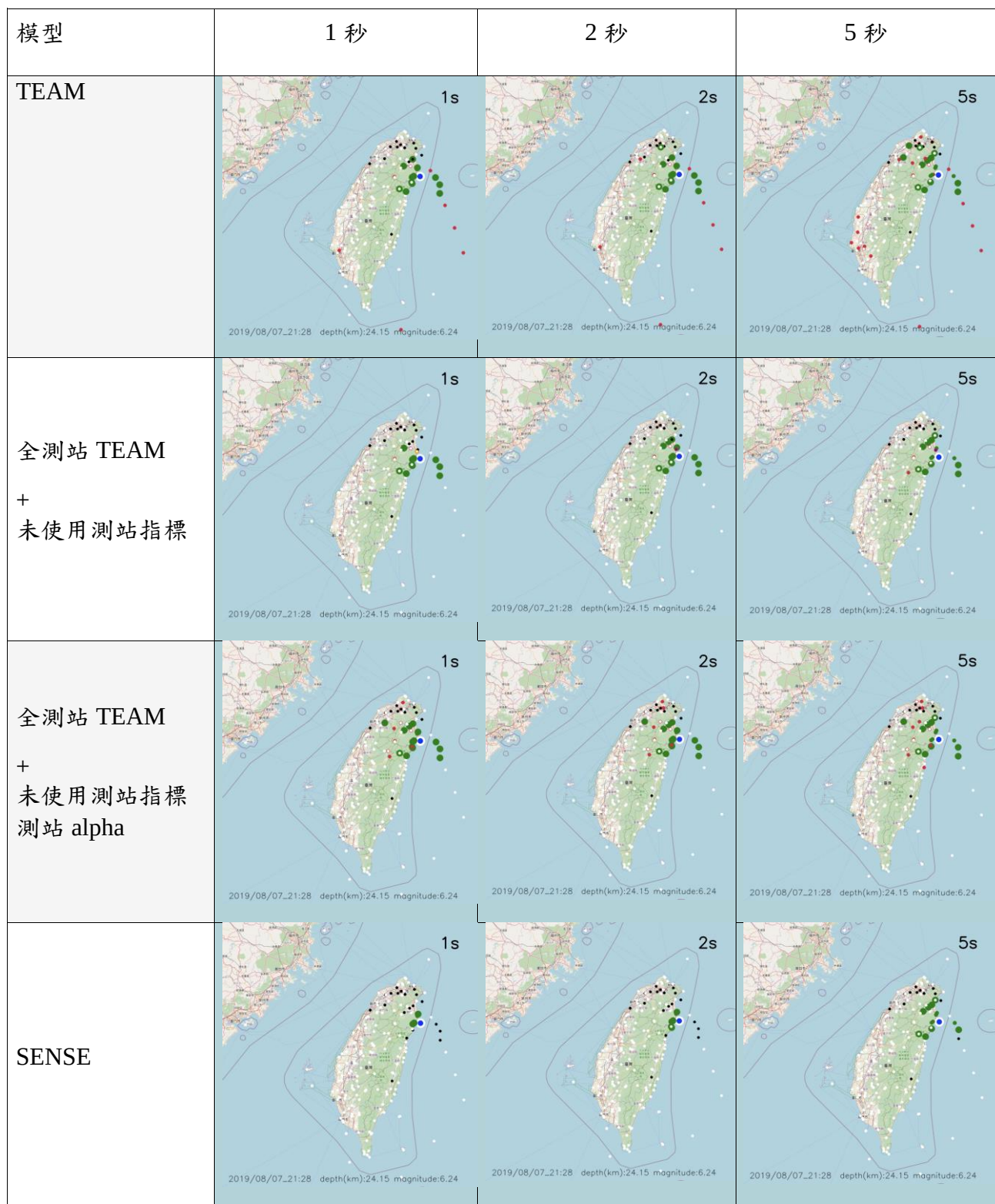


表 4.15: 事件時間 2019 年 8 月 7 日 21 點 28 分，各模型自第一個測站偵測到體波後的模型預測狀況

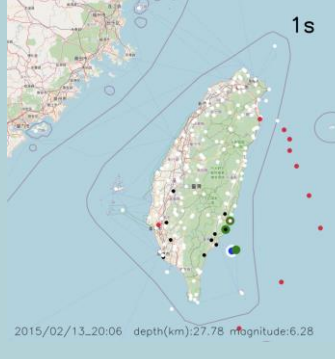
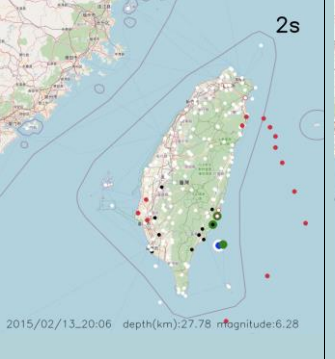
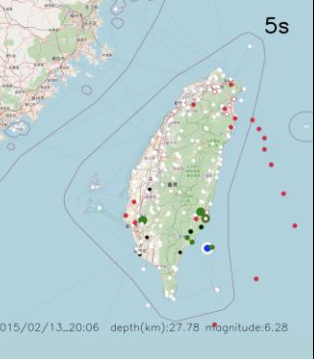
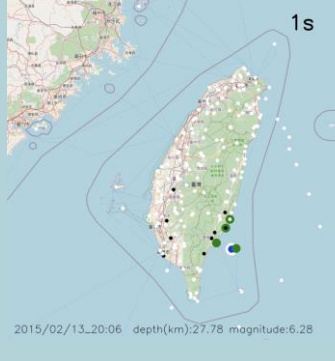
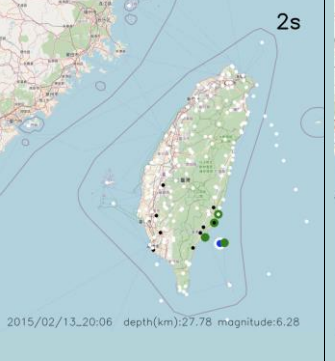
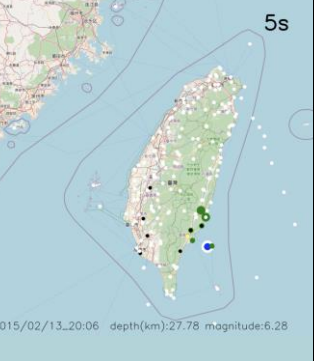
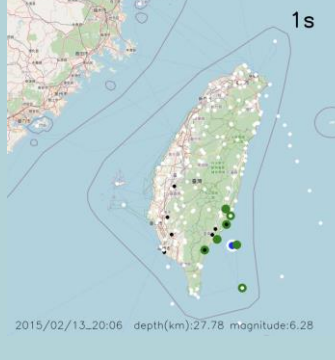
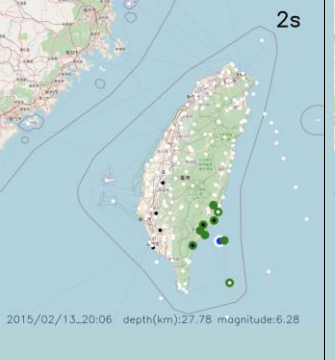
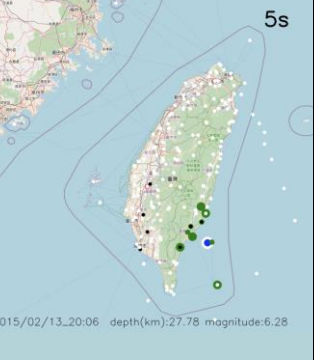
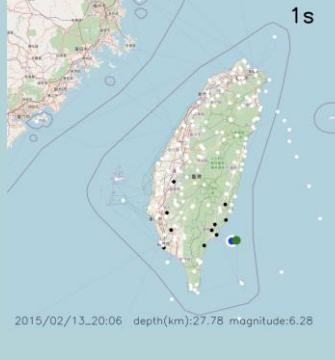
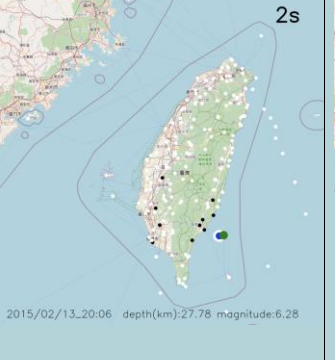
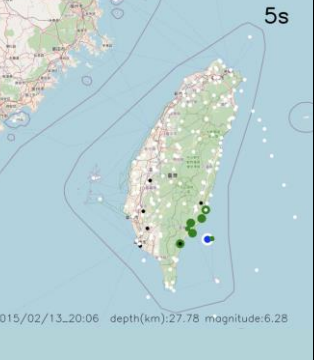
模型	1 秒	2 秒	5 秒
TEAM	 1s 2015/02/13_20:06 depth(km):27.78 magnitude:6.28	 2s 2015/02/13_20:06 depth(km):27.78 magnitude:6.28	 5s 2015/02/13_20:06 depth(km):27.78 magnitude:6.28
全測站 TEAM + 未使用測站指標	 1s 2015/02/13_20:06 depth(km):27.78 magnitude:6.28	 2s 2015/02/13_20:06 depth(km):27.78 magnitude:6.28	 5s 2015/02/13_20:06 depth(km):27.78 magnitude:6.28
全測站 TEAM + 未使用測站指標 測站 alpha	 1s 2015/02/13_20:06 depth(km):27.78 magnitude:6.28	 2s 2015/02/13_20:06 depth(km):27.78 magnitude:6.28	 5s 2015/02/13_20:06 depth(km):27.78 magnitude:6.28
SENSE	 1s 2015/02/13_20:06 depth(km):27.78 magnitude:6.28	 2s 2015/02/13_20:06 depth(km):27.78 magnitude:6.28	 5s 2015/02/13_20:06 depth(km):27.78 magnitude:6.28

表 4.16: 事件時間 2015 年 2 月 13 日 20 點 06 分，各模型自第一個測站偵測到體波後的模型預測狀況

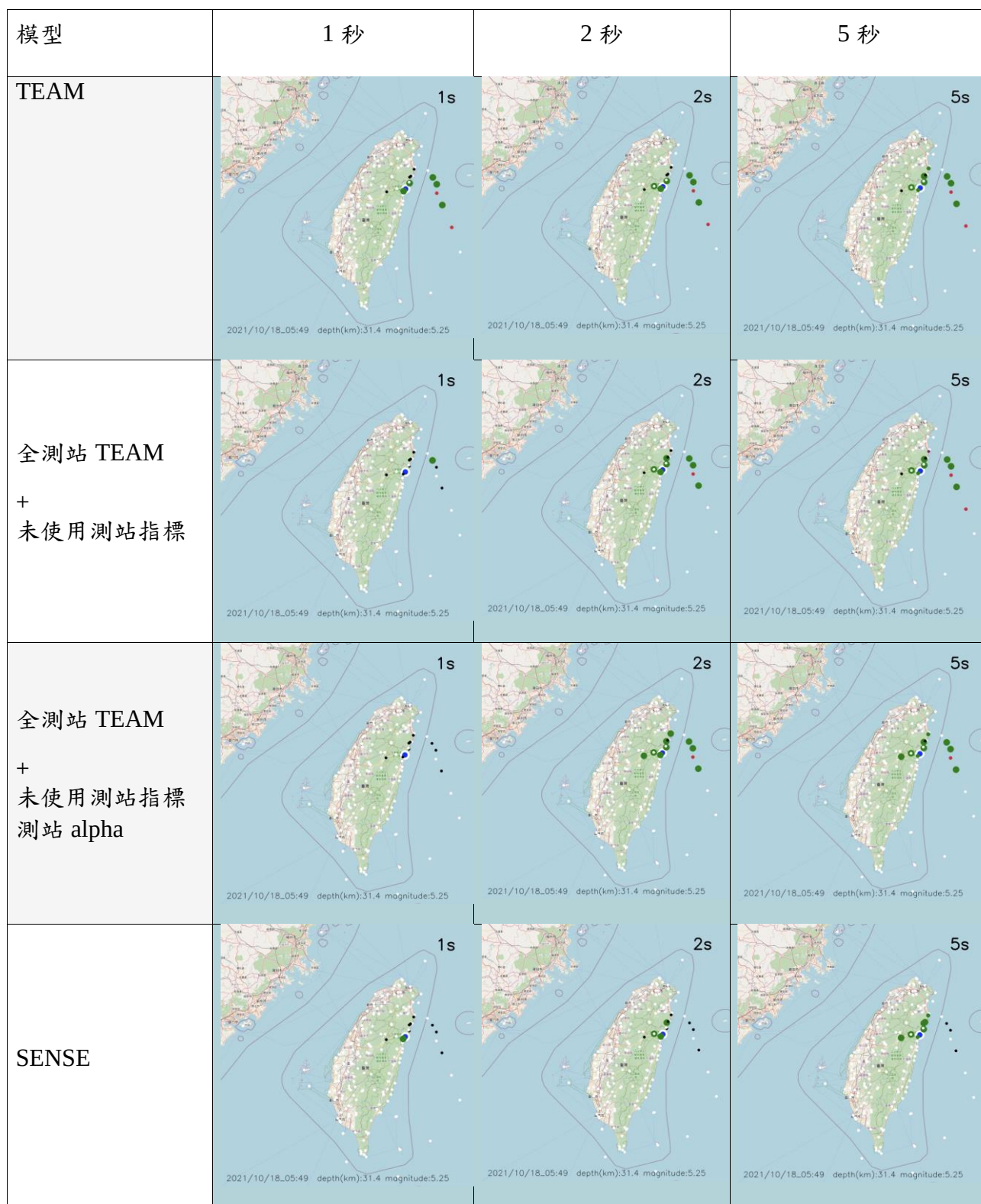


表 4.17: 事件時間 2021 年 10 月 18 日 5 點 49 分，各模型自第一個測站偵測到體波後的模型預測狀況

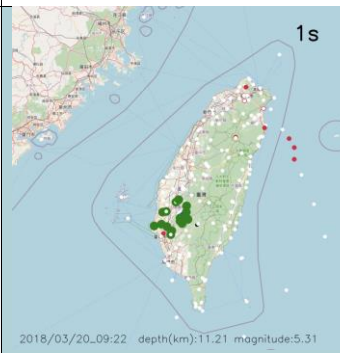
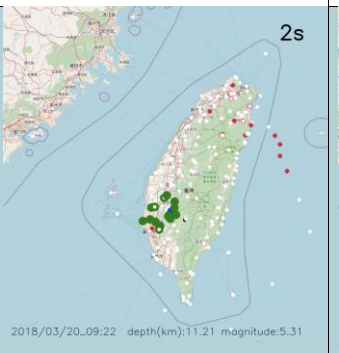
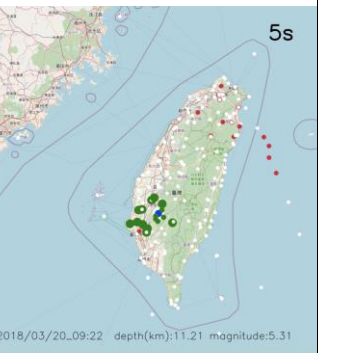
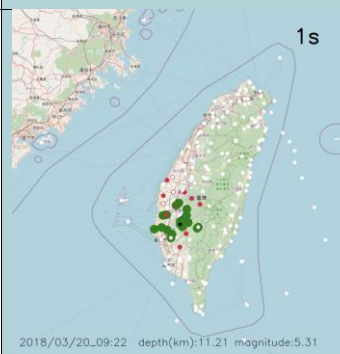
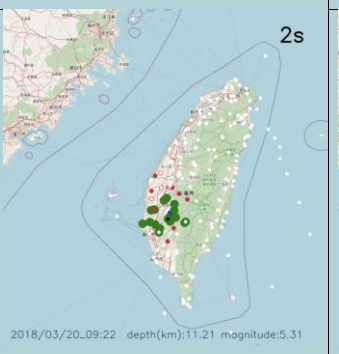
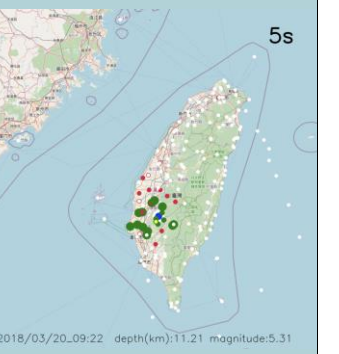
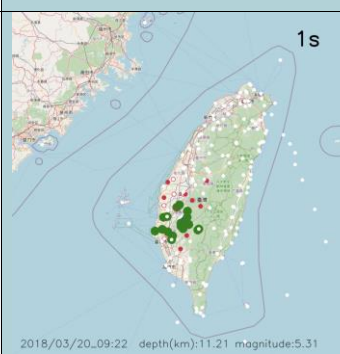
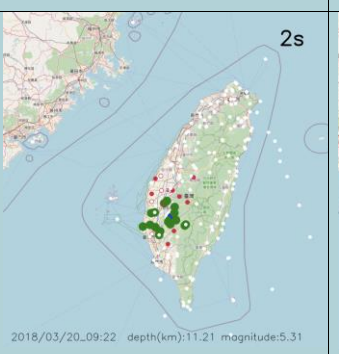
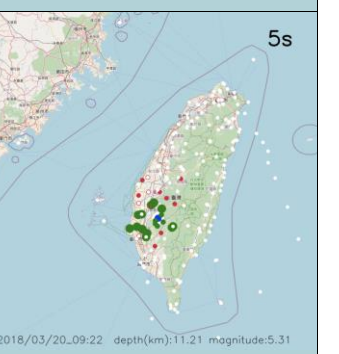
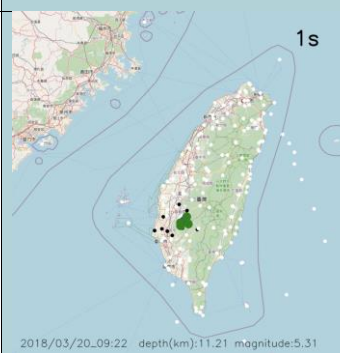
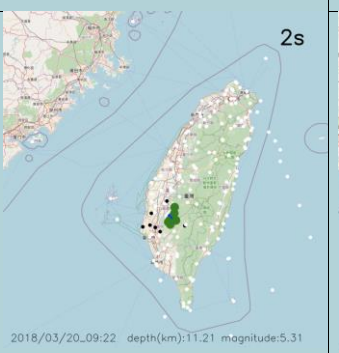
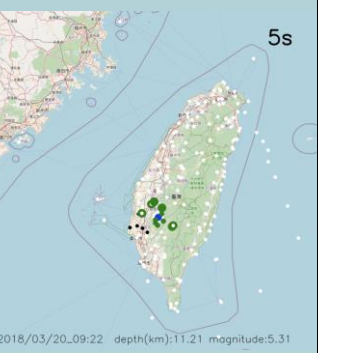
模型	1 秒	2 秒	5 秒
TEAM			
全測站 TEAM + 未使用測站指標			
全測站 TEAM + 未使用測站指標 測站 alpha			
SENSE			

表 4.18: 事件時間 2018 年 3 月 20 日 9 點 22 分，各模型自第一個測站偵測到體波後的模型預測狀況

4.3.5 訓練細節

區域型預警是一個對於資料集有較大要求的任務。為了避免缺值補零造成與實際運用不相符的現象發生，必須在一個事件中的各個測站皆有相同切割時間點的資料集，因此在中央氣象署地震觀測網(CWASN)和臺灣強地動觀測網(TSMIP)中，我們最終選擇使用中央氣象署地震觀測網的資料作為模型訓練集。並且，因本實驗想要以針對性的模型訓練為目標，沒有選擇非臺灣的資料集。

在資料集的處理上，挑選了 2012 至 2021 年在中央氣象署地震觀測網的所有 249 個加速度測站。在測站漏值補零部分，我們修改做法，只將波形非零的部分做歸零校正，這樣的做法在實際應用幾乎不會造成計算延誤的狀況，可以在很短的時間內將所有測站波形處理完，並且不會再發生零填充變為凸波的情形。如表 4.1 所示，僅是修正資料集，訓練成績提升約 10 至 20%。

在處理資料不平衡上，我們做了以下幾點的處理：

- 規模越大的事件，上採樣越多的次數
- 接收測站數越多的事件，上採樣越多的次數
- 對模型的損失函數（loss function）給予權重：依照標籤值之 PGA 大小，給予每個測站輸出不同的權重比例，劃分等級依照臺灣震度區間。

為了讓模型在三十秒中的各個時間點皆能產生預測，訓練階段會在 3000 個採樣點中隨機挑選一個採樣點，並將該採樣點後的所有波形變為零，以還原實際情況中只能看到一部分波形這點。

4.4 實驗結果

在訓練新資料集在初始 TEAM 模型上，我們做了一些實驗，原先資料集的標籤值為公尺每平方秒，在經過標籤值取 log 的處理後，如表 4.19，模型能有較好的學習效果。我們以表 4.19 中，TEAM 模型的最終訓練方法訓練 TEAM 全測站模型，首先先以單純右邊 249 測站訓練，再將未使用測站指標加入模型之中，結果如表 4.20 所示，在幾組訓練有發現到，表現普遍較沒有使用該指標的模型來的好，猜想是因為未使用測站在訓練上會造成模型小幅度的混亂，舉例兩地在一次事件中的標籤值是相同的震度，但另一起事件卻因為廢棄或尚未開始使用而造成某一個測站標籤值為零，而每個測站都會計算損失值，這樣的損失值差異便是可能造成模型成績稍差的原因。

表 4.19: 前四項為使用 TEAM 模型各種處理下之「部分測站」以測試集挑選 alpha 之 F1 score 成績，以部分測站計算成績是為了驗證資料集的正確性，因為原版本的 TEAM 是以部分測站成績做計算，除了此表其餘報告之成績皆是全測站成績，最後一項為全測站 TEAM 模型

處理	三級	四級	類五弱	類五強	類六弱	類六強	類七級
TEAM 未修正資料集	0.264	0.166	0.133	0.143	0.161	0.148	0.170
TEAM 修正資料集	0.394	0.324	0.237	0.237	0.234	0.197	0.228
TEAM log(標籤值)	0.490	0.443	0.349	0.380	0.410	0.438	0.461
TEAM 10*log(標籤值)	0.515	0.519	0.488	0.479	0.462	0.531	0.571
全測站 TEAM	0.588	0.593	0.532	0.525	0.492	0.571	0.500

表 4.20: 全測站以驗證集挑選 alpha 之 F1 score 成績

處理	三級	四級	類五弱	類五強	類六弱	類六強	類七級
TEAM	0.363	0.373	0.361	0.362	0.327	0.444	0.372
全測站 TEAM	0.515	0.516	0.497	0.487	0.427	0.508	0.428
全測站 TEAM + 未使用測站指 標	0.519	0.544	0.482	0.494	0.485	0.521	0.420

本實驗嘗試過一些其他模型方法，希望能幫助模型更加提升，很可惜皆未展現出足夠顯著的表現，以下表 4.21 羅列出嘗試過之模型更改。實驗主要圍繞在取消

原先在目標測站注意力矩陣上的遮罩，讓目標測站能彼此互相進行注意力機制，嘗試許多種訓練方式，發現成效皆不彰。此外，先前有取得全臺的場址效應 VS30 分佈表，實驗使用網格法將每個測站以外一定範圍內的 VS30 以固定排列順序排列成 250 維的向量，連接在原先特徵向量之後，但實際效果並不顯著。

表 4.21: 未有顯著效果方法之全測站以測試集挑選 alpha 之 F1 score 成績，將全測站 TEAM+未使用測站指標稱為「上游模型」

處理	三級	四級	類五弱	類五強	類六弱	類六強	類七級
未遮罩目標測站注意力矩陣之全測站 TEAM	0.319	0.243	0.221	0.254	0.309	0.333	0.350
凍結上游模型+下游 2 層全測站自注意力	0.524	0.530	0.487	0.498	0.480	0.508	0.470
凍結上游模型+下游 4 層全測站自注意力	0.474	0.548	0.503	0.496	0.467	0.508	0.482
上游模型+下游 2 層全測站自注意力	0.479	0.538	0.484	0.472	0.511	0.560	0.500
上游模型+前期特定地點向量	0.016	0.550	0.496	0.482	0.489	0.535	0.550
上游模型+前後期特定地點向量	0.020	0.559	0.490	0.472	0.459	0.571	0.500
凍結上游模型+下游 2 層全測站自注意力+前期特定地點向量	0.477	0.533	0.484	0.505	0.496	0.500	0.400
凍結上游模型+下游 4 層全測站自注意力+前期特定地點向量	0.516	0.545	0.496	0.498	0.478	0.500	0.538
凍結上游模型+下游 2 層全測站自注意力+前後期特定地點向量	0.466	0.535	0.488	0.484	0.486	0.444	0.425

凍結上游模型+下游 4 層全測站自注意力+前後期特定地點向量	0.494	0.545	0.490	0.492	0.485	0.485	0.444
特徵向量變為 750 維度連接法之上游模型	0.494	0.541	0.527	0.516	0.484	0.560	0.500
特徵向量變為 750 維度連接法之上游模型+下游 2 層全測站自注意力	0.016	0.548	0.517	0.516	0.470	0.566	0.500
特徵向量變為 750 維度連接法之上游模型+下游 4 層全測站自注意力	0.015	0.522	0.466	0.455	0.385	0.454	0.333
輸入測站特徵向量融入 VS30 之上游模型	0.510	0.512	0.490	0.459	0.503	0.566	0.538
輸入、目標測站特徵向量融入 VS30 之上游模型	0.016	0.548	0.517	0.516	0.470	0.566	0.500
凍結上游模型+下游 4 層全測站自注意力 + VS30	0.014	0.539	0.494	0.498	0.476	0.518	0.425

參考文獻

- [1] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, “Attention is all you need,” *Advances in neural information processing systems*, vol. 30, 2017.
- [2] A. Gulati, J. Qin, C.-C. Chiu, N. Parmar, Y. Zhang, J. Yu, W. Han, S. Wang, Z. Zhang, Y. Wu, et al., “Conformer: Convolution-augmented transformer for speech recognition,” *arXiv preprint arXiv:2005.08100*, 2020.
- [3] J. Münchmeyer, D. Bindi, U. Leser, and F. Tilmann, “The transformer earthquake alerting model: A new versatile approach to earthquake early warning,” *Geophysical Journal International*, vol. 225, no. 1, pp. 646–656, 2021.
- [4] C.-T. Lee and B.-R. Tsai, “Mapping vs30 in taiwan,” *TAO: Terrestrial, Atmospheric and Oceanic Sciences*, vol. 19, no. 6, p. 6, 2008.
- [5] J. Semblat and A. Pecker, “Waves and vibrations in soils: earthquakes, traffic, shocks, construction works (i. press, ed.),” 2009.
- [6] S. Bloemheuvel, J. v. d. Hoogen, D. Jozinović, A. Michelini, and M. Atzmueller, “Multivariate time series regression with graph neural networks,” *arXiv preprint arXiv:2201.00818*, 2022.
- [7] M. P. van den Ende and J.-P. Ampuero, “Automated seismic source characterization using deep graph neural networks,” *Geophysical Research Letters*, vol. 47, no. 17, p. e2020GL088690, 2020.
- [8] D.-Y. Chen, N.-C. Hsiao, and Y.-M. Wu, “The earthworm based earthquake alarm reporting system in taiwan,” *Bulletin of the Seismological Society of America*, vol. 105, no. 2A, pp. 568–579, 2015.
- [9] D. J. Wald, V. Quitoriano, T. H. Heaton, and H. Kanamori, “Relationships between peak ground acceleration, peak ground velocity, and modified mercalli intensity in california,” *Earthquake spectra*, vol. 15, no. 3, pp. 557–564, 1999.
- [10] L. Breiman, “Bagging predictors,” *Machine learning*, vol. 24, pp. 123–140, 1996.
- [11] Y.-J. Chiang, T.-L. Chin, and D.-Y. Chen, “Neural network-based strong motion prediction for on-site earthquake early warning,” *Sensors*, vol. 22, no. 3, p. 704, 2022.
- [12] S. M. Mousavi, W. L. Ellsworth, W. Zhu, L. Y. Chuang, and G. C. Beroza, “Earthquake transformer— an attentive deep-learning model for simultaneous earthquake detection and phase picking,” *Nature communications*, vol. 11, no. 1, p. 3952, 2020.
- [13] Zhou, X., J. Wang, Z. Cui, S. Zhang, Z. Yan, J. Zhou, and C. Zhou (2022). Mmspeech: Multi-modal multi-task encoder-decoder pre-training for speech recognition. *arXiv preprint arXiv:2212.00500*.
- [14] Watanabe, S., T. Hori, S. Kim, J. R. Hershey, and T. Hayashi (2017). Hybrid ctc/attention architecture for end-to-end speech recognition. *IEEE Journal of Selected Topics in Signal Processing* 11(8), 1240–1253.
- [15] Mousavi, S. M., Y. Sheng, W. Zhu, and G. C. Beroza (2019). Stanford

earthquake dataset (stead): A global data set of seismic signals for ai. IEEE Access 7, 179464–179476.

- [16] Zhu, W. and G. C. Beroza (2019). Phasenet: a deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International* 216(1), 261–273.
- [17] Liao, W.-Y., E.-J. Lee, D.-Y. Chen, P. Chen, D. Mu, and Y.-M. Wu (2022). Red-pan: Real-time earthquake detection and phase-picking with multitask attention network. *IEEE Transactions on Geoscience and Remote Sensing* 60, 1–11.
- [18] Schneider, S., Baevski, A., Collobert, R., & Auli, M. (2019). wav2vec: Unsupervised pre-training for speech recognition. *arXiv preprint arXiv:1904.05862*.
- [19] Baevski, A., Zhou, Y., Mohamed, A., & Auli, M. (2020). wav2vec 2.0: A framework for self-supervised learning of speech representations. *Advances in neural information processing systems*, 33, 12449-12460.
- [20] Hsu, W. N., Bolte, B., Tsai, Y. H. H., Lakhota, K., Salakhutdinov, R., & Mohamed, A. (2021). Hubert: Self-supervised speech representation learning by masked prediction of hidden units. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 29, 3451-3460.
- [21] Yang, S. W., Chi, P. H., Chuang, Y. S., Lai, C. I. J., Lakhota, K., Lin, Y. Y., & Lee, H. Y. (2021). Superb: Speech processing universal performance benchmark. *arXiv preprint arXiv:2105.01051*.
- [22] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- [23] Saad, O. M., Chen, Y., Savvaidis, A., Fomel, S., & Chen, Y. (2022). Real-time earthquake detection and magnitude estimation using vision transformer. *Journal of Geophysical Research: Solid Earth*, 127(5), e2021JB023657.
- [24] Zhou, T., Niu, P., Wang, X., Sun, L., & Jin, R. (2023). One Fits All: Power General Time Series Analysis by Pretrained LM. *arXiv preprint arXiv:2302.11939*.

機器學習技術建立現地型地震預警系統相關研究(II)

子計畫三

多尺度多域 P 波訊號輸入之 CNN 單站法預測 PGA 技術開發及其應用於 Earthworm 系統之研究

許丁友 陳柏睿 陳威樑 趙振源 陳婷萱 陳昱伯 官竑昱 田禹容 曾冠証
國立臺灣科技大學

中文摘要

現地型強震預警系統主要目的在於提前發出精準的警報，讓近震央區域有時間做好預防措施，以減少強震波帶來的傷亡與財損。目前現地型強震預警系統需要先判斷 P 波到時，之後利用 P 波到時後之訊號預測 PGA，據以對單一測站發出警報。

先前的研究已初步開發出基於 CNN 的單站法預測 PGA 技術。這個技術採用了 P 波後的兩秒內的原始加速度，經過多尺度和多域處理，作為 CNN 模型的輸入；並利用 TSMIP 的地震資料來訓練和驗證這個模型，且在歷次地震的離線驗證中已獲得不錯的成效。本研究旨在將這個 CNN 模型實際整合到 Earthworm 系統中，並結合 Twin 機制作為發報準則，以確保該模型在氣象署實時運行時，能提供更迅速且準確的預警。目前本研究已將 CNN 模型整合到 Earthworm 系統，在氣象署所提供之設備進行線上運算，並於過程中改善程式的問題，提高其運算之即時性與準確度，初步驗證該模型於實際地震事件預測成效不錯。

此外，目前判斷 P 波到時之模型以及預測 PGA 之模型係互相獨立運作。本研究另一目標為開發一模型可以同時處理這兩個問題，並設定在即時系統上運行。本研究以 Liao, et al. [1]所開發之即時標記地震波相模型架構為基礎，開發一可以同時標記地震波相並進行單站法強震預警之模型，並探討三軸向原始加速度做標準分數(Z-score)，並搭配三軸向原始加速度在不同輸入地表加速度歷時尺度組合之效果，所考慮的組合包括三軸原始加速度歷時、Z-score、Z-score&0.8gal、Z-score&2.5gal、Z-score&8gal、Z-score&25gal、Z-score&80gal、Z-score&250gal、Z-score&800gal 和 Z-score&原始加速度歷時。評估模型效果有四種方式，包括：(一)、觀察每個輸入在 Ground Truth 後 1 秒、2 秒、3 秒間，對 P 波到時之準確率、召回率和 F1-score，判斷訊號是否為地震訊號之準確率、召回率和 F1-score，預測 PGA 是否大於 25gal 之準確率、召回率和 F1-score。選出 Z-score&0.8gal、Z-score&2.5gal、Z-score&25gal 做以下評估方式之比較，(二)、各輸入對於小地震之誤報率。(三)、各輸入對於大地震之漏報率。(四)、評估各輸入在 Ground Truth 後每秒發報情形，根據(二)到(四)之結果，模型輸入為 Z-score&25gal 之效果比其餘兩種輸入效果好。透過分別跟 RED-PAN[1]和柏鈞模型[2]比較在標記波相和預測 PGA 是否大於 25gal 之二元混淆矩陣結果，驗證所提出之模型輸入 Z-score&25gal 確實可同時標記地震

波相及進行單站法強震預警，其準確率、召回率等與上述個別之模型相差不大。所以最後選擇 Z-score $\geq$ 25gal 當作模型之輸入。將其應用在 2022 年 9 月 18 日之池上地震結果為準確率為 84.43%、召回率為 41.53%、F1-score 為 55.67%。

關鍵詞：卷積神經網路、多任務學習、U-Net、單站法、波相標注、強震預警

Abstract

The primary objective of earthquake early warning systems is to issue precise alerts in advance, providing regions near the epicenter with sufficient time to implement precautionary measures and thereby mitigate the potential casualties and economic losses caused by strong seismic waves. Currently, On-site earthquake early warning systems relies on the P arrival, followed by it to predict Peak Ground Acceleration (PGA). Subsequently, based on this information, alerts are issued for individual stations.

Previous research had preliminarily developed a CNN model for predicting PGA. This technique utilizes raw acceleration data within two seconds after P arrival, undergoes multi-scale and multi-domain processing, inputting to the CNN model. The model is trained and validated using seismic data from TSMIP, demonstrating favorable performance in offline verifications across various seismic events. This study aims to practically integrate the CNN model into the Earthworm system in Central Weather Administration, incorporating a Twin mechanism as the reporting criterion to ensure swift and accurate warnings during real-time operation. The CNN model has been successfully integrated into the Earthworm system in the current study, utilizing equipment provided by the Central Weather Administration for online computation. Throughout this process, improvements have been made to address program issues, enhancing the real-time performance and accuracy of the system. The forthcoming discussion will present recent online prediction results for earthquake events as part of the research findings.

The current models for picking P arrival and predicting Peak Ground Acceleration (PGA) operate independently. The second objective of this report is to develop a model that can simultaneously handle both tasks and is designed for real-time system operation. The study builds on the real-time seismic phase picking model structure developed by Liao et al. [1], to create a model that can pick seismic phases and predict PGA for single station. It investigates the effectiveness of using the three-axis acceleration time history through standard scores (Z-scores) combined with different input scales of ground acceleration duration. The combinations considered include three-axis acceleration time history without standard scoring or normalization, Z-score, Z-score & various thresholds of acceleration (0.8gal to 800gal), and Z-score & three-axis acceleration time history. Evaluation of model performance involves four methods: (1) Observing the accuracy, recall, and F1-score of P-wave arrival times within 1, 2, and 3 seconds after the Ground Truth, assessing the accuracy, recall, and F1-score of seismic signal identification, and predicting the accuracy, recall, and F1-score of PGA exceeding 25gal. The selected criteria for comparison include Z-score $\geq$ 0.8gal, Z-score $\geq$ 2.5gal, and Z-score $\geq$ 25gal. (2) Assessing the false alarm rate for each input in detecting small earthquakes. (3) Evaluating the missed detection rate for large earthquakes in each input. (4) Assessing the reporting situation of each input after the Ground Truth for each second. Based on the results of (2) to (4), the model input Z-score $\geq$ 25gal demonstrated superior

performance compared to the other two input methods. Comparisons were made with the RED-PAN[1] and the model proposed by Bojun[2], utilizing binary confusion matrices for marking wave phases and predicting PGA exceeding 25gal. The results validated that the proposed Z-score&25gal input method can effectively mark seismic wave phases and perform single-station strong earthquake early warning. Its accuracy, recall, and other metrics show negligible differences compared to the individual models mentioned earlier. Therefore, Z-score&25gal was ultimately chosen as the input for the model.

Keywords: Convolutional Neural Network、Multi-Task、U-Net、Single Station、Phase picking、earthquake early warning

1 前言

快速且精確地預測地震震度是強震預警技術研發的核心目標之一。儘管多站法（區域型）強震預警技術能迅速且確切地估算各地的震度，但需要多個測站接收到 phase 資訊才能進行預測。因此，對於接近震央的地區，往往在強震波抵達後預警才發出，這使得該地將成為預警盲區。為此，近震央地區特別適合採用現地法預測。先前的研究已初步開發出基於 CNN 的單站法預測 PGA 技術。這個技術採用了 P 波後的兩秒內的原始加速度，經過多尺度和多域處理，作為 CNN 模型的輸入；而本研究之目標其一為將所初步開發之 CNN 模型應用於 Earthworm 系統。

本報告分兩部分，第一部分旨在將所初步開發之 CNN 模型應用於 Earthworm 系統，期能增加近震央地區的預警時效，以縮小盲區。同時，為了減少誤報的發生，本研究嘗試使用多測站預測的 PGA 閾值作為測站發布條件。此部分主要於第二章進行介紹，其內容包含(1)CNN 模型相關架構及內容；(2)現地形警報與發報邏輯；(3)統計準確性指標-二元混淆矩陣；(4)模型即時運行上的問題與解決；(5)近期事件討論；(6)小結。

本計畫另一目標為開發一個可同時具備標注 P、S 波和強震預警之模型，由於以往的研究幾乎都是分別針對辨別地震波的 P 波到時，以及預測地表最大加速度 (PGA) 等兩個動作，發展各自之預測模式。且關於辨別地震波的 P、S 波到時之機器學習模型，幾乎都是使用地震過後的資料辨別 P、S 波到時，很少針對運用在即時系統上標記地震波相之模型。近年開始有論文針對在即時系統上標記地震波相之模型[1]。至於運用 P 波到時後幾秒的資訊預測地震之震度或 PGA 等之模型，幾乎都假設 P 波到時已經被順利找到，運用準確之 P 波到時後之資訊來進行預測。因此本研究據 Liao, et al. [1] 提出之模型為基礎，進而開發一模型，可即時分析單一測站之加速度歷時資料並得知以下三個目標，(一)、辨別地震波相之 P、S 波，(二)、不易受非地震事件影像而誤判，(三)、預測該地震事件的 PGA 是否超過 25gal。

2 時頻域多尺度預測 PGA 之 CNN 模型

本研究中所使用的卷積神經網路 (CNN) 模型架構與輸入資料轉換方式參考黃昭文[3]所提出之時頻域多尺度之 CNN 預測 PGA 模型。時頻域指的是時間域與頻率域，多尺度則是不同的標準化尺度。以下將針對 CNN 訓練模型資料、模型架構、警報發布邏輯、統計指標、地震模擬分析進行說明。

建置模型所使用的資料來源為臺灣中央氣象署強震紀錄資料庫(TSMIP)，涵蓋 1992 年至 2006 年的地震紀錄，包含 921 集集地震、1022 嘉義地震等具代表性的強震，共計 91140 筆地震歷時資料，資料採樣頻率為 200 赫茲。為維持不同震度資料的平衡，僅選取其中的 10000 筆地震加速度歷時資料；且為確保訓練資料的代表性，將最大地表加速度對應舊制分級的 1、5、6、7 級以上的資料全部取用，而剩餘的資料則由舊制分級的 2、3、4 級資料平均分配數量。各加速度區間資料量分布如下表 2.1-1 所示。

表 2.1-1 訓練資料各加速度區間資料分布

加速度 (cm/ s ²)	0.8~2.5	2.5~8	8~25	25~80	80~250	250~400	400 以上	總和
筆數	117	2254	2254	2254	2845	198	78	10000

建構好地震資料後，依照模型需求將 P 波抵達後前 1、2 與 3 秒的三軸向加速度歷時資料分別進行時頻域多尺度轉換，最後標籤上相對應的最大地表加速度值。我們選擇 8000 筆資料進行訓練與驗證，並將 2000 筆作為測試資料，以完成訓練與測試資料的準備。

2.1 模型架構

根據黃昭文的研究[3]，其 CNN 模型係結合頻譜資訊與加速度歷時資料至同一矩陣中作為新的輸入，在時間域之尺度為 $\pm 2.5\text{gal}$ 、 $\pm 25\text{gal}$ 、 $\pm 250\text{gal}$ 三種，頻率域之尺度為 0~1、0~20 兩種，如此可將同一筆地震資料以多種不同尺度範圍表示，並將其合併至同一矩陣中，使單一輸入資料中能包含多種不同的資訊。

P 波到時後 3 秒模型的架構共有三組卷積層與池化層，分別是 150*1 的卷積核配合 3*1 的池化層、5*3 的卷積核配合 3*1 的池化層、1*3 的卷積核配合 3*1 的池化層。在資料經過卷積層與最大池化層後，通常為三維的資料。若要將資料輸入全層連接層，則需要將三維的資料透過平坦層展開為一維的資料後再輸入全層連接層，最後輸出 PGA 值。詳細結構如圖 2.1-1，模型詳細參數如

表 2.1-1 所示。訓練時 Optimizer 採用 Adam，因為其收斂速度快且能自行調整學習速率；Loss Function 採用均方根對數誤差(RMSLE)，基本運算如式(2.1-1)。此法將數值取對數能使數值大小降低到適當的範圍內，降低數值大的資料權重影響；Activation Function 採用 ReLU，能幫助卷積運算更有效率的萃取影像特徵，batch size 為 64。本研究中會使用 P 波到時後 2 秒模型對現地型警報進行評估。

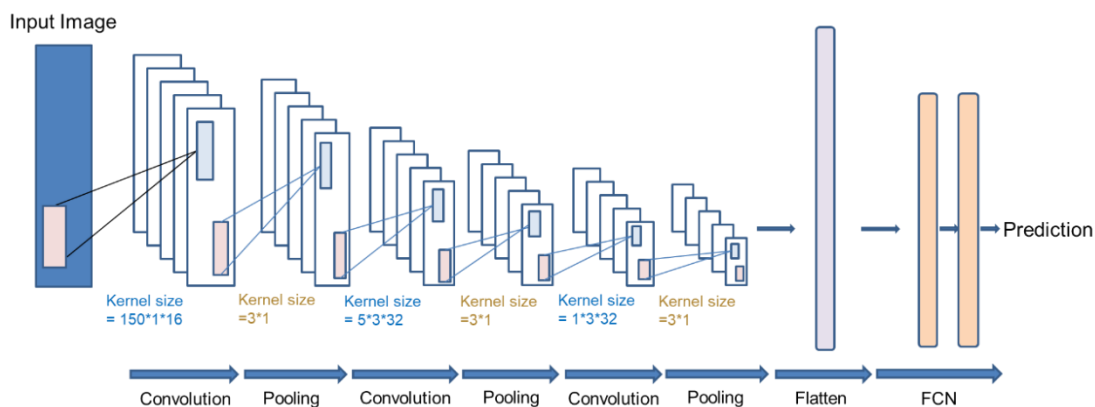


圖 2.1-1 卷積神經網路結構

表 2.1-1 模型參數表

Optimizer	Adam
Activation Function	ReLU
Loss Function	RMSLE
Batch Size	64

$$RMSLE = \sqrt{\frac{1}{n} \sum_{i=1}^n [\log(y_{i,true}) - \log(y_{i,predict})]^2} \quad (2.1-1)$$

2.2 警報發布邏輯

2.2.1 現地型警報

臺灣的 TSMIP 測站共設有 855 個測站，中央氣象署計畫將所有測站都改為即時測站，而目前其中有 553 個測站能夠傳輸即時資料。本研究以此為前提，預期在所有測站都改為即時測站後，現地型警報的成效將會提升。故本研究使用臺灣本島的 839 個測站進行統計，並假設所有測站都能傳輸即時資料。測站在臺灣的分布如圖 2.2-1 所示。在進行地震模擬時，若測站偵測到 P 波，則擷取 P 波到時後 2 秒的三軸向加速度歷時訊號，再進行時頻域多尺度轉換並輸入至 CNN 模型中，CNN 模型的輸出為測站的預測 PGA。

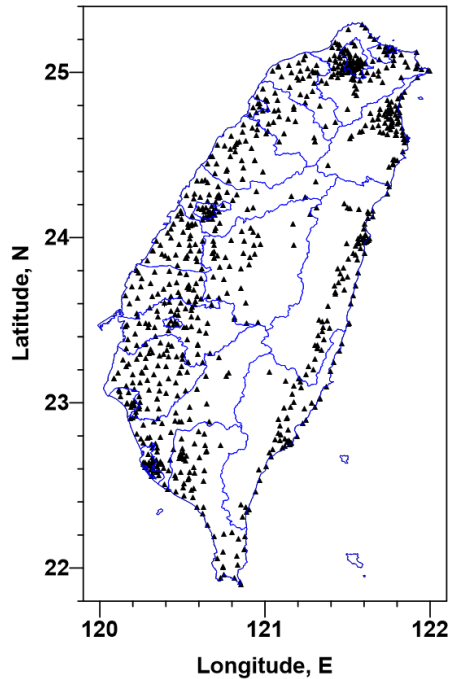


圖 2.2-1 TSMIP 測站分布

2.2.2 警報發布邏輯以及 Twin 發報機制

現地型警報發布邏輯可以分為單站法和多站判斷法。單站法是直接以測站的預測 PGA 是否大於門檻值 25gal 作為發報標準。由於單站法的發報條件過於簡單，容易出現誤報。因此，於上期期末報告中，本研究提出了多站判斷法來減少誤報的發生。多站判斷法的概念是當多個測站的預測 PGA 都大於門檻值時，才會予以發布警報。

本研究提出的多站判斷法機制為當距離最近的兩個測站，於設定時間區間內（目前設定為一分鐘），預測結果皆大於門檻值 25 gal 時，才會予以發報，本研究稱此機制為 Twin。示意如圖 2.2-2，圖中三角形皆為測站，若以紅色三角形為基準（是否要發布警報的測站），其最近距離之測站為橘色三角形測站，當事件發生時，假如紅色三角形測站預測 PGA 大於 25gal，則僅當在設定時間區間內，橘色三角形測站預測 PGA 也大於 25gal，才會發報；若是僅其餘藍色三角形測站預測 PGA 大於 25gal，則不會發報。



圖 2.2-2 多站判斷法與鄰近測站相關示意圖

此外，目前發報是以縣市為單位，因此在建立最近測站關係時，應以同縣市為

條件。如:當 A 測站最近距離之測站為 B 測站，但 A 測站位於花蓮縣，但 B 測站位於臺東縣，則在 Twin 機制中不能將這兩測站視為最近距離之測站，而是要在同縣市找出最近距離之測站。

舉例來說，當測站 A 最近測站為測站 B，且皆位於相同縣市內，則發報流程如圖 2.2-3。

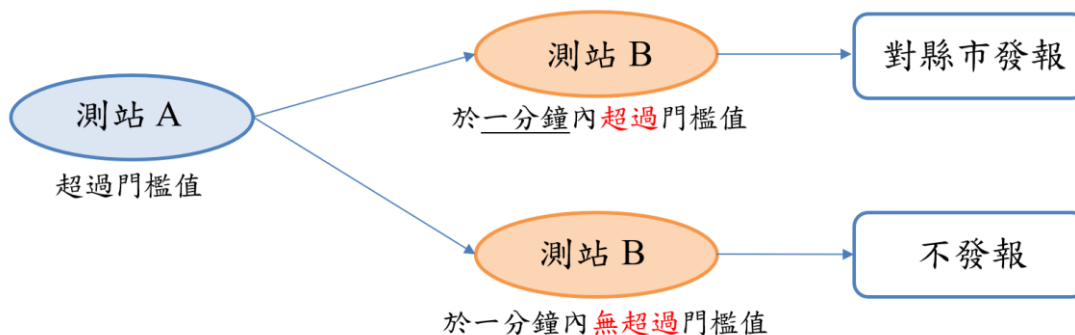


圖 2.2-3 實際發報流程圖

2.3 統計指標

2.3.1 統計指標 – 二元混淆矩陣(Confusion Matrix)

單站法統計會以單一測站為單位進行地震警報發布統計，是否需要發布地震警報以及是否成功發布預警之定義如下：計算出各測站真實 PGA，若大於門檻值，該測站則須發報；Twin 機制會以臺灣本島 17 個縣市為單位進行地震警報發布統計，是否需要發布地震警報以及是否成功發布預警之定義如下：計算出各測站真實 PGA，若大於門檻值，該縣市則須發報。

本研究使用二元混淆矩陣(Confusion Matrix)對警報發布的正確性進行判斷，如圖 2.3-1。本研究以 25gal 為警報門檻值，並可考慮適當之容許的誤差範圍，容許誤差的意思為若有發出警報，但是實際 PGA 在 8~25gal 之間，雖然實際 PGA 沒有高於門檻值未達到發報標準，但是 8~25gal 的 PGA 在某些區域可能會有民眾感到明顯震動，因此仍然將此事件歸類為正確發布；若未發出警報，但是實際 PGA 在 25~80gal 之間，雖然實際 PGA 已經高於門檻值達到發報標準，但是 25~80gal 的 PGA 還不會對建築物以及民眾造成危害，因此仍然將此事件歸類為正確未發布，含容許誤差之二元混淆矩陣，如圖 2.3-2。

二元混淆矩陣判斷，對應本研究指標為: TP 為正確發布，意即該地區依照真實 PGA 達發報標準，且預測 PGA 也達發報標準；TN 為正確不發布，意即該地區依照真實 PGA 未達發報標準，且預測 PGA 也未達發報標準；FP 為錯誤發布，意即該地區依照真實 PGA 未達發報標準，但預測 PGA 達發報標準；FN 為錯誤不發布，意即該地區依照真實 PGA 達發報標準，但預測 PGA 未達發報標準。最後在由 TP、FP、TN 與 FN，計算出精確率(Precision)、召回率(Recall)公式如式(2.3-1)、(2.3-2)。

		25 gal	
Predicted	Positive	FP	TP
	Negative	TN	FN
		Negative	Positive
		Real	

圖 2.3-1 二元混淆矩陣

		8 gal	
Predicted	Positive	FP	TP
	Negative	TN	FN
		Negative	Positive
		Real	

圖 2.3-2 容許誤差之二元混淆矩陣

$$Precision = \frac{TP}{TP + FP} \tag{2.3-1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2.3-2}$$

2.4 Earthworm 系統

目前於氣象署運行之 Earthworm 系統流程，如圖 2.4-1 所示，從各測站輸入之

波型資料，會寫入共享記憶體 WAVE_RING，接著經過 PICKER 標記 P 波到時後，亦會寫入共享記憶體 PICK_RING。

現地型警報可以利用 Python 套件，從 WAVE_RING 與 PICK_RING 中得到地震波型資料與 P 波到時，經過資料處理後，進行預測。而預測完成後，將結果寫入共享記憶體 OUTPUT_RING，再經過 TCPD 程式進行 Twin 機制篩選與發報。

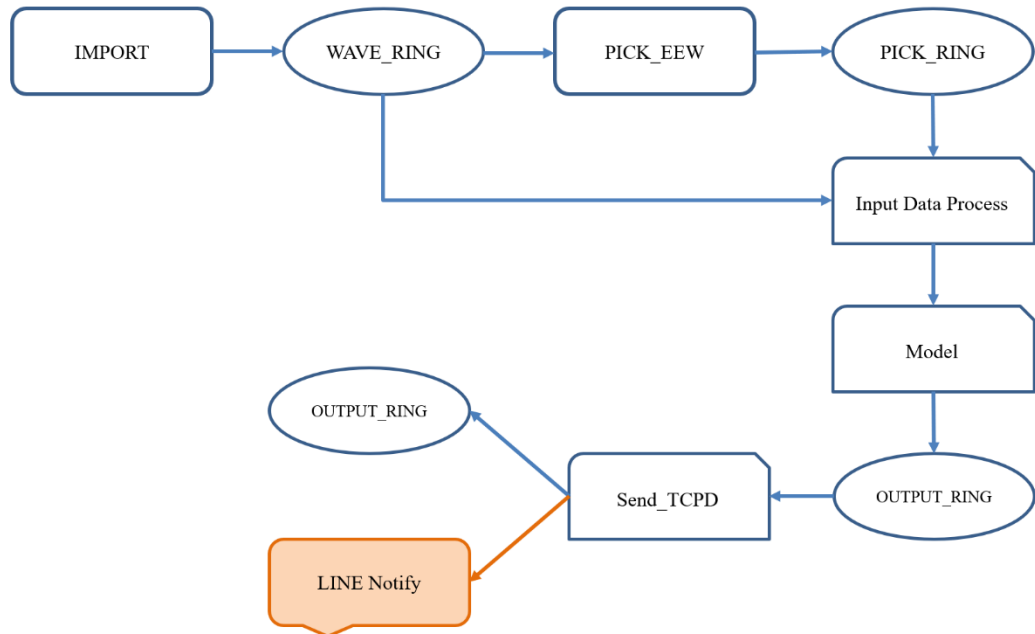


圖 2.4-1 氣象署 Earthworm 系統流程

為了建立與修改預測程式，在研究室建立 Earthworm 系統，並使用 Earthworm 系統的 TANKPLAYER 模組回放波形檔來模擬地震事件，流程如圖 2.4-2 所示。

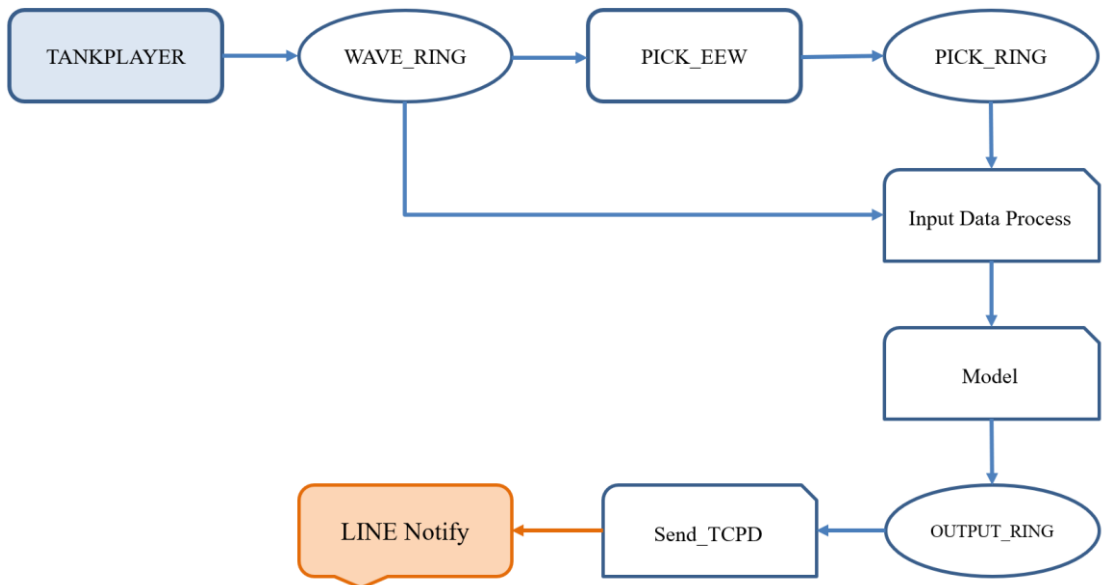


圖 2.4-2 研究室運行之 Earthworm 系統流程

2.5 建立與測試運行之 Python 程式

為了於研究室建立與測試程式運行狀況，在此以發生於 2022 年 09 月 17 日，

規模 6.6 之關山地震做測試。此次事件採用臺灣地震與地球物理資料管理系統 (GDMS) 下載之 TSMIP 測站波型資料製作回放檔。經過統計後，此次事件，共有 402 測站有波型資料，震央與測站分布如圖 2.5-1 所示，震度大於 4 級縣市如圖 2.5-2 所示。

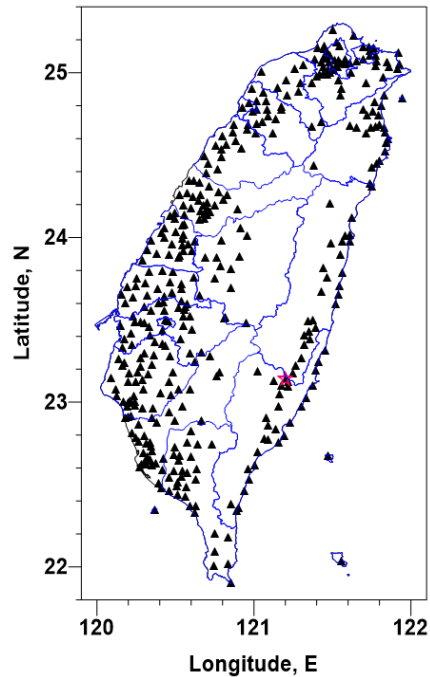


圖 2.5-1 關山事件震央與測站分布

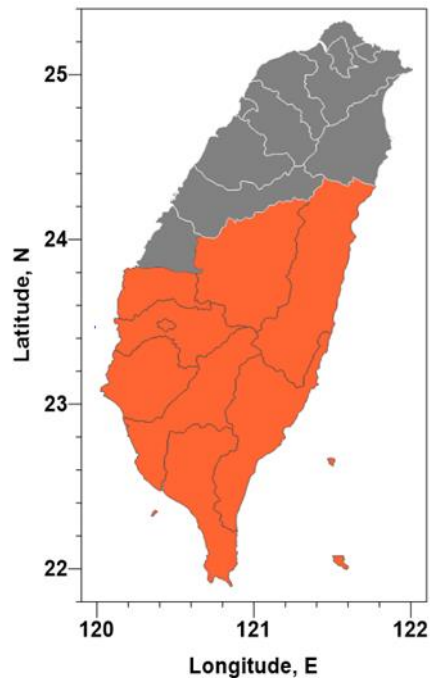


圖 2.5-2 關山事件震度大於 4 級縣市

2.5.1 初期版本-程式離線回放成果

此次模擬中，經由 pick_eww 觸發之測站總數為 161 個測站，而模型預測出來

PGA 大於 25 gal 之測站總數為 87 個測站，預測 PGA 小於 25 gal 的測站共有 74 個測站，分布如圖 2.5-3。

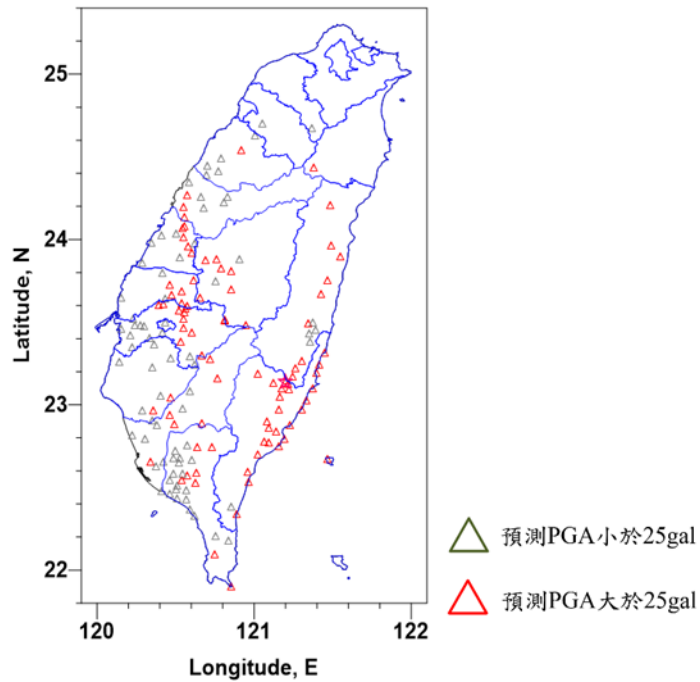


圖 2.5-3 預測 PGA 測站分布

以單站法統計結果如下，觸發的 161 個測站中，實際 PGA 大於 25 gal 的測站共有 92 個，實際 PGA 小於 25 gal 的測站共有 69 個。單站之二元混淆矩陣之判斷指標分布如圖 2.5-4，計算出準確率為 59.3%，召回率為 55.4%。

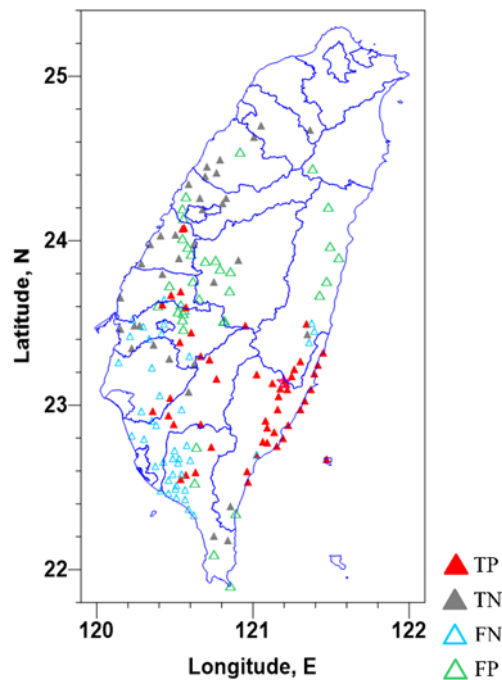


圖 2.5-4 二元混淆矩陣分布

單一測站預測 PGA 與實際 PGA 比較對數圖如圖 2.5-5 所示，縱軸為預測 PGA，橫軸為實際 PGA。從圖中可以明顯觀察到，預測結果偏差較大。

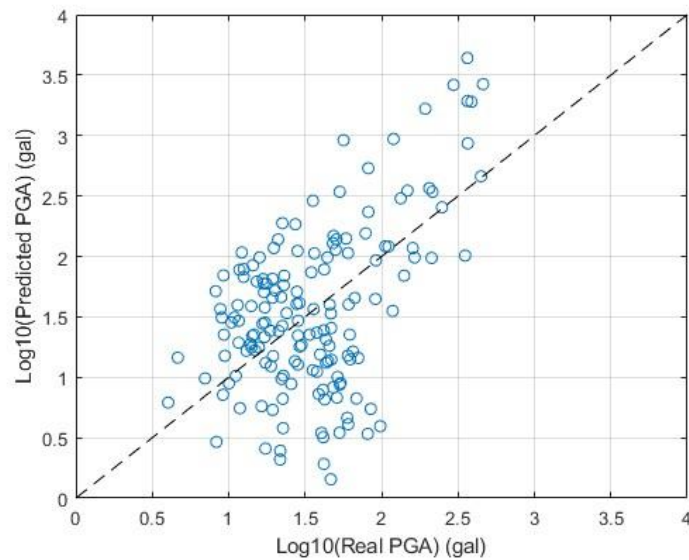


圖 2.5-5 預測 PGA 與實際 PGA 比較對數圖

為了評估程式的運算時間，分析了從接收到 P 波開始到預測完成的時間。從第 1 筆到第 161 筆的每一筆預測所需的時間進行了統計，並將這些數據繪製成趨勢圖，如圖 2.5-6。結果顯示，平均接收到 P 波開始到預測完成的時間為 9.62 秒。值得注意的是，事件剛開始預測的耗時相對較長。這一現象可能源於地震初始階段時多個測站同時被觸發，造成系統資源分配壓力增大，導致程式運行阻塞。這樣的資源分配壓力可能導致多筆預測資料處於等待狀態，從而影響了預測執行的效率。

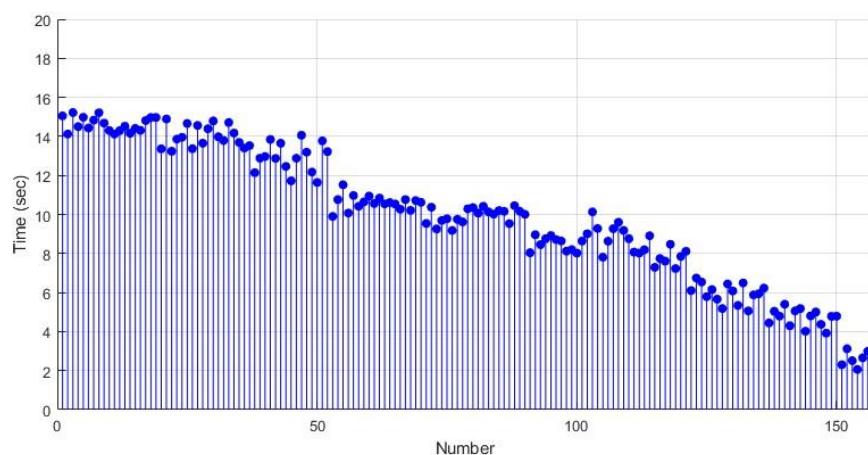


圖 2.5-6 程式預測耗時趨勢圖

2.5.2 程式問題與解決

在前述章節的分析中，單站之二元混淆矩陣之判斷指標與預測 PGA 與實際 PGA 相差比預期還大，檢討程式後發現訓練模型時資料是採用 Matlab 進行資料處理而後再放入 Python 做預測；而線上則是全由 Python 做處理，而當中的資料處理順序與使用之函數有些微出入，造成預測結果與離線分析有所差異。另外亦發現，初期程式在資料處理中，會在同一筆預測中，調用不同時間上之波型資料，導致預測不準確。此外，許多誤報是因為波型資料中，偶有存在 Inf、Nan 或異常大值得情形，導致資料處理時，如零均值（Zero-mean）等，出現偏差，導致預測失準。

而由統計程式的運算時間可以發現，初期程式存在預測耗時太久的問題，原因為初期程式並未導入多線程(多執行緒)概念，使得預測僅能一筆一筆順序預測，導致程式可能出現堵塞問題。

因此，本研究改進了以上發現的問題，說明如下。

針對預測失準解決方向：

- 1、 因應訓練與實際上線程式差異，需重新調整運算的架構與函數調用，確保計算的順序與方法與訓練模型時保持一致。
- 2、 為避免不同時間段的波型資料在同一筆預測中調用，須確保在即時系統中，資料處理中擷取的波形數據的時間一致性。
- 3、 建立一套機制，以解決波型資料中遇到雜訊，使得出現 Inf、Nan 或異常大值時，造成在進行零均值（Zero-mean）等數據預處理時的誤差，從而提高預測的準確度。

針對預測過慢解決方向：

- 1、 為避免程式堵塞，嘗試引入執行緒池（Thread pool）概念，使每筆預測都能在獨立的執行緒中進行，這樣即使在大量資料同時到達時也不會造成系統堵塞。同時，設定執行緒數量的最大上限，以避免系統過載。

2.5.3 現今版本程式離線回放成果

此次模擬中，經由 pick_ew 觸發之測站總數為 161 個測站，而模型預測結果，PGA 大於 25 gal 之測站總數為 94 個測站，PGA 小於 25 gal 之測站總數為 67 測站，分布如圖 2.5-7。

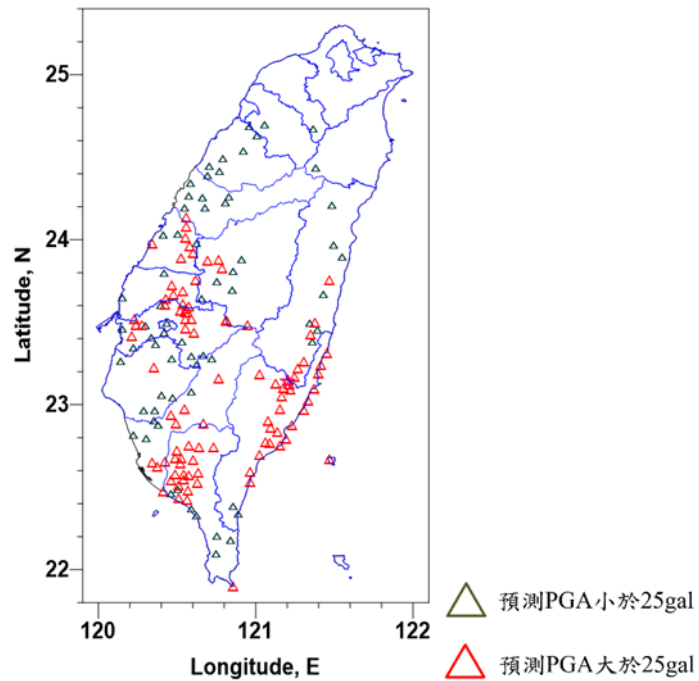


圖 2.5-7 預測 PGA 測站分布

以單站法統計結果如下，觸發的 161 個測站中，實際 PGA 大於 25 gal 的測站共有 92 個，實際 PGA 小於 25 gal 的測站共有 69 個。單站之二元混淆矩陣分布，在無容許誤差情況下，計算出準確率為 80.2%，召回率為 77.8%，如圖 2.5-8；若是容許誤差情況下，準確率為 100%，召回率為 97.9%，如圖 2.5-9。

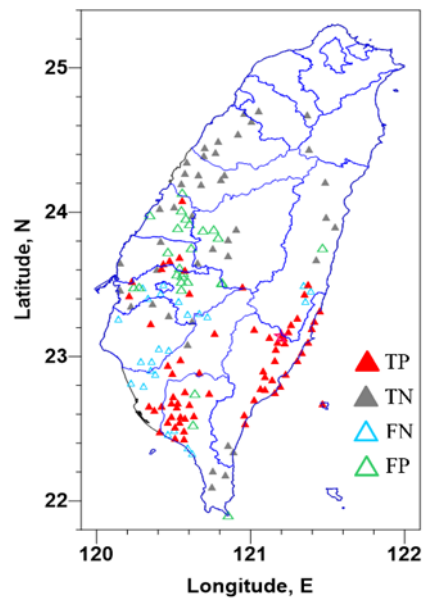


圖 2.5-8 二元混淆矩陣分布(無容許誤差)

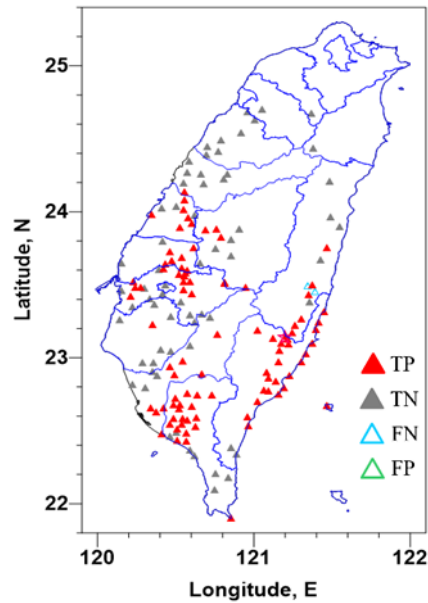


圖 2.5-9 二元混淆矩陣分布(含容許誤差)

單一測站預測 PGA 與實際 PGA 比較對數圖，如圖 2.5-10 所示，縱軸為預測 PGA 取對數，橫軸為實際 PGA 取對數，從圖中可以發現，相較於先前版本，當前的預測模型在 PGA 值的預測上有顯著的改善。

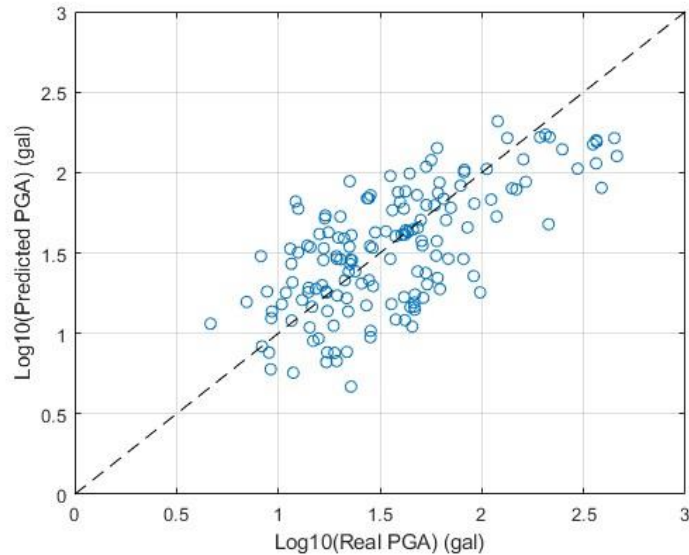


圖 2.5-10 預測 PGA 與實際 PGA 比較對數圖

下圖 2.5-11 為此事件每筆預測從接收到 p 波資訊到預測完成所花費的時間圖，此地震事件的預測平均時間為 1.59 秒(圖 2.5-11 的 Mean)。圖上的 Data Process 為耗費於資料處理的時間，Predict 為啟動模型預測到輸出結果的時間，能發現預測階段的平均耗時為 0.64 秒(圖 2.5-11 上的 Mean (Predict stage))，且沒有明顯因事件剛開始，大量波型資料進入造成預測堵塞問題。

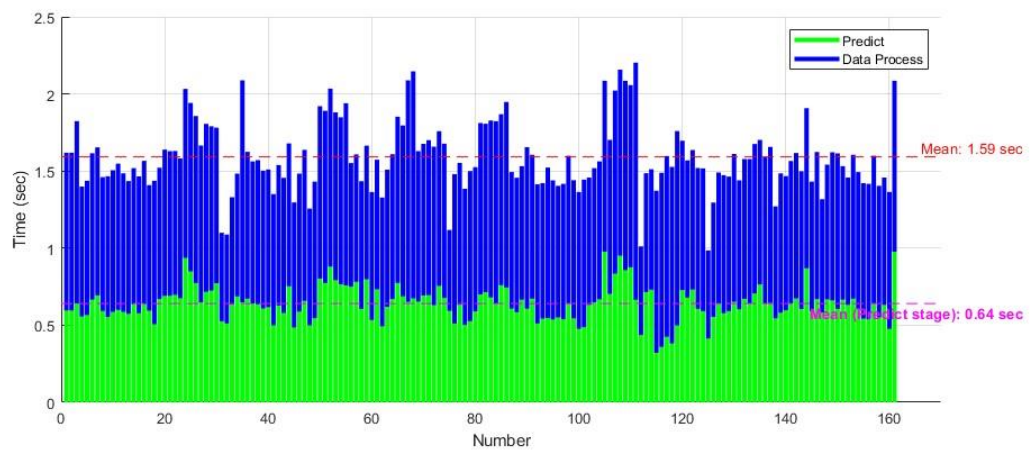


圖 2.5-11 程式預測耗時趨勢圖

採用 Twin 機制對於 17 個縣市發報結果，黃色縣市為正確發報，紅色縣市為誤報，粉色縣市為漏報。未含容許誤差下，準確率為 85.7%，召回率為 75%，如圖 2.5-12；含容許誤差下，準確率為 100%，召回率為 87.5%，如圖 2.5-13。

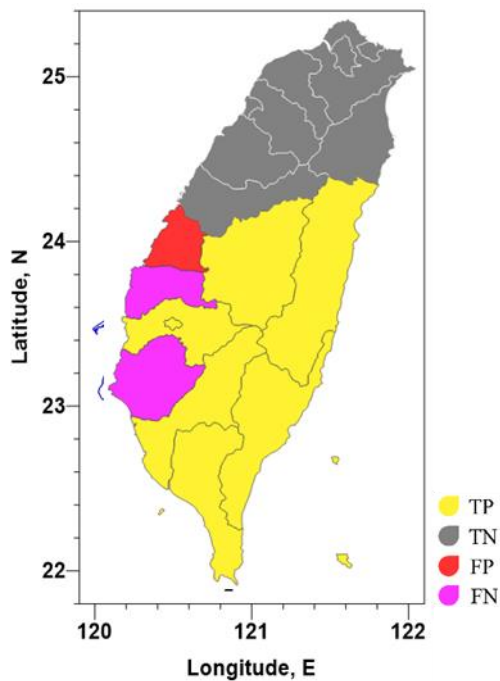


圖 2.5-12 Twin 機制對於縣市發報結果(無容許誤差)

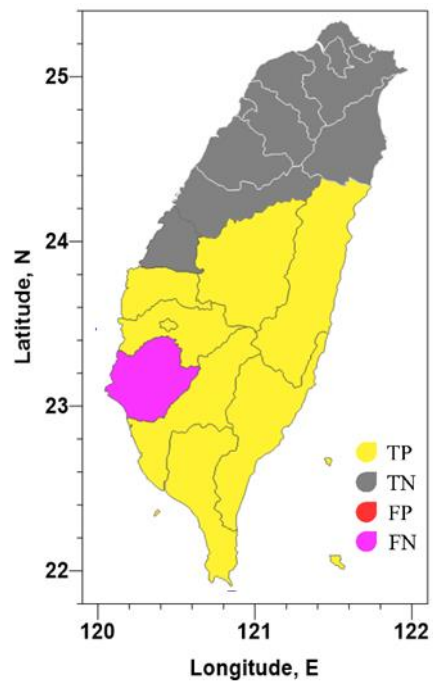


圖 2.5-13 Twin 機制對於縣市發報結果(含容許誤差)

其中，臺南市不論有無考慮容許誤差的情況下，皆為漏報，其中一個原因可能為研究室 Earthworm 的 pick_eww 模組參數調整問題，導致臺南市觸發的測站過少，導致難以達到 Twin 機制要求，臺南市觸發測站，如圖 2.5-14 所示。

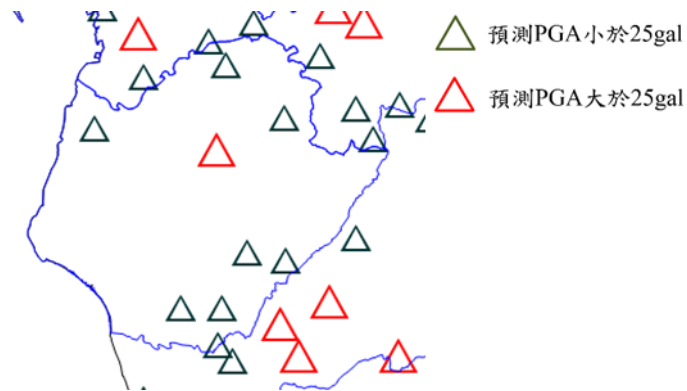


圖 2.5-14 臺南市觸發測站

Twin 正確發報之縣市發報時間軸如圖 2.5-15，以氣象署公布的關山地震實際發震時間作為事件開始的基準點，並將其轉換為實驗室數據回放的相對時間以便對照。Picking 為該縣市記錄到的第一個 P phase 時間，Model Alert 為此縣市符合 Twin 機制發報的時間，此外，地面加速度（PGA）的到達時間以及當加速度超過 25 gal 的到達時間是根據各測站的回放波型資料計算所得。

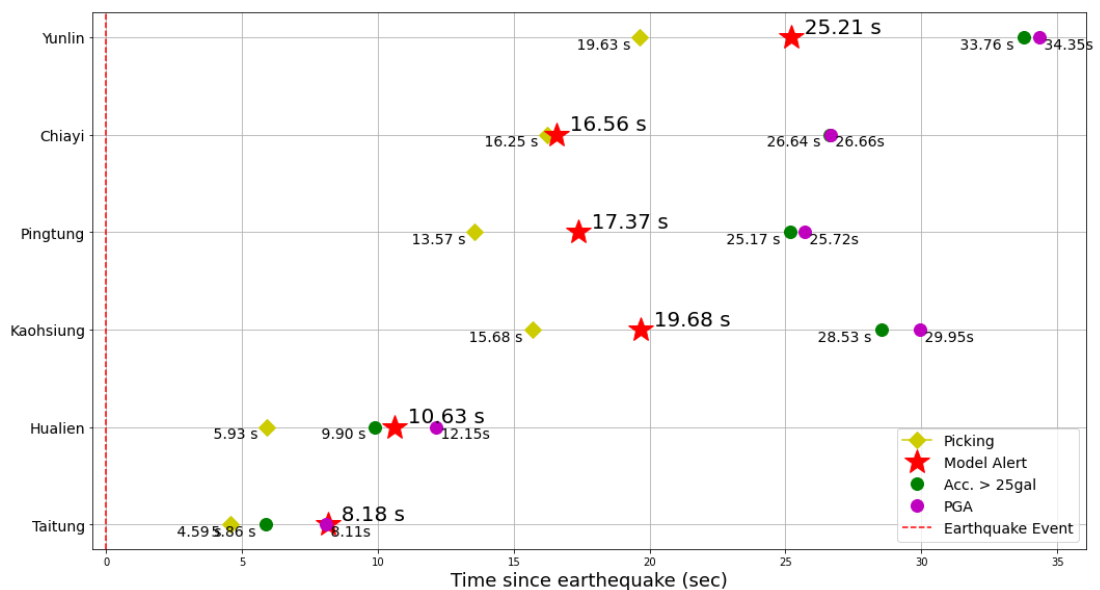


圖 2.5-15 Twin 正確發報之縣市發報時間軸

星型標示時間減去菱形標示時間，即可算出從得到 p-phase 資訊後，多久對該縣市發報。在正確發報的情況下，除了臺東震央以外，皆可以在 PGA 到時前發出預警，且 Twin 機制在 Picker 給予 P 波資訊後，最快 0.31 秒，最慢 5.58 秒符合 Twin 機制發報。

本節的成果表明，先前的版本的問題，皆已經獲得了改善。

2.6 於氣象署運行情況

更新後的程式近期才在氣象署上線，在此提出上線後遇到的事件作為成果討論。此事件發生於 2023 年 11 月 11 日 00:54:39 規模 5.4，震央位於花蓮縣政府南方 64.9 公里 (臺灣東部海域)。各地震度如圖 2.6-1 所示，震度大於四級之縣市分為別為臺東縣與花蓮縣。

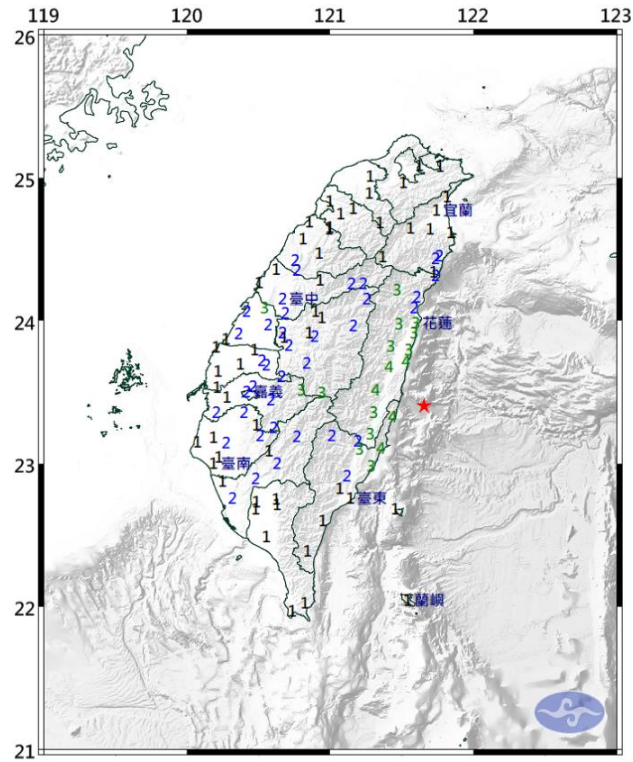


圖 2.6-1 事件各地震度

觸發之測站總數為 80 個測站，而模型預測結果，PGA 大於 25 gal 之測站總數為 17 個測站，PGA 小於 25 gal 之測站總數為 63 測站，分布如圖 2.6-2 所示。

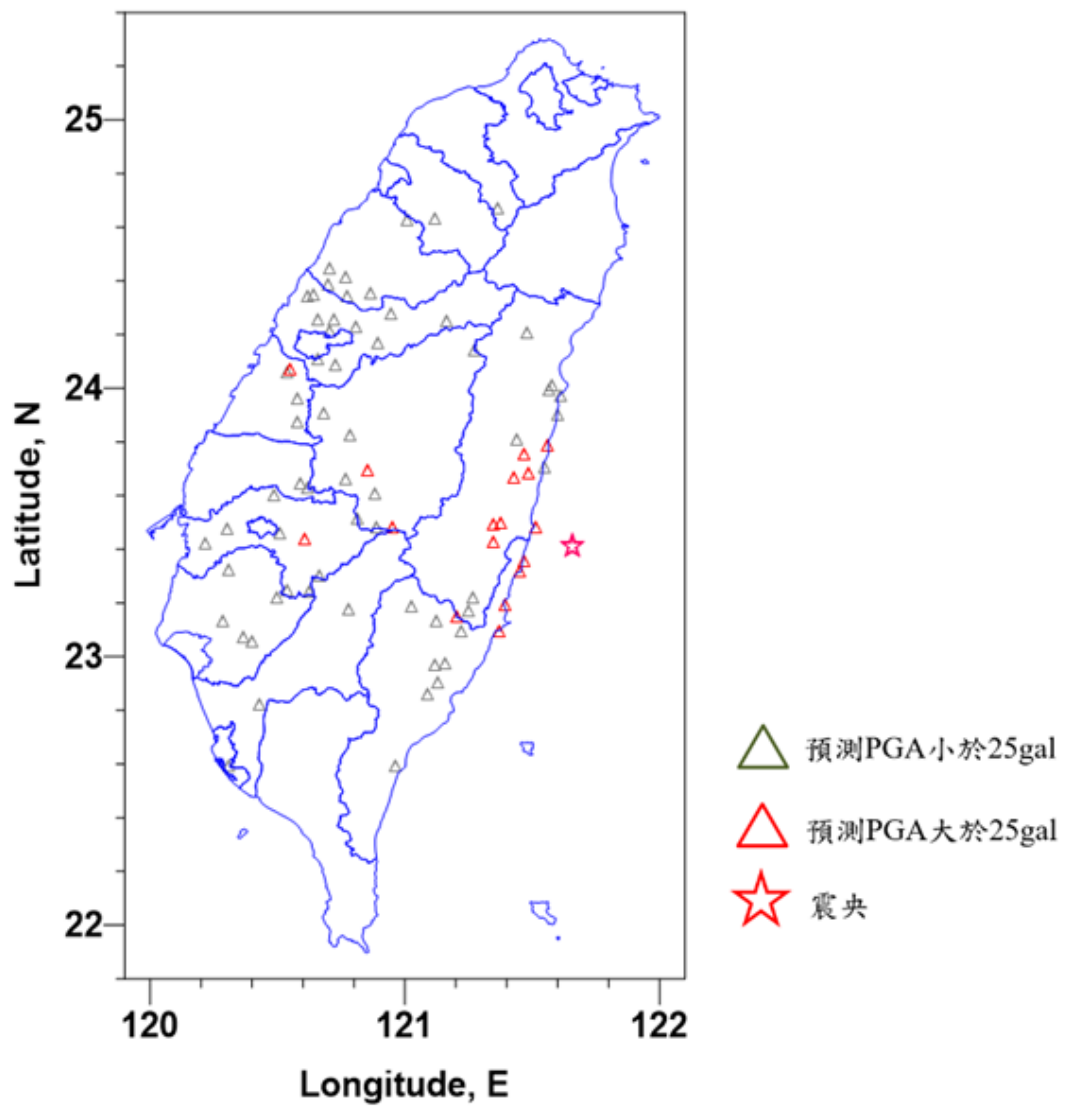


圖 2.6-2 觸發測站

以單站法統計結果如下，觸發的 80 個測站中，實際 PGA 大於 25 gal 的測站共有 12 個，實際 PGA 小於 25 gal 的測站共有 68 個。單站之二元混淆矩陣分布，在無容許誤差情況下，計算出準確率為 75%，召回率為 100%，如圖 2.6-3；若是容許誤差情況下，準確率為 93.8%，召回率為 100%，如圖 2.6-4。

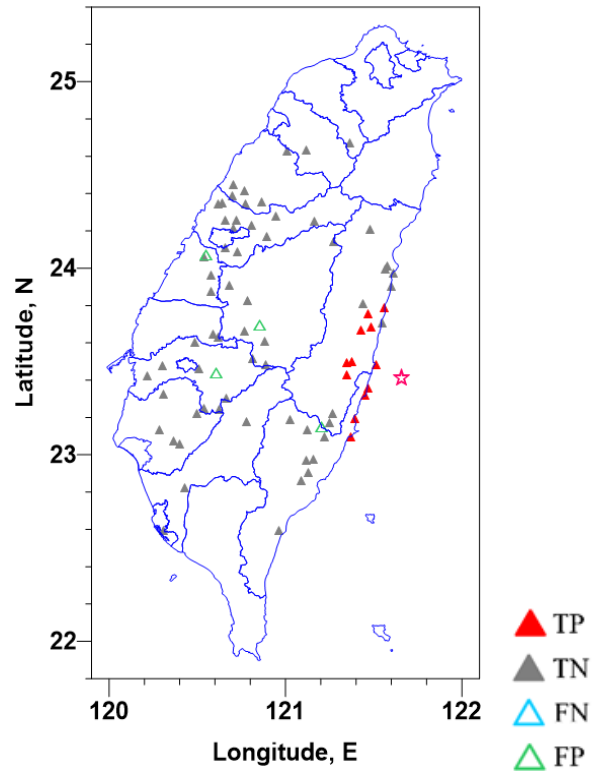


圖 2.6-3 二元混淆矩陣(無容許誤差)

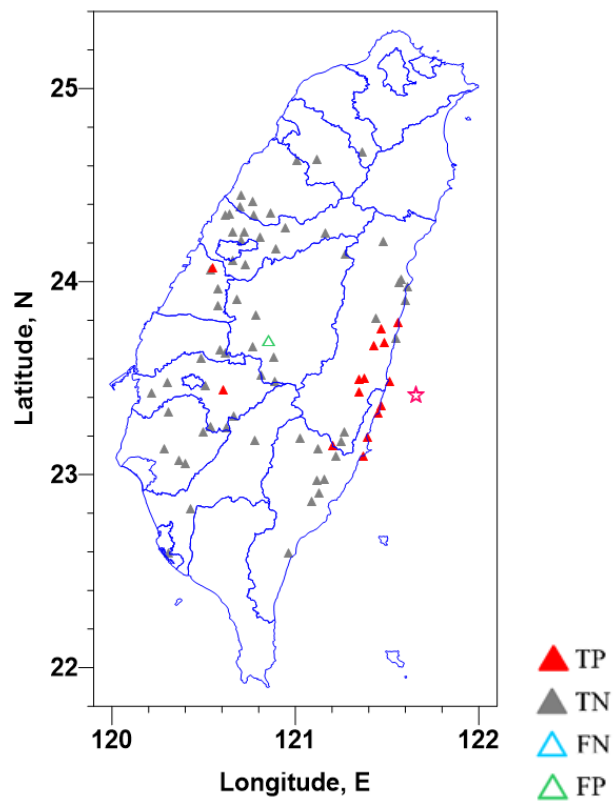


圖 2.6-4 二元混淆矩陣(含容許誤差)

單一測站預測 PGA 與實際 PGA 比較對數圖，如圖 2.6-5 所示，縱軸為預測 PGA 取對數，橫軸為實際 PGA 取對數。

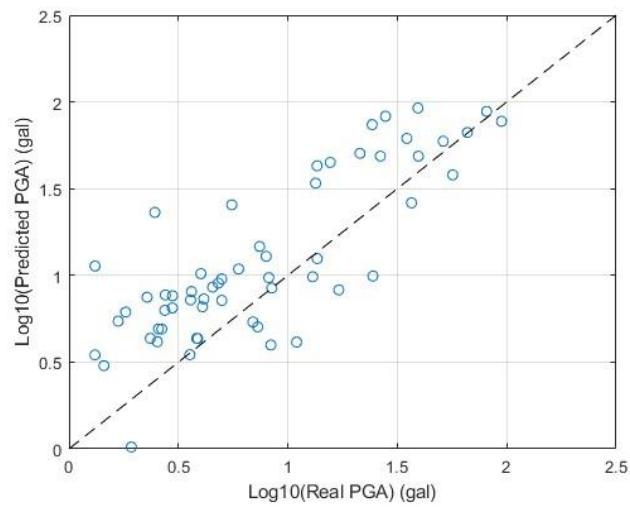


圖 2.6-5 單一測站預測 PGA 與實際 PGA

採用 Twin 機制對於 17 個縣市發報結果，黃色縣市為正確發報，紅色縣市為誤報，粉色縣市為漏報。未含容許誤差下，準確率為 100%，召回率為 100%，如圖 2.6-6。

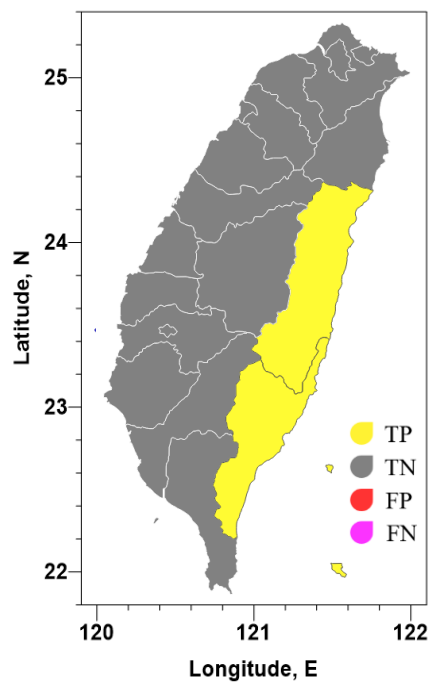


圖 2.6-6 對於縣市發報結果

Twin 正確發報之縣市發報時間軸如圖 2.6-7，以氣象署公布的地震實際發震時間作為事件開始的基準點。P_{arrival} 為該縣市記錄到的第一個 P phase 時間，

Get_Pick_Message 為程式收到 P phase 資訊的時間，Model_Alert 為此縣市符合 Twin 機制發報的時間，地面加速度（PGA）的到達時間以及當加速度超過 25 gal 的到達時間是根據各測站的波型資料計算所得。

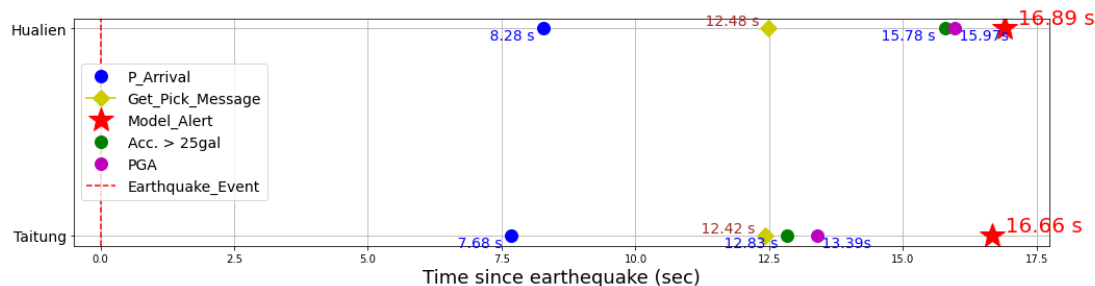


圖 2.6-7 發報時間軸

由圖 2.6-7 可以發現，P 波到時與程式得到 P 波到時資訊的時間差為 4.74 及 4.20 秒，發報時間分別為 P 波到時後 8.98 秒及 8.61 秒，但發報時兩地皆已超過 PGA 到時。不論是接收到 P 波資訊的時間與預測時間皆須再提升速度。舉例來說，若是在 P 波到時後 3 秒內得到 P 波到時資訊，預測時間加快至兩秒以內，則兩地皆有可能在當地加速度超過 25 gal 前，對地方發出預警。

下圖 2.6-8 為此事件每筆預測從接收到 p 波資訊到預測完成所花費的時間圖，此地震事件的預測平均時間為 4.56 秒(圖 2.6-8 的 Mean)。圖上的 Data Process 為耗費於資料處理的時間，Predict 為啟動模型預測到輸出結果的時間，能發現預測階段的平均耗時為 1.90 秒(圖 2.6-8 上的 Mean (Predict stage))。這與實驗室中統計的運算時間相比，顯示出一定的差距。此差異顯示了進一步調整程式效能的必要性，以縮短預測階段的處理時間，從而達到更接近實驗室條件下的理想運算表現。

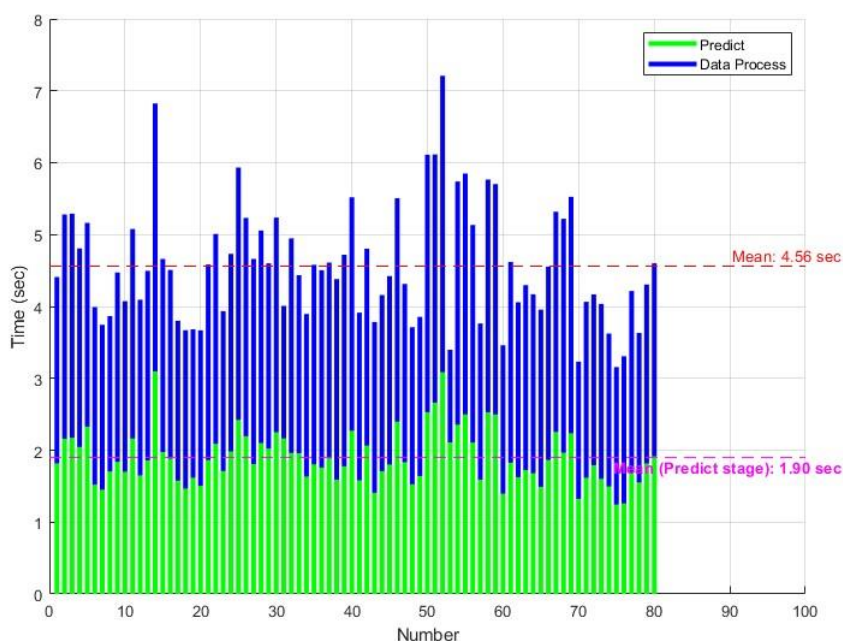


圖 2.6-8 預測耗時

2.7 小結

本研究旨在將現地型警報加入到 Earthworm 系統中，以獲得更快速和精確的強震預警。本研究假設 TSMIP 的所有 839 個測站都可以接收即時資料，以瞭解未來實際應用的可能情況。

在實驗室環境測試下，發現了諸多問題，使得模型預測失準以及預測耗時太長，而大多數問題皆在期末版本程式解決。

根據在氣象署運行的模型，由事件成果顯示預測結果成效不錯，在 Twin 機制下，準確率為 100%，召回率為 100%。但其運算時間性能尚未達到預期。在預測耗時方面，相較於實驗室環境，花費較多時間，推測可能由於硬體環境的差異所導致。因此，需進行深入討論以釐清如何提升運算效率，包括是否利用 GPU 加速運算的可行性。

此外，實驗室環境中的地震波形 Picking 是透過 pick_eww 模組進行，未來計畫將實驗室現有的 Picking 模組更新為 AI picker，並制定相應的驗證流程，以測試 AI Picker 與本模型的整合潛力。

3 單站法即時地震波到時標記及強震預警模型開發

3.1 模型架構介紹

多任務學習(Multi-Task Learning, MTL)係在一個網絡中執行多個任務且任務間可共享訊號之特徵(Ruder [4],Crawshaw [5])。多任務學習被廣泛應用於電腦視覺和自然語言處理等領域。而多任務學習在不同任務間共享訊號之特徵會面臨兩個挑戰，(一)、如何共享訊號之特徵，(二)、如何平衡不同任務間之損失(Loss)[6]。本研究參考 Liao, et al. [1]，由遞迴殘差 U 型網(Recurrent Residual U-Net)作為共享訊號特徵之網路，並搭配三個用於特定任務的注意力子網路(Attention sub-net)，此架構請參考圖 3.1-1、圖 3.1-2。

圖 3.1-1 之計算流程如下，輸入進模型之矩陣先進入一維卷積層且步數(Stride)為一，再進入 RRC operation，輸出 RRC operation 後之矩陣 R^0 再分別經過三個 Attention gate，輸出矩陣為 $\hat{a}_I^0$ 、 $\hat{a}_{II}^0$ 、 $\hat{a}_{III}^0$ ，I、II、III 分別為三個 Attention sub-net，I 為地震波相標記(Earthquake Phase Picking)、II 為判斷訊號是否為地震訊號(Earthquake Waveform Detection)、III 為預測 PGA 是否大於 25gal(Predict whether PGA is greater than 25 gal)。 R^0 再經過一維卷積層且步數為一和 RRC operation 為 R^1 ， R^1 再跟 R^0 、各別之 $\hat{a}^0$ 分別經過 Attention gate 得到 $\hat{a}_I^1$ 、 $\hat{a}_{II}^1$ 、 $\hat{a}_{III}^1$ ，各別之 $\hat{a}^1$ 再各別經過一維卷積層且步數為五為各別之 $\hat{a}_{down}^1$ 。 R^1 經過一維卷積層且步數為五後為 R_{down}^1 ， R_{down}^1 經過一維卷積層且步數為一和 RRC operation 為 R^2 ， R^2 再跟 R_{down}^1 、各別之 $\hat{a}_{down}^1$ 分別經過 Attention gate 得到 $\hat{a}_I^2$ 、 $\hat{a}_{II}^2$ 、 $\hat{a}_{III}^2$ ，各別之 $\hat{a}^2$ 再各別經過一維卷積層且步數為五為各別之 $\hat{a}_{down}^2$ 。 R^2 經過一維卷積層且步數為五後為 R_{down}^2 ， R_{down}^2 經過一維卷積層且步數為一和 RRC operation 為 R^3 ， R^3 再跟 R_{down}^2 、各別之 $\hat{a}_{down}^2$ 分別經過 Attention gate 得到 $\hat{a}_I^3$ 、 $\hat{a}_{II}^3$ 、 $\hat{a}_{III}^3$ ，各別之 $\hat{a}^3$ 再各別經過一維卷積層且步數為五為各別之 $\hat{a}_{down}^3$ 。 R^3 經過一維卷積層且步數為五後為 R_{down}^3 ， R_{down}^3 經過一維卷積

層且步數為一和 RRC operation 為 R^4 ， R^4 再跟 R_{down}^3 、各別之 $\hat{a}_{down}^3$ 分別經過 Attention gate 得到 $\hat{a}_I^4$ 、 $\hat{a}_{II}^4$ 、 $\hat{a}_{III}^4$ ，各別之 $\hat{a}^4$ 再各別經過一維上取樣(Upsampling)讓矩陣沿著長邊重複五次為各別之 $\hat{a}_{up}^4$ 。 R^4 經過一維上取樣(Upsampling)讓矩陣沿著長邊重複五次為 R_{up}^4 ， R_{up}^4 和 R^3 短邊接在一起經過一維卷積層且步數為一和 RRC operation 為 R^5 ， R^5 再跟 R_{up}^4 、各別之 $\hat{a}_{up}^4$ 分別經過 Attention gate 得到 $\hat{a}_I^5$ 、 $\hat{a}_{II}^5$ 、 $\hat{a}_{III}^5$ ，各別之 $\hat{a}^5$ 再各別經過一維上取樣(Upsampling)讓矩陣沿著長邊重複五次為各別之 $\hat{a}_{up}^5$ 。 R^5 經過一維上取樣(Upsampling)讓矩陣沿著長邊重複五次為 R_{up}^5 ， R_{up}^5 和 R^2 短邊接在一起經過一維卷積層且步數為一和 RRC operation 為 R^6 ， R^6 再跟 R_{up}^5 、各別之 $\hat{a}_{up}^5$ 經過 Attention gate 得到 $\hat{a}_I^6$ 、 $\hat{a}_{II}^6$ 、 $\hat{a}_{III}^6$ ，各別之 $\hat{a}^6$ 再各別經過一維上取樣(Upsampling)讓矩陣沿著長邊重複五次為各別之 $\hat{a}_{up}^6$ 。 R^6 經過一維上取樣(Upsampling)讓矩陣沿著長邊重複五次為 R_{up}^6 ， R_{up}^6 和 R^1 短邊接在一起經過一維卷積層且步數為一和 RRC operation 為 R^7 ， R^7 再跟 R_{up}^6 、各別之 $\hat{a}_{up}^6$ 經過 Attention gate 得到 $\hat{a}_I^7$ 、 $\hat{a}_{II}^7$ 、 $\hat{a}_{III}^7$ ，各別之 $\hat{a}^7$ 分別經過一維卷積層和 Softmax 即為三種目標之輸出。

三種目標輸出之機率隨時間之函數 $Prob(t)$ 為(一)、標注地震相位(Earthquake Phase Picking)，其中包含 P 波到時、S 波到時、其他(Others);(二)、地震波形檢測(Earthquake Waveform Detection)，其中包含遮罩(Mask)、反遮罩(Unmask);(三)、預測 PGA 是否大於 25gal(Predict whether PGA is greater than 25 gal)，其中包含 PGA 是否超過 25gal、反 PGA 是否超過 25gal，此目標為 Liao, et al. [1]額外增加之項目。Liao, et al. [1]將三軸向波形資料之輸入做的標準化為 Z-score，而本研究之輸入除了將三軸向加速度歷時做 Z-score 外，還有對其做不同加速度區間之正規化。

圖 3.1-2 中有三個注意力子網路(Attention sub-net)，Attention sub-netI 的功能為標記地震波相(Earthquake Phase Picking)、Attention sub-netII 的功能為判斷訊號是否地震訊號(Earthquake Waveform Detection)、Attention sub-netIII 的功能為預測 PGA 是否大於 25gal(Predict whether PGA is greater than 25 gal)，此 Attention sub-net 為本研究在 Liao, et al. [1]之模型架構中額外增加之注意力子網路(Attention sub-net)。

圖 3.1-3 中的上半部分為 Recurrent Residual Convolution operation(RRC operation)，輸入矩陣會分別經過兩個一維卷積層且步數(Stride)為一，上面流程名稱為 Recurrent，下面流程名稱為 Residual。先說明 Recurrent，經過一維卷積層之矩陣和輸入矩陣相加，相加後之矩陣再跟輸入矩陣相加再經過一維卷積層後之矩陣，再跟輸入矩陣相加，重複此動作三次輸出之矩陣即為經過 Recurrent 後之矩陣，再跟經過 Residual 後之矩陣相加，以上動作即為 Recurrent Residual Convolution operation 內之計算流程。而 Attention gate 為圖 3.1-3 下半部分，上層經過 RRC operation 後之矩陣 R' 和上層經過 Attention gate 後之矩陣 $\hat{a}'$ ，兩個連接在一起經過 ReLU、Sigmoid 後之矩陣跟此層經過 RRC operation 後之矩陣 R 相乘，此處的相乘為矩陣中之元素逐個相乘，相乘後之矩陣為經過 Attention gate 後之矩陣 $\hat{a}$ 。模型詳細之參數設定如表 3.1-1。

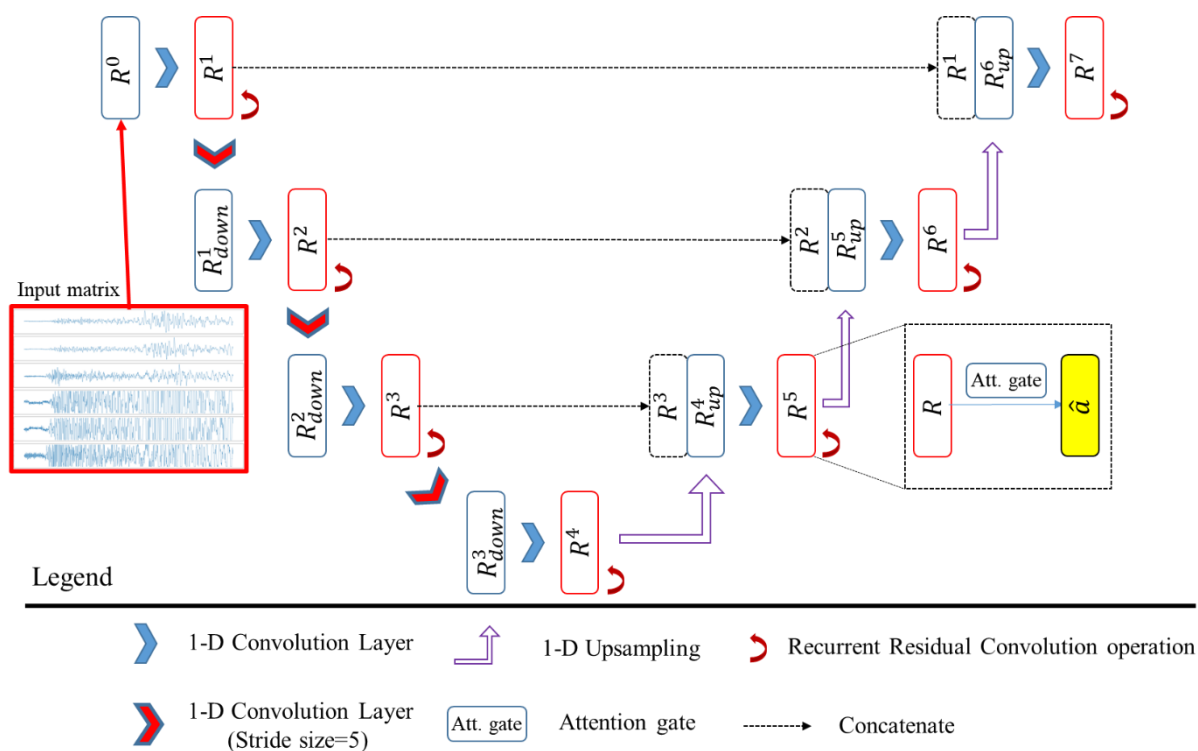


圖 3.1-1 Recurrent-residual U-Net 架構

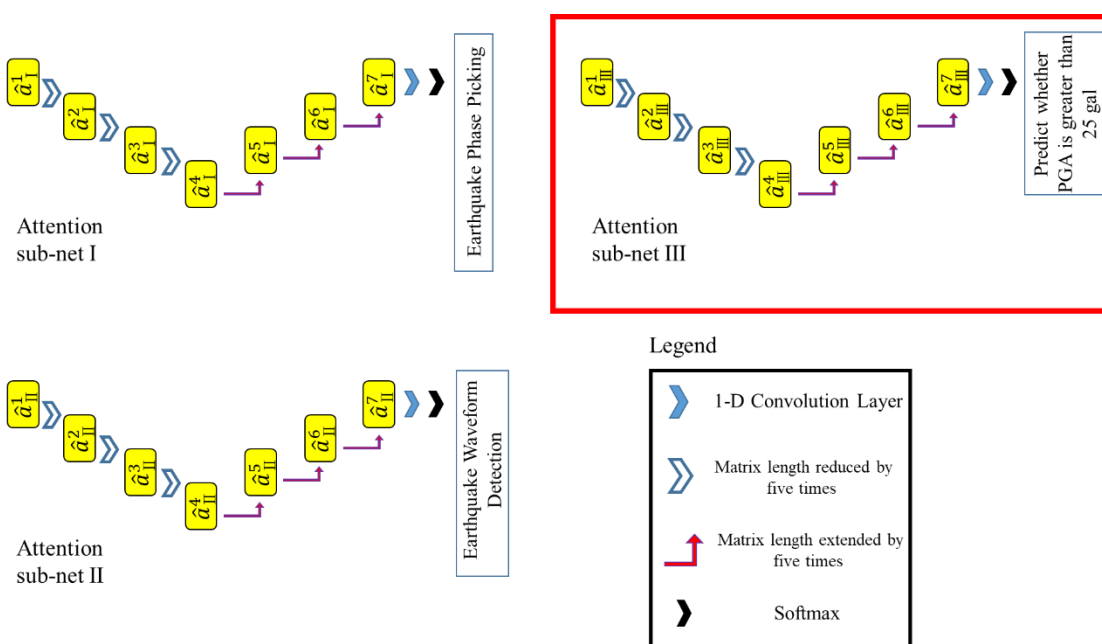


圖 3.1-2 特殊任務注意力網路(Task-Specific Attention Network)

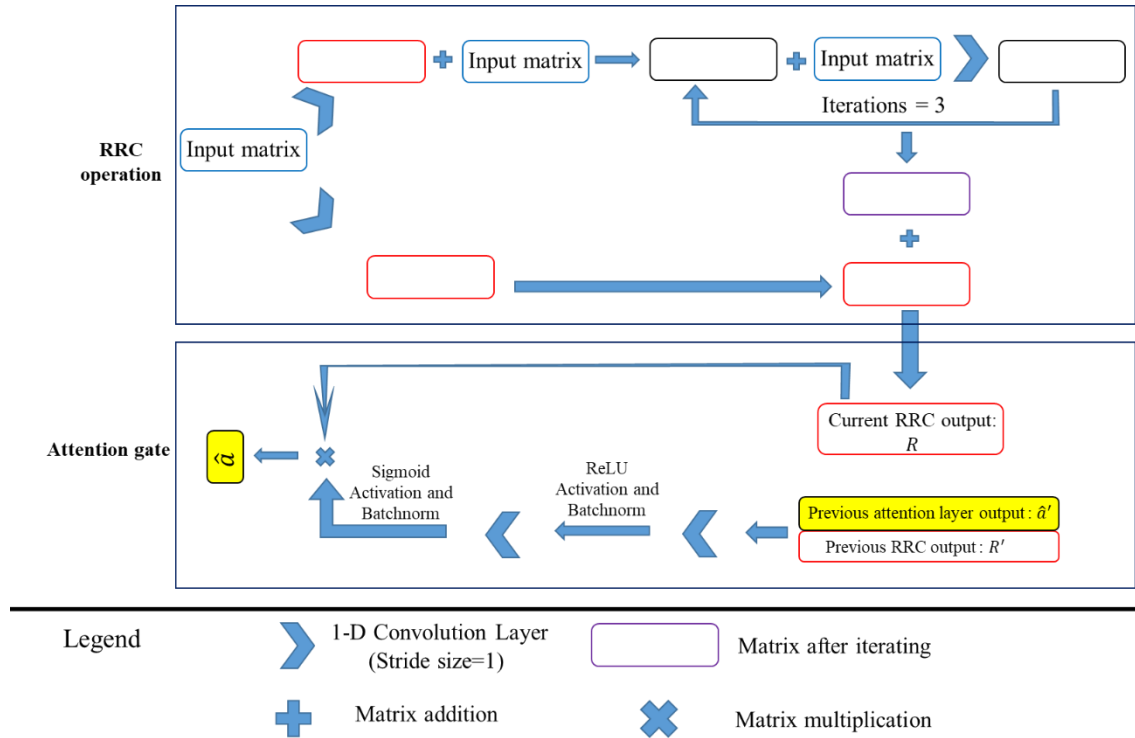


圖 3.1-3 RRC Operation 和 Attention Gate 詳細計算流程

本研究之損失函數為 Categorical Cross Entropy，為交叉熵(Cross Entropy)之其中一種，計算三種任務之損失函數 H 是根據 Softmax 正規化之預測函數 q_c 和相應的真實目標函數 p_c 之間的交叉熵(Cross-Entropy)，其定義如下：

$$H_{\text{EQ Phase Picking}}(p, q) = -\sum_c \sum_x p_c(x) \log q_c(x) \quad (3.1-1)$$

$$H_{\text{EQ Detection}}(p, q) = -\sum_c \sum_x p_c(x) \log q_c(x) \quad (3.1-2)$$

$$H_{\text{EQ Prediction}}(p, q) = -\sum_c \sum_x p_c(x) \log q_c(x) \quad (3.1-3)$$

在(3.1-1)、(3.1-2)、(3.1-3)中，為在 Liao, et al. [1]原本損失函數額外增加之損失函數， x 代表每個資料點， c 代表該目標中每個任務之數量。

表 3.1-1 模型使用參數

Optimizer	Rectified Adam[7]
Output Activation Function	Softmax
Loss Function	Categorical Cross Entropy
Batch Size	40

3.1.1 平衡不同任務間之損失

在多任務學習中，如何讓不同任務間預測之函數 $Prob(t)$ 和真實之函數 $Prob(t)$ 計算之損失值達到平衡是個難題。而本報告參考 Liao, et al. [1]、Liu [6] 等人之研究成果，採用動態權重平均(Dynamic Weight Average, DWA)權衡不同任務間之損失值。 L_k 是特定任務之損失值， α_k 是權衡特定任務損失值的權重，定義如下：

$$\alpha_k(I) = \frac{K * \exp(w_k(I-1)/T)}{\sum_i \exp(w_i(I-1)/T)} \quad (3.1-4)$$

$$w_k(I-1) = \frac{L_k(I-1)}{L_k(I-2)} \quad (3.1-5)$$

I 代表每個訓練週期(Epoch)，本研究設定第一和第二個訓練週期之 w_k 為 1。 w_k 是相對下降率。 T 為溫度調節因子(Temperature Scaling Factor)，其愈大在不同任務間之權重越平均，而本研究設定此值為 7。 K 為任務總數，所以在本研究中， K 為 3。 α_k 為權衡不同任務損失值之權重。因此，本研究計算損失值時是考慮三種任務之預測之機率密度函數和真實之機率密度函數，根據以下公式計算在每個訓練週期之整體損失值：

$$L_{total} = \sum_k^3 \alpha_k(I) H_k \quad (3.1-6)$$

其中 $k=1,2,3$ 分別表示地震相位標記(Earthquake Phase Picking)、地震波形檢測(Earthquake Waveform Detection)、預測 PGA 是否大於 25gal(Predict whether PGA is greater than 25 gal)跟 Liao, et al. [1] 相比，此目標為額外增加之目標。

3.2 模型資料集介紹

本研究所使用的資料來自兩個臺灣觀測網，(一)、2012~2019 年中央氣象署地震觀測網(Central Weather Administration Seismographic Network, CWASN)，(二)、1992~2006 年臺灣強地動觀測網(Taiwan Strong Motion Instrument Program Network, TSMIP)，兩個觀測網之頻率統一為 100 赫茲且資料長度都為 30 秒之波形資料。由於本研究是以 Liao, et al. [1] 之模型(RED-PAN)為基礎並增加強震預警之功能，所以為了讓本研究之模型有如 RED-PAN 標記地震相位之能力，故本研究資料集有包含 Liao, et al. [1] 之資料集。本研究資料集有考慮噪音(Noise)的資料樣本，如圖 3.2-1，此圖之資料長度為 30 秒，資料來源為 STEAD(STanford Earthquake Dataset)資料集 [8]，訓練資料集有 3000 筆、驗證資料集有 750 筆、測試資料集來自臺灣資料集，是由 [9] 提供，共 23260 筆。

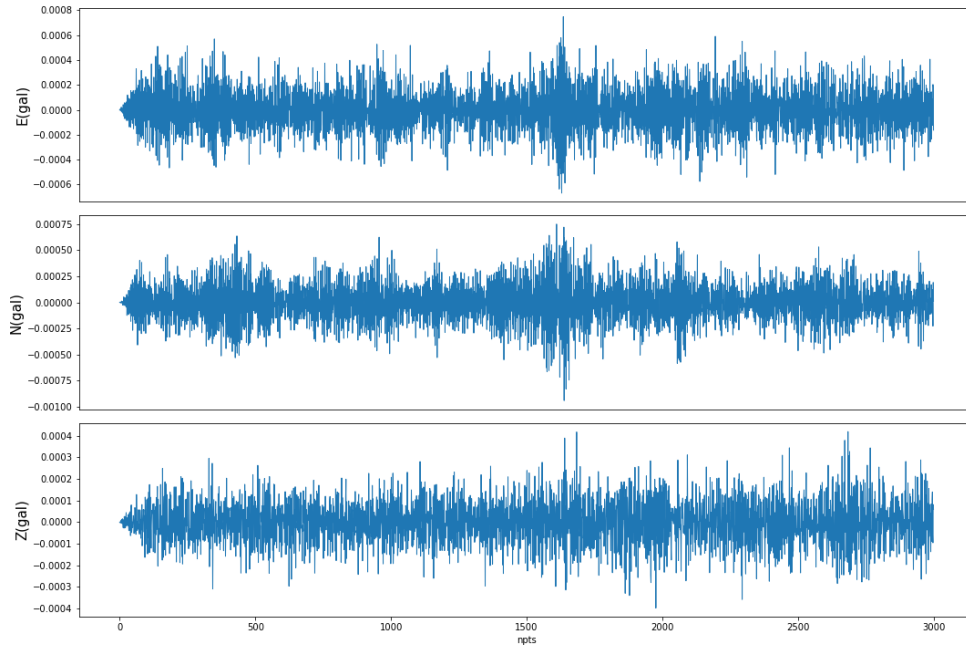


圖 3.2-1 STEAD 資料集中的 noise

3.2.1 訓練、驗證資料集

訓練和驗證資料集都有分兩種資料，(一)、P、S 波都含之單一事件，(二)、只含一小段 P 波訊號。P、S 波都含之單一事件之波形資料為從 P 波到時前 3 秒開始之 30 秒地震訊號，如圖 3.2-3，圖 3.2-2 為圖 3.2-3 切之前之原始訊號，此為 1994 年 1 月 20 日下午 1 點 50 分，觸發測站為 TSMIP 中的 HWA016。

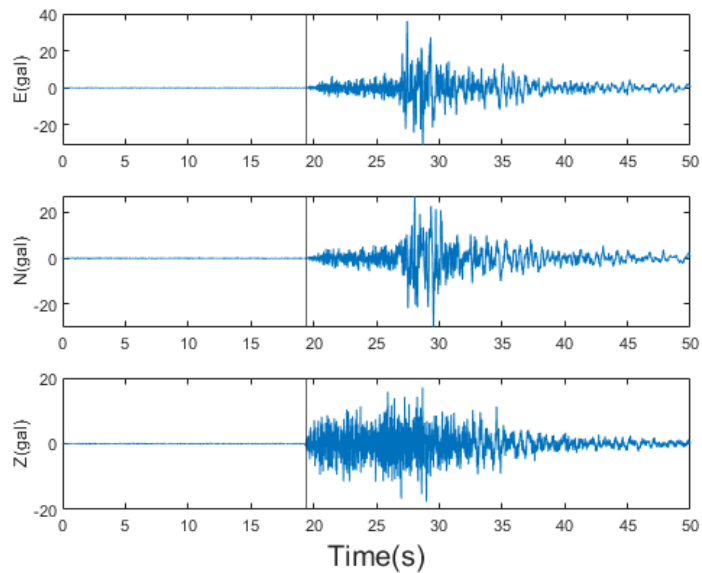


圖 3.2-2 P、S 波都含之單一事件切之前之原始地動加速度歷時

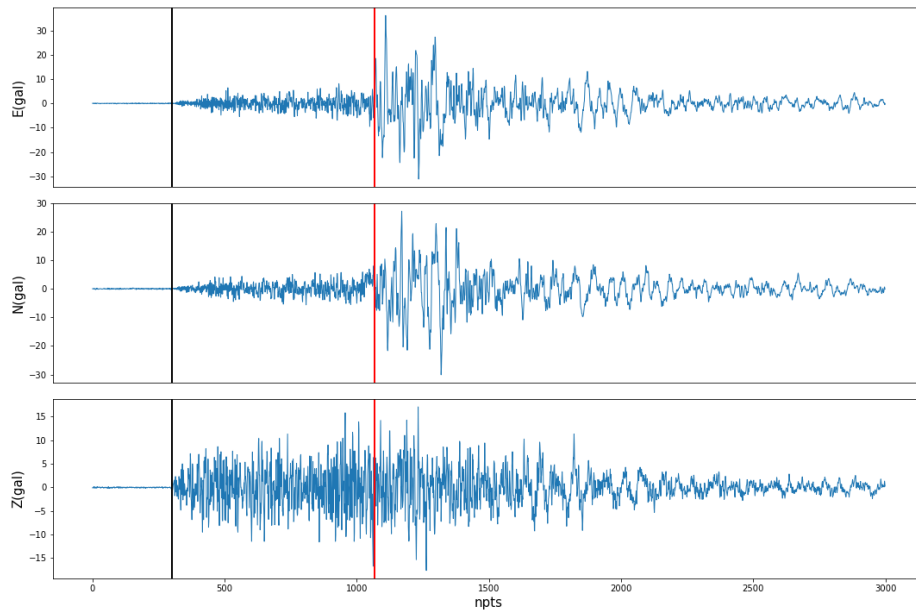


圖 3.2-3 P、S 波都含之單一事件切之後之地動加速度歷時

只含一小段 P 波訊號之地動加速度歷時，如圖 3.2-4，在 P、S 波都含之單一事件中取 P、S 波到時間隨機選一點往前 30 秒之地震訊號，若地震訊號長度不足 30 秒，會重複從頭到 P 波到時前 1 秒之訊號，將資料補到 30 秒。

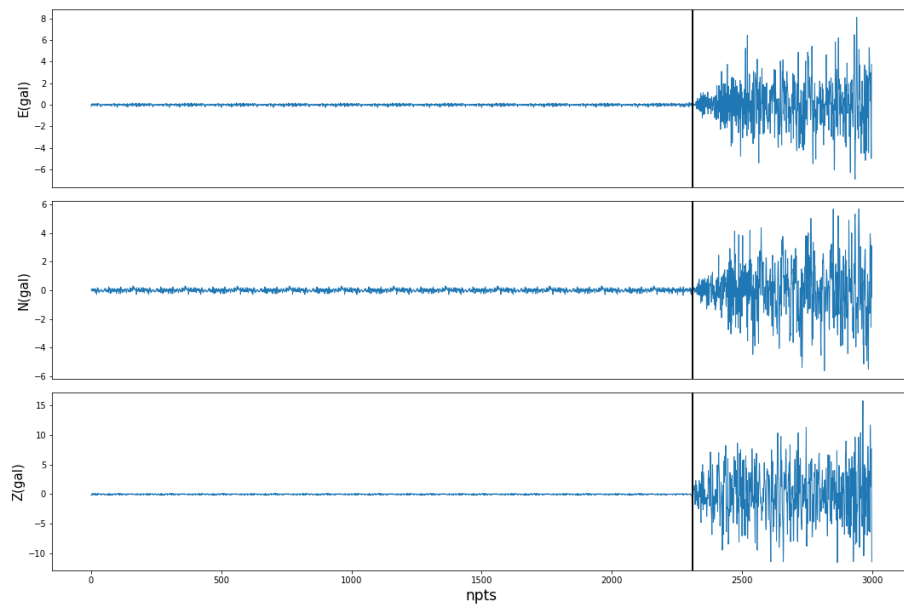


圖 3.2-4 只含一小段 P 波訊號之地動加速度歷時

表 3.2-1 為 P、S 波都含之單一事件和只含一小段 P 波訊號在訓練和驗證資料集之數量：

表 3.2-1 按照兩種資料之訓練和驗證資料數量

	訓練資料集	驗證資料集
P、S 波都含之單一事件	28567 筆	13214 筆
只含一小段 P 波訊號	33551 筆	8393 筆

表 3.2-2、表 3.2-3 為訓練和驗證資料集之 PGA 區間分布，此兩個資料集之 PGA 區間分布是以地震事件為單位統計。0.8gal 以下之地震資料都為 Liao, et al. [1] 訓練模型之地震資料，且這些資料幾乎都來自速度計。為了讓本研究之模型擁有如 RED-PAN 標記地震相位之能力，所以本研究不刪減 0.8gal 以下之資料。為了在比較各種不同模型時更快，故本研究在訓練、驗證資料集取每個區間之 30%。

表 3.2-2 訓練資料集 PGA 區間分布

PGA 區間	0.8gal 以下	0-0.8gal	2.5-8.0gal	8.0-25gal	25-80gal	80-250gal	250-800gal	800gal 以上
數量	19015	5375	5669	5669	5669	1797	215	30

表 3.2-3 驗證資料集 PGA 區間分布

PGA 區間	0.8gal 以下	0-0.8gal	2.5-8.0gal	8.0-25gal	25-80gal	80-250gal	250-800gal	800gal 以上
數量	4714	1330	1418	1418	1418	450	54	8

3.2.2 測試資料集

本研究之測試資料集分別為 P 波到時後 0.1 秒到 3.0 秒，表 3.2-4 為測試資料集之數量：

表 3.2-4 按照兩種資料之測試資料數量

	測試資料集
Ground Truth 後 0.1 秒到 3 秒	12545-12563 筆

表 3.2-5、表 3.2-6、表 3.2-7 為測試資料集之 PGA 區間分布，此資料集也是以地震事件統計各 PGA 區間數量。0.8gal 以下之地震資料也都為 Liao, et al. [1] 之測試資料集，為了評估本研究之模型標記地震相位能力，所以不刪減 0.8gal 以下之地震資料。

表 3.2-5 測試資料集 Ground Truth 後 0.1 到 1.0 秒之資料數量

PGA 區間	0.8gal 以下	0- 0.8gal	2.5- 8.0gal	8.0- 25gal	25- 80gal	80- 250gal	250- 800gal	800gal 以上
0.1 秒	6384	1409	1607	1237	1418	447	53	6
0.2 秒	6383	1408	1606	1238	1418	447	53	6
0.3 秒	6380	1409	1606	1238	1418	447	53	6
0.4 秒	6381	1409	1607	1238	1418	447	53	6
0.5 秒	6380	1409	1606	1238	1418	447	53	6
0.6 秒	6379	1408	1606	1238	1418	447	53	6
0.7 秒	6370	1406	1607	1238	1418	447	53	6
0.8 秒	6384	1409	1607	1238	1418	447	53	6
0.9 秒	6379	1408	1606	1238	1418	447	53	6
1.0 秒	6381	1409	1606	1238	1418	447	53	6

表 3.2-6 測試資料集 Ground Truth 後 1.1 到 2.0 秒之資料數量

PGA 區間	0.8gal 以下	0- 0.8gal	2.5- 8.0gal	8.0- 25gal	25- 80gal	80- 250gal	250- 800gal	800gal 以上
1.1 秒	6370	1408	1606	1238	1418	447	53	6
1.2 秒	6379	1409	1607	1238	1418	447	53	6
1.3 秒	6380	1409	1607	1238	1418	447	53	6
1.4 秒	6385	1407	1607	1238	1418	447	53	6
1.5 秒	6380	1408	1607	1237	1418	447	53	6
1.6 秒	6379	1408	1607	1236	1418	447	53	6
1.7 秒	6375	1408	1607	1238	1418	447	53	6
1.8 秒	6386	1409	1606	1238	1418	447	53	6
1.9 秒	6375	1409	1606	1237	1418	447	53	6
2.0 秒	6376	1409	1607	1238	1418	447	53	6

表 3.2-7 測試資料集 Ground Truth 後 2.1 到 3.0 秒之資料數量

PGA 區 間	0.8gal 以下	0- 0.8gal	2.5- 8.0gal	8.0- 25gal	25- 80gal	80- 250gal	250- 800gal	800gal 以上
2.1 秒	6380	1409	1607	1238	1418	447	53	6
2.2 秒	6379	1409	1607	1238	1418	447	53	6
2.3 秒	6381	1408	1607	1237	1418	447	53	6
2.4 秒	6383	1409	1607	1238	1418	447	53	6
2.5 秒	6385	1407	1607	1238	1418	447	53	6
2.6 秒	6379	1409	1607	1238	1418	447	53	6
2.7 秒	6381	1409	1606	1238	1418	447	53	6
2.8 秒	6379	1409	1607	1238	1418	447	53	6
2.9 秒	6380	1408	1607	1238	1418	447	53	6
3.0 秒	6383	1409	1606	1238	1418	447	53	6

3.3 模型之輸入及輸出

3.3.1 模型輸入

由於地震歷時訊號的值範圍較大，故輸入資料的尺度變得較為重要。本研究探討綜合使用不同縮放尺度方式的資料作為輸入預測模型所造成的影響，主要包括標準化，以及裁切固定範圍訊號並進行正規化等兩類。標準化的好處在於可以觀察資料整體的形狀，因為全部的歷時訊號的形狀都可被觀察，但缺少了輸入資料振幅大小的資訊。而裁切固定範圍訊號並進行正規化的方式，其好處在於可以提供輸入資料振幅大小的資訊，且當裁切較小範圍時，可觀察較為細緻的地震歷時訊號，而當裁切較大範圍時，可觀察整體的地震歷時訊號特性。以下詳細介紹。

標準分數(Standard Score, Z-score)計算公式如式(3.3-1)， $Acc(t_i)$ 為每個時間點之加速度值， $Acc(t_{all})$ 為整段加速度歷時，長度為 3000 個時間點， $mean(Acc(t_{all}))$ 為整段加速度歷時之平均值， $\sigma(Acc(t_{all}))$ 為整段加速度歷時之標準差。

$$Acc(t_i)_{zscore} = \frac{Acc(t_i) - mean(Acc(t_{all}))}{\sigma(Acc(t_{all}))}, i = 1 \sim 3000 \quad (3.3-1)$$

採用標準分數之原因是經過此標準化後，可以讓模型辨識加速度波形形狀之特徵，例如本研究在訓練時加入多筆地動加速度歷時，讓模型可辨識其之 P 波到時和 S 波到時，也有加上多筆噪音(Noise)之加速度歷時，讓模型辨識其特徵而不要錯誤地對其標記 P 波到時、S 波到時，但此標準化沒有加速度振幅大小之概念。

而本研究之模型需要預測最大地動加速度值(PGA)是否大於 25gal，所以需要額外加上有關加速度振幅大小的資訊，讓模型可以預測 PGA 是否大於 25gal。所

以本研究將標準分數加上以下其中一個加速度區間，(一)、正負 0.8gal，(二)、正負 2.5gal，(三)、正負 8.0gal，(四)、正負 25.0gal，(五)、正負 80gal，(六)、正負 250gal，(七)、正負 800gal，讓模型有加速度振幅大小之概念，這些加速度區間是根據中央氣象署(Central Weather Administration)新制之地震震度階級對照最大地動加速度範圍表而訂定。本研究還有將標準分數加上原始地動加速度歷時，以及標準分數加上以上個別七個加速度區間之結果做比較。原始地動加速度歷時的意思是不對地動加速度歷時做加速度區間之正規化。

不同加速度區間正規化之公式如下式(3.3-2)，式中之加速度區間值為 0.8gal、2.5gal、8.0gal、25gal、80gal、250gal、800gal。式(3.3-2)說明長度為 3000 之加速度歷時中，加速度值大於 0 時，加速度值大於等於加速度區間值，其值就等於加速度區間值，並除以加速度區間值等於 1，若加速度值小於加速度區間值，其值就等於原加速度值，並除以加速度區間值。加速度值小於 0 時，加速度值小於等於負加速度區間值，其值就等於負加速度區間值，並除以加速度區間值等於 -1，若加速度值大於負加速度區間值，其值就等於原加速度值，並除以加速度區間值。

$$\left\{ \begin{array}{l} \text{Acc}(t_i) > 0, i = 1 \sim 3000, \left\{ \begin{array}{l} \text{Acc}(t_i) \geq \text{加速度區間值}, \text{Acc}(t_i) = \frac{\text{加速度區間值}}{\text{加速度區間值}} = 1 \\ \text{Acc}(t_i) < \text{加速度區間值}, \text{Acc}(t_i) = \frac{\text{Acc}(t_i)}{\text{加速度區間值}} \end{array} \right. \\ \text{Acc}(t_i) < 0, i = 1 \sim 3000, \left\{ \begin{array}{l} \text{Acc}(t_i) \leq -\text{加速度區間值}, \text{Acc}(t_i) = \frac{-\text{加速度區間值}}{\text{加速度區間值}} = -1 \\ \text{Acc}(t_i) > -\text{加速度區間值}, \text{Acc}(t_i) = \frac{\text{Acc}(t_i)}{\text{加速度區間值}} \end{array} \right. \end{array} \right. \quad (3.3-2)$$

圖 3.3-1、圖 3.3-2、圖 3.3-3 都為 1993 年 12 月 16 日凌晨 5 點 49 分 43 秒在 TCU060 之加速度歷時，其 PGA 為 6.24gal。圖 3.3-1 為 TCU060 之原始加速度歷時、圖 3.3-2 是顯示 TCU060 原始加速度歷時正負 2.5gal 間之訊號、圖 3.3-3 是顯示 TCU060 原始加速度歷時正負 250gal 間之訊號。觀察圖 3.3-2、圖 3.3-3 可以發現小地震在正負 2.5gal 間可以觀察波形中較細緻之形狀，但把加速度尺度放大到 250gal 間，就看不到加速度歷時。圖 3.3-4、圖 3.3-5、圖 3.3-6 都為 1993 年 12 月 16 日凌晨 5 點 49 分 43 秒在 CHY015 之加速度歷時，其 PGA 為 290.28gal。圖 3.3-4 為 CHY015 之原始加速度歷時、圖 3.3-5 是顯示 CHY015 原始加速度歷時正負 2.5gal 間之訊號、圖 3.3-6 是顯示 CHY015 原始加速度歷時正負 250gal 間之訊號。觀察圖 3.3-5、圖 3.3-6 可發現大地震在正負 2.5gal 間，加速度歷時密密麻麻，看不出加速度歷時之波形形狀，但把尺度放大到 250gal 可以看出整體加速度歷時之波形形狀。比較圖 3.3-2、圖 3.3-5 即可發現同一個尺度下，大小地震之差別。所以本研究認為裁切固定範圍訊號並進行正規化的方式，可讓模型擁有地震大小之概念，而要在本研究設定之加速度區間值中選擇哪個尺度，即為本研究之重點。

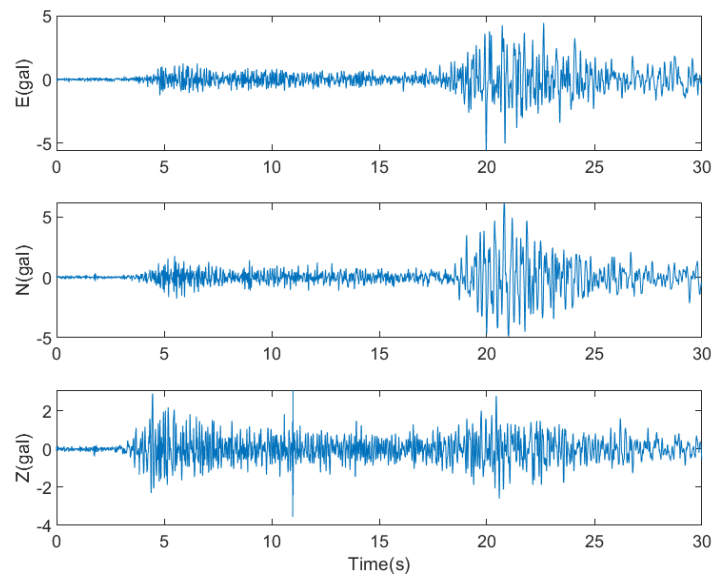


圖 3.3-1 TCU060 之原始加速度歷時

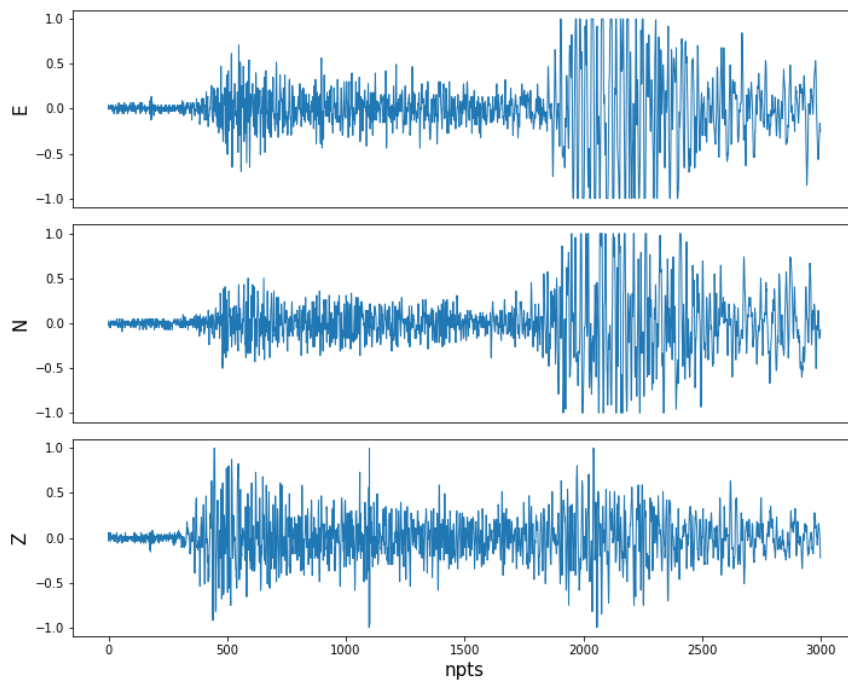


圖 3.3-2 小地震在正負 2.5gal 間之加速度歷時

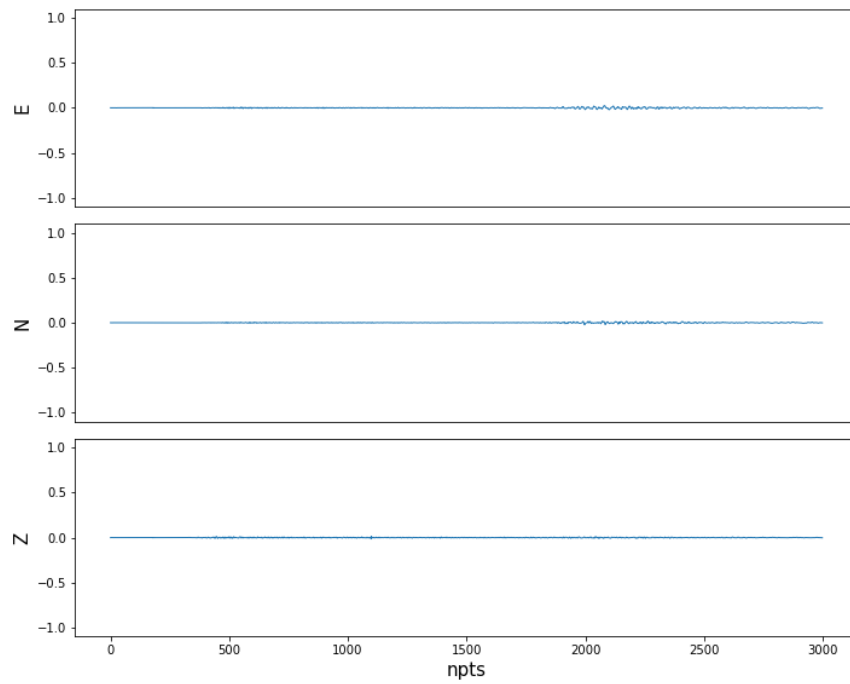


圖 3.3-3 小地震在正負 250gal 間之加速度歷時

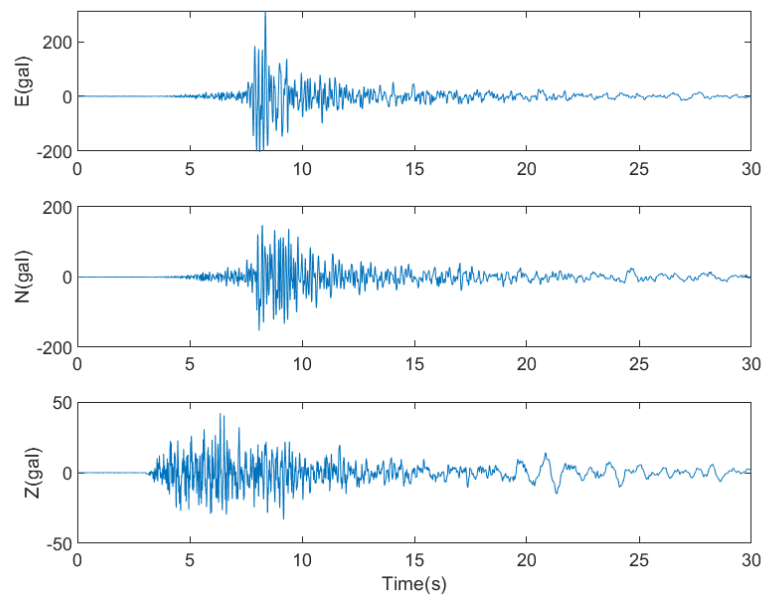


圖 3.3-4 CHY015 之原始加速度歷時

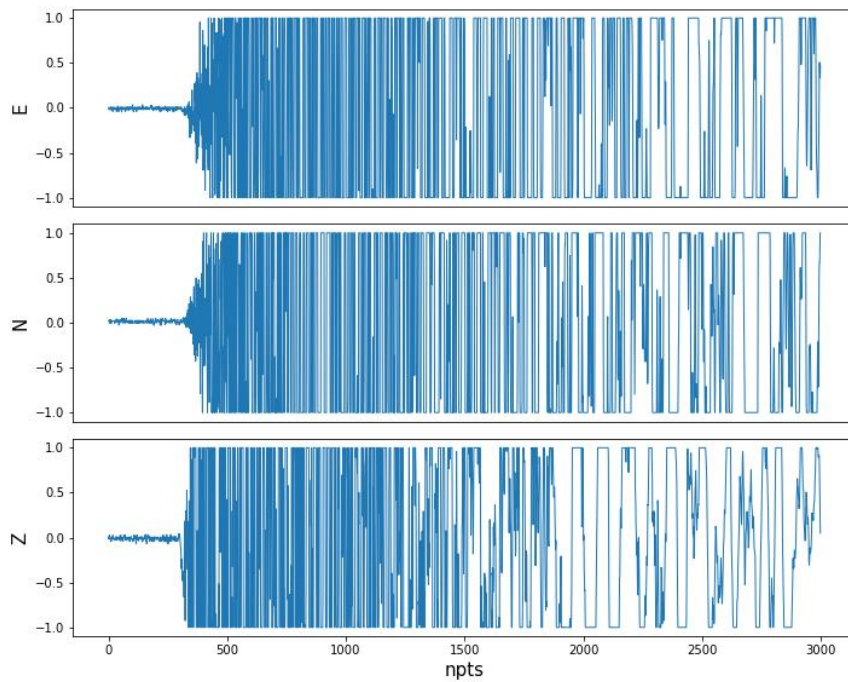


圖 3.3-5 大地震在正負 2.5gal 間之加速度歷時

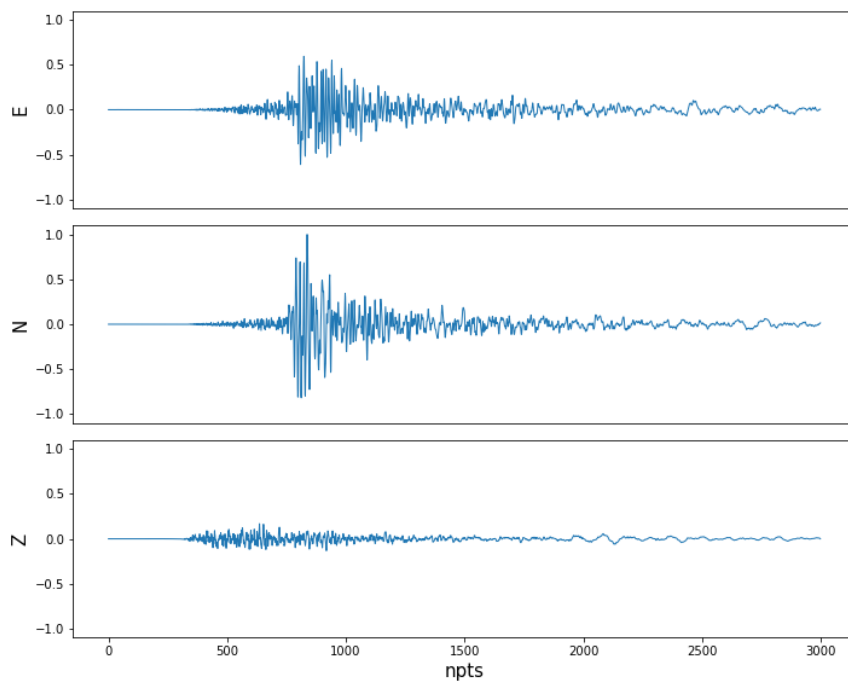
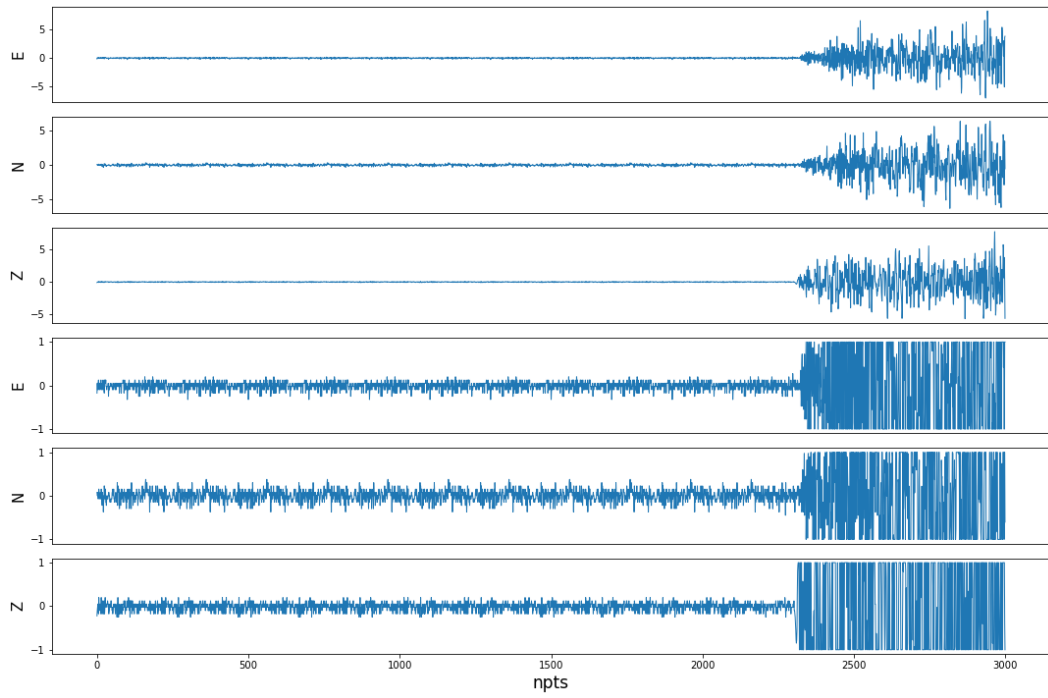
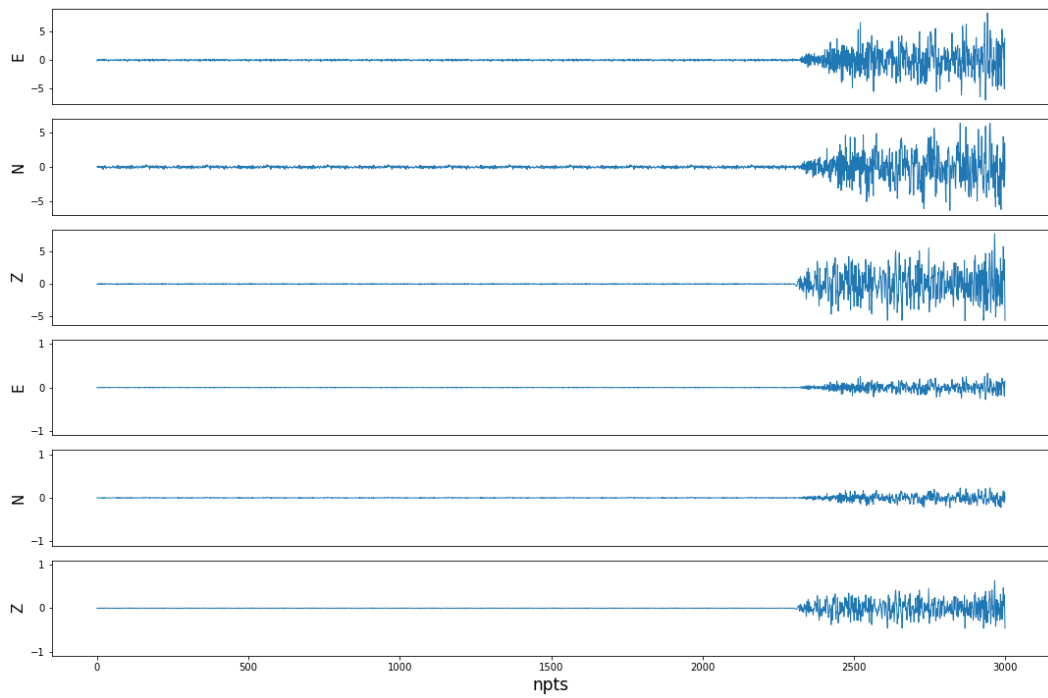


圖 3.3-6 大地震在正負 250gal 間之加速度歷時

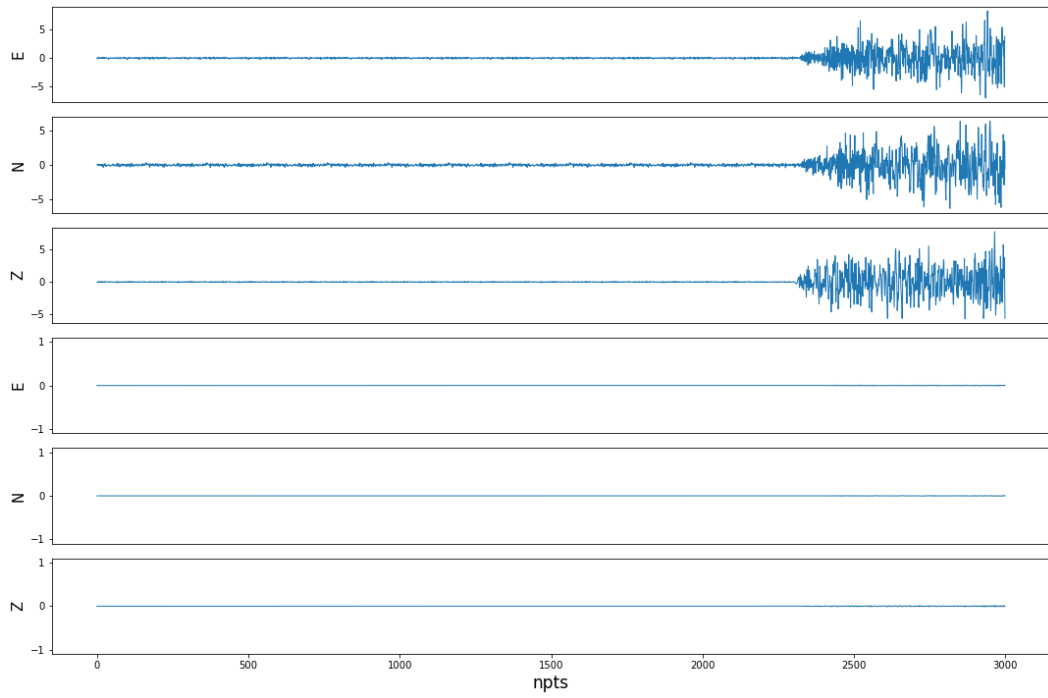
圖 3.3-7 (a)(b)(c)(d)(e)(f) 兩種波形資料和三種加速度區間為典型之輸入資料，包括 P、S 波都含之單一事件、只含一小段 P 波訊號且輸入為標準分數+正負 0.8gal、標準分數+正負 25gal、標準分數+正負 800gal 等。上面三行分別為 E、N、Z 且做 Z-score，下面三行為 E、N、Z 且做三個加速度區間之正規化。



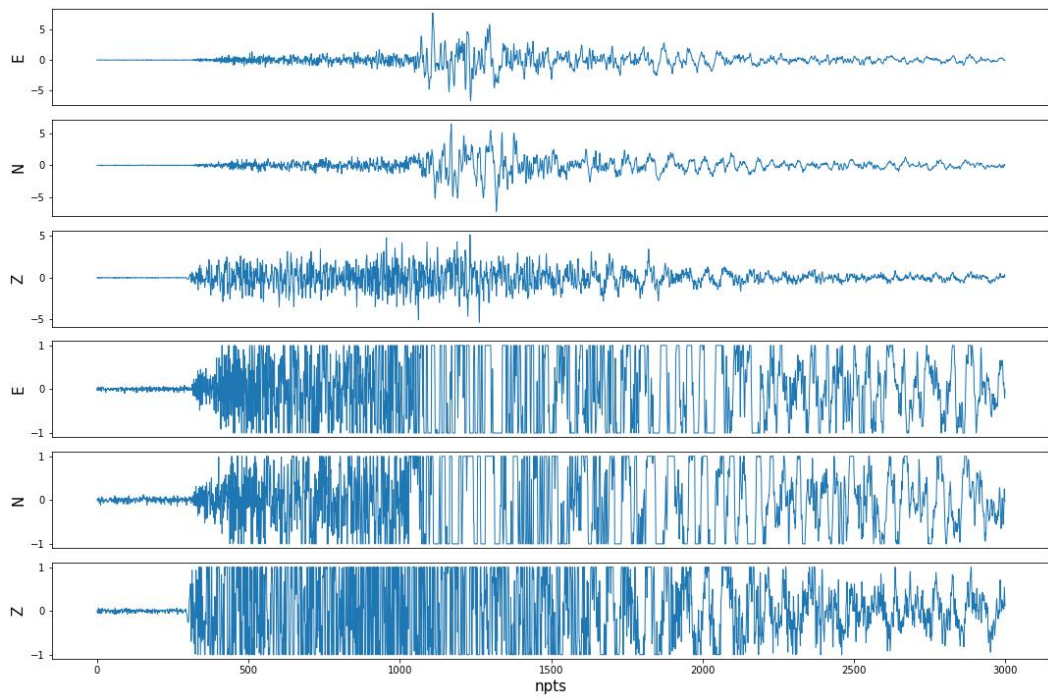
(a)、只含一小段 P 波訊號且輸入為標準分數+正負 0.8gal



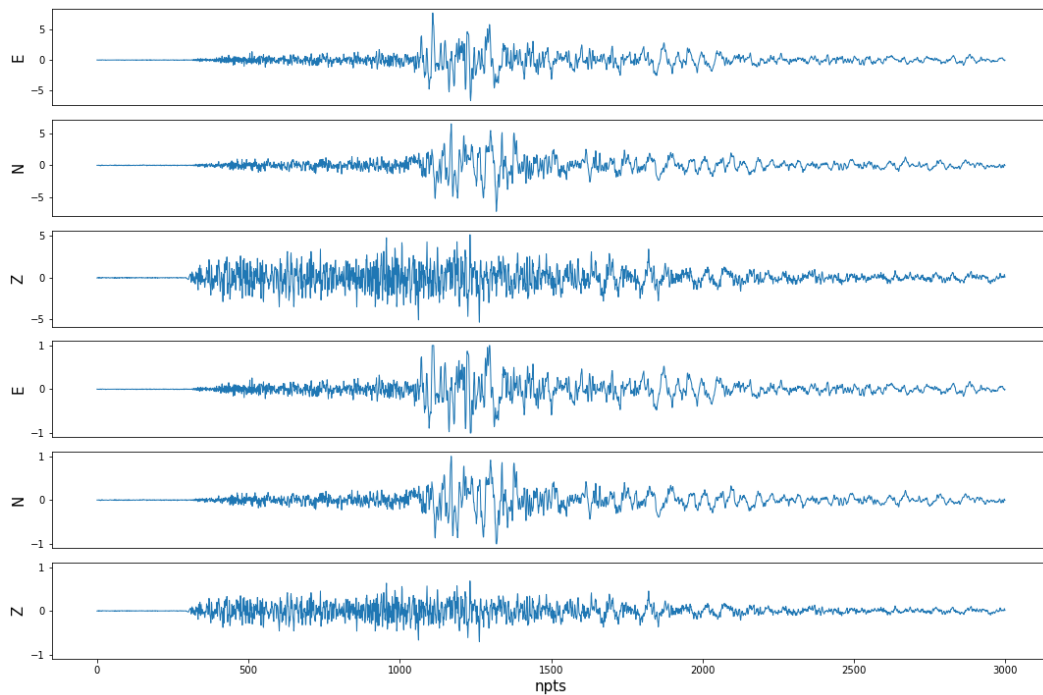
(b)、只含一小段 P 波訊號且輸入為標準分數+正負 25gal



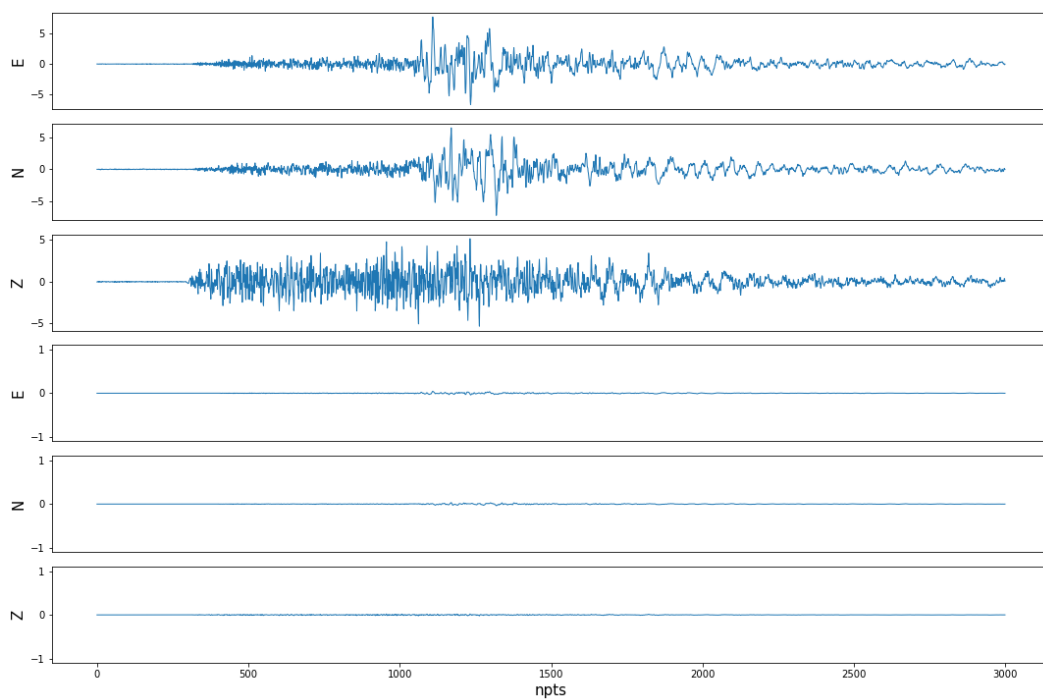
(c)、只含一小段 P 波訊號且輸入為標準分數+正負 800gal



(d)、P、S 波都含之單一事件且輸入為標準分數+正負 0.8gal



(e)、P、S 波都含之單一事件且輸入為標準分數+正負 25gal



(f)、P、S 波都含之單一事件且輸入為標準分數+正負 800gal

圖 3.3-7 (a)(b)(c)(d)(e)(f) 兩種波形資料和三種加速度區間

3.3.2 模型輸出

本報告之模型輸出分為三種輸出，分別為標記地震相位(Earthquake Phase Picking)、地震波形檢測(Earthquake Waveform Detection)、預測 PGA 是否大於 25gal(Predict whether PGA is greater than 25 gal)。三種輸出更細部之分類如下：

- 標記地震相位(Earthquake Phase Picking)
- P 波到時機率隨時間函數 $Prob_{P\ phase}(t)$
- S 波到時機率隨時間函數 $Prob_{S\ phase}(t)$
- 其他 $Prob_{Others}(t)$
- 地震波形檢測(Earthquake Waveform Detection)
- 遮罩 $Prob_{Mask}(t)$
- 反遮罩 $Prob_{Unmask}(t)$
- 預測 PGA 是否大於 25gal(Predict whether PGA is greater than 25 gal)
- 預測 PGA 是否超過 25gal 之機率隨時間函數 $Prob_{PGAUpOrSmall\ 25\ gal}(t)$
- 反預測 PGA 是否超過 25gal 之機率隨時間函數 $Prob_{UnPGAUpOrSmall\ 25\ gal}(t)$

接下來分別介紹各個輸出，圖 3.3-8 中之 Prob_P 為 P 波到時機率隨時間函數 $Prob_{P\ phase}(t)$ 、Prob_S 為 S 波到時機率隨時間函數 $Prob_{S\ phase}(t)$ ，Prob_others 為其他 $Prob_{Others}(t)$ 。P 波到時機率隨時間函數 $Prob_{P\ phase}(t)$ 、S 波到時機率隨時間函數 $Prob_{S\ phase}(t)$ 採用截斷高斯分布(Truncated Gaussian Function)，其為常態分佈(Normal Distribution)保留 ground truth 正負標準差之內的值，以外的值皆為零，P 波到時機率隨時間函數 $Prob_{P\ phase}(t)$ 之標準差為 0.2 秒[10]；S 波到時機率隨時間函數 $Prob_{S\ phase}(t)$ 之標準差為 0.3 秒[10]；其他 $Prob_{Others}(t)$ 為 1.0 減掉 P 波到時機率隨時間函數 $Prob_{P\ phase}(t)$ 和 S 波到時機率隨時間函數 $Prob_{S\ phase}(t)$ ；圖 3.3-9 中之 Prob_mask 為遮罩 $Prob_{Mask}(t)$ 、Prob_unmask 為反遮罩 $Prob_{Unmask}(t)$ 。遮罩 $Prob_{Mask}(t)$ 為 P 波到時機率隨時間函數 $Prob_{P\ phase}(t)$ 加上 S 波到時機率隨時間函數 $Prob_{S\ phase}(t)$ 且在 P 波到時和 S 波到時之間的值為 1.0 [11]；反遮罩 $Prob_{Unmask}(t)$ 為 1.0 減掉遮罩 $Prob_{Mask}(t)$ [11]；圖 3.3-10、圖 3.3-11 中的 Prob_PGAUpto25galornot 為預測 PGA 是否超過 25gal 之機率隨時間 $Prob_{PGAUpOrSmall\ 25\ gal}(t)$ 、Prob_unPGAUperto25galornot 為反預測 PGA 是否超過 25gal 之機率隨時間 $Prob_{UnPGAUpOrSmall\ 25\ gal}(t)$ 。PGA 為超過 25gal 之機率隨時間 $Prob_{PGAUpOrSmall\ 25\ gal}(t)$ 是若此地震事件 PGA 超過 25gal，PGA 超過 25gal 之機率隨時間 $Prob_{PGAUpOrSmall\ 25\ gal}(t)$ 就等於遮罩(mask)，而反 PGA 超過 25gal 之機率隨時間 $Prob_{UnPGAUpOrSmall\ 25\ gal}(t)$ 為 1.0 減 PGA 超過 25gal 之機率隨時間 $Prob_{PGAUpOrSmall\ 25\ gal}(t)$ ，如圖 3.3-10；若此地震事件之 PGA 沒有超過 25gal，PGA 超過 25gal 之機率隨時間 $Prob_{PGAUpOrSmall\ 25\ gal}(t)$ 就等於零，反 PGA 超過 25gal 之機率隨時間 $Prob_{UnPGAUpOrSmall\ 25\ gal}(t)$ 為 1.0，如圖 3.3-11。

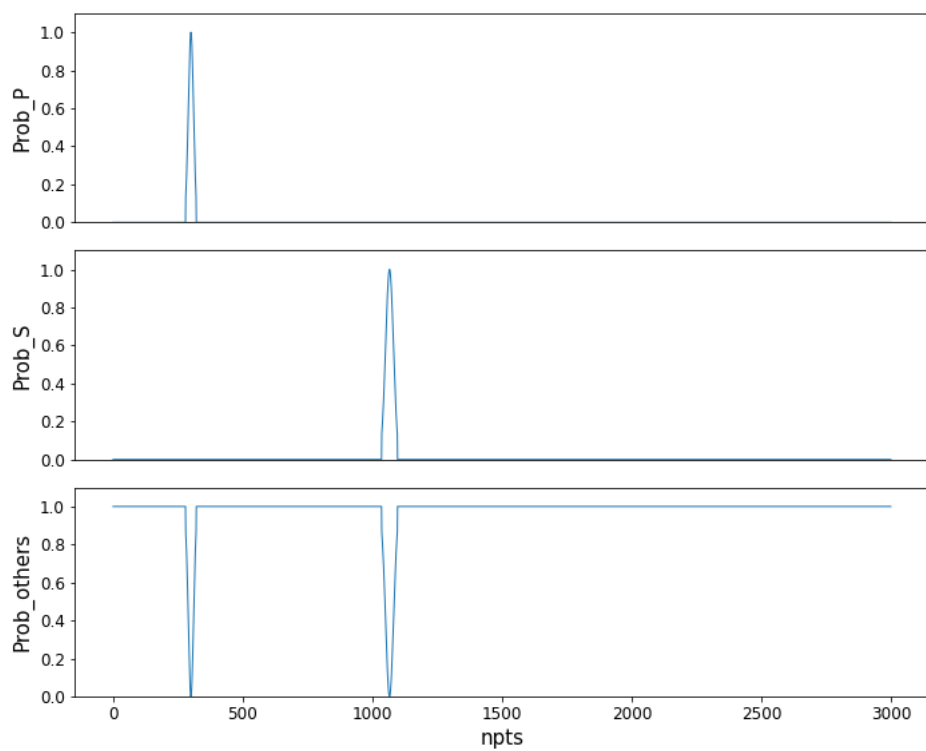


圖 3.3-8 標記地震相位

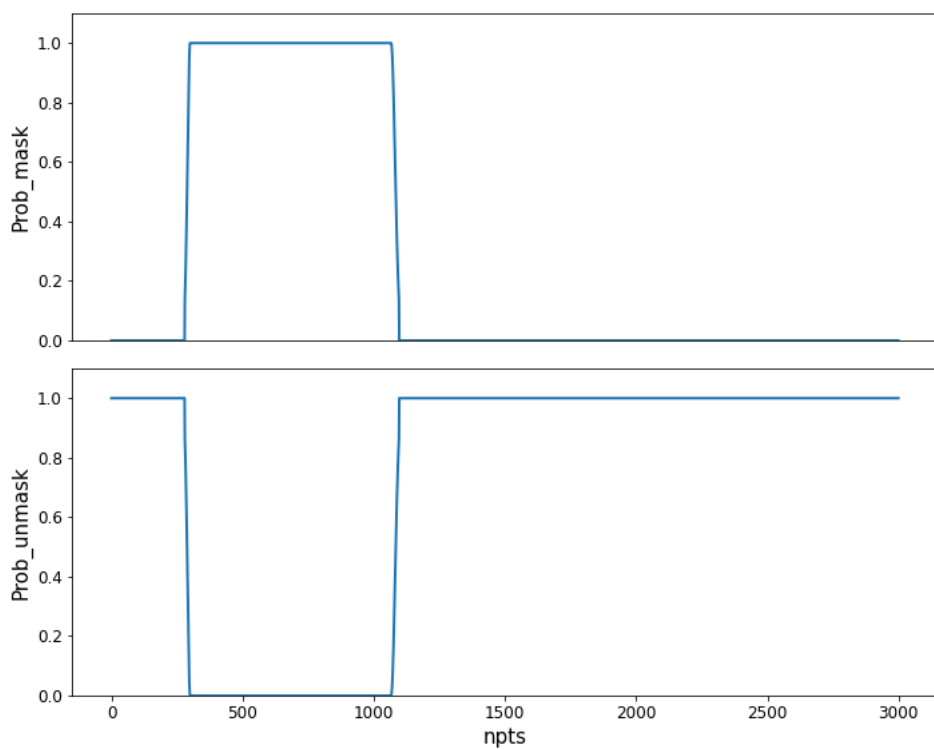


圖 3.3-9 地震波形檢測

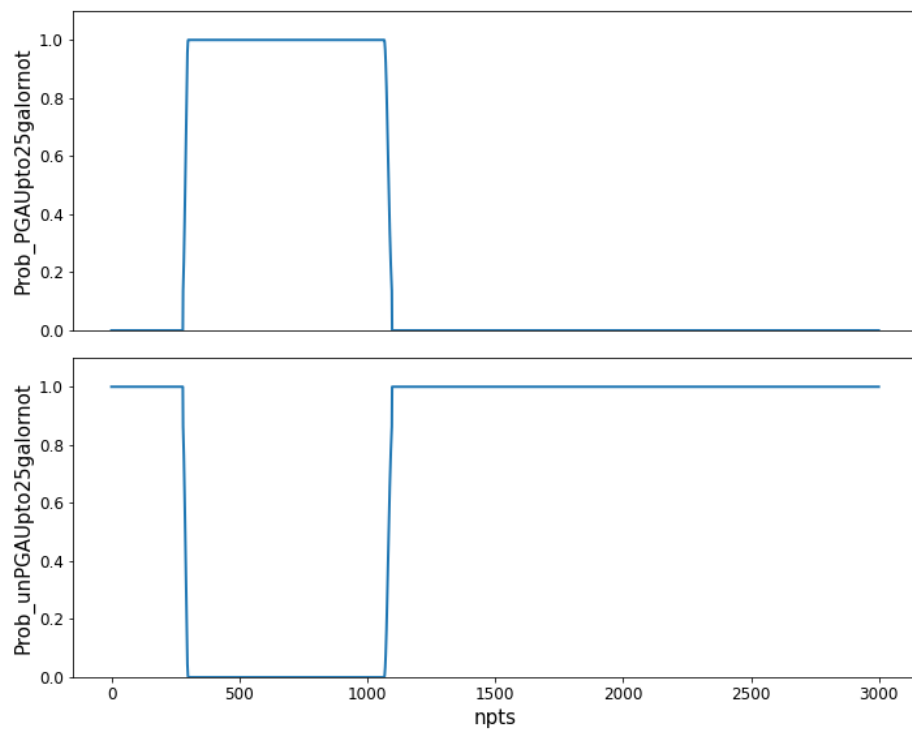


圖 3.3-10 預測 PGA 是否超過 25gal (PGA 大於 25gal)

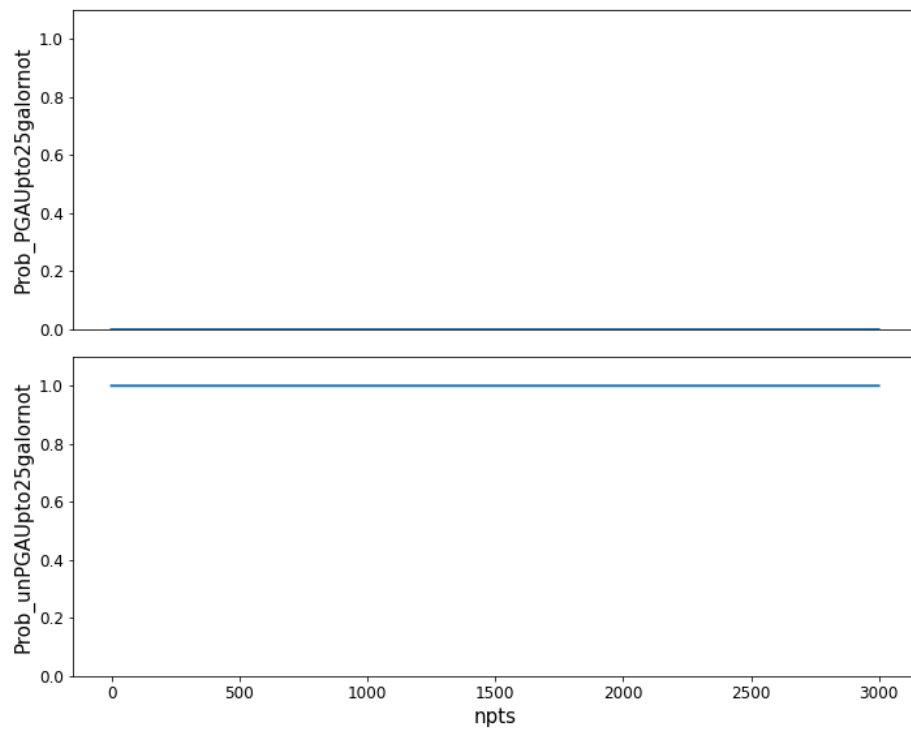


圖 3.3-11 預測 PGA 是否超過 25gal (PGA 小於 25gal)

3.4 說明三種目標之二元混淆矩陣

本研究用二元混淆矩陣評估模型對於三種目標的正確性。準確來說是利用二元混淆矩陣中的 TP(True Positive, 真陽性)、TN(True Negative, 真陰性)、FP(False Positive, 偽陽性)、FN(False Negative, 偽陰性)計算準確率(Precision)、召回率(Recall)、F1-score, 並採用這三個值評估模型成效, 這三個值都是越大越好, 公式如(3.4-1)、(3.4-2)、(3.4-3)。後面三個小節會定義各目標之二元混淆矩陣。

$$Precision = \frac{TP}{TP + FP} \quad (3.4-1)$$

$$Recall = \frac{TP}{TP + FN} \quad (3.4-2)$$

$$F1-Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (3.4-3)$$

本研究以圖 3.4-1 說明模型如何輸出, 圖 3.4-1 為 2012 年 3 月 11 日凌晨 2 點 53 分觸發之地震資料, 測站為 NTC, PGA 為 38.98gal。地震相位標記為運用模型預測之 $Prob_{P\ phase}(t)$ 、 $Prob_{S\ phase}(t)$ 之最大值和最大值位置跟本研究設定之機率值門檻值 0.3 和 ground truth 正負 0.5 秒, 以上兩個值參考 Liao, et al. [1]。透過二元混淆矩陣評估八種模型對於地震相位標記之表現, $Prob_{P\ phase}(t)$ 、 $Prob_{S\ phase}(t)$ 為圖 3.4-1 中之 P、S, 詳細說明會在 3.4.1。地震波形檢測為運用模型預測之 $Prob_{Mask}(t)$ 判斷訊號是否為地震訊號, 如何運用 $Prob_{Mask}(t)$ 判斷是否為地震並比較判斷結果和真實結果, 在 3.4.3 詳細說明, $Prob_{Mask}(t)$ 為圖 3.4-1 中之 mask。單站強震預警為運用 $Prob_{PGAUpOrSmall25gal}(t)$ 判斷 PGA 是否大於 25gal, 如何運用 $Prob_{PGAUpOrSmall25gal}(t)$ 判斷 PGA 是否大於 25gal 並比較判斷結果和真實結果, $Prob_{PGAUpOrSmall25gal}(t)$ 為圖 3.4-1 中之 PGApob, 在 3.4.4 詳細說明。

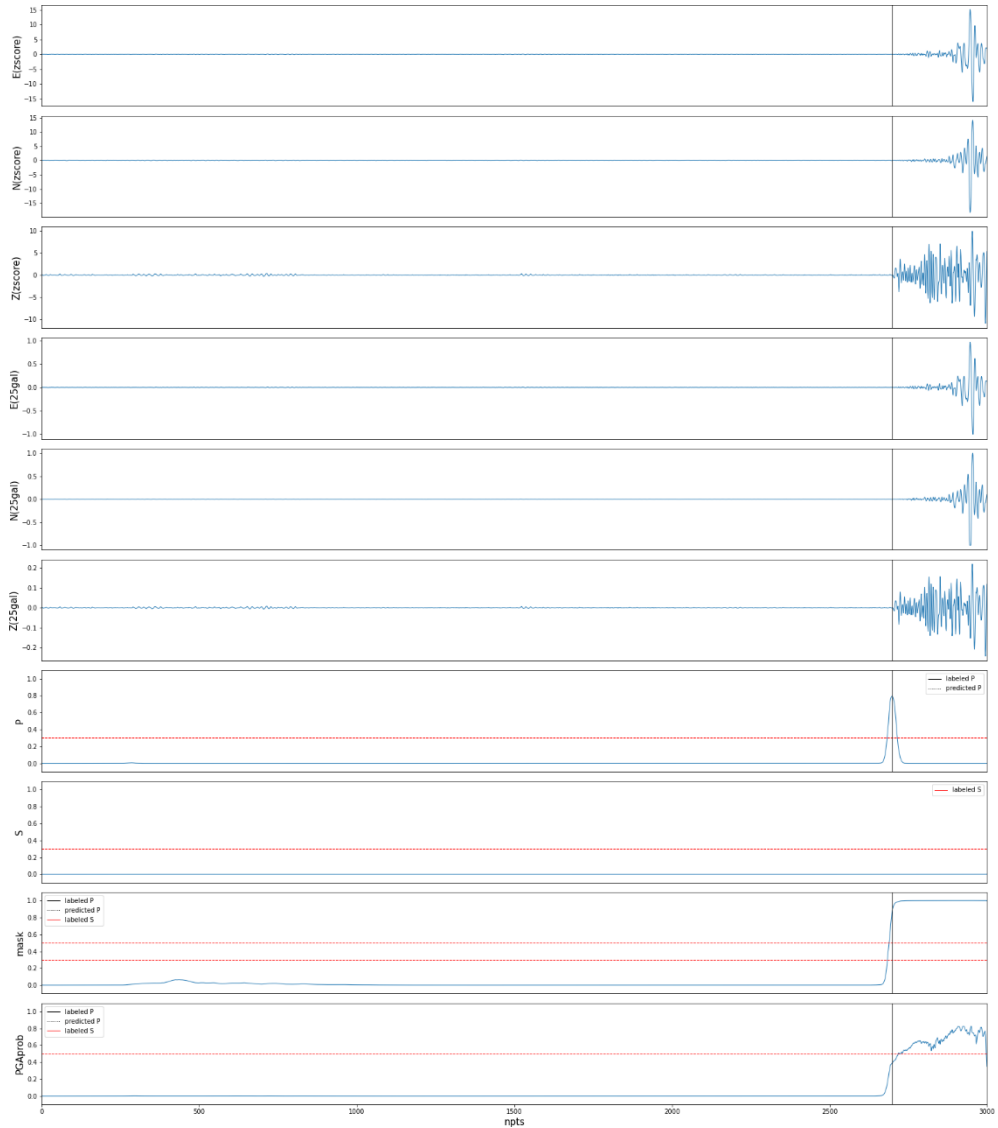


圖 3.4-1 模型輸入和模型輸出結果

3.4.1 地震相位標記(Earthquake Phase Picking)之二元混淆矩陣

本研究之測試資料集含有地震資料和非地震資料，而此目標是針對地震資料中的 P、S 波到時，本研究運用模型預測之 $Prob_{P\ phase}(t)$ 、 $Prob_{S\ phase}(t)$ 中之最大值和最大值位置評估八個模型對於地震相位標記之表現。圖 3.4-2 為地震相位標記之二元混淆矩陣。本研究設定模型預測之 $Prob_{P\ phase}(t)$ 、 $Prob_{S\ phase}(t)$ 中之最大值高於 0.3，即認定模型認為 ground truth 在此位置，若此機率最大值位置在本研究設定之 ground truth 正負 0.5 秒間，則為 TP；若模型預測之 $Prob_{P\ phase}(t)$ 、 $Prob_{S\ phase}(t)$ 中之最大值低於 0.3，表示模型不認為 ground truth 在此位置，但此機率最大值位置在 ground truth 正負 0.5 秒間，表示模型漏掉這個 phase，此情況為 FN；若模型預測之 $Prob_{P\ phase}(t)$ 、 $Prob_{S\ phase}(t)$ 中之最大值位置在 ground truth 正負 0.5 秒外且此機率最大值大於 0.3，表示模型失誤認定此位置有 phase，但實際上此位置沒有 phase，所以此情況即為 FP；若模型預測之 $Prob_{P\ phase}(t)$ 、 $Prob_{S\ phase}(t)$ 中之最大值低於 0.3，

此機率最大值位置在 ground truth 正負 0.5 秒外，表示模型認為此位置沒有 phase，實際上此位置也沒有 phase，此情況為 TN。

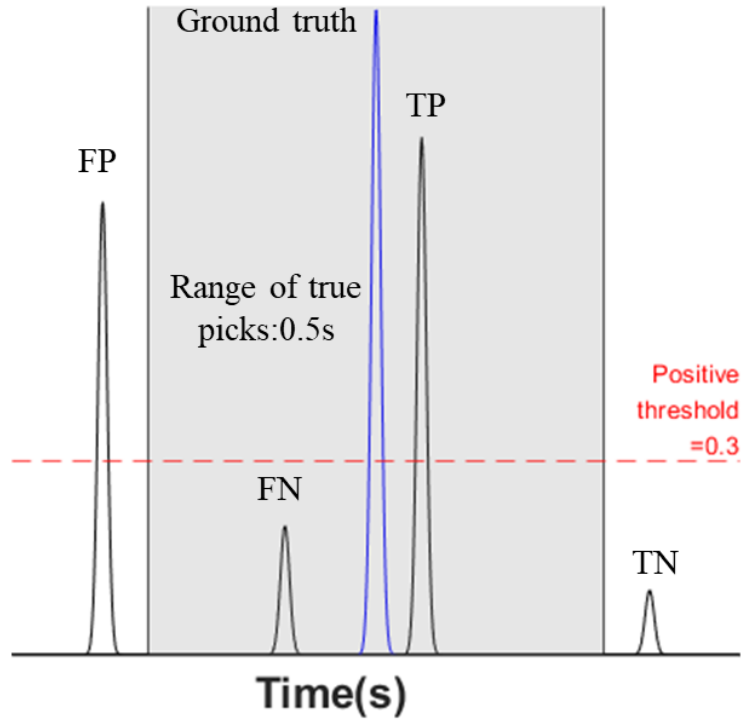


圖 3.4-2 地震相位標記之二元混淆矩陣

3.4.2 模型對非地震訊號標記之錯誤率

3.4.1 是模型針對地震訊號中 P、S 波標記結果之二元混淆矩陣，此小節針對若模型對非地震訊號標記 P 波之情況。因為非地震訊號不會有 P 波到時，所以如果 $Prob_{P\ phase}(t)$ 中最大值大於 0.3，本研究認為模型誤標記 P 波到時，因為 $Prob_{P\ phase}(t)$ 中最大值大於 0.3 表示模型認為 ground truth 在此最大值位置，但非地震訊號不會有 ground truth，所以本研究定義如果 $Prob_{P\ phase}(t)$ 中最大值大於 0.3，就算模型對非地震訊號標記 P 波到時，反之，模型沒有對非地震訊號標記 P 波到時，如式(3.4-4)。

$$\begin{cases} \max(P_{P\ phase}(t)) \geq 0.3 \Rightarrow False \\ \max(P_{P\ phase}(t)) < 0.3 \Rightarrow True \end{cases} \quad (3.4-4)$$

3.4.3 地震波形檢測(Earthquake Waveform Detection)之二元混淆矩陣

$Prob_{Mask}(t)$ 大於 0.3 之時間點為 $t_{prob>0.3}$ ，隨著時間往前走， $Prob_{Mask}(t)$ 會有小於 0.3 之時間點 $t_{prob<0.3}$ ，本研究取 $t_{prob>0.3}$ 、 $t_{prob<0.3}$ 間 $Prob_{Mask}(t)$ 之平均值，若此平均值大於等於 0.5，即表示模型認為此訊號為地震訊號，反之為非地震訊號，以圖 3.4-1 中 mask 為例， $Prob_{Mask}(t)$ 大於 0.3 之資料點在 2650，從此點開始到最後一個資料點計算 $Prob_{Mask}(t)$ 之平均值，因為 mask 沒有下降低於 0.3，所以就會一直算到最後

一個資料點，而此例子之平均值大於 0.5，所以模型認為此訊號為地震訊號。 $Prob_{Mask}(t)$ 為 $Prob_{P\ phase}(t)$ 加 $Prob_{S\ phase}(t)$ 且在 P、S 波到時間的值為 1，此物理意義為檢視此訊號是否為 P 波。若此訊號為地震訊號，模型也認為是地震訊號，此情況為 TP；若此訊號為地震訊號，模型認為是非地震訊號，此情況為 FN；若此訊號為非地震訊號，模型認為是地震訊號，此情況為 FP；若此訊號為非地震訊號，模型認為是非地震訊號，此情況為 TN，如表 3.4-1。

表 3.4-1 地震波形檢測之二元混淆矩陣

		Real	
		地震訊號	非地震訊號
Predict	地震訊號	TP	FP
	非地震訊號	FN	TN

3.4.4 預測 PGA 是否大於 25gal 之二元混淆矩陣

預測 PGA 是否大於 25gal 之結果可參考圖 3.4-1 中之 P、PGAprob。如果 $Prob_{P\ phase}(t)$ 最大值大於等於 0.3，最大值之時間點為 $t_{\max Prob_{P\ phase}}$ ，然後在 $Prob_{PGAUpOrSmall\ 25gal}(t)$ 從 $t_{\max Prob_{P\ phase}}$ 開始每 0.1 秒，對 $Prob_{PGAUpOrSmall\ 25gal}(t)$ 取平均值，直到 $t_{\max Prob_{P\ phase}}$ 後 3 秒，如果平均值大於等於 0.5 即預測 PGA 大於 25gal，反之為預測 PGA 小於 25gal。若平均值大於 0.5，模型即認為此地震大於 25gal，實際上，此地震也有大於 25gal，即為 TP；若平均值大於 0.5，模型即認為此地震大於 25gal，而實際上，此地震小於 25gal，表示模型失誤認為此地震大於 25gal，即為 FP；若平均值小於 0.5，模型即認為此地震小於 25gal，實際上，此地震小於 25gal，即為 TN；若平均值小於 0.5，模型即認為此地震小於 25gal，實際上，此地震大於 25gal，表示模型漏掉此大於 25gal 之地震，即為 FN，如表 3.4-2。

圖 3.4-3 為考慮容許誤差之二元混淆矩陣，相較於無容許誤差，TP 之數量增加真實 PGA 在 8-25gal 之結果，TN 之數量增加真實 PGA 在 25-80gal 之結果。

表 3.4-2 預測 PGA 是否大於 25gal 之二元混淆矩陣(無容許誤差)

		Real	
		小於 25gal	大於 25gal
Predict	大於 25gal	FP	TP
	小於 25gal	TN	FN

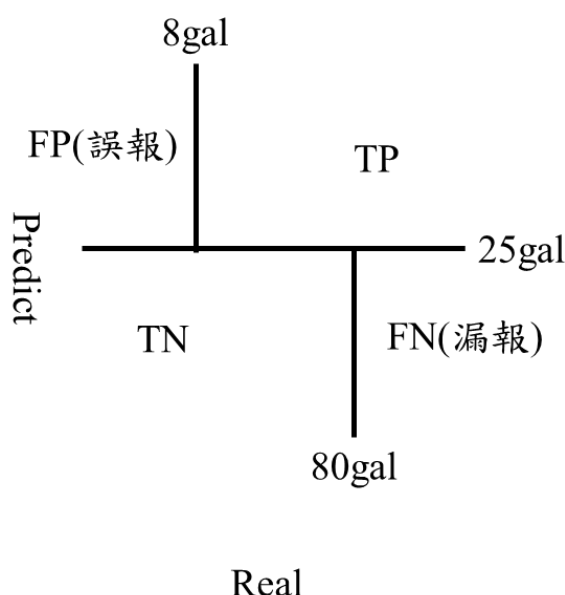


圖 3.4-3 預測 PGA 是否大於 25gal 之二元混淆矩陣(有容許誤差)

3.5 模型測試結果

3.5.1 各輸入對標記 P 波到時、判斷訊號是否為地震訊號、預測 PGA 大於 25gal 在每秒之 precision、recall、F1-score 和針對非地震訊號誤標記率

此小節是模擬各測站觸發後每秒，模型在不同輸入下，標記 P 波到時、判斷訊號是否為地震訊號、預測 PGA 大於 25gal 之準確率、召回率、F1-score，各測站觸發之定義為 Ground Truth 之 P 波到時。各測站觸發後三秒，每秒模型都會輸出三種結果，(一)P 波到時、(二)、判斷訊號是否為地震訊號、(三)、預測此波形資料之 PGA 是否大於 25gal。因為是事後分析，所以有人工標記之 P 波到時，此訊號是地震訊號還是非地震訊號、真實 PGA 是否大於 25gal，並分別參考圖 3.4-2、表 3.4-1、表 3.4-2 得知 TP、TN、FP、FN，並計算出準確率、召回率、F1-score。

圖 3.5-1 到圖 3.5-6 之橫軸 t_w 為 Ground Truth 後幾秒，像是 $t_w = 1.0s$ 表示 Ground Truth 後 1.0 秒之意思，縱軸為準確度(Precision)、召回率(Recall)、F1-score 之百分比，九個圖例分別為 Z-score 為對三軸向加速度歷時作標準分數(Z-score)，用紅色空心圓圈表示、Z-score & 0.8gal 為對三軸向加速度歷時作標準分數(Z-score)搭配對三軸向加速度歷時作正負 0.8gal 間之正規化，用青藍色空心正方形表示、Z-score & 2.5gal 為對三軸向加速度歷時作標準分數(Z-score)搭配對三軸向加速度歷時作正負 2.5gal 間之正規化，用藍色空心菱形表示、Z-score & 8gal 為對三軸向加速度歷時作標準分數(Z-score)搭配對三軸向加速度歷時作正負 8gal 間之正規化，用綠色十字表示、Z-score & 25gal 為對三軸向加速度歷時作標準分數(Z-score)搭配對三軸向加速度歷時作正負 25gal 間之正規化，用粉紅色空心上三角表示、Z-score & 80gal 為對三軸向加速度歷時作標準分數(Z-score)搭配對三軸向加速度歷時作正負 80gal 間之正規化，用淺棕色空心下三角表示、Z-score & 250gal 為對三軸向加速度歷時作標準分數(Z-score)搭配對三軸向加速度歷時作正負 250gal 間之正規化。

化，用黑色空心左三角表示、Z-score&800gal 為對三軸向加速度歷時作標準分數(Z-score)搭配對三軸向加速度歷時作正負 800gal 間之正規化，用棕色空心右三角表示、Z-score&OriginAcc 為對三軸向加速度歷時作標準分數(Z-score)搭配不作正規化之原始加速度歷時，用紫色空心五角表示、OriginAcc 為不做標準分數(Z-score)和不對加速度歷時尺度做正規化之三軸原始加速度歷時。

圖 3.5-1 到圖 3.5-6 之意義為將 Ground Truth 後每秒之波形資料輸入各模型，直到 Ground Truth 後 3 秒，並對其預測此波形資料之 PGA 是否大於 25gal，所以 Ground Truth 後每秒都會有 PGA 是否大於 25gal 的機率值，若此機率值大於 0.5，表示模型預測 PGA 大於 25gal，反之小於 25gal。

得知模型預測 PGA 是否大於 25gal 即可用表 3.4-2 得知 TP(True Positive, 真陽性)、TN(True Negative, 真陰性)、FP(False Positive, 偽陽性)、FN(False Negative, 偽陰性)，有以上數值，即可根據(3.4-1)、(3.4-2)、(3.4-3)得到準確度(Precision)、召回率(Recall)、F1-score。而準確度越小，表示模型誤報越多，召回率越小，表示模型漏報越多。

圖 3.5-1 到圖 3.5-3 為無容許誤差的情況下，在 Ground Truth 後 1 秒到 3 秒間，不同輸入對於預測 PGA 是否大於 25gal 之準確度(Precision)、召回率(Recall)、F1-score。圖 3.5-4 到圖 3.5-6 為有容許誤差的情況下，在 Ground Truth 後 1 秒到 3 秒間，不同輸入對於預測 PGA 是否大於 25gal 之準確度(Precision)、召回率(Recall)、F1-score。因為氣象署希望誤報越少越好，表示準確度(Precision)越高越好，而本計畫認為同時考慮準確度和召回率之 F1-score 也很重要，因此本計畫主要根據圖 3.5-1、圖 3.5-3、圖 3.5-4、圖 3.5-6 選擇在 Ground Truth 後 1 秒到 3 秒間百分比前五高的輸入，以下進一步說明。

無容許誤差的情況下，不同輸入對於預測 PGA 是否大於 25gal 之準確度(Precision)，如圖 3.5-1，在 Ground Truth 後 1 秒到 3 秒間，百分比前五高的輸入分別為三軸原始加速度歷時(OriginAcc)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 25gal 間之正規化(Z-score&25gal)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 2.5gal 間之正規化(Z-score&2.5gal)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 0.8gal 間之正規化(Z-score&0.8gal)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 250gal 間之正規化(Z-score&250gal)。

無容許誤差的情況下，不同輸入對於預測 PGA 是否大於 25gal 之 F1-score，如圖 3.5-3，在 Ground Truth 後 1 秒到 3 秒間，百分比前五高的輸入分別為三軸原始加速度歷時(OriginAcc)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 25gal 間之正規化(Z-score&25gal)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 2.5gal 間之正規化(Z-score&2.5gal)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 8gal 間之正規化(Z-score&8gal)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 0.8gal 間之正規化(Z-score&0.8gal)。

有容許誤差的情況下，不同輸入對於預測 PGA 是否大於 25gal 之準確度(Precision)，如圖 3.5-4，在 Ground Truth 後 1 秒到 3 秒間，百分比前五高的輸入分別為三軸原始加速度歷時(OriginAcc)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 25gal 間之正規化(Z-score&25gal)、三軸向加速度歷時作 Z-score

搭配對三軸向加速度歷時作正負 2.5gal 間之正規化(Z-score&2.5gal)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 0.8gal 間之正規化(Z-score&0.8gal)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 8gal 間之正規化(Z-score&8gal)。

有容許誤差的情況下，不同輸入對於預測 PGA 是否大於 25gal 之 F1-score，如圖 3.5-3，在 Ground Truth 後 1 秒到 3 秒間，百分比前五高的輸入分別為三軸原始加速度歷時(OriginAcc)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 25gal 間之正規化(Z-score&25gal)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 2.5gal 間之正規化(Z-score&2.5gal)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 8gal 間之正規化(Z-score&8gal)、三軸向加速度歷時作 Z-score 搭配對三軸向加速度歷時作正負 0.8gal 間之正規化(Z-score&0.8gal)。

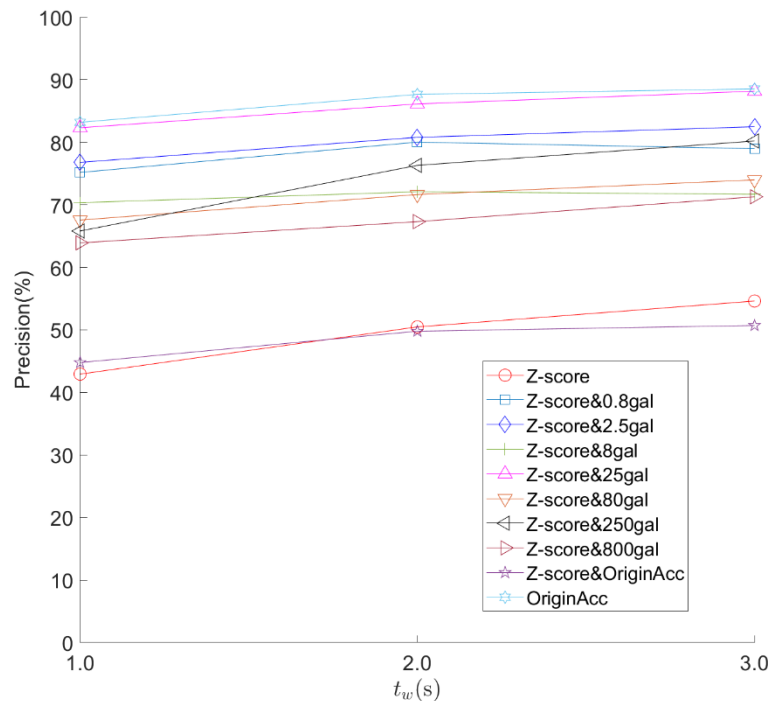


圖 3.5-1 預測 PGA 是否大於 25gal 之準確率(無容許誤差)

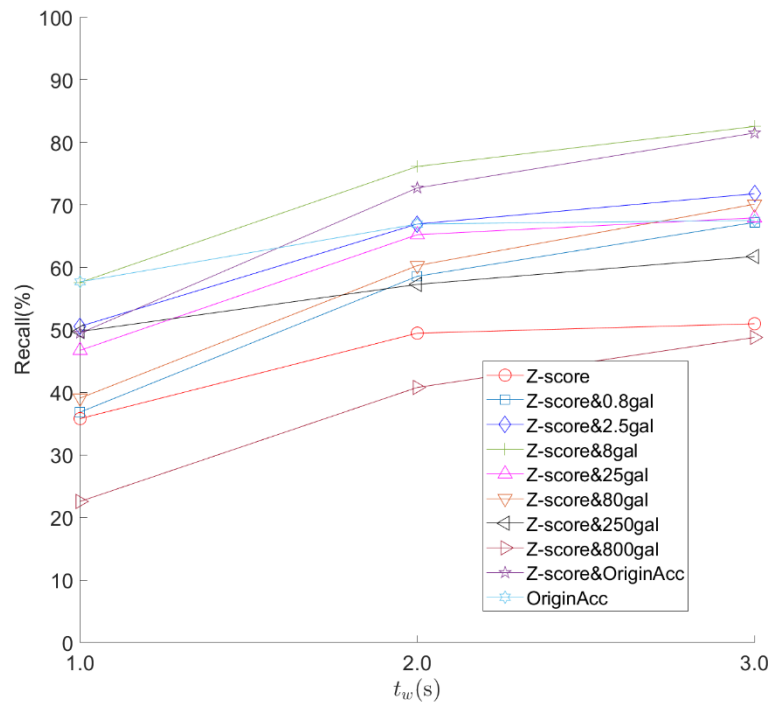


圖 3.5-2 預測 PGA 是否大於 25gal 之召回率(無容許誤差)

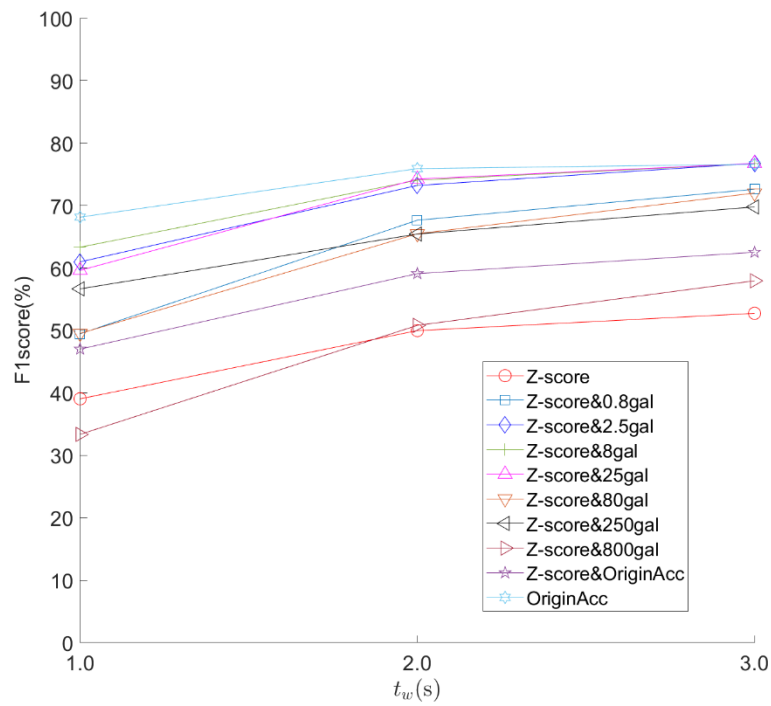


圖 3.5-3 預測 PGA 是否大於 25gal 之 F1-score(無容許誤差)

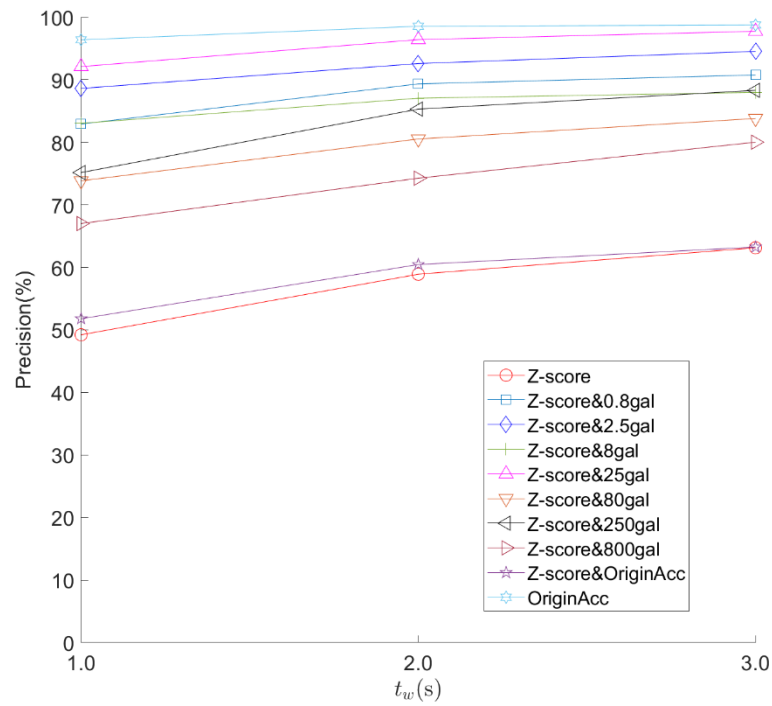


圖 3.5-4 預測 PGA 是否大於 25gal 之準確率(有容許誤差)

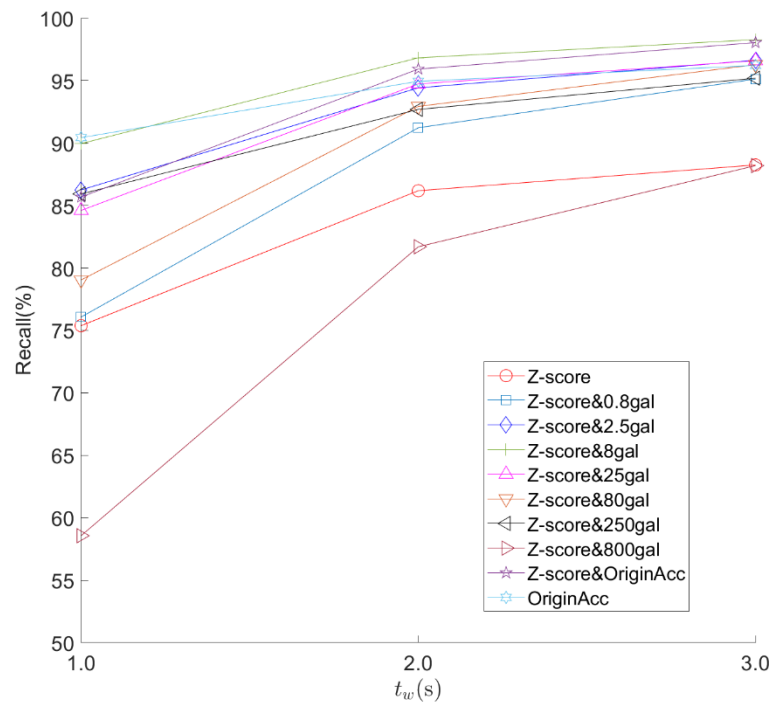


圖 3.5-5 預測 PGA 是否大於 25gal 之召回率(有容許誤差)

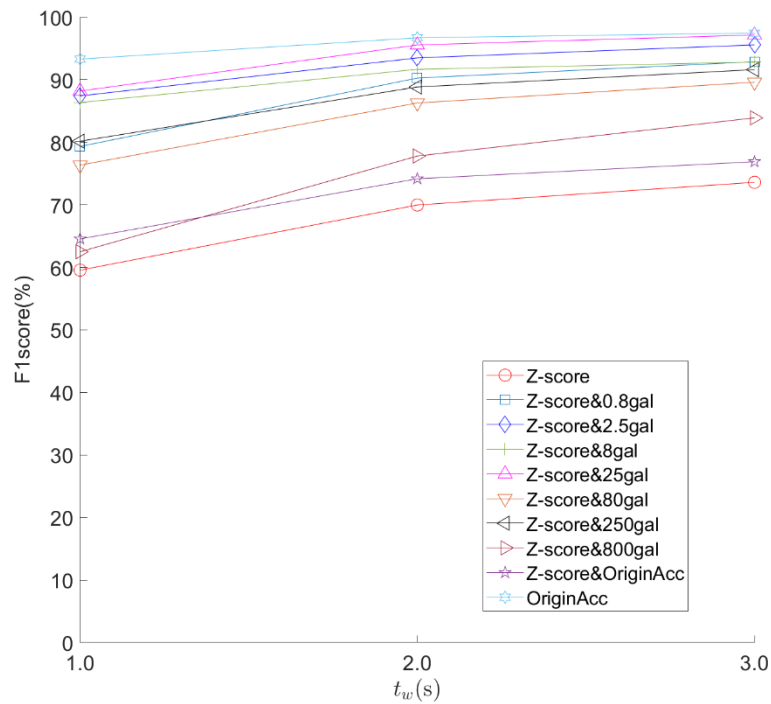


圖 3.5-6 預測 PGA 是否大於 25gal 之 F1-score(有容許誤差)

圖 3.5-7 到圖 3.5-9 分別為不同輸入對於模型判斷訊號是否為地震之準確率、召回率、F1-score。觀察圖 3.5-8 可知，輸入三軸原始加速度歷時(OriginAcc)容易讓模型把地震訊號判斷為非地震訊號。圖 3.5-10 到圖 3.5-12 分別為模型標記 P 波到時之準確率、召回率、F1-score。觀察圖 3.5-11 可知，輸入三軸原始加速度歷時(OriginAcc)容易讓模型漏標記 P 波到時，也就是說模型標記時間在 Ground Truth 正負 0.5 秒內，但機率值小於門檻值。因此三軸原始加速度歷時(OriginAcc)並不適合。

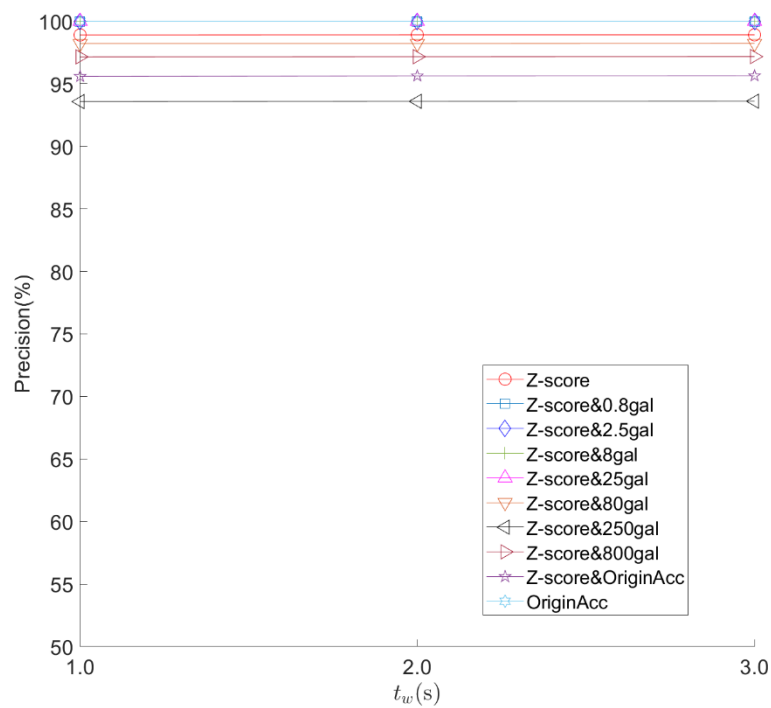


圖 3.5-7 訊號是否為地震之準確率

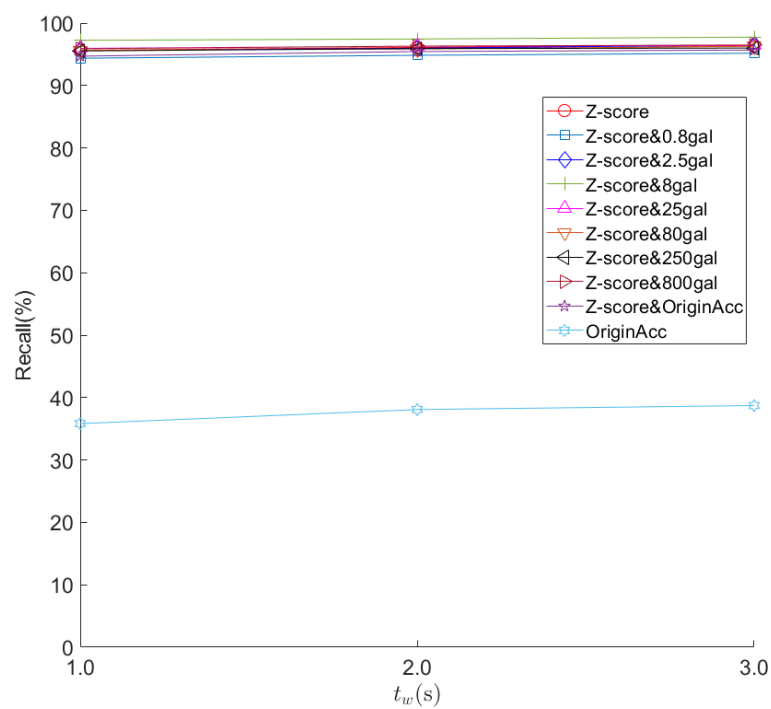


圖 3.5-8 訊號是否為地震之召回率

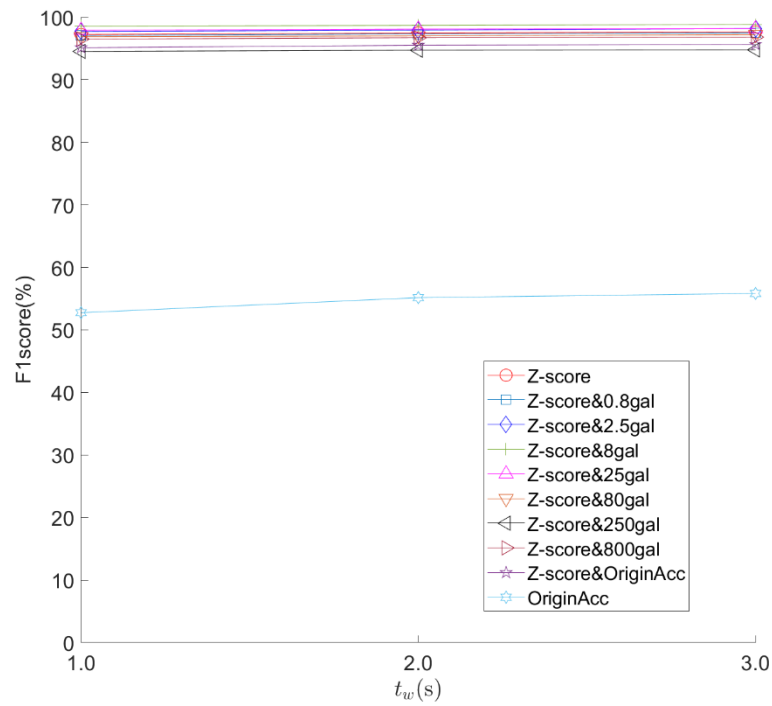


圖 3.5-9 訊號是否為地震之 F1-score

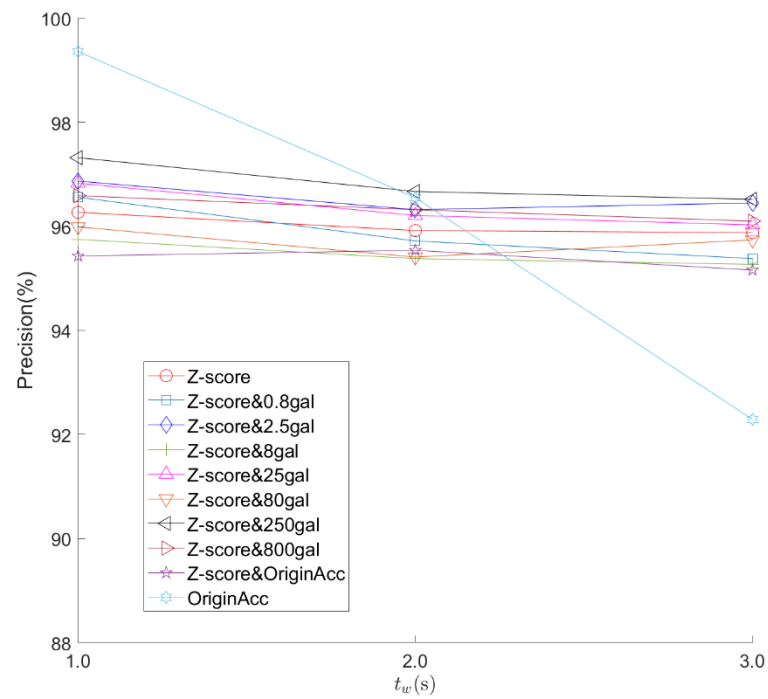


圖 3.5-10 標記 P 波到時之準確率

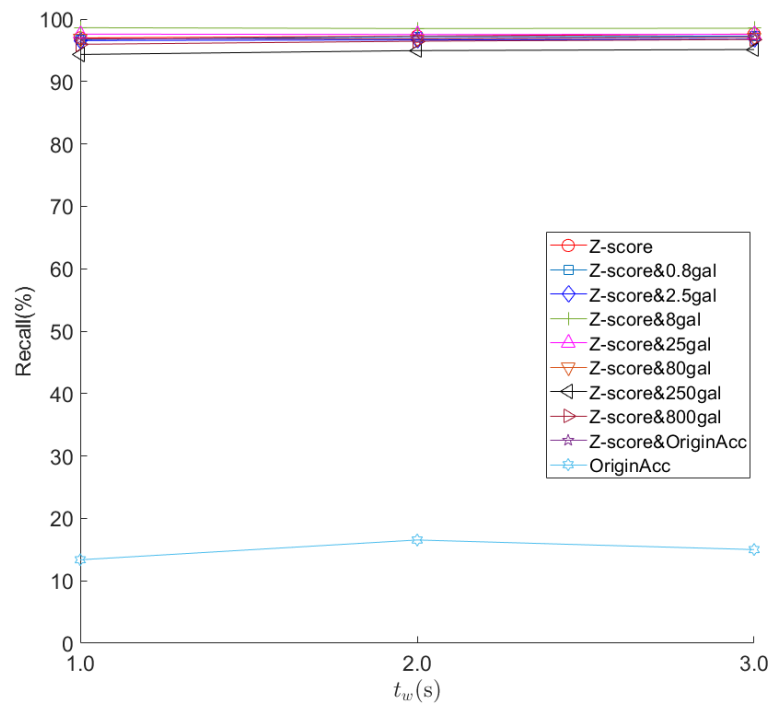


圖 3.5-11 標記 P 波到時之召回率

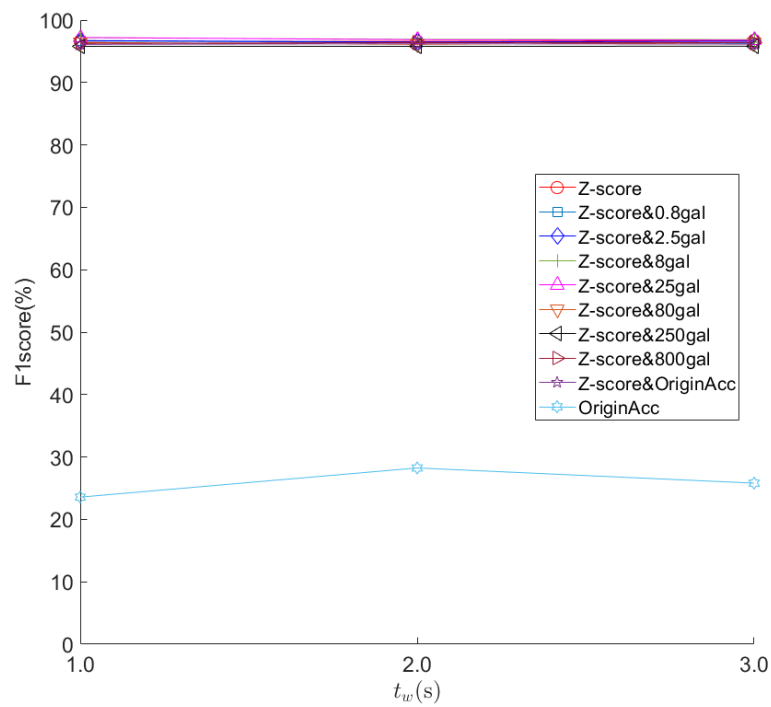


圖 3.5-12 標記 P 波到時之 F1-score

據圖 3.5-1、圖 3.5-3、圖 3.5-4、圖 3.5-6 選擇 OriginAcc、Z-score&25gal、Z-score&2.5gal、Z-score&8gal、Z-score&0.8gal。但據圖 3.5-8、圖 3.5-11，本計畫

認為不能選 OriginAcc，因此不將其列入考慮。

圖 3.5-13 為在不同輸入的情況下，模型對非地震訊號誤標記之錯誤率，非地震訊號來自 CWASN，共 23260 筆。觀察圖 3.5-13 中的 Z-score&25gal、Z-score&2.5gal、Z-score&8gal、Z-score&0.8gal 可知，對三軸向波形資料做 Z-score&8gal 較其他三個輸入容易讓模型對 noise 誤標記，所以本計畫將以下輸入 Z-score&25gal、Z-score&2.5gal、Z-score&0.8gal 繼續比較。

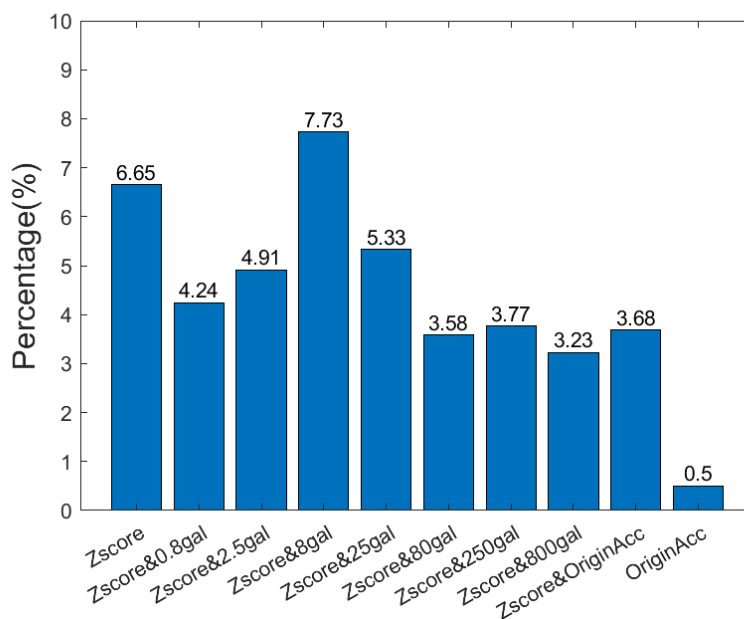


圖 3.5-13 各輸入對非地震訊號誤標記之錯誤率

3.5.2 探討三種輸入在現地型強震預警之表現

三種輸入對於小地震之誤報率

此小節展示三種輸入在 Ground Truth 後 1 秒到 3 秒之情況下，模型對小地震之誤報率，亦即模型對 PGA 小於 25 之加速度歷時預測其 PGA 大於 25gal。小地震定義為 PGA 小於 25gal 之加速度歷時，誤報率計算方式如(3.5-1)，FP、TN 之定義參考表 3.4-2。

$$\frac{FP}{FP + TN} \quad (3.5-1)$$

圖 3.5-14 為在三種輸入之情況下，模型對於 PGA 在 0-8gal 間之誤報率。觀察此圖可知三種輸入隨著每秒波形資料增長，誤報率都有下降。其中 Z-score&25gal 在 Ground Truth 後 1 秒到 3 秒間之誤報率，都小於其餘兩種輸入。三種輸入在 Ground Truth 後 1 秒到 3 秒間之誤報率都小於 1.8%

圖 3.5-15 為在三種輸入之情況下，模型對於 PGA 在 8-25gal 間之誤報率。觀察此圖可知 Z-score&2.5gal、Z-score&0.8gal 隨著每秒波形資料增長，誤報率都上升。而 Z-score&25gal 在 Ground Truth 後 3 秒的誤報率小於其 Ground Truth 後 2 秒的誤報率。三種輸入對 PGA 在 8 到 25gal 間之波形資料之誤報率，都大於 PGA 在 0 到 8gal 間之波形資料。

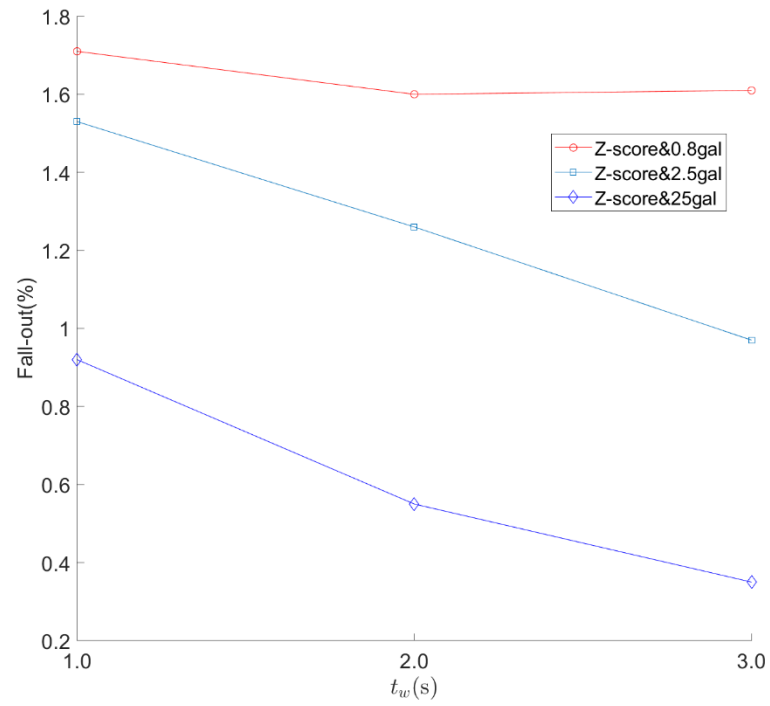


圖 3.5-14 不同輸入對於 PGA 在 0-8gal 間之誤報率

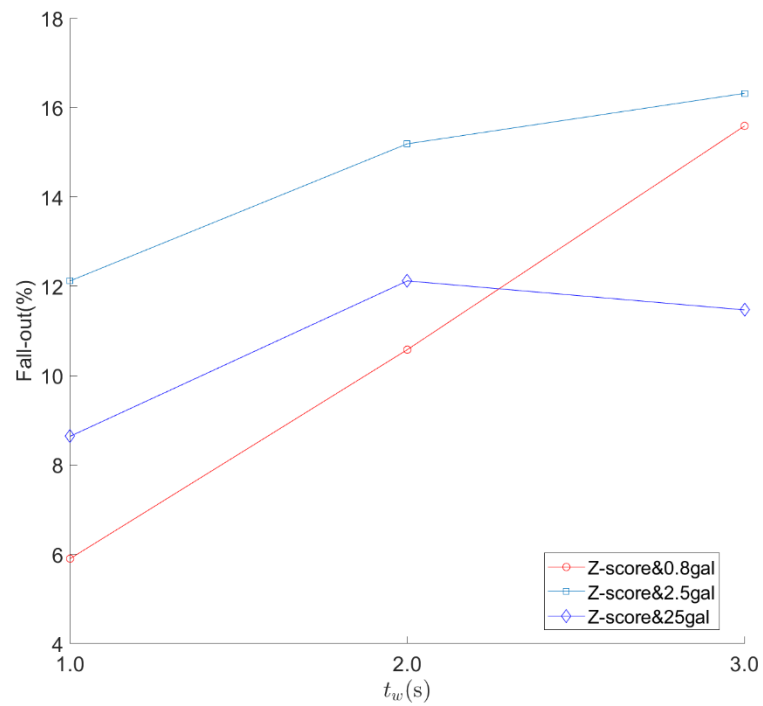


圖 3.5-15 不同輸入對於 PGA 在 8-25gal 間之誤報率

三種輸入對於大地震之漏報率

此小節展示三種輸入在 Ground Truth 後 1 秒到 3 秒之情況下，模型對大地震的漏報率，即為模型對 PGA 大於 25 之加速度歷時預測其 PGA 小於 25gal，大地震之定義為 PGA 大於 25gal 以上之加速度歷時。漏報率計算公式如(3.5-2)，TP、FN 之定義參考表 3.4-2。

$$\frac{FN}{TP + FN} \quad (3.5-2)$$

圖 3.5-16 為在三種輸入之情況下，模型對於 PGA 在 25-80gal 間之漏報率。觀察此圖可知三種輸入隨著波形資料越長，漏報率都有下降，但最低之漏報率為 Z-score&2.5gal 在 Ground Truth 後 3 秒之 35%，表示將 Z-score&2.5gal 當作模型輸入會對 PGA 在 25-80gal 間之加速度歷時漏報 35%。Z-score&2.5gal 在 Ground Truth 後 1 秒到 3 秒間之漏報率都小於其餘兩種輸入。

圖 3.5-17 為在三種輸入之情況下，模型對於 PGA 在 80-250gal 間之漏報率。觀察此圖可知三種輸入隨著波形資料越長，漏報率都有下降，但最低之漏報率為 Z-score&25gal 在 Ground Truth 後 3 秒之 10%左右。Z-score&2.5gal 和 Z-score&25gal 在 Ground Truth 後 1 秒到 3 秒間之漏報率表現差不多。

圖 3.5-18 為在三種輸入之情況下，模型對於 PGA 在 250-800gal 間之漏報率。觀察此圖可知三種輸入隨著波形資料越長，漏報率都有下降，但最低之漏報率為 Z-score&25gal 在 Ground Truth 後 3 秒之 5%左右。Z-score&25gal 在 Ground Truth 後 1 秒到 3 秒間之漏報率都小於其餘兩種輸入。

圖 3.5-19 為在三種輸入之情況下，模型對於 PGA 在 800gal 以上之漏報率。觀察此圖可知 Z-score&2.5gal、Z-score&25gal 在 Ground Truth 後 1 秒到 3 秒間之漏報率都為零。而 Z-score&0.8gal 在 Ground Truth 後 2 秒到 3 秒之漏報率都在 16% 左右。

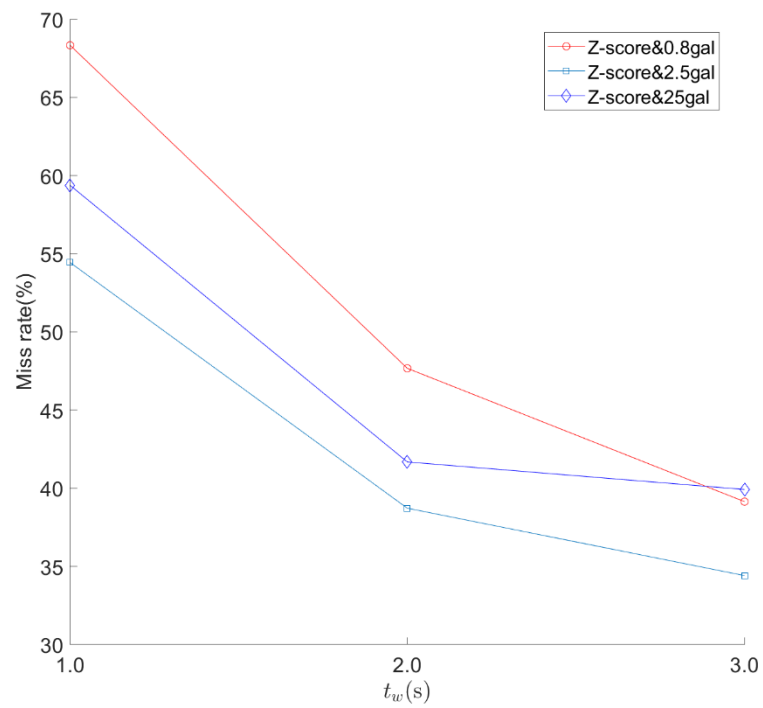


圖 3.5-16 不同輸入對於 PGA 在 25-80gal 間之漏報率

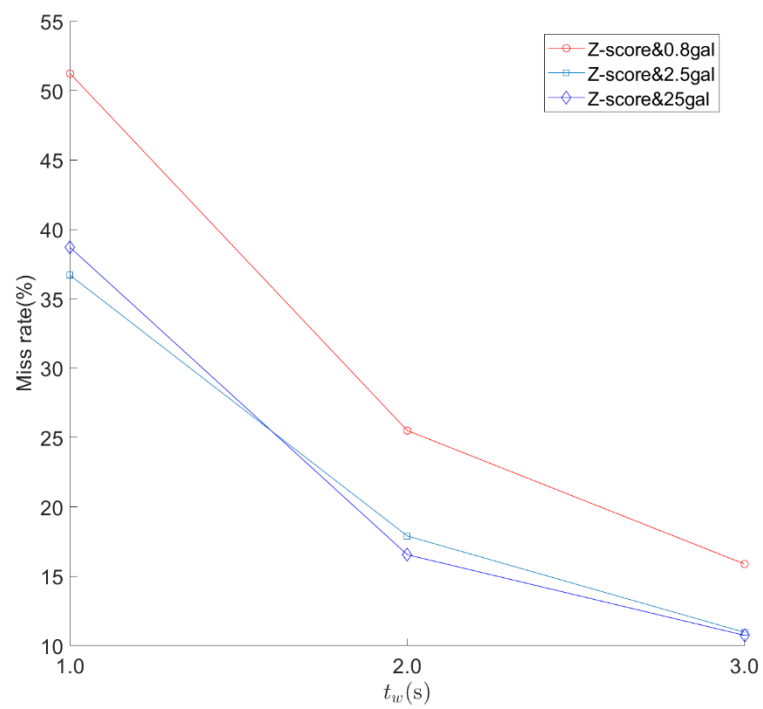


圖 3.5-17 不同輸入對於 PGA 在 80-250gal 間之漏報率

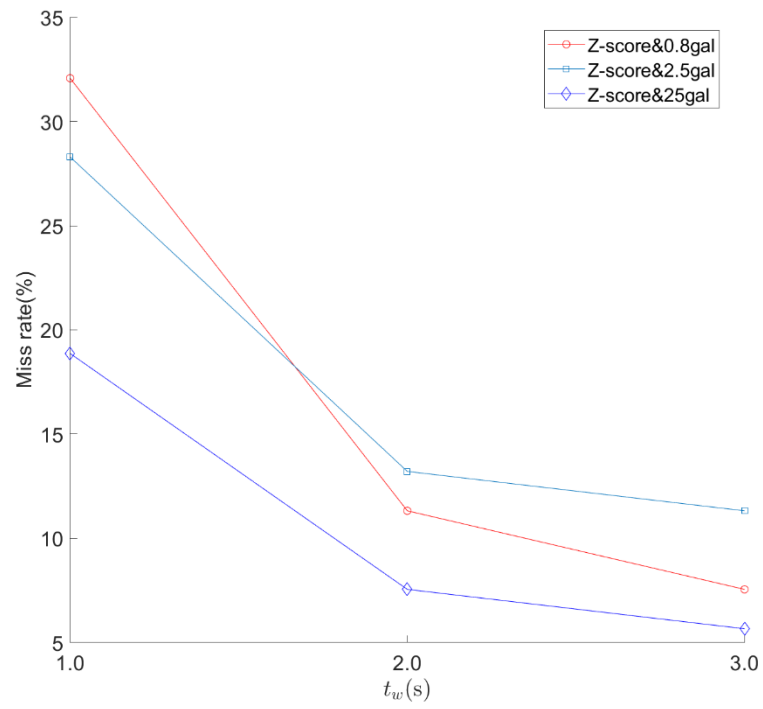


圖 3.5-18 不同輸入對於 PGA 在 250-800gal 間之漏報率

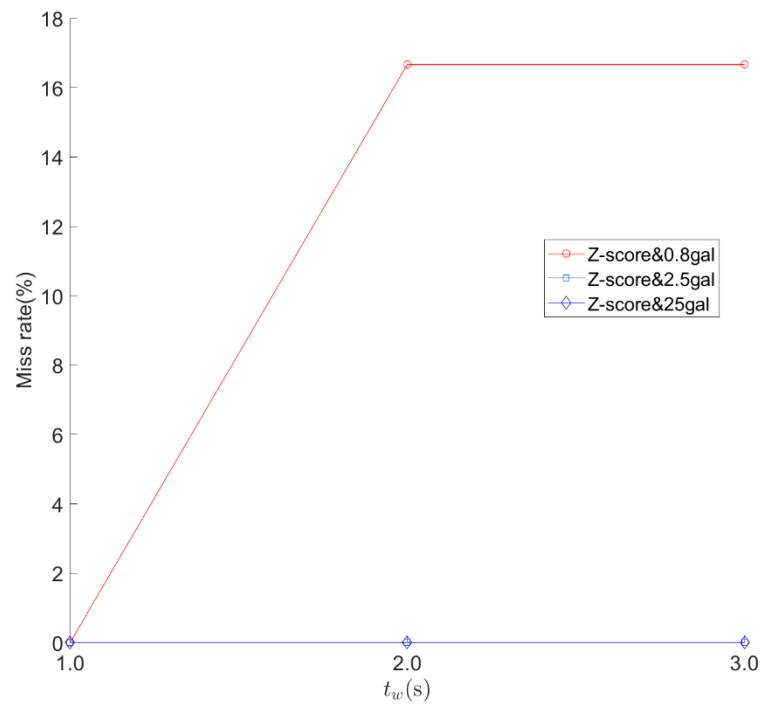


圖 3.5-19 不同輸入對於 PGA 在 800gal 以上之漏報率

三種輸入對於非地震訊號之誤報率

圖 3.5-20 為在不同輸入的情況下，模型對非地震訊號預測 PGA 大於 25gal 之誤報率。非地震訊號來自 CWASN，共 23260 筆。觀察此圖可知在三種輸入中，Z-score&25gal 之誤報率最小，為 0.18%。

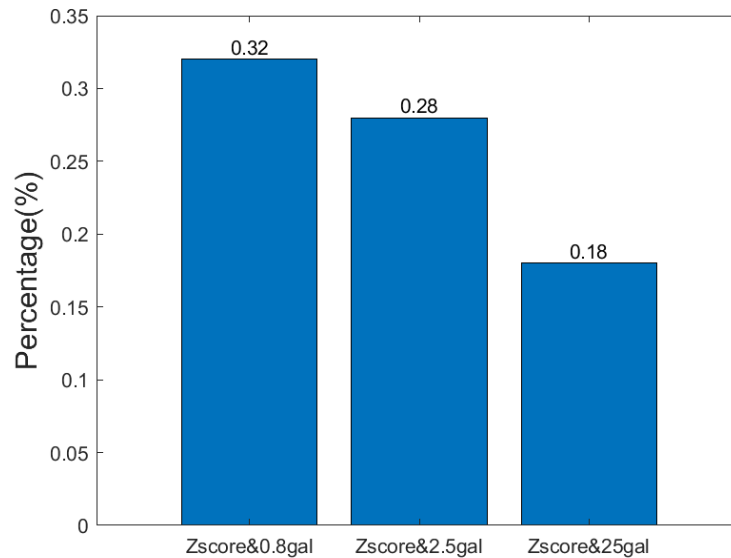


圖 3.5-20 不同輸入對非地震訊號預測 PGA 大於 25gal 之誤報率

評估三種輸入在 Ground Truth 後每秒發報情形

此小節是模擬每個測站在各測站觸發後三秒每秒發報之情形，而各測站觸發即為 Ground Truth 之時間點。各測站觸發後三秒每秒都會有一個預測結果，此預測結果為此波形資料之 PGA 是否大於 25gal，如果預測 PGA 大於 25gal 立即對此測站發出警報。因為警報發出無法收回，所以某秒之後此測站就一直是發報狀態。舉例，如果某測站在 Ground Truth 後 1 秒被預測 PGA 大於 25gal，某測站在 Ground Truth 後 1 秒到 3 秒間都算發報。因為是事後分析，所以每筆波形資料 PGA 都已知，故可以參考表 3.4-2 得知 TP、TN、FP、FN，進而計算出準確率、召回率、F1-score。

圖 3.5-21 為無容許誤差下，三種輸入在 Ground Truth 後每秒發報之準確率。觀察此圖可知 Z-score&0.8gal、Z-score&25gal 隨著波形資料增長，準確率也隨著上升，Z-score&25gal 在 Ground Truth 後 1 秒到 3 秒間之準確率高於 Z-score&0.8gal。Z-score&2.5gal 之準確率維持定值。

圖 3.5-22 為無容許誤差下，三種輸入在 Ground Truth 後每秒發報之召回率。觀察此圖可知 Z-score&0.8gal、Z-score&25gal 隨著波形資料增長，召回率也隨著下降，但 Z-score&2.5gal 之召回率維持定值。

圖 3.5-23 為無容許誤差下，三種輸入在 Ground Truth 後每秒發報之 F1-score。觀察此圖可知 Z-score&0.8gal 隨著波形資料增長，F1-score 也隨著上升，但 Z-score&2.5gal 之 F1-score 維持定值。

圖 3.5-24 為有容許誤差下，三種輸入在 Ground Truth 後每秒發報之準確率。

觀察此圖可知 Z-score&0.8gal、Z-score&25gal 隨著波形資料增長，準確率也隨著上升，Z-score&25gal 在 Ground Truth 後 1 秒到 3 秒間之準確率高於 Z-score&0.8gal，其準確率最高可達 98% 左右。Z-score&2.5gal 之準確率維持定值。

圖 3.5-25 為有容許誤差下，三種輸入在 Ground Truth 後每秒發報之召回率。觀察此圖可知 Z-score&0.8gal、Z-score&25gal 隨著波形資料增長，召回率也隨著下降，但 Z-score&2.5gal 之召回率維持定值。

圖 3.5-26 為有容許誤差下，三種輸入在 Ground Truth 後每秒發報之 F1-score。觀察此圖可知 Z-score&0.8gal、Z-score&25gal 隨著波形資料增長，F1-score 也隨著上升，Z-score&25gal 在 Ground Truth 後 1 秒到 3 秒間之 F1-score 高於 Z-score&0.8gal，其 F1-score 最高可達 97% 左右。Z-score&2.5gal 之 F1-score 維持定值。

根據上述之探討，發現在三種輸入對於小地震之誤報率、三種輸入對於大地震之漏報率、三種輸入對於非地震訊號之誤報率、評估三種輸入在 Ground Truth 後每秒發報情形等評估方式中，Z-score&25gal 表現比其餘兩種輸入效果好。

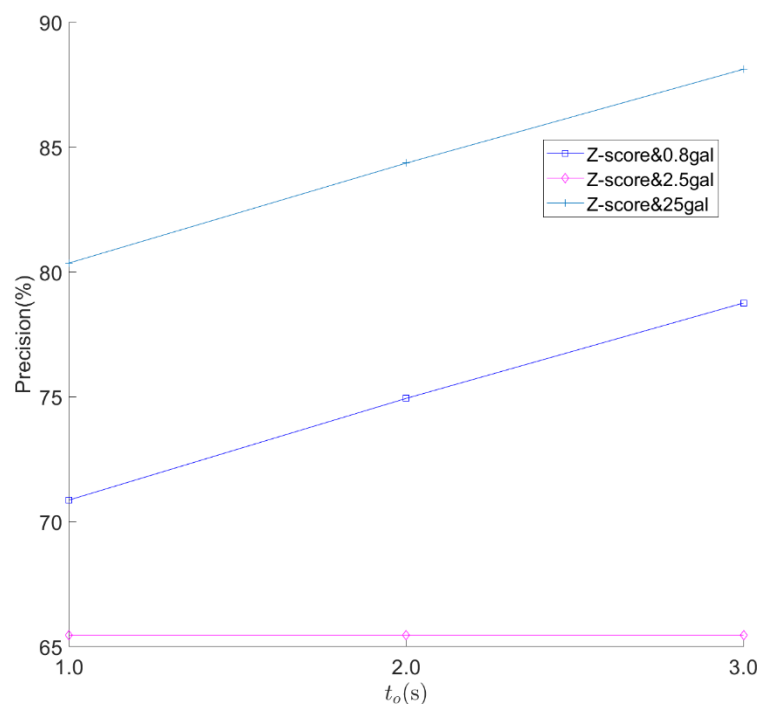


圖 3.5-21 三種輸入在 Ground Truth 後每秒發報之 Precision(無容許誤差)

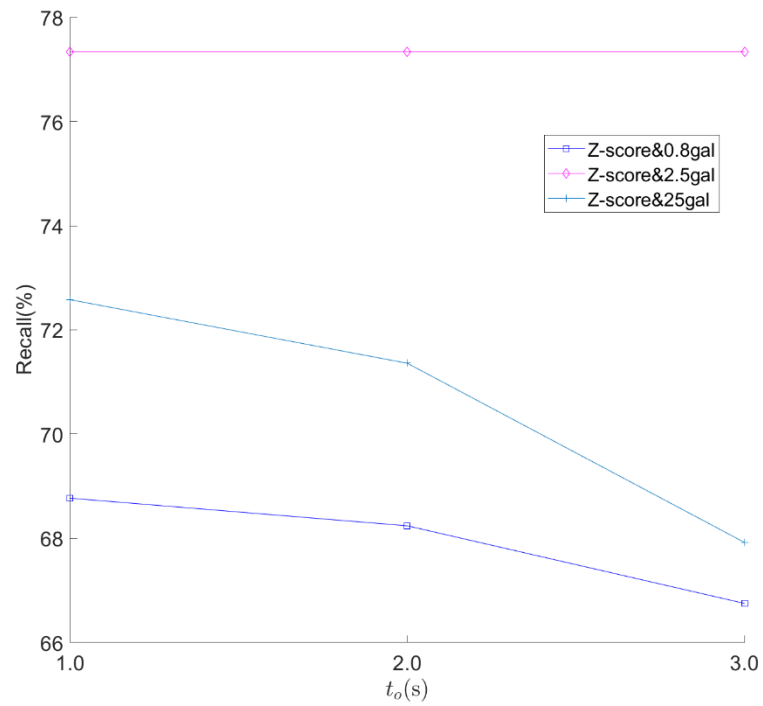


圖 3.5-22 三種輸入在 Ground Truth 後每秒發報之 Recall(無容許誤差)

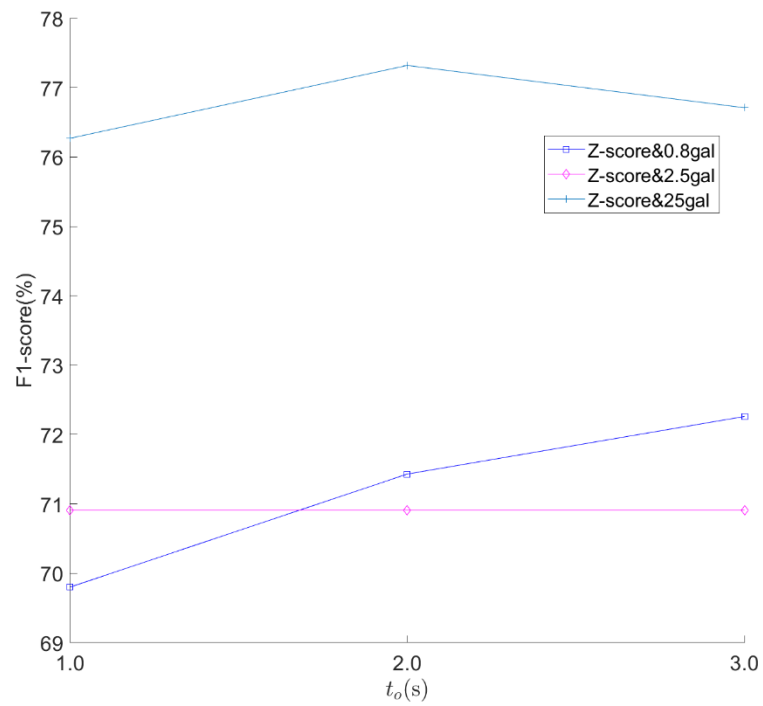


圖 3.5-23 三種輸入在 Ground Truth 後每秒發報之 F1-score(無容許誤差)

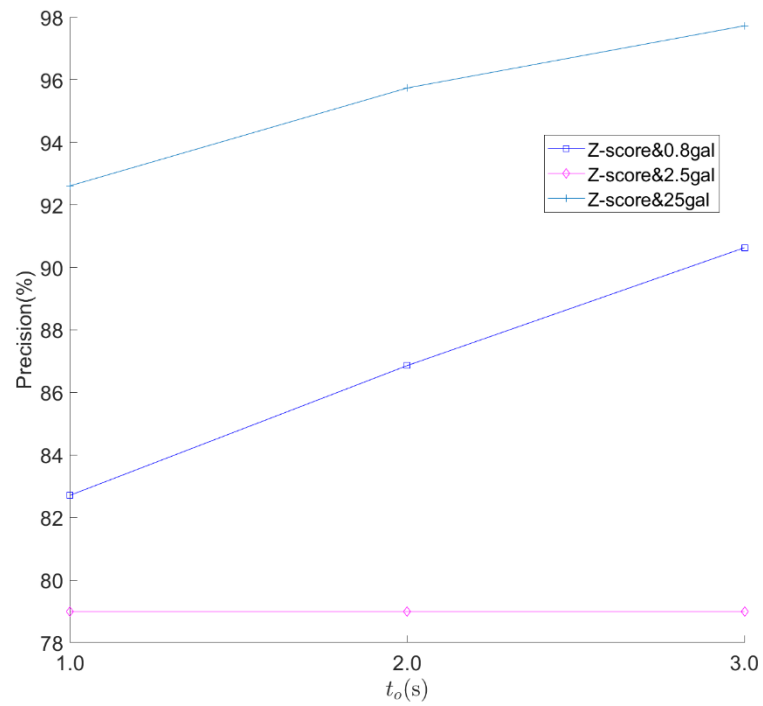


圖 3.5-24 三種輸入在 Ground Truth 後每秒發報之 Precision(有容許誤差)

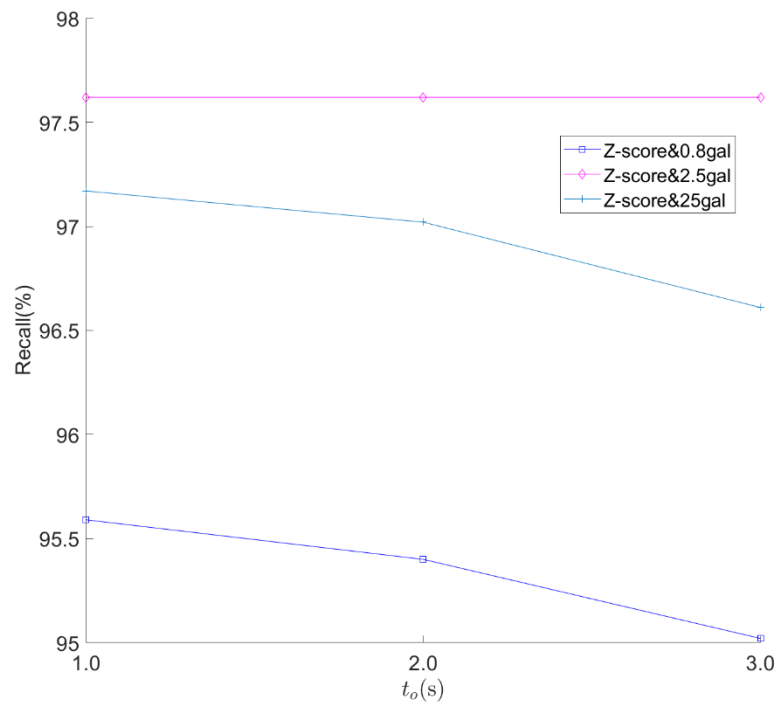


圖 3.5-25 三種輸入在 Ground Truth 後每秒發報之 Recall(有容許誤差)

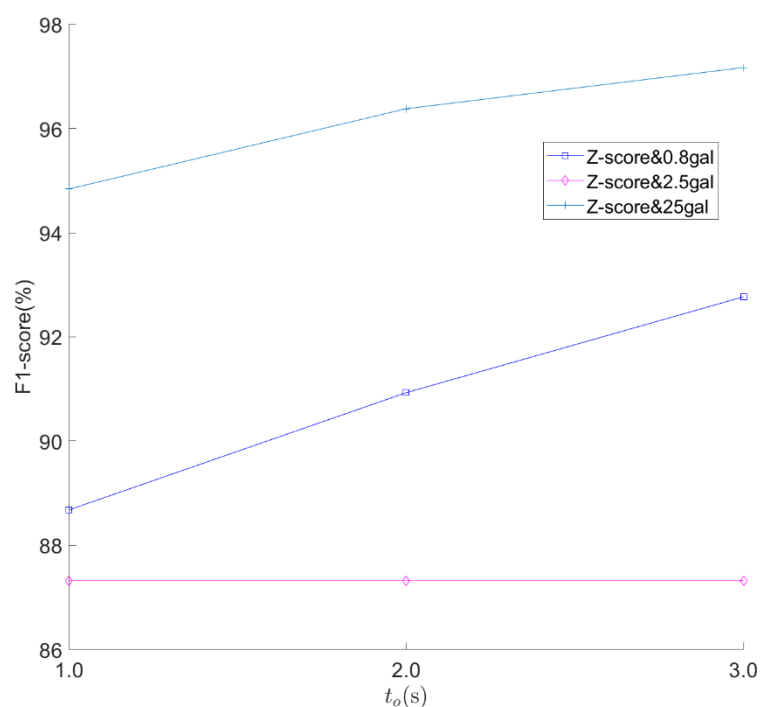


圖 3.5-26 三種輸入在 Ground Truth 後每秒發報之 F1-score(有容許誤差)

3.5.3 與 RED-PAN 和柏鈞模型比較

跟 RED-PAN[1]比較之資料集為 2019 年中央氣象署地震觀測網(Central Weather Administration Seismographic Network, CWASN)，共 122356 筆，此資料集為 Liao, et al. [1]提供，每筆波形資料長度都為 30 秒訊號，且 P 波到時和 S 波到時都含在 30 秒內，取樣頻率為 100Hz，但本報告之模型是使用 P 波到時後之訊號預測 PGA 是否大於 25gal，所以只跟 RED-PAN 比較標記 P 波到時之召回率、F1-score、與 Ground Truth 秒差之平均值(Mean)、與 Ground Truth 秒差之標準差(std)。Liao, et al. [1]也提供 RED-PAN[1]針對此資料集標記 P 波到時之召回率、F1-score、與 Ground Truth 秒差之平均值(Mean)、與 Ground Truth 秒差之標準差(std)，結果如表 3.5-1。觀察此表可知 RED-PAN 在各項指標，表現都高於三種輸入。若用綜合準確率和召回率之 F1-score 比較三種輸入可知，Z-score&25gal 高於其餘兩種輸入。若用 Mean 比較三種輸入可知，Z-score&2.5gal 高於其餘兩種輸入。若用 std 比較三種輸入可知，Z-score&25gal 高於其餘兩種輸入。

表 3.5-1 與 RED-PAN 比較標記 P 波到時之比較結果

模型	準確率 (unit : %)	召回率 (unit : %)	F1-score (unit : %)	Mean (unit : s)	std(s) (unit : s)
RED-PAN	<u>98.89</u>	<u>98.77</u>	<u>98.83</u>	<u>0.03</u>	<u>0.07</u>
Z-	96.76	96.54	96.65	0.09	0.09

score&0.8gal					
Z-score&2.5gal	97.28	97.16	97.22	0.05	0.12
Z-score&25gal	96.98	97.94	97.46	0.09	0.08

跟柏鈞模型[2]比較之資料集為模型測試資料集中臺灣強地動觀測網(Taiwan Strong Motion Instrument Program Network, TSMIP)，共 4056 筆，輸入本報告之模型每筆波形資料長度都為 30 秒訊號，其含 Ground Truth 後 2 秒之訊號，且 Ground Truth 後 2 秒之訊號是在 30 秒訊號中最後兩秒，其餘為背景訊號，取樣頻率也為 100Hz。而輸入進柏鈞模型之訊號為 P 波到時後 2 秒之波形資料經過時頻域轉換[2]後之訊號，取 P 波到時之方式為使用 STA/LTA 對 Ground Truth 後之訊號取 P 波到時。柏鈞模型之準確率、召回率、F1-score 表現會較本報告之模型好之原因可能為柏鈞模型之輸入有考慮時間域和頻率域，但本報告之模型只有考慮時間域，所以未來會嘗試加入頻率域的資訊，像是對時間域做 STFT(短時距傅立葉轉換, Short-time Fourier Transform)之訊號等。無容許誤差結果如表 3.5-2、有容許誤差結果如表 3.5-3。觀察表 3.5-2 可知柏鈞模型之 F1-score、召回率與準確率高於三種輸入。觀察表 3.5-3 也有此現象。

表 3.5-2 與柏鈞模型比較預測 PGA 大於 25gal 之結果(無容許誤差)

模型	準 確 率 (unit : %)	召 回 率 (unit : %)	F1- score (unit : %)
柏鈞模型	<u>84.75</u>	<u>69.04</u>	<u>76.09</u>
Z-score&0.8gal	79.03	36.57	50.00
Z-score&2.5gal	77.98	43.49	55.84
Z-score&25gal	78.71	50.41	61.46

表 3.5-3 與柏鈞模型比較預測 PGA 大於 25gal 之結果(有容許誤差)

模型	準 確 率 (unit : %)	召 回 率 (unit : %)	F1- score (unit : %)
柏鈞模型	<u>99.33</u>	<u>95.82</u>	<u>97.55</u>
Z-score&0.8gal	96.92	76.33	85.40
Z-	97.57	81.17	88.62

score&2.5gal			
Z-score&25gal	97.46	86.22	91.50

根據表 3.5-1、表 3.5-2、表 3.5-3 之結果顯示，本報告所提出之模型可同時進行標記地震波相和預測 PGA 是否大於 25gal 此兩項任務。本報告所提出之模型對於標記地震波相之表現略遜於 RED-PAN[1]，但對於預測 PGA 是否大於 25gal 之表現與柏鈞模型[2]相距盛遠。

3.6 小結

根據 3.5.1、3.5.2 節之結果，在三種輸入中 Z-score&25gal 表現較其餘兩種輸入效果好。而在 3.5.3 節與兩種個別進行標記地震波相和預測 PGA 是否大於 25gal 模型之比較中，再次驗證當模型輸入 Z-score&25gal 時，其表現較其餘兩種輸入為佳。

3.7 模擬 2022 年 9 月 18 號池上地震

2022 年 9 月 18 號池上地震之資料是從中央氣象署臺灣地震與地球物理資料管理系統(Taiwan Seismological and Geophysical Data Management System, GDMS)下載，顯著有感地震內 2022-09-18 06:44:15 之地震，觀測網為臺灣強地動觀測網(Taiwan Strong Motion Instrumentation Program network, TSMIP)，其資料長度為 20 分鐘。本報告從發震時間前 30 秒開始切每段 30 秒之三軸向加速度訊號，直到發震後 2 分鐘。並將各測站每段訊號輸入進模型，模型會對每一段訊號標記 P 波到時、S 波到時、預測 PGA 是否大於 25gal 之機率值。本報告對單站發報條件為：若模型對單站波型資料標記 P 波到時後第二、三個封包，其中一個封包預測 PGA 大於 25gal，此測站就會發報。圖 3.7-1 為在池上地震中，各測站發報結果，圖中藍色實心三角形為 TP(正確發報)、藍色空心三角形為 TN(正確不發報)、紅色實心三角形為 FP(誤報)、紅色空心三角形為 FN(漏報)。觀察此圖可知在近震央地區都是 TP(正確發報)，而雲林、嘉義、臺南、高雄幾乎都為 FN(漏報)，在桃園、新竹、苗栗、臺中有零星幾個測站為 FP(誤報)。在圖 3.7-1 中，共有 103 個測站為 TP(正確發報)、97 個測站為 TN(正確不發報)、19 個測站為 FP(誤報)、145 個測站為 FN(漏報)，準確率為 84.43%、召回率為 41.53%、F1-score 為 55.67%。漏報原因主要可能是因為本次測試僅使用 P 波到時後第二、三個封包進行預測，但有些測站模型對某一測站標記 P 波到時後第三個封包以後之封包才預測其 PGA 大於 25gal，例如若是在標記 P 波到時後第四個封包進來後才預測 PGA 大於 25gal，則會變成漏報的情況。

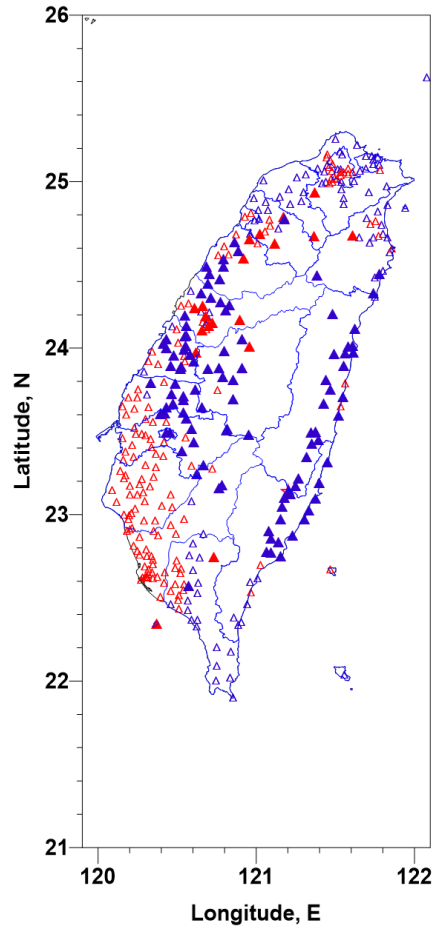


圖 3.7-1 各測站在池上地震發報結果

3.8 未來規劃

未來規劃將模型應用 MQTT 進行測試，其架構圖如圖 3.8-1，IMPORT 為各儀器之波形資料，WAVE_MQTT 為儲存 IMPORT 之資料，Model 為標記 WAVE_MQTT 內之波型資料並預測此波形資料之 PGA 是否大於 25gal，DCSN 為製作地震報告，REPORT 為將地震報告發布給用戶。由於在即時系統上做資料前處理可能會導致模型預測不理想，所以將資料前處理包在模型內。

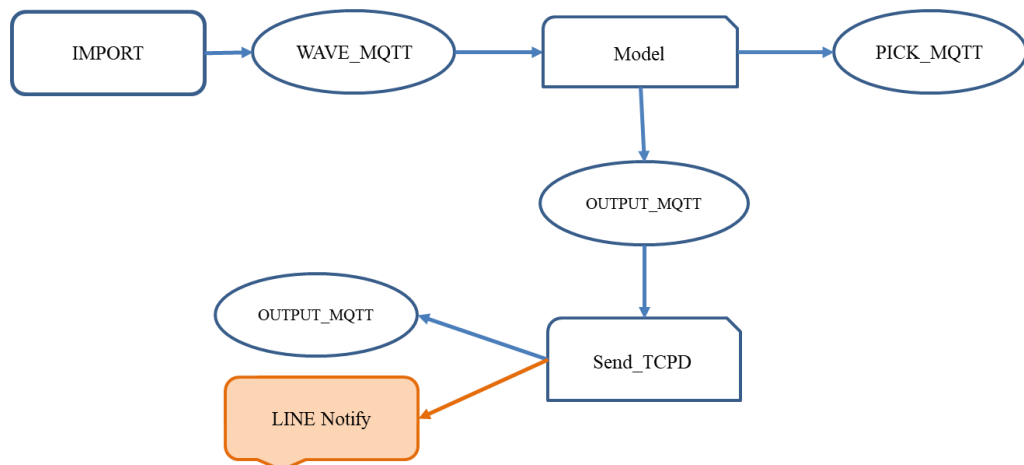


圖 3.8-1 模型應用 MQTT 之架構圖

結論

本研究提出將卷積神經網路(CNN)結合現地型強震預警應用於氣象署之 Earthworm 系統中，希望可以提升近震央區域預警的速度並縮小盲區，同時也要減少誤報的發生率。目前已完成建置實驗室 Earthworm 環境即時模擬，並在實驗室將模型接應程式建構完成，後續希望能在研究室加入 AI picker 進行相關測試與整合。而程式在氣象署運行，初步驗證該模型於實際地震事件預測成效不錯。唯運算能力不符合預期，應還能改善運算效率，包括是否利用 GPU 加速運算的可行性。

本計畫之模型是結合標記地震波相和強震預警之模型，並探討此模型架構在以下 10 種不同輸入之表現，10 種不同輸入包括：(一)、不做標準分數和不對加速度歷時尺度做正規化之三軸原始加速度歷時、(二)、Z-score、(三)、Z-score&0.8gal、(四)、Z-score&2.5gal、(五)、Z-score&8gal、(六)、Z-score&25gal、(七)、Z-score&80gal、(八)、Z-score&250gal、(九)、Z-score&800gal 和(十)、Z-score&原始加速度歷時。所探討之表現為在 Ground Truth 後 1 秒到 3 秒間，對標記 P 波到時、判斷訊號是否為地震訊號、預測 PGA 大於 25gal 每秒之 precision、recall、F1-score 和針對非地震訊號之誤標記率。根據上述各方面之表現，本研究在十種輸入中挑選出表現較佳之 Z-score&0.8gal、Z-score&2.5gal、Z-score&25gal。之後在上述三種輸入的情況下，評估模型對於小地震之誤報率、對於大地震之漏報率、對於非地震訊號之誤報率和 TSMIP、CWASN 測站在 Ground Truth 後每秒發報情形，發現 Z-score&25gal 之表現較佳；且在與兩種個別進行標記地震波相和預測 PGA 是否大於 25gal 模型之比較中，再次驗證當模型輸入為 Z-score&25gal 時，其表現較其餘兩種輸入為佳，因此最後選擇 Z-score&25gal 作為模型之輸入。而將其應用在 2022 年 9 月 18 日之池上地震結果為準確率為 84.43%、召回率為 41.53%、F1-score 為 55.67%。

後續將所選擇之 Z-score&25gal，以離線方式模擬實際地震事件之表現。未來規劃將模型放進中央氣象署 Earthworm 進行測試，以了解線上運作之表現情形。

參考文獻

- [1] W.-Y. Liao, E.-J. Lee, D.-Y. Chen, P. Chen, D. Mu, and Y.-M. Wu, "RED-PAN: Real-Time Earthquake Detection and Phase-Picking With Multitask Attention Network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-11, 2022, doi: 10.1109/tgrs.2022.3205558.
- [2] 施柏均, "單站法卷積神經網路預測 PGA 應用於 Earthworm 系統之研究," Master, Department of Civil and Construction Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, 2022.
- [3] 黃昭文, "卷積神經網路應用於初達波預估當站最大地表加速度之研究," 碩士, 營建工程系, 國立臺灣科技大學, 臺北市, 2020.
- [4] S. Ruder, "An Overview of Multi-Task Learning in Deep Neural Networks," *arXiv:1706.05098*, 2017.
- [5] M. Crawshaw, "Multi-task Learning With Deep Neural Networks A Surve," *arXiv:2009.09796*, 2020.
- [6] S. Liu, "End-to-End Multi-Task Learning with Attention," *arXiv:1803.10704*, 2019.
- [7] L. Liu, "On the Variance of the Adaptive Learning Rate and Beyond," *arXiv:1908.03265*, 2021.
- [8] S. M. Mousavi, Y. Sheng, W. Zhu, and G. C. Beroza, "STanford EArthquake Dataset (STEAD): A Global Data Set of Seismic Signals for AI," *IEEE Access*, vol. 7, pp. 179464-179476, 2019, doi: 10.1109/access.2019.2947848.
- [9] K.-W. Tang, "CWA: A Benchmark for Seismic Research from Taiwan," Master, Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, 2023.
- [10] W.-Y. Liao, E.-J. Lee, D. Mu, P. Chen, and R.-J. Rau, "ARRU Phase Picker: Attention Recurrent-Residual U-Net for Picking SeismicP- and S-Phase Arrivals," *Seismological Research Letters*, vol. 92, no. 4, pp. 2410-2428, 2021, doi: 10.1785/0220200382.
- [11] W.-Y. Liao, E.-J. Lee, D. Mu, and P. Chen, "Toward Fully Autonomous Seismic Networks: Backprojecting Deep Learning-Based Phase Time Functions for Earthquake Monitoring on Continuous Recordings," *Seismological Research Letters*, vol. 93, no. 3, pp. 1880-1894, 2022, doi: 10.1785/0220210274.

機器學習技術建立現地型地震預警系統相關研究(II)

子計畫四

地震預警技術研究-高精度快速地震定位演算法研發

李恩瑞 吳秉芬 廖勿渝 蔡文淮

國立成功大學

摘要

面對臺灣地區地震活動的頻繁性，本研究透過機器學習模型提升了即時 P 波相挑選的精確度，並利用測站間 P 波到時差 (Equal Differential Time, EDT) 進行震源定位，從而增強地震預警系統的效能，以達到防災減災的目的。本研究採用了一種即時地震檢測與波形識別多任務注意網絡 (Real-time Earthquake Detection and Phase-picking with multi-task Attention Network model, RED-PAN) 機器學習模型，這一模型在即時資料上進行 P 波監測，其在識別 P 波的效能上明顯優於傳統的短時窗平均除以長時窗平均 (STA/LTA) 方法，為地震預警系統提供了一個快速且準確的 P 波到時估計。利用 RED-PAN 模型在不同測站即時測得的 P 波到時，進行基於 EDT 的快速震源定位，該方法結合了臺灣的三維速度模型，計算潛在震源到鄰近測站的 P 波到時，從而能夠在短時間內準確評估震源位置。在 2023 年臺灣地區的實際地震事件應用測試中，我們的方法不僅提供了可靠的震源位置評估，也顯示出能夠減少目前氣象署地震預警系統的潛在漏報。對於幾乎同時發生且地點接近的地震事件，我們透過分析實際波形資料，證明了我們方法在精確定位方面的優異性能。這些成果的潛在應用在於能夠對時間和位置接近的多個地震事件進行同時定位，從而提高地震預警系統的精確性。展望未來，我們計劃將這些成果整合至氣象署的即時監測系統，進行更全面的效能和精確度測試，並評估其在實際地震預警中的應用潛力。

關鍵詞：地震定位、地震預警系統、地震災害

Abstract

In response to the high seismic activity in Taiwan, this study has enhanced the accuracy of real-time P-wave phase selection through machine learning models and has utilized the Equal Differential Time (EDT) between stations for source localization, thereby improving the performance of earthquake early warning systems for the purpose of disaster prevention and mitigation. The study employed a Real-time Earthquake Detection and Phase-picking with multi-task Attention Network model (RED-PAN), a machine learning model that monitors P-waves on real-time data. This model significantly outperforms the traditional Short-Term Average over Long-Term Average (STA/LTA) method in identifying P-waves, providing the earthquake early warning system with a rapid and accurate estimation of P-wave arrivals. Using the RED-PAN model to measure the immediate P-wave arrival times at different stations, the method for rapid source localization based on EDT incorporates Taiwan's three-dimensional velocity model to calculate the potential source to the nearby stations' P-wave arrival times, thus allowing for an accurate assessment of the source location in a short time frame. In the application tests on actual seismic events in the Taiwan region in 2023, our method not only provided reliable source location assessments but also demonstrated the potential to reduce the current Central Weather Administration's (CWA) earthquake early warning system's potential underreporting. For seismic events that occur nearly simultaneously and close in location, our analysis of the actual waveform data has proven the superior performance of our method in precise localization. The potential applications of these results lie in their ability to simultaneously locate multiple seismic events that are close in time and location, thereby enhancing the accuracy of the earthquake early warning system. Looking forward, we plan to integrate these findings into the CWA's real-time monitoring system, to conduct more comprehensive performance and accuracy tests, and to evaluate their potential application in actual earthquake early warnings.

Keywords: Earthquake location, earthquake early warning system, seismic hazard

一、前言

地震預警系統 (Earthquake Early Warning System, EEWS) 的目標是即時監測大地震並在強烈地面震動到達前向民眾發出警報，以最大限度地減少人員傷亡和基礎設施損壞(Allen, 2003; Gasparini et al., 2007; Allen and Melgar, 2019; Cremen and Galasso, 2020)。目前，地震預警系統主要利用地震發生後傳播速度最快的 P 波來估計地震源參數，例如震源位置和規模，進而評估各地區的最大震度，並決定是否發出預警 (Hsiao et al., 2009; Chen et al., 2019)。因此，精確的震源參數評估，在地震預警系統中扮演著重要角色。

在地震活動頻繁的地區 (圖一)，地震預警系統至關重要。臺灣位於活躍的太平洋環火帶上，過去一世紀內由於嚴重地震導致了巨大的經濟和人員損失(Wu et al., 2007; Wu and Kanamori, 2008; Hsiao et al., 2009)。臺灣的地震預警系統由中央氣象署建立，並依賴於遍布全島的密集地震監測網來監測地震 (圖二)，向公眾提供早期警告。臺灣目前的地震預警系統的工作流程主要包括：使用 STA/LTA 方法在連續紀錄中挑選地震 P 波、通過選取空間和時間上相近的 P 波到時進行震源定位，以及使用初到 P 波的峰值位移 (Pd) 來評估地震規模 (Chen et al., 2015, 2019)。儘管臺灣擁有密集的地震監測網和即時的地震信號傳輸能力，但由於複雜的斷層系統和高地震活動率，當多起事件同時發生時，目前的地震預警系統在 P 波的挑選上存在精確度下降的問題，且無法同時定位多起地震事件，導致可能需要發出警報的地震事件漏報。因此，有必要開發一個更高效且精確的地震預警系統來監測大地震，以減少可能的傷亡。

為了建立一個精確可靠的地震預警系統，必須迅速且正確地檢測到大地震產生的 P 波。近年來，機器學習在地震信號識別和波相挑選方面展現了良好的效果 (Zhu and Beroza, 2018; Mousavi et al., 2020; Liao et al., 2021, 2022)。由於機器學習模型在訓練時可以通過加入異常信號和背景噪音的數據來減少誤報，因此，許多基於機器學習的方法在結果上優於傳統方法，如 STA/LTA。此外，在面對多變的地震情境時，經過訓練的機器學習模型能夠改善識別能力，尤其在同時發生的多個地震事件和即時 P 波到時挑選方面(Liao et al., 2022)，顯著提升了性能，這對於地震預警至關重要。

要建立一個高效的地震預警系統，我們必須在快速檢測到大地震產生的 P 波後，精確進行地震定位，評估地震規模及潛在的破壞性，以此作為是否發布預警的決策依據。若能提高後續地震定位的效率及精準度，將有助於提升地震預警系統的可信度。本研究將利用機器學習方法獲得的 P 波到時，計算兩兩測站之間 P 波到時差，即等時差 (Equal Differential Time, EDT)，作為地震定位算法的基礎(Satriano et al., 2007, 2008)，以建立一個精確且快速的震源定位方法。為了提高地震定位的精確度，本研究將使用三維速度模型計算並儲存震源格點到鄰近測站的理論 P 波到時，並將計算結果依測站分類儲存。利用機器學習方法進行 P 波到

時挑選後，計算被觸發測站間的 P 波 EDT，並將各個測站儲存的潛在震源位置進行交集，以估算震源位置的機率。隨著觀測數量的增加，在多次交集運算後，將會減少震源格點的數量，從而降低後續震源定位計算所需的時間，達到快速且精確的定位效果，以提供更及時、更精確的地震預警信息，減少地震帶來的損失。

二、資料

1. **機器學習模型訓練及測試波形資料：**在本項研究中，我們利用臺灣地震資料和 STEAD 地震數據集(Mousavi et al., 2019)來訓練、驗證和測試基於機器學習方法的波相挑選模型。臺灣地震資料主要來自於中央氣象署地震觀測網、臺灣強地動觀測網和臺灣寬頻地震網的紀錄，其中 2012 年至 2018 年的地震紀錄被用於模型訓練和驗證，而 2019 年的地震紀錄則用於模型測試。為了確保模型能夠處理不同時間段的地震波形，我們保留了臺灣地震紀錄中較長的波形，這樣就可以進行隨機波形切片測試。我們還採用了資料增強方法來處理臺灣地震資料，以提升模型在連續紀錄和地震預警中波相挑選的時效性和可靠性。然而，由於 STEAD 數據集中地震波形紀錄的長度受限，這些資料難以用於資料增強。此外，我們還將臺灣測站和 STEAD 數據集中的環境噪聲資料納入訓練集，以增強模型識別背景噪聲的能力。為了在只有單分量地震記錄儀可用或記錄中斷的情況下也能進行波相識別，我們對波形進行了隨機時段的零值填充，或者選擇完全丟棄某一分量的紀錄。
2. **震源定位波形測試資料：**在震源定位波形測試方面，為了評估快速震源定位演算法對波形資料的監測能力，我們根據 2023 年現有的測站分布進行地震定位配置。選取了中央氣象署於 2023 年記錄的符合地震預警標準的地震事件，作為離線地震預警系統進行快速震源定位測試的事件。

三、方法

本研究採用了一種基於機器學習的即時 P 波挑選演算法，RED-PAN(Liao et al., 2022)，它是一個基於多任務學習（Multi-Task Learning, MTL）框架的模型。這個模型通過在不同任務間共享資料特徵，能夠在單一網路中執行多重任務(Ruder, 2017; Crawshaw, 2020)。RED-PAN 同時學習兩項任務：在連續地震紀錄上的地震訊號偵測以及地震 P 波的挑選。模型的輸入是經 Z 分數標準化處理的三分量地震波形，採樣率為 100 Hz，持續時間為 30 秒，輸出則是地震訊號偵測和波相挑選的機率函數。地震波相挑選的目標函數以地震的 P 波和 S 波到時為中心。為了使機器學習模型能在連續紀錄上識別多重地震信號，我們採用資料增強方法來提升訓練資料的多樣性。具體操作是將多個地震事件的波形剪切並疊加，生成短時內含多重地震事件的連續紀錄。我們將不同規模的地震波形隨機疊加在同一測站紀錄上，形成合成地震波形以供模型訓練，這增強了模型在地震紀錄中識別多事件的能力。此外，為了在地震預警系統中快速且準確地識別 P 波到時，

我們將包含早期 P 波的波形紀錄納入訓練資料，從而提高了模型在僅有前部 P 波波形時準確判定 P 波到時的能力。

在測試 RED-PAN 模型即時識別地震 P 波的能力方面，我們採用不同的時間窗進行測試。當模型在最初的 5 秒內提供的 P 波機率超過 0.6 時，我們將其作為門檻值（圖三）。為了減少隨機或人為噪聲造成的 P 波誤檢測，並排除輕微地震的 P 波，我們為 P 波識別後的振幅設定了一個門檻。由於 P 波的主要特徵多在垂直分量上，我們對來自每個測站的連續紀錄進行以下處理以設定垂直分量的門檻：首先對連續波形進行去均值和去趨勢處理，然後取其絕對值。對振幅應用 0.3 秒的移動平均，從而獲得移動平均結果的中位數，這樣我們就能觀察到背景噪聲水平的變化趨勢。在此基礎上，我們將每個測站 P 波的振幅門檻設為所有中位數的 5 倍，確保檢測到的 P 波具有高於背景噪聲的振幅。

在本研究中，我們利用基於等時差（Equal Differential Time，EDT）演算法（Zhou, 1994; Font et al., 2004; Satriano et al., 2007, 2008）的快速震源定位方法，並應用於地震預警系統（Earthquake Early Warning Systems，EEWS）。EDT 演算法通過觀測到的相位到達時間差($O t_i - O t_j$)和相對應的計算到達差($C t_j - C t_i$)來評估潛在震源位置的機率，這些時間差是在測站 i 和測站 j 之間，對於一潛在震源點 g 來說，可以表達為：

$$D t_{g,i,j} = (O t_i - O t_j) - (C t_{g,i} - C t_{g,j})$$

這兩者差異相等的區域代表潛在的震源，稱為 EDT 面。在這項研究中，我們假設震源的機率和 Dt 遵循高斯機率分布，表示為：

$$H_{g,i,j} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{D t_{g,i,j}^2}{2\sigma^2}}$$

在此， σ 是一個可調參數，可根據研究區域的各種不同因素修改機率模型，如相位到達的不確定性，以及速度模型的準確性。對於這項研究，我們將 σ 的值指定為 1.0。在地震預警系統（EEWS）的框架內，我們可以利用每個測站觸發的 P 波到達差異來評估震源的機率。隨著觸發的 P 波到達次數的增加，我們可以通過將所有對的機率函數相乘，逐步更新震源的機率 Q：

$$Q_{g,i,j} = \prod_{i=1}^{N-1} \prod_{j=i+1}^N H_{g,i,j}$$

Or,

$$Q_{g,i,j} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\sum_{i=1}^{N-1} \sum_{j=i+1}^N D t_{g,j,i}^2}{2\sigma^2}}$$

在上述方程中，'N' 表示一地震事件觸發 P 波到時數。基於觸發 P 波到時估算出的最高機率位置，被認為是該事件的最佳震源位置。

由於等時差（Equal Differential Time，EDT）方法的可靠性，已被用於地震預警系統即時定位地震（Satriano 等人，2007、2008）。在具有顯著地形差異和複雜速度結構的地區，如臺灣，採用 3D 速度模型可以提高震源定位的精度。因此，我們已將 EDT 方法結合到 3D 速度模型中，用於即時震源確定。臺灣的地震監測範圍由東經 119°至 123°和北緯 21°至 26°界定。我們將這個範圍劃分為網格點，這些網格點作為 EDT 定位的潛在震源。在水平面上，網格點基於站點密度進行劃分。對於臺灣站點分布密集的區域，水平網格間隔為 0.4°，而對於其他地區，間隔為 0.8°。垂直方向，網格在 2 至 20km 的深度上以 2km 為間隔劃分，在 24 至 60km 的深度上以 4km 為間隔，在 68 至 300km 的深度上以 8km 為間隔。為了提高地震活躍區域定位的準確性，我們將 1995 年至 2023 年 10 月間發生的 4.0 級以上地震震源 10 公里半徑內的水平網格間隔減半。圖二顯示了 10 公里深度網格點的分布。這些網格點在 EDT 定位過程中作為潛在震源，並計算在 3D 速度模型內，這些潛在震源到鄰近站點的 P 波到達時間。

為了及時災害預防需要迅速傳播警報信號的關鍵需求，地震預警系統主要利用最接近震源測站的 P 波到達時間進行即時震源評估。在本研究中，我們從震源網格點計算 P 波到達時間至最近的 30 個站點，大約佔所有站點的 5%。如果第一個站和最後一個站之間的 P 波到達時間差小於 2.0 秒，則用於計算的站點數量將擴大，但不超過 60 個站點（大約 10% 的所有站點）。每個網格點的 P 波到時以測站分類存儲，以供後續震源定位使用。

在地震預警系統的應用上，我們的快速震源定位流程如下（圖四），當即時地震紀錄通過 RED-PAN 辨識到一個 P 波同時滿足振幅和機率門檻值時，則進入地震事件識別。

- (1) 如果沒有地震事件，則以觸發站建立新事件。
- (2) 如果存在已知事件，則判斷觸發的 P 相位是否屬於這些地震事件。

在地震預警系統中，同一震源觸發的 P 波到時，其時間和空間關係密切相關可篩選後用於震源定位。在目前的研究中，P 波到時時間和空間篩選標準如下：

- (1) 如果觸發 P 波的站已經存在於一個地震事件中，則它不屬於該事件。

- (2) 如果觸發站 s 位於事件中站點的鄰近站點集交集之內，並且 P 波到達時間差在所有站點的最小和最大時間差範圍內，則它將被添加到事件中。
- (3) 如果觸發站的 P 波到時，在估計震源的 P 波到時 1.0 秒的時間誤差之內，則被添加到事件中。時間誤差可依照研究進行調整。

如果觸發的 P 波到時不屬於任何現有事件，則建立新事件。相反，觸發的 P 波到達用於估算或更新潛在震源機率。在更新潛在震源機率後，若滿足地震發佈報告條件，則回報系統。目前使用以下條件：

- (1) 地震事件應包含至少 5 個 P 波到時。
- (2) 在機率最高的潛在震源格點，其觀測和計算的 P 波到時均方根（RMS）差應低於門檻值。本研究中，門檻值設定為 0.3 秒。

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n (Ot_i - Ct_{g,i})^2}$$

其中 n 是用於震源估計的 P 波到時數量。此外，為了移除過時的地震事件，系統每秒檢查現有事件。如果一個現有事件超過 30 秒沒有添加任何 P 波到時，該事件就會被刪除。

四、結果

為了測試本研究開發的快速震源定位方法對地震預警地震事件震源定位的效能及精確度方面，我們以 2023 年氣象署現有的地震觀測網作為格點計算基礎，並選取了 2013 年 1 月至 10 月達地震預警要求的地震事件進行測試。本研究的地震定位結果顯示，對於臺灣島內及鄰近地區的地震事件，地震震源的誤差大多都在 10 公里以內（圖五）。然而，在蘭嶼地區發生的地震可能由於速度構造的差異及事件位於測網邊緣，因此僅利用少數 P 波進行定位的結果與地震目錄定位結果之間有較大的差異（圖五）。整體而言，本研究即時震源定位方法具有相當的精確度。

對於發生在測站分布較密集地區的地震事件，由於測站在不同方位有良好的覆蓋度，本研究的快速震源定位方法通常能夠在使用 5 到 6 個測站的 P 波到時資料後，提供相當可靠的地震定位結果。例如，2003 年 9 月 5 日發生在嘉義地區的芮氏規模 5.5 地震（圖二中 1 號地震），由於該地區測站分布較密集，演算法能在第 5、6 個測站的 P 波觸發後，對震源位置進行相當可靠的評估（圖六）。由

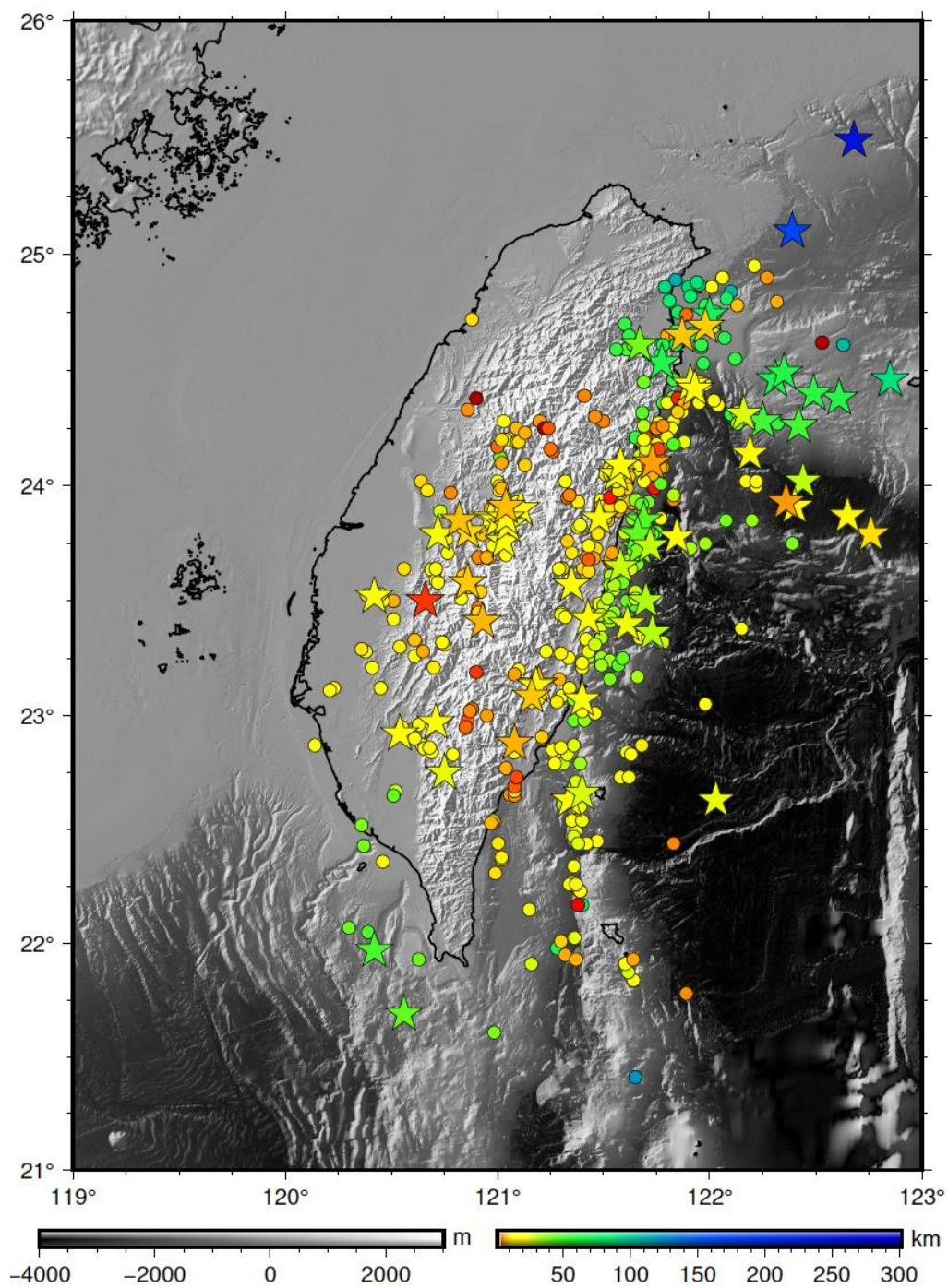
於氣象署的地震測站分布與人口分布高度相關，這有助於對人口較密集地區發生的地震事件進行快速且精確的震源評估，從而提高地震預警事件的震源判斷的效率。

然而，當地震事件發生於測站覆蓋率較差的地區時，往往需要更多的觀測資料來獲得更精確的震源位置評估。例如，2023 年 10 月 11 日發生於花蓮地區富里的芮氏規模 5.8 地震（圖二中 2 號地震），由於震源接近中央山脈，而主要的測站位於花東縱谷地區，當僅有 6 個測站的 P 波到時資料時，其震源定位結果在深度控制上較為不佳（圖七）。當觀測資料達到 8 至 9 個測站時（圖七），震源的深度評估得到了較好的控制，與氣象署地震預警系統的震源定位相比，本研究提供的地震震源評估與地震目錄位置的差異較小。

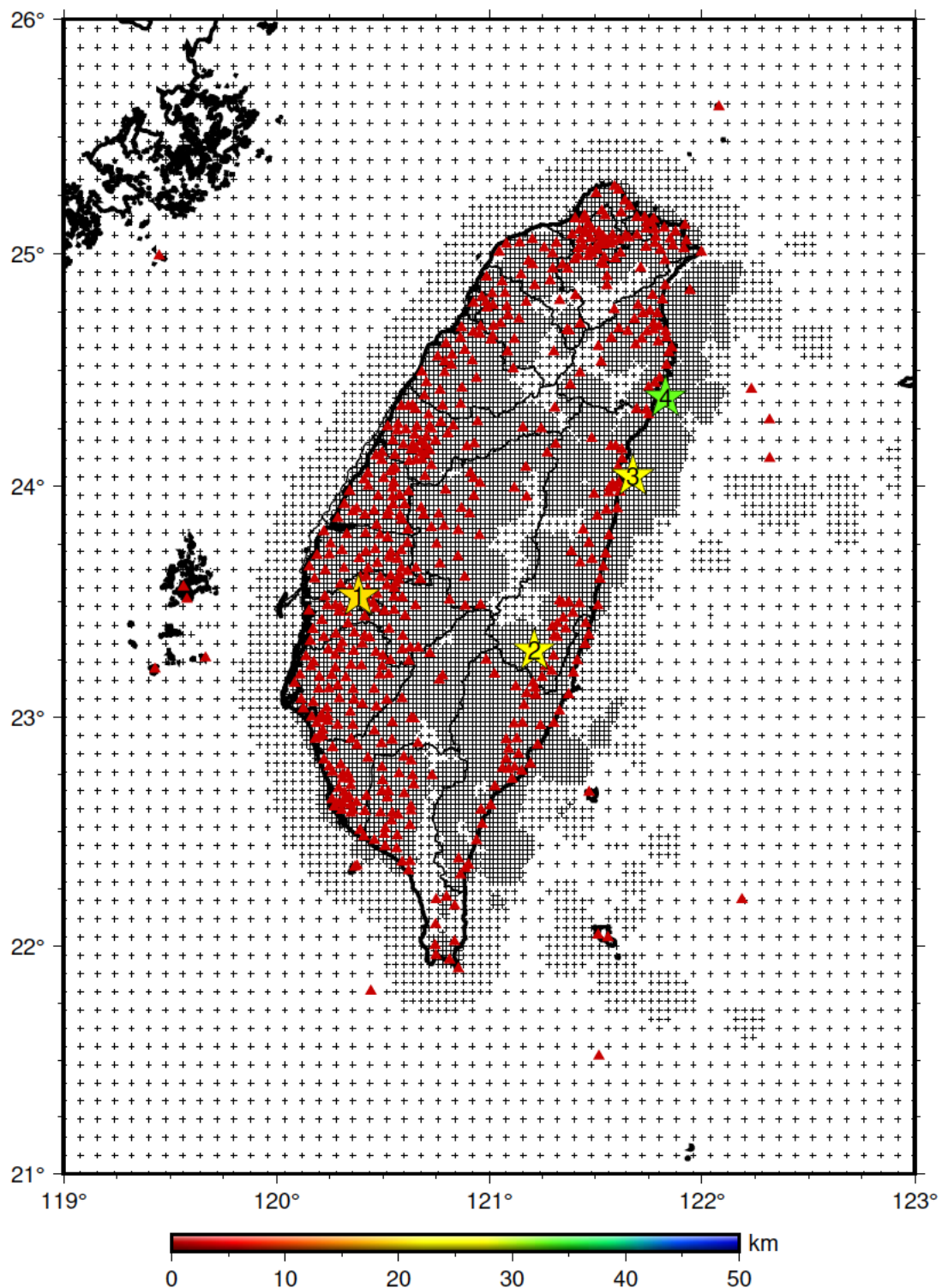
臺灣位於活躍的造山帶，其地質構造複雜且地震活動頻繁，因此在相近的時間和距離上，有可能會發生一起以上的地震事件。目前氣象署的地震預警系統只能對單一地震事件進行震源定位，因此當在接近的時間內發生多起地震時，會造成地震預警事件的漏報。例如，2023 年 9 月 15 日花蓮外海發生的芮氏規模 4.4 地震（圖二中 3 號地震，圖八），其後 5 秒內在其北方的宜蘭外海發生了芮氏規模 5.3 地震（圖二中 4 號地震，圖九）。由於兩起地震發生時間相近，其波形有重疊現象。目前氣象署採用基於 STA/LTA 的 P 波挑選方法，大多只能挑選出第一起地震事件的 P 波，並對其進行震源定位，無法同時對第二起事件進行定位，導致第二起達到地震預警標準的事件卻未發出預警。本研究的快速震源定位能在大多數測站上識別出兩個地震事件的 P 波到時，並對兩起發震時間相近的地震進行震源定位，這有助於降低臺灣地震活動頻繁情況下類似地震事件漏報的可能性。

五、結論

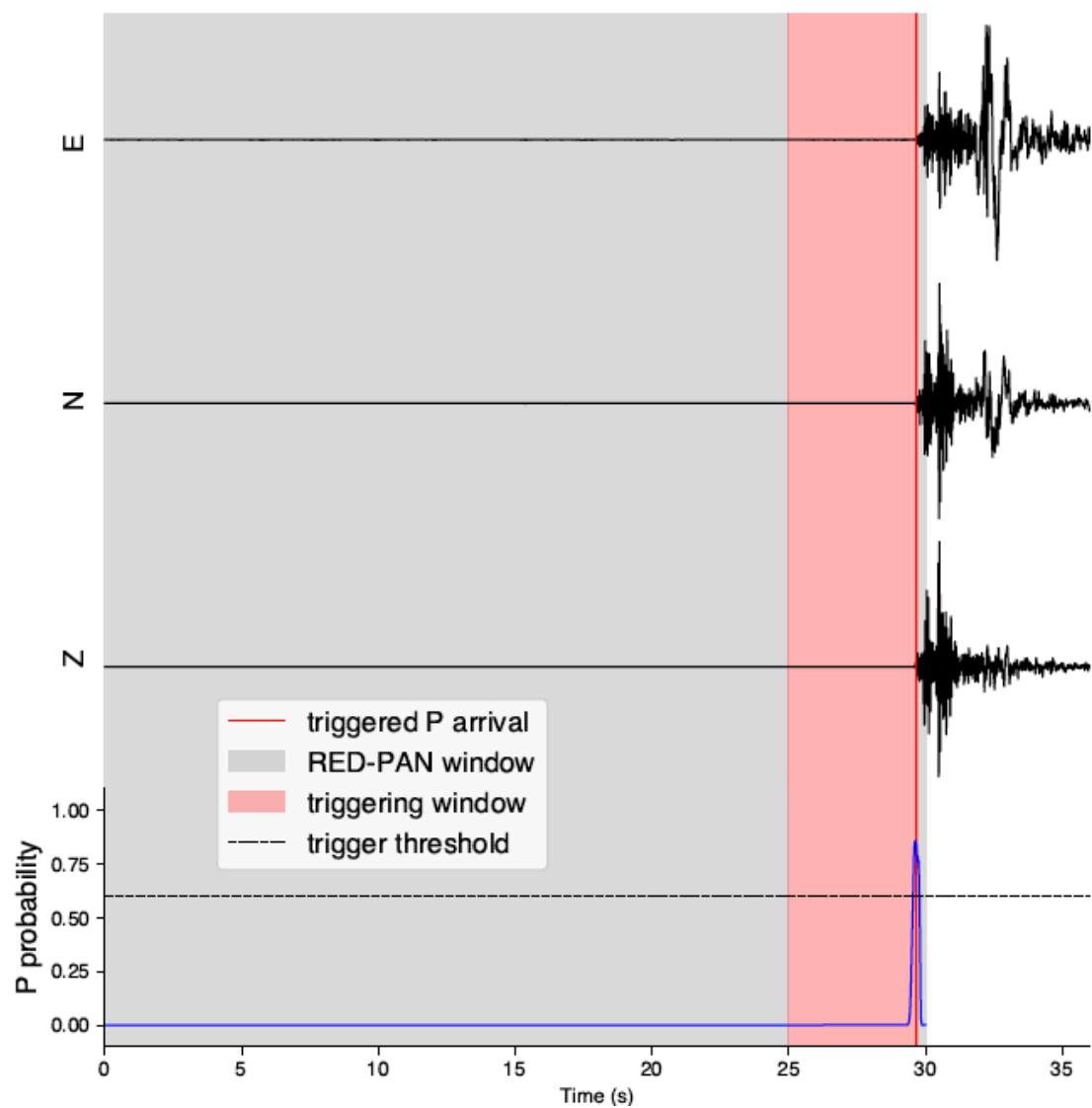
本研究採用機器學習方法對連續地震紀錄上的地震 P 波進行即時識別，並利用測站間的即時 P 波到時差來進行地震預警的震源定位。採用的機器學習模型是基於多任務學習的 RED-PAN 演算法，能夠同時進行地震信號的偵測和 P 波挑選。模型在訓練時採用了資料增強方法，以提高模型在連續資料以及即時到達的 P 波的識別能力。當多個測站偵測到觸發 P 波到時後，則利用測站間 P 波到時差，在三維速度模型下計算潛在震源格進行震源定位。本研究的即時震源定位方法以臺灣的地震波形資料進行測試，結果顯示對於達到地震預警標準的地震事件上，展現出良好的效能和精確度。在測站分布密集的地區，即使只有少量的 P 波到時資料，也能達到可靠的定位結果。相對地，在測站覆蓋率較低的地區則需要更多的資料來提高精確度。此外，本研究的震源定位方法能對一起以上的地震事件進行定位，對於臺灣和其他地震活動頻繁的地區，本研究的即時震源定位方法可有效減少地震預警的漏報情況。



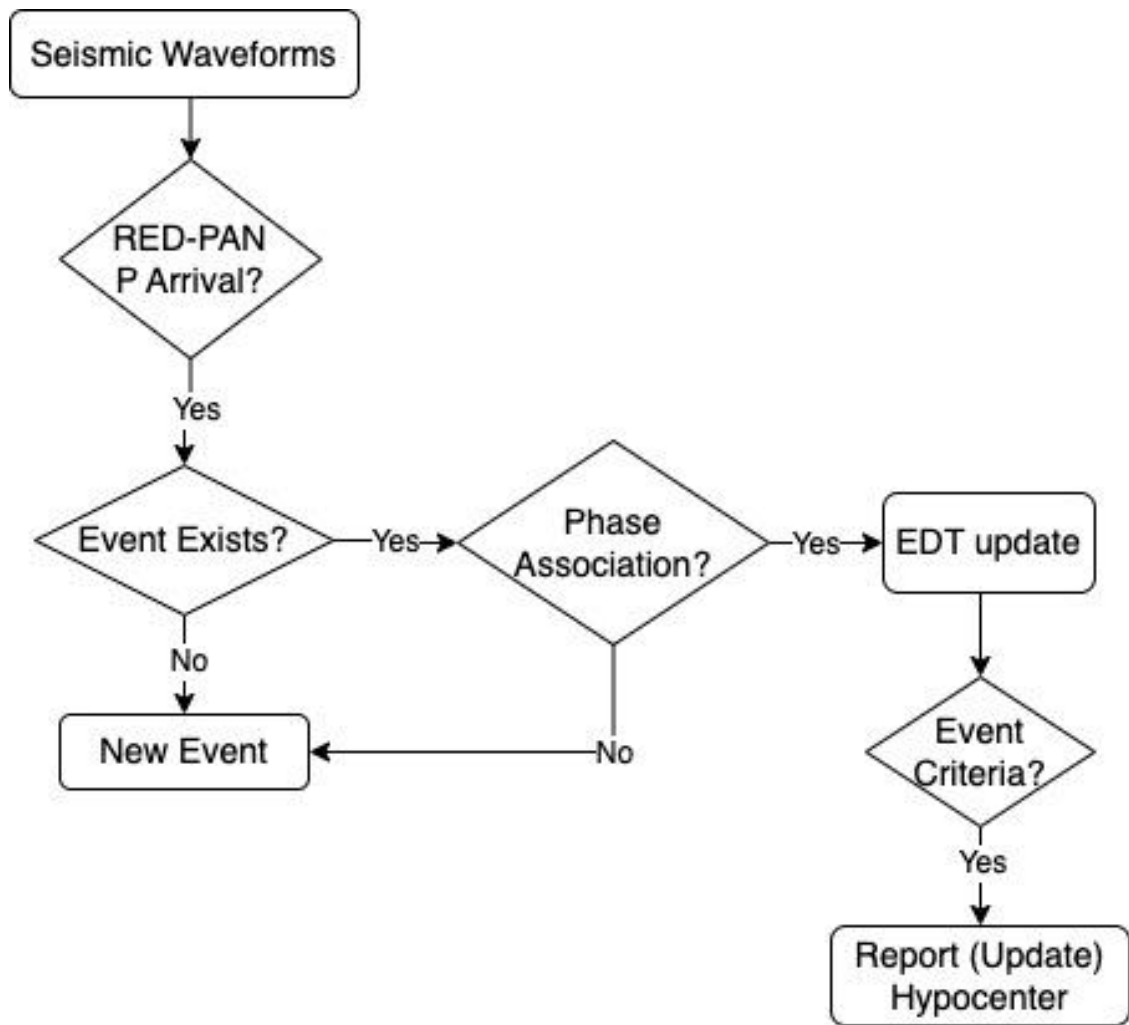
圖一、1990 至 2023 年 10 月地震預警事件分布圖。此圖展示了 1990 年至 2023 年 10 月期間符合氣象署地震預警標準的地震事件分布。圖中，規模小於 6 的地震事件以圓形表示，規模大於 6 的地震事件則以星形表示。



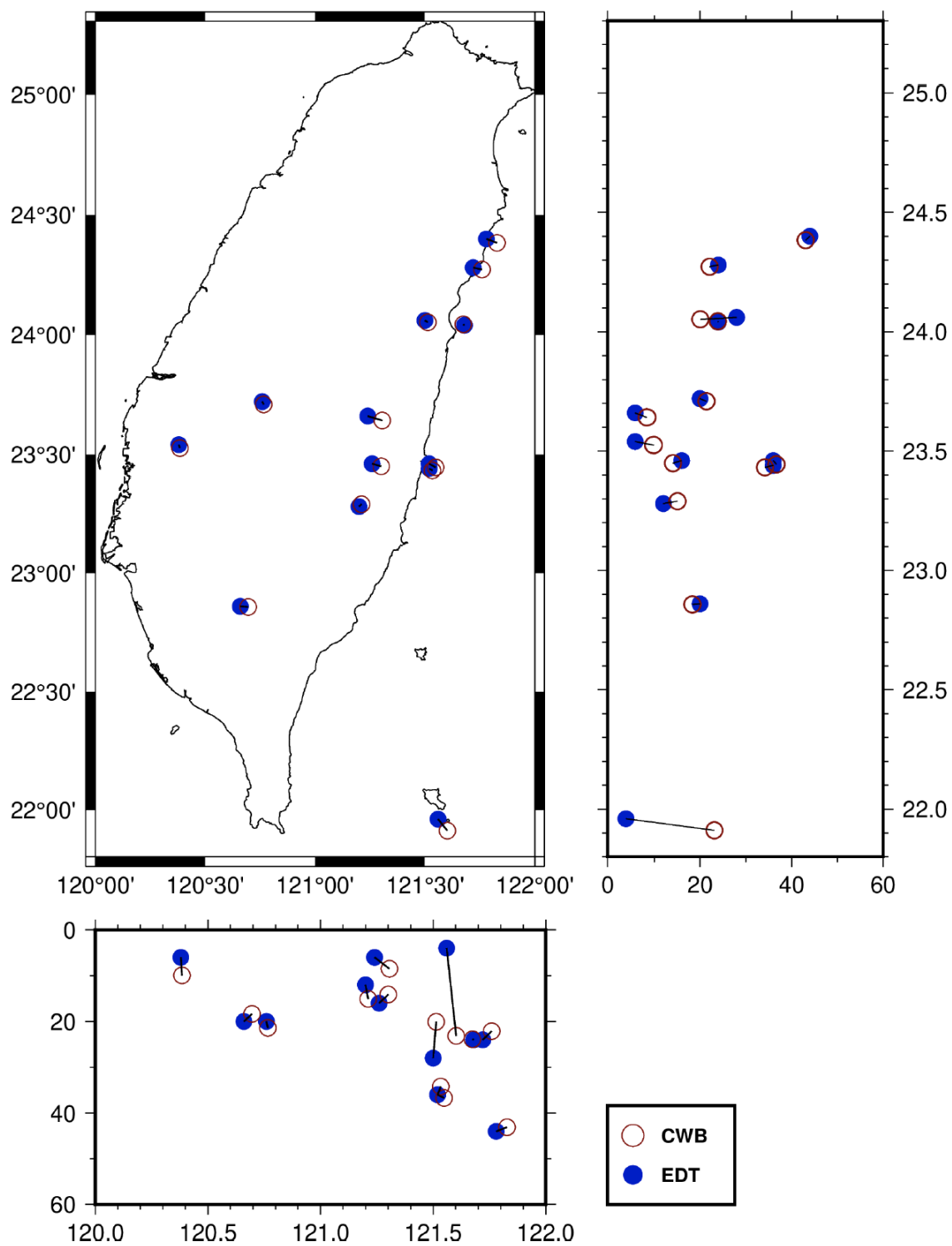
圖二、本研究潛在震源格點分布示意圖。圖中顯示了 10 公里深度下本研究考慮的潛在震源格點分布情況。在臺灣島內或歷史地震較密集的區域，格點分布較為密集。相對於島外或測站覆蓋率較低的地區，格點間隔較大。圖中星號標示了本研究用於展示地震定位的特定地震事件（圖五～圖八）。



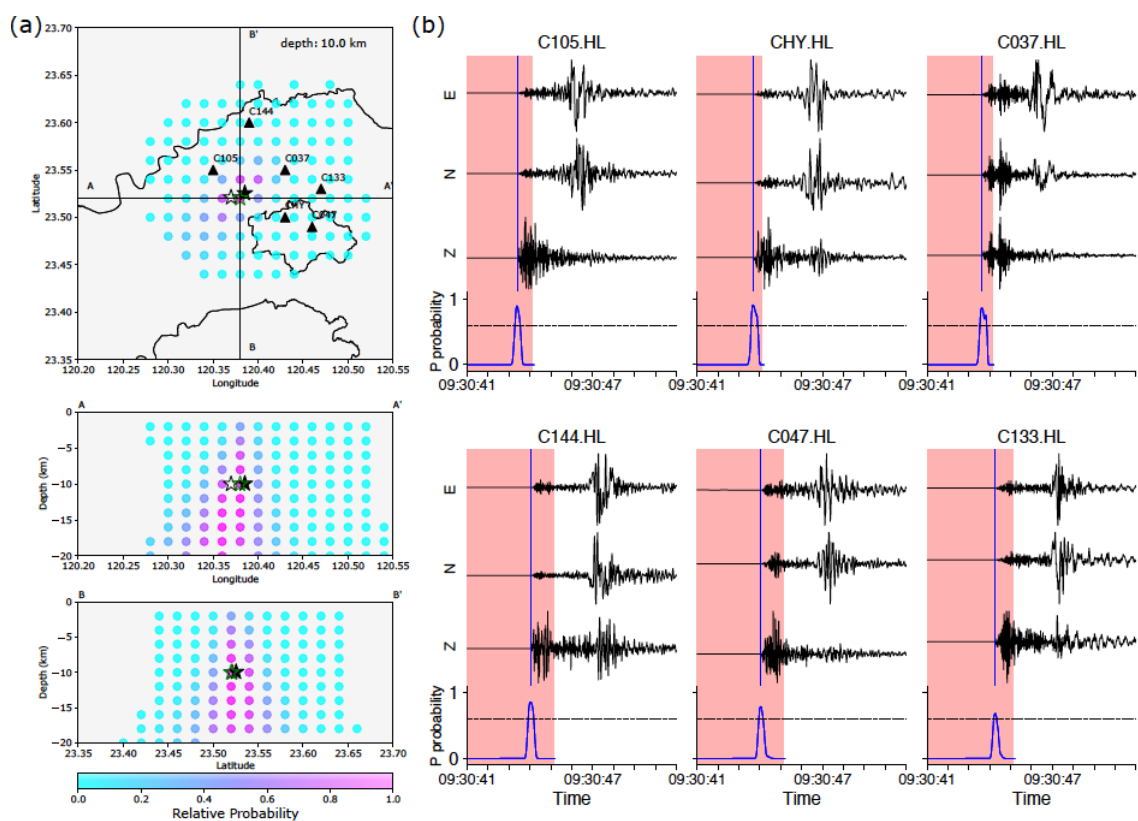
圖三、機器學習模型即時挑選 P 波到時的示意圖。灰色時間窗口標示了本研究使用的機器學習模型 RED-PAN 的預測範圍。粉紅色區域為 P 波觸發窗口，當 P 波達到預定的觸發門檻機率（虛線）且其振幅達到觸發標準時，則判定為 P 波到時（紅線）。



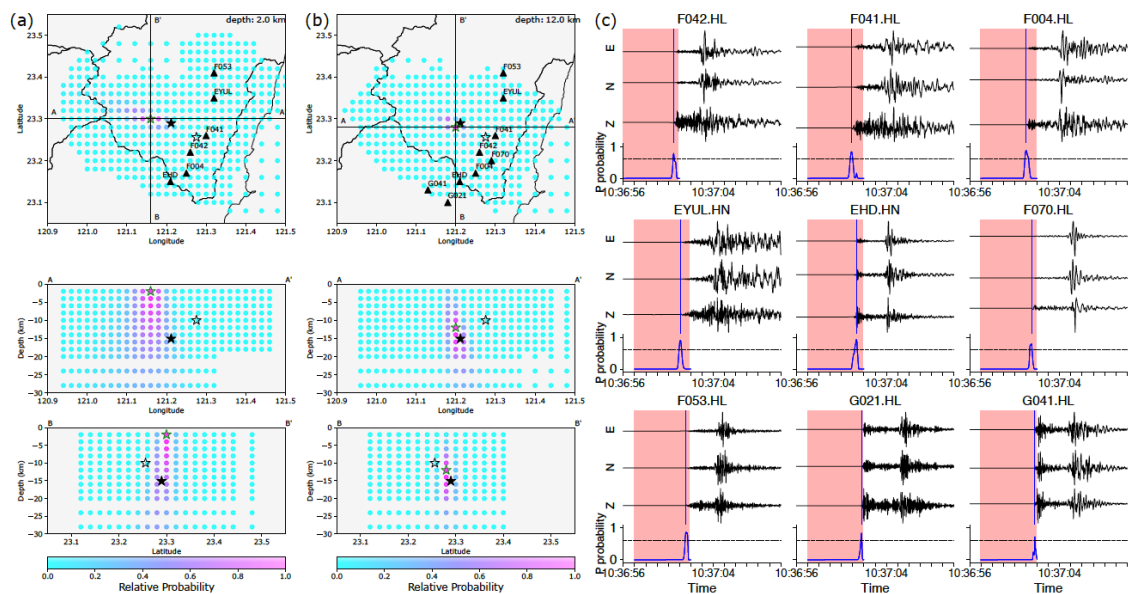
圖四、本研究的即時地震定位方法流程圖。



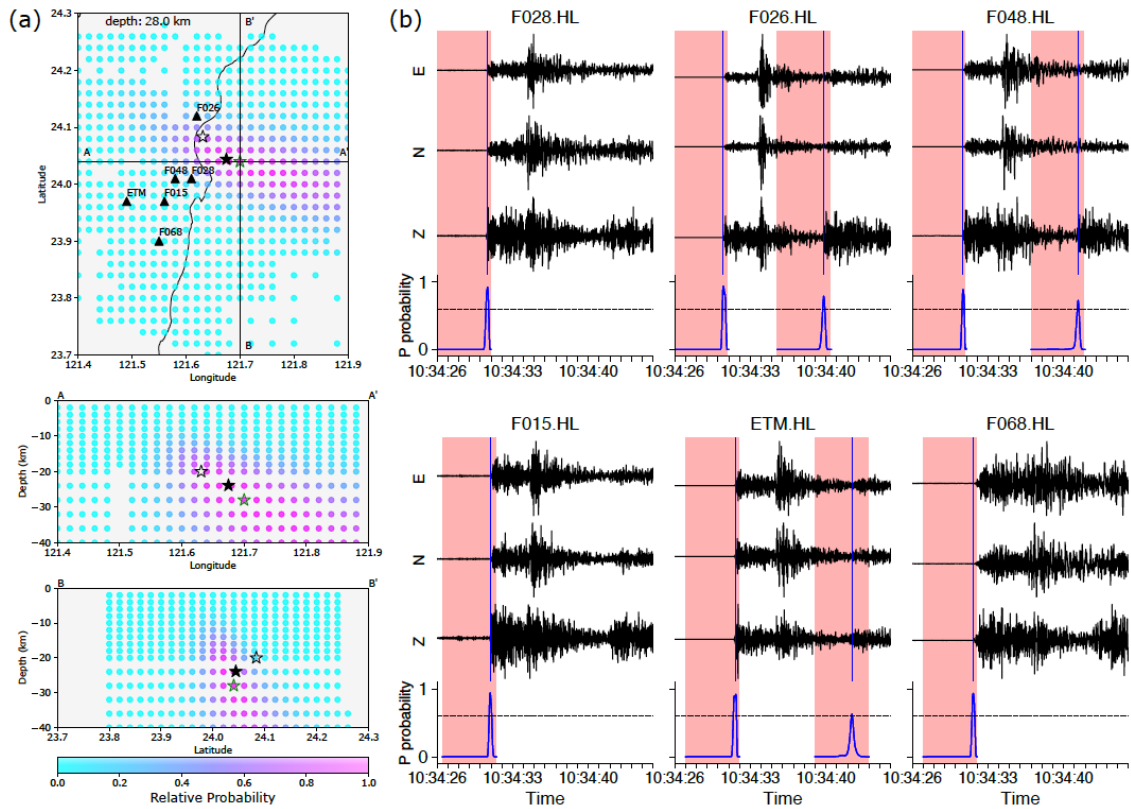
圖五、本研究即時定位結果與氣象署地震目錄震源比較圖。本研究即時震源定位對 2023 年 1 月至 10 月的地震預警事件定位結果（實心圓）和氣象署地震目錄（空心圓）之間的結果比較。



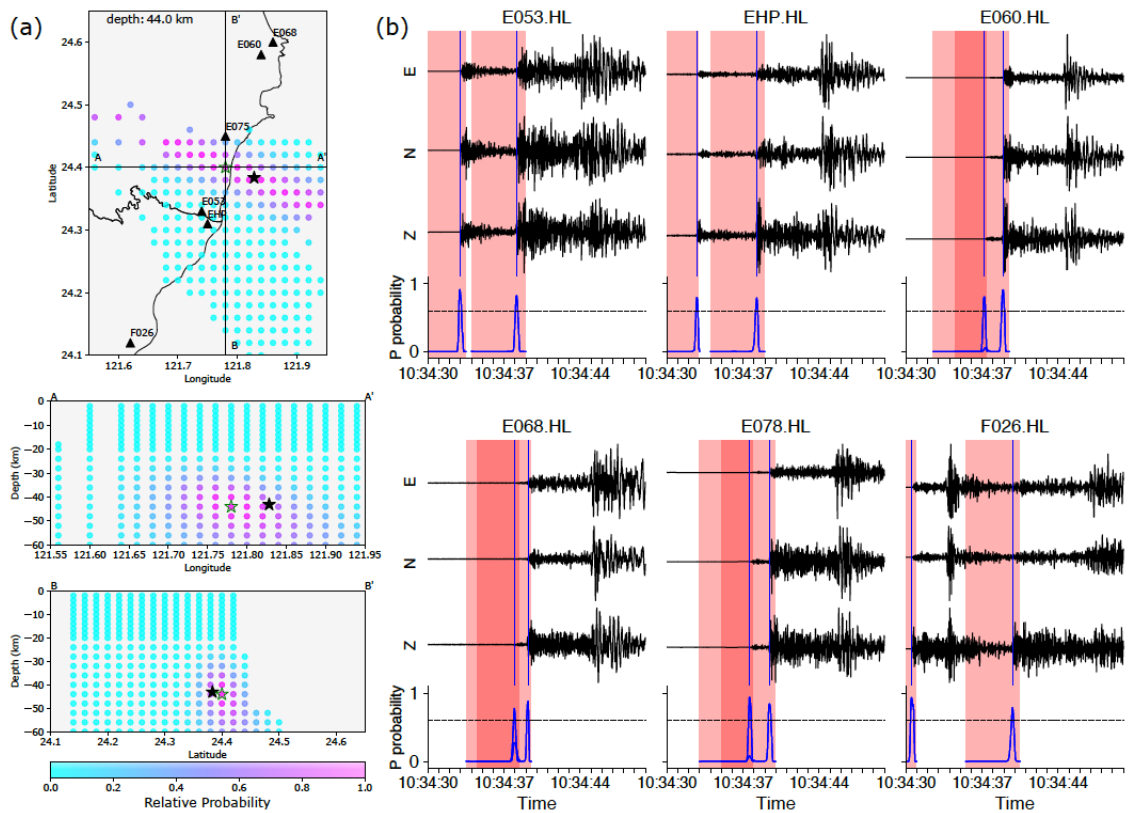
圖六、2023 年 9 月 5 日芮氏規模 5.5 嘉義地震事件震源定位結果。(a)本研究即時震源定位機率圖及震源位置比較。氣象署地震目錄震源(黑色星號)、氣象署地震預警震源位置(黑色空心星號)、本研究即時震源位置(綠色空心星號)。(b)即時震源定位使用到的地震紀錄及其觸發的 P 波到時。



圖七、2023 年 10 月 11 日芮氏規模 5.8 花蓮地區富里地震事件震源定位結果。(a) 本研究使用 6 個測站紀錄得到的即時震源定位機率圖及震源位置比較。(b) 本研究使用 9 個測站紀錄得到的即時震源定位機率圖及震源位置比較。氣象署地震目錄震源（黑色星號）、氣象署地震預警震源位置（黑色空心星號）、本研究即時震源位置（綠色空心星號）。(c) 即時震源定位使用到的地震紀錄及其觸發的 P 波到時。



圖八、2023 年 9 月 15 日芮氏規模 4.4 花蓮外海地震事件定位結果。(a)本研究即時震源定位機率圖及震源位置比較。氣象署地震目錄震源（黑色星號）、氣象署地震預警震源位置（黑色空心星號）、本研究即時震源位置（綠色空心星號）。(b)即時震源定位使用到的地震紀錄及其觸發的 P 波到時。



圖九、2023 年 9 月 15 日芮氏規模 5.3 宜蘭外海地震事件氣象署地震預警系統漏報定位結果。(a)本研究即時震源定位機率圖及震源位置比較。氣象署地震目錄震源（黑色星號）和本研究即時震源位置（綠色空心星號）。(b)即時震源定位使用到的地震紀錄及其觸發的 P 波到時。

參考文獻

- Allen, R. M., 2003, The Potential for Earthquake Early Warning in Southern California, *Science*, 300, no. 5620, 786–789, doi: 10.1126/science.1080912.
- Allen, R. M., and D. Melgar, 2019, Earthquake Early Warning: Advances, Scientific Challenges, and Societal Needs, *Annu. Rev. Earth Planet. Sci.*, 47, no. 1, 361–388, doi: 10.1146/annurev-earth-053018-060457.
- Chen, D., N. Hsiao, and Y. Wu, 2015, The Earthworm Based Earthquake Alarm Reporting System in Taiwan, *Bulletin of the Seismological Society of America*, 105, no. 2A, 568–579, doi: 10.1785/0120140147.
- Chen, D.-Y., T.-L. Lin, H.-C. Hsu, Y.-C. Hsu, and N.-C. Hsiao, 2019, An approach to improve the performance of the earthquake early warning system for the 2018 Hualien earthquake in Taiwan, *Terr. Atmos. Ocean. Sci.*, 30, no. 3, 423–433, doi: 10.3319/TAO.2018.12.23.02.
- Crawshaw, M., 2020, Multi-Task Learning with Deep Neural Networks: A Survey, doi: 10.48550/arXiv.2009.09796.
- Cremen, G., and C. Galasso, 2020, Earthquake early warning: Recent advances and perspectives, *Earth-Science Reviews*, 205, 103184, doi: 10.1016/j.earscirev.2020.103184.
- Font, Y., H. Kao, S. Lallemand, C.-S. Liu, and L.-Y. Chiao, 2004, Hypocentre determination offshore of eastern Taiwan using the Maximum Intersection method, *Geophysical Journal International*, 158, no. 2, 655–675, doi: 10.1111/j.1365-246X.2004.02317.x.
- Gasparini, P., G. Manfredi, and J. Zschau, 2007, *Earthquake Early Warning Systems*, Springer.
- Hsiao, N.-C., Y.-M. Wu, T.-C. Shin, L. Zhao, and T.-L. Teng, 2009, Development of earthquake early warning system in Taiwan, *Geophysical Research Letters*, 36, no. 2, doi: 10.1029/2008GL036596.
- Liao, W.-Y., E.-J. Lee, D.-Y. Chen, P. Chen, D. Mu, and Y.-M. Wu, 2022, RED-PAN: Real-Time Earthquake Detection and Phase-Picking With Multitask Attention Network, *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–11, doi: 10.1109/TGRS.2022.3205558.
- Liao, W.-Y., E.-J. Lee, D. Mu, P. Chen, and R.-J. Rau, 2021, ARRU Phase Picker: Attention Recurrent-Residual U-Net for Picking Seismic *P* - and *S* -Phase Arrivals, *Seismological Research Letters*, 92, no. 4, 2410–2428, doi: 10.1785/0220200382.
- Mousavi, S. M., W. L. Ellsworth, W. Zhu, L. Y. Chuang, and G. C. Beroza, 2020, Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking, *Nature Communications*, 11, no. 1, doi: 10.1038/s41467-020-17591-w.

- Mousavi, S. M., Y. Sheng, W. Zhu, and G. C. Beroza, 2019, STanford EArthquake Dataset (STEAD): A Global Data Set of Seismic Signals for AI, IEEE Access, 7, 179464–179476, doi: 10.1109/ACCESS.2019.2947848.
- Ruder, S., 2017, An Overview of Multi-Task Learning in Deep Neural Networks, doi: 10.48550/arXiv.1706.05098.
- Satriano, C., A. Lomax, and A. Zollo, 2007, Optimal, Real-time Earthquake Location for Early Warning, in *Earthquake Early Warning Systems* P. Gasparini, G. Manfredi, and J. Zschau (Editors), Springer, Berlin, Heidelberg, 85–96.
- Satriano, C., A. Lomax, and A. Zollo, 2008, Real-Time Evolutionary Earthquake Location for Seismic Early Warning, Bulletin of the Seismological Society of America, 98, no. 3, 1482–1494, doi: 10.1785/0120060159.
- Wu, Y.-M., N.-C. Hsiao, W. H. K. Lee, T. Teng, and T.-C. Shin, 2007, State of the Art and Progress in the Earthquake Early Warning System in Taiwan, in *Earthquake Early Warning Systems* P. Gasparini, G. Manfredi, and J. Zschau (Editors), Springer, Berlin, Heidelberg, 283–306.
- Wu, Y.-M., and H. Kanamori, 2008, Development of an Earthquake Early Warning System Using Real-Time Strong Motion Signals, Sensors, 8, no. 1, 1–9, doi: 10.3390/s8010001.
- Zhou, H., 1994, Rapid three-dimensional hypocentral determination using a master station method, J. Geophys. Res., 99, no. B8, 15439–15455, doi: 10.1029/94JB00934.
- Zhu, W., and G. C. Beroza, 2018, PhaseNet: A Deep-Neural-Network-Based Seismic Arrival Time Picking Method, Geophysical Journal International, doi: 10.1093/gji/ggy423.