

機器學習技術建立現地型地震預警系統相關研究(III)

子計畫一

以人工智慧技術進行台灣各地區即時地震最大震度預測

金台齡 吳祐甄 張朋楫 陳彥翰 李世璽

國立臺灣科技大學

摘要

地震是臺灣面臨的主要自然災害之一，其突發性和破壞性對人民的生活安全和經濟活動造成嚴重威脅。因此，如何有效減少地震所造成的災害一直是人們關注的焦點。傳統的地震預警方法主要依靠地震波的速度差異進行預警，但這些方法在預警時間和精確性上往往有限制。隨著機器學習技術的發展，我們將其與地震預警系統結合，利用機器學習更精確地分析地震數據，並透過地震預警系統提前警告地震的發生，從而增加反應時間，減少損失。本研究計畫旨在開發和改進基於機器學習技術的地震預警系統，涵蓋三個主要方向。首先，利用近震央測站的地震波形數據，預測遠震央目標位置的峰值地面加速度(PGA)，本研究提出兩個模型預測 PGA，一是利用多視角圖注意力網路、二是將圖神經網路的節點特徵傳播方法結合 Transformer 的注意力機制，使模型學習空間關係。其次，結合卷積神經網路和 Transformer 技術進行自動拾取 P 波的研究，通過自動拾取 P 波的技術，加速地震事件的識別和分析。最後，將現有 Earthworm 系統的波形數據傳輸方式改為 MQTT 協定，以提高數據傳輸效率和穩定性，並實際應用於地震預警系統上，將數據傳輸至預測模型以進行後續預測任務，並使用 Line Notify 進行警報的通知。這三項研究的成果預期將提升臺灣地震預警系統的精確性、速度和可靠性，從而減少地震災害對生命和財產的威脅。

關鍵字：類神經網路、地震預警系統、最大加速度預測、P 波拾取任務

Abstract

Earthquakes are one of the critical threats faced by Taiwan, posing sudden and destructive dangers to public safety and economic activities. Therefore, effectively reducing damages caused by earthquakes has been a focus of public concern. Traditional earthquake early warning methods typically rely on the observed characters of seismic waves, but these methods often have limitations in terms of warning time and accuracy. By integrating machine learning technologies, earthquake early warning systems may perform better data analysis and issue alerts earlier. Thus, the reaction time can be increased and potential losses can be reduced.

This project aims to develop enhanced earthquake early warning systems based on machine learning technologies. The major investigated problems are on the following three folds. First, seismic waveform data from the seismic stations near the epicenter are used to predict the peak ground acceleration (PGA) at target locations farther from the epicenter. We propose two models for predicting PGA: one utilizes a multi-view graph attention network, while the other combines the node feature propagation on graph neural networks with the attention mechanism of Transformers to enable PGA predictions for remote locations. Second, Transformer technologies are exploited to conduct automated P-wave picking in order to accelerate identifying the earthquake events. Finally, a realtime system is developed with the transmission protocol MQTT to enhance transmission efficiency and scalability. The warning alerts are sent via Line Notify.

The results of these three studies are expected to improve the accuracy, speed, and reliability of Taiwan's earthquake early warning systems and reduce the threats of damages caused by earthquakes.

Keywords : Neural Network · Earthquake Early Warning System · PGA prediction · P-wave Picking

1 前言

臺灣位於環太平洋地震帶，地震活動頻繁。特別是最近發生的芮氏規模 7.2 的強震，造成了震後一個多月內地震頻率的顯著上升，且經常伴隨著破壞性地震，對民眾的財產與心理造成了巨大損失。雖然當前無法精確預測地震的時間與位置，但透過地震預警系統，可以有效地減少災害的影響並降低損失。該系統能在地震發生之初，還未造成嚴重破壞時，提前發出警告，為民眾提供更多時間來採取防範措施，從而減輕災難造成的損害。因此，積極應用地震預警系統並不斷強化防災意識和應對措施，是減少地震災害影響的關鍵途徑。

地震預警系統主要分為兩類：現地型預警與區域型預警。現地型預警系統專注於震央附近區域的警報，盡可能縮小預警盲區，以應對震央周圍地區的需求。而區域型預警系統則是通過多個近震央測站的數據，根據探測到的強烈地震波向較遠區域發送警報。這樣的設置能夠確保遠離震央的重點區域也能夠提前收到警報，降低地震帶來的風險。結合兩類預警系統的運作，能夠全方位提升地震防災的效果，保障更多民眾的生命與財產安全。

第四、五、六、七節將對不同的研究進行詳盡介紹並分析其成果。第四節將探討如何利用圖注意力網路(Graph-attention)，第五節是將圖神經網路結合Transformer之圖轉換器(Graph Transformer)技術進行空間分析，通過近震央測站的地震波形數據來預測位於遠震央的目標位置的峰值地面加速度(PGA)。第四及第五節之研究旨在運用圖神經網路技術來實現區域型地震預警系統，以提高地震強度的預測準確度，並為相關區域提供更及時的警報，從而有效降低地震災害風險。第六節將介紹結合Transformer編碼器與UNET架構的TransUNet模型，該模型能通過自動學習捕捉地震波形數據中的特徵來實現P波的自動識別，從而加速地震事件的辨識與分析。這項研究有助於提高地震預警系統的反應速度與預測精度，對未來的地震預警技術研究具有重要意義。第七節探討將現有Earthworm系統的波形數據傳輸協議更換為MQTT協定，這一改進將大幅提升數據傳輸的效率與穩定性，並實際應用於地震預警系統中，透過Line Notify進行預警訊息通知。

這四項研究的結果預期將為地震預警技術的發展和實際應用提供有力的支持，進一步提高地震災害管理的效果與效率。報告將深入探討這些方法、數據及實驗結果，並分析它們在地震預警領域的潛在應用，為地震預警技術的改進與發展提供寶貴的參考，進一步降低地震災害帶來的風險。

2 類神經網路應用於區域型地震預警I

2.1 簡介

本研究利用P波到達前5秒至P波後數秒內的三軸加速度波形資料，結合類神經網路技術，來預測目標測站的PGA（峰值地動加速度）。研究的目標是希望藉由近震央的幾個測站，能在地震發生後，地震波尚未到達其他縣市區域前，快速且準確地預測出該區域的PGA。同時，期望模型能夠實時預測，隨著地震事件的進展，透過獲取更多的波形資訊，使預測更加精確，以實現區域性預警的效果。

2.2 模型介紹

目前採用的模型核心是多視角圖注意力網路(Multi-Perspective Graph Attention Network, MP-GAT)，透過多個視角的注意力機制，為每個節點分配不同的權重，從而更全面地捕捉節點之間的關聯，增強重要資訊的傳遞效果。本次研究將地震測站視為節點，利用測站偵測到的波形與經緯度位置作為節點的特徵、測站間距離的關係作為兩節點相連之邊的權重，以此構建出圖結構數據。圖注意力網路(Graph Attention Network, GAT)是一種用於處理圖結構數據的深度學習模型[1]。它透過引入注意力機制，根據每個節點與鄰居節點的關聯性分配權重，從而選擇性地關注圖中重要的節點和邊。具體計算方式是先將節點 i 和節點 j 透過評分函數 e 計算分數，該分數表示節點 i 與其鄰居節點 j 的關聯程度，如下式(1)，

$$e(h_i, h_j) = \sigma(a^\top \cdot [Wh_i \parallel Wh_j]) \quad (\text{式 1})$$

h_i 、 h_j 為節點 i 和節點 j 的特徵向量； σ 為非線性激活函數； a 、 W 為可學習的參數矩陣，用於計算兩個節點間的注意力分數； $\parallel$ 表示特徵拼接。將節點 i 與所有鄰居節點 $j \in N_i$ 的注意力分數使用 softmax 進行正規化，得到的注意力函數定義，如下式(2)，

$$\alpha_{ij} = \text{softmax}(e(h_i, h_j)) = \frac{\exp(e(h_i, h_j))}{\sum_{k \in N_i} \exp(e(h_i, h_k))} \quad (\text{式 2})$$

α_{ij} 表示節點 i 和節點 j 之間正規化後的注意力分數； N_i 為節點 i 的鄰居節點集合。然後，使用該分數作為鄰居節點的權重，更新節點 i ，如下式(3)。

$$h_i = \sigma \left(\sum_{j \in N_i} \alpha_{ij} Wh_j \right) \quad (\text{式 3})$$

在 GATv2 中，為了克服原本 GAT 評分函數中的限制，進行了一些修改，使其更加靈活和表現更佳。GATv2 改進的關鍵點是將非線性操作應用在兩個線性變換之中，增加模型的表現，如下式(4)。

$$e(h_i, h_j) = a^\top \sigma(W \cdot [h_i \parallel h_j]) \quad (\text{式 4})$$

MP-GAT [2]在 GATv2 基礎上，不只關注於所有鄰居節點之間的注意力分數來捕捉全局信息，另外加入局部特徵，通過去除注意力分數較低的鄰居節點，專注於最相關的局部節點捕捉到關鍵的信息。以節點 i 為例，先通過評分函數得到所有鄰居節點的注意力分數，並選擇與節點 i 最相關的前 M 個節點信息作為局部的注意力分數，如下式(5)。

$$\alpha_{im} = \frac{\exp(e(h_i, h_m))}{\sum_{n \in N_{i,m}} \exp(e(h_i, h_n))} \quad (\text{式 5})$$

$N_{i,m}$ 代表與節點 i 最相關的 M 個鄰近節點的集合，且 $m, n \in N_{i,m}$ 。接下來，將 M 個節點特徵聚合，以獲得節點 i 的局部特徵表示，如下式(6)。

$$h_i^L = \sigma \left(\sum_{m \in N_{i,m}} \alpha_{im} Wh_m \right) \quad (\text{式 6})$$

最後，MP-GAT 整合的多視角特徵，如下式(7)， h'_i 為更新後節點 i 的特徵， h_i^G 為節點 i 的全局特徵； h_i^L 為節點 i 的局部特徵。

$$h'_i = h_i^G \parallel h_i^L \quad (\text{式 7})$$

2.3 模型架構

本研究模型架構如圖 1，由輸入、卷積神經網路(Convolutional Neural Network, CNN)、位置嵌入(Position embedding)、多視角圖注意力網路(Multi-Perspective Graph Attention Network, MP-GAT)和輸出主要的五個部分組成，接下來將介紹各部分的功能。

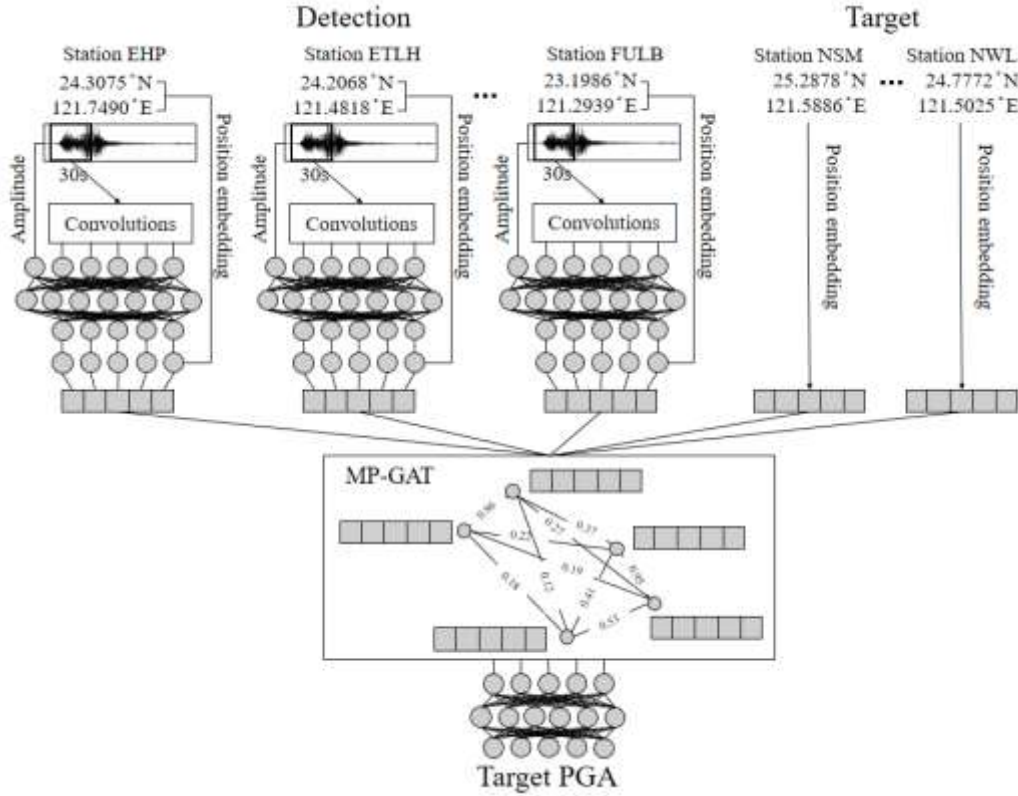


圖 1、模型架構圖

在輸入部分，模型會將地震測站依照輸入資訊的不同，分為偵測與目標左右兩個部分。其中，偵測部分為花蓮地區的地震測站，輸入為各測站的 30 秒三軸加速度波形和測站經緯度；目標部分為臺北地區的地震測站，輸入為各測站經緯度。在左半部分會先將花蓮測站的三軸加速度波形利用 2 層 2 維卷積神經網路和 1 層 1 維卷積神經網路進行特徵提取，kernel size 依序分別為 (1,1)、(1,3)、5，filter 分別為 8、32、64，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)，另外，為了保留波形中的關鍵信息，將波形中最大絕對振幅合併到卷積神經網路後的特徵中，提供一個能捕捉全局訊號強度的重要特徵，再來模型會通過 2 層全連接層 (Fully Connected Layer)，units 分別為 1024 和 400，並在全連接層後加入機率為 0.1 的 Dropout，形成具有各測站波形的特徵矩陣 M ， $M \in \mathbb{R}^{n \times d}$ ， n 為偵測測站數， d 為特徵維度。針對位置資訊，本研究將花蓮測站的經緯度做位置嵌入，將偵測部分的測站經緯度分別映射到一個 d 維的向量，形成一個矩陣 P ， $P \in \mathbb{R}^{n \times d}$ ，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，形成具有各測站的混合特徵矩陣 E ， $E \in$

$\mathbb{R}^{n \times d}$ ，如下式(8)。

$$E_{n \times d} = M + P \quad (\text{式 } 8)$$

模型的右半部分則為目標輸出的測站，不會有波形資料輸入，只有測站位置的資訊，因此，本研究將目標測站的經緯度做位置嵌入，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 為目標測站數，提取目標測站位置的特徵。接著，將花蓮和目標測站的特徵合併，形成一個具有各測站特徵的矩陣 X ， $X \in \mathbb{R}^{(n+k) \times d}$ ，如下式(9)。

$$X_{(n+k) \times d} = [E \parallel T] \quad (\text{式 } 9)$$

為了考慮各測站的距離關係，本研究參考相關論文計算各測站的距離權重[3]，並以鄰接矩陣(Adjacency matrix) A ， $A \in \mathbb{R}^{(n+k) \times (n+k)}$ 的方式表示，計算方式為利用經緯度計算兩兩測站的距離，根據距離的大小比例分配權重，如下式(10)， A_{ij} 為鄰接矩陣元素； $D_{i,j}$ 為測站 i 和測站 j 的距離； D_{max} 為各測站的最大距離。

$$A_{ij} = 1 - \frac{D_{i,j}}{D_{max}} \quad (\text{式 } 10)$$

特徵矩陣和鄰接矩陣會透過兩層多視角圖注意力網路學習測站之間的距離關係和地震波能量衰減與峰值地面加速度的關係。多視角圖注意力網路的前向傳播如下式(11)，

$$h_i^{(l+1)} = \sigma \left(\sum_{j \in N_i} \alpha_{ij} A_{ij} W^{(l)} h_j^{(l)} + \sum_{m \in N_{i,m}} \alpha_{im} A_{im} W^{(l)} h_m^{(l)} \right), \quad (\text{式 } 11)$$

$$h_i^{(0)} = X_i$$

$h_i^{(l+1)}$ 為測站 i 在第 l 層的特徵； σ 為激活函數； α_{ij} 表示測站 i 和測站 j 之間正規化後的注意力分數； N_i 為節點 i 的鄰居節點集合； W 為可學習的權重矩陣； $N_{i,m}$ 代表與節點 i 最相關的 M 個鄰近節點的集合。

多視角圖注意力神經網路的核心操作是圖注意力層(Graph Attention Layer)，它通過為每個節點計算與其鄰近節點的注意力權重來更新節點的特徵表示。相比於圖卷積層(Graph Convolutional Layer)，圖注意力層會根據注意力機制為鄰居節點分配不同的重要性權重，然後將這些加權的特徵進行總和，再通過非線性變換來更新節點特徵。這種加權計算基於圖結構中的鄰接關係，並且通過兩層圖注意力層的堆疊，模型可以學習到更複雜的圖上空間特徵。每層圖注意力層後都接有整流線性單位函數(Rectified Linear Unit, ReLU)。最後，將目標測站的特徵通過全連接層(Fully Connected Layer)進行回歸(regression)，輸出目標測站的PGA。

2.4 資料處理

本次研究使用的地震事件為 2012-2023 發生在花蓮的地震事件，並且只關注臺北測站有偵測到 PGA 的地震，以第一個偵測到 P 波的測站作為標準，統一取其 P 波到時前 5 秒到後 25 秒的 30 秒時間窗。為了使模型能夠即時預測，本研究對資料集進行了調整。在訓練過程中，取時間點 T 的三軸加速度波形去做訓練，且將時間點 T 之後的三軸加速度波形補 0，以匹配模型的輸入，T 為 P 波到時後 3 秒開始，以 1 秒作為間隔，持續到 P 波到後 25 秒。本次研究使用測站為 CWA 地震觀測網中臺北、新北與花蓮的地震測站，測站分布如下圖 2。在訓練集使用 2012-2020 年的地震事件，規模 2-3 有 69 筆；規模 3-4 有 3289 筆，規模 4-5 有 4025 筆，規模 5 以上有 690 筆，總共 8073 筆地震事件。測試集使用 2021-2023 年的地震事件，規模 3-4 有 943 筆，規模 4-5 有 2116 筆，規模 5 以上有 851 筆，總共 3910 筆地震事件，規模分佈如下圖 3。另外，為了因應大地震在本次資料及過少的問題，進行了數據增強。本次研究的數據增強方法，主要針對地震事件數據進行處理，目的是通過增強數據來解決數據集不平衡問題，並且通過隨機變換波形來增加數據多樣性。具體增強步驟如下：對每個事件的臺北地區測站的數據進行分析，根據 PGA 的範圍，將事件分為以下五類： <0.8 、 $0.8\sim 2.5$ 、 $2.5\sim 8$ 、 $8\sim 25$ 、 >25 ，這些類別決定了每個事件數據的增強權重（即該事件將被複製多少次），根據預設權重進行數據平衡。較少出現的類別會有更高的增強倍率。為了增加數據多樣性，我們對每個事件的波形數據進行隨機變換。具體來說，對花蓮地區測站的波形數據進行隨機選擇，並將這些測站的三分量波形同時進行正負顛倒（即沿基線隨機翻轉）。圖 4 為原始數據集與資料增強後數據集的 PGA 分佈。

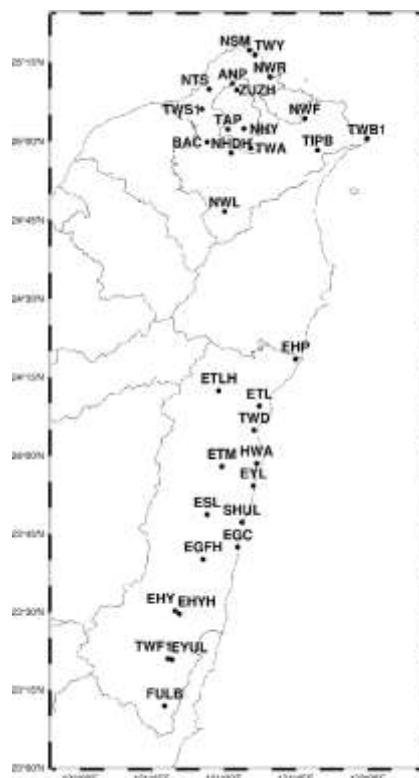


圖 2、測站分布圖

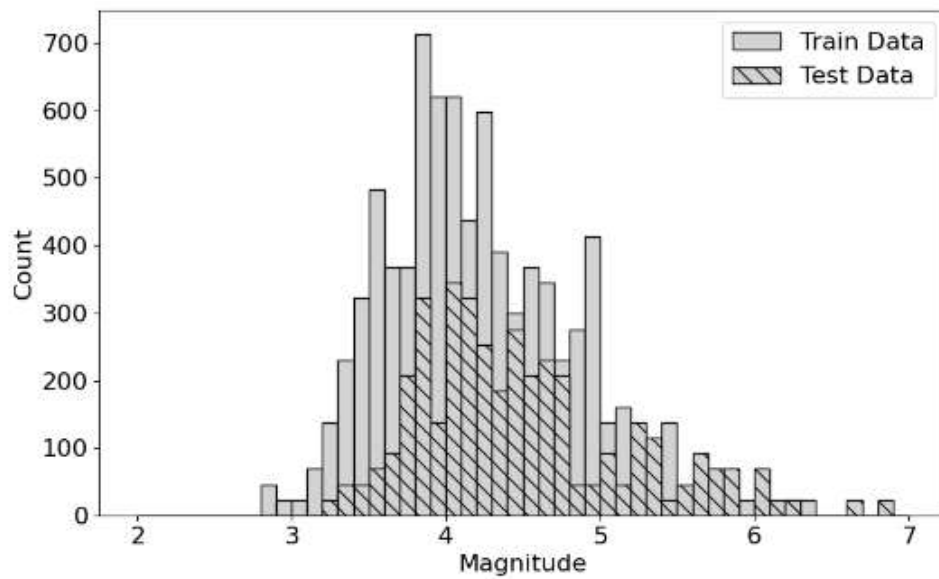


圖 3、訓練集和測試集規模分布圖

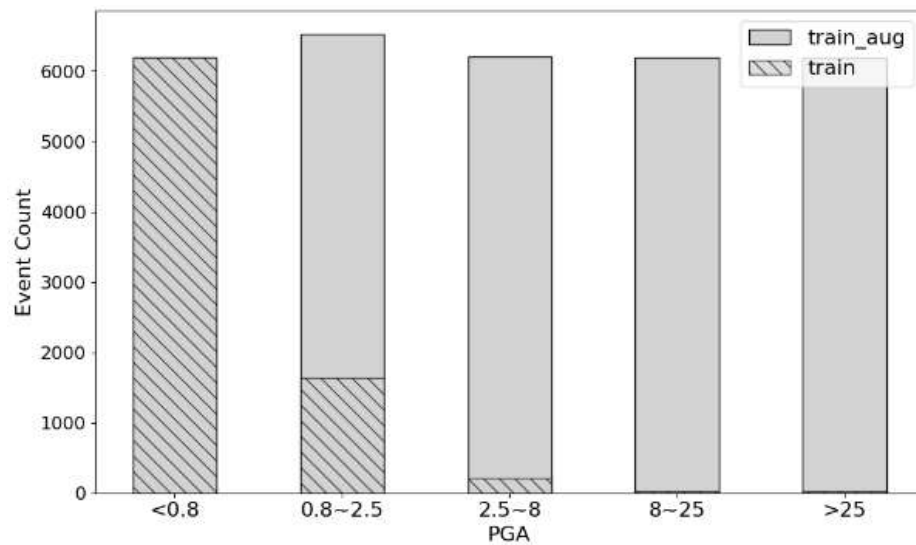


圖 4、原始資料集與數據增強後資料集 PGA 分布圖

2.5 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 with Weight Decay (AdamW)，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，並將權重衰減從梯度更新中分離出來，改進在正則化上的效果。學習率是設定為 10^{-5} ，批次大小為 32，總訓練次數設為 150 epoch，權重衰減設為 10^{-5} ，早停(Early Stopping)的耐心值(patience)設為 15，避免模型過擬合，損失函數為均方誤差(Mean Square Error, MSE)，其為實際值和預測值的平方差之平均，另外，為了使模型更關注於低估

部分的誤差，本研究採用了一種加權損失策略，即將低估的部分損失加倍處理。具體而言，當預測值低於真實值時，損失函數中的對應誤差項將乘以 2，從而增強模型對低估情況的懲罰力度。將偵測到的三軸加速度波形輸入到模型後進行訓練，模型會輸出目標測站的 PGA，並計算預測值和實際時的誤差，以此數據對模型進行更新和調整，優化預測結果。

2.6 具體成果

2.6.1 第 019 號花蓮地震事件預警測試

在此研究中，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，訓練臺北分區模型。選取 2024 年 4 月 3 日 7 時 58 分 9 秒，規模 7.2，深度 15.5 公里，震央位置位於北緯 23.86 度、東經 121.58 度的第 019 號花蓮地震。

本次研究從該測站 P 波到後的第三秒開始預測，預測結果如下圖 5。橫軸代表臺北地區的地震測站，排列順序是從北到南；縱軸則顯示預測與實際的 PGA 分布。空心圓表示實際的 PGA 值，而叉號則代表模型預測的 PGA 值。點線與虛線分別代表震度大於三級 ($PGA > 8$) 與震度大於四級 ($PGA > 25$) 的閾值 (threshold)。下圖 6、7、8、9 和 10，為第 5、10、15、20 和 25 秒的預測結果。隨著時間的推移，預測結果的準確度逐漸提高，並且在第 20 秒後預測值趨於穩定，且大部分測站皆有達到發布警報的閾值，但整體模型的預測還是偏低估。低估的原因可能是由於數據集中缺乏與這次大地震規模相當的事件。未來可以將此次事件加入訓練數據，以增強模型對大規模地震的預測能力。

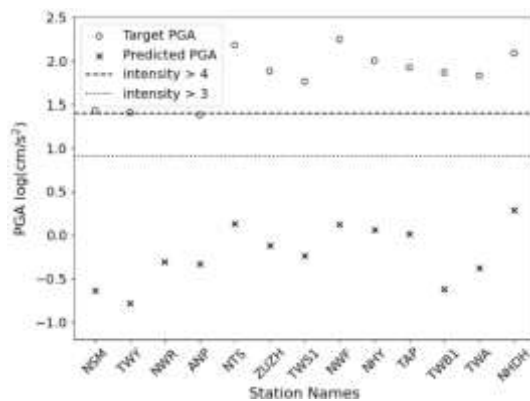


圖 5、第 3 秒預測結果

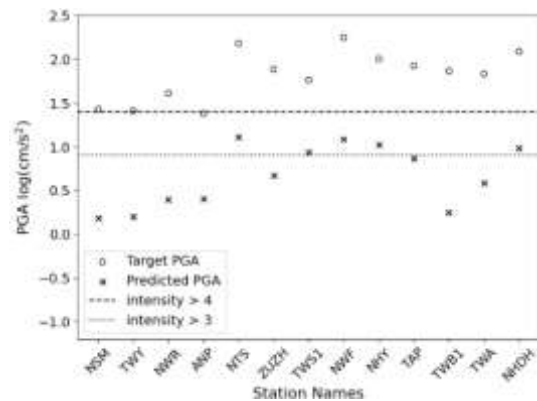


圖 6、第 5 秒預測結果

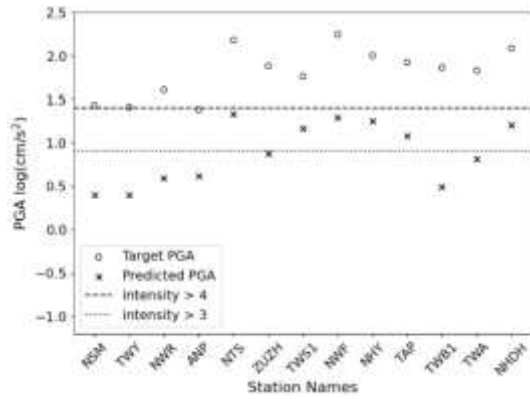


圖 7、第 10 秒預測結果

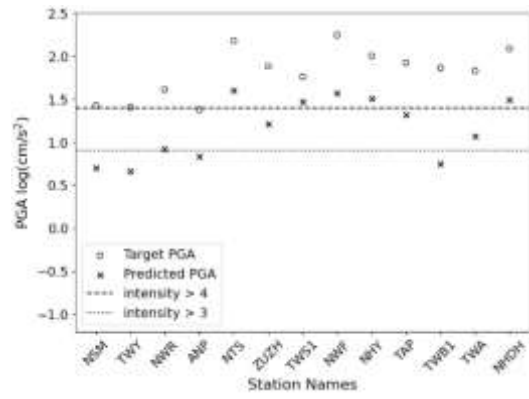


圖 8、第 15 秒預測結果

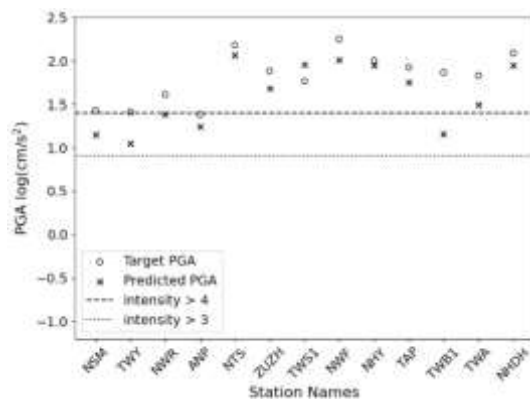


圖 9、第 20 秒預測結果

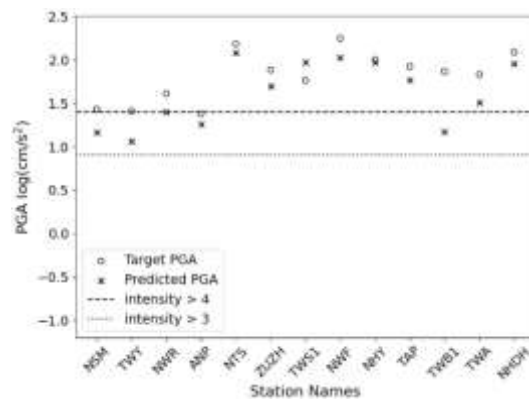


圖 10、第 25 秒預測結果

本次研究中，各測站的模型預測領先時間(leading time)如圖 11-23 所示，測站依照從北到南的順序排列。橫軸為時間軸，統一將第一個偵測到地震測站的 P 波到時設為 0，並且將每個測站的 PGA 大於閾值的時間點用垂直虛線表示。實心圓點為模型預測之 PGA 值，水平點線與水平虛線分別代表震度大於三級與震度大於四級的閾值。從圖中可以看出，靠近北邊的測站雖然模型預測的領先時間較長，但也更容易出現小於四級閾值的情況；相對地，靠近南邊的測站則有較短的領先時間，但在預測大於四級閾值的狀況下，表現得較好。當以震度大於三級為閾值時，平均領先時間為 16.68 秒；而以震度大於四級為閾值時，平均領先時間則為 18.32 秒。

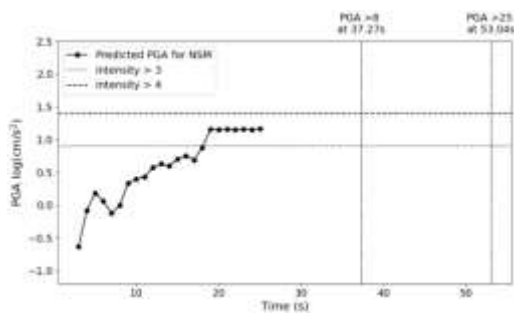


圖 11、NSM 測站預測

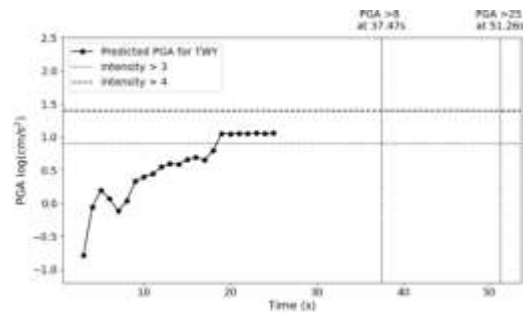


圖 12、TWY 測站預測

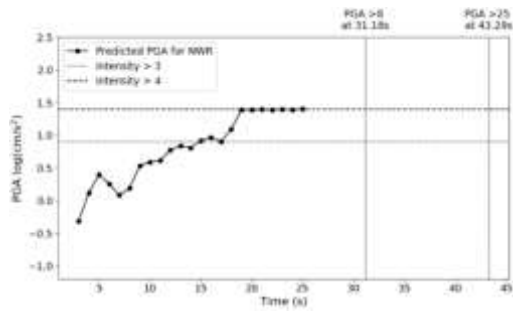


圖 13、NWR 測站預測

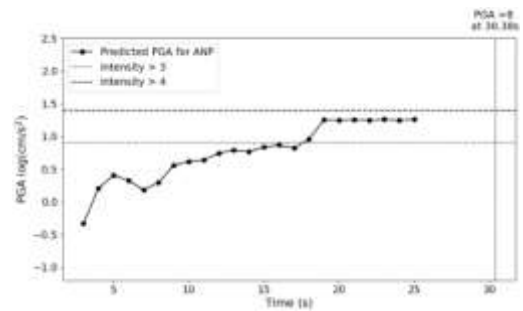


圖 14、ANP 測站預測

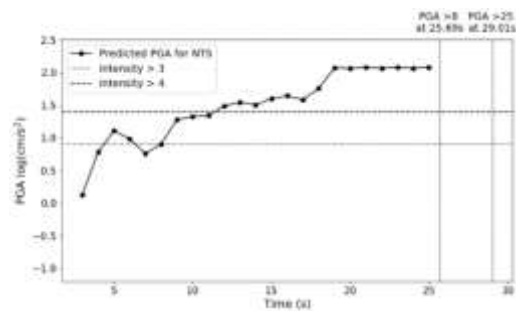


圖 15、NTS 測站預測

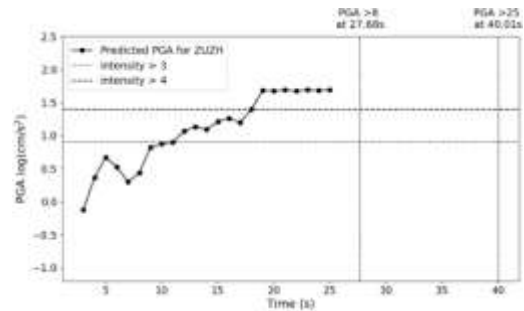


圖 16、ZUZH 測站預測

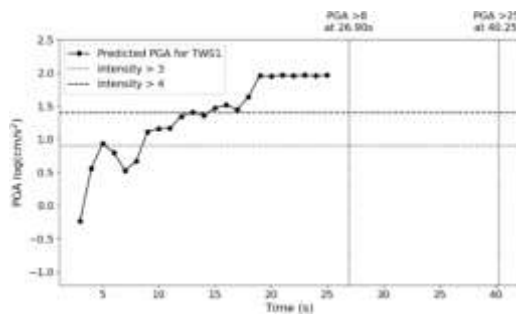


圖 17、TWS1 測站預測

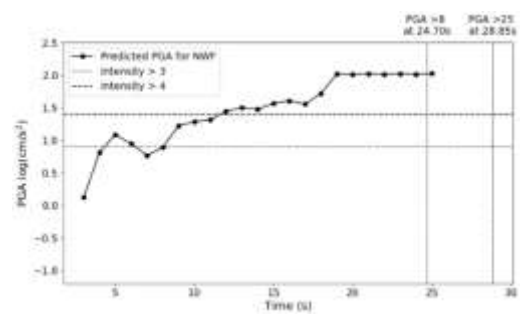


圖 18、NWF 測站預測

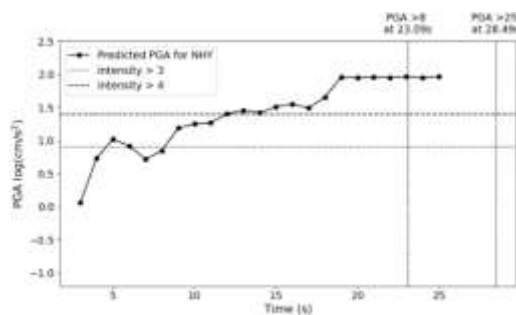


圖 19、NHY 測站預測

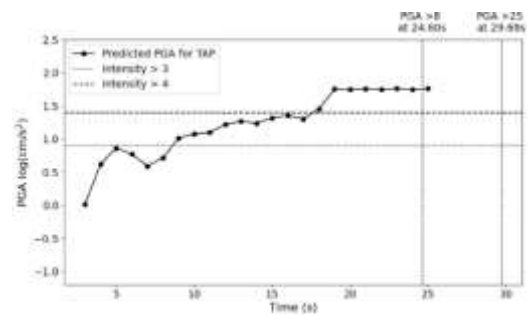


圖 20、TAP 測站預測

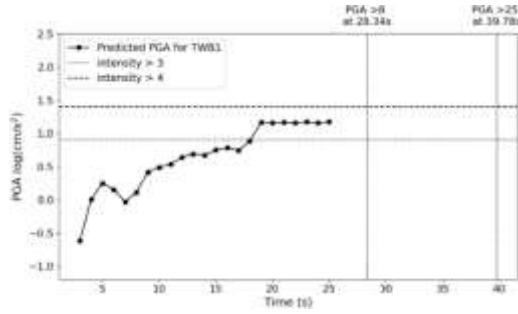


圖 21、TWB1 測站預測

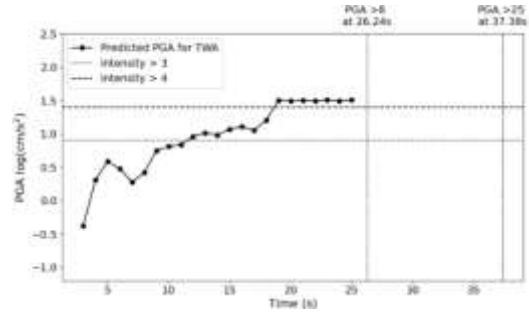


圖 22、TWA 測站預測

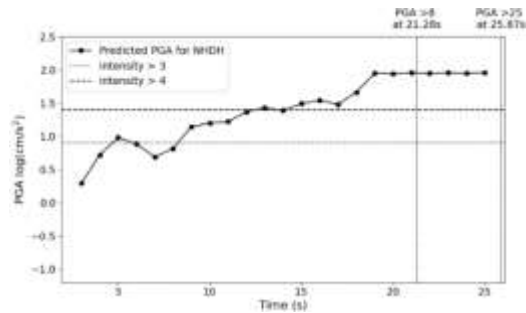


圖 23、NHDH 測站預測

2.7 模型評估

為了評估模型效能，本研究將模型在多視角圖注意力層(Multi-Perspective Graph Attention layer, MP-GAT)的部分替換成變換器層(Transformer layer)進行比較，並在第 3、5、10、15、20 和 25 秒時，呈現實際與預測值的分布，對應的結果如圖 24 到 29 所示。從結果可以看出，在前 5 秒，兩個模型都無法有效預測；但在第 10 秒後，MP-GAT 顯著收斂，並且其回歸線相比於 Transformer 更接近實際值。在均方誤差方面，MP-GAT 的值為 2.5638，而 Transformer 的值為 2.6912。由此可知，本次研究中，MP-GAT 模型的預測精度優於 Transformer 模型。

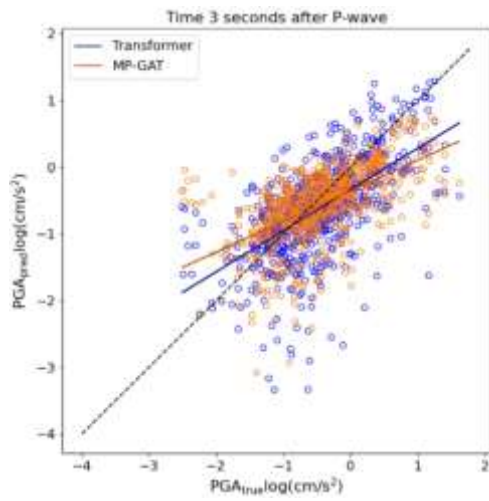


圖 24、第 3 秒預測結果

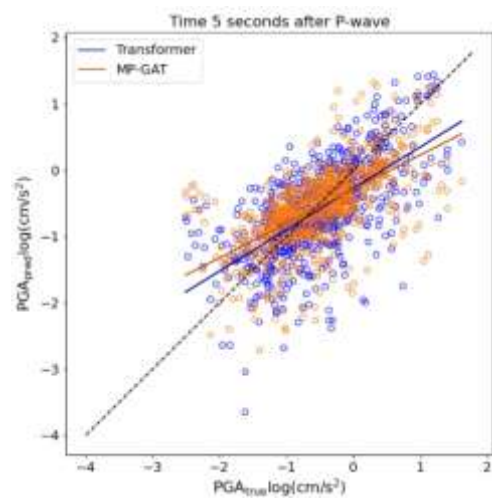


圖 25、第 5 秒預測結果

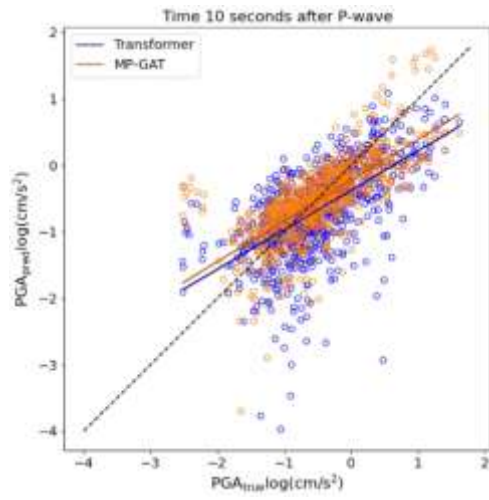


圖 26、第 10 秒預測結果

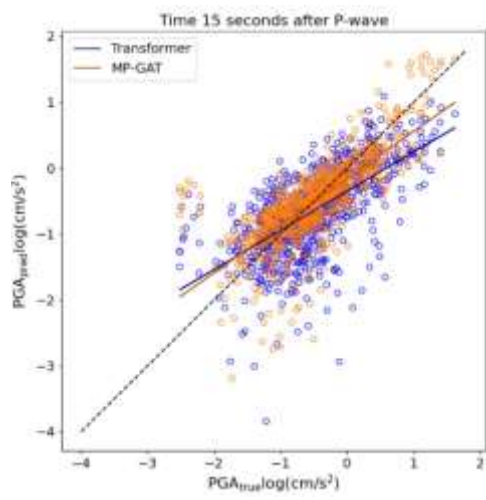


圖 27、第 15 秒預測結果

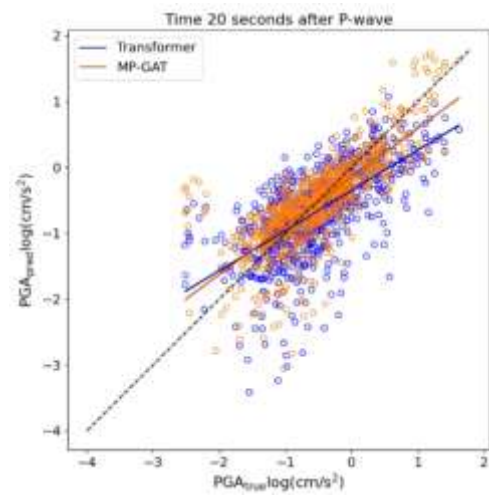


圖 28、第 20 秒預測結果

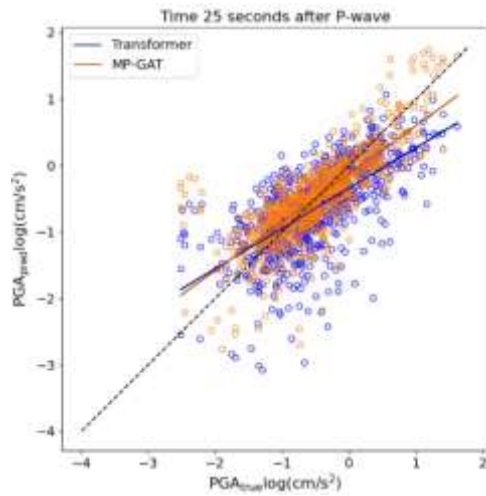


圖 29、第 25 秒預測結果

2.8 結論

本研究利用卷積神經網路提取三軸加速度波形的特徵，接著利用多視角圖注意力神經網路，根據測站間距離的遠近，賦予不同權重來聚合各測站的特徵，並對數據進行增強，以改善低估情況。

未來，我們將針對相同資料集進行不同分區模型的效果分析，探討造成模型效能差異的原因，並觀察是否與地區的地形結構等因素有關。近期地震發生較為頻繁，對於模型的訓練與測試提供了寶貴的數據。我們將進一步調整模型架構，使其更具泛化能力，提升預測的準確性和速度，並加強模型的預測能力。

3 類神經網路應用於區域型地震預警II

3.1 簡介

本研究針對特定的來源區域及目標區域進行預測，來源地區選擇地震頻繁的花蓮縣，臺北市和新北市則為目標地區。利用 P 波初達後 30 秒內來源地區中 16 個測站的波形資料以及測站所在之經緯度位置作為輸入資料，以卷積神經網路(CNN)擷取波形中的時間序列關係，以測站位置作為圖結構的節點、測站間的距離經正規化後作為邊權重，並結合圖神經網路(GNN)及 Transformer 中的注意力機制學習測站間的空間關係，進行目標地區 15 個位置的 PGA 預測任務，實現區域型地震預警系統。

在地震發生後，當震央附近偵測到第一個測站之 P 波時，以此時間為 T，擷取自 T 前 5 秒至 T 後 25 秒共 30 秒的時間窗，並利用滑動窗格的方式以 10 秒為窗格大小、1 秒為步模擬實際接收到波形的情況，使模型每秒都能夠預測出一個結果，每個事件有自 T 後 5 秒至 T 後 25 秒共 21 個結果。

在第 2 小節中講述資料篩選及處理的過程、第 3 小節中為模型架構細節、第 4 小節為訓練過程相關參數設定及訓練結果、第 5 小節中選擇了 2024 年 4 月 3 日的四個地震事件進行模型表現評估及分析，並且統計每個事件及每個測站的提前預警時間，第 6 小節則為本研究之結論及未來方向。

3.2 資料處理

本研究使用的資料集來自中央氣象署 CWASN 觀測網，以測站作為觀測位置。由於本研究需要針對特定來源地區和目標地區進行地震預測，因此選定地震頻繁的花蓮縣作為來源地區，並以臺北市和新北市（首都圈）為目標地區。測站分布如圖 30 所示，來源地區的測站共有 16 個，測站位置以圓表示；目標地區的測站共有 15 個，測站位置以三角形表示，並附有對應之測站編號。

地震事件的資料範圍涵蓋 2012 年 1 月 1 日至 2023 年 5 月 20 日，共計 6791 筆有感地震紀錄。本研究進一步篩選來源地區發生的事件，篩選條件是確認第一個偵測到 P 波的測站位置是否位於來源地區。經過篩選後，最終得到 4404 筆地震事件，位置分布如圖 31 所示；其中，圓圈大小代表地震規模，顏色深淺則表示地震深度。

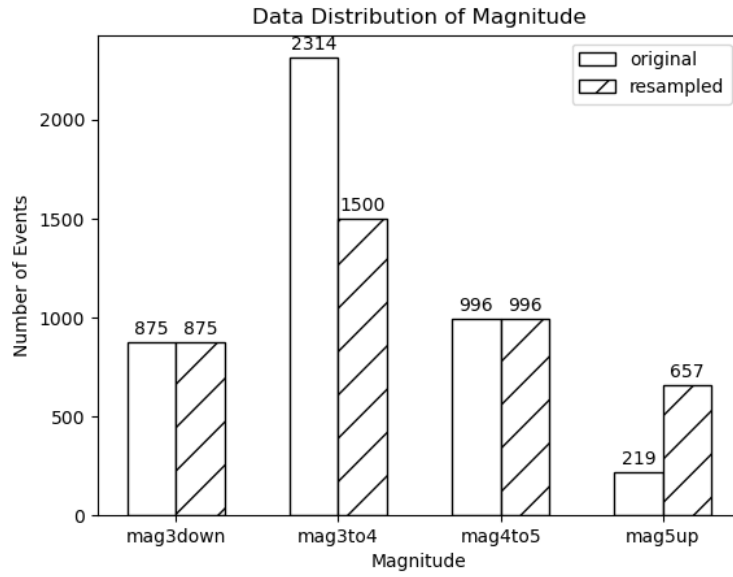


圖 32、數據分布圖

3.3 模型架構

本研究結合圖神經網路和 Transformer 模型，將注意力機制應用於圖結構資料中。圖結構由節點與邊構成，以測站位置為節點，測站間的距離經正規化後作為邊權重，使模型能學習空間上的關係。為提升模型學習能力，本研究對 PGA 取對數作為輸出，以縮小數值範圍。

如圖 33 所示，模型架構分為四個部分：輸入層 (Input)、特徵提取層 (Feature Extraction)、特徵合併層 (Feature Combination) 以及預測輸出層 (Output)。首先，輸入層包括來源地區與目標地區的特徵向量及圖結構的鄰接矩陣。對於來源地區，輸入三軸波形和經緯度位置，以第一個測站接收到 P 波的時間 T 為基準，擷取 T 前 5 秒至 T 後 25 秒（共 30 秒）的波形數據，如圖 34 所示，虛線框為滑動窗格，並以 10 秒為窗格大小、1 秒為步模擬實際接收到波形的情况。如此一來，每個事件依時間進行預測，自 T 後 5 秒至 T 後 25 秒共產生 21 個預測結果；而目標地區則僅輸入對應測站的經緯度。鄰接矩陣大小為 $(N \times N, 1)$ ，其中 N 為來源和目標的節點總數，矩陣中的元素代表測站間的絕對距離 D，並將 D 正規化後乘上負號如式(12)，

$$d_{(i,j)} = -(D_{(i,j)} \div 100) \quad (\text{式 } 12)$$

以使距離近的測站權重大、距離遠的權重小。

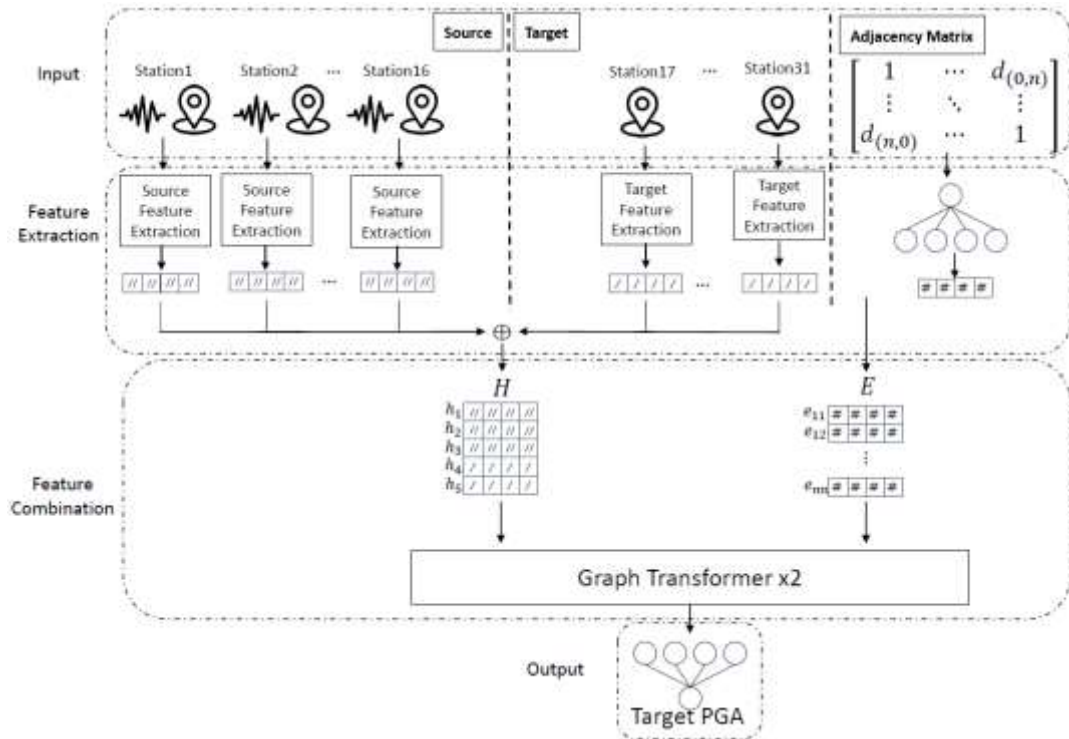


圖 33、模型架構圖

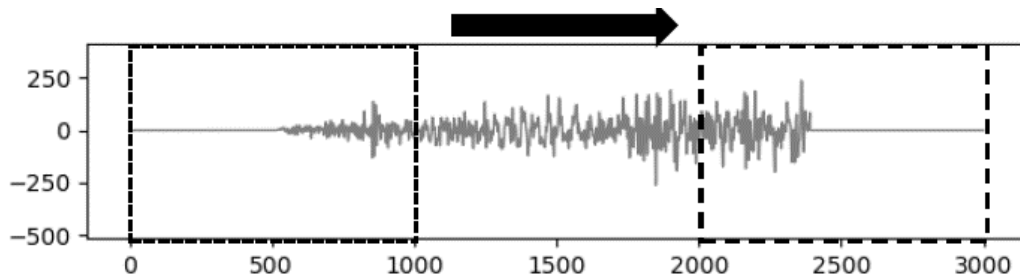


圖 34、輸入波形示意圖

在特徵提取層中，對每個測站的波形資訊、經緯度資訊及鄰接矩陣進行處理。來源地區之輸入資料會進入 Source Feature Extraction Block 如圖 35、目標地區資料則進入 Target Feature Extraction Block 如圖 36。對於波形資訊會經過 2 層二維卷積神經網路(Convolutional Neural Network, CNN)和 5 層一維 CNN 提取時間序列的特徵向量，每層 CNN 的 kernel size、stride、filter 參數如表 1 所示，每層 CNN 後都會接整流線性單位函數(Rectified Linear Unit, ReLU)，在前三層一維 CNN 會接最大池化層(Max pooling)保留重要的特徵同時減少計算量。特徵經攤平(Flatten)後再經過 2 層全連接層(Fully Connected Layer)進行線性轉換，最終形成長度為 400 的特徵向量。經緯度資訊則是使用 Positional Embedding，將 (緯度:1, 經度:1)擴增成(緯度 sin:100, 緯度 cos:100, 經度 sin:100, 經度 cos:100)總長度為 400 的位置特徵向量。在 Source Feature Extraction Block 將提取完的波形特徵以及位置特徵相加，得到每個測站長度為 400 的混合特徵向量。Target

Feature Extraction Block 僅將經緯度資訊經 Positional Embedding 生成位置特徵向量，鄰接矩陣則經過全連接層轉換成長度為 400 的邊特徵。如圖 37 所示，來源地區的混和特徵向量(以//表示)及目標地區的位置特徵向量(以/表示)進行合併(Concatenate)的處理，並作為圖結構的節點特徵 H 與邊特徵 E (以#表示)，將兩特徵共同輸入模型，其中 $H = \{h_1, h_2, \dots\}$ 、 $E = \{e_{11}, e_{12}, \dots\}$ 。

表 1、CNN 參數

CNN layer	In channel	Out channel	Kernel size	Stride
2dConv1	1	8	(5, 1)	(5, 1)
2dConv2	8	32	(16, 3)	(1, 3)
1dConv1	32	64	16	1
1dConv2	64	128	16	1
1dConv3	128	32	8	1
1dConv4	32	32	8	1
1dConv5	32	16	4	1

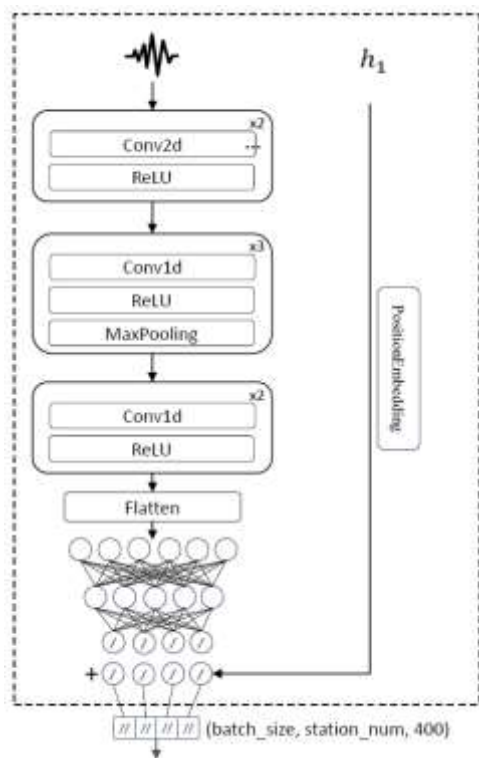


圖 35、Source Feature Extraction

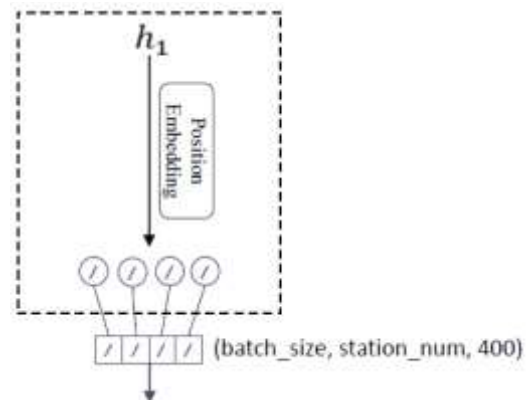


圖 36、Target Feature Extraction

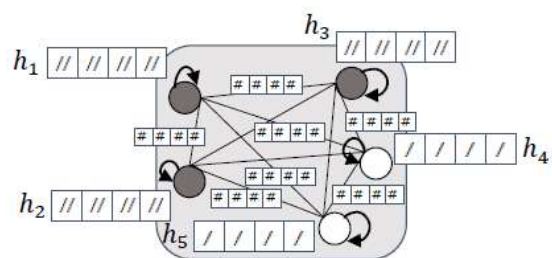


圖 37、Graph structure

特徵合併層為主要模型層，由 2 層 Graph Transformer 組成 [4]，Graph Transformer 和原始 Transformer [5] 同樣採用注意力機制(Attention Mechanism)。原始 Transformer 應用於自然語言處理(Natural language processing, NLP)，透過 Query (式 13)、Key (式 14)和 Value (式 15)向量計算句詞間的注意力分數以識別重要的句詞 (式 16)， W_Q 、 W_K 、 W_V 為模型訓練的權重矩陣， X 為輸入特徵。

$$Q = W_Q \cdot X \quad (\text{式 13})$$

$$K = W_K \cdot X \quad (\text{式 14})$$

$$V = W_V \cdot X \quad (\text{式 15})$$

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V \quad (\text{式 16})$$

Graph Transformer 則將注意力機制延伸至圖結構數據中，Query (式 17)、Key (式 18)、Value (式 19)由節點特徵計算而來，並且在注意力分數計算過程中加入 Edge (式 20)用來加強節點間的關聯性，其由邊特徵計算得，注意力分數 α 公式如式(21)，適用於圖形數據的特徵傳播與學習。

$$q_i = W_q \cdot h_i + b_q, Q = \{q_0, q_1, \dots, q_N\} \quad (\text{式 17})$$

$$k_i = W_k \cdot h_i + b_k, K = \{k_0, k_1, \dots, k_N\} \quad (\text{式 18})$$

$$v_i = W_v \cdot h_i + b_v, V = \{v_0, v_1, \dots, v_N\} \quad (\text{式 19})$$

$$e_{ij}' = W_e \cdot e_{ij} + b_e, E' = \{e_{00}', e_{01}', \dots, e_{NN}'\} \quad (\text{式 20})$$

$$\alpha_{ij} = \frac{\langle q_i, k_j + e_{ij}' \rangle}{\sum_{u \in N(i)} \langle q_i, k_u + e_{iu}' \rangle} \quad (\text{式 21})$$

Graph Transformer 在保留了圖神經網路(GNN) 中消息傳遞 (Message Passing) 特性的同時，引入了注意力機制來精確地建模節點之間的相互依賴關係。圖 38 為本研究特徵合併層及預測輸出層，在特徵合併層的每層中，會針對每個節點取其特徵向量(h_i)計算與鄰居間的注意力分數(α)後，節點特徵會基於鄰居節點的資訊進行聚合 (Aggregation) 與更新 (Update)以此更新每個節點的特徵，節點更新公式如式(22)，

$$h_i' = \sum_{j \in N(i)} \alpha_{ij} * (v_j + e_{ij}') \quad (\text{式 22})$$

Graph Transformer 結合邊特徵可以更好地表達節點間的結構化關聯。注意力機制不僅能自適應地調整不同節點間的權重，還能根據圖結構動態更新節點特徵，

使模型能夠更精細地捕捉複雜的空間依賴性。該模型同時整合了時間與空間上的資訊，從而在節點間實現高效的信息交流，充分利用測站間的距離與位置資訊，顯著提升地震 PGA 預測的準確性。最後在預測輸出層中，將每個節點的節點特徵經過全連接層輸出預測值。

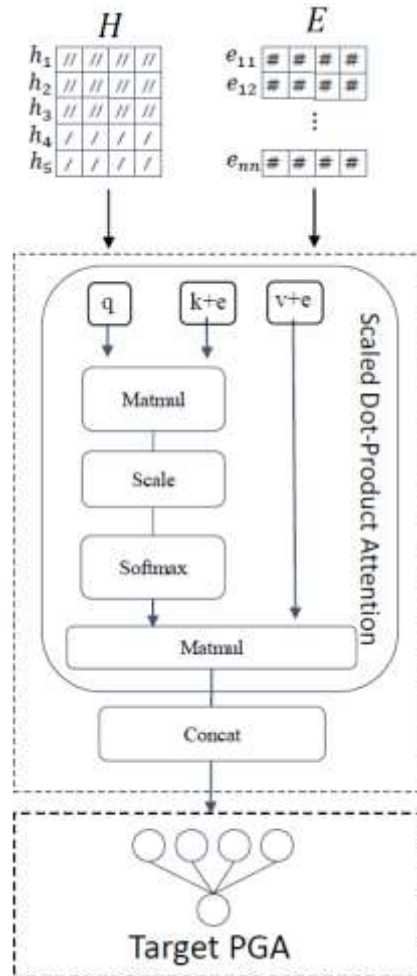


圖 38、Graph Transformer 及 Output

3.4 訓練過程

在訓練過程中，本研究採用改進型的 AdamW 優化器(Adaptive Moment Estimation with Weight Decay Fix) [6]，透過權重衰減(Weight Decay)而非傳統 L2 正則化來控制模型參數，以改善不穩定的權重更新問題。初始學習率設定為 10-6，batch size 設為 32，訓練 200 個 epoch，並加入早停條件(Early Stopping)以避免過擬合，其中 patience 設為 15。訓練損失函數使用均方誤差(Mean Square Error, MSE)，該函數計算預測值與實際值的平方差平均值。

訓練流程包括將波形、位置和圖結構資訊作為模型輸入，進行訓練並輸出目標位置的 PGA 預測對數值。隨後，根據損失函數進行反向傳播，以調整模型

參數並優化預測結果。訓練及驗證過程中的損失變化如圖 39 所示，實線為訓練損失、虛線為驗證損失，最終模型達到收斂狀態。

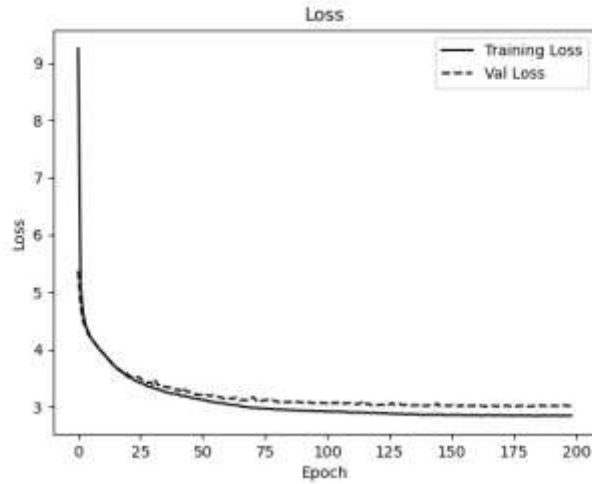


圖 39、訓練及驗證損失變化

3.5 具體成果

本研究以 2024 年 4 月 3 日發生的四個地震事件評估模型表現並統計提前預警時間(Leading time)，編號/芮氏規模分別為第 19 號/7.2、第 26 號/5.7、第 45 號/6.2、第 71 號/4.8 之地震事件。此研究中預測值皆為 PGA 之對數值，表 2 為各震級 PGA 閾值，預測結果之秒數皆以來源地區第一個測站偵測到 P 波的時間 T 以後之秒數。另外，本研究還統計了每事件每測站的提前預警時間，其定義如式(23)，

$$\text{Leading time} = t_{\text{true}} - t_{\text{pred}} \quad (\text{式 } 23)$$

t_{true} 為真實波超過各閾值的時間、 t_{pred} 為預測值最早超過各閾值的時間。

表 2、各震級 PGA 閾值

	1 級	2 級	3 級	4 級
$PGA \text{ (cm/sec}^2\text{)}$	0.8	2.5	8	25
$\log_{10} PGA$	-0.097	0.398	0.903	1.398

3.5.1 第 19 號花蓮地震事件預警測試

在此實驗中，我們選取 2024 年 4 月 3 日 7 時 58 分 9 秒，芮氏規模 7.2，深

度 22.5 公里的第 19 號地震事件，圖 40、圖 41、圖 42、圖 43、圖 44 分別代表 T 後 5、10、15、20、25 之預測結果，以三角形為預測值、圓為真實值、以實線為 1 級、以-線為 2 級、以--線為 3 級、以·線為 4 級，若真實值為空則以-3 表示。在此事件中真實值中有 3 個測站為空、10 個測站超過 4 級；預測值大部分皆有超過 2 級，不過在預測結果最大為 3 級，因此在此事件中有低估的情況。由於此事件為少數的大規模事件，儘管有對規模 5 以上之數據進行過採樣的處理，大規模數據仍舊占少數，這也是造成此事件低估的原因。

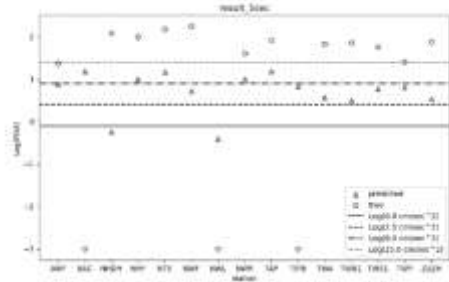


圖 40、T 後 5 秒預測/真實值

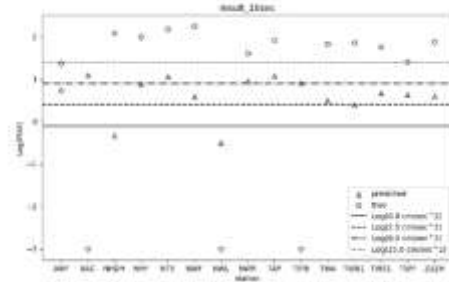


圖 41、T 後 10 秒預測/真實值

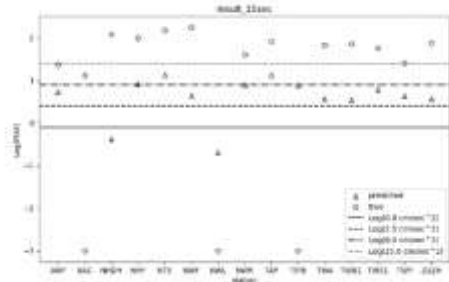


圖 42、T 後 15 秒預測/真實值

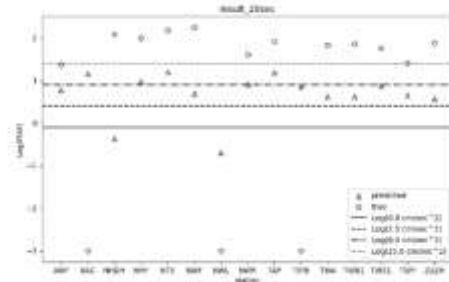


圖 43、T 後 20 秒預測/真實值

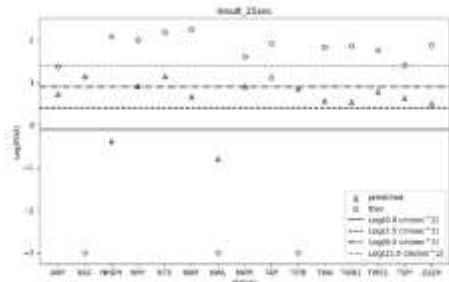


圖 44、T 後 25 秒預測/真實值

圖 45、圖 46、圖 47、圖 48、圖 49 及圖 50 分別代表 T 後 5、10、15、20、25 及真實之雙北地區預測結果。圖中的星號為震央、黑圓點為 P 波已達的測站及圓點旁的數字為 P 波到達的順序、三角形為目標位置。可由顏色看出在 T 後 5 秒 25°00'N 以北呈現灰色較深的狀態，這代表雙北地區已達 2 至 3 級的標準。

和真實結果相比，由於有 3 個測站真實值為空，因此在緯度區間[24°45'N, 25°00'N]經度區間[121°30'E, 122°00'E]有較明顯的差異。

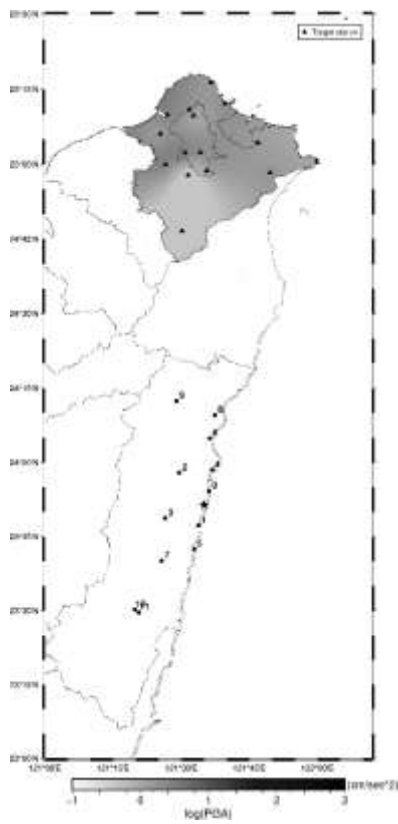


圖 45、T 後 5 秒地區預測結果

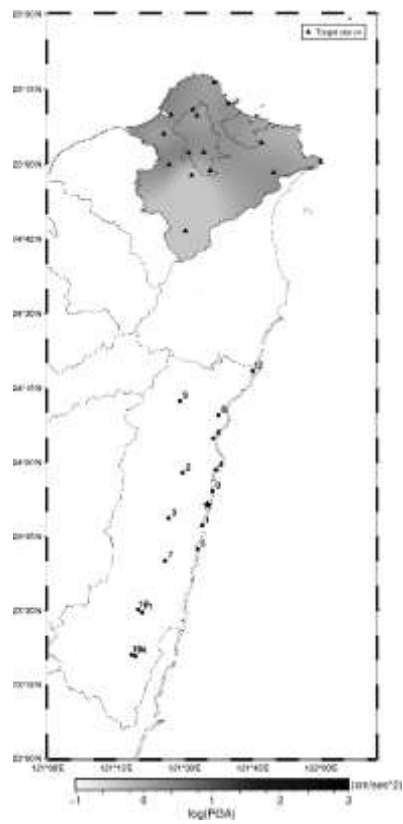


圖 46、T 後 10 秒地區預測結果

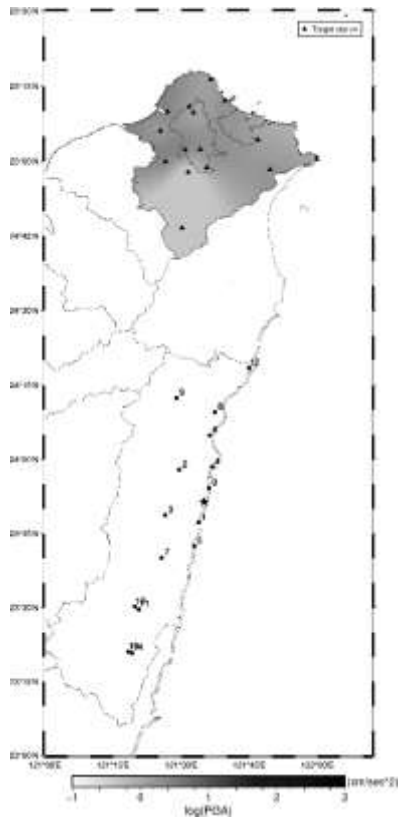


圖 47、T 後 15 秒地區預測結果

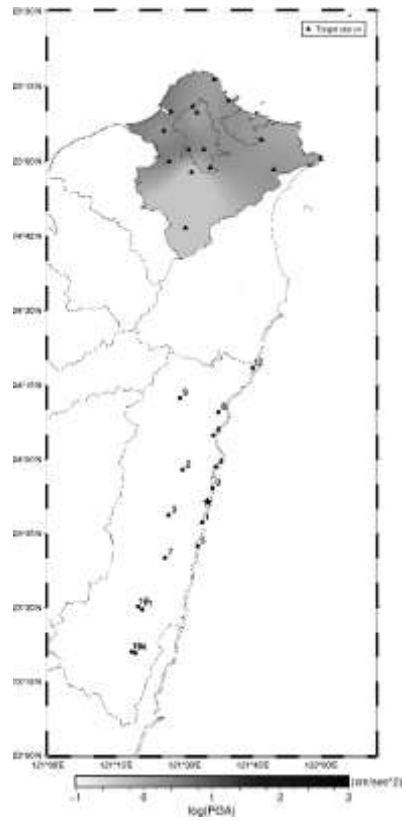


圖 48、T 後 20 秒地區預測結果

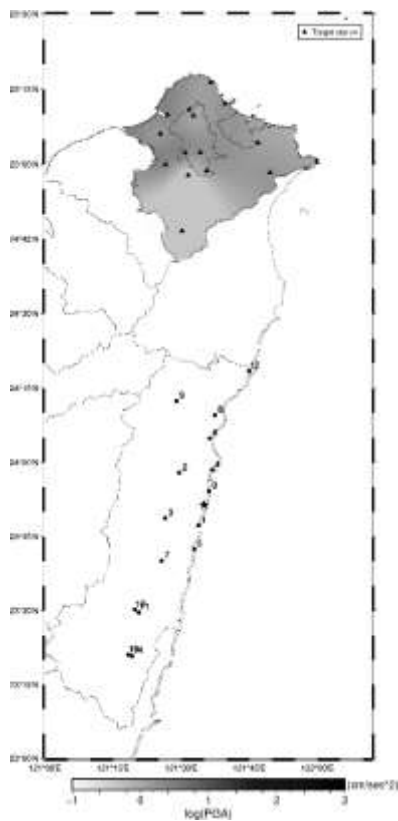


圖 49、T 後 25 秒地區預測結果

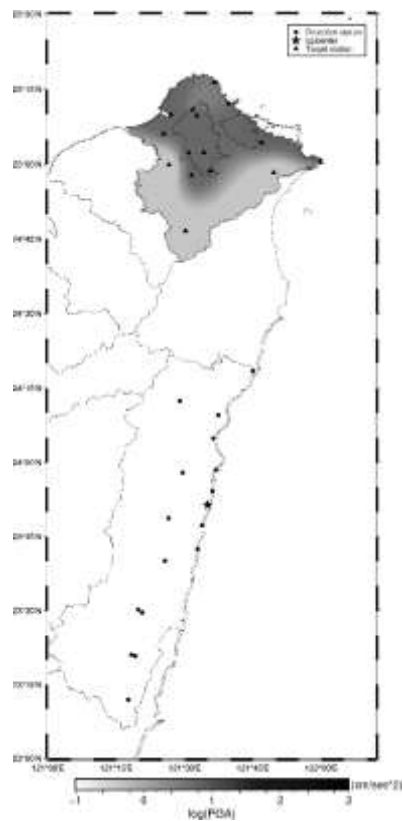


圖 50、地區真實結果

另外，表 3 統計了此事件各測站之提前預警時間，表中第一欄分別為各測站名稱、count 為各震級提前預警之測站數量。由統計結果得知，在此事件中共有 11 個測站預測超過 1 級、11 個測站預測超過 2 級、4 個測站超過 3 級，且皆至少有 15 秒的提前預警時間。

表 3、第 19 號提前預警時間(秒)

	>1 級	>2 級	>3 級	>4 級		>1 級	>2 級	>3 級	>4 級
ANP	20	23	0	0	TAP	16	18	22	0
BAC	0	0	0	0	TIPB	0	0	0	0
NHDH	0	0	0	0	TWA	16	18	0	0
NHY	15	17	19	0	TWB1	18	20	0	0
NTS	17	19	21	0	TWS1	18	20	0	0
NWF	17	18	0	0	TWY	23	26	0	0
NWL	0	0	0	0	ZUZH	18	20	0	0
NWR	21	23	27	0	count	11	11	4	0

3.5.2 第 26 號花蓮地震事件預警測試

在此實驗中，我們選取 2024 年 4 月 3 日 8 時 35 分 36 秒，芮氏規模 5.7，深度 7.3 公里的第 26 號地震事件，圖 51、圖 52、圖 53、圖 54、圖 55 分別代表 T 後 5、10、15、20、25 之預測結果。在此事件中真實值中有 3 個測站為空、大部分測站皆在 1 至 3 級之間；預測值大部分皆有超過 1 至 2 級，且預測及真實差異也大幅縮小，在 T 後 15 秒時有 8 個位置預測及真實值差異小於一個震級。低估的情況在此事件中有明顯改善，且預測結果也隨時間推移愈加精準，至 T 後 15 秒達最佳預測。

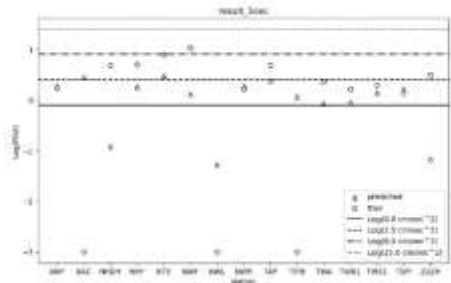


圖 51、T 後 5 秒預測/真實值

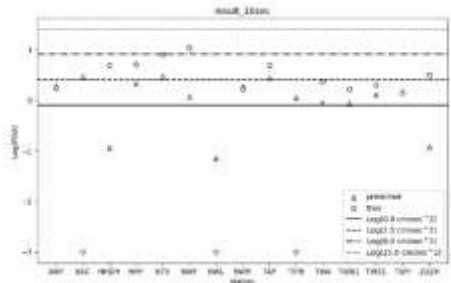


圖 52、T 後 10 秒預測/真實值

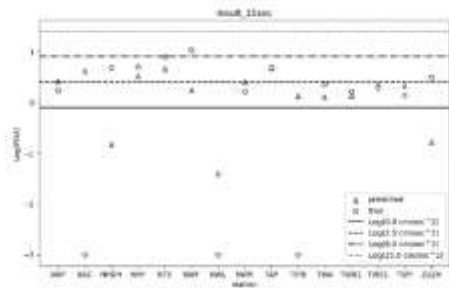


圖 53、T 後 15 秒預測/真實值

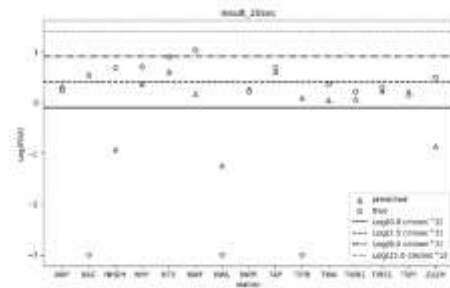


圖 54、T 後 20 秒預測/真實值

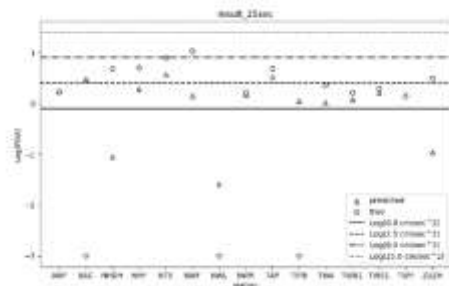


圖 55、T 後 25 秒預測/真實值

圖 56、圖 57、圖 58、圖 59、圖 60 及圖 61 分別代表 T 後 5、10、15、20、25 及真實之雙北地區預測結果。可由顏色看出在 T 後 5 秒在緯度區間 $[25^{\circ}00'N, 25^{\circ}20'N]$ 灰色稍微深一點，即預測值超過 1 級的情況，隨時間推進至 T 後 15 秒時， $25^{\circ}00'N$ 以北灰色較深的情況，此時最接近真實值。

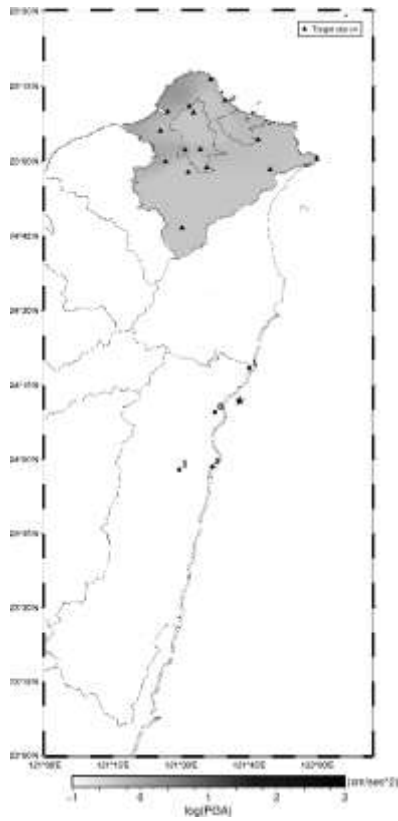


圖 56、T 後 5 秒地區預測結果

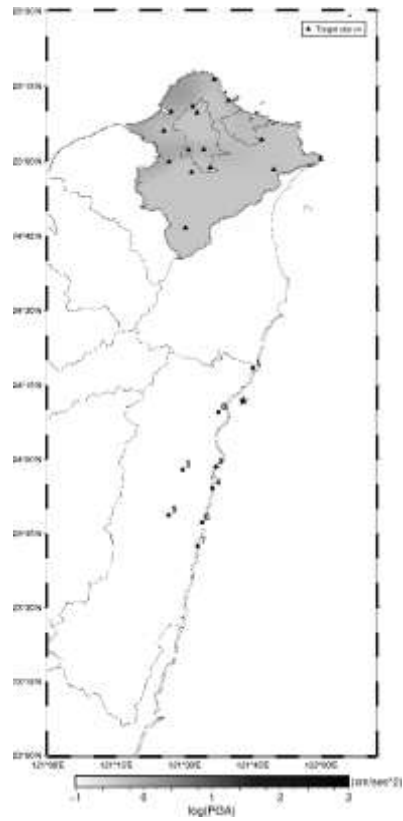


圖 57、T 後 10 秒地區預測結果

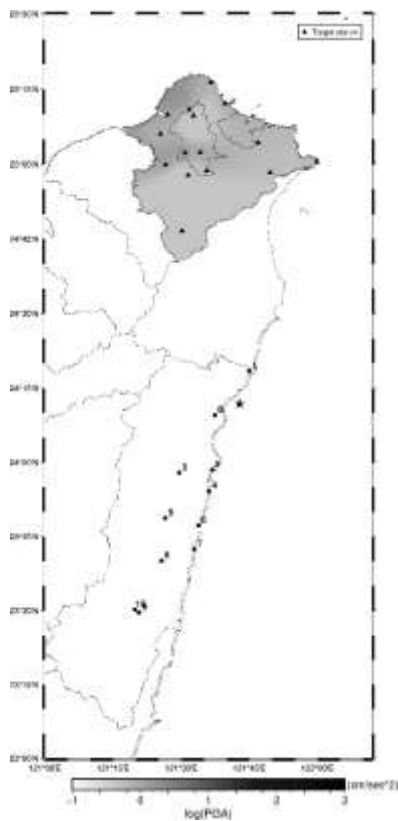


圖 58、T 後 15 秒地區預測結果

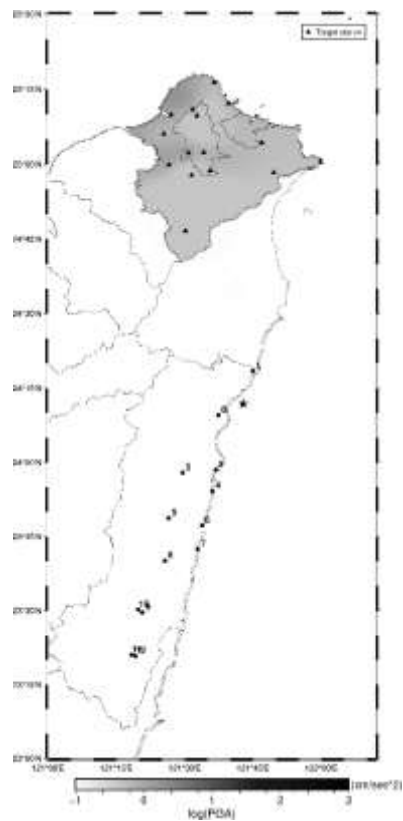


圖 59、T 後 20 秒地區預測結果

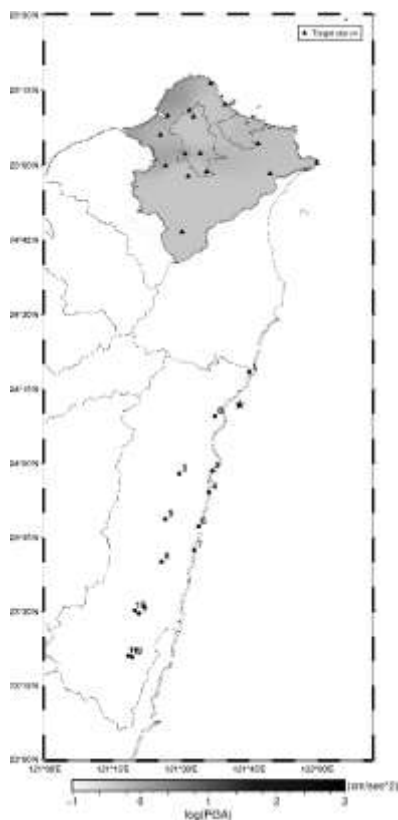


圖 60、T 後 25 秒地區預測結果

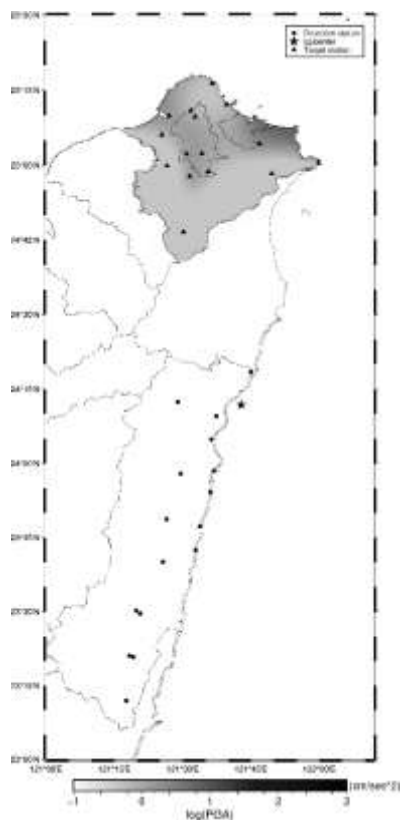


圖 61、地區真實結果

另外，表 4 統計了此事件各測站之提前預警時間。由統計結果得知，在此事件中共有 10 個測站預測超過 1 級、有 3 個測站預測超過 2 級，且皆有至少 9 秒的提前預警時間。

表 4、第 26 號提前預警時間(秒)

	>1 級	>2 級	>3 級	>4 級		>1 級	>2 級	>3 級	>4 級
ANP	19	0	0	0	TAP	11	0	0	0
BAC	0	0	0	0	TIPB	0	0	0	0
NHDH	0	0	0	0	TWA	16	0	0	0
NHY	11	0	0	0	TWB1	23	0	0	0
NTS	14	9	0	0	TWS1	15	0	0	0
NWF	12	30	0	0	TWY	35	0	0	0
NWL	0	23	0	0	ZUZH	0	0	0	0
NWR	17	0	0	0	count	10	3	0	0

3.5.3 第 45 號花蓮地震事件預警測試

在此實驗中，我們選取 2024 年 4 月 3 日 10 時 14 分 34 秒，芮氏規模 6.2，深度 26.9 公里的第 45 號地震事件，圖 62、圖 63、圖 64、圖 65、圖 66 分別代表 T 後 5、10、15、20、25 之預測結果。在此事件中真實值中有 5 個測站為空、大部分測站皆在 2 至 3 級之間；預測值大部分皆有超過 1 至 2 級，在 T 後 10 秒至 15 秒時有 5 個位置預測及真實值差異小於一個震級。在此事件中大部分測站在 T 後 5 秒已預測超過 1 至 2 級，評估結果和第 26 號事件相似，低估的情況和第 19 號相比有所改善，且預測結果也隨時間推移愈加精準，至 T 後 10 至 15 秒達最佳預測。

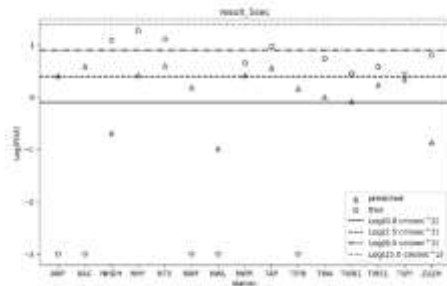


圖 62、T 後 5 秒預測/真實值

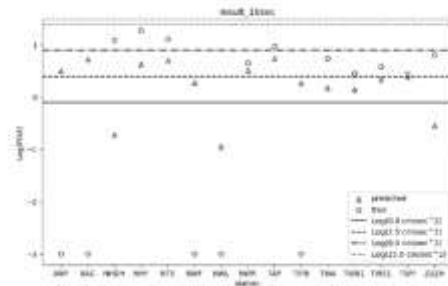


圖 63、T 後 10 秒預測/真實值

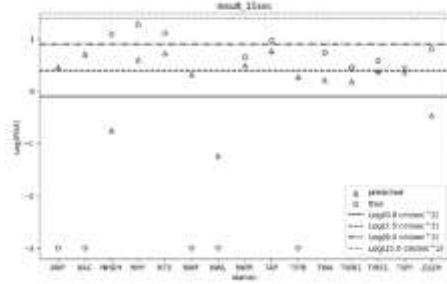


圖 64、T 後 15 秒預測/真實值

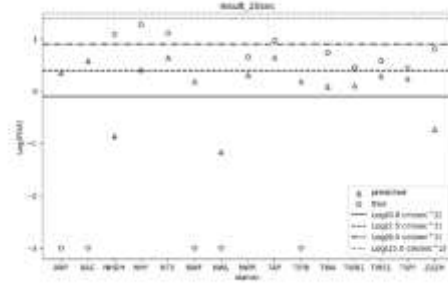


圖 65、T 後 20 秒預測/真實值

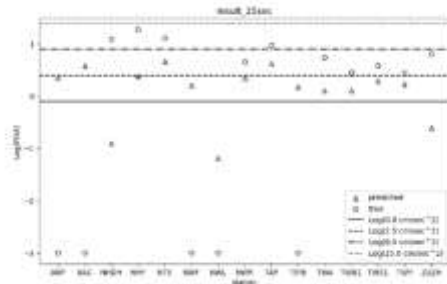


圖 66、T 後 25 秒預測/真實值

圖 67、圖 68、圖 69、圖 70、圖 71 及圖 72 分別代表 T 後 5、10、15、20、25 及真實之雙北地區預測結果。可由顏色看出在 T 後 5 秒緯度區間 $[24^{\circ}55'N, 25^{\circ}20'N]$ 灰色有稍微深一點，即預測值超過 1 級的情況，隨時間推進至 T 後 10 至 15 秒時， $24^{\circ}55'N$ 以北皆有較深灰色，此時最接近真實值，且此事件有預測結果之區域和真實區域較相似。

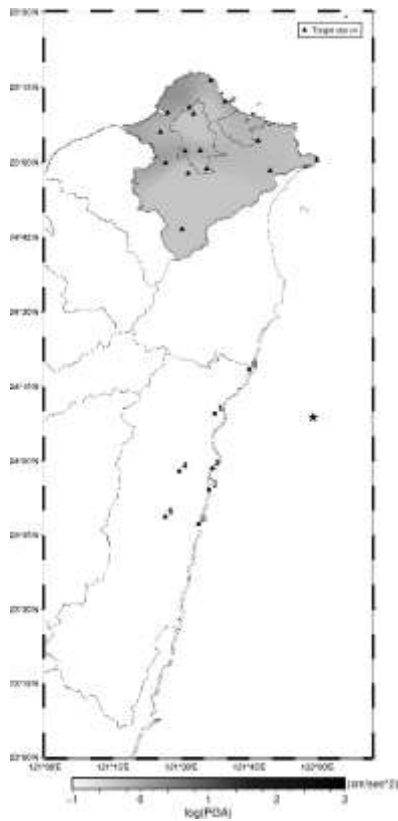


圖 67、T 後 5 秒地區預測結果

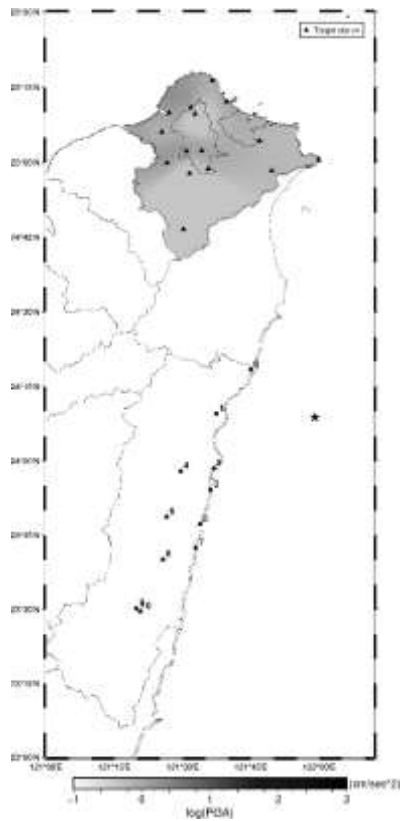


圖 68、T 後 10 秒地區預測結果

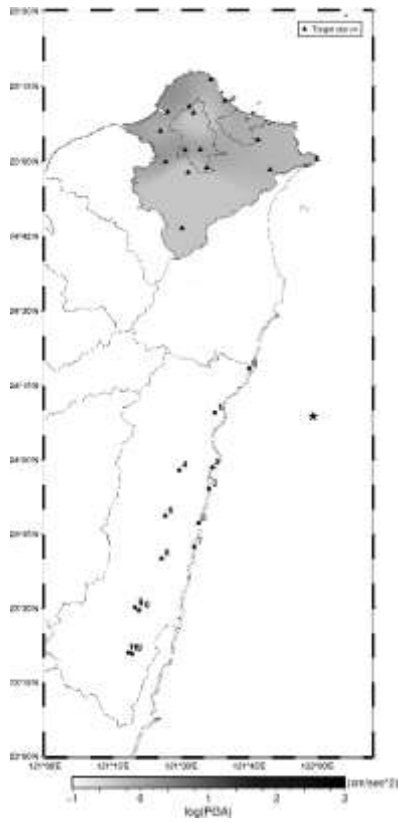


圖 69、T 後 15 秒地區預測結果

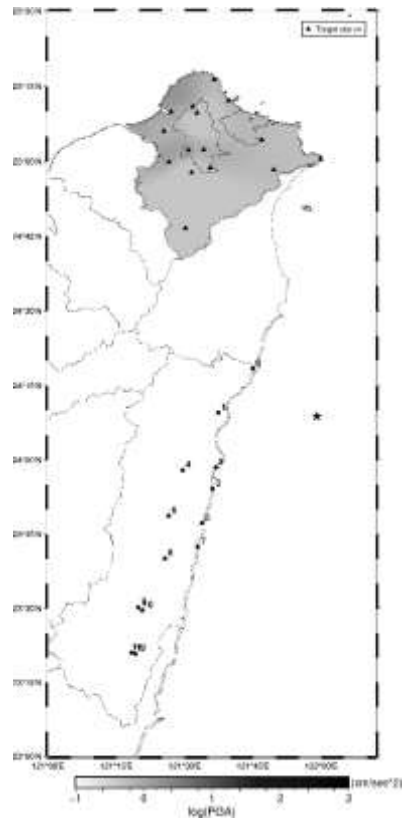


圖 70、T 後 20 秒地區預測結果

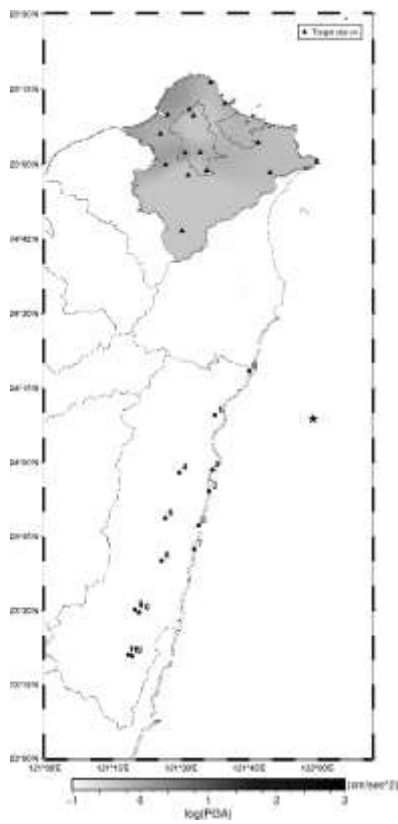


圖 71、T 後 25 秒地區預測結果

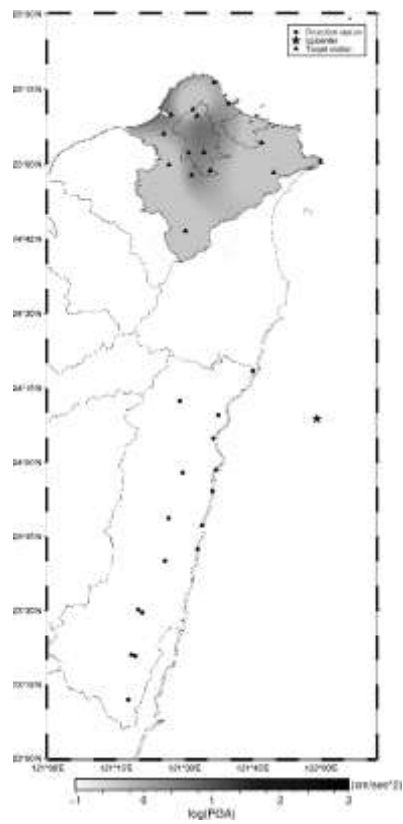


圖 72、地區真實結果

另外，表 5 統計了此事件各測站之提前預警時間。由統計結果得知，在此事件中共有 8 個測站預測超過 1 級、有 6 個測站預測超過 2 級，且皆有至少超過 9 秒的提前預警時間。

表 5、第 45 號提前預警時間(秒)

	>1 級	>2 級	>3 級	>4 級		>1 級	>2 級	>3 級	>4 級
ANP	0	0	0	0	TAP	11	15	0	0
BAC	0	0	0	0	TIPB	0	0	0	0
NHDH	0	0	0	0	TWA	10	0	0	0
NHY	10	10	0	0	TWB1	9	0	0	0
NTS	14	17	0	0	TWS1	13	18	0	0
NWF	0	0	0	0	TWY	17	31	0	0
NWL	0	0	0	0	ZUZH	0	0	0	0
NWR	14	26	0	0	count	8	6	0	0

3.5.4 第 71 號花蓮地震事件預警測試

在此實驗中，我們選取 2024 年 4 月 3 日 14 時 33 分 57 秒，芮氏規模 4.8，深度 5 公里的第 71 號地震事件，圖 73、圖 74、圖 75、圖 76、圖 77 分別代表 T 後 5、10、15、20、25 之預測結果。在此事件中真實值中有 4 個測站為空、1 個測站超過 1 級、大部分測站皆在 0 至 1 級；預測值在 T 後 5 秒大部分位置皆有超過 1 級，在 T 後 20 至 25 秒時有 5 個位置預測及真實值差異小於一個震級，整體預測以及真實差異較其他 3 個事件最小。由於此事件規模較小，低估狀況在此事件有良好的改善，預測以及真實的最大震級相符。

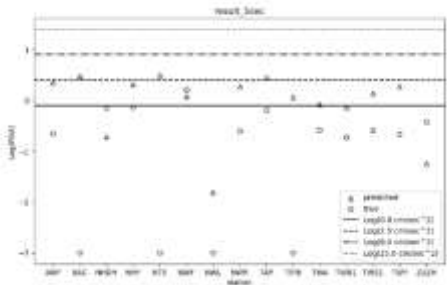


圖 73、T 後 5 秒預測/真實值

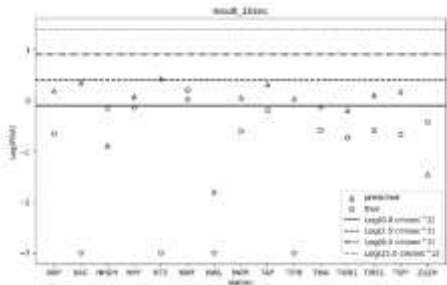


圖 74、T 後 10 秒預測/真實值

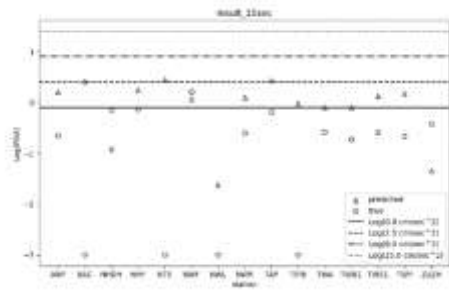


圖 75、T 後 15 秒預測/真實值

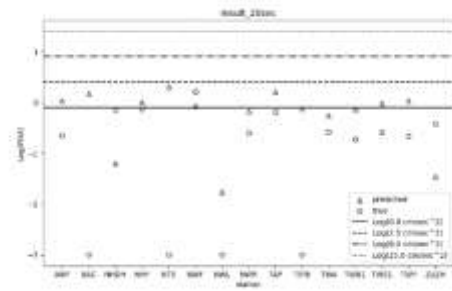


圖 76、T 後 20 秒預測/真實值

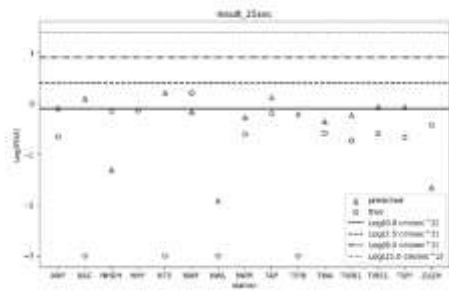


圖 77、T 後 25 秒預測/真實值

圖 78、圖 79、圖 80、圖 81、圖 82 及圖 83 分別代表 T 後 5、10、15、20、25 及真實之雙北地區預測結果。在 T 後 5 秒時緯度區間 $[25^{\circ}00'N, 25^{\circ}20'N]$ 經度區間 $[121^{\circ}15'E, 121^{\circ}30'E]$ 有稍微高估的情況，不過由於震級很小因此高估情況不算嚴重。

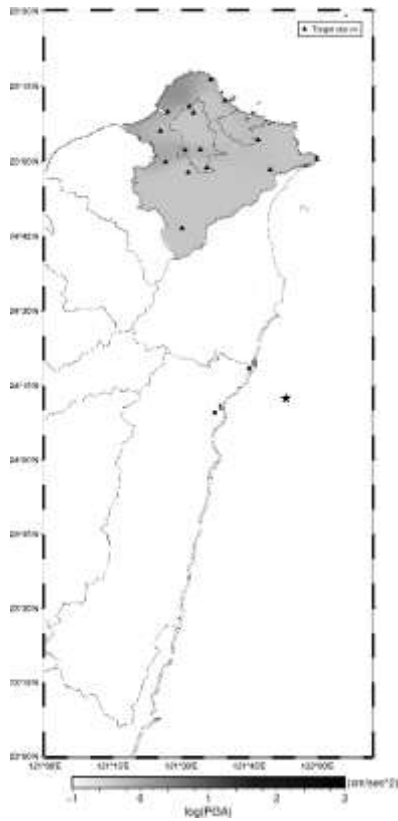


圖 78、T 後 5 秒地區預測結果

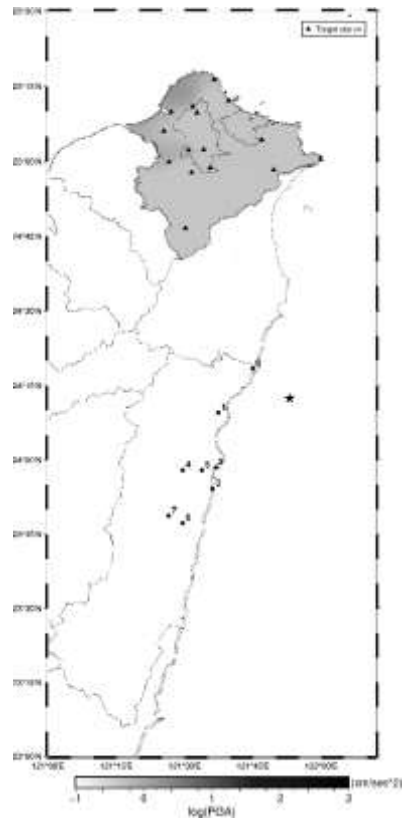


圖 79、T 後 10 秒地區預測結果

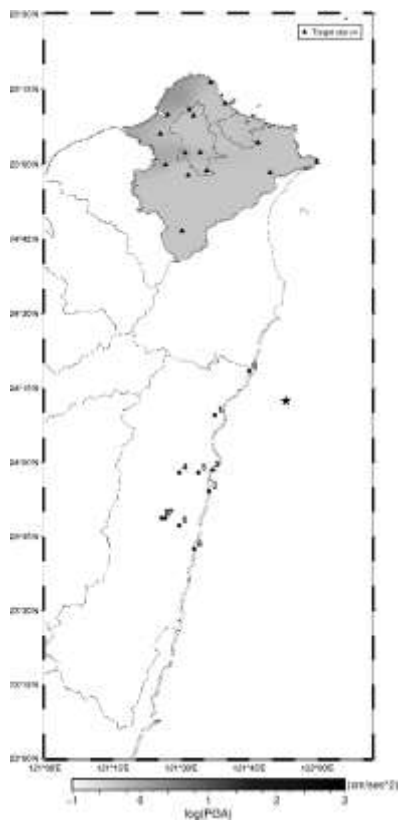


圖 80、T 後 15 秒地區預測結果

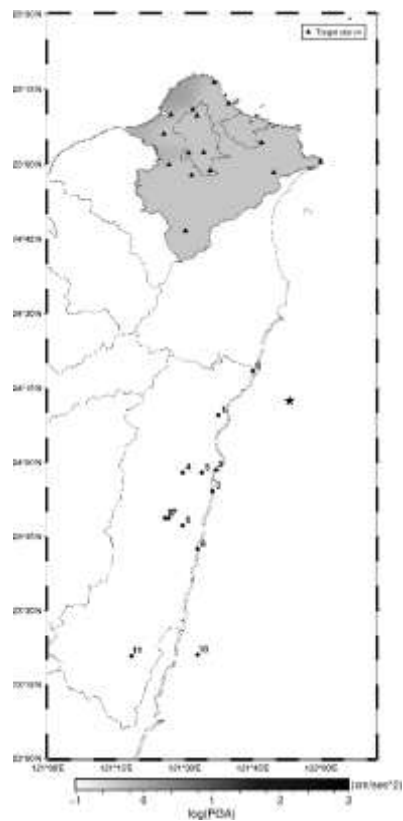


圖 81、T 後 20 秒地區預測結果

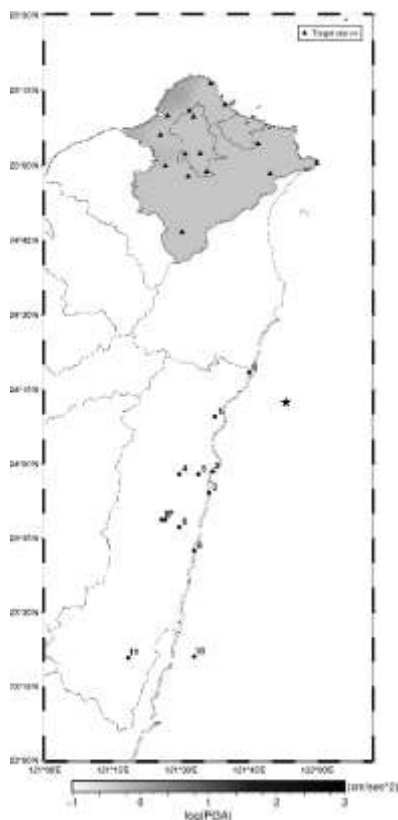


圖 82、T 後 25 秒地區預測結果

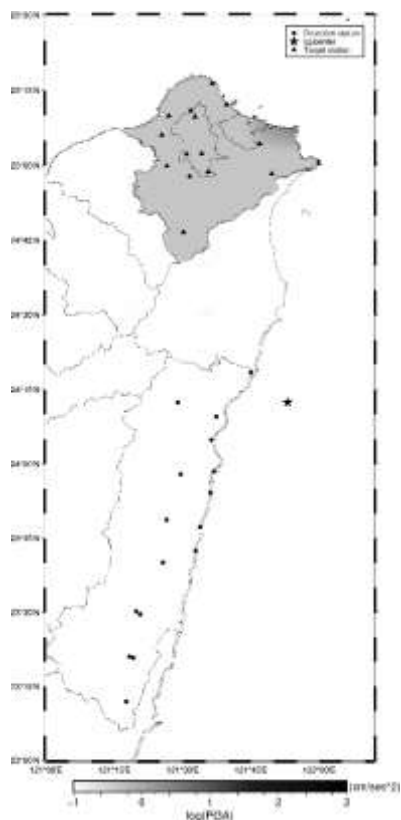


圖 83、地區真實結果

另外，表 6 統計了此事件各測站之提前預警時間。由統計結果得知，在目標位置中只有 NWF 是有真實超過 1 級且有提前 22 秒預警的情況，其餘測站真實值皆無超過 1 級因此不列入提前預警時間的計算。

表 6、第 71 號提前預警時間(秒)

	>1 級	>2 級	>3 級	>4 級
ANP	0	0	0	0
BAC	0	0	0	0
NHDH	0	0	0	0
NHY	0	0	0	0
NTS	0	0	0	0
NWF	22	0	0	0
NWL	0	0	0	0
NWR	0	0	0	0

	>1 級	>2 級	>3 級	>4 級
TAP	0	0	0	0
TIPB	0	0	0	0
TWA	0	0	0	0
TWB1	0	0	0	0
TWS1	0	0	0	0
TWY	0	0	0	0
ZUZH	0	0	0	0
count	1	0	0	0

3.5.5 各測站提前預警時間統計

此小節統計了上述四個事件之各測站平均提前預警時間，如表 7，計算方式為「各測站提前預警時間總和」與「有提前預警之事件總數」相除，例如 ANP 測站在第 19 號事件提前 20 秒預警及第 26 號事件提前 19 秒預警，則預警時間總和為 39 秒，事件總數為 2，相除過後得平均提前預警秒數 19.5 秒，有兩種情況為 NaN，一種是此測站預測未達該震級，亦或是此測站真實值為空，兩者之一發生就沒有計算提前預警時間。

表 7、各測站平均提前預警時間(秒)

	>1 級	>2 級	>3 級	>4 級		>1 級	>2 級	>3 級	>4 級
ANP	19.5	23	NaN	NaN	TAP	12.67	18.67	22	NaN
BAC	NaN	NaN	NaN	NaN	TIPB	NaN	NaN	NaN	NaN
NHHDH	NaN	NaN	NaN	NaN	TWA	14	18	NaN	NaN
NHY	12	12	19	NaN	TWB1	16.67	20	NaN	NaN
NTS	15	22	21	NaN	TWS1	15.33	19	NaN	NaN
NWF	17	18	NaN	NaN	TWY	25	28.5	NaN	NaN
NWL	NaN	NaN	NaN	NaN	ZUZH	18	20	NaN	NaN
NWR	17.33	24.5	27	NaN					

3.5.6 結論

本研究利用卷積神經網路 (CNN) 提取波形特徵，並結合圖神經網路 (GNN) 與 Transformer 中的注意力機制來學習測站間的空間關係。研究中以花蓮的 16 個測站資訊作為輸入，對雙北地區 15 個位置進行 PGA 預測任務。為獲得更均衡的數據分布，數據集依地震規模進行欠採樣(Undersampling)和過採樣(Oversampling)處理。此外，本研究具備實時(real-time) 預測能力，隨時間推進，每秒可更新一次預測，模擬實際接收波形的情境。

針對 2024 年 4 月 3 日的四個地震事件進行的評估結果顯示，對於第 19 號事件，由於規模較大，模型的預測結果普遍低估；而對第 26 號及第 45 號事件，低估情況有所改善，且預測值與真實值之間的差異顯著縮小。對於第 71 號事件，雙北地區的最大震級為 1 級，預測的最大震級與實際情況相符。在這四個事件的提前預警時間統計中，本研究的模型架構成功達到提前預警效果，無論地震規模大小，至少可提前 9 秒進行預測。

未來改進方向包括三個面向：首先，現模型使用的資料平衡方式仍無法解決大規模事件數據缺失的問題，因此將會改進數據增強技術；其次，調整模型架構以減少低估現象，縮小預測值與真實值的差異；最後，將分區預測擴展至

全臺，期望打破僅依賴特定來源地區測站和目標地區預測的限制。希望未來能進一步強化模型性能與泛化能力，為地震預警系統提供有力支援。

4 類神經網路應用於 P 波到時偵測

4.1 簡介

在地震檢測領域，P 波的精確定位對於快速預警系統的效能至關重要。地震初期的 P 波檢測能有效區分地震事件，並發出預警，為災害應對贏得寶貴的時間。然而，P 波檢測任務面臨諸多挑戰，尤其是在噪聲環境或低訊雜比(SNR)條件下。傳統的基於閾值和簡單特徵提取的算法，往往難以精準識別 P 波訊號，並且容易導致誤報，從而影響地震預警的準確性與可靠性。為了提升 P 波檢測的準確度，本研究採用了改良的 TransUNET 架構，專注於精準識別 P 波在時間序列中的位置。

TransUNET 結合了 UNET [7]和 Transformer，具備 UNET 的局部特徵提取能力及 Transformer 的長距離依賴建模優勢，特別適合處理時間序列數據。在該架構中，UNET 負責提取地震波形的局部特徵，而 Transformer 則通過建模時間序列的上下文關係，增強對 P 波位置的識別能力。然而，僅依賴此一架構仍難以充分應對地震數據中的噪聲干擾。為進一步強化模型的特徵選擇能力，本研究將 CBAM (卷積塊注意力模組) [8]整合至 TransUNET 架構中，以增強模型對空間與通道資訊的選擇性注意力，提升特徵提取的精確度。

CBAM 模組的引入使模型在處理噪聲數據時更具適應性。CBAM 透過學習自適應權重分配，提高模型的區分能力，能有效強化模型對 P 波的聚焦，並減少噪聲的干擾。結合 TransUNET 和 CBAM 的優勢，本研究提出的模型在 P 波檢測的準確性和特徵提取的精確度方面具有提升的潛力。透過結合 TransUNET 和 CBAM 的優勢，模型預期能增強地震預警系統在噪聲環境下的適應性和穩定性，有望為實現更精確的地震早期預警提供技術支持。

4.2 模型架構

模型架構如圖 84，輸入資料是三軸加速度的採樣值，時間窗每秒滑動一次，假設 t 是時間窗的大小，輸入可以表示為二維矩陣 $X \in \mathbb{R}^{3 \times t}$ ，二維矩陣 X 可以被定義為 (x_{ij}) $1 \leq i \leq 3, 1 \leq j \leq t$ ，其中 x_{ij} 是時間窗中 i 軸向的第 j 個加速度樣本 (Sample)。在我們訓練的資料中， t 被設定為 1000，意即有 10 秒的資料，各軸每秒有 100 個加速度的採樣值。

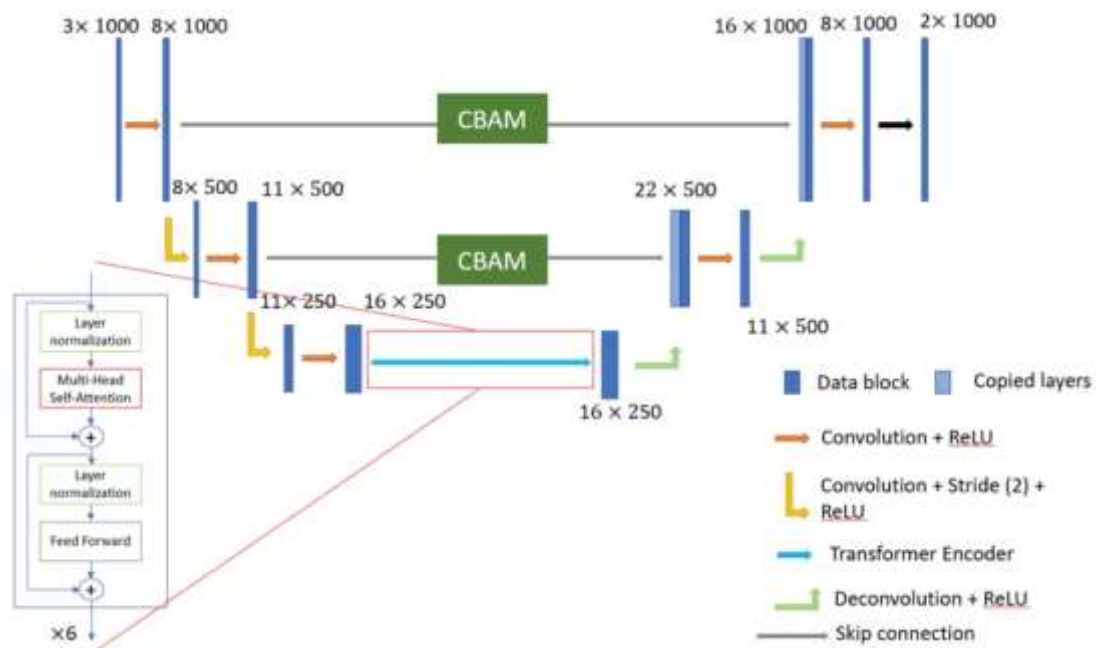


圖 84、模型架構圖

在特徵提取過程中，卷積層利用其在數據中捕捉複雜特徵的能力，成為提取有效特徵的關鍵層。本模型採用一維卷積(Conv1D)，用於對三分量加速度波形進行卷積運算，以提取出有助於挑選 P 波到時的相關特徵。為提升模型的收斂速度並減少梯度消失的風險，卷積層後加入了批次標準化(Batch Normalization)和整流線性單位函數(Rectified Linear Unit, ReLU)作為激活函數。批次標準化增強了訓練的穩定性和速度，而整流線性單位函數則提供非線性轉換，使模型更精確地學習數據特徵。儘管經卷積層提取的特徵可能欠缺直接的物理意義，這些高度抽象的特徵卻賦予模型分析多種類型波形資料的能力，使其能夠識別和提取出具有實際意義的數據規律。

經過卷積層提取的特徵隨後輸入至 Transformer 編碼器。Transformer 編碼器結合了注意力機制、前饋神經網路(Feedforward Neural Network, FFN)及殘差學習(Residual Learning)等技術。注意力機制使得模型能夠同時關注序列中的不同位置，從而捕捉輸入序列中的全局依賴關係。這一特性使得模型能夠針對序列中每個部分動態調整其關注度，以提取最具價值的特徵。此外，編碼器內的前饋神經網路對來自注意力機制的特徵進行進一步的非線性轉換，通常由多層全連接層組成，並在層間加入激活函數，進而提升模型在捕捉複雜特徵上的效能。同時，殘差連接通過將輸入與網路中間層輸出相加，有效減輕了深層網路的梯度消失或梯度爆炸問題，並加速了模型的收斂。這些特性使得 Transformer 編碼器成為一個強大的特徵提取器，能夠充分捕捉時間序列中的關鍵模式，進一步增強模型的準確性和泛化能力。

特徵提取完成後，模型通過反卷積層將特徵還原至與目標相同的維度，確保輸出能夠準確反映原始數據的特徵。反卷積的過程將提取的特徵進行放大和補充，使最終輸出的尺寸與目標保持一致。在此過程中，我們通過跳躍連結(skip

connection)將卷積層提取的特徵與反卷積後的特徵進行合併，以保持特徵信息的連續性和完整性。

為了進一步增強特徵傳遞，我們在跳躍連結中引入了卷積式注意力模組(Convolutional Block Attention Module, CBAM)，包括通道注意力(Channel Attention)和空間注意力(Spatial Attention)兩個子模組。通道注意力自動分配各通道的重要性權重，而空間注意力強化局部特徵，抑制不相關或干擾的特徵，使模型能聚焦於影響最終輸出的重要特徵。CBAM 的詳細架構如圖 85。

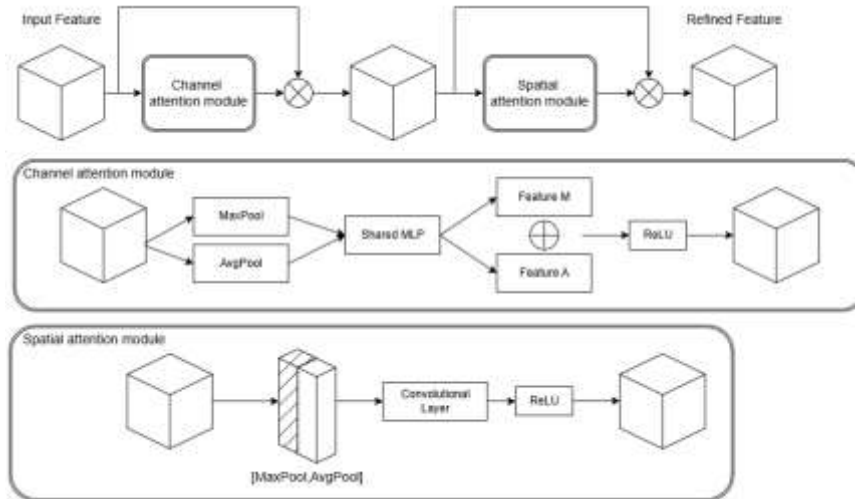


圖 85、卷積式注意力模組

在反卷積和 CBAM 處理後，模型的最後一層使用 softmax 函數輸出每個時間點 P 波到達的機率分佈。這樣的設計能夠使模型更直觀地給出每個時間點是否為 P 波到達的可能性，從而準確地檢測 P 波到達時刻。透過 CBAM 模組的整合，模型能夠更精確地捕捉和傳遞細微特徵，減少信息丟失，提升特徵提取和還原過程的效果，進而增強最終預測的精度與準確性。

4.3 資料處理

在本研究中，我們針對不同年份的地震數據進行系統化的訓練、驗證及測試，具體而言，我們將 2012 年至 2018 年間的 CWASN 地震數據作為訓練集，2019 年作為驗證集，並使用 2020 至 2021 年的資料作為最終測試集，以模擬模型在實際應用中的表現，CWASN 的測站分布如圖 86。為了提升資料品質並確保特徵準確性，前處理步驟中首先對數據進行了零均值化處理(zero-mean)，該步驟有效消除了數據中的偏移量，使波形均勻分布於基準線周圍，從而提高模型在數據變異性下的穩定性。接著，我們進一步對數據進行了帶通濾波(bandpass filtering)，該步驟能夠去除目標頻段外的噪聲成分，從而有效提升訊噪比(SNR)，確保所提取的特徵在不同頻率範圍內的代表性，資料集的訊噪比分布如圖 87。

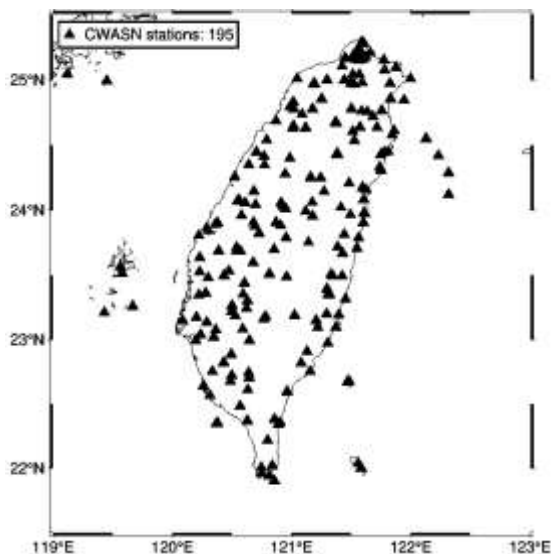


圖 86、測站分布

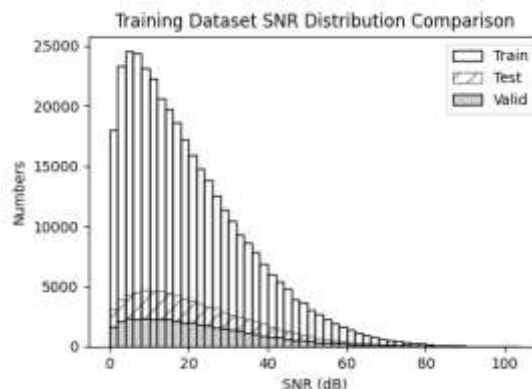


圖 87、資料集 SNR 分布

為了增強模型對多樣波形特徵的學習能力，我們在訓練過程中隨機設定 P 波到時的位置，將其擺放於 10 秒的時間窗內。這種隨機位置分布的設置不僅增加了訓練數據的多樣性，還在訓練階段模擬了現實地震波形的隨機性，從而提升了模型的泛化能力和實用性。在目標標注方面，我們採用了基於高斯分布的標注方法，如式(24)所示：

$$y = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (\text{式 } 24)$$

其中， μ 表示 P 波的到時位置， σ 則設定為 0.1 sec。透過此分布設置，P 波到時位置被賦予了最高的標注機率，並隨著位置的偏離而呈現逐步遞減的特性。這種設計在很大程度上反映了 P 波真實到達過程中的機率分布特性，使得模型在訓練時能更準確地學習到 P 波的典型特徵。另外，我們也設定噪聲的機率，噪聲的機率依據式(25)進行設定：

$$Prob(N) = 1 - Prob(P) \quad (\text{式 } 25)$$

圖 88 展示了該標注方法的實際應用範例。此方法不僅提升了模型在 P 波到時特徵上的識別精度，還能在不同的波形與環境中保有高度適應性。此外，該設計在標注過程中體現了對 P 波和背景噪聲的清晰區分，進一步提高了模型在真實場景中的辨識效果。此過程所帶來的多樣性和彈性，使模型能夠在不同的地震波形特徵下保持穩定的偵測能力，從而提升模型的總體性能和在實際應用中的有效性。

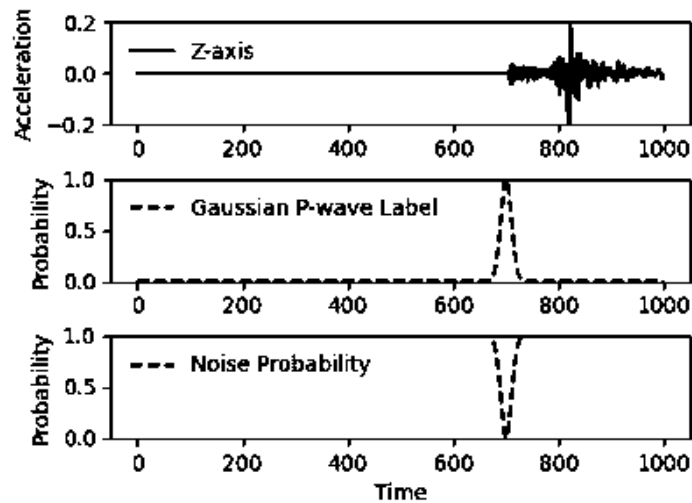


圖 88、標注方法

4.4 訓練過程

我們的訓練設備選用了 i5-13500 的 CPU、48GB 的 DDR5 記憶體，以及具備 16GB VRAM 的 RTX4080 顯示卡，以支持高效能的訓練過程。由於我們的模型在運行時需要處理大量地震數據，這樣的硬體配置確保了運算的穩定性和模型的流暢訓練。在參數設置方面，我們將 batch size 設定為 32，以平衡訓練的速度和計算的需求；學習率則設置為 10^{-5} ，這樣的較低學習率可以促進模型的穩定學習，避免過快收斂可能導致的錯誤結果。我們使用交叉熵(Cross Entropy)作為損失函數(Loss function)，該函數能有效計算出模型預測結果與實際標籤之間的誤差，從而為優化過程提供精確的調整依據。

為了避免模型在訓練過程中出現過度擬合(Overfitting)，我們引入了早停(EarlyStopping)機制，該機制會監控驗證損失的變化情況，若在連續 8 個 epoch 中驗證損失無下降，即判定模型可能已達到最佳表現，從而自動停止訓練，以避免過度擬合對泛化能力的影響。針對模型的優化，我們採用了 Adam 優化器(Optimizer)，該優化器結合了動量和自適應學習率的優勢，有助於提升模型的收斂速度和穩定性，避免傳統梯度下降方法中的停滯問題。

在訓練過程中，我們將每筆資料的 10 秒三軸加速度波形數據輸入模型，模型隨即為每一個樣本點生成 P 波與噪聲的機率輸出。透過損失函數的計算，模型能夠獲得預測結果與目標標籤的熵差，並依此調整權重參數，以持續優化其預測效果。本研究中的最終模型經過充分訓練後達到穩定的表現，並有效提高了 P 波檢測的準確性。在訓練過程中，僅記錄驗證損失下降的情況，以確保模型參數在每次改進時得到最佳化，訓練損失和驗證損失的變化趨勢如圖 89 所示。

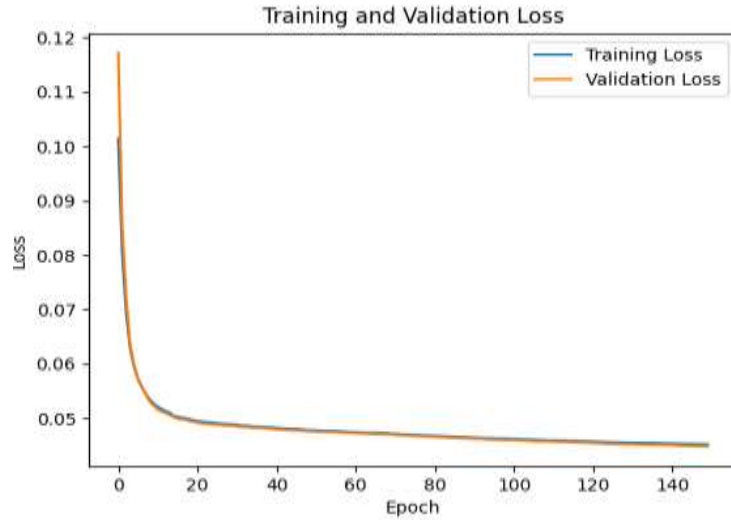


圖 89、損失函數

4.5 模型結果

我們在評估地震事件的偵測性能時，將事件與噪音分開進行評量，這樣的做法有助於深入了解模型在兩種不同情況下的表現。具體而言，我們分別計算事件的準確率(Accuracy) 以及噪音偵測的準確率，以確保模型能在偵測地震波的到達時間上具有足夠的精確度，同時也能有效地過濾背景噪音，避免誤報。

首先，對於地震事件的評估指標，我們使用式(26)計算事件的準確率：

$$Acc_E = \frac{E}{AE} \quad (\text{式 } 26)$$

其中，E 代表偵測正確的事件數，AE 代表事件的總數，若模型偵測出的 P 波到達時間與實際到達時間的差異小於 ± 0.5 秒，且該預測值超過設定的閾值(threshold)，則我們認為這次偵測是成功的，則 E 增加 1。這樣的準確率計算方式可以幫助我們了解模型對於事件的反應速度和精度，尤其是在地震預警系統中，準確拾取 P 波到達時間至關重要，因為它直接影響到後續的地震預警和減災措施的啟動。其次，為了評估模型在噪音偵測方面的性能，我們使用式(27)來計算噪音的準確率：

$$Acc_N = \frac{N}{AN} \quad (\text{式 } 27)$$

其中，N 代表預測正確的噪音數，AN 代表噪音的總數，如果模型對噪音部分的預測結果未超過設定的閾值，則我們認為這是一個成功的噪音偵測，N 增加 1。這樣的評估方式有助於我們確保模型不會將噪音誤認為地震事件，從而減少誤報率。在實際應用中，噪音的準確偵測對於系統的穩定性和可靠性具有重要意義，因為過多的誤報會導致不必要的警報，進而影響系統的可信度。

此外，選擇適當的閾值對於模型的準確性有很大的影響。在選擇合適的閾

值時，我們參考了模型的 ROC 曲線 (Receiver Operating Characteristic Curve)，以確保事件偵測的準確性和穩定性。圖 90 顯示了模型的 ROC 曲線以及各個閾值下的偵測性能。我們通過觀察 ROC 曲線上的變化來選擇最佳的閾值，找到了一個在達到較高 TPR (True Positive Rate, 真陽性率) 同時保持較低 FPR (False Positive Rate, 偽陽性率) 的平衡點。其中 TPR 為 Acc_E ，其意義為在實際為事件的樣本中，被正確偵測為事件的比例，FPR 為 $1 - Acc_N$ ，其意義為在實際為噪音的樣本中，被錯誤偵測為事件的比例。

從圖中可以看到，當閾值增加時，TPR 逐漸降低。在分析中，我們發現當閾值設定為 0.2 時，可以保持較高的偵測率，同時控制 FPR 不過高。因此，我們選擇閾值為 0.2 作為最終的設定，以達到準確偵測事件與減少誤報之間的最佳平衡。虛線標示了所選擇的閾值位置，即 0.2，這個值是在 ROC 曲線顯示 TPR 顯著下降前的最佳點，能有效地在最大化偵測正確事件數量的同時，降低噪音帶來的誤報。

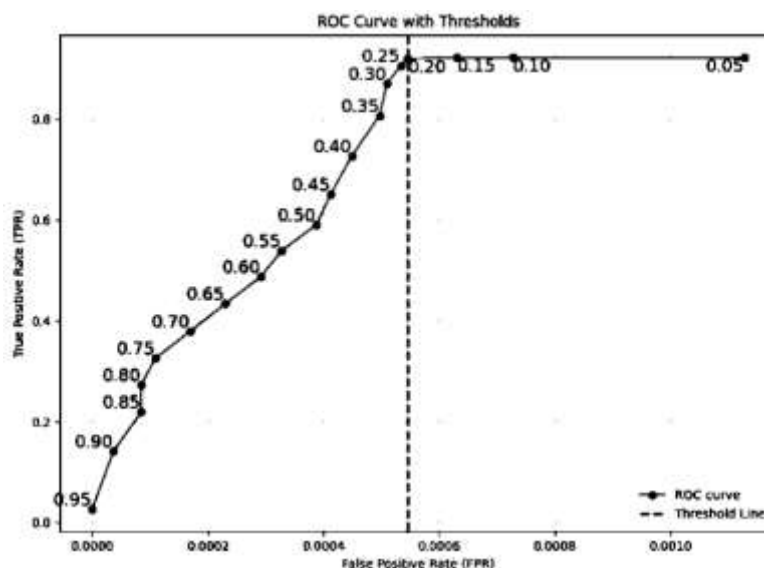


圖 90、ROC 曲線圖

表 8 顯示了在閾值設定為 0.2 的情況下，模型的分類性能。結果顯示，對於實際的地震事件，模型檢測率達到 92.17%，而誤判為噪音的比例為 7.83%；對於實際噪音，模型正確檢測率高達 99.93%，而誤判為地震事件的比例為 0.07%。此外，平均絕對誤差為 0.125 秒，平均誤差為 -0.035 秒，標準差為 0.174 秒。這些結果表明模型在噪音與地震事件之間具有較高的區分能力，但也存在一定的誤判。

表 8、模型表現

	Events(detect)	Noise(detect)
Events(actual)	92.17%	7.83%
Noise(actual)	0.07%	99.93%

4.6 事件評估

在這個章節中，本研究選用了 109 年 1 月 12 日發生的小區域有感地震，震央位於臺東，地震深度 26.6 公里，芮氏規模為 3.5。為了評估模型的性能，我們將 P 波到時時間固定在時間窗的第 700 個索引值(index)，並展示挑選正確 P 波到時位置、挑選錯誤 P 波到時位置，以及被誤判為噪音的案例。

在圖 91 中，我們選用了此事件 CWASN 測站網的阿里山(ALS)測站，這是一個被正確偵測的地震，誤差時間為-0.02 秒。在圖 92 中，我們選用了花蓮(HWA)測站，這同樣是一個被正確偵測的地震，誤差時間為 0.12 秒，即使在背景噪音的干擾下，模型依舊能準確挑選出 P 波到時位置，表明模型在這些挑戰性條件下具有良好的穩健性。在圖 93 中，我們選用了鹽寮(EYL)測站，儘管模型的預測值超過了閾值，但由於訊號中的背景噪音干擾，導致模型挑選了錯誤的 P 波到時位置，誤差時間為 0.58 秒。而在圖 94 中，南山(NNS)測站的地震訊號則被模型視為噪音，這是一則漏報案例。

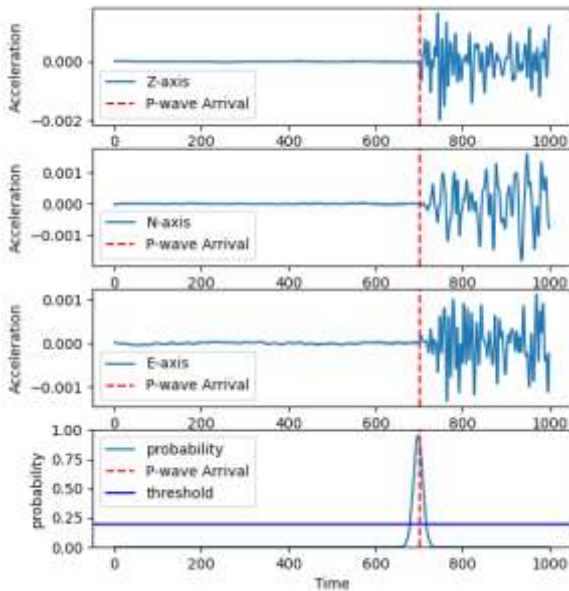


圖 91、阿里山測站結果

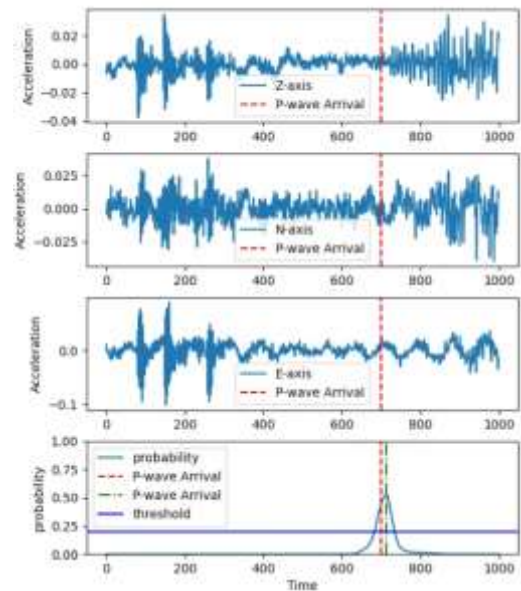


圖 92、花蓮測站結果

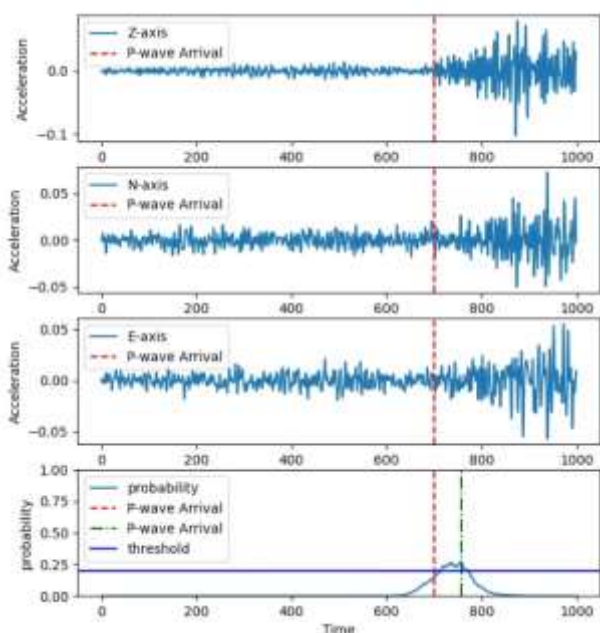


圖 93、鹽寮測站結果

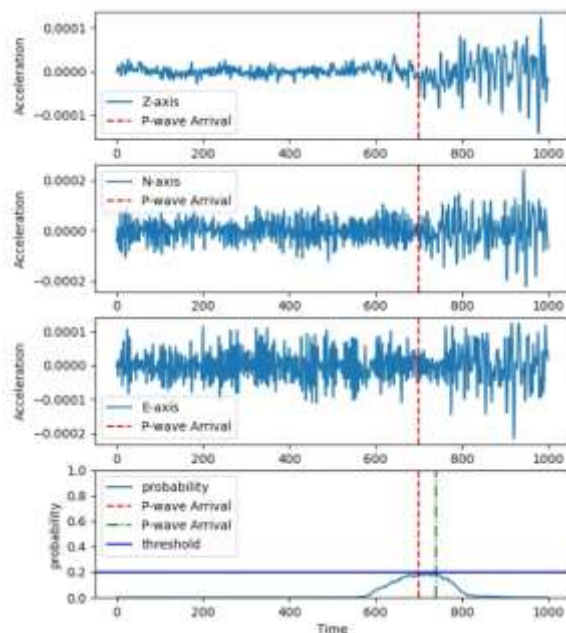


圖 94、南山測站結果

本研究認為，震幅大小與地震訊號的明顯程度對模型的偵測結果有顯著影響。在鹽寮測站的案例中，訊號中的背景噪音干擾導致模型無法正確辨識 P 波位置，這顯示模型在地震訊號不夠顯著或噪音訊號較為複雜的情況下，其可靠性仍有待提升。而南山測站的漏報案例顯示，該地震訊號的特徵可能與常見的噪音特徵存在一定的相似性，使得模型難以有效區分。我們認為，這種錯誤可能與該測站的觀測條件或地震訊號特徵不夠明顯有關。

總體來說，模型在高訊噪比的情況下表現良好，能夠有效忽略噪音並準確地挑選 P 波到時位置；然而，在訊噪比較低或噪音較為複雜的情況下，模型的表現尚有不足，這為我們未來的改進方向提供了具體的指引。

4.7 結論

本研究提出了一種結合 TransUNET 與 CBAM 的改良模型，旨在提高 P 波到時的精確檢測。TransUNET 利用 UNET 和 Transformer 的結合，具備局部特徵提取和全局依賴建模的優勢，而 CBAM 的引入進一步提升了模型在空間和通道資訊上的選擇性注意力能力。在多個實驗測站的測試中，模型展示了在高訊噪比情況下的良好表現，能夠有效忽略背景噪音並準確檢測到 P 波到達時間，顯示出模型在噪音條件下的穩健性。

然而，實驗結果也表明，模型在低訊噪比或訊號特徵不明顯的情況下存在一定的不足。在一些案例中，模型無法準確區分地震訊號和噪音，導致誤報或漏報的發生。這些結果顯示了現有模型在面對複雜噪音環境時的局限性。因此，未來的工作將專注於進一步優化模型，包括增加對低訊噪比數據的訓練，並引入地震的物理特徵，以提升模型在不同地震條件下的穩定性和準確性。

5 即時地震預警系統

5.1 簡介

地震發生時，能夠在破壞性震波到達前發出有效的預警至關重要。迅速且準確的地震預警不僅能減少財產損失，還能為民眾和公共設施爭取寶貴的反應時間。隨著人工智慧 (AI) 技術的發展，即時分析測站數據的需求日益增長，應用 AI 模型進行地震資料的分析已被證明能顯著提高預警系統的準確性和反應速度。然而，實現 AI 模型與現有系統的整合需要克服數據傳輸效率及時效性的挑戰。

本系統整合了 Earthworm 系統與 AI 模型，並採用 Message Queuing Telemetry Transport (MQTT) 協議作為數據傳輸方案，以建立更高效、穩定的即時地震預警系統。此架構透過將 P-Alert 觀測網的即時波形數據與 P 波到時資訊從 Earthworm 系統的共享記憶體中發布至 MQTT 上，使多個訂閱端能夠接收最新數據並進行即時處理。AI 模型通過訂閱這些數據並即時進行預測分析，結果隨後被發布至 MQTT，為後續的地震預警決策提供可靠的參考依據。

5.2 MQTT 與 Earthworm 的協作

Earthworm 系統由美國地質調查局(USGS)開發，是一套功能完善的地震預警系統，已廣泛應用於全球多個地震監測機構中，包括臺灣的中央氣象署。Earthworm 系統的架構主要由兩個元件組成：RING 和 MODULE。RING 作為共享記憶體空間，用於存放波形資料和地震事件資訊，而 MODULE 則通過讀取 RING 中的資料來執行後續操作，例如篩選 P 波到時、觸發測站和發出警報等。傳統的 Earthworm 系統流程中，波形數據由各地測站接收並存入 WAVE_RING，隨後由 PICK_EEW 模組判斷是否需要觸發測站並將結果存入 PICK_RING，之後 TCPD 模組根據觸發頻率判斷是否發出警報，並計算震央和地震規模等資訊，最終將事件資訊存入 HYPO_RING，以便進行地震警報的發布。儘管 Earthworm 系統在穩定性和數據處理流程方面已相當完善，但隨著即時地震預警需求的增長，僅依靠 Earthworm 系統難以完全滿足日益增長的數據處理速度需求。因此，本系統引入 MQTT 協議，旨在提升數據傳輸的靈活性和效率，同時實現 Earthworm 與 AI 模型的無縫集成，圖 95 為本系統的架構圖。加入 MQTT 的目的是將地震數據的傳輸與處理進一步架構化，使多個模組和訂閱端可以同時接收最新數據，並即時進行分析和預測，從而減少延遲，並為系統未來的擴展提供更高的靈活性。

在 MQTT 與 Earthworm 的協作流程中，Python 程式首先通過 PyEarthworm 套件的 get_wave() 方法從 WAVE_RING 中接收波形封包，並使用 paho-mqtt 套件將其發布到 MQTT_WAVE 主題，該主題用於存放來自各測站的即時波形數據，讓其他系統能夠即時訂閱並進行後續處理。同時，系統持續監聽 PICK_RING 的數據，當 PICK_EEW 模組觸發某測站時，Python 程式將該測站的 P 波到時資訊即時發布到 MQTT_PICK 主題，這樣 AI 模型可以迅速接收觸發資訊，並進行地震預測。AI 模型透過訂閱 MQTT_WAVE 和 MQTT_PICK 主題，實時接收波形數據和測站觸發資訊，進行地震預測。預測

結果會發布至 MQTT_EEW 主題，以供後續的預警系統參考。當預測結果達到預設警報標準時，透過 LINE Notify 向相關人員發送即時警報，以確保警報訊息能迅速傳達。

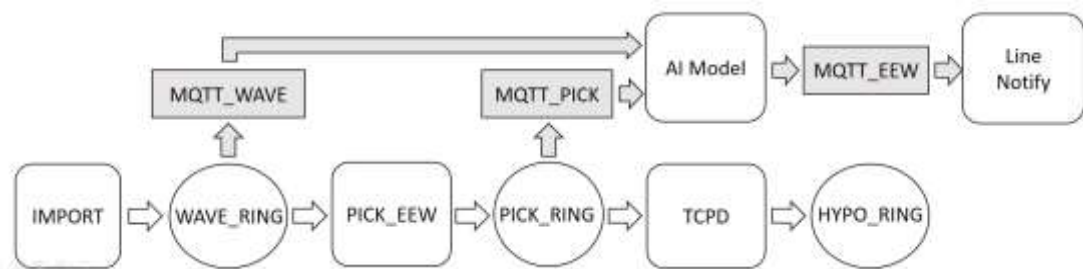


圖 95、系統架構圖

5.3 Line Notify

本系統透過 LINE Notify 實現即時警報的推送，以便將 AI 模型的地震預測結果快速發送給相關人員或群組。LINE Notify 是 LINE 提供的一項服務，允許應用程式透過 API 向用戶發送訊息。以下將說明設定 LINE Notify 所需的步驟及如何結合至地震預警系統中。

- I. 登入 LINE Notify 官方網站
首先，請前往 LINE Notify 官方網站 並使用您的 LINE 帳號進行登入。
- II. 進入個人頁面
登入後，點擊右上角的「個人頁面」連結，進入您的 LINE Notify 個人設定頁面。(圖 96)

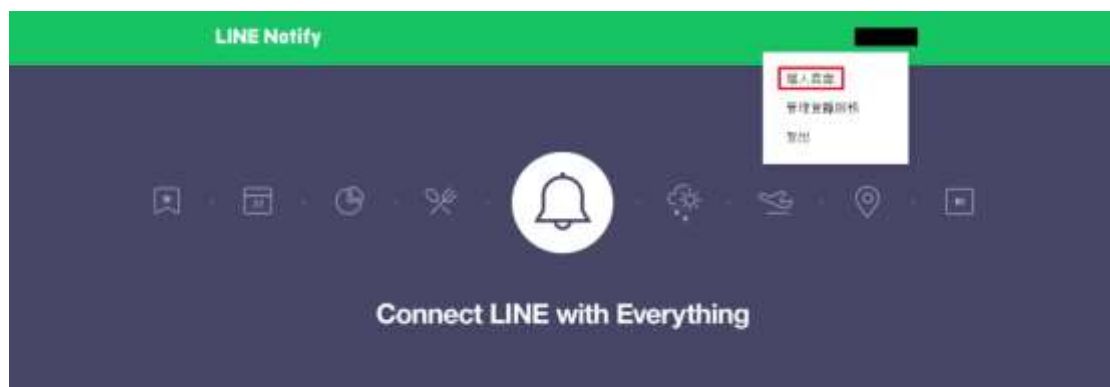


圖 96、介面示意圖

- III. 發行存取權杖
在個人頁面的底部，您會看到「發行存取權杖（開發人員用）」區域。點擊「發行權杖」按鈕，開始進行權杖的設定。(圖 97)

發行存取權杖(開發人員用)

若使用個人存取權杖，不須登錄網站服務，即可設定通知。



圖 97、存取權杖介面示意圖

IV. 設定權杖名稱與通知對象

設定權杖名稱，選擇想接收通知的聊天室後，點擊「發行」。若選擇的聊天室為群組，則需將 LINE Notify 官方帳號邀請至該群組中。(圖 98)

圖 98、權杖設定介面示意圖

V. 完成權杖發行並取得 Token

設定完成後，點擊「發行」按鈕。系統將生成一組存取權杖(Token)。請務必將此權杖妥善保存，因為出於安全考量，該頁面只會顯示一次，日後無法再查看同一組 Token。(圖 99)



圖 99、權杖設定介面示意圖

如此一來就能取得的 Token 配置至系統設定中。這樣，當地震預警系統偵測到地震事件並達到預警標準時，系統將自動使用該 Token 透過 LINE Notify API 發送通知。這使得地震預警訊息能夠以即時、可靠的方式傳遞給相關人員，提高了整個預警系統的反應速度和效率。

5.4 事件預測

當本系統接收到來自花蓮測站的 P 波到時的觸發訊號(PICK)後，會立即啟動地震事件的數據處理流程。首先，系統會提取該測站 P 波到時前後的波形數據，並將這些數據儲存到 NumPy 陣列 waveform 中。為了確保數據的精確性和可用性，系統會對這些波形數據進行一系列預處理操作，包括檢查並移除異常值（例如在數據中出現的重複 9999 值，通常代表缺失或錯誤的數據點）。此外，系統會從波形數據中去除平均值，以消除基線漂移的影響，從而使得 AI 模型可以專注於地震波形中的真實振幅變化。

在完成預處理之後，處理過的波形數據會被輸入到 AI 模型進行地震強度的預測。此 AI 模型經過訓練，能夠根據波形中的特徵識別出 PGA 值的變化，從而判定地震的強度。模型輸出的是臺北測站的預測 PGA 值，當其中任何一個測站的 PGA 值超過 8 Gal（代表震度三級）時，系統認為該臺北地區的地震

強度達到需要警報的程度。為了避免單一測站的誤判，系統還會對多個測站的預測結果進行交叉比對，確保超過閾值的測站數據是一致的，這樣的多測站驗證可以有效減少誤報，並提高警報的準確性。

當系統確認警報條件成立時，會生成一份地震速告，速報參考國家級警報內容，警示臺北地區發生有感地震。最終，透過 LINE Notify 將警報訊息傳送到指定的 LINE 帳號或群組，如圖 100。目前為簡單的文字通知，這樣的警報讓接收者可以快速了解地震事件的基本情況，未來也可以考慮進一步結構化報告內容，使得訊息更清晰易懂。

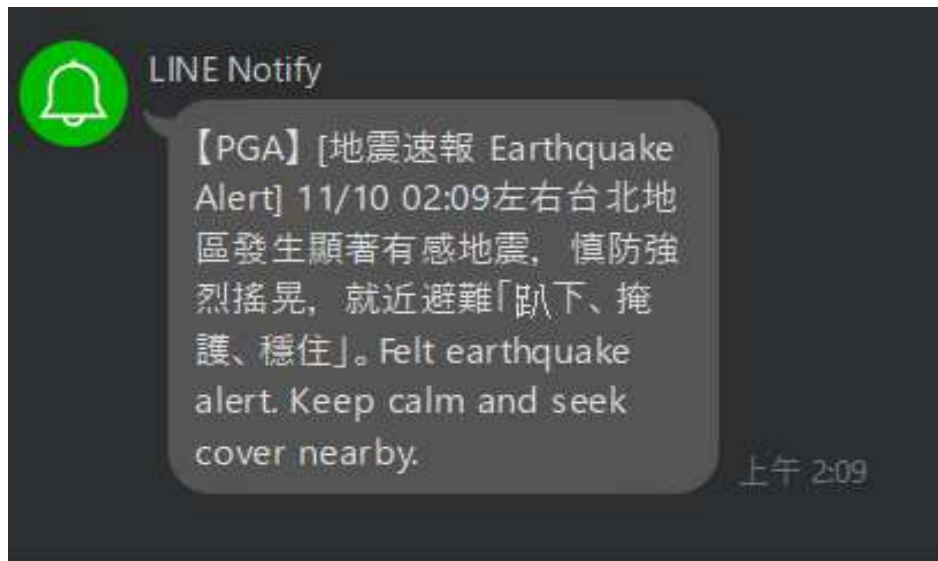


圖 100、警報通知範例

5.5 結論

本系統完成了一套整合 AI 模型與傳統地震預警系統的即時地震監測與預警平台。系統透過花蓮測站的即時波形數據，利用 AI 模型來預測臺北地區的地震強度(PGA)，當超過設定的閾值時，系統會即時透過 LINE Notify 發出地震警報。此系統結合了 MQTT 協議以實現測站數據的即時接收與傳輸，並以結構化的數據處理流程來提升模型的預測效率。未來的改進方向包括優化數據處理、完善警報訊息格式、強化多測站數據的整合及優化 AI 模型，以進一步提升系統的效能。

参考文献

- [1] S. Brody, U. Alon and E. Yahav, "How attentive are graph attention networks?" *arXiv:2105.14491*, 2021.
- [2] L. Zhou, H. Wang, "MST-GAT: A multi-perspective spatial-temporal graph attention network for multi-sensor equipment remaining useful life prediction", *Information Fusion*, vol. 110, p. 102462, 2024.
- [3] S. Bloemheuvel, J. van den Hoogen, D. Jozinović, A. Michelini and M. Atzmueller, "Graph Neural Networks for Multivariate Time Series Regression with Application to Seismic Data", *International Journal of Data Science and Analytics*, pp. 1-16, 2022.
- [4] Y. Shi, Z. Huang, W. Wang, H. Zhong, S. Feng and Y. Sun, "Masked label prediction: Unified message passing model for semi-supervised classification", *International Joint Conference on Artificial Intelligence (IJCAI)*, 2021.
- [5] A. Vaswani et al., "Attention is all you need," *Neural Information Processing Systems (NIPS)*, pp. 5998–6008, 2017.
- [6] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *International Conference on Learning Representations*, 2019.
- [7] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 234–241, 2015.
- [8] S. Woo, J. Park, J.-Y. Lee and I. S. Kweon, "CBAM: Convolutional block attention module", *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 3-19, 2018.

機器學習技術建立現地型地震預警系統相關研究(III)

子計畫二

使用多測站資料與地震基石模型的前瞻研究

陳冠宇 朱冠霖 林汶蔚 林怡萱 徐旻暄

國立臺灣科技大學

摘要

臺灣位於歐亞大陸板塊與菲律賓海板塊的交界處，是全球地震活動最為頻繁的地區之一。由於地質構造複雜，每年平均會發生約兩萬次大大小小的地震，其中不乏具有破壞性甚至致命的地震事件。這樣的地震頻率與強度，對於地震預警及災害防治是巨大的挑戰，也促使相關研究持續推進。近年來，深度學習技術以其卓越的數據處理與識別能力，在各領域得到廣泛應用。在地震研究領域，深度學習已被運用於多種任務，例如地震波到時的 P 波自動拾取 (P-wave picking)、地表最大加速度 (PGA) 的預測等，均展現了不凡的效果。在本次研究中，我們基於深度學習技術，聚焦於四個主要研究主題：第一，針對大型地震數據進行預訓練模型的研究，期望建立具泛用性的地震預訓練模型；第二，探索單測站地震預警模型的預訓練技術，提升小型或單站資料應用的準確性與穩健性；第三，發展區域型地震預警模型，以適應臺灣多變的地質條件；第四，研究並改進 P 波自動拾取模型，以提高拾取的精準度與效能。此外，我們針對現行部署於中央氣象署 MQTT 架構上的 P 波拾取模型進行改良，引入 Decision Making 模型，進一步降低實務應用中的誤報率。透過這項改進，我們希望提升模型在實際地震預警系統中的可靠性，為地震災害的快速反應與減災工作提供更強健的模型。

關鍵字：深度學習、預訓練模型、單測站、多測站

Abstract

Taiwan is located at the boundary between the Eurasian continental plate and the Philippine Sea plate, making it one of the most seismically active regions in the world. Due to its complex geological structure, approximately 20,000 earthquakes of varying magnitudes occur annually, some of which are destructive and even fatal. This high frequency and intensity of seismic activity pose significant challenges to earthquake early warning systems and disaster mitigation efforts, driving continuous advancements in related research. In recent years, deep learning technology has been widely applied across various fields due to its exceptional capabilities in data processing and pattern recognition. In the field of earthquake research, deep learning has been employed for various tasks, such as P-wave picking and peak ground acceleration (PGA) prediction, achieving remarkable results. In this study, we leverage deep learning technology and focus on four main research topics. Pre-training models on large-scale seismic datasets to develop generalized earthquake analysis models. Exploring pre-training techniques for single-station earthquake early warning models to enhance the accuracy and robustness of applications using limited or single-station data. Developing regional earthquake early warning models to adapt to Taiwan's highly variable geological conditions. Investigating and improving automatic P-wave picking models to enhance their precision and performance. Furthermore, we have refined the P-wave picking model currently deployed in the Central Weather Administration's MQTT system by integrating a Decision-Making model. This improvement aims to reduce the false alarm rate in practical applications. Through this enhancement, we seek to improve the reliability of our P-wave picking model in real-world earthquake early warning systems, providing a stronger foundation for rapid response and disaster mitigation efforts.

Keywords : Deep Learning, Pre-trained Models, Single-station, Multi-station

第一章：預訓練大型地震模型之研究

1.1 背景

地震監測和預警是災害預警系統中至關重要的組成部分。隨著深度學習的發展，許多研究已探索這些方法在相關任務中的應用 [10][11][12][13]。其中，自監督學習因其能夠從未標註答案的資料中學習有用的特徵，而成為一項極具潛力的技術，並引起了廣泛關注 [16]。通過這種方式訓練的模型可以對參數進行微調，以適應特定任務，通常能夠獲得更好的效果，或顯著減少所需的標註資料。儘管已有大量關於深度學習應用於地震監測和預警的研究，但在此領域中對自監督學習的應用仍然較少。因此我們提出了一種大型地震模型（Large Earthquake Model, LEM），以探討自監督學習在構建地震預訓練模型方面的有效性和效率。

近年來，自監督學習的預訓練模型在許多研究領域取得了顯著成效，如自然語言處理和語音處理。這些預訓練模型最具吸引力的特點是訓練時不需要使用標註答案的資料。換句話說，它們通過簡單的目標（如掩碼預測和對比學習）利用資料本身來學習具有信息量和通用性的表示（Representations）。基於這些表示，可以進行微調以獲得任務導向的模型，通常能在幾乎所有下游任務中達到最先進的性能。鑑於這些優勢，越來越多的基礎模型被提出並證明可以提升各類任務的性能，同時減少了數據標記需求，降低了複雜或耗時任務的入門門檻。

1.2 方法

對比學習方法已在語音處理中展現了出色的表現。Wav2Vec 2.0 [17]使用對比學習作為自監督學習的訓練目標，通過提取特徵後掩蓋部分語音特徵，並對掩蓋特徵和其他區域的量化向量 [21]，進行對比學習，在多種語音任務中達到卓越的效果。

延續此一研究，我們提出了一種新穎的大型地震模型（LEM），基於先進的深度學習技術。LEM 以大量未經過標註答案的地震波形進行自監督預訓練，旨在學習通用且具信息量的地震特徵。經過預訓練後，LEM 允許高效利用大量未標註地震資料進行地震監測任務，無需為每個特定地震任務從頭開始訓練新模型。換句話說，LEM 在經過少量任務導向的標註數據微調後，可有效適應各種地震監測和預警任務。

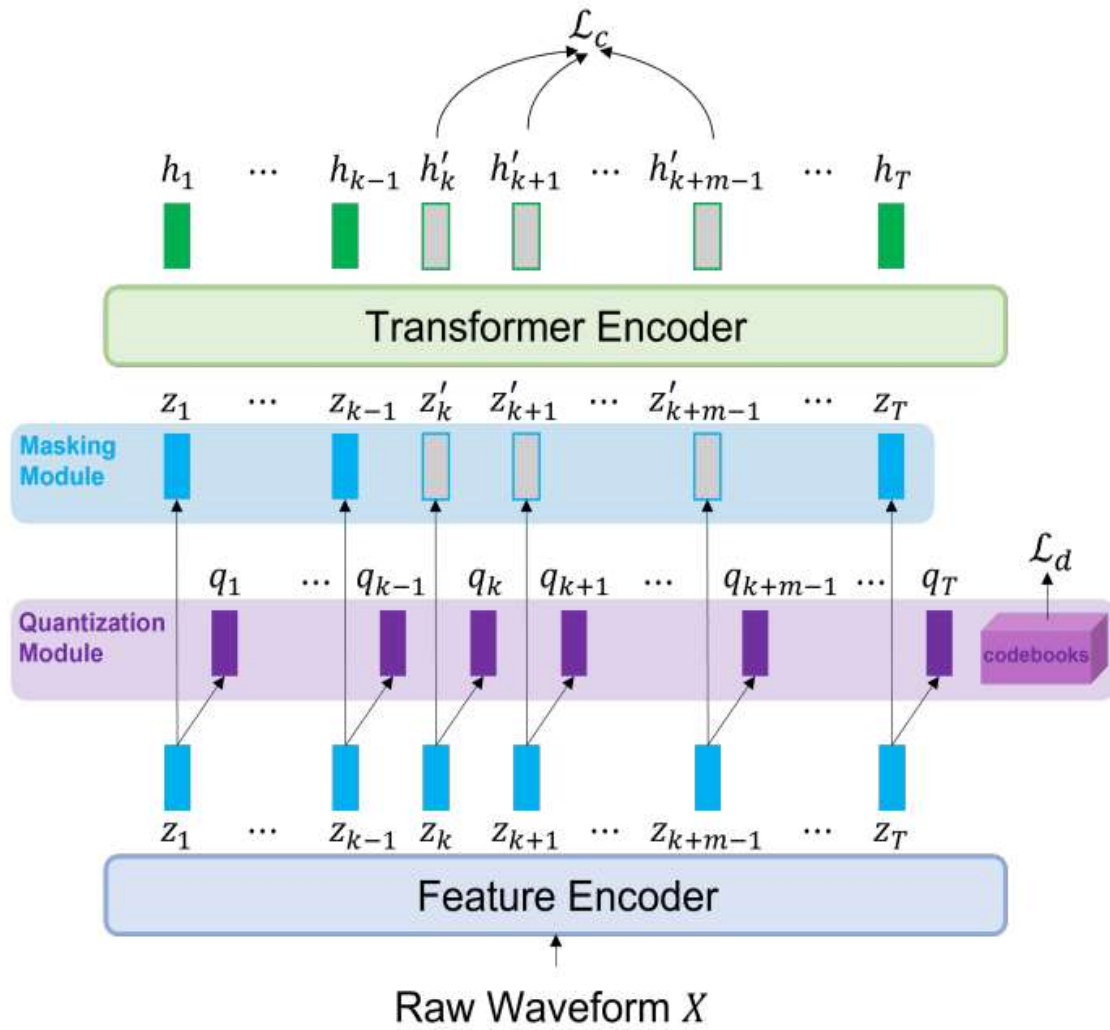


圖 1.1：LEM 模型架構圖

LEM 模型由特徵編碼器 (Feature Encoder)、上下文編碼器 (Transformer Encoder)、量化模組 (Quantization Module) 和遮罩模組 (Masking Module) 組成，如圖 1.1 所示。特徵編碼器用於將原始波形轉換為一系列的向量特徵，由兩層包含層正規化和 GELU 激活函數的時間卷積組成。特徵編碼器將波形數據轉換為連續 embedding，這些 embedding 捕捉地震信號的局部結構，並作為後續處理的基礎；上下文編碼器位於 LEM 的末端，引導這些向量學習上下文資訊，該編碼器包含 12 層 Transformer；我們使用量化模組來離散化密集特徵，以模仿自然語言處理中學習字符 (token) 預測，先前的研究顯示此步驟有助於表徵學習；受大型基礎模型訓練成功的啟發，我們引入遮罩預測任務，遮罩模組引入了一個特殊向量，以遮罩部分嵌入，使模型能以自監督方式進行訓練。

1.3 實驗結果

微調過程對於將預訓練期間學習的通用特徵適應於特定的下游任務至關重要。為了緩解不同任務的不同目標，我們將在 LEM 上堆疊一個簡單的任務導向解碼器。本研究提出並比較了三種微調策略：整個模型訓練 (Finetune)、僅解碼器訓練 (Freeze) 以及加權和僅解碼器 (Weighted Sum) 訓練 [22]。我們選擇了 P 波到時拾取和規模估算任務作為範例，來評估預訓練的 LEM 模型的能力。對於 P 波到時拾取，我們選用 EQTransformer [15] 和 PhaseNet [14] 作為我們的基線模型。對於規模估算，我們選用 MagNet [20] 作為我們的基線模型。我們的實驗顯示，LEM 不僅能夠提供具有競爭力甚至更優於當前最先進模型的結果，且僅需 0.01% 的數據即可達到滿意的性能。除其卓越的性能外，LEM 還提供了一種可能的解決方案，能有效應對地震數據標記耗時且費力的挑戰，從而使用極少的標記數據完成任務。

表 1.1：使用完整資料集時，P 波到時拾取的實驗結果

原始實驗模型效能比較		P 波 Picking 成績					
Model	Dataset	Precision	Recall	F1	Mean	Std.	MAE
EQTransformer	Dev	0.9379	0.9824	0.9596	0.0009	0.0210	0.0051
	Test	0.9118	0.9741	0.9419	0.0017	0.0252	0.0064
PhaseNet	Dev	0.9320	0.9453	0.9386	0.0008	0.0259	0.0061
	Test	0.9106	0.9393	0.9247	0.0010	0.0280	0.0067
LEM Freeze	Dev	0.9439	0.9668	0.9552	0.0005	0.0222	0.0056
	Test	0.9257	0.9513	0.9383	0.0007	0.0251	0.0064
LEM Weighted Sum	Dev	0.9381	0.9821	0.9596	0.0015	0.0216	0.0053
	Test	0.9234	0.9698	0.9460	0.0010	0.0249	0.0064
LEM Finetune	Dev	0.9521	0.9926	0.9719	0.0001	0.0162	0.0038
	Test	0.9286	0.9871	0.9570	0.0003	0.0199	0.0047

表 1.2：使用完整資料集時，規模估算的實驗結果

原始實驗模型 效能比較		所有規模值實驗結果				規模值大於 4 之實驗結果			
		MAE	MSE	Std.	R2	MAE	MSE	Std	R2
MagNet	Dev	0.3063	0.1818	0.2965	0.7137	0.3471	0.2291	0.3296	0.5339
	Test	0.4141	0.2919	0.3470	0.4511	0.4965	0.4004	0.3924	0.4523
LEM Freeze	Dev	0.2258	0.0934	0.2059	0.8630	0.2581	0.1199	0.2308	0.7289
	Test	0.2863	0.1374	0.2355	0.7789	0.3212	0.1717	0.2618	0.6069
LEM Weighted Sum	Dev	0.1943	0.0702	0.1801	0.8912	0.2209	0.0862	0.1933	0.7668
	Test	0.2488	0.1044	0.2062	0.8258	0.2792	0.1285	0.2248	0.6752
LEM Finetune	Dev	0.3010	0.1424	0.2276	0.7666	0.3625	0.1975	0.2571	0.4765
	Test	0.3970	0.2237	0.2572	0.5998	0.4536	0.2854	0.2821	0.2548

表 1.3：使用 0.01%資料集時，P 波到時拾取的實驗結果

僅使用 0.01% 資料實驗模型效能比較		P 波 Picking 成績					
		Precision	Recall	F1	Mean	Std.	MAE
EQTransformer	Dev	0.2399	0.8670	0.3759	0.0010	0.0557	0.0152
	Test	0.2329	0.8597	0.3665	0.0012	0.0560	0.0154
PhaseNet	Dev	0.1935	0.9889	0.3237	0.1971	0.0728	0.2094
	Test	0.1863	0.9820	0.3133	0.1983	0.0725	0.2105
LEM Freeze	Dev	0.9183	0.9418	0.9299	0.0059	0.0349	0.0102
	Test	0.9132	0.9202	0.9167	0.0062	0.0367	0.0110
LEM Weighted Sum	Dev	0.9211	0.9626	0.9414	0.0077	0.0346	0.0110
	Test	0.9089	0.9475	0.9278	0.0082	0.0373	0.0120
LEM Finetune	Dev	0.9276	0.9673	0.947	0.0006	0.0298	0.0076
	Test	0.9034	0.9563	0.9291	0.0007	0.0328	0.0082

表 1.4：使用 0.01%資料集時，規模估算的實驗結果

僅使用 0.01% 資料實驗模型效能比較		Normal				M4 (mag>=4)			
		MAE	MSE	Std.	R2	MAE	MSE	Std	R2
MagNet	Dev	0.6365	0.6783	0.5226	-39.144	0.7577	0.9154	0.5842	-28.589
	Test	0.6226	0.6833	0.5438	-64.500	0.7650	0.9616	0.6135	-55.756
LEM Freeze	Dev	0.5233	0.4464	0.4154	0.3815	0.6350	0.6258	0.4718	0.0729
	Test	0.6046	0.5922	0.4760	0.0814	0.7323	0.8138	0.5267	-0.2885

LEM Weighted Sum	Dev	0.4191	0.3013	0.3545	0.4719	0.5149	0.4251	0.4000	0.0065
	Test	0.4887	0.3961	0.3966	0.2331	0.5922	0.5384	0.4332	-0.2811
LEM Finetune	Dev	0.5062	0.4214	0.4065	0.3323	0.6183	0.5942	0.4603	-0.1303
	Test	0.5924	0.5577	0.4548	-0.0076	0.7117	0.7573	0.5007	-0.5914

第二章：預訓練單測站地震預警模型之研究

2.1 背景

過去幾年的研究已成功在音訊和語音識別等領域中應用預訓練技術，如“BEATs: Audio Pre-Training with Acoustic Tokenizers”這篇論文所提出的 BEATs[1] 模型。BEATs 利用聲學分詞器將音訊轉化為高維度的語音表徵，進而在無標籤數據上進行大規模的自監督學習，有效捕捉了音訊的時頻特徵與語義關係。這樣的技術具備強大的特徵學習能力，能夠提取豐富的時序資訊，並且具有良好的泛化能力。

鑒於地震數據與音訊數據在時序和頻譜特性上有相似之處，本研究將嘗試將 BEATs 應用於地震數據的預訓練模型，期待透過此模型的自監督學習能力，有效地擷取地震波形中的特徵，使區域型地震預警能夠更加準確。

2.2 方法

BEATs (Bidirectional Encoder representation from Audio Transformers) 預訓練框架分成 Tokenizer 和 SSL model 兩個部分，它們會在每一個迭代中不斷優化。每次迭代，Tokenizer 會使用未標記的地震波型資料來生成離散標籤，再拿離散標籤來訓練 SSL model，等到 SSL model 收斂後，把 SSL model 當成老師來訓練一個新的 Tokenizer，重複此流程，如圖 2.1 所示。

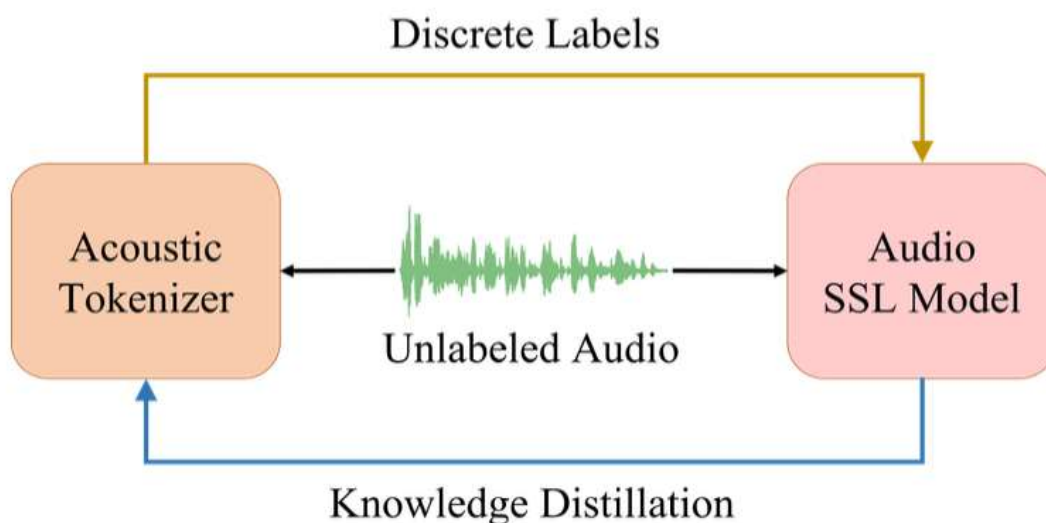


圖 2.1：BEATs 的迭代音訊預訓練示意圖

第一次迭代，由於沒有老師模型可以來訓練 Tokenizer，因此採用 Random-Projection Tokenizer，包含了一個 linear projection layer 和一組 codebook embeddings。從輸入的地震波型資料中提取 patch sequence，使用隨機初始化的 projection layer 將 patch sequence 轉換成投影向量，然後從隨機初始化的 codebook embeddings 向量中，尋找與投影向量最近的鄰居向量，將離散標籤定義為鄰居向量的索引，如圖 2.2(a)所示。

從第二次迭代開始，利用上一次迭代的 SSL model 作為老師模型，來訓練當前迭代的 Tokenizer，稱為 Self-Distilled Tokenizer。首先將 input patches 輸入進十二層的 Transformer encoder，獲得編碼的向量序列，每個編碼向量再從 codebook embeddings 中找到最近鄰居向量來進行量化，最後將量化後的向量序列作為輸入，使用一個三層 Transformer estimator 來預測老師模型的最後一層輸出。訓練目標為 tokenizer estimator 的輸出序列和老師模型的輸出序列之間的余弦相似度，以及編碼向量序列和量化向量序列之間的均方誤差，如圖 2.2(b)所示。

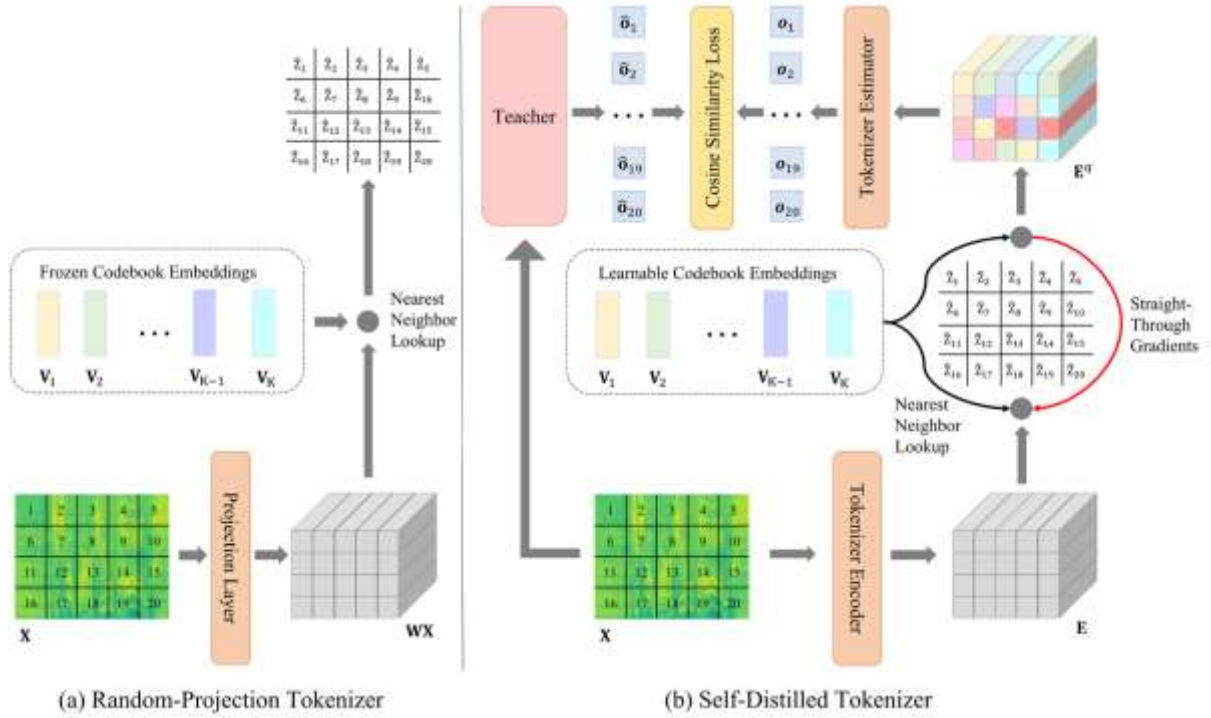


圖 2.2：用於生成離散標籤的 Tokenizer 模型架構圖

SSL model 以 ViT[2] 架構作為主幹網絡，提出一個 Masked Audio Modeling(MAM)任務用於 SSL model 的預訓練。一開始先隨機遮罩 75% 的 input patch sequence，然後將未遮罩的 patch sequence 輸入 ViT encoder，獲得 encoded representations。再將未遮罩 encoded representations 和遮罩 patch features 組合，最後輸入進 label predictor 來預測離散標籤。而訓練目標為遮罩的 cross entropy loss，如圖 2.3(a)所示。

而在 fine-tuning 中，SSL model 丟棄了 label predictor，並再加上六層 Transformer[3] encoder 和 linear layer，來預測 PGA。輸入為單測站的地震波型資料，然後隨機遮罩 50% 後，將其分為 patch sequence，與預訓練不同，會將整個 patch sequence 輸入進預訓練好的 SSL model(無 label predictor)，生成 encoded representations，除了將 encoded representations 的輸入進六層 Transformer encoder 外，額外再加入一個向量輸入，用來預測 PGA，最後將新加入的向量的輸出輸入到 linear layer，產生預測的 PGA 值，Loss 用真實 PGA 和預測 PGA 的 MSE 和 MAE 相加來做訓練，如圖 2.3(b)所示。此外，baseline model 為六層 Transformer encoder 和 linear layer(即移除微調模型中預訓練的 SSL model)。

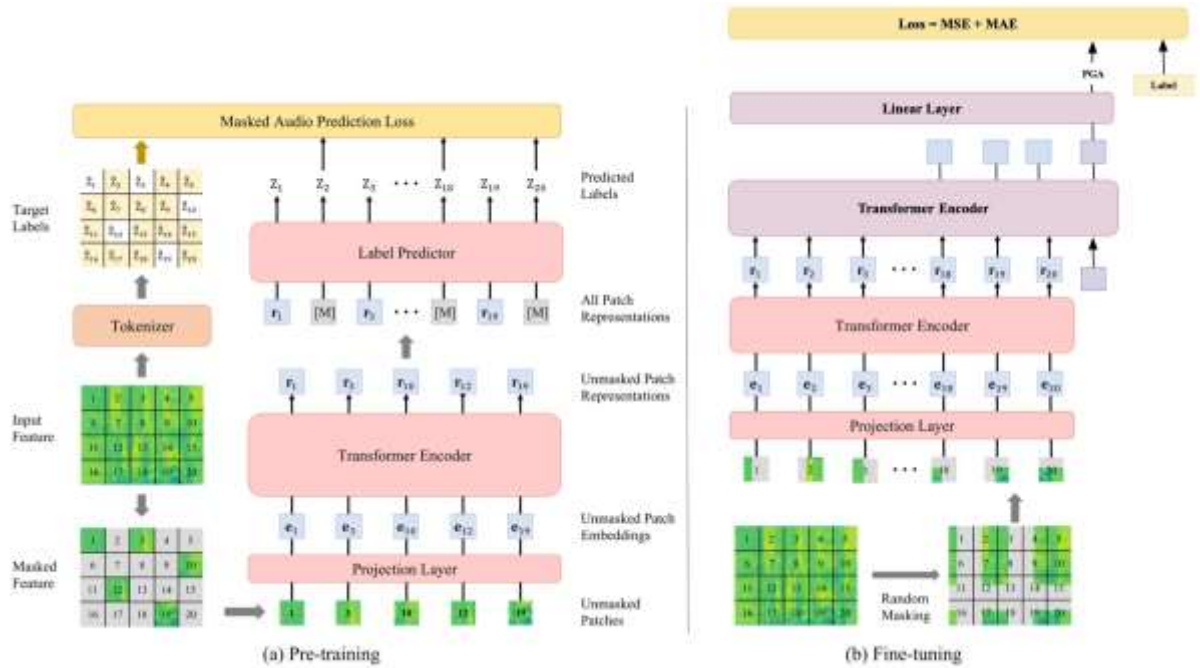


圖 2.3：SSL model 在預訓練和微調的模型架構圖

2.3 實驗結果

在預訓練和微調中，使用 P 波到達後 1s、2s、3s、4s、5s 的波型資料，個別對每個波型時間訓練特定模型，且針對每個特定模型也訓練了輸入資料有無被 mask 的兩種模型。

表 2.1：在測試集上，每個模型的 MSE 分數

Time window	1s	2s	3s	4s	5s
Baseline	0.21989	0.20731	0.22036	0.228870	0.228875
Baseline (mask)	0.22865	0.21459	0.17920	0.20492	0.19813
Pre-training	0.21056	0.17527	0.17208	0.17201	0.16504
Pre-training (mask)	0.18304	0.17663	0.18565	0.18050	0.15354

表 2.2：在測試集上，每個模型的 MAE 分數

Time window	1s	2s	3s	4s	5s
Baseline	0.02336	0.02347	0.02346	0.02358	0.02359
Baseline (mask)	0.02356	0.023338	0.02217	0.02309	0.02428
Pre-training	0.01979	0.01837	0.01796	0.01721	0.01627
Pre-training (mask)	0.01954	0.01855	0.01898	0.01776	0.01791

由表 2.1 和表 2.2 可看出，在 1s~5s 的 MSE 和 MAE 分數，Pre-training model 的表現皆比 Baseline model 好，而在 Pre-training model 成績中，大部分輸入沒有被 mask 的模型結果比輸入被 mask 的模型好。

為了更貼近實際的預警情況，因此還需要評估模型的 Accuracy、Precision、Recall、F1-score。FP 為預測 PGA 和實際 PGA 均 ≥ 25 Gal；TN 為預測 PGA 和實際 PGA 均 < 25 Gal；FP 為預測 PGA ≥ 25 Gal，但實際 PGA < 25 Gal；FN 為如果預測 PGA < 25 Gal，但實際 PGA ≥ 25 Gal。再使用 FP、TN、FP、FN 去計算出 Accuracy、Precision、Recall、F1-score 來進行比較。

表 2.3：在測試集上，每個模型的 Accuracy 分數

Time window	1s	2s	3s	4s	5s
Baseline	0.99	0.99	0.99	0.99	0.99
Baseline (mask)	0.99	0.99	0.99	0.99	0.99
Pre-training	0.992	0.993	0.994	0.994	0.994
Pre-training (mask)	0.991	0.993	0.993	0.993	0.993

表 2.4：在測試集上，每個模型的 Precision 分數

Time window	1s	2s	3s	4s	5s
Baseline	1.0	1.0	1.0	0.0	0.0
Baseline (mask)	1.0	1.0	1.0	1.0	1.0

Pre-training	0.837	0.733	0.871	0.789	0.879
Pre-training (mask)	0.568	0.788	0.879	0.849	0.952

表 2.5：在測試集上，每個模型的 Recall 分數

Time window	1s	2s	3s	4s	5s
Baseline	0.014	0.006	0.003	0.0	0.0
Baseline (mask)	0.003	0.007	0.044	0.006	0.099
Pre-training	0.288	0.524	0.438	0.564	0.504
Pre-training (mask)	0.487	0.409	0.355	0.432	0.35

表 2.6：在測試集上，每個模型的 F1-score 分數

Time window	1s	2s	3s	4s	5s
Baseline	0.027	0.013	0.006	0.0	0.0
Baseline (mask)	0.006	0.014	0.085	0.013	0.18
Pre-training	0.429	0.611	0.583	0.658	0.64
Pre-training (mask)	0.524	0.538	0.506	0.573	0.512

由表 2.3、表 2.4、表 2.5、表 2.6 得知，Pre-training model 的 Accuracy、Recall、F1-score 成績皆高於 Baseline model，雖然在 Precision 上，Baseline model 分數高於 Pre-training model，但在 Recall 中，Baseline model 明顯低於 Pre-training model，表示 Baseline model 有很多 FN 的情況。此外，Pre-training model 在 2-5s 的成績，輸入沒有被 mask 的模型表現比輸入被 mask 的模型好。由上述結果可以推論，預訓練模型能夠有效地幫助單測站地震預警模型提高預警的準確性。

第三章：基於檢索增強之區域型地震預警模型

3.1 介紹

本研究聚焦於臺灣地區的區域型地震預警模型。基於先前計劃所構建的全測站模型，進一步引入歷史事件檢索技術，以提升模型對地震事件的預測準確性。具體而言，通過檢索歷史事件庫，篩選出與當前地震事件最相似的歷史事件，並將這些歷史數據中的特徵融合至當前事件的預測過程中。這種方法不僅能夠保持過去模型對地震預測的靈活性與適應性，還顯著提升了預警準確率。

3.2 背景與目的

地震作為一種破壞性極強的自然災害，對全球經濟和人類生命安全造成了深遠的影響。由於地震的不可預測性和巨大破壞力，地震研究在減少災害風險和提升基礎設施抗震能力方面顯得尤為重要。在此背景下，早期地震預警系統(Earthquake Early Warning System, EEWs)逐漸成為當代減災技術中的一個關鍵組成部分。早期地震預警系統基於地震波不同類型(P波和S波)的傳播速度差異，通過快速檢測初始的P波來預估破壞性較強的S波到達時間，並向目標區域發送警報，為應急行動爭取寶貴時間。

近年來，自然語言處理領域的研究逐漸以預訓練模型為核心，而檢索增強生成技術(Retrieval-Augmented Generation, RAG)更成為提升模型效能的熱門方法[4][5][6]。檢索增強生成技術模型將預訓練好的生成式模型與檢索系統相結合，透過從外部知識庫或資料庫中檢索相關訊息輔助生成內容，從而提升生成結果的準確性和相關性。此技術的優勢在於能夠利用外部知識來補足僅依賴語言模型的不足，特別適用於需要專業知識的場景。

檢索增強生成技術的工作流程通常分為兩個階段：首先是「檢索階段」，該階段根據輸入內容查詢大型資料庫以找到相關資訊；隨後進入「生成階段」，生成式模型在檢索結果的基礎上生成更加準確且具備豐富資訊的文句。這種雙階段的方法結合了檢索與生成的優勢，使模型在處理長文句或特定領域問題時更具效果。

在地震領域的應用中，檢索增強生成技術方法展現了十足的潛力。由於地震研究涉及大量歷史資料、震源機制與地質結構等專業知識，僅依賴通用模型進行震度預測可能無法提供更高品質的結果。透過檢索相關的地震資料，檢索預測可以在應用時提供更精準的預測，提升在地震領域的應用價值。

過去的研究計劃中已訓練出一套全測站模型，該模型以Transformer作為基底，確保了測站數量的擴充性以及地震預警的準確性。在此基礎上，本研究結合檢索增強生成技術與地震資訊，實現了區域型地震預警的即時應用。不僅在區域型震度預估上保持了高提前預警時間，還顯著提升了預測準確性，為地震震度預測模型的設計提供了一種創新思路。

3.3 方法

本研究的方法始於地震事件資料的收集與整理，其中地震事件總數以 E 表示。針對每個事件，根據測站的座標軸來統計參與觀測的測站數量，記為 N ，以反映觀測資料的空間分布。

每個波形 $W \in R^{3 \times T}$ 的總長度為 T 個樣本，並包含三個通道 (channel)，對應到地震儀器偵測的三個維度。為了保證所有測站資料的一致性，我們根據第一個測站的體波到時 (primary wave pick time) 對所有測站的波形進行對齊。對於每條波形，選取第一個測站到時前 5 秒以及到時後 25 秒的波形段。

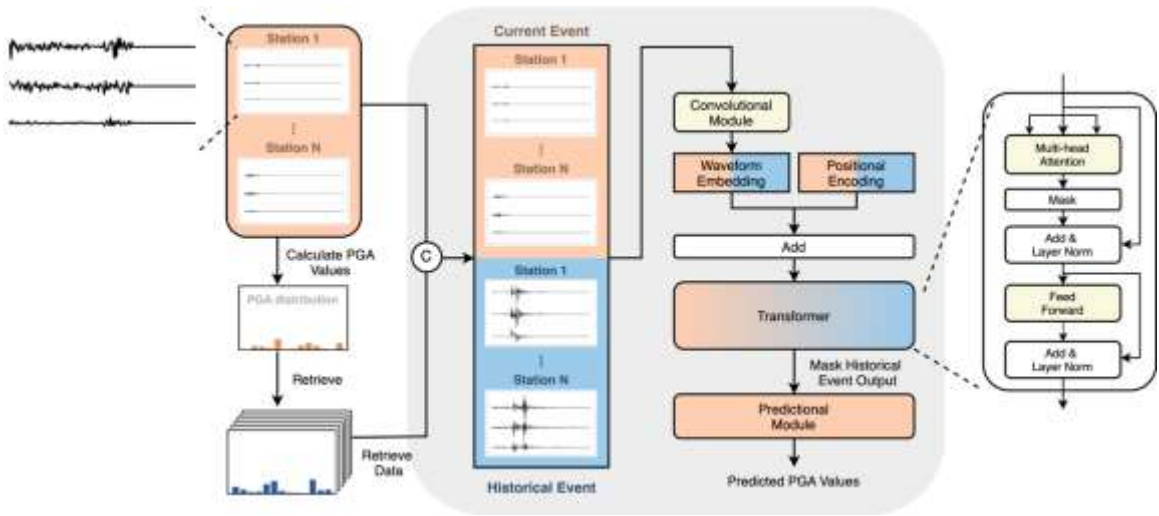


圖 3.1 : eq-RAG 模型架構圖

本方法分為兩個主要階段：檢索和預測。在檢索階段，根據當前事件與歷史地震事件的相似性，從資料庫中檢索出相關的歷史地震數據。接著在預測階段，模型利用當前的波形數據以及檢索到的歷史事件訊息進行地震預測。模型架構如圖 3.1 所示。

1. 檢索階段

- (1) 建立每個事件在各個時間點的時序歷史事件資料庫，由於任務是做在即時系統之上，我們希望檢索到的歷史事件是基於相同的間隔時間，因此我們每隔一段時間就建立一個時間點的時序歷史事件資料庫，共建立 U 次。 $D = \{D_i\}_{i=1}^U$ ，間距為 T/U 。其中 $D_i = \{e_1, e_2, \dots, e_E\}$, $e_k \in R^N$ 。
- (2) 針對每個測站的三軸波形，根據當下的時序歷史事件資料庫所處時間點應用遮罩操作，將時間點之後的波形設置為零。接著，對波形計算其三軸的

L2 正規化，從中選取最大值。通過這一過程，我們生成了向量 e_i ，代表事件的測站 PGA 測站分佈。以下公式具體展示了單個測站進行波形遮罩的過程， $w(j)$ 代表第 j 個採樣點 (sample) 的波。

$$W_i(j) = \begin{cases} w(j), & 1 \leq j \leq \frac{T}{U} \cdot i \\ 0, & \frac{T}{U} \cdot i < j \leq T \end{cases}$$

(3) 計算當下事件向量 c_i ，計算 PGA 流程相同。

(4) 計算最相關事件：在相似度計算中，存在多種方法，每種方法都有其不同的特性和應用場景。由於我們無法預先確定哪種方法能夠使模型表現最佳，因此選用了三種不同的相似度度量方法，如下列公式所示。具體而言，L2 距離越小表示兩個事件越相似，而 cosine 相似度 (Cosine Similarity) 和內積相似度 (Inner Product) 越大則代表相似度越高。

$$L2 \text{ Distance} = \sqrt{\sum_{i=1}^N (c_i - e_i)^2}$$

$$Cosine \text{ Similarity} = \frac{\sum_{i=1}^N c_i e_i}{\sqrt{\sum_{i=1}^N c_i^2} \cdot \sqrt{\sum_{i=1}^N e_i^2}}$$

$$Inner \text{ Product} = \sum_{i=1}^N c_i e_i$$

在模型訓練階段 (training phase)，我們將當前事件與所有歷史事件進行相似度計算，並從相似度最高的前 K 個事件中隨機選取一個作為後續模型輸入的選項，其中是有機率挑選到當次事件。這一策略旨在增加模型訓練的數據多樣性，從而提高模型的泛化能力。在解碼階段 (decoding phase)，則直接選擇與當前事件最相似的歷史事件作為參考。

為了實現高效的相似度搜索，整個過程依賴於 Facebook 於 2020 年提出的 FAISS 庫 [7]，該工具顯著縮短了實時運算的時間，提升了系統的處理效率。

2. 預測階段

此階段，我們專注於兩個事件如何在神經網路模型架構上做融合。我們把兩事件全部測站的波形和座標表示為 $\{W_{11}, \dots, W_{1N}\}, \{W_{21}, \dots, W_{2N}\}, \{G_{11}, \dots, G_{1N}\}, \{G_{21}, \dots, G_{2N}\}$ 。

(1) 合併層：

$$W_{con} = \text{concat}([W_{11}, \dots, W_{1N}]; [W_{21}, \dots, W_{2N}])$$

$$G_{con} = \text{concat}([G_{11}, \dots, G_{1N}]; [G_{21}, \dots, G_{2N}])$$

(2) 卷積模組 (Convolutional Module) 與地理測站做標嵌入：

此模組具有多層卷積神經網路，所有測站皆經過相同的卷積模組中，這項操作將波形的高維特徵進行降維處理，轉換為低維度的嵌入向量。這一過程有助於提取波形的關鍵特徵，同時降低計算成本。

$$W'_{con} = \text{ConvolutionalModule}(W_{con})$$

與此同時，另一部分特徵則被用來處理測站座標嵌入 (station coordination embedding)。該嵌入過程將測站的地理位置資訊進行編碼 (encode)，為了後續捕捉不同測站之間的空間關聯性。

$$G'_{con} = \text{PositionalEncoding}(G_{con})$$

這兩部分嵌入特徵最終將融合在一起，為後續的 Transformer 提供輸入，從而更好地捕捉波形與測站位置資訊之間的關係。

$$H_{con} = \text{ADD}(W'_{con}; G'_{con})$$

(3) 變形模組 (Transformer)：

在變形模組模型中，我們使用多頭注意力機制，這允許模型從不同的特徵空間中學習注意力權重。其中，我們把變形模組的輸入稱作 X ， $X \in \mathbb{R}^{2N \times d_{model}}$ 。這裡 $2N$ 代表串接過後的總測站數， d_{model} 是 X 中每個特徵向量的維度。注意力機制會將查詢矩陣 (Q)、鍵矩陣 (K) 和值矩陣 (V) 分成多個頭進行計算，然後將各頭的輸出拼接起來。對於每一個頭 $h = 1, 2, \dots, H$ ，查詢、鍵和值矩陣的計算為：

$$Q_h = XW_{Q_h}, \quad K_h = XW_{K_h}, \quad V_h = XW_{V_h}$$

其中：

$$W_{Q_h} \in \mathbb{R}^{d_{model} \times d_k}$$

$$W_{K_h} \in \mathbb{R}^{d_{model} \times d_k}$$

$$W_{V_h} \in \mathbb{R}^{d_{model} \times d_v}$$

這裡， $d_k = d_v = \frac{d_{model}}{H}$ ，每個頭的維度被劃分為 $\frac{d_{model}}{H}$ ，然而，當輸入資料中存在無效或冗餘的部分時，直接計算這些位置的權重會導致無效資訊干擾模型的學習過程，我們的實驗也驗證了這一點。因此，我們採用了遮罩機制來排除這些無效位置，保證模型只關注有效資料。具體而言，對於每個時間步，對輸入的三軸波形應用遮罩操作，該操作將時間點之後的波形設為無效。使用 mask M 來表示這些位置是否應被遮罩，定義如下：

$$M = \begin{cases} 1, & \text{if the entire waveform is 0} \\ 0, & \text{otherwise} \end{cases}$$

此遮罩操作應用於查詢與鍵的點積操作之後。在進行 softmax 計算之前，我們對遮罩位置的點積值減去一個非常大的值，這確保了在 softmax 計算中，這些位置的輸出權重趨近於零。這一過程可具體表示為：

$$A_{ij} = \frac{\exp\left(\frac{Q_i K_j^T}{\sqrt{d_k}} + M_{ij} \cdot (-\infty)\right)}{\sum_j \exp\left(\frac{Q_i K_j^T}{\sqrt{d_k}} + M_{ij} \cdot (-\infty)\right)}$$

其中， Q_i 和 K_j 分別表示查詢向量 (query) 和鍵向量 (key)，而 d_k 是鍵向量的維度。當 $M_{ij}=1$ 時，將相應的位置設置為負無窮大，從而在 softmax 中將這些無效位置的權重計算為零。

3. 訓練細節

如圖 3.2 所示，此任務的 PGA 分佈非常不均，為了緩解 PGA 值分佈不均的問題，本研究採用了三種策略來進行處理，分別是大規模事件的上採樣、高測站數事件的上採樣，以及加權損失 (Weighted loss)。

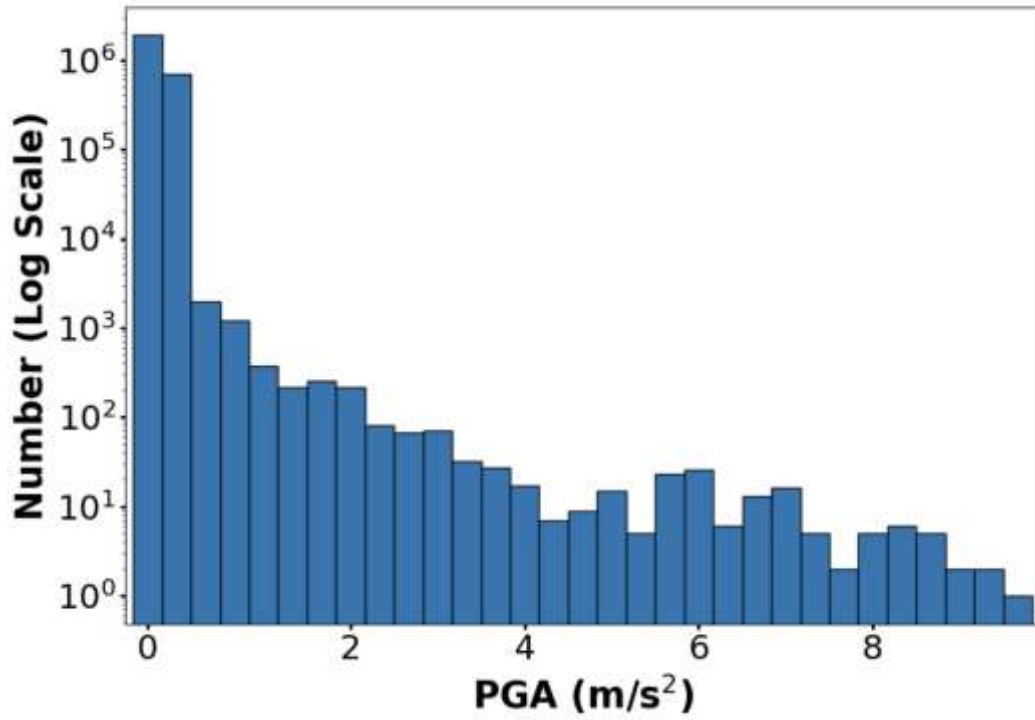


圖 3.2: 所有事件之 PGA 統計分佈圖。橫軸為測站 PGA 值，縱軸為取對數之統計次數。

(1) 大規模事件的上採樣：

首先，我們對每個事件的規模進行分區。規則可以表示為：

$$\text{bin}(\text{mag}_i) = \lfloor \text{mag}_i \rfloor$$

根據每個分區 $\text{bin}(\text{mag}_i)$ ，我們採取不同的上採樣次數。具體的做法如下：

$$\text{upsample_number} = \lceil (\text{resampling_factor})^{\text{bin}} \rceil$$

這裡的 resampling_factor 是一個可調參數，用來控制上採樣比例；而 bin 是規模的分區結果。

(2) 高測站數事件的上採樣：

$$\text{upsample_number} = \left\lceil \frac{n_{\text{event_stations}}}{st_resample_factor} \right\rceil$$

這裡的 $n_{\text{event_stations}}$ 代表事件中收錄的測站總數， $st_resample_factor$ 為可調參數，用來控制上採樣比例。

(3) 加權損失 (Weighted loss)：

我們根據各測站的標籤值進行分類，並根據不同類別賦予相應的權重。具體而言，對於接收到較大波幅的測站數據，我們賦予其更高的權重，以便模型在訓練過程中更加關注這些具有高 PGA 值的測站，從而平衡不同測站間的影響，改善模型在不均衡數據下的表現。具體作法我們根據測站的 PGA 值，將其分配到不同的類別。這些類別由一組閾值 $h_1, h_2, \dots, h_k$ 定義。每個測站的 PGA 值 PGA_i 被分到類別 c_i ，其分類函數 $f(PGA_i)$ 定義如下：

$$c_i = f(PGA_i) = \begin{cases} 0, & \text{if } PGA_i < h_1 \\ 1, & \text{if } h_1 \leq PGA_i < h_2 \\ \dots \end{cases}$$

我們將每個類別 c_i 對應一個權重 $w(c_i)$ ，根據分類結果，我們可以將各測站損失值 $Loss_i$ 給予不同權重值。具體的加權損失公式如下：

$$Weighted_Loss = \frac{1}{\sum_{i=1}^N w(c_i)} \sum_{i=1}^N w(c_i) \cdot Loss_i$$

3.4 資料集

本計畫採用了 CWASN 資料集。該資料集涵蓋了 2012 年至 2021 年間發生的 10,392 個地震事件，測站歷年統計共有 249 個，包含地表儀器和井下儀器和海底儀器，為研究提供了豐富的地震波形資料。資料劃分方式，我們將臺灣資料集依時間順序進行劃分。其中，2012 年 1 月 1 日至 2017 年 12 月 31 日的地震事件被用於訓練，2018 年 1 月 1 日至 2019 年 12 月 31 日的事件作為驗證集，而 2020 年 1 月 1 日至 2021 年 12 月 31 日的事件則作為測試集。此外，由於部分大地震事件會有測站斷網，導致波形資料為零的漏值情況。為了使學習更貼近真實情形，我們對於這些波形為零的資料，利用插值法對標籤震度進行補值處理。具體來說，我們將測站利用人工的方式挑選各測站最鄰近的 3 個平地測站，並盡可能確保位於三個鄰近測站圍繞形成的三角形中心點，以確保融合各個方位的標籤值。另外，當測站的標籤值為 0 時，對其最鄰近的 3 個測站進行非零平均。作法如下所示，假設某測站 S_i 的標籤震度為 y_i ，並且該測站因斷網導致波形資料為零。為了進行補值，我們選擇其鄰近的 3 個測站 S_{i1}, S_{i2}, S_{i3} ，並對這些測站的標籤震度 y_{i1}, y_{i2}, y_{i3} 進行非缺失值平均 (nanmean) 處理，得到補值 y' ，具體公式如下：

$$y' = \frac{1}{N'} \sum_{j=1}^3 y_{ij}$$

其中 N' 為參與計算的非缺失值 (NaN) 的測站數量。

3.5 實驗結果

成績部分我們和 TEAM[8] 作比較，同樣是基於 Transformer 架構下的區域型地震預警模型方法。此外，成績皆過濾掉海底以及井下儀器，只保留平地儀器做評估。

本實驗之混淆指標計算法如表 3.1 所示。

表 3.1: 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	
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

1. 主成績

表 3.2 展示了不同震度下，TEAM 和 eq-RAG 模型在 Precision、Recall 和 F1-score 指標上的整體表現差異。整體而言，eq-RAG 在各個比例上表現均優於 TEAM。在 Precision 上，整體皆優於 TEAM 方法；而 Recall 和 F1 則是在低震度（如 0.81% 和 2.5%）時，優勢更加顯著。

表 3.2：模型在各震度之 Precision、Recall、F1 成績

	TEAM			eq-RAG		
	Precision	Recall	F1	Precision	Recall	F1
0.81%	0.327	0.292	0.308	0.543	0.322	0.405
2.5%	0.372	0.353	0.363	0.590	0.444	0.507
8.1%	0.463	0.324	0.382	0.577	0.388	0.464
14%	0.294	0.248	0.269	0.500	0.318	0.389
25%	0.111	0.150	0.127	0.333	0.183	0.236

2. 距離劃分之成績

表 3.3 和 表 3.4 展示了 TEAM 和 eq-RAG 模型在不同震度（0.8%g 和 2.5%g）以及震中距離範圍（0-50 km、50-100 km、100-200 km 等）中的表現。可以看到，eq-RAG 模型在 Precision、Recall 和 F1-score 指標上的整體表現較

好，並且在各距離範圍和震度下的 **準確度 (Precision)** 均顯著優於 TEAM，尤其在遠距離表現更加突出，減少了誤報的可能性。這表示 eq-RAG 模型在處理事件時能更有效地降低遠距離的誤報問題，在應用於地震預警時具備更高的可靠性和穩定性。

表 3.3：在 0.8% g 震度下，距震央遠近對應之 Precision、Recall、F1 成績

	TEAM			eq-RAG		
	Precision	Recall	F1	Precision	Recall	F1
0-50 km	0.565	0.529	0.546	0.599	0.580	0.590
50-100 km	0.350	0.278	0.310	0.479	0.305	0.372
100-200 km	0.190	0.116	0.144	0.506	0.129	0.206
200-300 km	0.029	0.063	0.040	0.540	0.061	0.109
> 300 km	0.000	0.000	0.000	0.000	0.000	0.000

表 3.4：在 2.5% g 震度下，距震央遠近對應之 Precision、Recall、F1 成績

	TEAM			eq-RAG		
	Precision	Recall	F1	Precision	Recall	F1
0-50 km	0.538	0.451	0.491	0.599	0.532	0.563
50-100 km	0.283	0.349	0.312	0.561	0.447	0.498
100-200 km	0.077	0.038	0.051	0.576	0.118	0.196
200-300 km	-	-	-	-	-	-
> 300 km	-	-	-	-	-	-

3. PGA 預測分佈圖

表 3.5 和 表 3.6 展示了兩模型在不同時間下的 PGA 預測與真實值的回歸分佈圖。橫軸為真實的 PGA 值，縱軸為模型的預測值，對角線表示理想的預測情況。

從圖中可以看到，eq-RAG 整體相較 TEAM 更貼近對角線，表現出良好的回歸趨勢。

此外，在較短的觀測時間中，模型已表現出良好的回歸趨勢。這說明模型在短時間內提供一定的預測準確性。隨著時間延長到 8 秒和 16 秒，分佈進一步靠近對角線，顯示模型在更長觀測時間下的穩定性和精確度有所提升，但短時間的結果已經顯示出接近長時間的預測能力。

表 3.5 : eq-RAG 模型在各時間對應之 PGA 預測分佈圖

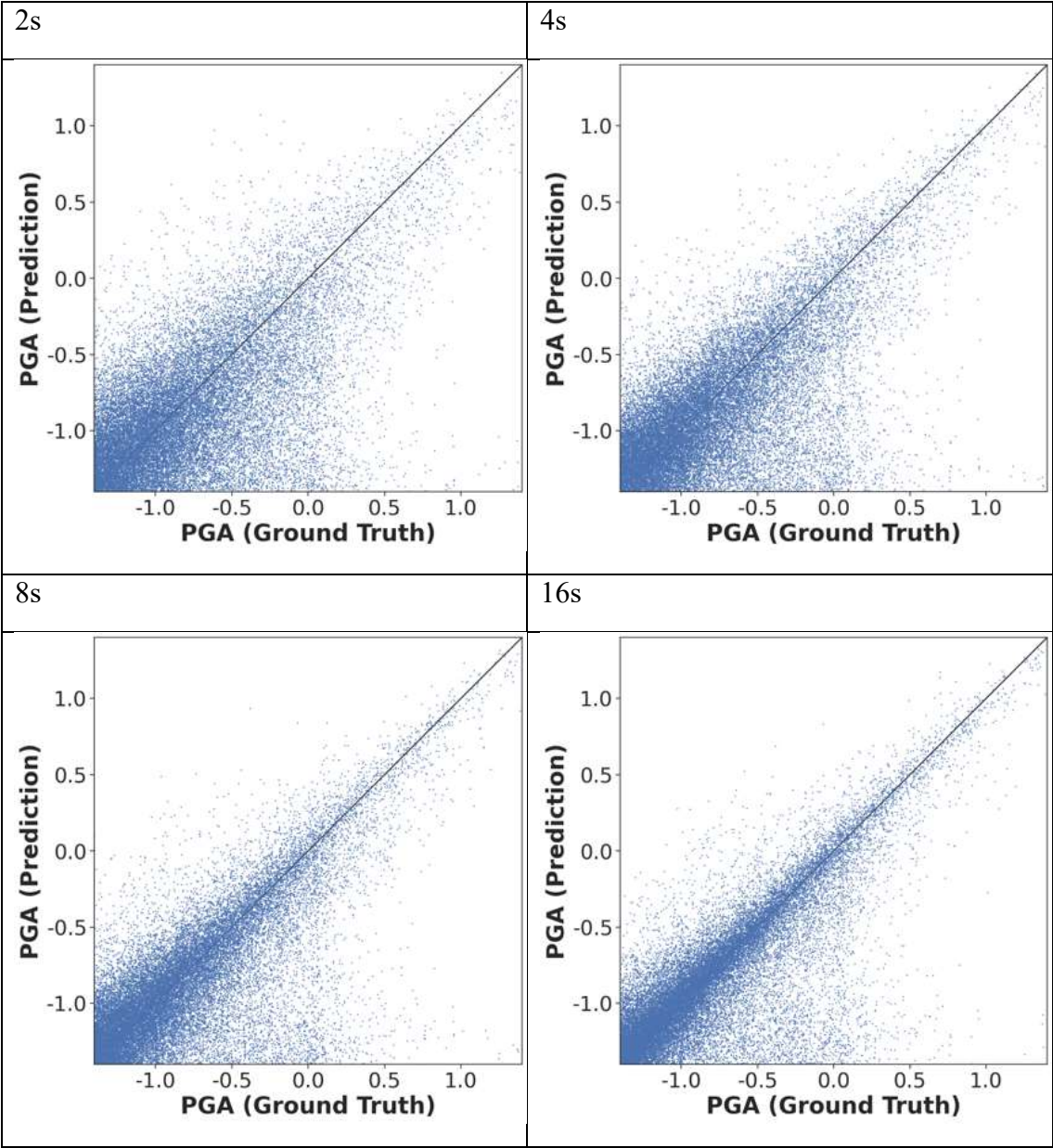
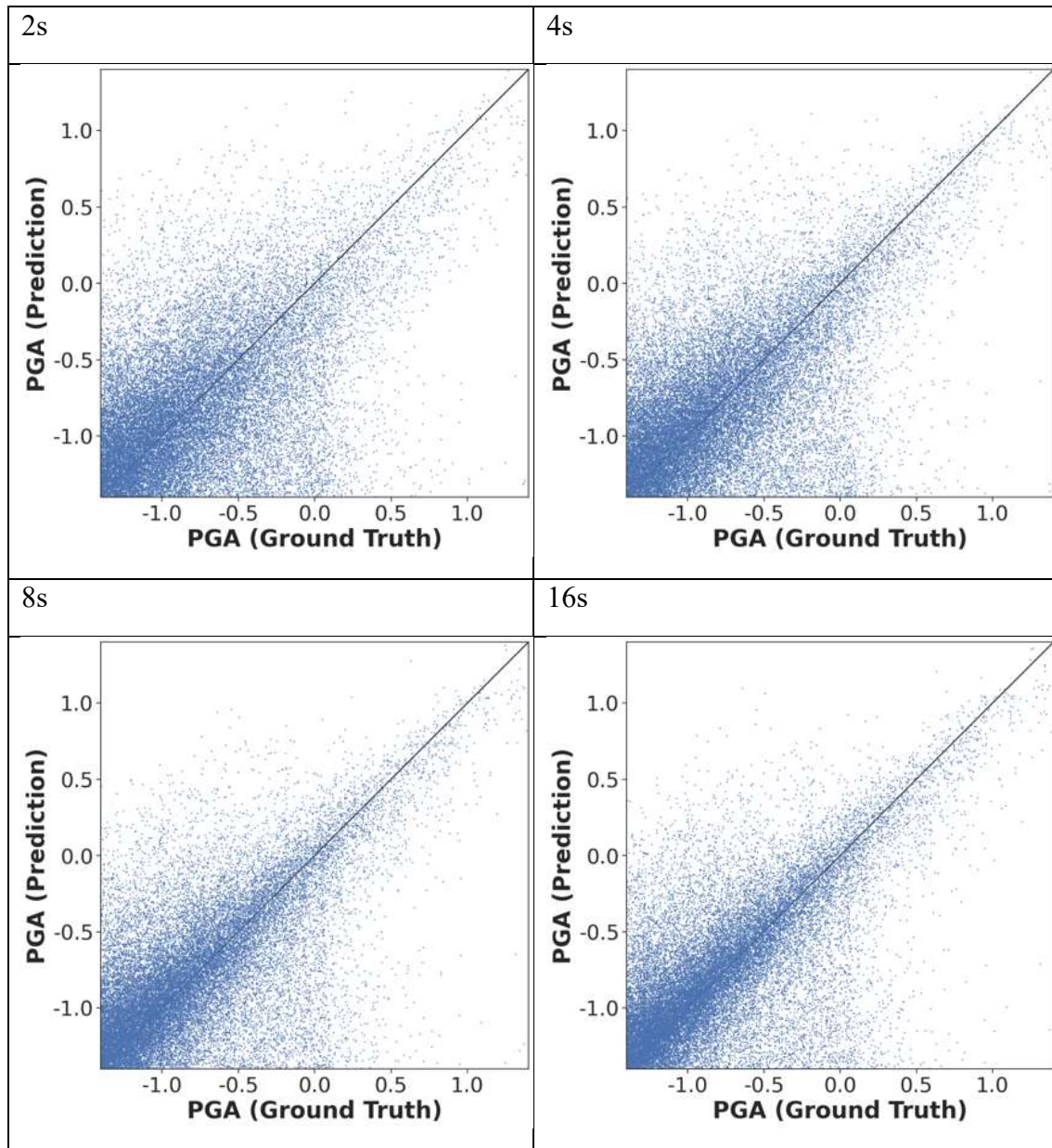


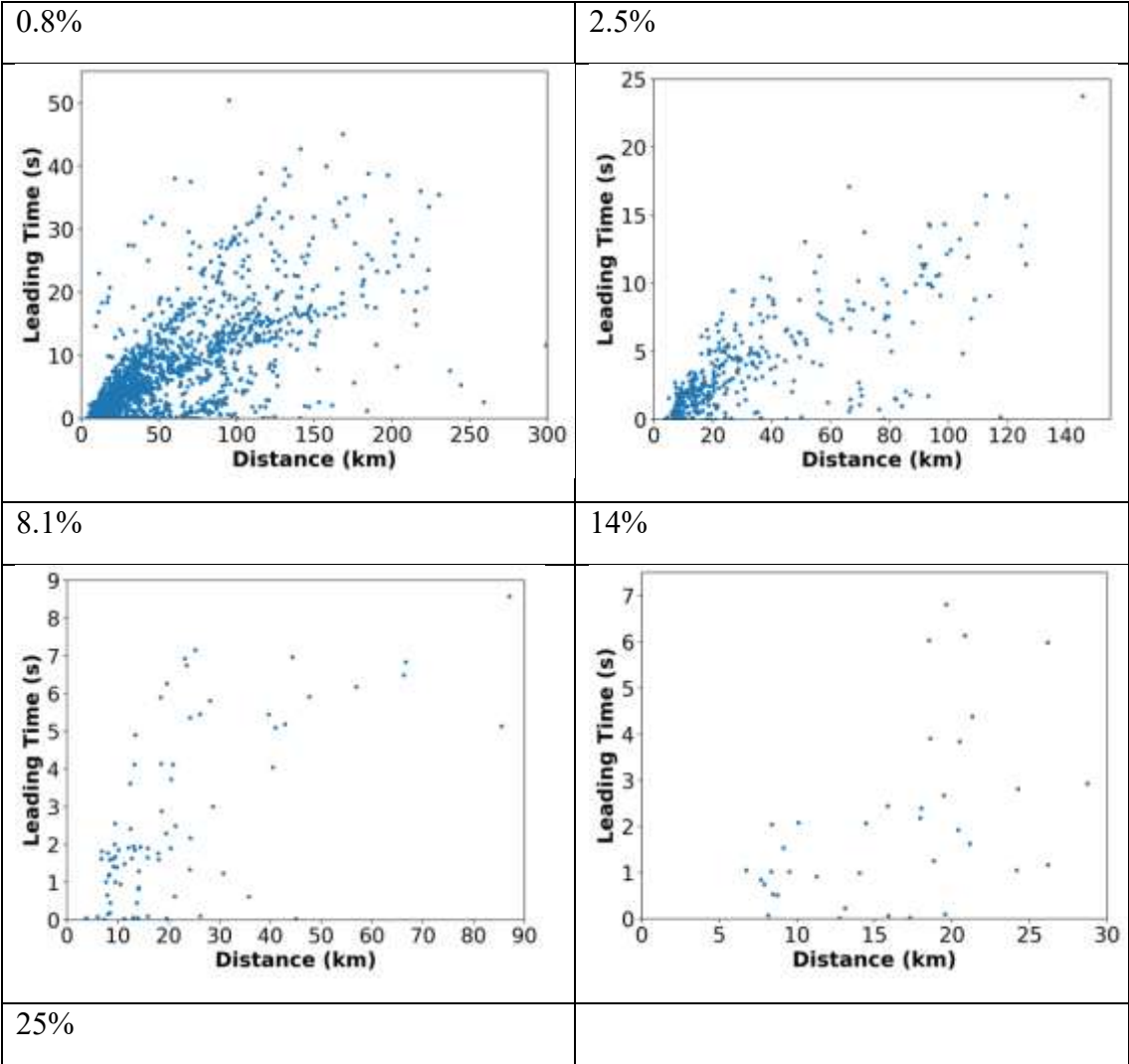
表 3.6 : TEAM 模型在各時間對應之 PGA 預測分佈圖



4. 不同距離下的提前警報時間分佈

表 3.7 和 表 3.8 展示了兩模型不同震度下，距離與提前警報時間之間的關係。可以看到隨著距離震央越遠，警報時間可以越長，eq-RAG 和 TEAM 具有相近的提前警報時間能力。

表 3.7 : eq-RAG 模型在不同震度下對應之提前警報時間分佈



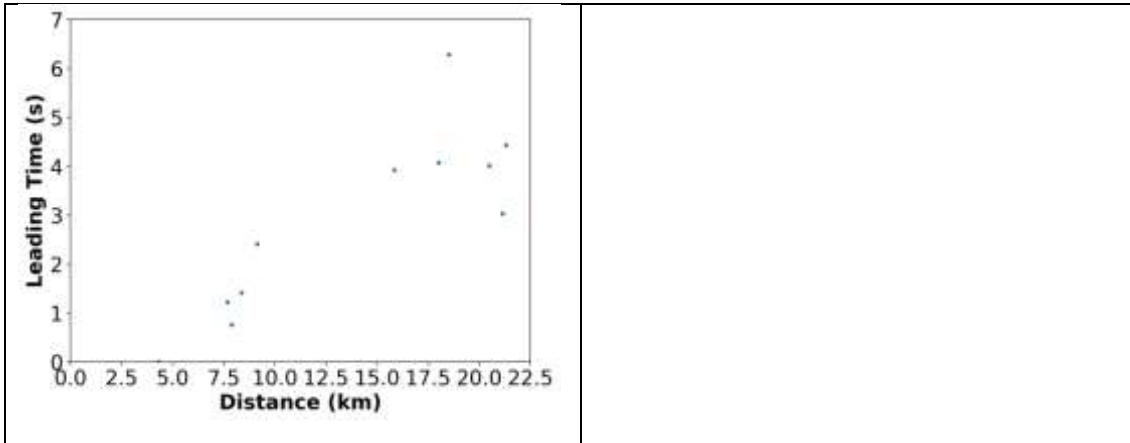
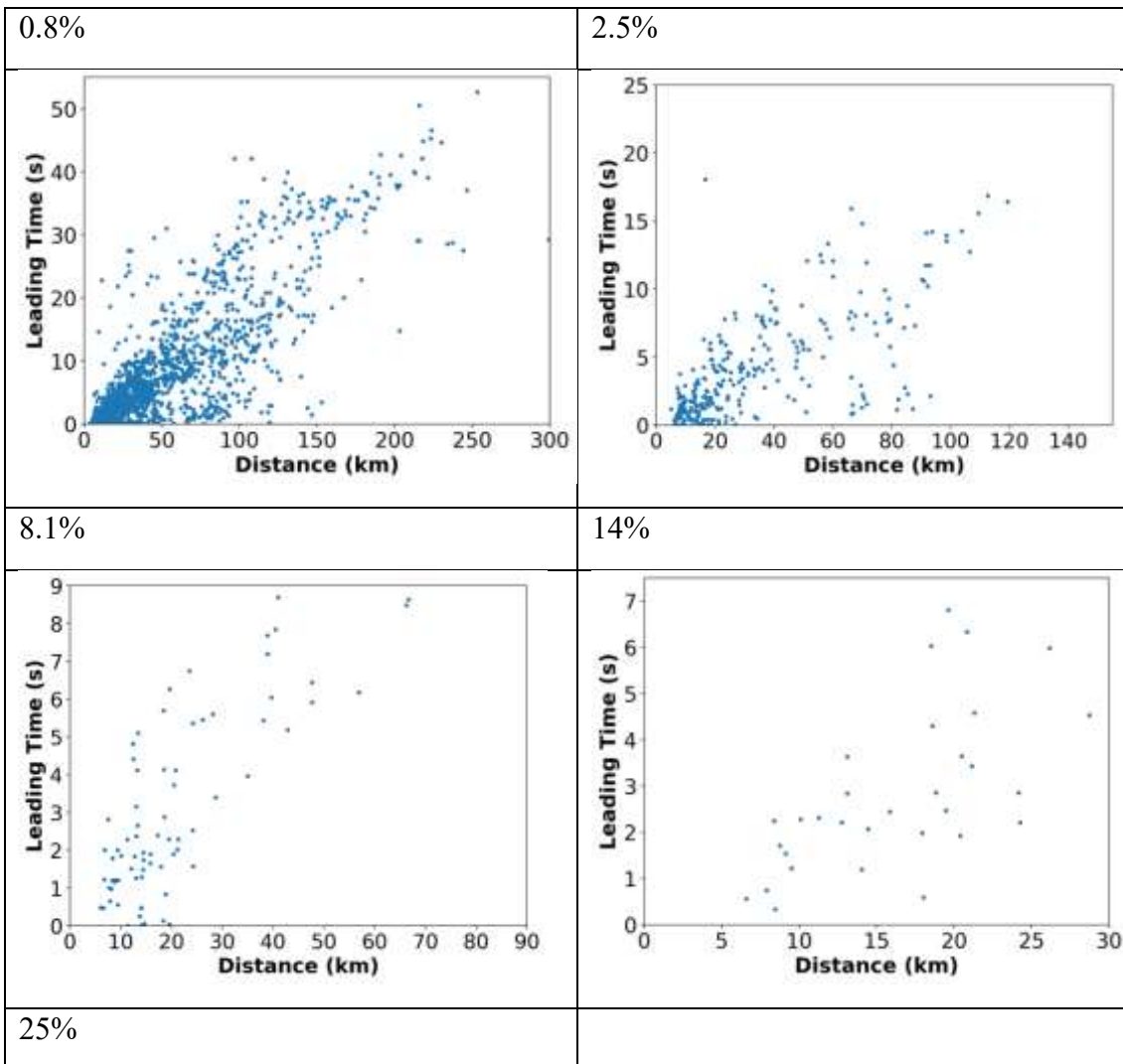
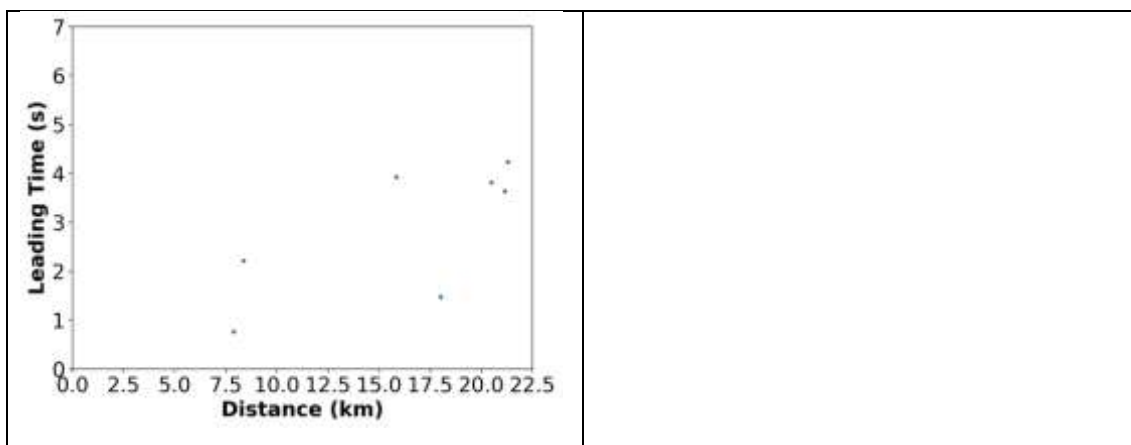


表 3.8：TEAM 模型在不同震度下對應之提前警報時間分佈





3.6 討論 - retrieval 公式選擇

表 3.9 展示了在不同震度下，三種 retrieval 公式（L2、Cosine Similarity 和 MIP）在 F1 分數上的比較。在較低震度比例下，Cosine Similarity 的 F1 分數表現相對較佳，超越了其他兩種方法。隨著震度比例增加，所有方法的 F1 分數均有所下降，但 Cosine Similarity 在各個震度比例下的表現仍較為穩定。基於這些結果，選擇 Cosine Similarity 作為 retrieval 的主要公式。

表 3.9：各檢索方法在各震度下之 F1 成績

	L2	Cosine Similarity	MIP
0.81%	0.404	0.405	0.404
2.5%	0.490	0.507	0.500
8.1%	0.446	0.464	0.450
14%	0.370	0.389	0.356
25%	0.208	0.236	0.195

dropout 層在第三卷積區塊後，為保留輸入維度，不使用池化層，使 ViT 能聚焦於 P 波到達時間的位置，卷積核大小為 11，三個卷積區塊的特徵圖數量分別為 10、20、40，ViT 的 patch 大小為 40 樣本，與三角形標籤大小一致，此設計引導 ViT 將最高權重放在 P 波到達時間上，利用注意力機制，在此配置下，patch 數為 150，嵌入層的投影維度為 40，模型使用四個 Transformer 編碼器(圖 4.1(f))及四個多頭注意力(見圖 4.1(g))，我們未在 Transformer 編碼器後使用 MLP 網路，而是將最終編碼器的輸出展平成一維，Transformer 編碼器的輸出維度為 150×40 ，展平後為 6000，與目標標籤尺寸一致，我們添加了一個一維卷積層，使用 sigmoid 激活函數，卷積核大小為 15，作為輸出層，此外，Transformer 編碼器後還增加了數個卷積區塊(圖 4.1(b))，進一步提取與 P 波到達時間相關的穩健特徵，增強注意力機制，提升模型表現。

表 4.1 輸入數據為 30 秒的結果

P-arrival	Mean	σ	Precision	Recall	F1-score
750	0.0360	0.1883	0.9880	0.9984	0.9932
1500	0.0288	0.1829	0.9886	0.9998	0.9942
2000	0.0273	0.1965	0.9890	0.9998	0.9943
2500	0.0285	0.1998	0.9886	0.9998	0.9942
2750	0.1556	0.4975	0.8810	0.9831	0.9293

4.2 實驗結果

表 4.1 總結了基於 ViT 模型在 STEAD 數據集上進行 P 波挑選的性能表現，分別測試了 P 波固定在第 750、1500、2000、2500 及 2750 個 sample 點的成績，分析了模型的精確率 (Precision)、召回率 (Recall) 和 F1 分數 (F1-score)，此外，我們還關注了平均誤差 (Mean) 和標準差 (σ)，以評估不同數據集規模下模型的準確性和穩定性，評估標準為：

正確預測(TP)：若模型預測的 P 波到達時間在正確答案的 ± 0.5 秒範圍內，視為正確預測(True Positive)。

誤報(FP)：若模型預測的 P 波到達時間超出 ± 0.5 秒範圍，但實際並無事件發生，則視為誤報(False Positive)。

漏報(FN)：若模型未能偵測到實際發生的 P 波到達事件，則視為漏報(False Negative)。

正確拒絕(TN):若模型正確判斷並無事件發生,則視為正確拒絕(True Negative)。

此模型表現出穩健的性能,當數據集規模介於 750 至 2500 樣本時達到最佳準確性,這個結果表明,Vision Transformer 在地震學應用中具有潛力,尤其在 P 波精確偵測上能夠取得優異的效果。

第五章：MQTT 的即時系統- Decision Making

在地震信號處理中,精確識別地震 P 波,對於地震檢測、早期預警及後續分析至關重要,為了提升地震速報系統 (EEW) 的檢測精度與穩定性,我們整合了多個不同的地震波到時 (picking) 方法,包括 EQTransformer、STA/LTA、GRADUATE 等不同的 picker 工具,以提供多角度的訊號判斷,並將其結果結合於 Decision Making 模型中,最後生成統一格式的報告,目前使用的是機器學習中的 XGBoost 方法,當 P 波檢測信號從各測站傳遞到系統時,系統將接收到的數據放入訊息隊列,接著系統使用 25 公里內的鄰近測站數據來獲取多站訊號的穩定性,再將從鄰近測站收集的數據傳入 XGBoost 模型,模型會根據輸入特徵生成一個預測結果,若預測結果符合發布條件,系統會生成並發送 P 波到達的通知。

参考文献

- [1] Sanyuan Chen, Yu Wu, Chengyi Wang, Shujie Liu, Daniel Tompkins, Zhuo Chen, Furu Wei, “BEATs: Audio Pre-Training with Acoustic Tokenizers,” arXiv preprint arXiv:2212.09058, 2022.
- [2] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby, “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net, 2021.
- [3] 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.
- [4] M. Lewis, W. Yih, T. Rocktäschel, S. Riedel, D. Kiela, “Retrieval- Augmented Generation for Knowledge-Intensive NLP Tasks”, *Neural Information Processing Systems*, 2020
- [5] Y. Gao, Y. Xiong, X. Gao, K. Jia, J. Pan, Y. Bi, Y. Dai, J. Sun, M. Wang, H. Wang, “Retrieval-Augmented Generation for Large Language Models: A Survey”, arXiv, 2312.10997, 2024
- [6] G. Izacard and E. Grave, “Leveraging Passage Retrieval with Generative Models for Open. Domain Question Answering,” in *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, 2021, pp. 874-880.
- [7] J. Johnson, M. Douze, and H. Jégou, “Billion-scale similarity search with GPUs,” *IEEE Transactions on Big Data*, vol. 7, no. 3, pp. 535-547, 2021.
- [8] 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.
- [9] O. M. Saad, Y. Chen, A. Savvaiddis, S. Fomel, and Y. Chen, “Real-time earthquake detection and magnitude estimation using vision transformer,” *Journal of Geophysical Research: Solid Earth*, vol. 127, no. 5, p.e2021JB023657, 2022.
- [10] S. M. Mousavi and G. C. Beroza, “Deep-learning seismology,” *Science*, vol. 377, no. 6607, eabm4470, 2022.
- [11] S. M. Mousavi and G. C. Beroza, “Machine learning in earthquake seismology,” *Annual Review of Earth and Planetary Sciences*, vol. 51, pp. 105–129, 2023.
- [12] J. Woollam, J. Munchmeyer, F. Tilmann, “ et al., “Seisbench—a toolbox for machine. learning in seismology,” *Seismological Society of America*, vol. 93, no. 3, pp. 1695–

1709, 2022.

- [13] E. Pardo, C. Garfias, and N. Malpica, “Seismic phase picking using convolutional networks,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 7086–7092, 2019.
- [14] W. Zhu and G. C. Beroza, “Phasenet: A deep-neuralnetwork-based seismic arrival-time picking method,” *Geophysical Journal International*, vol. 216, no. 1, pp. 261–273, 2019.
- [15] 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.
- [16] A. Mohamed, H.-y. Lee, L. Borgholt, et al., “Selfsupervised speech representation learning: A review,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 6, pp. 1179–1210, 2022.
- [17] A. Baevski, Y. Zhou, A. Mohamed, and M. Auli, “Wav2vec 2.0: A framework for self-supervised learning of speech representations,” *Advances in neural information processing systems*, vol. 33, pp. 12 449– 12 460, 2020.
- [18] Central Weather Administration, Central weather administration seismographic network, 2012. DOI: 10 . 7914/SN/T5.
[Online]. Available: <https://www.fdsn.org/networks/detail/T5/>.
- [19] 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. 179 464–179 476, 2019.
- [20] S. M. Mousavi and G. C. Beroza, “A machine learning approach for earthquake magnitude estimation,” *Geophysical Research Letters*, vol. 47, no. 1, e2019GL085976, 2020.
- [21] H. Jegou, M. Douze, and C. Schmid, “Product quantization for nearest neighbor search,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 1, pp. 117–128, 2010.
- [22] Z.-C. Chen, C.-L. Fu, C.-Y. Liu, S.-W. D. Li, and H.-y. Lee, “Exploring efficient-tuning methods in selfsupervised speech models,” in *2022 IEEE Spoken Language Technology Workshop (SLT)*, IEEE, 2023, pp. 1120–1127.

機器學習技術建立現地型地震預警系統相關研究(III)

子計畫三

同時進行標記 P 波到時與單站法地震預警之研究

許丁友 陳柏睿 田禹容 曾冠証 官竣昱 李泓翰 吳錡霖 許瑜宸 王之寬

國立臺灣科技大學

摘要

地震是一種難以預測且具有毀滅性的自然災害，對人類生命財產造成巨大威脅。因此，準確且快速地預測地震震度成為強震預警技術研發的首要目標之一。

雖然多站法（區域型）的強震預警技術可快速且準確地預測各地的震度，但由於需要數個測站接收到 P 波訊號後才能進行預測，導致近震央區域往往在強震波到達後才發出預警，其「盲區」半徑約為 50 至 60 公里。

先前研究已初步開發完成 CNN 單站法預測 PGA 的技術，並將 CNN 模型整合至 Earthworm 系統，在氣象署提供的設備上進行即時預警。然而，根據 0403 花蓮地震事件的分析發現，模型接收相位資訊的時間或是模型預測的時間皆耗時太長。本研究使用 Triton Inference Server 嘗試解決模型預測時間過長之問題，且在研究室運行時，模型預測時間相比於線上模型在 0403 事件之模型預測時間減少了約 2.3 秒。而模型接收相位資訊的時間過長之問題，本研究嘗試使用 MQTT 通訊協定處理，而在研究室運行時，模型接收相位資訊相比於線上模型在 0403 事件之模型預測時間反而多了約 0.7 秒。

此外，為了統計目前 CNN 模型於線上運行的狀況，本研究針對 2024 年 1 至 9 月的 9 筆地震事件進行單站分析，統計各事件的準確率與召回率，結果顯示 Precision 約為 78.89%、Recall 約為 67.56%。此外，並分析漏報與誤報的原因。漏報多由地震事件特性所致，部分事件初期波形較小，使得在 P 波到達後 2 秒內的預測結果偏低；此外，某些事件在主震前發生較小地震導致觸發位置並非主震，亦使得預測結果偏低。而誤報的主要原因為某些事件中程式記錄的 P 波起始點位因更新延遲而導致延後，使得預測值偏大，此部分程式之缺點已改善完成，等待之後的地震事件加以驗證。

關於網格化地震預警範圍，本研究探討了依據區域面積、人口數量及測站數量等方式劃分的可行性。結果顯示，若適當地利用面積和測站數量作為劃分基準，較能有效提升預測準確性，雖然存在少數漏報情形，但這些測站的 PGA 值仍在容許誤差範圍內。然而，過度細化網格可能導致形狀不規則，影響發報結果；若以人口

數作為劃分基準，則因人口分布較為不均，較不適用於地震預警系統。綜合考量後，建議同時考量區域面積和測站數量進行網格劃分，以提升發報準確性。

此外，目前判斷 P 波到時的模型以及預測 PGA 的模型是互相獨立運作的。本研究開發一個能同時處理這兩個問題的模型，並設定在即時系統上運行。在前期開發的模型中，PGA 大於 25 的預測性能仍需改善。本期研究將輸入從單一時域更改為時域及時頻域的雙輸入，而在增加時頻域特徵後，對於預測 PGA 是否超過 25gal 之任務有顯著改善漏報之情況，但對地震相位標記之任務則無顯著改善。

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、網格化

Abstract

Earthquakes are unpredictable and devastating natural disasters that pose a significant threat to human lives and property. Therefore, accurately and promptly predicting earthquake intensity has become one of the primary objectives in the development of strong-motion early warning technology.

Although regional strong-motion early warning systems utilizing the multi-station method can quickly and accurately predict the intensity at various locations, they rely on receiving P-wave signals from multiple stations to make predictions. This often results in delayed warnings for areas near the epicenter, as the warning is issued only after the strong motion arrives, creating a "blind zone" with a radius of approximately 50 to 60 kilometers.

Previous studies have preliminarily developed a CNN single-station method for predicting PGA (Peak Ground Acceleration) and integrated the CNN model into the Earthworm system for real-time warnings on equipment provided by the Central Weather Bureau. However, analysis of the April 3 Hualien earthquake revealed that both the time taken to receive phase information and the model's prediction time were excessively long. This study attempted to address the issue of prolonged prediction time by utilizing the Triton Inference Server, reducing the model's prediction time by approximately 2.3 seconds compared to the online model during the April 3 event when run in the laboratory. To address the issue of prolonged phase information reception, the MQTT protocol was used. However, during laboratory operations, the phase information reception time was approximately 0.7 seconds longer compared to the online model's prediction time in the April 3 event.

Additionally, to statistically evaluate the current CNN model's online performance, this study conducted single-station analyses on nine earthquake events between January and September 2024. Accuracy and recall rates for each event were computed, and the causes of missed and false alerts were analyzed. Missed alerts were often due to the characteristics of the earthquake events, where initial waveforms were small, leading to low predictions within 2 seconds of P-wave arrival. In some cases, smaller pre-mainshock events triggered the system, resulting in low predictions for the mainshock. False alerts were primarily caused by delays in updating the recorded P-wave onset time in some events, leading to overestimated predictions. These issues have been resolved in the

program, awaiting validation in future earthquake events.

For grid-based seismic warning coverage, this study explored the feasibility of dividing areas based on region size, population, and the number of stations. Results indicated that using area size and station density as criteria for division significantly improved prediction accuracy. While a few missed alerts occurred, the PGA values at these stations remained within acceptable error margins. However, excessive grid refinement could result in irregular shapes, impacting warning outcomes. Using population as a criterion was less suitable due to uneven population distribution. Therefore, it is recommended to consider both area size and station density for grid division to enhance warning accuracy.

Furthermore, the current models for detecting P-wave arrival and predicting PGA operate independently. This study developed a unified model capable of handling both tasks simultaneously, designed for real-time system operation. In the earlier model development, the predictive performance for PGA values exceeding 25 gal required improvement. In this phase of the study, the input was changed from a single time-domain input to dual inputs comprising time-domain and time-frequency domain features. Incorporating time-frequency domain features significantly reduced missed alerts for predicting PGA exceeding 25 gal, though it did not lead to notable improvements in earthquake phase identification tasks.

Keywords: Convolutional Neural Network 、 Multi-Task 、 U-Net 、 Single Station 、

Phase picking 、 earthquake early warning 、 gridding

1 前言

強震預警系統研發技術的核心目標之一是能更快速且更精確地預測地震震度。儘管多站法（區域型）強震預警技術能快速且精確地預測各地的震度，但需要多個測站都收到 phase 資訊才能進行預測。因此，對於近震央地區，常常會在強震波到達後才發出預警，使得該地成為預警盲區，因此近震央地區特別適合採用現地法預測。先前的研究已開發出基於 CNN 的單站法預測 PGA 技術，該技術採用了 P 波後兩秒的加速度波形，經過多尺度和多域處理，作為 CNN 模型的輸入；而本研究目標之一為對於在氣象署線上運行於 Earthworm 系統之 CNN 模型進行 2024 年 1 月至 9 月震度大於 5 級以上之地震事件的分析，並針對預測時效問題進行改善，此外，亦探討針對網格化區域發布警報的可行性，期望能使 CNN 模型對於發報表現可以更完善。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)線上模型在 0403 事件發生之時效性問題；(3)統計指標-二元混淆矩陣；(4)統計線上模型在 2024 年 1 月至 9 月震度大於 5 級以上之單站分析結果；(5)在研究室環境、氣象署環境之 MQTT 架構；(6) 針對網格化區域發布警報。

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，由於前期的研究結果顯示該模型在預測 PGA 大於 25gal 之任務表現不盡理想，因此本期主要將先前的以時域特徵作為單輸入的模型，改為加上時頻域特徵作為雙輸入之模型，以期模型能學習時頻域特徵，改善預測性能。此部分主要在第三章進行介紹，其內容包含(1)研究動機；(2)模型架構；(3)資料集；(4)模型輸入及輸出；(5)評估模型表現；(6)短時傅立葉變換參數分析和不同時頻輸入架構特定區間對數正規化；(7) DAPPER (Dual-domain Attentive Phase Picker and Earthquake Reporter)各任務門檻值設定和與特定任務模型比較；(8) 20240403 花蓮事件之離線分析。

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[1]所提出之時頻域多尺度之 CNN 預測 PGA 模型。時頻域是指時間域與頻率域，多尺度則是指不同的標準化尺度。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、氣象署模型運行情況、地震離線回放成果進行說明。

模型建置所使用的資料來源為臺灣中央氣象署強震紀錄資料庫(TSMIP) 1992 年至 2006 年的地震紀錄，其中包含 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 所示。

表 2-1 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。我們從 10,000 筆資料中，選擇 8,000 筆進行訓練與驗證，分別取 PGA 小於 25 gal 的資料和 PGA 大於等於 25 gal 的資料各八成；另外選取 2,000 筆作為測試資料，分別取 PGA 小於 25 gal 的資料和 PGA 大於等於 25 gal 的資料各兩成，以完成訓練與測試資料的分配。訓練和測試資料加速度區間資料量分布如下表 2-2 所示。

表 2-2 訓練和測試資料加速度區間資料分布

加速度(cm/ s ²)	0.8~25	25 以上	總和
訓練資料	3700	4300	8000
測試資料	925	1075	2000

2.1 模型架構

根據黃昭文的研究[1]，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，在時間域尺度有 $\pm 2.5\text{gal}$ 、 $\pm 25\text{gal}$ 、 $\pm 250\text{gal}$ 三種，頻率域尺度有 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 2 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，模型詳細參數如表 2-3 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)。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。

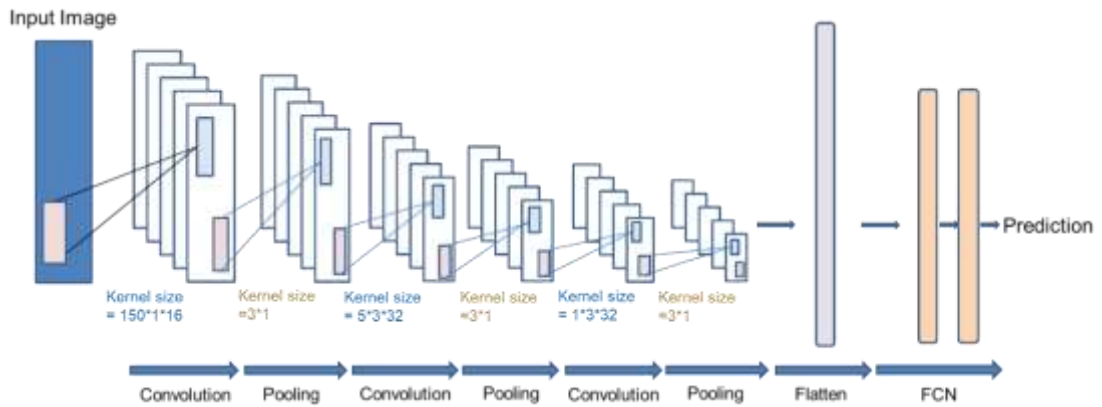


圖 2-1 卷積神經網路結構

表 2-3 模型參數表

Optimizer	Adam
Activation Function	ReLU
Loss Function	RMSLE
Batch Size	64

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

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，分布如圖 2-2 所示。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。假設所有測站都能傳輸即時資料，且能夠準確即時的標記 P 波到時。在進行地震模擬時，當測站偵測到 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。

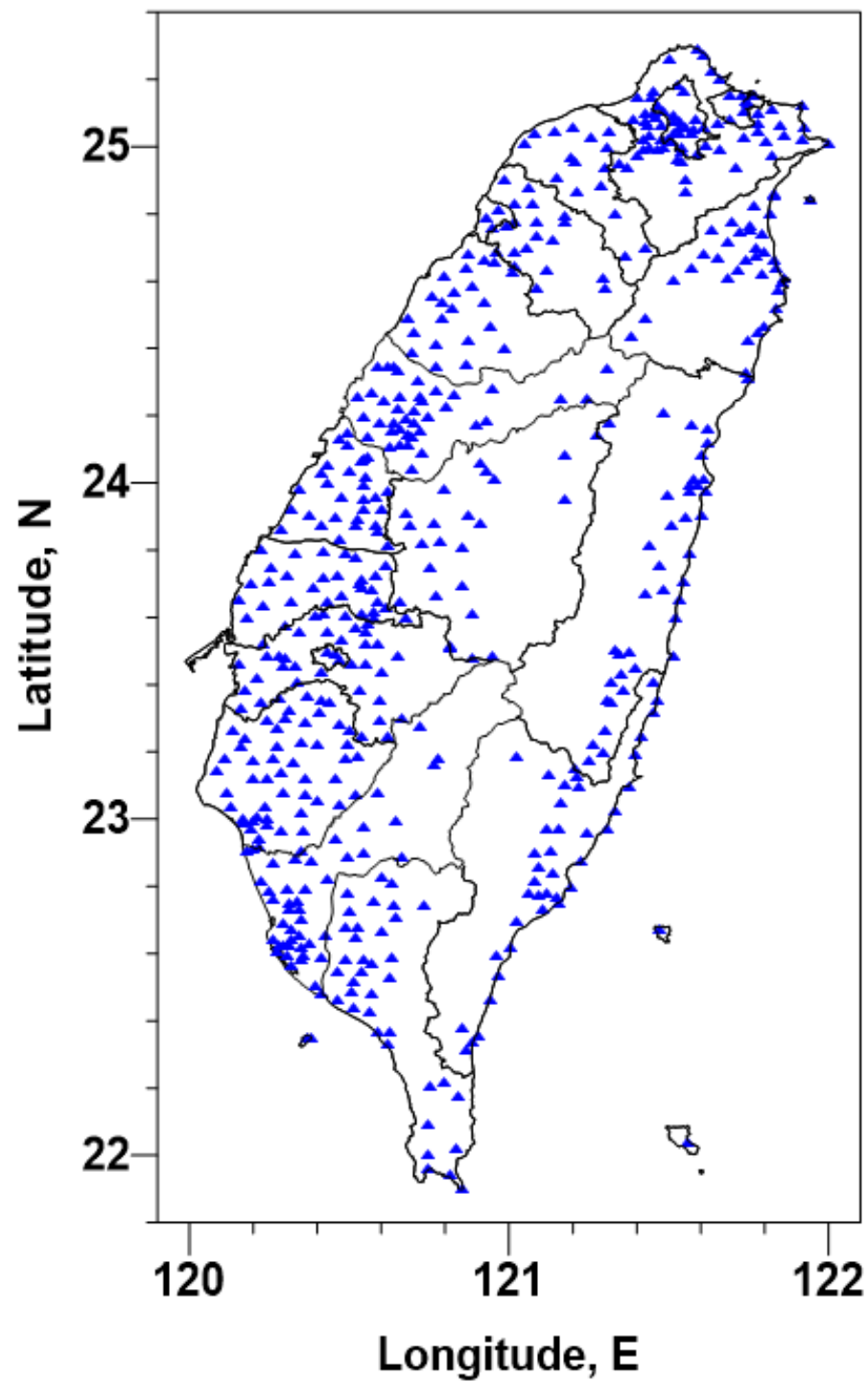


圖 2-2 TSMIP 測站分布

2.2 統計指標

2.2.1 統計指標 – 二元混淆矩陣(Confusion Matrix)

對於是否需要發布地震警報以及是否成功發布預警，以單站法作為統計。單站

法統計會以單一測站為單位對地震進行發布統計，若各測站計算出的真實 PGA 大於門檻值，該測站就必須發報。

本研究使用二元混淆矩陣(Confusion Matrix) 對於警報發布的正確性進行判斷。本研究以 25gal 作為門檻值進判斷，如圖 2-3 所示，並以正負一級作為警報之容許誤差範圍，如圖 2-4 所示。容許誤差的意思為若有發布警報，但實際 PGA 在 8~25gal 之間，雖然未達門檻值，但在此區的人民可能會感到明顯晃動，因此將此事件歸為正確發報；若在沒有發報，但實際 PGA 在 25~80gal，雖達門檻值，但還不足以對人民造成危害，因此將此事件歸為正確不發報。

二元混淆矩陣判斷指標為：TP 為正確發報，指真實 PGA 和預測 PGA 皆達發報標準；TN 為正確不發報，指真實 PGA 和預測 PGA 皆未達發報標準；FP 為誤報，指真實 PGA 未達發報標準，但預測 PGA 達發報標準；FN 為漏報，指真實 PGA 達發報標準，但預測 PGA 未達發報標準。得到 TP、TN、FP、FN 四個指標後，在計算出精確率(Precision)和召回率(Recall)，公式如式(2-2)和式(2-3)所示。

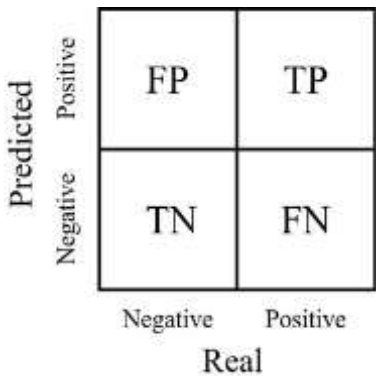


圖 2-3 二元混淆矩陣(無容許誤差)

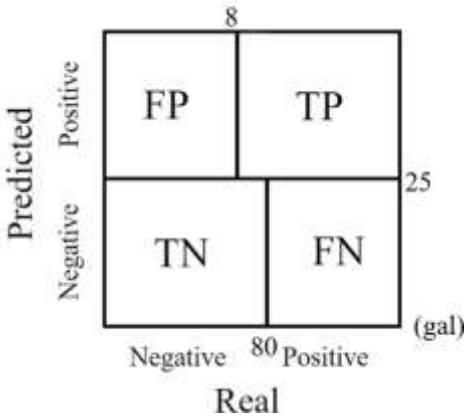


圖 2-4 二元混淆矩陣(含容許誤差)

$$Precision = \frac{TP}{TP + FP} \quad (2-2)$$

$$Recall = \frac{TP}{TP + FN} \quad (2-3)$$

2.3 2024 年 1 月至 9 月震度大於 5 級線上模型運行情況

2.3.1 20240403 花蓮地震事件分析

本報告先使用 2024 年 4 月 3 日 07:58:09 規模 7.2 之花蓮縣地震事件做檢討分析。此次地震事件震央位於花蓮縣政府南南東方 25 公里(臺灣東部海域)。各地地震震度分布如圖 2-5 0403 事件各地震度所示，全臺縣市最大震度皆大於四級。

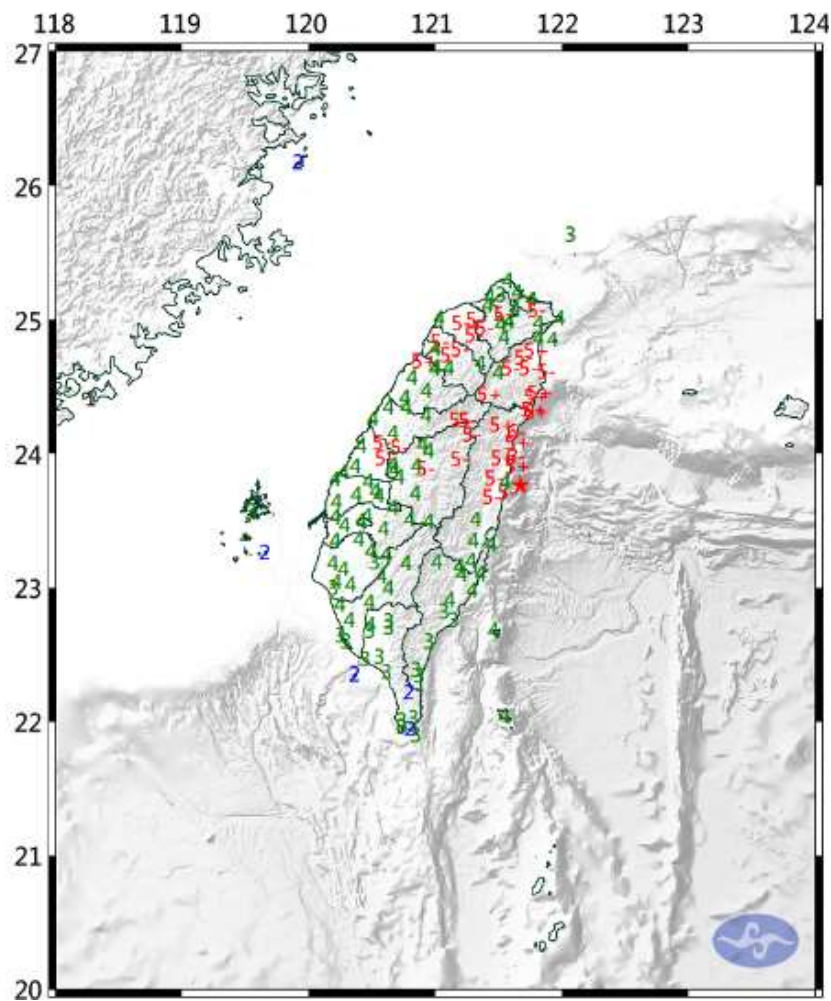


圖 2-5 0403 事件各地震度

此次事件成果分析採用即時預測的留存紀錄檔進行。PGA Ground Truth 使用氣象署下載之測站資料(SAC 檔)回放進行計算。預測結果的比較，採線上與 SAC 檔共有的 390 個測站進行分析，圖 2-6 為共有之測站分布圖。由兩者皆有之測站，做單站的檢討；標記出 Twin 測站，作為發報之依據。

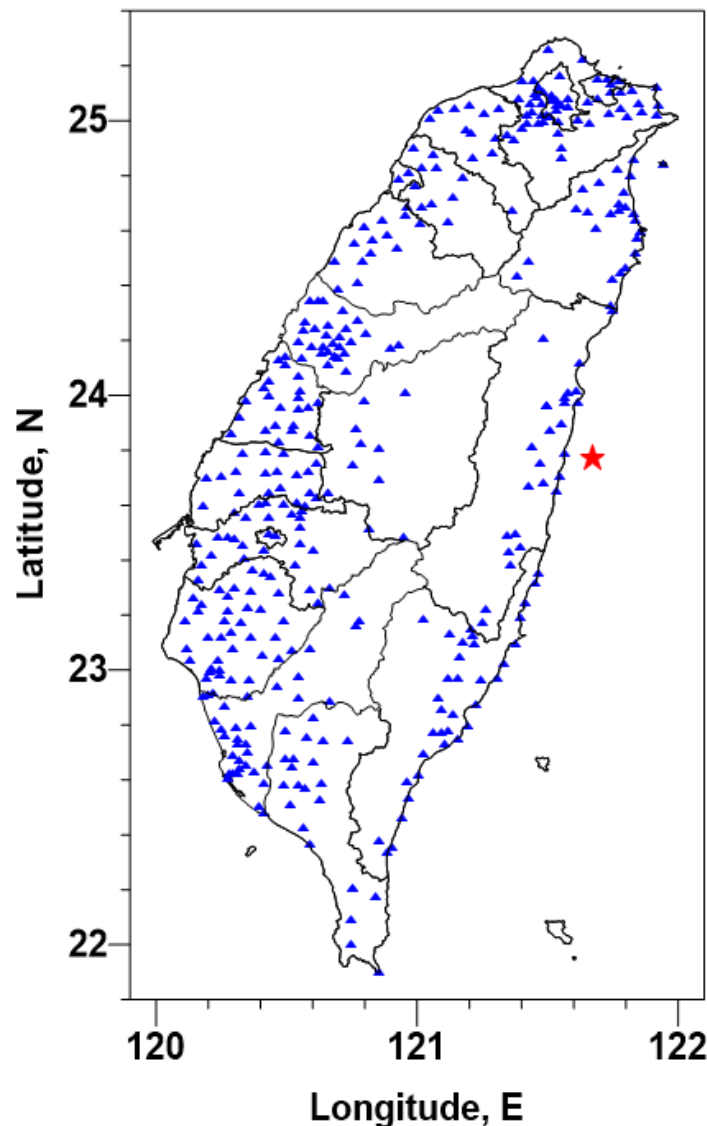


圖 2-6 線上與 SAC 檔共有之測站分布圖

單站法統計結果如下，該模型線上即時運算時採用氣象署觸發程式所提供之資訊，結果發現觸發的測站數目非常少。以下分析僅考慮線上和 SAC 檔共有之測站。觸發的 28 個測站中，實際 PGA 大於 25gal 之測站數量為 26 個；實際 PGA 小於 25gal 之測站數量為 2 個。而模型預測結果，PGA 大於 25gal 之測站總數為 13 個測站；PGA 小於 25gal 之測站總數為 15 個測站，分布如圖 2-7 所示，圖中紫色三角形測站表示模型有觸發，但是 SAC 檔中並沒有此測站。單站之二元混淆分佈矩陣分布，在無正負 1 級的容許誤差情況下，計算出準確率為 100%，召回率為 50%，分布結果如圖 2-8 所示；若是在有正負 1 級的容許誤差情況下，準確率為 100%，召回率為 61.9%，分布結果如圖 2-9 所示。這次地震分析的過程中還有發現觸發測站數量不足的問題，僅有 28 個測站觸發，剩下的 361 個測站皆未觸發，如圖 2-10 所示，這會嚴重影響模型的表現效果。

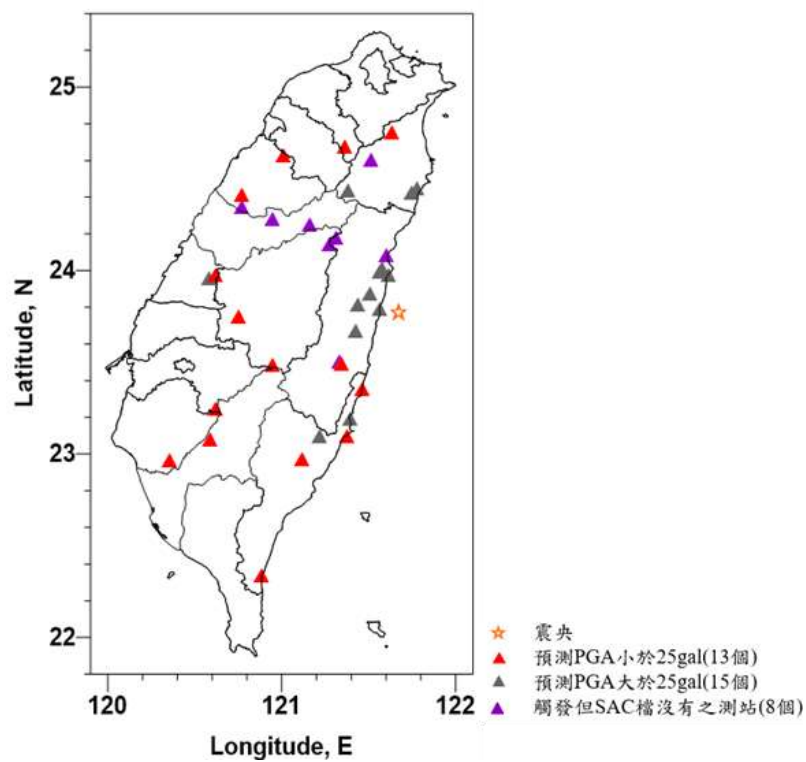


圖 2-7 觸發測站分布圖

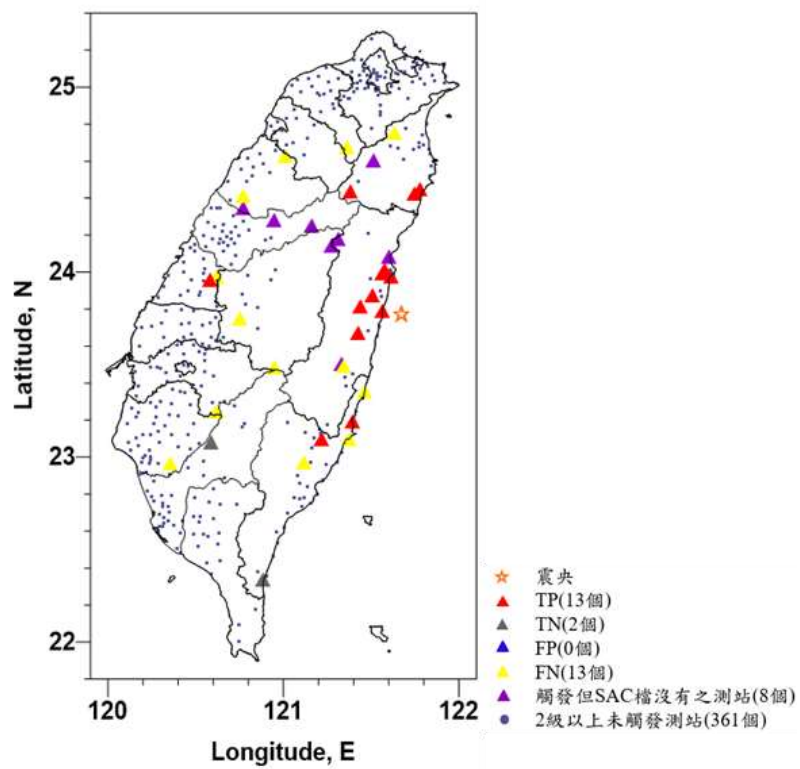


圖 2-8 二元混淆矩陣(無容許誤差)

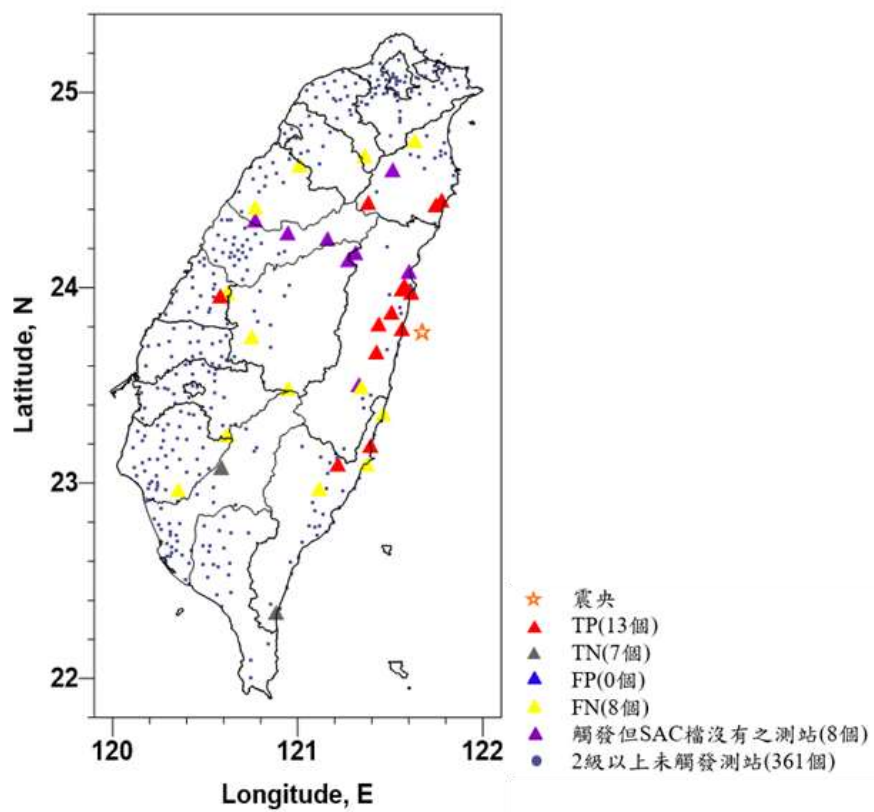


圖 2-9 二元混淆矩陣(含容許誤差)

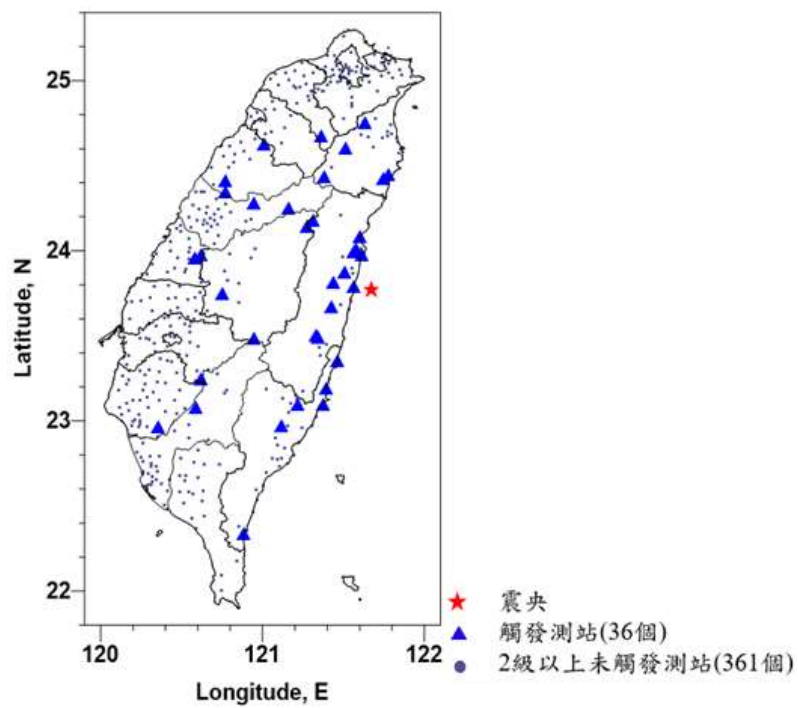


圖 2-10 未觸發測站和觸發測站分布圖

為了評估程式的運算時間，我們統計了模型接收相位資訊所需的時間和模型預測所需的時間(以有觸發的 36 個測站為例)，繪製了長條圖如圖 2-11 和圖 2-12 所示，模型平均接收相位資訊所需的時間為 5.14 秒，而模型平均預測所需的時間為 2.33 秒，不論是地震後到模型接收相位資訊的時間或是模型預測的時間皆耗時太長。我們也統計了包括模型取得相位資訊與預測完成所需的整體時間，如圖 2-13 所示，平均需要 8.7 秒才能完成。關於對於運算時間的改善狀況，請見 2.4 節。

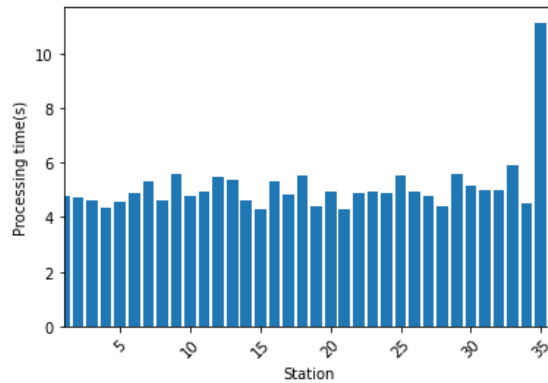


圖 2-11 模型接收相位資訊時間

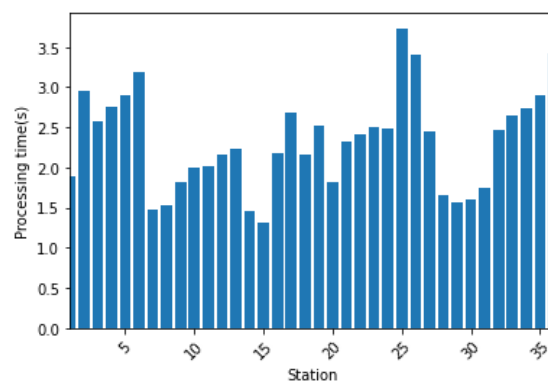


圖 2-12 模型完成預測時間

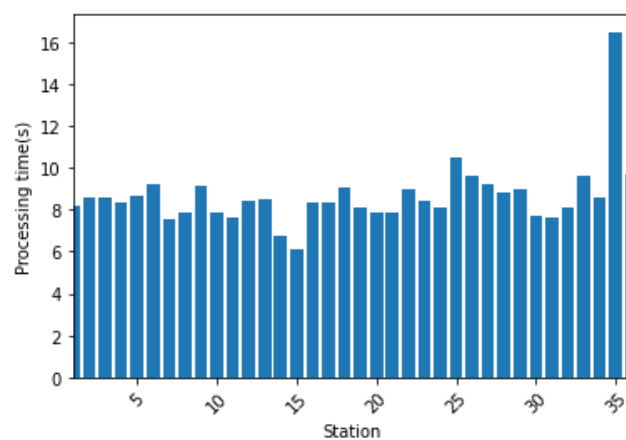


圖 2-13 模型取得相位資訊到預測完成總時間

2.3.2 2024 年 1-9 月最大震度 5 級以上事件分析

為統計 CNN 模型在單站線上預測中的效能，本研究除了 0403 地震事件外，還針對 2024 年 1 月至 9 月間最大震度達 5 級以上有編號地震事件進行分析，並評估各單站的預測結果。此次分析共 9 筆地震事件(三筆事件震央相近)，其震央分布如圖 2-14 所示。

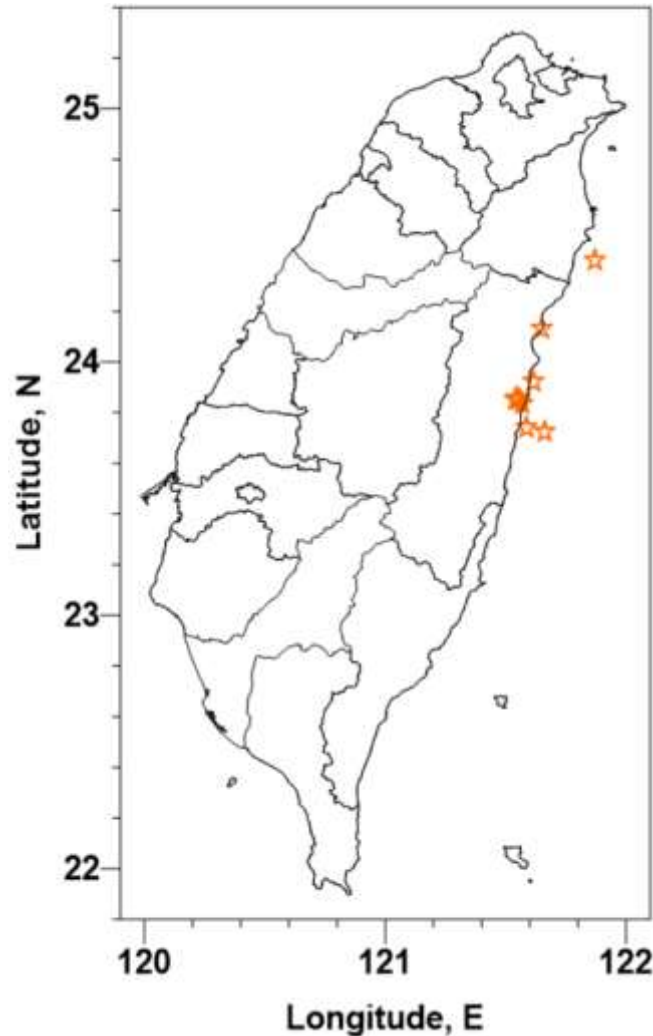


圖 2-14 1-9 月最大震度 5 級以上且有編號之事件震央分布

以下為 1-9 月最大震度 5 級以上 9 筆地震事件之單站預警二元混淆矩陣分布結果：

- 2024 年 4 月 3 日 (編號 021) 08:11:26 發生於花蓮近海的地震事件，其單站二元混淆矩陣分布結果在無正負 1 級容許誤差的情況下如圖 2-15(a)所示。其中，有 1 個測站發生漏報 (標示為黃色三角形)，計算得到準確率為 100%，召回率為 80%。若在考慮正負 1 級容許誤差的情況下，分布結果如圖 2-15(b)所示，仍有 1 個測站發生漏報 (黃色三角形)，準確率為 100%，召回率依然為 80%。

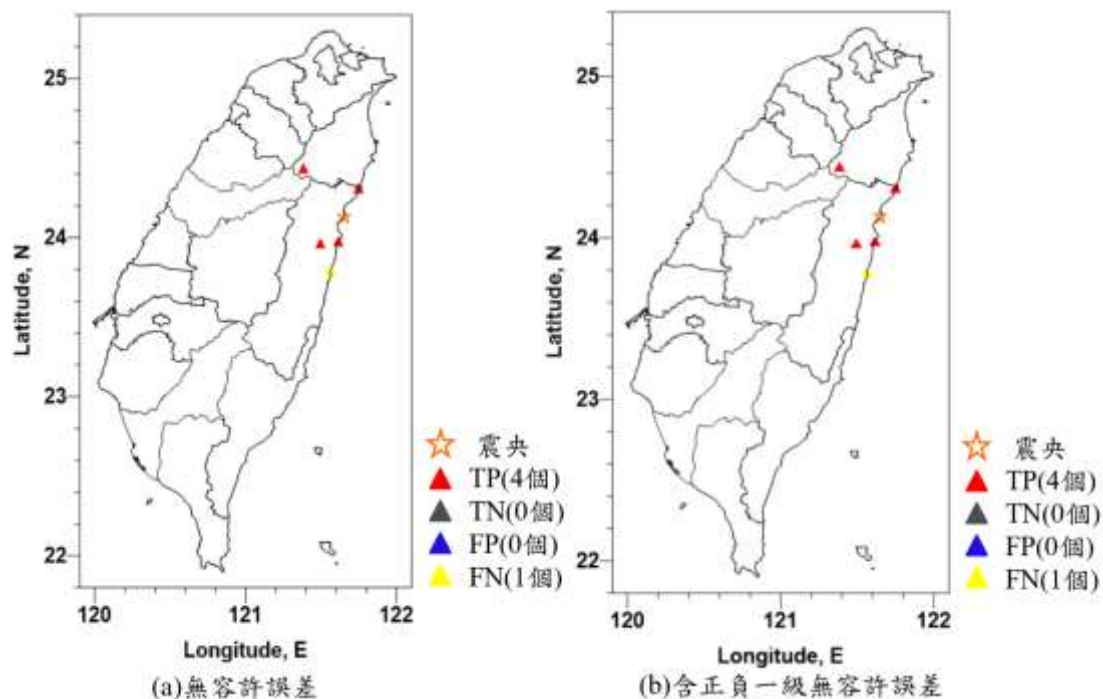


圖 2-15 0403(編號 021)地震之二元混淆矩陣

- 2024 年 4 月 23 日 (編號 266) 02:26:51 發生於臺灣東部海域的地震事件，其單站二元混淆矩陣分布結果在無正負 1 級容許誤差的情況下如圖 2-16(a)所示。其中，有 4 個測站發生漏報 (標示為黃色三角形)，計算得到準確率為 100%，召回率為 64%。若考慮正負 1 級的容許誤差，分布結果如圖 2-16(b)所示，則有 3 個測站發生漏報 (黃色三角形)，準確率仍為 100%，召回率提高至 70%。

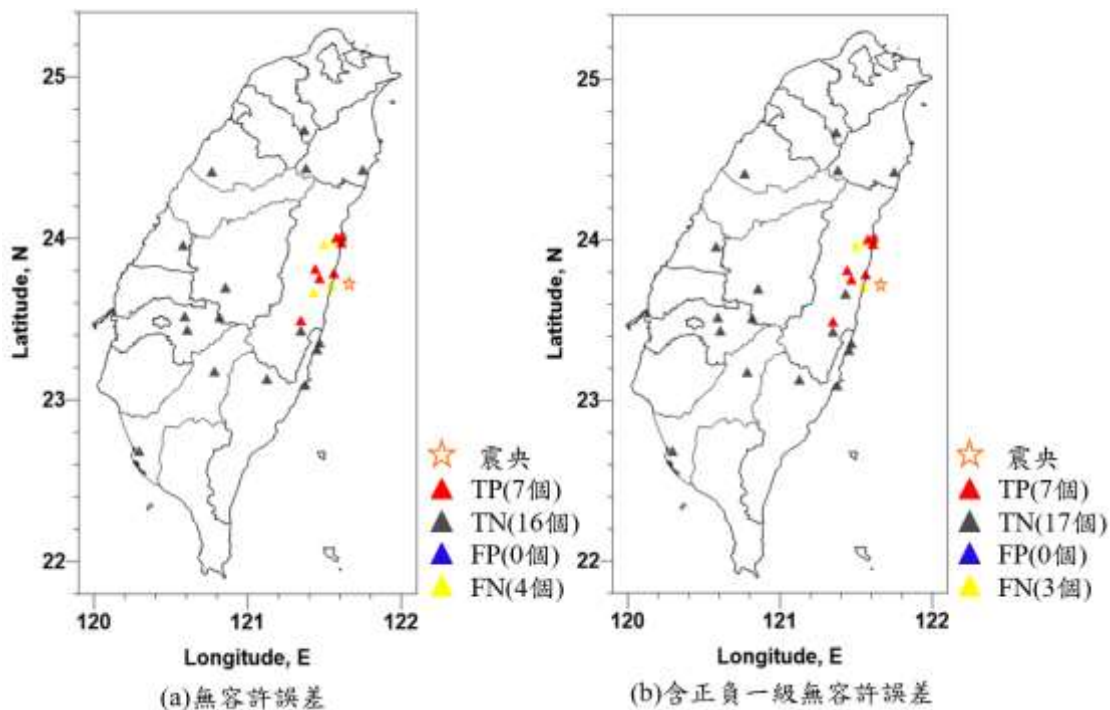


圖 2-16 0423(編號 266)地震之二元混淆矩陣

- 2024 年 4 月 23 日（編號 267）02:32:49 發生於花蓮縣壽豐鄉的地震事件，其單站二元混淆矩陣分布結果在無正負 1 級容許誤差的情況下如圖 2-17(a)所示。其中，有 1 個測站發生漏報（標示為黃色三角形），並有 5 個測站誤報（標示為藍色三角形），計算得到準確率為 55%，召回率為 86%。若考慮正負 1 級的容許誤差，分布結果如圖 2-17(b)所示，則有 1 個測站漏報（黃色三角形）和 1 個測站誤報（藍色三角形），準確率和召回率皆提高至 91%。

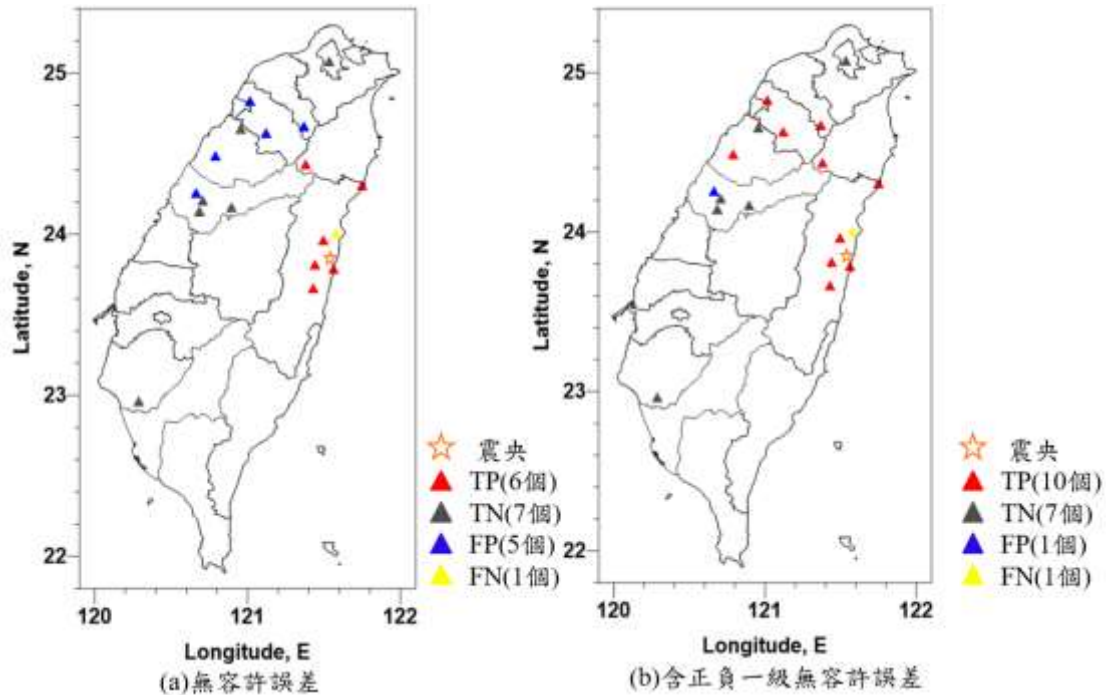


圖 2-17 0423(編號 267)地震之二元混淆矩陣

- 2024 年 4 月 23 日（編號 300）05:31:39 發生於花蓮縣壽豐鄉的地震事件，其單站二元混淆矩陣分布結果在無正負 1 級容許誤差的情況下如圖 2-18(a)所示。其中，有 5 個測站發生漏報（標示為黃色三角形），計算得到準確率為 100%，召回率為 44%。若考慮正負 1 級的容許誤差，分布結果如圖 2-18(b)所示，則有 1 個測站漏報（黃色三角形），準確率仍為 100%，召回率提升至 80%。

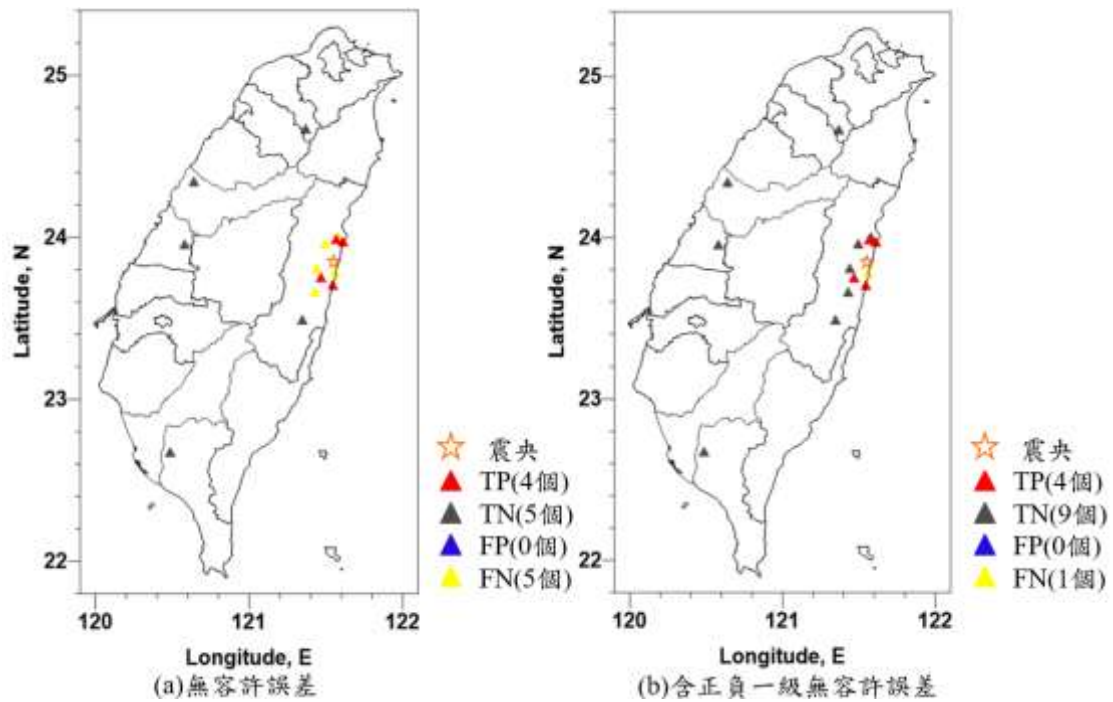


圖 2-18 0423(編號 300)地震之二元混淆矩陣

- 2024 年 4 月 23 日 (編號 307) 08:04:05 發生於花蓮縣壽豐鄉的地震事件，其單站二元混淆矩陣分布結果在無正負 1 級容許誤差的情況下如圖 2-19(a)所示。其中，有 7 個測站發生漏報 (標示為黃色三角形)，1 個測站發生誤報 (標示為藍色三角形)，計算得到準確率為 67%，召回率為 22%。若考慮正負 1 級的容許誤差，分布結果如圖 2-19(b)所示，則有 5 個測站發生漏報 (黃色三角形)，1 個測站發生誤報 (藍色三角形)，準確率仍為 67%，召回率提升至 29%

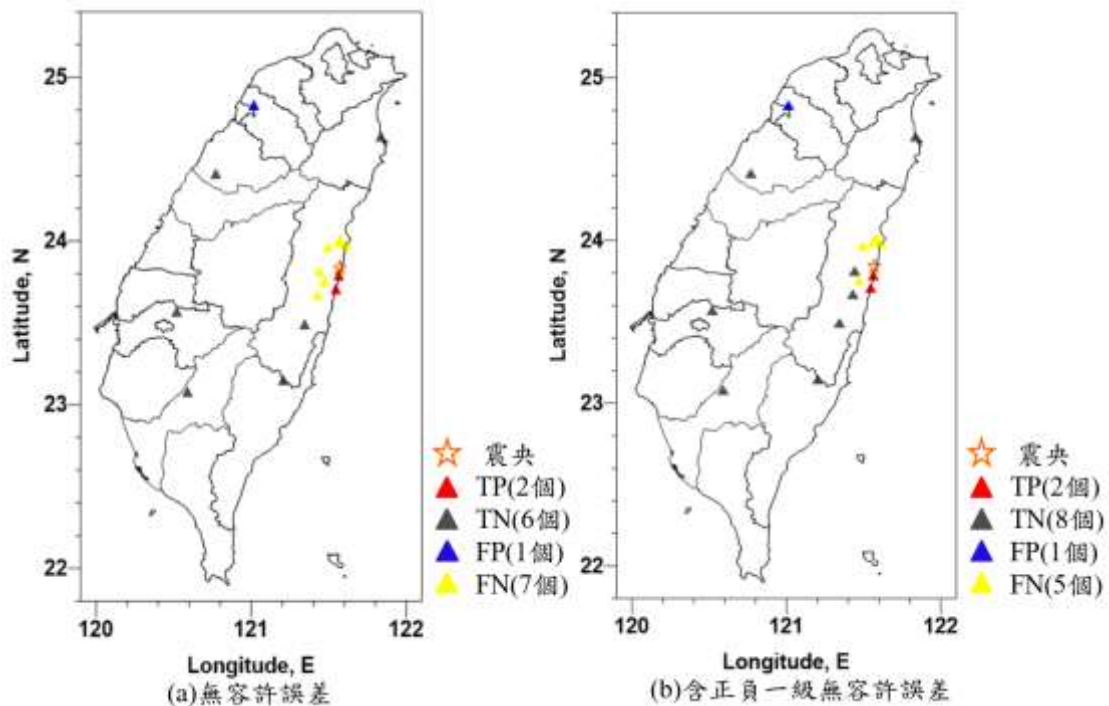


圖 2-19 0423(編號 307)地震之二元混淆矩陣

- 2024 年 4 月 23 日 (編號 318) 10:21:00 發生於花蓮縣壽豐鄉的地震事件，其單站二元混淆矩陣分布結果在無正負 1 級容許誤差的情況下如圖 2-20(a)所示。結果顯示，有 6 個測站發生漏報 (標示為黃色三角形)，1 個測站發生誤報 (標示為藍色三角形)，計算得出準確率 83%，召回率為 45%。若考慮正負 1 級的容許誤差，分布結果如圖 2-20(b)所示，則有 3 個測站發生漏報 (黃色三角形)，且無誤報情況，準確率為 100%，召回率提升至 67%。

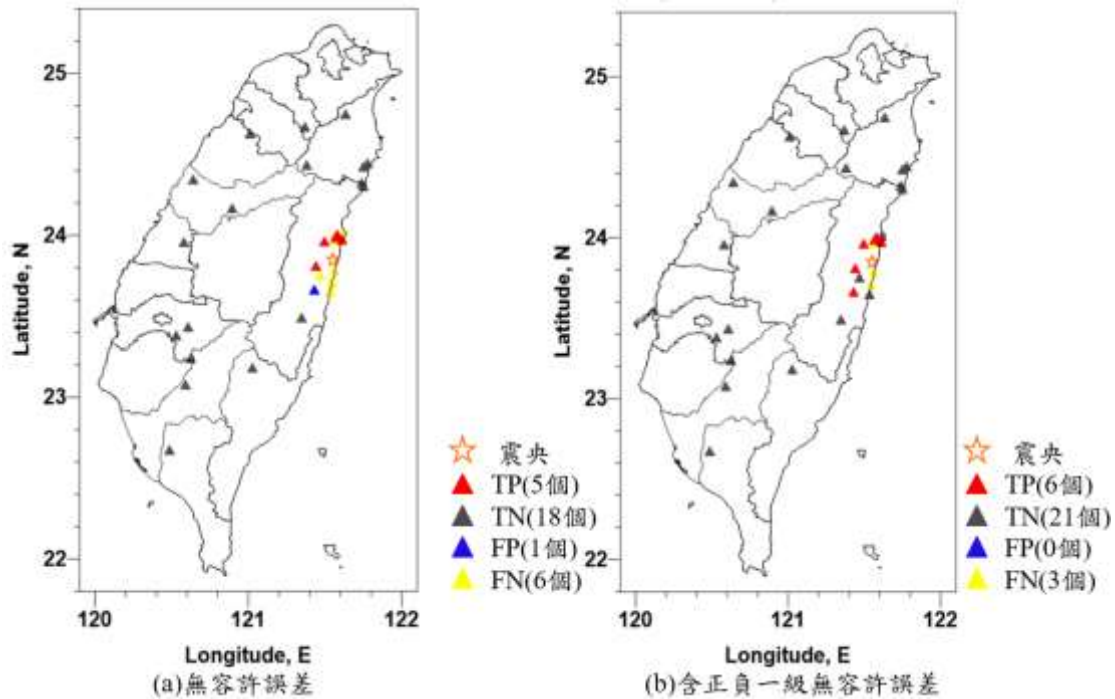


圖 2-20 0423(編號 318)地震之二元混淆矩陣

- 2024 年 4 月 25 日 (編號 341) 07:53:34 發生於花蓮縣近海的地震事件，其單站二元混淆矩陣分布結果在無正負 1 級容許誤差的情況下如圖 2-21(a)所示。結果顯示，有 1 個測站發生誤報(標示為藍色三角形)，計算得出準確率為 50%，召回率為 100%。若考慮正負 1 級的容許誤差，分布結果如圖 2-21(b)所示，所有發報皆正確，準確率和召回率皆達 100%。

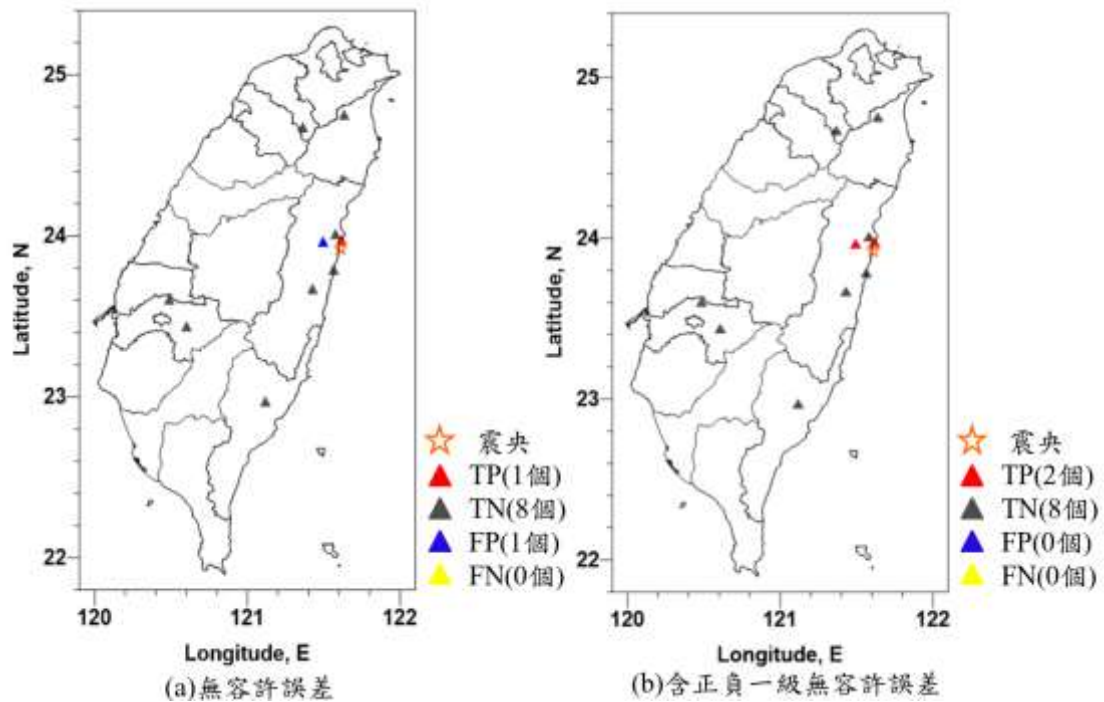


圖 2-21 0425(編號 341)地震之二元混淆矩陣

- 2024 年 5 月 8 日（編號 377）16:40:43 發生於花蓮縣近海的地震事件，其單站二元混淆矩陣分布結果在無正負 1 級容許誤差的情況下如圖 2-22(a)所示。結果顯示，有 1 個測站發生漏報（標示為黃色三角形），計算得出準確率為 100%，召回率為 75%。若考慮正負 1 級的容許誤差，分布結果如圖 2-22(b)所示，所有發報皆正確，準確率和召回率皆達 100%。

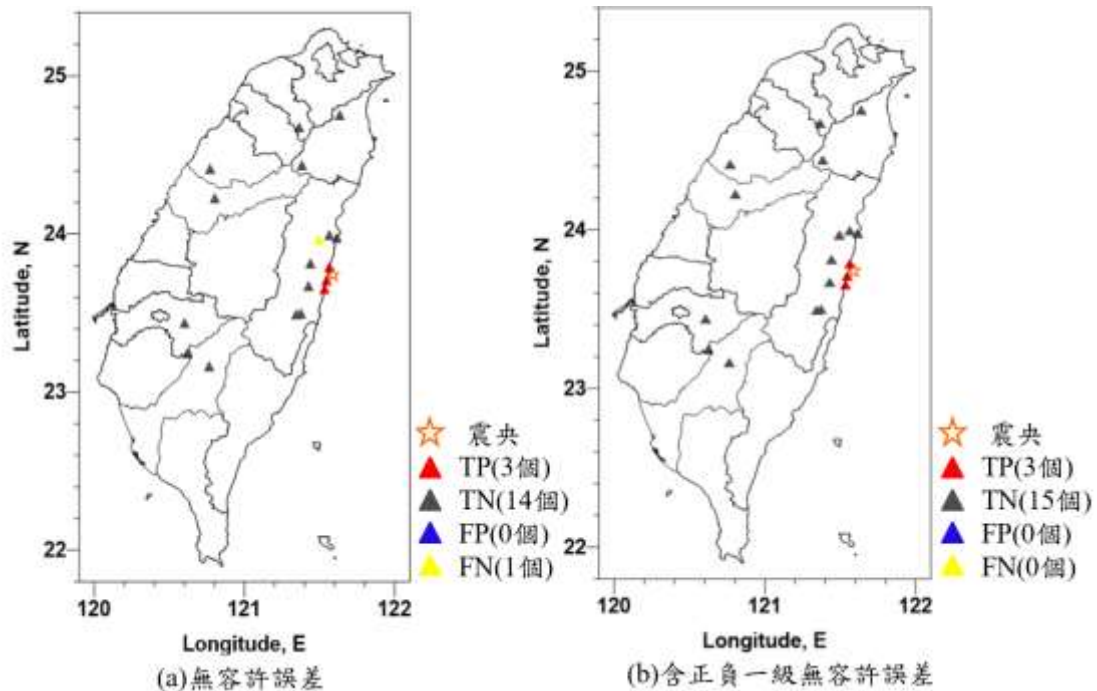


圖 2-22 0508(編號 377)地震之二元混淆矩陣

- 2024 年 8 月 15 日（編號 446）17:06:46 發生於宜蘭縣近海的地震事件，其單站二元混淆矩陣分布結果在無正負 1 級容許誤差的情況下如圖 2-23(a)所示。結果顯示，有 1 個測站發生漏報（標示為黃色三角形），且有 10 個測站為誤報（標示為藍色三角形），計算得出準確率為 55%，召回率為 92%。若考慮正負 1 級的容許誤差，分布結果如圖 2-23(b)所示，所有發報皆正確，準確率和召回率皆達 100%。

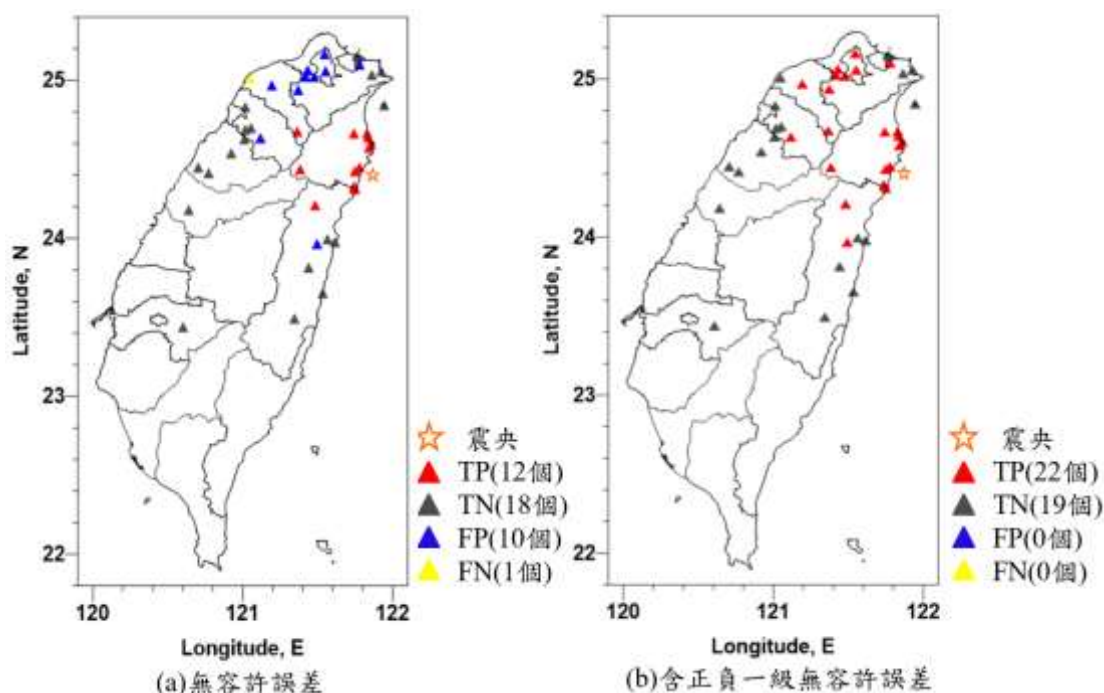


圖 2-23 0815(編號 442)地震之二元混淆矩陣

統計以上 9 筆地震事件之準確率和召回率，如表 2-4 所示，每筆地震事件皆有漏報或誤報的情況發生，我們分析了漏報和誤報發生的原因。漏報的原因大致上分為三類，第一個原因是地震事件的特性，部分事件在一開始的波形較小，而我們的模型是基於 P 波到時後 2 秒來進行預測，因此預測結果可能會較小，導致漏報的情況發生，如**錯誤！找不到參照來源**。所示，紅色區段為 P 波到時，而紫色區段為 P 波到時後 2 秒，本模型僅取此段區間進行預測；第二個原因是有些事件在主震前已有較小地震，這使得所使用的波段為較小地震的事件，因此預測結果較小，進而導致漏報，如圖 2-25 所示；而最後一個原因是某些測站並沒有被觸發到，如圖 2-26 所示，可以觀察到該測站對於主震是沒有被觸發的。所有漏報測站的波形圖可以參考附錄 A。而誤報的主要原因是某些事件中程式記錄的 P 波起始點位因更新延遲而導致延後，使得預測值偏大。具體來說，程式儲存的波形每 5 秒更新一次，且總長度固定為 6000 點，當程式確定 P 波點位後，會檢查該點是否位於 6000 點內，並確認 P 波到時後 2 秒內是否存在 inf，經過這些篩選後，才會保存當前的波形。然而，在這段篩選過程中，波形可能已經被更新一次，導致 P 波點位在更新後的波形會晚 500 點（即 5 秒鐘），這可能使得預測值過大，進而引發誤報。圖 2-27 為 log 檔所記錄之 P 波點位的標記，可以觀察到實際觸發的點位是正確的；而圖 2-28 為程式紀錄之 P 波點位的標記，可以看出晚了 500 點，導致預測的結果較大。所有誤報測站的波形圖可以參考附錄 B。此部分程式之缺點已改善完成，等待之後的地震事件加以驗證。

表 2-4 2024 年 1-9 月 5 級以上地震事件之 Precision 和 Recall

地震編號	Precision(%)	Recall(%)	Precision(%)	Recall(%)
------	--------------	-----------	--------------	-----------

			含容許誤差	含容許誤差
021	100	80	100	80
266	100	64	100	70
267	55	86	91	91
300	100	44	100	80
307	67	22	67	29
318	83	45	100	67
341	50	100	100	100
377	100	75	100	100
442	55	92	100	100
平均值	78.9	67.6	95.3	79.7

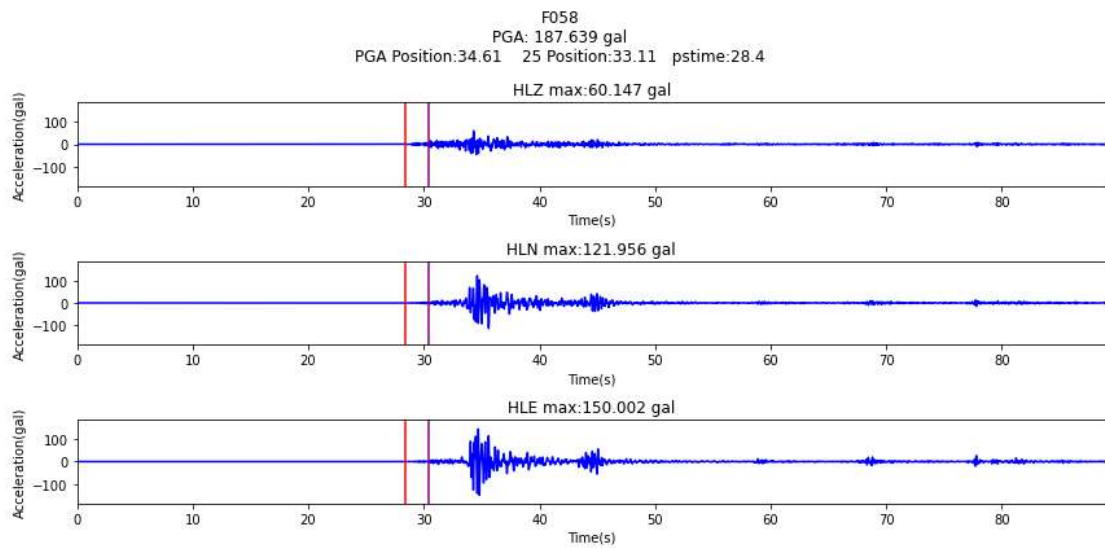


圖 2-24 F058 測站(編號 266 事件)波形圖

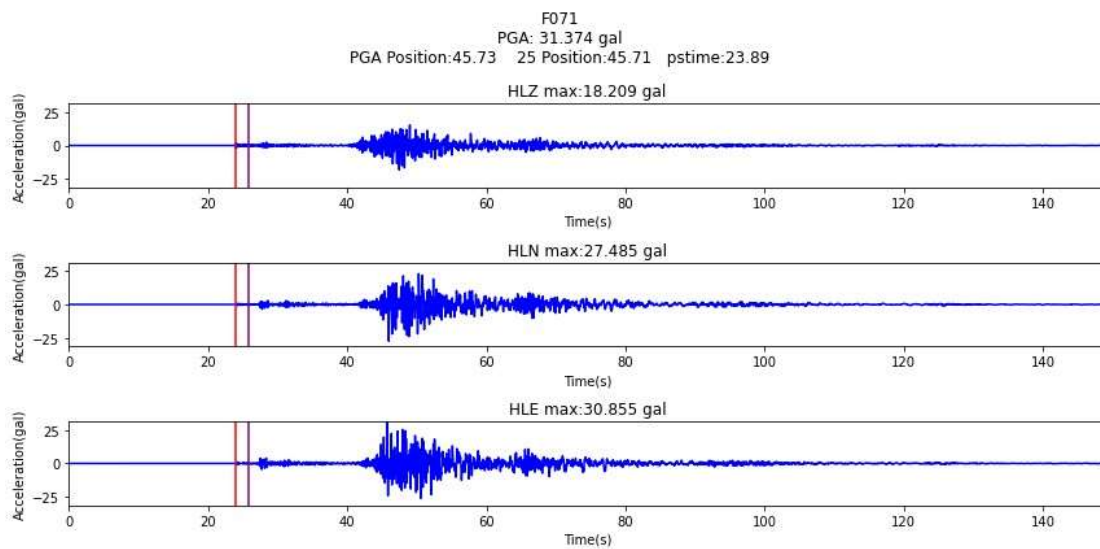


圖 2-25 F071 測站(編號 307 事件)波形圖

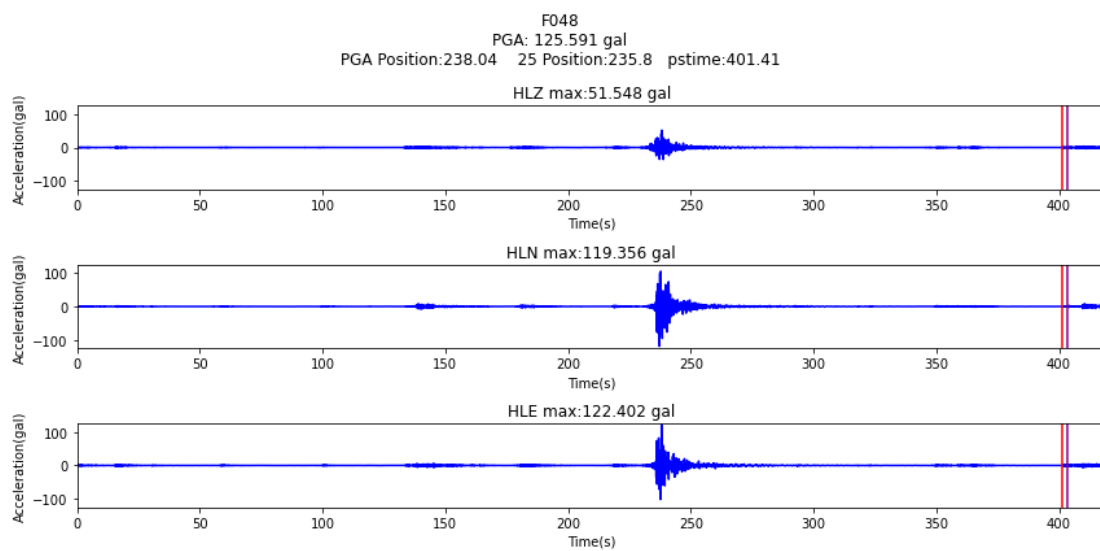


圖 2-26 F048 測站(編號 267 事件)波形圖

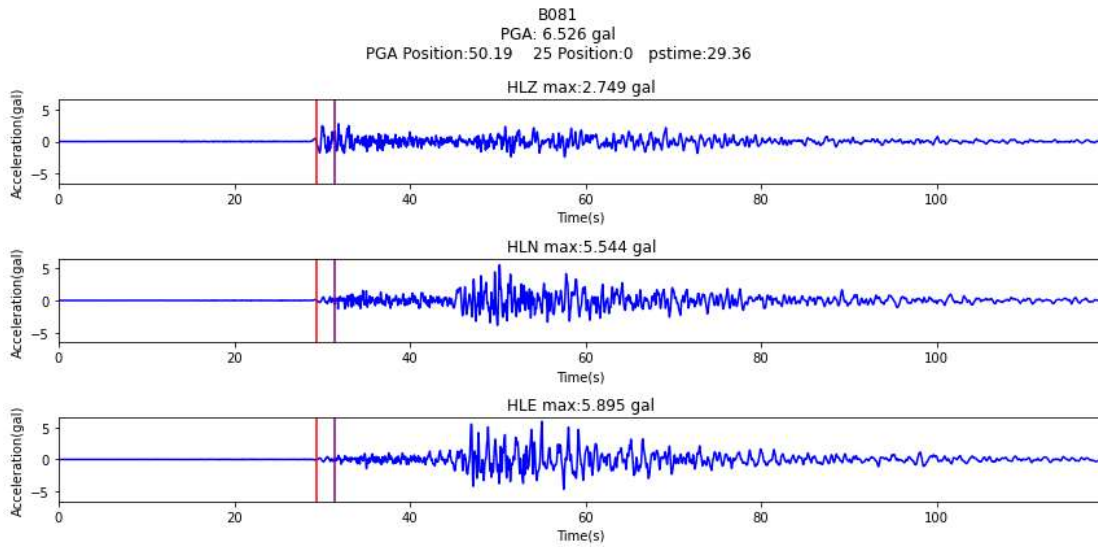


圖 2-27 B081 測站(編號 307 事件)波形圖(依 log 檔標記)

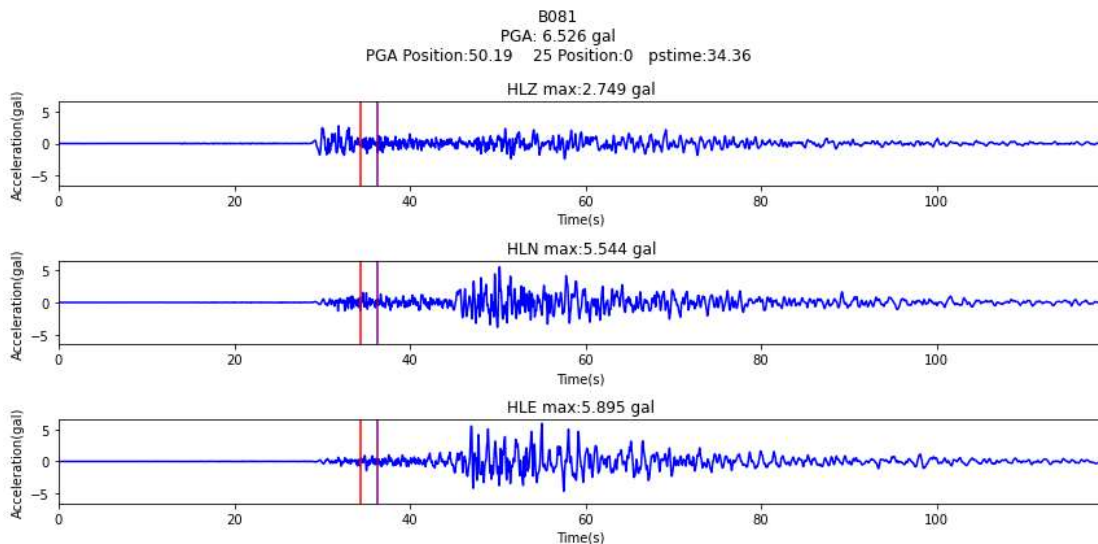


圖 2-28 B081 測站(編號 307 事件)波形圖(依程式標記)

2.4 研究室環境、氣象署環境之 MQTT 架構

2.4.1 MQTT 協議介紹

MQTT (Message Queuing Telemetry Transport) 是一種專為低頻寬和不穩定網路環境設計的輕量級消息傳遞協議，廣泛應用於物聯網 (IoT) 領域。以發布/訂閱 (publish/subscribe) 模型為基礎，MQTT 特別適合於需要高效數據交換的應用場景，如地震預警系統。

在地震預警應用中，MQTT 能有效協助波型資料的分發。系統可透過 MQTT 將波型資料發布至專用主題，讓所有相關模型訂閱該主題以接收資料。負責 P 波相位預測的模型會將預測的到達時間發布至指定的相位資訊主題。而負責地震預

警的模型則可訂閱波型和相位資訊主題，根據這些資訊進行進一步分析和預測，並將警報結果發布至預警主題，以供後續處理和應用。MQTT 的發布/訂閱機制確保了數據傳輸的靈活性，尤其適合本團隊需在不同主機上傳遞資料的特性。

2.4.2 研究室環境之 MQTT 架構

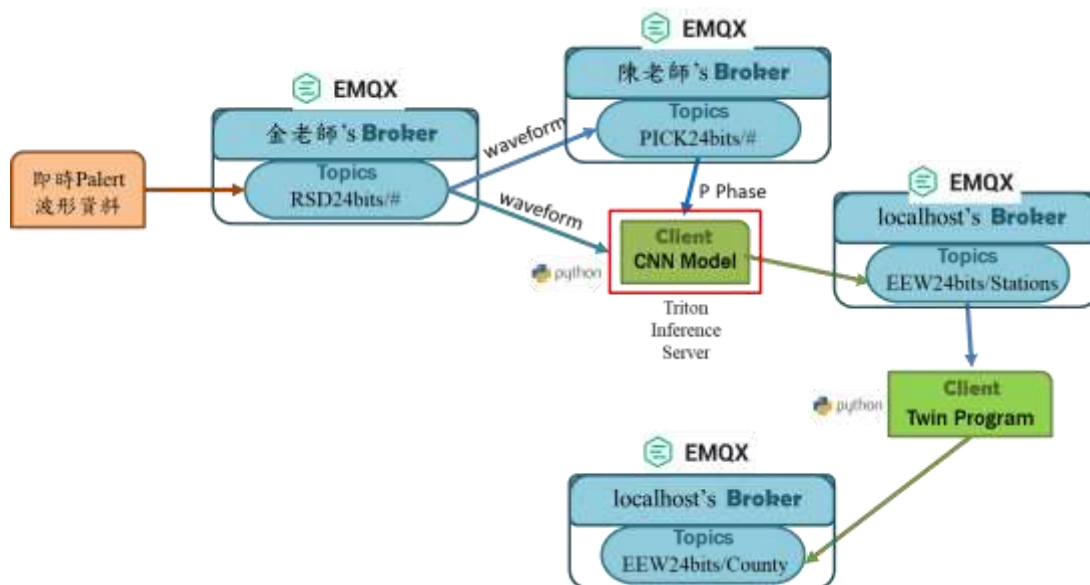


圖 2-29 在研究室之 MQTT 架構

圖 2-29 為建置在研究室之 MQTT 架構。整個架構的資料流為即時 Palert 波形資料發布至金老師的 topic-RSD24bits，本計畫之 CNN 模型和陳老師的 topic-PICK24bits 都訂閱金老師的 topic-RSD24bits 的波形資料。本計畫之 CNN 模型也訂閱陳老師的 topic-PICK24bits，獲得相位資訊。預測 PGA 大於 25gal 之測站會發布至 topic-EEW24bits/Station，另一方面，運用 Twin 機制對縣市發報之 python 程式會訂閱此 topic，以獲得預測 PGA 大於 25gal 之測站，並判斷需發報之縣市，並將結果會發布至 topic- EEW24bits/County。

因為本計畫之 CNN 模型預測時間過長，所以本計畫嘗試運用 Triton Inference Server 縮短 CNN 模型預測時間。

Triton Inference Server 是一個由 NVIDIA 開發的開源推理服務器，專為支持多種機器學習和深度學習模型的部署和推理而設計。Triton 旨在簡化 AI 模型的部署過程，提高推理效率和靈活性。它支持不同的框架和硬體加速選項，使得模型可以更有效地在生產環境中運行。

Triton Inference Server 的特點如下：

多模型支持：Triton 支持多種機器學習和深度學習框架，例如 TensorFlow、PyTorch、ONNX、TensorRT 等。這使得它可以同時部署不同框架下訓練的模型，

滿足多樣化的需求。

並行推理：Triton 支持多模型同時推理。這意味著它可以在同一服務器上運行多個模型，並提供高吞吐量的推理能力，適合大規模的生產環境。

批量推理：Triton 能夠自動將多個請求合併為批量請求，從而更高效地利用 GPU 資源進行並行推理，顯著提高了推理速度和效率。

動態批量處理：Triton 可以動態地調整批量大小，確保在推理延遲和吞吐量之間達到最佳平衡。這對於需要低延遲的應用特別有用。

多硬體支持：Triton 支持多種硬體，包括 NVIDIA GPU、CPU 和 DLA (Deep Learning Accelerator)。在 GPU 上運行時，Triton 能夠利用 CUDA 和 TensorRT 等技術進行加速。

其主要功能有：

模型自動管理：Triton 可以自動加載和卸載模型，支持在運行時更新和重新加載模型，無需重啟服務。

多種推理策略：支持同步和異步推理模式，可以根據應用場景選擇最適合的推理方式。

模型優化：通過 TensorRT 等技術進行模型優化，提供高效的推理性能。此外，它還支持精度優化（如半精度浮點數），以進一步提升推理速度。

儀表板和監控：提供 Prometheus 支持，用戶可以通過 Grafana 等工具監控 Triton 的推理性能指標，包括吞吐量、延遲、GPU 使用率等。

表 2-5 為運行在研究室 MQTT 架構和 Earthworm 架構之時間差。其中，接收相位之時間差為各觸發測站即時標記相位的時間至程式實際接收到相位資訊的時間；前處理所需之時間差為程式實際接收到各觸發測站相位資訊的時間，至前處理計算完畢並將結果輸入模型的時間；模型預測所需之時間為模型收到輸入資料的時間，至模型預測完成的時間。研究室 MQTT 架構使用 2024 年 9 月 20 日至 2024 年 10 月 31 日 log 檔所紀錄內容，並統計所有測站的所有觸發事件的平均值。Earthworm 架構使用氣象署線上運行在 20240403 花蓮地震事件記錄之 log 檔，並統計所有測站的所有觸發事件的平均值。從此表可發現使用 Triton Inference Server 可以讓模型預測所需之時間差平均值縮短約 2.33 秒，並縮短前處理所需之時間差平均值約 0.57 秒。因此，正將其架設在中央氣象署，期能縮短模型在氣象署運作環境下預測所需時間。然而，即使使用 MQTT 架構改變傳輸資料的方式，接收相位之時間差平均值並沒有縮短，甚至稍微延長。

表 2-5 研究室 MQTT 架構和 Earthworm 架構之時間差

	研究室 MQTT 架構 (運用 Triton Inference Server)	Earthworm 架構

接收相位之時間差平均值	5.82 秒	5.14 秒
前處理所需之時間差平均值	0.0062 秒	0.57 秒
模型預測所需之時間差平均值	0.0015 秒	2.33 秒

2.4.3 氣象署之環境架構

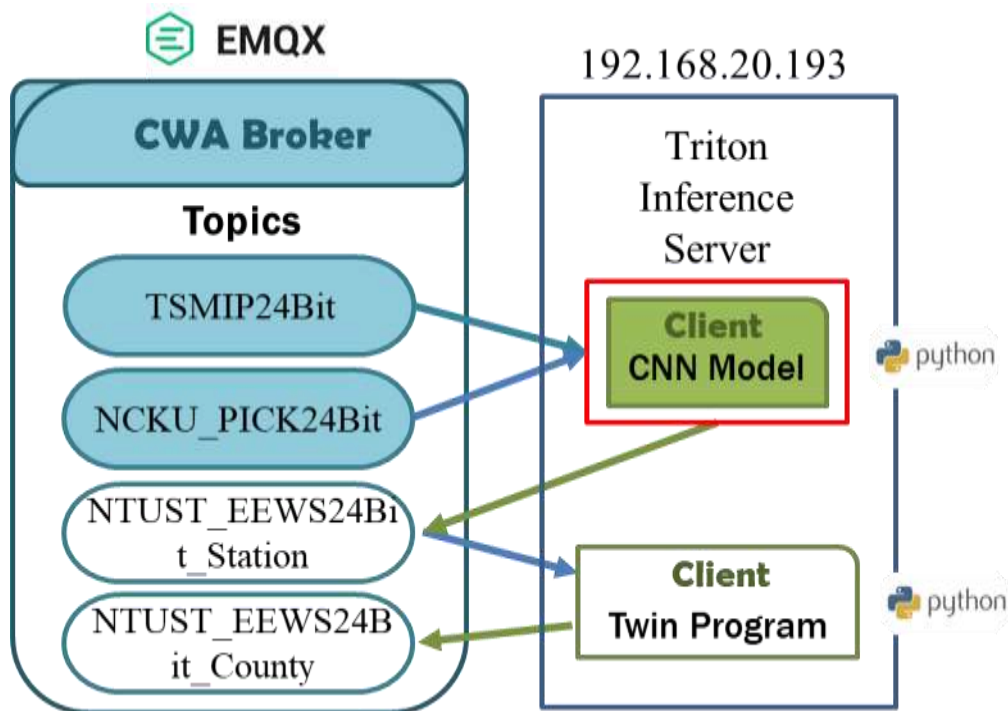


圖 2-30 氣象署線上之系統架構圖

圖 2-30 為 MQTT 架構應用在氣象署 TSMIP 即時測站線上系統之架構圖，於 2024 年 12 月開始運行。運行 CNN 模型之 python 程式會即時接收波形資料和 P 波相位資訊，分別從 topic-TSMIP24Bit 和 topic-NCKU_PICK24Bit。只要有接收到波形資料和 P 波相位資訊，模型就會對此觸發測站預測 PGA 是否大於 25gal，如果此觸發測站預測 PGA 大於 25gal，即發佈到 topic-NTUST_EEWS24Bit_Station。運行 Twin 機制之 python 程式透過訂閱 topic-NTUST_EEWS24Bit_Station，得到所有預測 PGA 大於 25gal 之測站，若這些測站有符合 Twin 機制，即會對此縣市發報。

2.5 縣市網格化

2.5.1 研究目的與目標

在分析地震事件時，發現對於地域範圍較大的縣市，若依現行地震預警系統以縣市為單位發佈警報，可能會產生不必要的警報。例如，當一地震震央位於某縣市的南端，且最大震度達四級並符合 Twin 機制，整個縣市將皆會收到預警通知，然而實際上北端居民可能完全沒有感受到此地震事件。因此，現行以縣市為單位的預警機制可能無法精確反映地震影響範圍，容易造成過度或不足的預警。為解決此問題，本研究嘗試以網格化的方式細化預警單位，以提升不同區域預警的準確度，並使其更能反映地震震度的真實分佈情形。

本研究以 2022 年 9 月 18 日 13:19:19 規模 6.1 之臺東縣地震事件為例，並針對區域較廣闊的花蓮縣進行網格的劃分，以觀察對縣市進行網格化後的發布情況。此次地震事件震央位於臺東縣池上鄉，各地震度分布圖如圖 2-31 所示，其中花蓮縣最大震度達 4 級以上，且符合 Twin 機制，若是依現行地震預警系統以縣市為單位發布警報，會對整個花蓮縣市進行發報，但實際上花蓮縣北部的居民可能並未感受到明顯的震動。因此，以下將根據面積大小、居住人口及測站數量劃分網格區域，分別分析不同劃分依據下的發報結果。

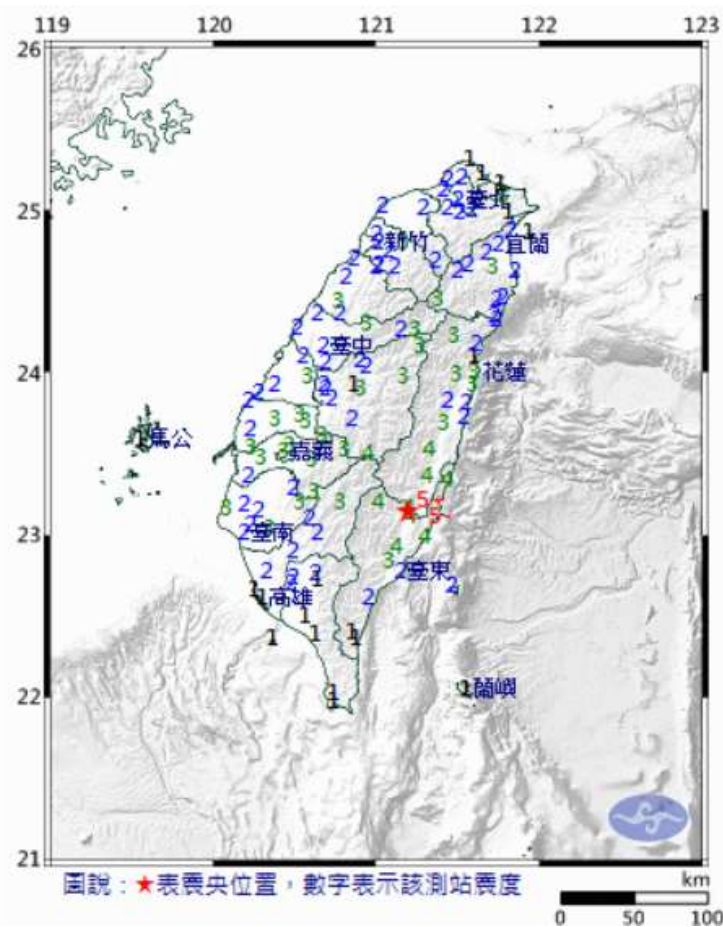


圖 2-31 0918 事件各地震度圖

2.5.2 以面積大小作為劃分依據

首先考量面積大小，圖 2-32 為花蓮縣各鄉鎮市區的面積分布（參考 112 年鄉鎮市區界線圖），將花蓮縣細分為不同大小的網格，依次設定為約 2000、1500、1000 及 500 平方公里，以便觀察網格大小對預警發報的影響。嘗試透過面積劃分，希望能夠更清楚地評估震源遠離的區域是否仍需發佈警報。這樣的劃分方式可能有助於避免無震感的區域收到不必要的預警，從而提升預警的準確性。

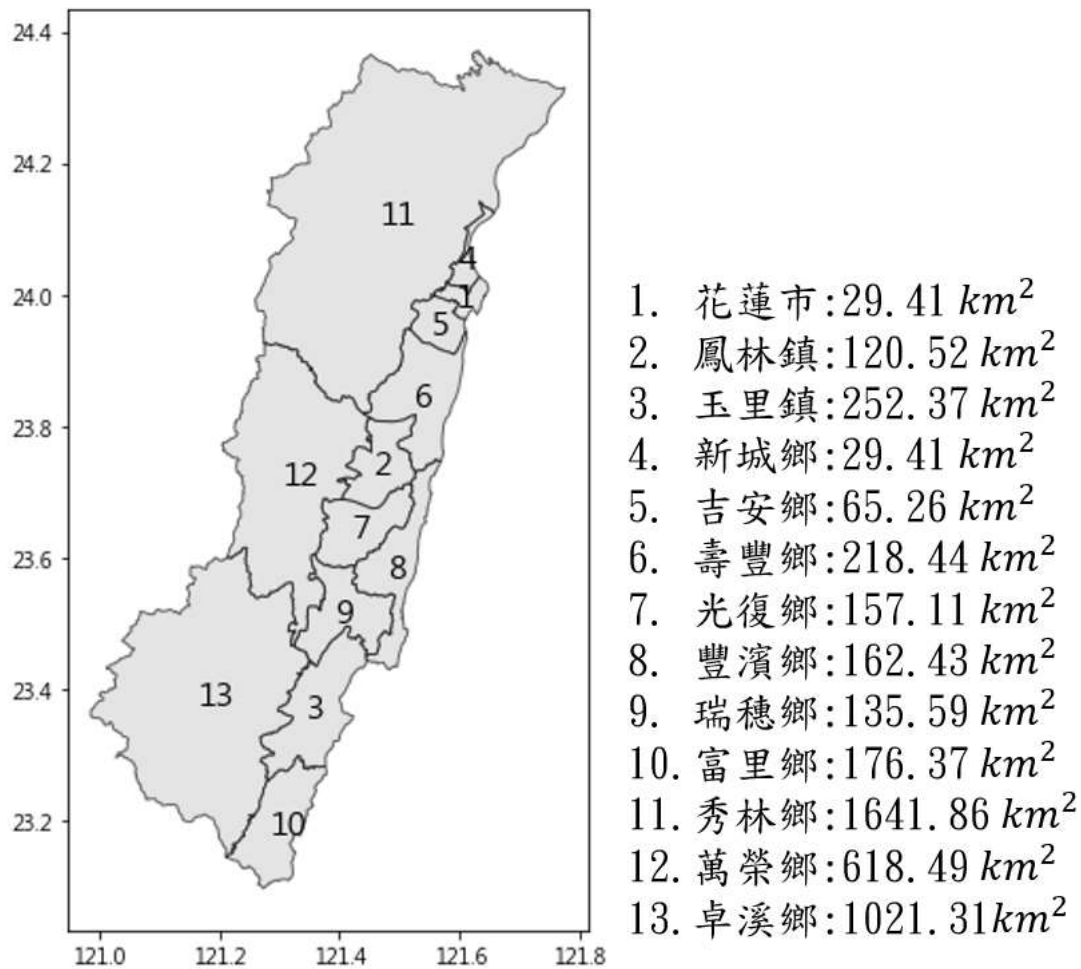


圖 2-32 花蓮鄉鎮市區面積大小分布圖

各面積劃分條件下的網格化後實際震度分布圖、網格化後預測震度分布圖及縣市發報的二元混淆矩陣分布圖分別如下所示。若以面積 2000 平方公里進行劃分，大致可分為兩個區域，如圖 2-33 所示。網格化後的發報的二元混淆分佈結果如圖 2-34，近震央區域（綠色）為正確發報，遠震央區域（黃色）為正確不報，兩個區域皆發報正確。



圖 2-33 以面積 2000km^2 劃分區域圖

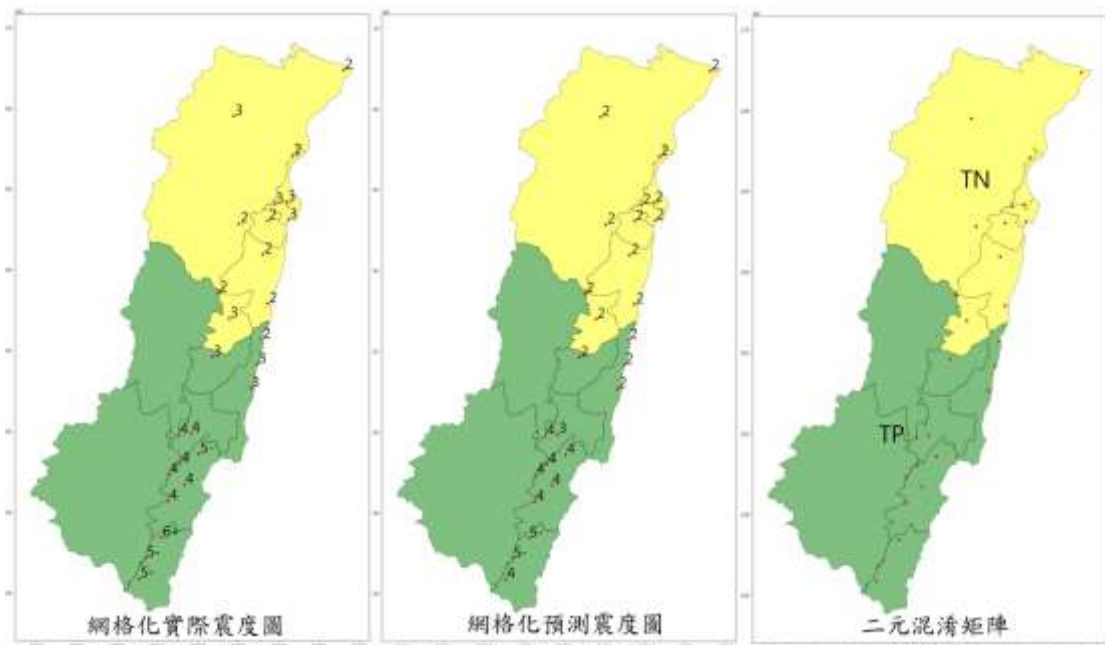


圖 2-34 以面積 2000km^2 劃分之網格化結果

若以 1500 平方公里劃分，大致可分為三個區域，如圖 2-35 所示。如圖 2-36 表示，藍色和黃色區域的發報皆正確，而綠色區域雖然最大震度有達 4 級，但因未滿足 Twin 機制，導致漏報。然而，若在正負一級的容許誤差情況下，該區域則可不需發報。

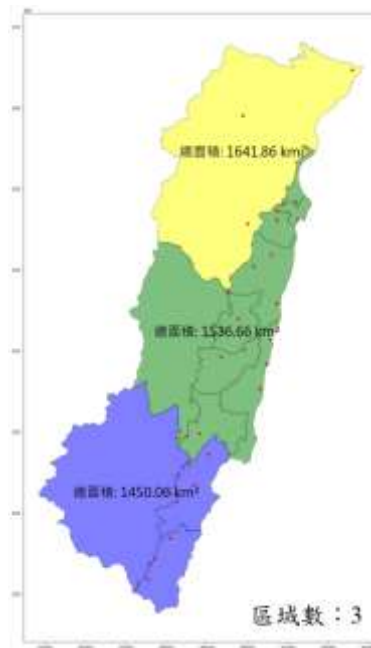


圖 2-35 以面積 1500km^2 劃分區域圖

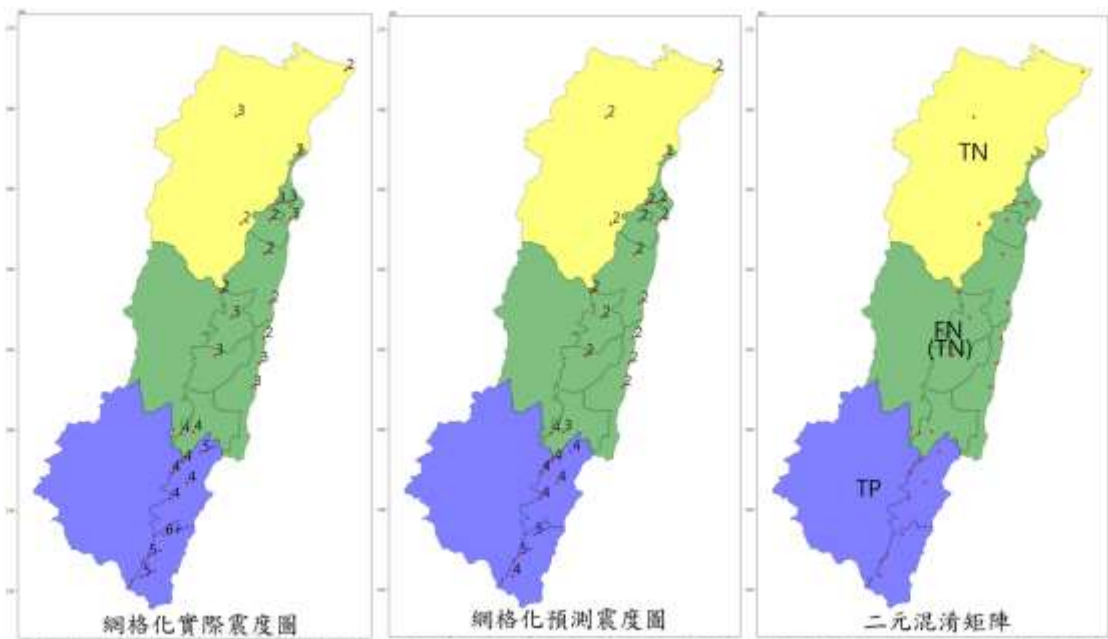


圖 2-36 以面積 1500km^2 劃分之網格化結果

若以 1000 平方公里劃分，大致可分為四個區域，如圖 2-37 所示。如圖 2-38 表示，藍色、橘色和黃色區域皆發報正確，而綠色區域雖最大震度達 4 級，但因不滿足 Twin 機制，造成漏報。然而，若有容許誤差，則該區域可不需發報。

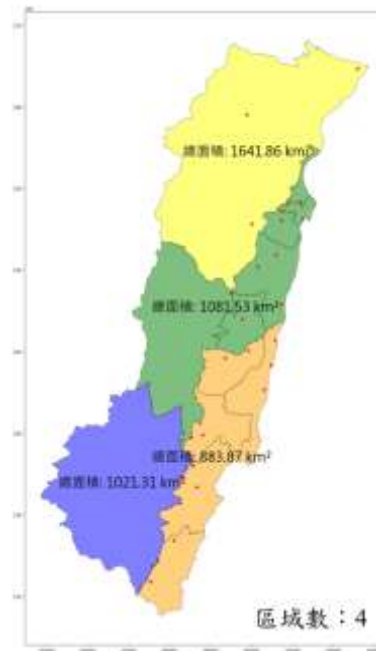


圖 2-37 以面積 1000km^2 劃分區域圖

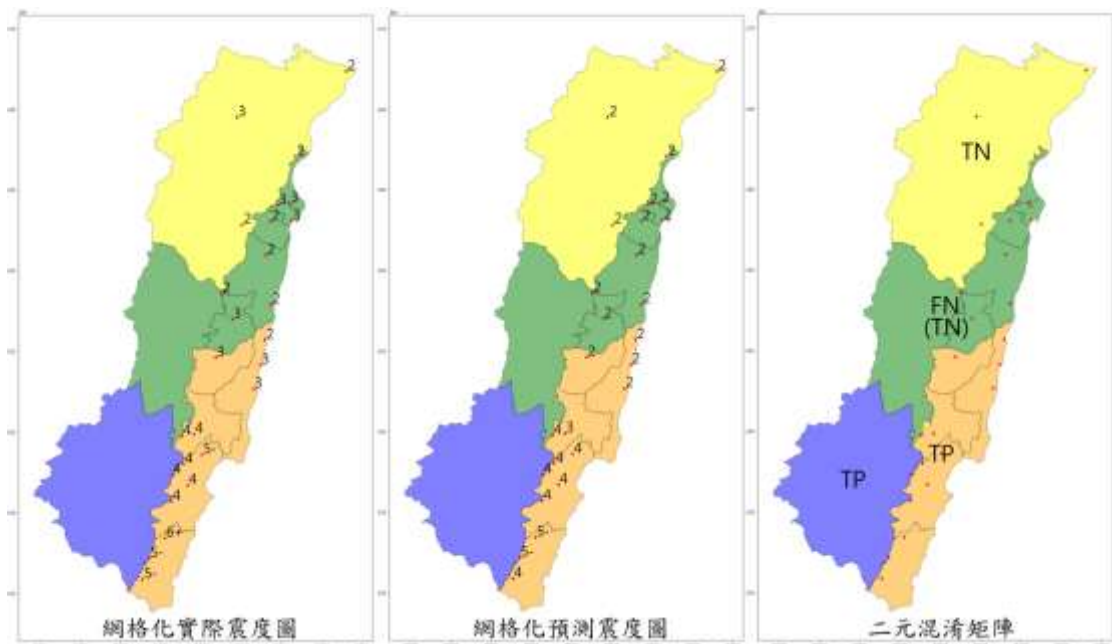


圖 2-38 以面積 1000km^2 劃分之網格化結果

若以 500 平方公里劃分，大致可分為六個區域，如圖 2-39 所示。圖 2-40 表示，藍色、橘色、白色和黃色區域皆發報正確，而綠色雖最大震度達 4 級，但因不滿足 Twin 機制，造成漏報。然而，綠色和灰色區域在有容許誤差的情況下，則會變成可以不需發報。



圖 2-39 以面積 500km^2 劃分區域圖

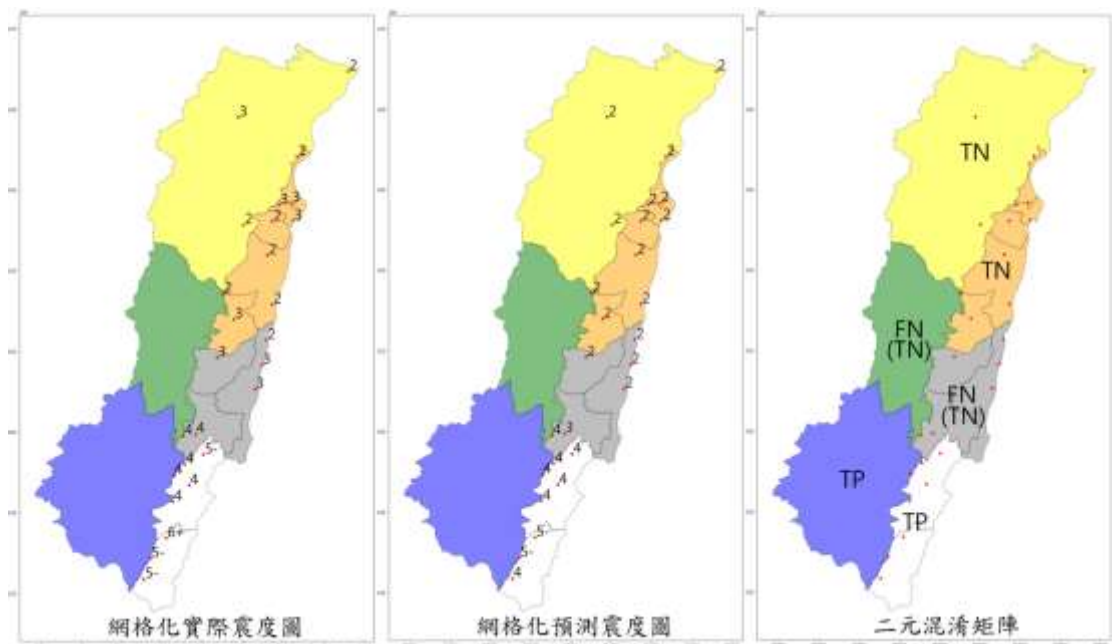


圖 2-40 以面積 500km^2 劃分之網格化結果

2.5.3 以居住人口作為劃分依據

考量人口密度差異，本研究也嘗試根據居住人口劃分網格區域，圖 2-41 顯示了花蓮縣各鄉鎮市區的居住人口分布（參考 109 年人口普查資料），分別設定每個區域內的居住人口約為 10 萬、5 萬及 1 萬人。此劃分方式將重點放在地震影響可能較大的高人口密度區域，避免低人口密度區域過度警報。嘗試透過人口導向的劃

分方式，希望能更有效地針對影響範圍進行準確的警報發送。

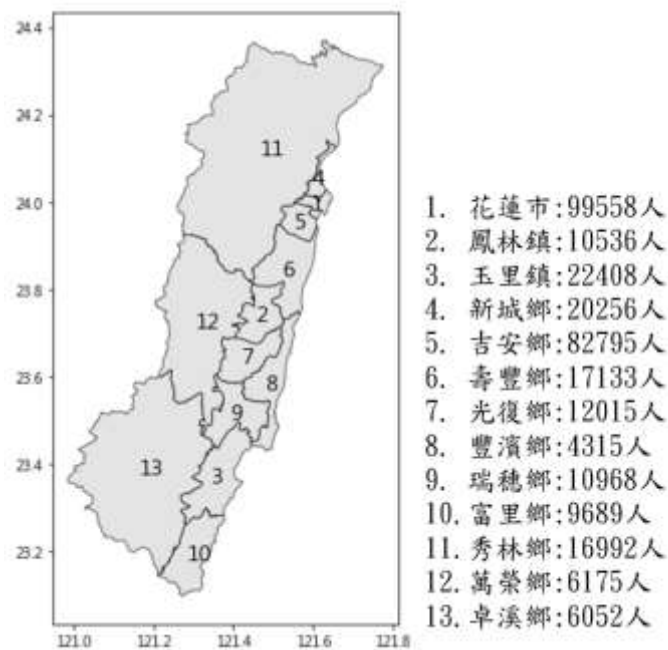


圖 2-41 花蓮鄉鎮市區居住人口分布圖

各居住人口劃分條件下的網格化後實際震度分布圖、網格化後預測震度分布圖及縣市發報的二元混淆矩陣分布圖分別如下所示。若以居住人口 10 萬人進行劃分，大致可分為三個區域，如圖 2-42 所示。網格化後的發報的二元混淆分佈結果如圖 2-43，三個區域皆發報正確。

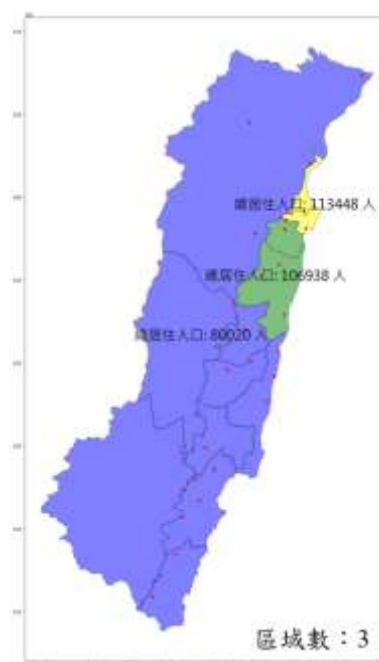


圖 2-42 以居住人口 10 萬人劃分區域圖

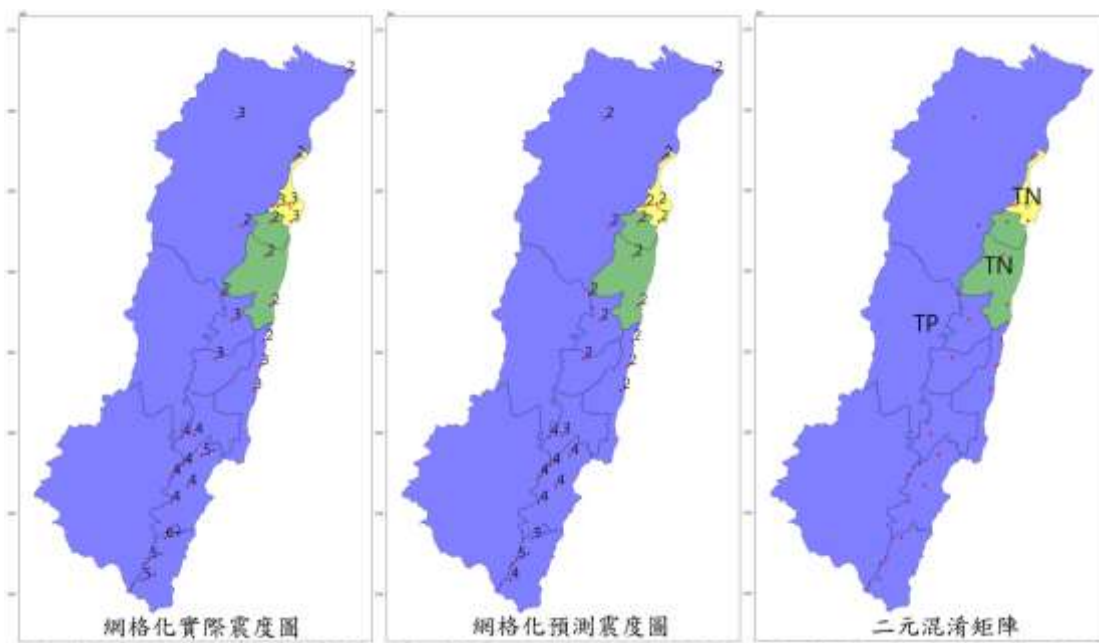


圖 2-43 以居住人口 10 萬人劃分之網格化結果

若以 5 萬人進行劃分，大致可分為四個區域，如圖 2-44 所示。圖 2-45 表示，除了藍色區域外的三個區域皆發報正確，而藍色區域雖然最大震度有達 4 級，但因未滿足 Twin 機制，導致漏報。然而，若在正負一級的容許誤差情況下，該區域則可不需發報。



圖 2-44 以居住人口 5 萬人劃分區域圖

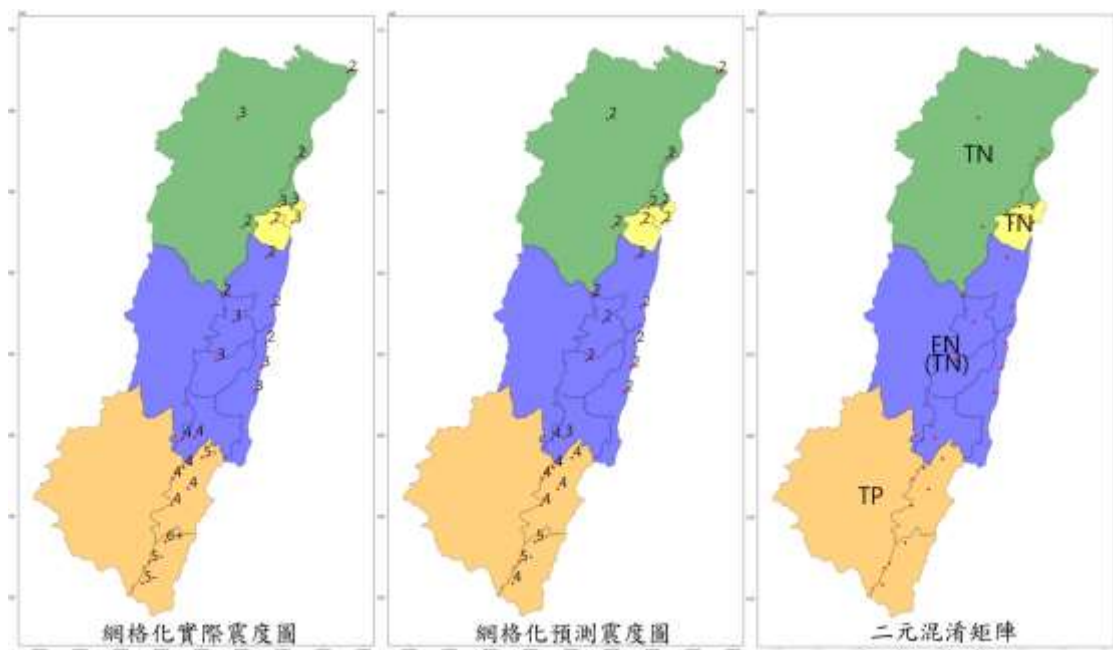


圖 2-45 以居住人口 5 萬人劃分之網格化結果

若以居住人口 1 萬人進行劃分，大致可分為七個區域，如圖 2-46 所示。圖 2-47 表示，除了橘色區域外的六個區域皆發報正確，而橘色區域雖最大震度達 4 級，但因不滿足 Twin 機制，造成漏報。然而，若有容許誤差，則該區域可不需發報。



圖 2-46 以居住人口 1 萬人劃分區域圖

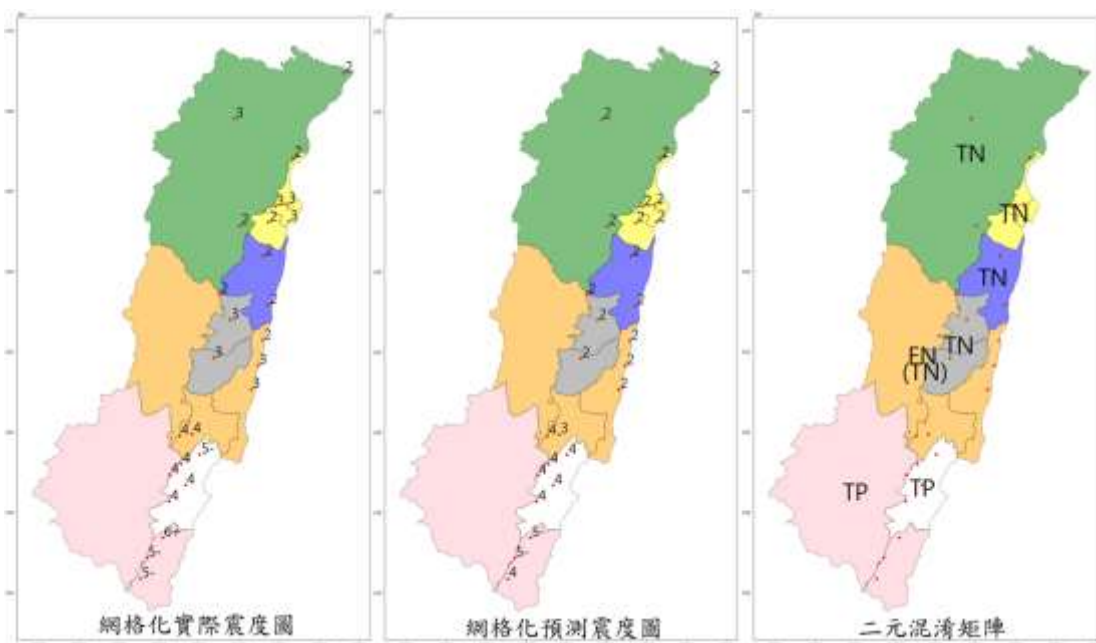


圖 2-47 以居住人口 1 萬人劃分之網格化結果

2.5.4 以測站數量作為劃分依據

此外，本研究也依測站數量進行劃分，圖 2-48 顯示了花蓮縣各鄉鎮市區的測站數量分布 (TSMIP 測網)，根據區域內的測站配置將網格劃分為包含約 15、10 和 5 個測站的網格單位。此方式能夠充分利用現有測站的監測數據，提升各區域震度監測的精確度。透過不同測站密度的網格劃分，能更快速而有效地掌握各區域的地震強度，從而提高預警系統的靈敏度與可靠性。

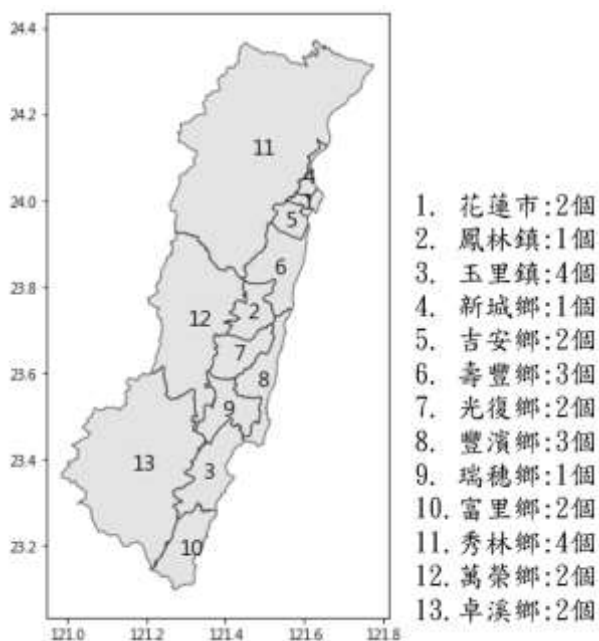


圖 2-48 花蓮鄉鎮市區測站數量分布圖

各測站數量劃分條件下的網格化後實際震度分布圖、網格化後預測震度分布圖及縣市發報的二元混淆矩陣分布圖分別如下所示。若以測站數量 15 個進行劃分，大致可分為兩個區域，如圖 2-49 所示。網格化後的發報的二元混淆分佈結果如圖 2-50，兩個區域皆發報正確。



圖 2-49 以測站數量 15 個劃分區域圖

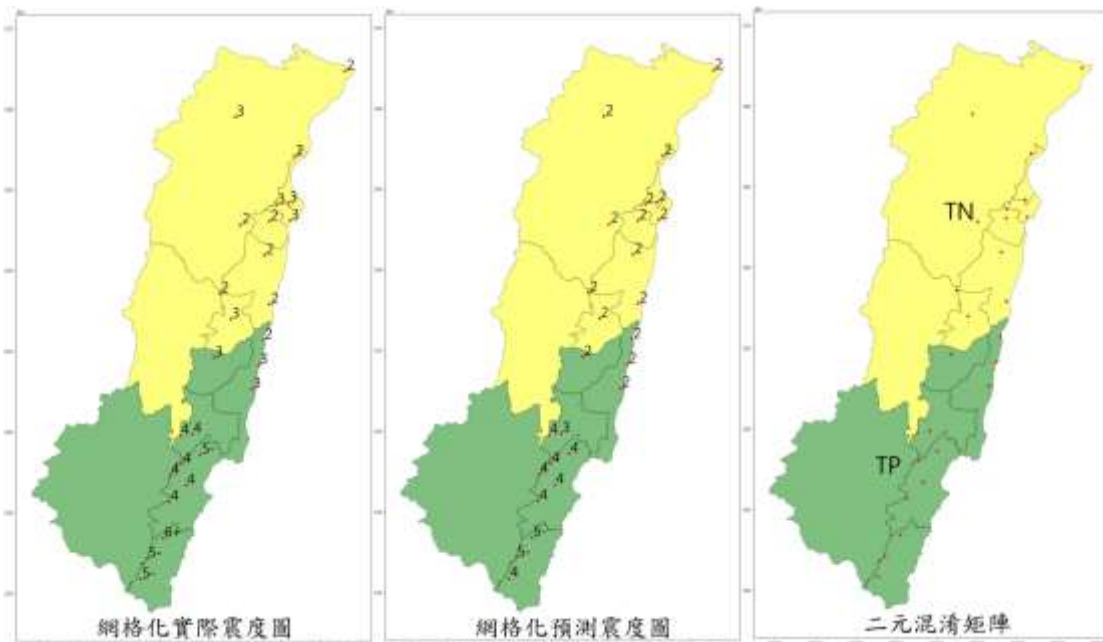


圖 2-50 以測站數 15 個劃分之網格化結果

若以 10 個測站進行劃分，大致可分為三個區域，如圖 2-51 所示。圖 2-52 表

示，除了綠色區域外的兩個區域皆發報正確，而綠色區域雖然最大震度有達 4 級，但因未滿足 Twin 機制，導致漏報。然而，若在正負一級的容許誤差情況下，該區域則可不需發報。



圖 2-51 以測站數量 10 個劃分區域圖

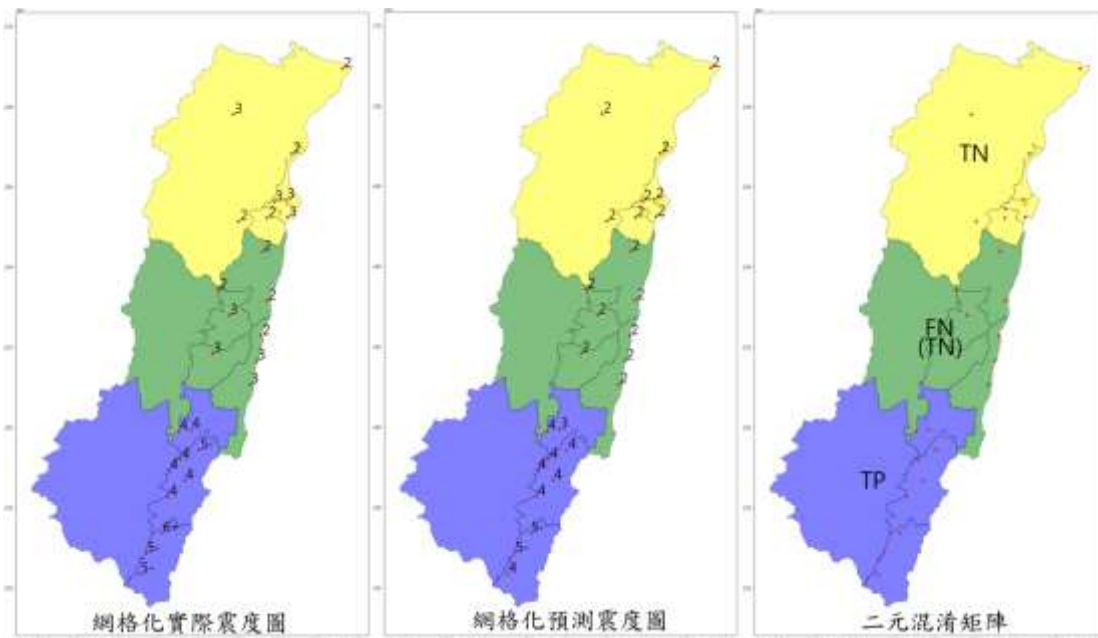


圖 2-52 以測站數 10 個劃分之網格化結果

若以 5 個測站進行劃分，大致可分為六個區域，如圖 2-53 所示。圖 2-54 表示，除了藍色區域外的五個區域皆發報正確，而藍色區域雖最大震度達 4 級，但因

不滿足 Twin 機制，造成漏報。然而，若在有容許誤差，則該區域可不需發報。



圖 2-53 以測站數量 5 個劃分區域圖

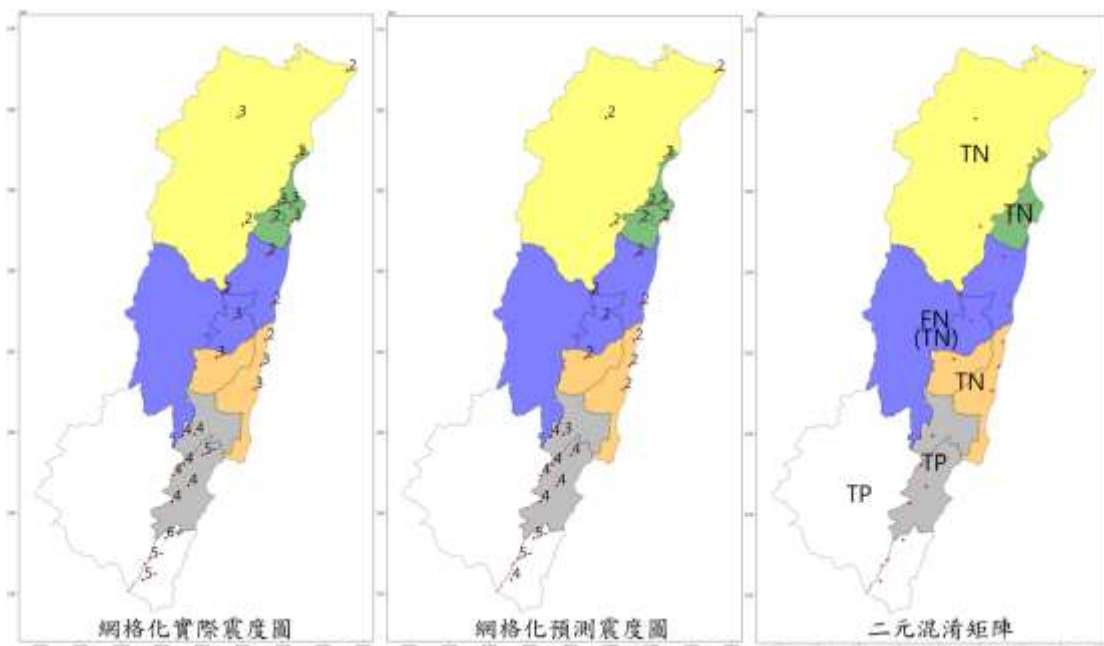


圖 2-54 以測站數 5 個劃分之網格化結果

本研究針對不同劃分依據進行地震預警系統的網格化分析，結果顯示若適當地利用面積大小和測站數量作為劃分依據，具有一定的可行性。在適當的網格劃分下，雖然存在少數漏報情形，但這些測站的 PGA 值仍在容許誤差範圍內，整體預測準

確度較高。然而，若將區域劃分過多，可能導致某些網格形狀狹長不規則，影響發報的準確性。另一方面，若以居住人口數作為劃分基準，會因人口分佈較為不均，導致網格大小差異過大，因此可能較不適用於地震預警系統。綜合以上考量，建議同時考量面積大小和測站數量作為網格化依據，這種劃分方式可能有助於提升發報準確性。

2.6 小結

第二章主要針對：(1)2024 年 4 月 3 日花蓮地震事件之問題、(2)應用 MQTT 架構和 Triton Inference Server 試著縮短模型預測之時間、(3)應委員期中之審查意見分析 2024 年 1 月至 9 月線上模型預測情況、(4)透過縣市網格化試著提升模型對幅員廣大縣市之預測效果等進行說明。

2024 年 4 月 3 日花蓮地震事件發生時，有發現接收相位時間和模型預測時間都過長之問題，故應用 MQTT 架構試著縮短接收相位時間，效果並不顯著。而應用 Triton Inference Server 試著縮短模型預測時間，確實可能使模型預測時間縮短。目前正將採用上述 MQTT 架構及 Triton Inference Server 運作方式的程式部屬至氣象署中。

2024 年 1 月至 9 月線上模型預測情況有漏報或誤報之情況，漏報的主因有三類：一是部分地震的波形在 P 波到達後的 2 秒內較小，導致模型預測值偏低；二是有些事件之前的較小地震使模型觸發過早，影響預測結果；三是有些測站的主震並沒有被觸發。誤報則多因程式中 P 波點位記錄不正確，因程式儲存的波形每 5 秒更新一次，當 P 波位置偏移 500 點時，易使預測值過高，引發誤報。誤報部分改進之程式已部屬至氣象署，等待之後地震驗證其成效。

關於網格化地震預警範圍，若適當地利用面積大小和測站數量作為劃分化分基準具有一定的可行性，即便存在少數漏報情形，但這些測站的 PGA 誤差仍在容許範圍內，因此能有效提升發報準確度；但若網格過多，可能出現不規則形狀，影響發報效果。相比之下，若以人口數作為劃分基準，因人口分布較為不均，可能造成網格大小差異較大，因此較不適合用於地震預警系統。建議同時考量面積和測站數量進行網格化劃分，以提升發報的準確性。

3 單站法即時地震波到時標記及強震預警模型開發

3.1 研究動機

開發高效能的強震預警系統，為本研究的核心目標。在強震預警領域中，準確偵測 P 波相位和預測地表最大加速度(PGA)為兩大關鍵任務。目前，大多數 PGA 或震度預測模型仰賴 P 波到時資訊作為重要參數輸入。然而，在先前預測模型的實際應用中，常因 P 波標記錯誤或遺漏而導致預測失準或漏報，或者因為標記資訊傳遞的延時導致時效不佳。這種高度依賴性凸顯了準確標記 P 波相位的重要性，將直接影響 PGA 預測的精準度和時效性。

為解決此一問題，本研究期望能提出一種多任務學習(Multi-Task Learning, MTL)模型，能同步執行 P 波相位偵測和 PGA 預測。多任務學習係在單一網路中執行多個任務，且任務間可共享特徵表示(Ruder [4]、Crawshaw [5])。本研究參考 Liao 等人[2]的遞迴殘差 U 型網(Recurrent Residual U-Net)作為特徵共享網路，並參考 Chen 等人[4]以 Z-score 標準化與 2.5gal 加速度歷時尺度作為輸入，設計兩個專用的注意力子網路(Attention sub-net)，分別用於 P 波相位偵測和 PGA 超過 25gal 的機率值預測。

於先前的研究中，Chen 等人[3]發現僅以 Z-score 與 2.5gal 加速度歷時尺度作為輸入的模型，對於預測 PGA 是否超過 25gal 存在低召回率的問題。不考慮容許誤差時，該模型在 P 波到時後 2-3 秒的整體發報準確率為 89.3%，召回率僅 38.5%，F1 分數為 53.8%。因此，本計畫引入時頻域特徵，將模型輸入由單一時間域輸入擴展為時域與時頻域雙輸入。此模型名稱為 DAPPER (Dual-domain Attentive Phase Picker and Earthquake Reporter)。

3.2 模型架構

本研究借鑒了 Liao 等人[2]的方法，採用遞迴殘差 U 型網 (Recurrent Residual U-Net) 作為主要架構，實現特徵共享。為了處理時間域與時頻域的資訊，並於前端加入了兩個分支，這些資訊經過融合層 (concatenate layer) 後作為 Mtan U-Net 的輸入。Mtan U-Net 不僅包含前述所說的遞迴殘差 U 型網，並配合兩個專門針對各自任務設計的注意力子網路 (Attention sub-net)，以提升各自任務處理的專注度與效能。模型架構圖，請參考圖 3-1、圖 3-2、圖 3-3、圖 3-4 及圖 3-5。

。

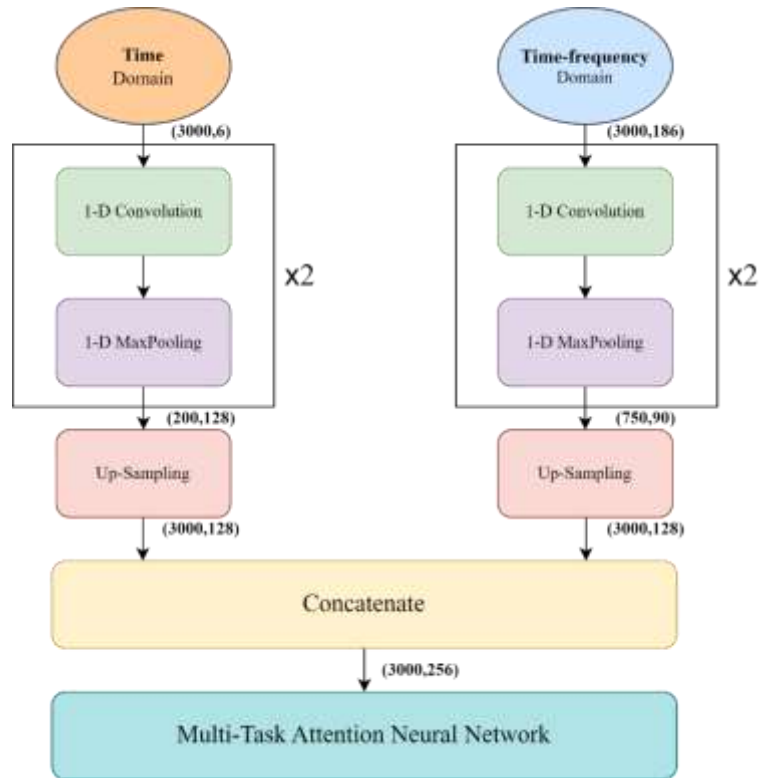


圖 3-1 時域與時頻域雙輸入架構

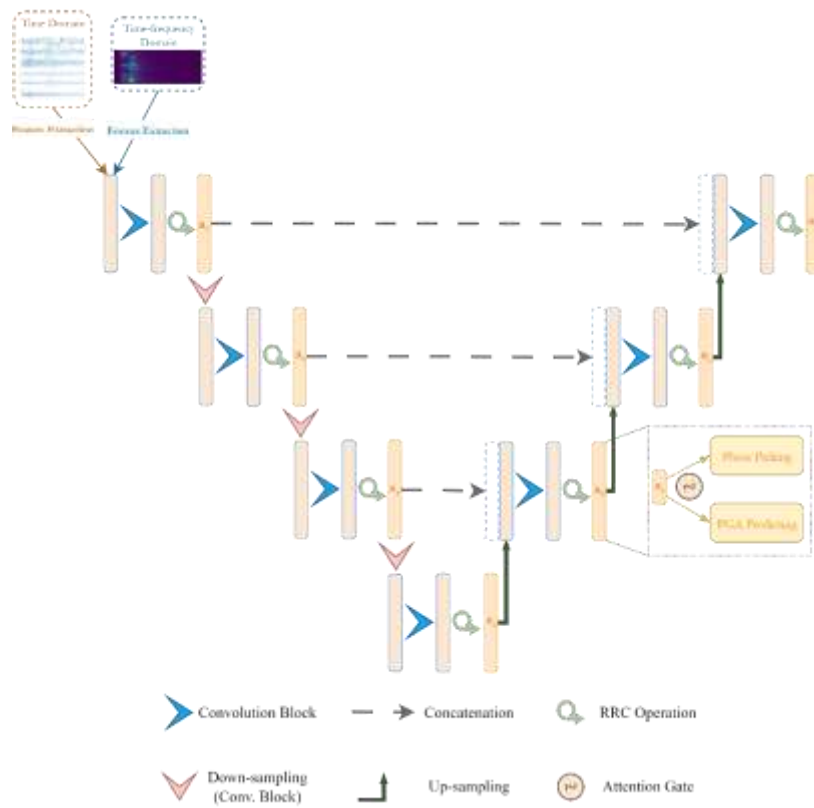


圖 3-2 Multi-task Attention Neural Network Architecture 架構

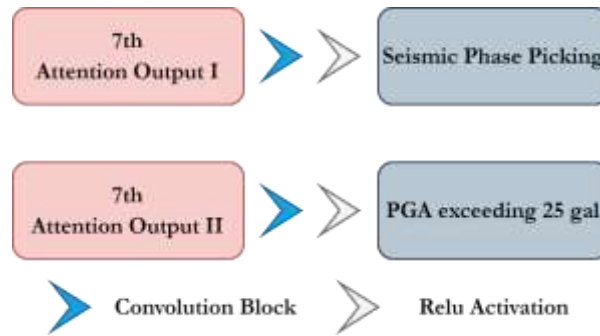


圖 3-3 特殊任務注意力網路(Task-Specific Attention Network)

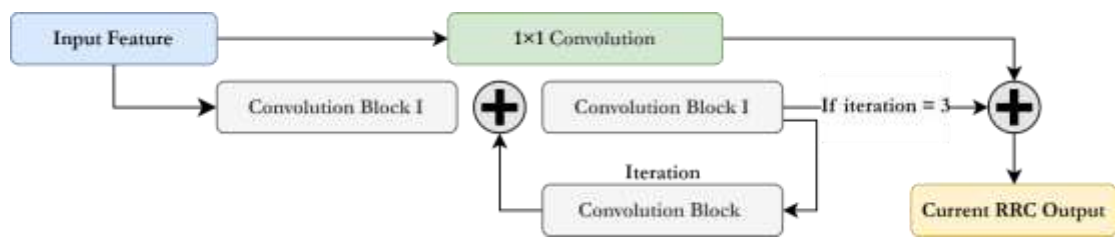


圖 3-4 RRC Operation

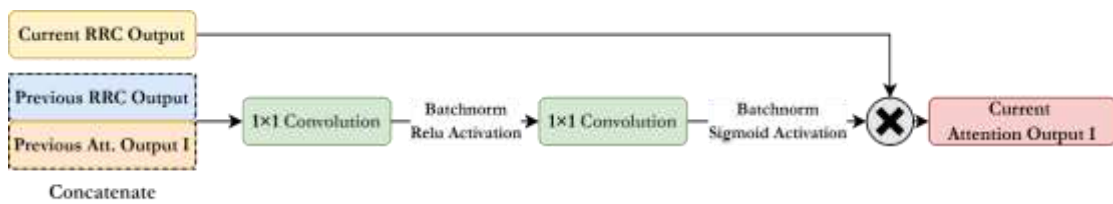


圖 3-5 Attention Gate

3.3 資料集

本研究利用的數據來源於兩個不同的地震觀測網絡：一是 2012 至 2019 年間的中央氣象署地震觀測網（Central Weather Administration Seismographic Network, CWASN）；二是 1992 至 2006 年間的臺灣強地動觀測網（Taiwan Strong Motion Instrument Program Network）。這兩個測網所收集的數據均統一為 100 赫茲的採樣率，且每份波形數據的長度為 30 秒。各測網資料集數量如下表 3-1，而訓練、測試與驗證資料集於各加速度區間分布如下圖 3-6、圖 3-7 及圖 3-8 所示。本研究資料集還加入了中央氣象署地震觀測網（Central Weather Administration Seismographic Network, CWASN）噪音資料集。

表 3-1 各測網資料數量

	Network	Training	Validation	Testing
--	---------	----------	------------	---------

	(Year)			
Waveform	CWASN (2012-2021)	22647	5774	5497
	TSMIP (1992-2006)	33194	8407	5490
Noise	CWASN	23260	5815	5000
Total	-	79101	19996	15987



圖 3-6 訓練資料集各加速度區間分布

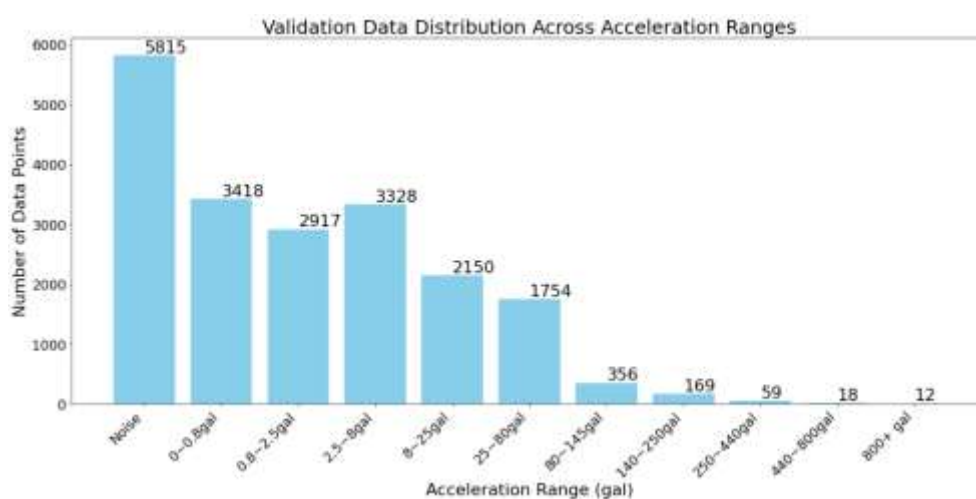


圖 3-7 驗證資料集各加速度區間分布

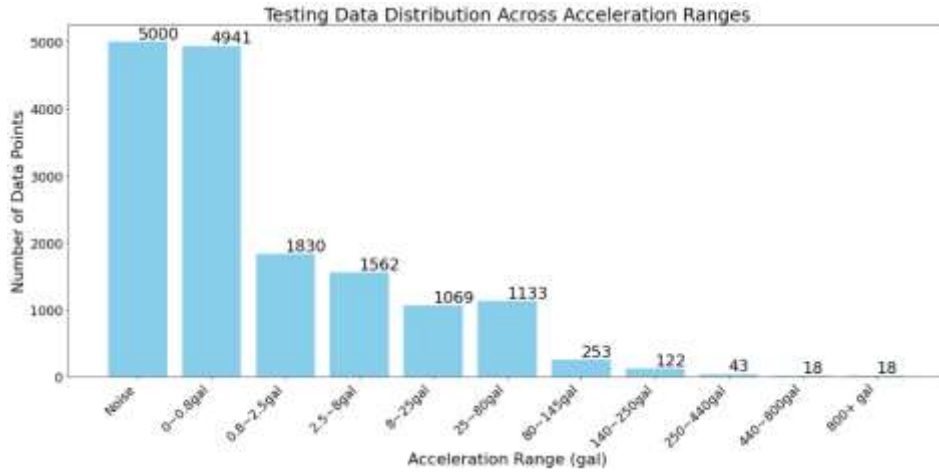


圖 3-8 測試資料集各加速度區間分布

3.4 模型輸入及輸出

3.4.1 模型輸入

● 時間域輸入處理

在本研究中，模型有一項任務需要預測的是最大地表面加速度（PGA）是否超過 25gal，故對振幅的資訊尤其重要。因此，參考了 Chen[3] 等人於先前的研究，結合標準分數（Z-score）和 2.5gal 的加速度尺度縮放，讓模型同時獲得振幅大小和訊號形狀的資訊。

標準分數(Standard Score, Z-score)計算公式如式(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。

$$Acc(t_i)_{zscore} = \frac{Acc(t_i) - mean(Acc(t_{all}))}{\sigma(Acc(t_{all}))}, i = 1 \sim 3000 \quad (3-1)$$

2.5gal 加速度區間正規化計算公式如式(3-2)，長度 3000 點之加速度歷時中，加速度大於 2.5gal 之值會轉化為 1，加速度小於-2.5gal 之值會轉化為-1，加速度位於 2.5gal 至-2.5gal 區間內，則直接除以 2.5，最終，會將加速度值正規化至[-1,1]的範圍內。

$$\left\{ \begin{array}{l} \text{Acc}(t_i) > 0, i = 1 \sim 3000, \\ \text{Acc}(t_i) < 0, i = 1 \sim 3000, \end{array} \right\} \left\{ \begin{array}{l} \text{Acc}(t_i) \geq \text{加速度區間值}, \text{Acc}(t_i) = \frac{\text{加速度區間值}}{\text{加速度區間值}} = 1 \\ \text{Acc}(t_i) < \text{加速度區間值}, \text{Acc}(t_i) = \frac{\text{Acc}(t_i)}{\text{加速度區間值}} \\ \text{Acc}(t_i) \leq -\text{加速度區間值}, \text{Acc}(t_i) = \frac{-\text{加速度區間值}}{\text{加速度區間值}} = -1 \\ \text{Acc}(t_i) > -\text{加速度區間值}, \text{Acc}(t_i) = \frac{\text{Acc}(t_i)}{\text{加速度區間值}} \end{array} \right. \quad (3-2)$$

時間域輸入如下圖 3-9 所示。

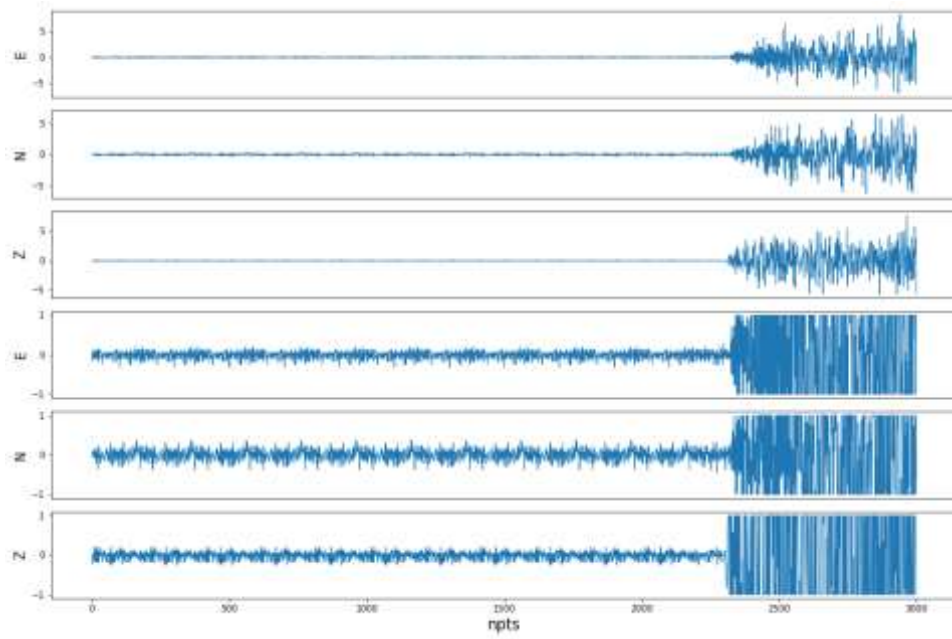


圖 3-9 時間域輸入

3.4.2 時頻域輸入處理

● 短時傅立葉變換 (Short-Time Fourier Transform, STFT)

在本研究中，為了進一步提升地震預測模型的性能，引入時頻域特徵。將時間歷時訊號進行時頻分析，期望模型同時觀察訊號在時間和頻率上的特徵變化。

本研究採用短時傅立葉變換 (Short-Time Fourier Transform, STFT) 來分析原始加速度訊號。短時傅立葉變換是一種常見的訊號處理方式，通常拿來描繪頻域與時域上的變化，為時頻分析中一種重要的工具。在離散的時間歷時資料中，歷時資料會被分割成數組等長訊號，而每組訊號乘上窗函數後再各自進行傅立葉轉換，即可獲得每個時間與頻率變化的大小與相位。數學上，STFT 如式(3-3)，其中 $x[n]$ 是待變換的訊號， $\omega[n-m]$ 是窗函數， N 是窗函數大小， m 是步進。可將其理解為使用窗函數和步進將訊號化分為多段訊號，並將每段訊號單獨進行快速傅立葉轉換 (Fast Fourier Transform, FFT)，然後按照順序地將其拼接成一個頻譜圖，如圖 3-10 所示。

$$STFT[m, k] \equiv X[m, k] = \sum_{n=-\infty}^{\infty} x[n] \omega[n - m] e^{-j \frac{2\pi k n}{N}} \quad (3-3)$$

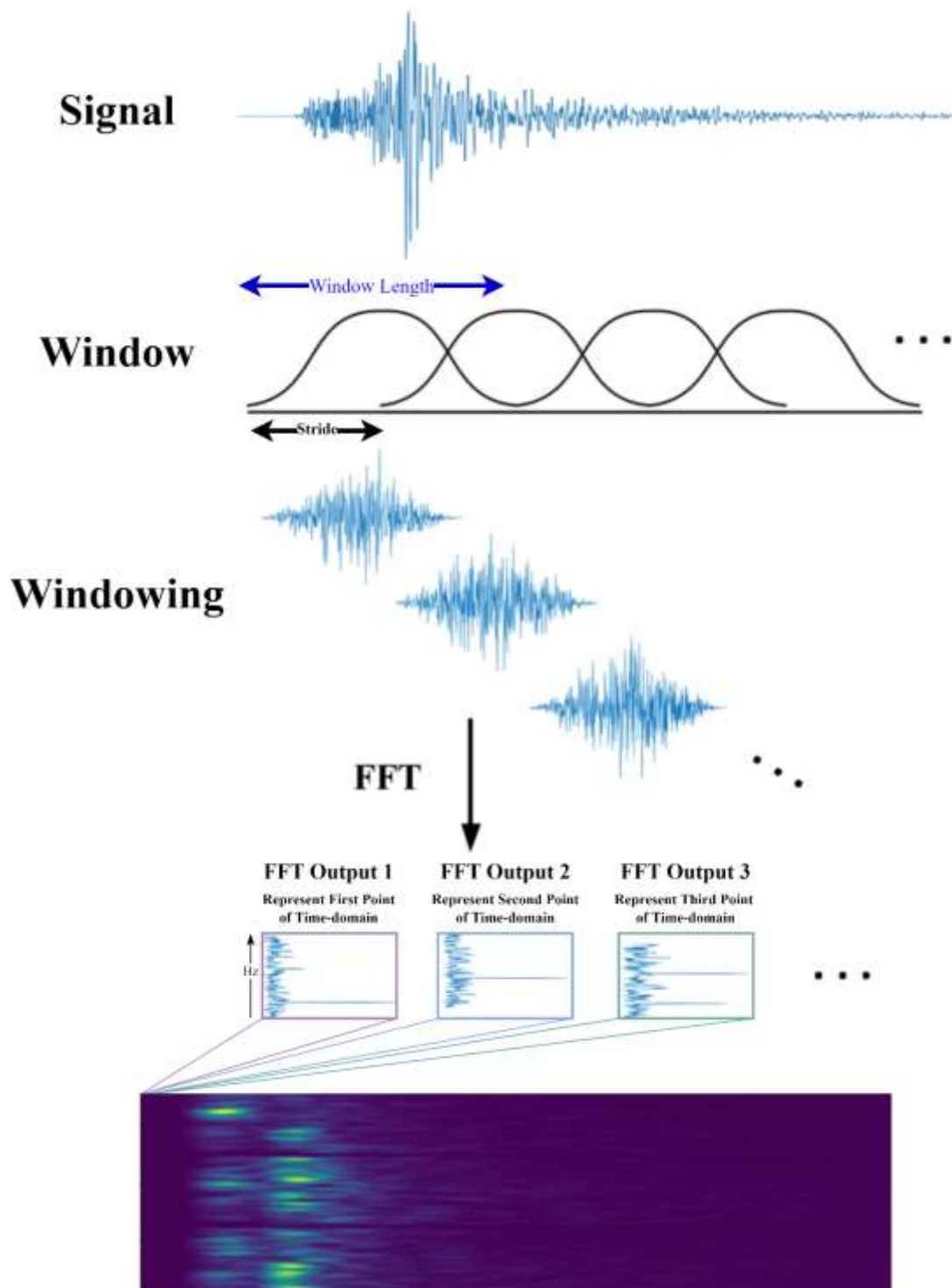


圖 3-10 STFT 概念圖

應用短時傅立葉變換，需要決定以下參數，窗函數、窗函數長度、步進。

窗函數一般常使用有以下幾種 Hanning (Hann) 窗函數、Hamming 窗函數、triangular 窗函數、boxcar 窗函數。本計畫選擇 Hamming 窗函數和 boxcar 窗函數做參數分析。

Hamming 窗函數：如圖 4.15 所示，其公式如 (4.24)。其在邊界處使用平滑過渡，降低了頻域中的旁瓣水平，增加了主瓣寬度，從而減少了頻譜洩漏。

boxcar 窗函數：如圖 4.16 所示，其公式如 (4.25)。其具有明顯的邊界，這會在頻域中產生較大的旁瓣，導致更明顯的頻譜洩漏。

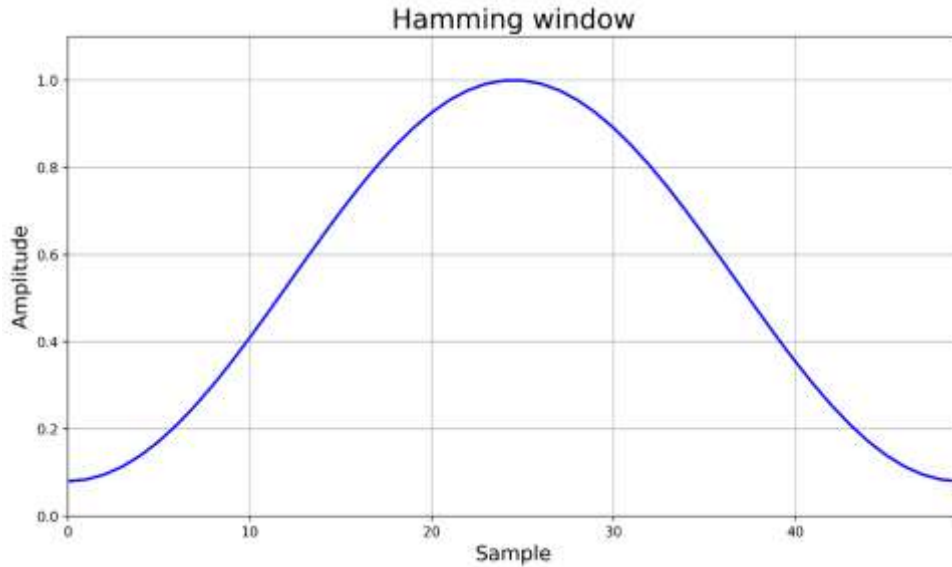


圖 3-11 Hamming 窗函數

$$\omega[n] = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right), & 0 \leq n \leq N-1 \\ 0, & \text{otherwise} \end{cases} \quad (3-4)$$

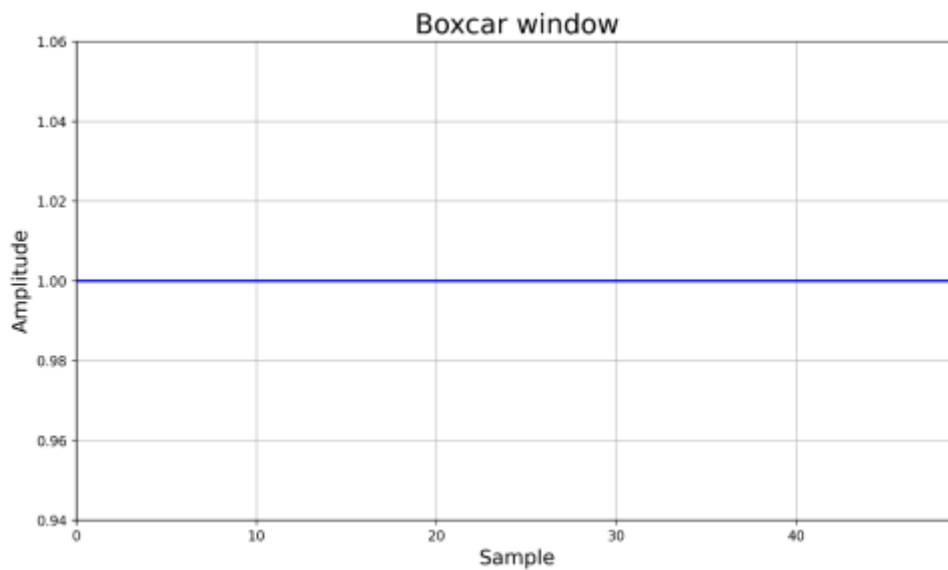


圖 3-12 boxcar 窗函數

$$\omega[n] = \begin{cases} 1, & 0 \leq n \leq N-1 \\ 0, & \text{otherwise} \end{cases} \quad (3-5)$$

窗函數長度和步進是影響時間域中信號長度和轉換後頻率解析度的兩個重要參數。窗口數量可以使用公式(3-6)計算，其中 L 是信號長度， N 是窗口大小（傅立葉變換的大小）， M 是步長。頻率解析度可以使用公式(3-7)計算，其中 Δf 是頻率解析度， f_s 是取樣率。

$$W = \left\lceil \frac{L-N}{M} \right\rceil + 1 \quad (3-6)$$

$$\Delta f = \frac{f_s}{N} \quad (3-7)$$

在本研究中，時頻域輸入的時間軸長度為 3000 點，每點代表 0.01 秒，取樣率為 100Hz。原始的加速度數據輸入同樣為 3000 點，為了保持此取樣率並確保變換後長度維持 3000 點，必須進行零填充。

本研究使用 boxcar 窗函數搭配 150 點的窗口長度、1 點的步長。根據公式(3-6)，這將僅產生 2851 點的資料輸出，因此需要額外填充 149 點以保持變換後的時間長度為 3000 點。然而，這 149 個零點的填充方式將影響頻譜圖中能量出現的位置。如圖 3-13 所示，若這些 149 點填充在起始處，能量出現的位置將延遲；若填充在末端，能量出現的位置將提前。本計畫選擇在前面補 72 個零；後面補 77 個零，如圖 3-14 中之 Default(72/77)。

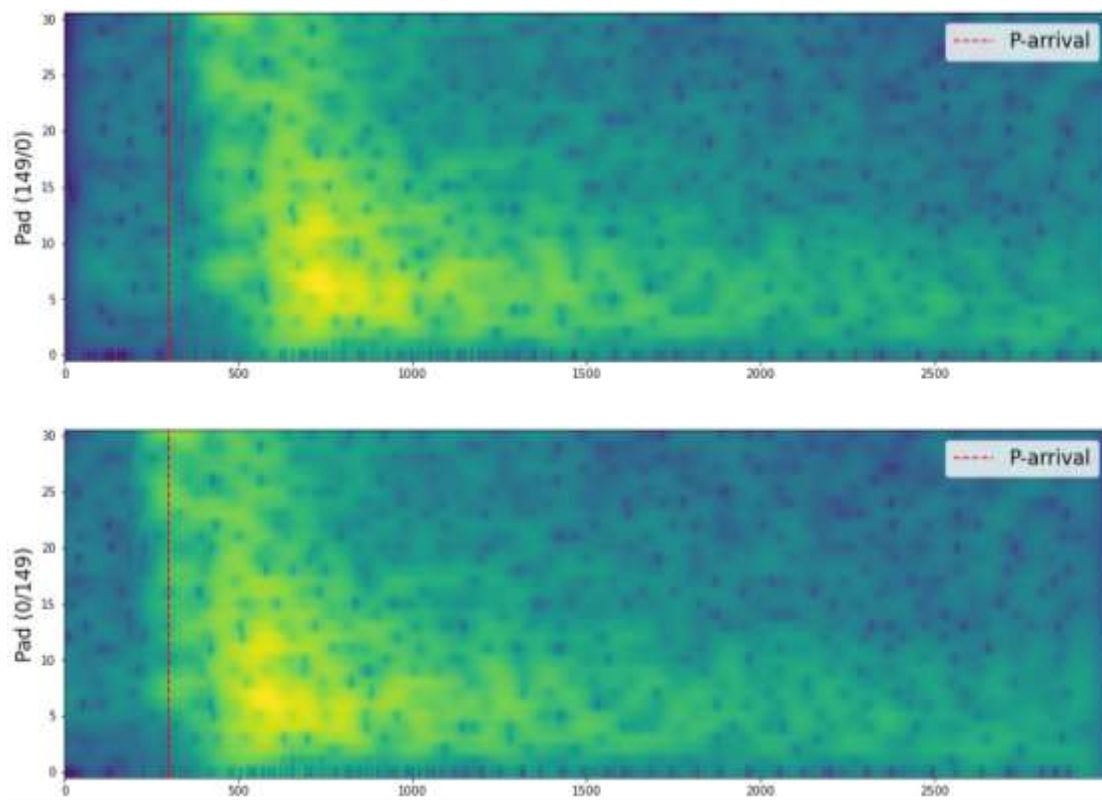


圖 3-13 把零全部補在前面或後面

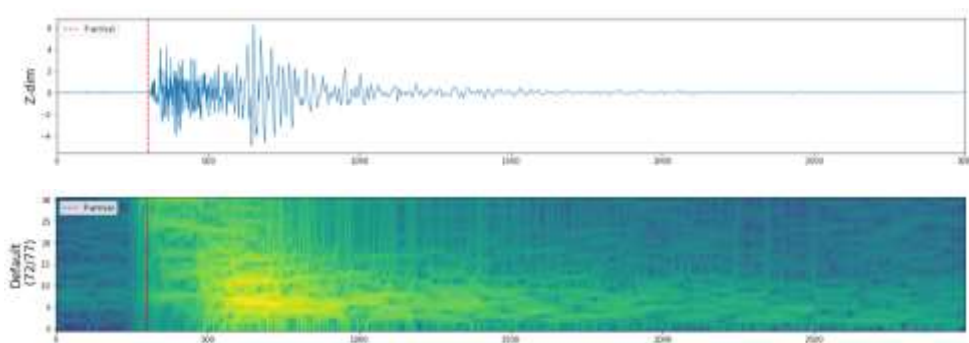


圖 3-14 使用 Boxcar 窗函數時，前面補 72 個零；後面補 77 個零之時頻域

由於地震事件的頻率資訊大多集中在 20Hz 以下，故本研究僅提取頻譜圖中 20Hz 以下的資訊以進行進一步的數據處理。由於頻率解析度為 $2/3(100/150)$ ，所以每軸之頻率有 31 個頻率，因為有三軸向，所以頻率軸共有 93 個頻率。

● 對時頻域訊號強度做對數正規化

由於所有事件的加速度分佈範圍較廣，經過 STFT 後的數值差異很大。在研究的早期階段，曾嘗試將經過 STFT 處理的原始加速度直接輸入模型進行訓練，但模型遇到了梯度爆炸的問題，無法收斂。因此，本研究提出了對數正規化方法來解決

此問題。為了了解 STFT 強度的分佈，我們對測試數據集進行了一些分析。首先，會分別計算訓練數據集中每個事件的最大三軸合成加速度。然後，對每個事件進行 STFT 以提取頻譜圖中的極值。接著，將最大三軸合成加速度和頻譜圖中的極值分別進行以 10 為底的對數變換，如下圖 3-15。可以發現對數化後的 STFT 頻譜分布介於[-5,4]之間，本研究進一步對此區間進行正規化，如(3-8)，並將三軸合併起來作為輸入，時頻域輸入如下圖 3-16。

$$STFT_{log-normalized} = \frac{STFT_{log_{10}Intensity} - (-5)}{4 - (-5)} \quad (3-8)$$

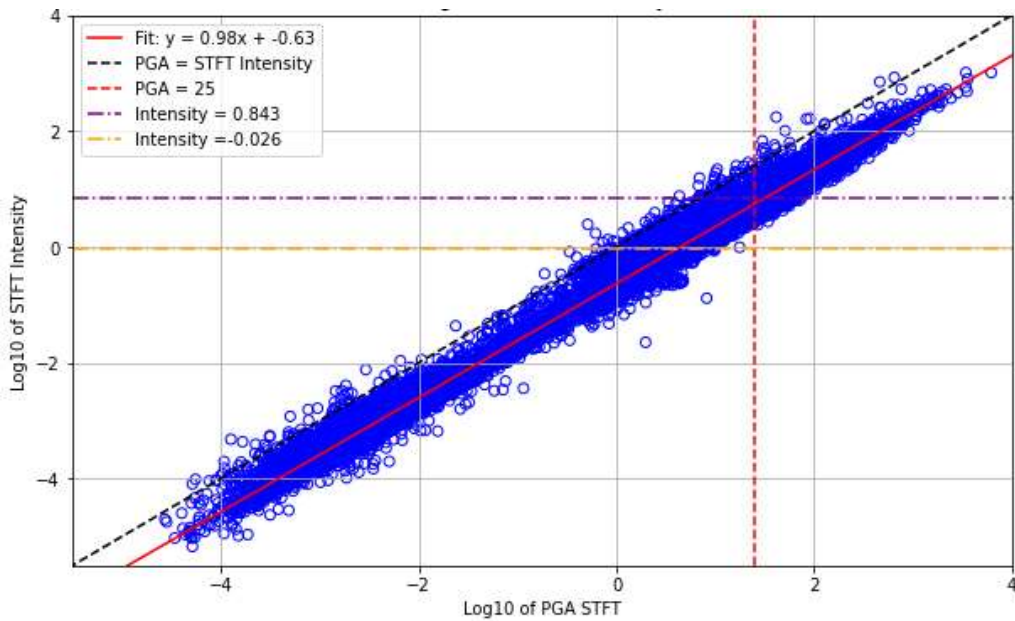


圖 3-15 加速度與頻譜對數(log10)分布圖

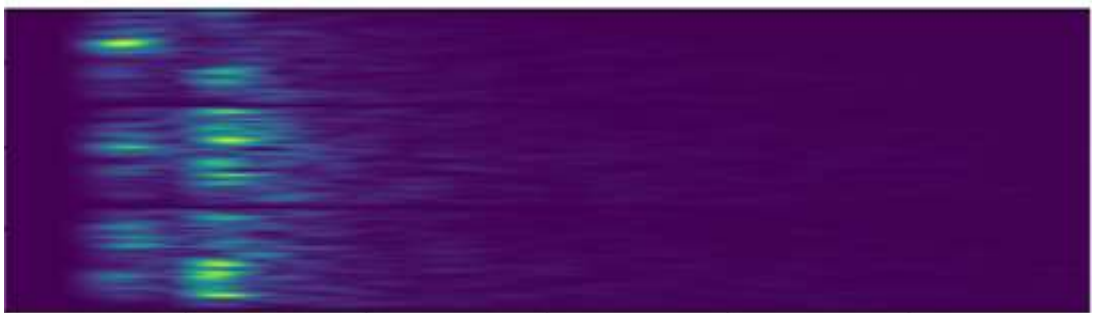


圖 3-16 時頻域輸入

除了對頻譜圖強度進行全範圍的對數正規化外，本計畫再添加特定區間的對數正規化。特定區間的選擇是基於對 4 級事件的 P 波進行 STFT 並以 10 為底的對數處理後的頻譜圖強度，然後進行分佈統計。針對 4 級事件原因是此模型對單站發報條件為預測 PGA 超過 25gal 之機率值超過門檻值。圖 3-17 的直方圖顯示，分

佈近似於常態分佈。本計畫對特定區間對數正規化之選擇為 ± 2 倍時頻域強度之標準差內。

所以本計畫從三軸向加速度訊號轉換成時頻域輸入訊號之步驟為：

1. 對每軸加速度訊號前方補 72 個零，在後方補 77 個零。矩陣大小為 3149×3 。
2. 對每軸做短時傅立葉變換，用 boxcar 窗函數搭配 150 點的窗口長度、1 點的步長。將每軸之加速度訊號轉換成時頻域訊號，並合併成三軸向時頻域訊號。矩陣大小為 3000×93 。
3. 將轉換後每軸之時頻域訊號強度做對數正規化，將時頻域訊號強度縮至 $[-5, 4]$ 之間，並保留此矩陣。矩陣大小為 3000×93 。
4. 將轉換後每軸之時頻域訊號強度做對數正規化，將時頻域訊號強度縮至 ± 2 倍時頻域強度之標準差內，並保留此矩陣。矩陣大小為 3000×93 。
5. 將 3.和 4.之兩個矩陣合併為大小為 3000×186 之矩陣。此矩陣即為輸入模型時頻域訊號。

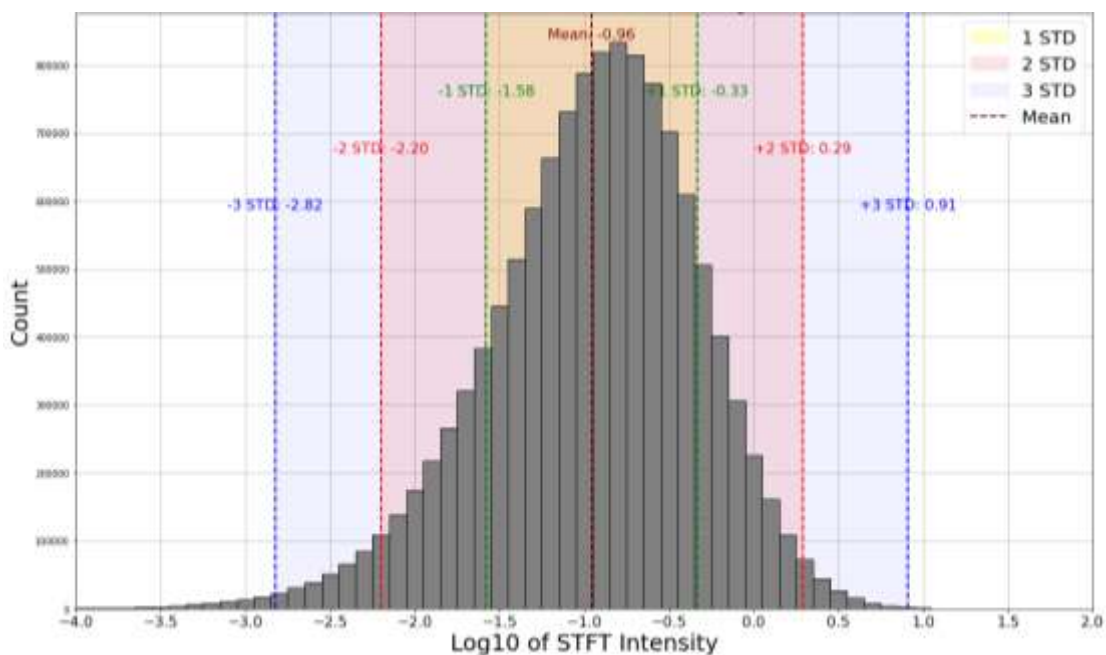


圖 3-17 4 級事件的 P 波時頻域強度分佈

3.4.3 標記模型輸出

本計畫之模型輸出分為兩種輸出，分別為標記地震相位(Earthquake Phase Picking)、標記 PGA 是否大於 25gal(Label whether PGA is greater than 25 gal)。

- 標記地震相位(Earthquake Phase Picking)

標記地震相位採用截斷高斯分佈(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，示意圖如圖 3-18，其中包含三種機率函數 $T_P(t)$ 、 $T_S(t)$ 、 $T_{other_phase}(t)$ 。

$$\begin{cases} T_P(t) = \exp\left(-\frac{(t - t_{P_arrival})^2}{2 \times 0.2^2}\right), & \text{if } t_{P_arrival} - 0.2 \leq t \leq t_{P_arrival} + 0.2 \\ T_P(t) = 0, & \text{otherwise} \end{cases} \quad (3-9)$$

$t_{P_arrival}$ 是 P 波相位。

$$\begin{cases} T_S(t) = \exp\left(-\frac{(t - t_{S_arrival})^2}{2 \times 0.3^2}\right), & \text{if } t_{S_arrival} - 0.3 \leq t \leq t_{S_arrival} + 0.3 \\ T_S(t) = 0, & \text{otherwise} \end{cases} \quad (3-10)$$

$t_{S_arrival}$ 是 S 波相位。

$$T_{other_phase}(t) = 1 - T_P(t) - T_S(t) \quad (3-11)$$

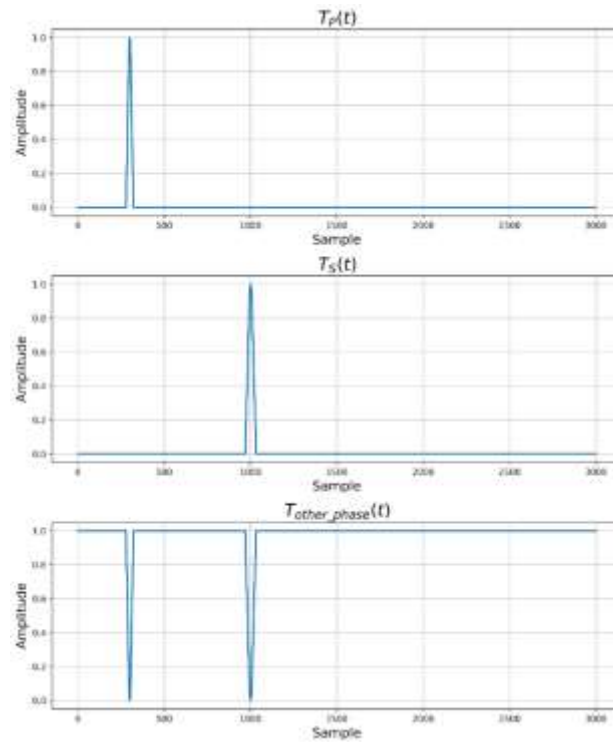


圖 3-18 標記地震相位

- 標記 PGA 是否大於 25gal(Label whether PGA is greater than 25 gal)

標記 PGA 大於 25gal 之輸出，如圖 3-19 所示，其中包含兩種機率函數 $T_{PGA_25}(t)$ 、 $T_{other_PGA}(t)$ 。

$$\begin{cases} T_{PGA_25}(t) = T_P(t), & \text{if } t_{P_arrival} - 0.2 \leq t \leq t_{P_arrival} \\ T_{PGA_25}(t) = 1, & \text{if } t_{P_arrival} \leq t \leq t_{S_arrival} \\ T_{PGA_25}(t) = T_S(t), & \text{if } t_{S_arrival} \leq t \leq t_{S_arrival} - 0.3 \\ T_{PGA_25}(t) = 0, & \text{otherwise} \end{cases} \quad (3-12)$$

$$T_{other_PGA25}(t) = 1 - T_{PGA_25}(t) \quad (3-13)$$

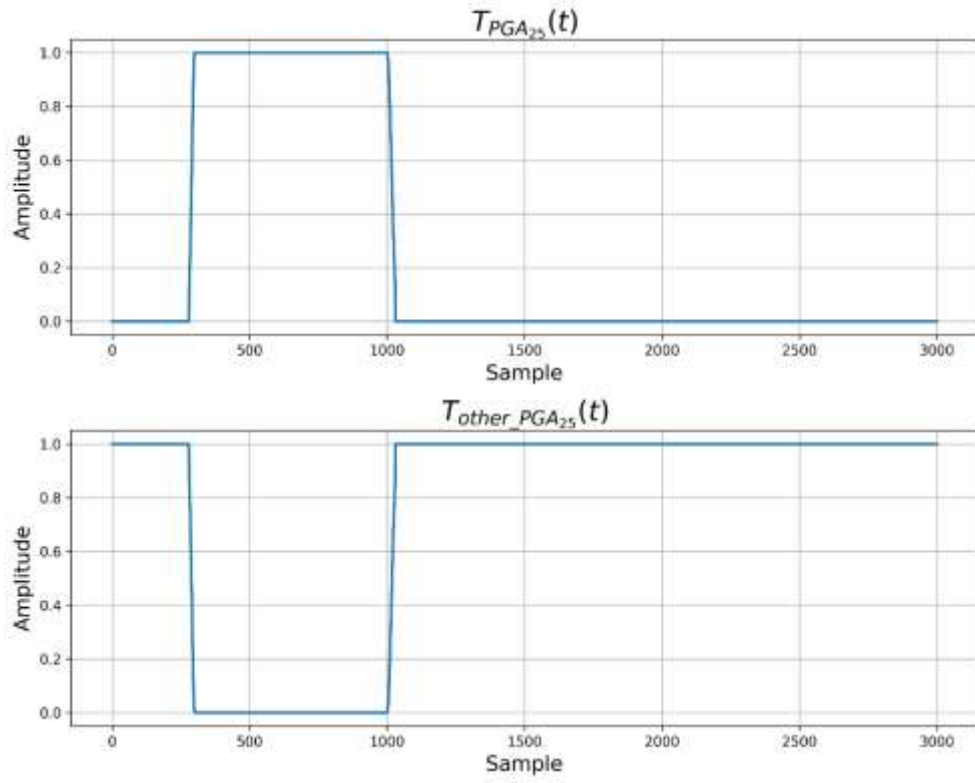


圖 3-19 標記 PGA 大於 25gal

標記 PGA 小於 25gal 之輸出，如所示，其中包含兩種機率函數 $T_{PGA_25}(t)$ 、 $T_{other_PGA}(t)$ 。

$$T_{PGA_25}(t) = 0 \quad (3-14)$$

$$T_{other_PGA}(t) = 1 \quad (3-15)$$

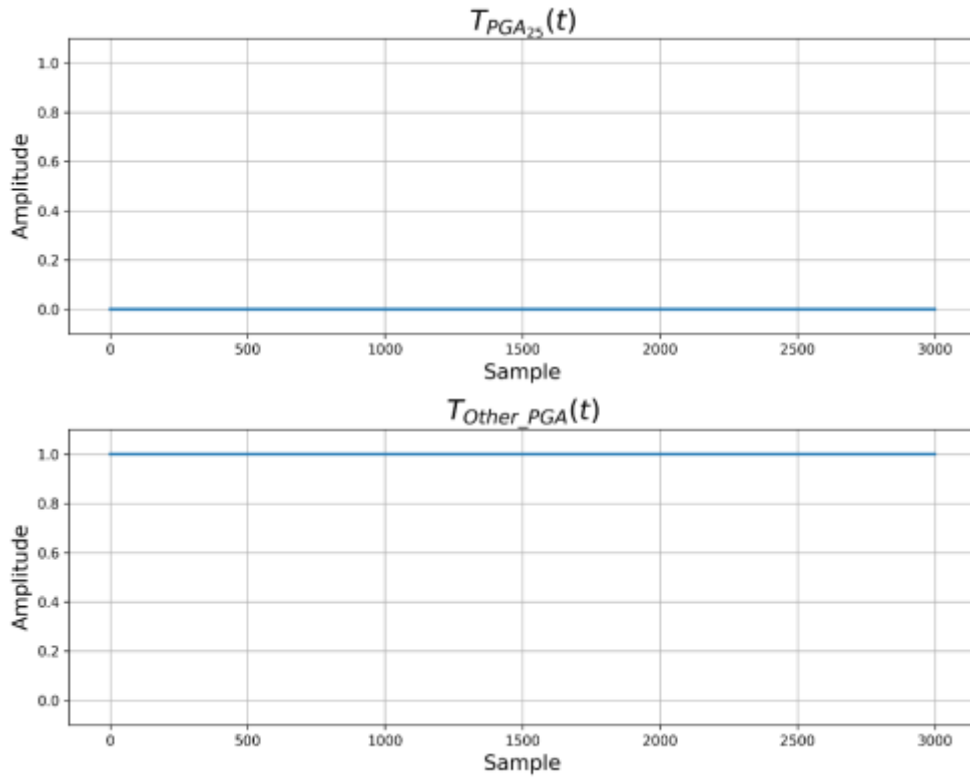


圖 3-20 標記 PGA 小於 25gal

3.5 評估模型表現

本計畫用二元混淆矩陣評估模型表現。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, 以評估模型成效, 公式如(3-16)、(3-17)、(3-18)。

$$Precision = \frac{TP}{TP + FP} \quad (3-16)$$

$$Recall = \frac{TP}{TP + FN} \quad (3-17)$$

$$F1-Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (3-18)$$

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

雖然本計畫中開發的模型具備 S 波挑選的能力, 但在實際應用中不會單獨使用, 因此不會評估 S 波挑選的性能。本研究使用模型預測 $T_p(t)$ 最大值及其出現的位置來評估模型在相位挑選中的表現。圖 3-21 顯示了相位挑選的二元混淆矩陣。本計畫有針對多個閾值 (0.05、0.1、0.15、0.2、0.25、0.3、0.35、0.4、0.45、0.5、0.55、0.6) 進行性能統計, 並選擇閾值為 0.3。當模型之 $T_p(t)$ 最大值超過設定的閾

值時，模型會將該時間點識別為 P 波相位到達時間。此外，根據 Liao 等人[2]的設置，正確的相位到達時間窗口設定為 $\text{Ground Truth} \pm 0.5$ 秒內。

對於地震波形資料，地震相位標記之二元混淆矩陣如下：

若模型標記最大 $T_p(t)$ 機率值在 $\text{Ground Truth} \pm 0.5$ 秒內且高於 0.3，則模型視為正確挑選到了相位到達時間，並在二元混淆矩陣中歸類為 True Positive(TP)。

若模型標記最大 $T_p(t)$ 機率值在 $\text{Ground Truth} \pm 0.5$ 秒內且低於 0.3，則模型被視為錯過了相位到達時間，並在二元混淆矩陣中歸類為 False Negative(FN)。

若模型標記最大 $T_p(t)$ 機率值在 $\text{Ground Truth} \pm 0.5$ 秒外且高於 0.3，則模型被視為錯誤地挑選到了相位到達時間，並在二元混淆矩陣中歸類 False Positive(FP)。

在純地震測試數據中，由於每個測試樣本都包含 Ground Truth，因此單個事件中不會存在不需要相位挑選的情況，故不存在 True Negative(TN)。

對於非地震波形資料，由於不存在可挑選的相位，因此也沒有 Ground Truth，相位挑選任務的二元混淆矩陣如下：

若模型標記最大 $T_p(t)$ 機率值高於 0.3，則模型被視為錯誤地挑選到了相位，並在二元混淆矩陣中歸類為 False Positive(FP)。

若模型標記最大 $T_p(t)$ 機率值低於 0.3，則模型被視為正確地識別出沒有相位，並在二元混淆矩陣中歸類為 True Negative(TN)。

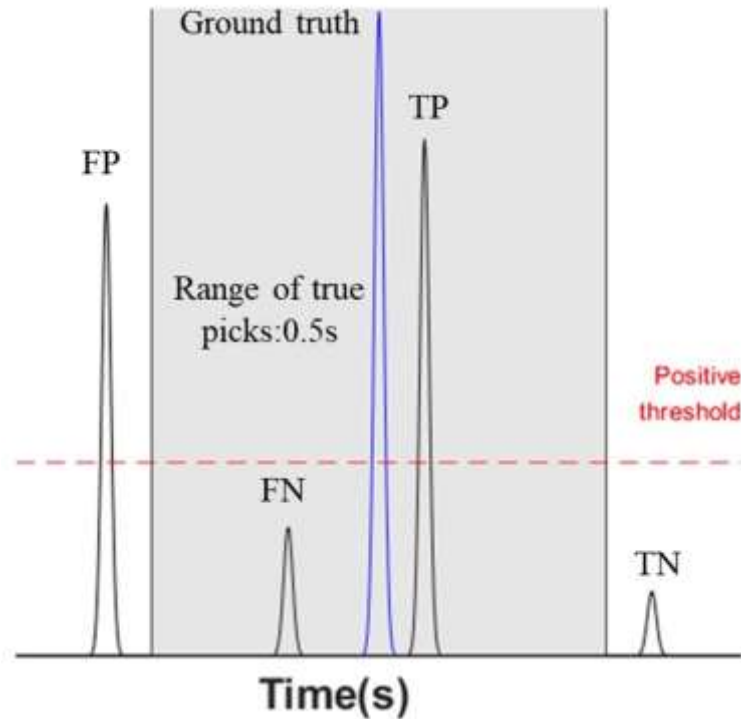


圖 3-21 地震相位標記之二元混淆矩陣

3.5.2 預測 PGA 是否大於 25gal 之二元混淆矩陣

本計畫使用模型輸出之 $T_{PGA}(t)$ 來評估模型在 PGA 預測性能。具體來說，在 P 波相位後陸續出現之 1 秒、2 秒、3 秒波形中， $T_{PGA}(t)$ 產生之最大機率值有超過閾值，即視為模型預測此測站之 PGA 大於 25gal。由於 P 波相位後 1 秒可能漏報或誤報之機率較其餘兩個秒數高，所以本計畫還有用 P 波相位後陸續出現之 2 秒、3 秒波形中， $T_{PGA}(t)$ 之最大機率值有超過閾值，即視為模型預測此測站之 PGA 大於 25gal。P 波相位是由氣象署人工標記。表 3-2 顯示 PGA 預測的二元混淆矩陣。本計畫針對多個閾值（0.45、0.50、0.55、0.60、0.65、0.70、0.75、0.80、0.85）進行性能統計，並選擇閾值為 0.6。

。表 3-2 PGA 預測是否大於 25gal 之二元混淆矩陣

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

對於地震和非地震波形，PGA 預測二元混淆矩陣如下：

如果實際測站 PGA 大於 25 gal，且 $T_{PGA}(t)$ 在 P 波相位後陸續出現之 1 秒、2 秒、3 秒波形(或 2 秒、3 秒波形)之最大值也大於閾值，則視為模型正確預測該測站 PGA 超過 25 gal，在二元混淆矩陣中歸類為 True Positive(TP)。

如果實際測站 PGA 大於 25 gal，但 $T_{PGA}(t)$ 在 P 波相位後陸續出現之 1 秒、2 秒、3 秒波形(或 2 秒、3 秒波形)之最大值小於閾值，則視為模型錯誤地預測該測站 PGA 小於 25 gal，在二元混淆矩陣中歸類為 False Negative(FN)。

如果實際測站 PGA 小於 25 gal，但 $T_{PGA}(t)$ 在 P 波相位後陸續出現之 1 秒、2 秒、3 秒波形(或 2 秒、3 秒波形)之最大值大於閾值，則視為模型錯誤地預測該測站 PGA 超過 25 gal，在二元混淆矩陣中歸類為 False Positive(FP)。

如果實際測站 PGA 小於 25 gal，且 $T_{PGA}(t)$ 在 P 波相位後陸續出現之 1 秒、2 秒、3 秒波形(或 2 秒、3 秒波形)之最大值也小於閾值，則視為模型正確預測該測站 PGA 小於 25 gal，在二元混淆矩陣中歸類為 True Negative(TN)。

3.5.3 現地型單站地震預警之二元混淆矩陣

本計畫開發之模型最終目標是利用單站方法，對目標區域或測站進行實時地震預警。因此，二元混淆矩陣將**地震相位標記(Earthquake Phase Picking)**與**預測 PGA 是否大於 25gal**兩者結果結合，以全面評估模型的性能表現。也就是說 P 波相位資訊和 PGA 預測資訊都由模型提供。所以評估模型表現為，在 P 波相位後陸續出現之 1 秒、2 秒、3 秒波形中， $T_{PGA}(t)$ 產生之最大機率值有超過閾值，即視為模型預測此測站之 PGA 大於 25gal；由於 P 波相位後 1 秒可能漏報或誤報之機率較其餘兩個秒數高，所以本計畫使用 P 波相位後陸續出現之 2 秒、3 秒波形中， $T_{PGA}(t)$ 之最大機率值超過閾值，即視為模型預測此測站之 PGA 大於 25gal。但 P 波相位是由模型提供。

發出警報的唯一標準是當 $T_P(t)$ 之最大值超過閾值，且 $T_{PGA}(t)$ 在 P 波相位後陸續出現之 1 秒、2 秒、3 秒波形(或 2 秒、3 秒波形)之最大值也超過閾值。如果任一任務或兩個任務皆未超過閾值，則模型不會發出警報。

現地型單站地震預警之二元混淆矩陣評估方式如下：

若實際測站 PGA 值大於 25 gal 且兩個任務的結果皆超過閾值，模型正確地

發出警報，此情況在二元混淆矩陣中歸類為 True Positive(TP)。

若實際測站 PGA 值大於 25 gal，但任一任務或兩個任務皆未超過閾值，模型未能發出警報，此情況在二元混淆矩陣中歸類為 False Negative(FN)。

若實際測站 PGA 值小於 25 gal，但兩個任務的結果皆超過閾值，模型錯誤地發出警報，此情況在二元混淆矩陣中歸類為 False Positive(FP)。

若實際測站 PGA 值小於 25 gal，且任一任務或兩個任務皆未超過閾值，模型正確地未發出警報，此情況在二元混淆矩陣中歸類為 True Negative(TN)。

		Ground Truth	
		PGA < 25 gal	PGA ≥ 25 gal
Predict	發報	FP	TP
	不發報	TN	FN

警報：兩個任務皆超過閾值。

不警報：任一或兩個任務皆未超過閾值。

3.6 短時傅立葉變換參數分析和不同時頻輸入架構特定區間對數正規化

3.6.1 短時傅立葉變換參數分析

由於本計畫採用短時傅立葉變換 (STFT)，因此需要針對 STFT 設定多個重要參數，例如所使用的窗函數、窗函數長度、步進、以及維持時間域長度所需的零填充長度及位置等。這些參數將影響頻率解析度與時間解析度，進而影響模型辨識特徵的能力。因此，本小節測試之模型架構如圖 3-22，來調整相關參數，訓練出表現較佳的模型，並確定 STFT 參數組合。

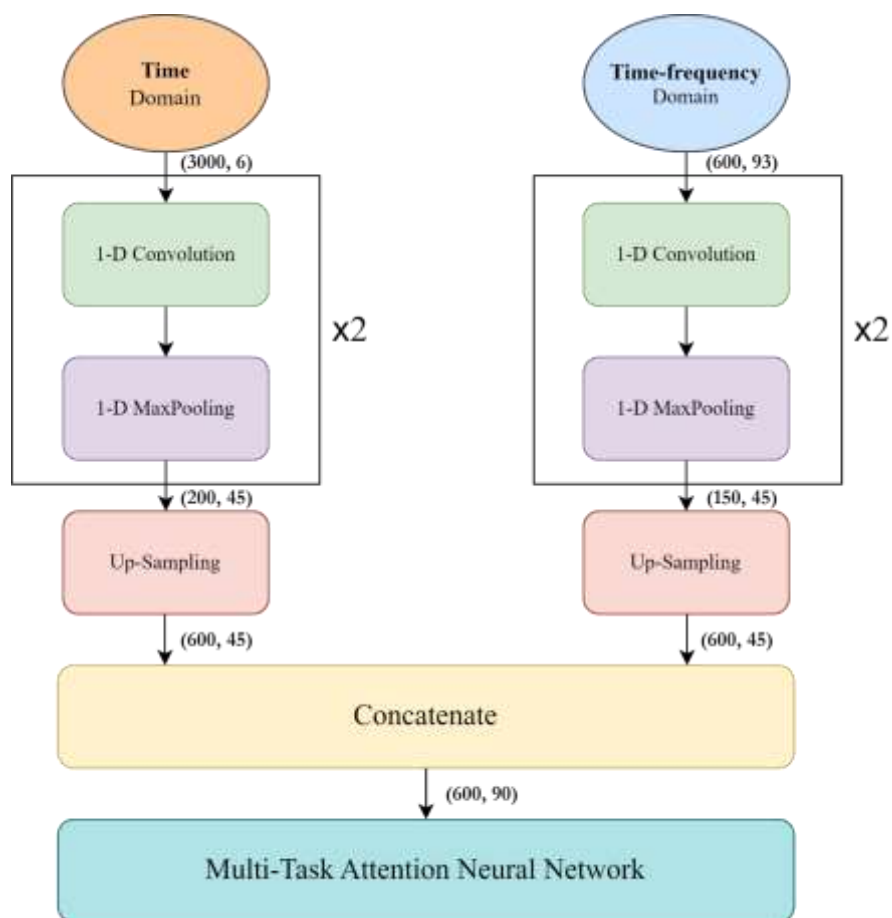


圖 3-22 STFT 參數分析之模型架構

本計畫採用 Hamming 窗函數與矩形 (Boxcar) 窗函數作為窗函數選項，以比較哪一種窗函數更適合模型。在本計畫中，時頻域的輸入有 3000 點的時間長度，每一點代表 0.01 秒，因此取樣頻率為 100 Hz。使用 150 點的窗長、1 點的步進，並填充 149 點的零值，以提取低於 20 Hz 的資訊。結合三軸後的時間-頻率域輸入尺寸將為 3000×93。

然而，初步的訓練嘗試顯示，使用 3000 點的時頻域輸入進行訓練過於耗時。因此，實驗將使用另一種縮小比例的輸入方式，以加速實驗並找出較佳的時間窗和零填充組合。窗長為 150 點，步長為 5 點，並填充 145 點的零值，僅提取低於 20 Hz 的資訊。結合三軸後的時頻域輸入尺寸將為 600×93。

時頻長度為 600 之輸入需要填充 145 個零值。本實驗將嘗試四種填充組合，將 145 點的零值分別填充至數據的起始(B)與結尾(E)，組合方式如下 (B/E)：(70/75)、(95/50)、(120/25) 以及 (140/0)。對應原始長度輸入所需填充的 149 個零值將分別以 (72/77)、(98/51)、(123/26) 及 (149/0) 的方式填充。這些組合將分別搭配兩種窗函數進行測試：Hamming 窗與 Boxcar 窗函數。

不同零填充組合在地震相位標記(Earthquake Phase Picking)中的結果，如表 3-3。各種類別中表現最佳的指標以粗體和底線標示。此外，計算模型標記 P 波相位與

Ground Truth 之間的差異，得到平均值和平均絕對誤差 (MAE)。根據結果可以觀察到，在各種零填充組合中，Boxcar 窗函數在 F1 分數上穩定優於 Hamming 窗函數。起始或結尾的零填充數量與模型的相位選取能力之間沒有明確的直接關聯。然而，值得注意的是，儘管在 Recall 方面 (95/50) 的 Boxcar 窗函數組合表現最佳，但在大多數其他情況下，(70/75) 的 Boxcar 窗函數組合提供最佳的整體表現。

表 3-3 不同零填充組合在地震相位標記中的結果

Pad-zero	Window	F1-score	Precision	Recall	Mean(s)	MAE(s)
145/0	Hamming	0.8812	0.8382	0.9288	-0.31	0.46
	Boxcar	0.9164	0.9058	0.9273	-0.15	0.25
120/25	Hamming	0.8792	0.8736	0.8849	-0.17	0.28
	Boxcar	0.9123	0.9045	0.9202	-0.17	0.28
95/50	Hamming	0.8917	0.9008	0.8829	-0.16	0.25
	Boxcar	0.9000	0.8709	<u>0.9312</u>	-0.23	0.35
70/75	Hamming	0.8753	0.8943	0.8571	-0.16	0.3
	Boxcar	<u>0.9225</u>	<u>0.9213</u>	0.9237	<u>-0.14</u>	<u>0.24</u>

不同零填充組合在預測 PGA 是否大於 25gal 中的結果，如表 3-4、表 3-5。表 3-4 顯示了 P 波相位後 1 秒、2 秒、3 秒內的整體表現，而表 3-5 則顯示了 P 波到達後 2 至 3 秒之間的表現。每個類別中表現最佳的指標以粗體和底線標示。

根據結果可以觀察到，Boxcar 窗函數在 F1 分數方面通常優於 Hamming 窗函數，除了在 23 測試集中的 (95/50) 零填充組合。針對不同的零填充組合，前後填充零值的數量與模型預測 PGA 大小的能力之間沒有明確的直接關聯。值得注意的是，在 P 波到達後 2 至 3 秒測試集中，(95/50) 的 Boxcar 窗函數組合在 Precision 上達到最佳表現，而 (145/0) 的 Boxcar 窗函數組合在 Recall 上達到最佳表現。在大多數其他情況下，(70/75) 的 Boxcar 窗函數組合提供了最佳的整體表現。

表 3-4 不同零填充組合在預測 PGA 是否大於 25gal 中的結果(P 波相位後 1、2、3 秒)

Pad-zero	Window	F1-score	Precision	Recall
145/0	Hamming	0.6528	0.6349	0.6717
	Boxcar	0.7082	0.6994	0.7174
120/25	Hamming	0.6493	0.6736	0.6268

	Boxcar	0.6667	0.6191	0.7221
95/50	Hamming	0.6805	0.6539	0.7093
	Boxcar	0.7003	0.7037	0.6969
70/75	Hamming	0.6778	0.6762	0.6794
	Boxcar	<u>0.7241</u>	<u>0.7196</u>	<u>0.7287</u>

表 3-5 不同零填充組合在預測 PGA 是否大於 25gal 中的結果(P 波相位後 2、3 秒)

Pad-zero	Window	F1-score	Precision	Recall
145/0	Hamming	0.7099	0.6475	0.7855
	Boxcar	0.7486	0.7157	<u>0.7847</u>
120/25	Hamming	0.7033	0.6494	0.7668
	Boxcar	0.7177	0.6897	0.7481
95/50	Hamming	0.7400	0.7153	0.7664
	Boxcar	0.7399	<u>0.7629</u>	0.7182
70/75	Hamming	0.7325	0.7053	0.7618
	Boxcar	<u>0.7604</u>	0.7557	0.7652

根據地震相位標記和預測 PGA 是否大於 25gal 結果，在各項比較指標中，起始填充 70 個零值、結尾填充 75 個零值，並搭配 Boxcar 窗函數的組合表現最佳。時頻域長度為 600 對應原始長度之零填充組合為起始填充 72 個零值、結尾填充 77 個零值，所以本計畫選擇 Boxcar 窗函、150 點的窗長、1 點的步進，並填充 149 點的零值，在加速度訊號前填充 72 個零值、後填充 77 個零值。

3.6.2 測試不同時頻輸入架構和決定特定區間對數正規化

本計畫測試兩種輸入架構，如圖 3-23、圖 3-24，並嘗試在時頻域輸入中加入額外的區間正規化。輸入為時頻域長度為 600。在模型的時頻域輸入部分，除了對 STFT 後獲得的頻譜圖強度進行全範圍對數正規化外，本實驗還嘗試加入特定區間的對數正規化。所以時頻域輸入大小會變為兩倍，所以時頻域輸入矩陣為 600*186。特定區間的選擇基於對 4 級事件 P 波進行 STFT 並取底數為 10 的對數後的頻譜圖強度，進行分佈統計。選擇 4 級事件原因是本計畫之模型是針對 PGA 大於 25gal 發報。圖 3-25 所示之直方圖顯示該分佈近似常態分佈。選取 ± 1 、 ± 2 和 ± 3 個標準差的區間進行進一步實驗。

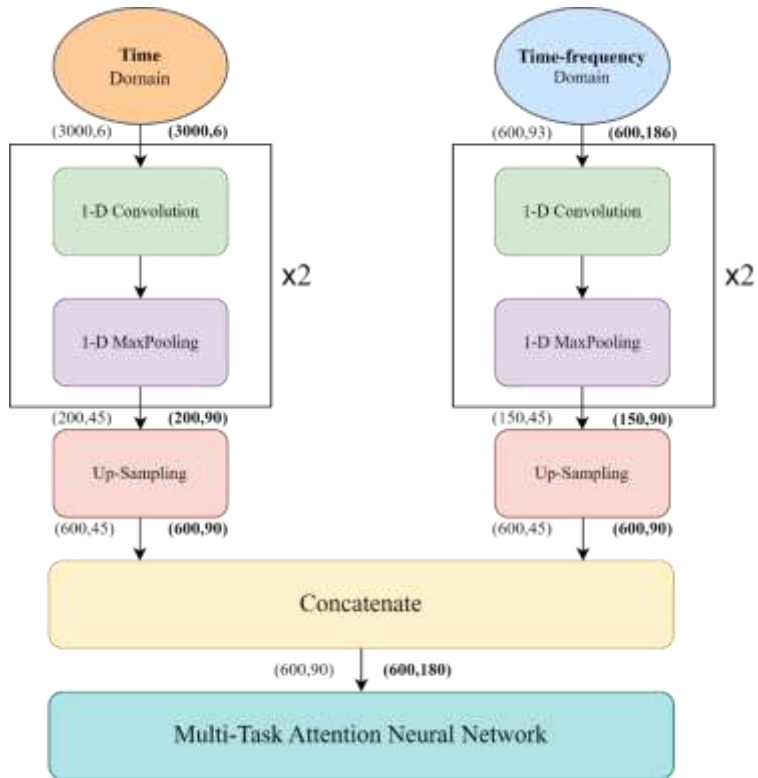


圖 3-23 1D Conv Encoder 架構

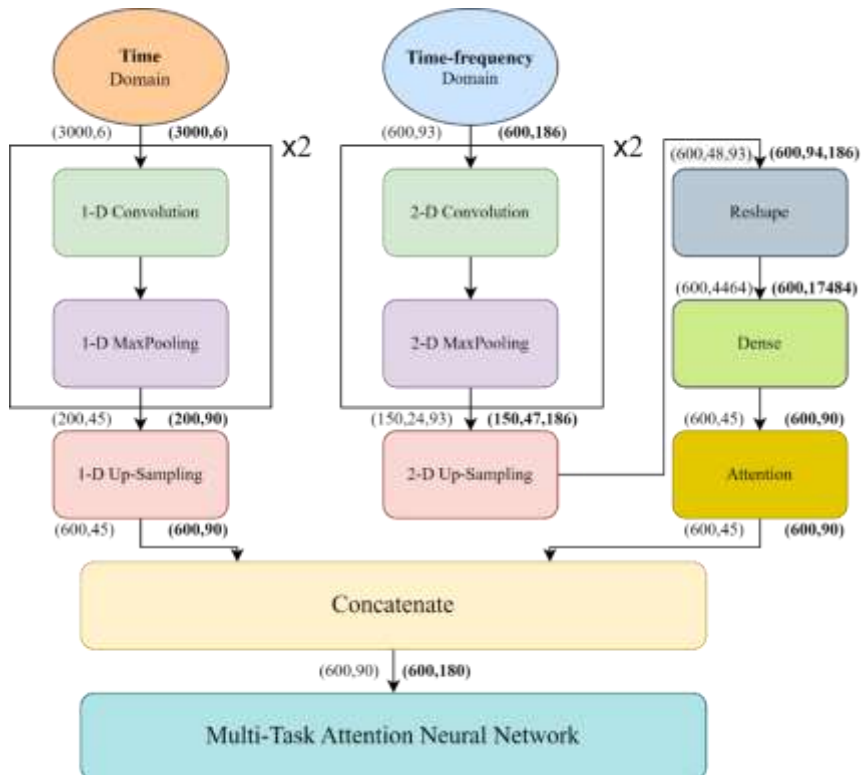


圖 3-24 Hybrid Conv-Attention Encoder 架構

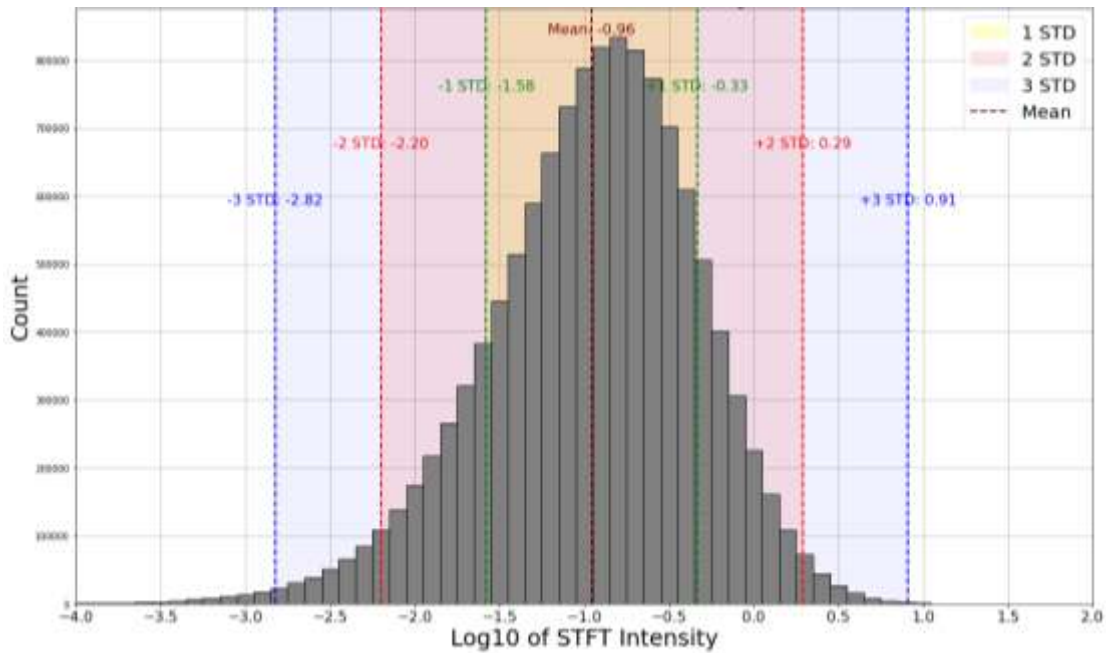


圖 3-25 4 級事件 P 波時頻域強度的分佈

不同輸入在地震相位標記中的結果，如表 3-6。顯示了不同輸入和架構在地震相位標記中的性能表現。每個類別中表現最佳的指標以粗體和底線標示。根據結果顯示，全範圍 (FR) 對數正規化結合 Hybrid Conv-Attention Encoder 架構在綜合評估中獲得最高的 F1 分數，且擁有最低的誤差率。在平均值和 MAE 方面，無論使用何種架構，使用全範圍對數正規化作為時頻域輸入的情況下，與 ground truth 的時間差異最小。

表 3-6 不同輸入和架構組合在地震相位標記中的性能表現

1D Conv. Encoder						
Combinations	F1-score	Precision	Recall	Mean(s)	MAE(s)	Error rate
FR	95.36	94.19	96.57	<u>-0.14</u>	<u>0.21</u>	0.062
FR±1STD	95.12	93.42	97.58	-0.21	0.28	0.077
FR±2STD	94.64	92.82	<u>97.94</u>	-0.19	0.25	0.090
FR±3STD	94.63	94.09	96.57	-0.22	0.28	0.060
Hybrid Conv-Attention Encoder						
FR	<u>95.83</u>	<u>94.91</u>	96.82	-0.17	0.22	<u>0.055</u>
FR±1STD	92.91	92.77	97.51	-0.34	0.39	0.084
FR±2STD	93.85	93.91	96.55	-0.34	0.39	0.066

FR$\pm$3STD	94.04	94.42	97.75	-0.24	0.29	0.081
-------------------------------	-------	-------	-------	-------	------	-------

不同輸入在預測 PGA 是否大於 25gal 中的結果如表 3-7 和表 3-8 所示。表 3-7 顯示了 P 波相位後 1、2、3 秒內的整體表現，而表 3-8 則顯示了 P 波相位後 2、3 秒之間的表現。每個類別中表現最佳的指標以粗體和底線標示。根據結果顯示，無論數據集的持續時間為何，全範圍和 ± 2 標準差 (FR $\pm$ 2SD) 的組合在 F1 分數上均穩定地達到最佳表現。雖然在第一秒中出現了較多的 false negatives，但在第二和第三秒中表現出顯著改善。

表 3-7 不同輸入和架構在預測 PGA 是否大於 25gal 中的結果(P 波相位後 1、2、3 秒)

	1D Conv. Encoder			Hybrid Conv-Attention Encoder		
Combinations	F1-score	Precision	Recall	F1-score	Precision	Recall
FR	73.69	71.96	75.51	74.57	75.30	73.84
FR+ $\pm$1STD	74.87	74.59	75.15	75.06	77.66	72.62
FR+ $\pm$2STD	<u>75.93</u>	<u>78.13</u>	73.84	74.73	72.25	77.39
FR+ $\pm$3STD	75.72	72.22	<u>79.58</u>	74.57	73.79	75.37

表 3-8 不同輸入和架構在預測 PGA 是否大於 25gal 中的結果(P 波相位後 2、3 秒)

	1D Conv. Encoder			Hybrid Conv-Attention Encoder		
Combinations	F1-score	Precision	Recall	F1-score	Precision	Recall
FR	75.57	74.85	76.31	76.89	77.39	76.39
FR+ $\pm$1STD	77.55	75.34	79.88	77.82	78.21	77.43
FR+ $\pm$2STD	<u>78.75</u>	<u>78.79</u>	78.72	77.50	76.64	78.39
FR+ $\pm$3STD	77.86	75.05	80.88	76.92	72.92	<u>81.38</u>

不同輸入在現地型單站地震預警中的結果如表 3-9 和表 3-10 所示。每個類別中表現最佳的指標以粗體和底線標示。

根據結果顯示，作為 現地型單站地震預警，全範圍和 ± 2 標準差 (FR + $\pm$ 2STD) 的輸入搭配 1D Conv 編碼器架構，在測試數據集中不同秒數範圍內均穩定地達到最佳 F1 分數。此組合將作為最終模型的參考配置。

表 3-9 不同輸入和架構在現地型單站地震預警中的結果(P 波相位後 1、2、3 秒)

	1D Conv. Encoder			Hybrid Conv-Attention Encoder		
Combinations	F1-score	Precision	Recall	F1-score	Precision	Recall
FR	73.78	72.20	75.42	74.84	75.94	73.77
FR+ ±1STD	74.92	74.72	75.13	75.21	78.01	72.60
FR+ ±2STD	<u>76.16</u>	<u>78.69</u>	73.79	74.83	72.44	77.38
FR+ ±3STD	75.82	72.41	<u>79.56</u>	74.66	73.99	75.35

表 3-10 不同輸入和架構在現地型單站地震預警中的結果(P 波相位後 2、3 秒)

	1D Conv. Encoder			Hybrid Conv-Attention Encoder		
Combinations	F1-score	Precision	Recall	F1-score	Precision	Recall
FR	75.71	72.62	79.08	77.20	78.10	76.31
FR+ ±1STD	77.61	75.47	79.87	77.98	78.56	77.41
FR+ ±2STD	<u>78.99</u>	<u>79.30</u>	78.68	77.64	76.91	78.38
FR+ ±3STD	77.98	75.29	80.87	77.09	73.25	<u>81.36</u>

3.7 DAPPER 各任務門檻值設定和與特定任務模型比較

3.7.1 DAPPER 各任務門檻值設定

本節將挑選 DAPPER 在地震相位標記、預測 PGA 是否超過 25 gal 和現地型地震預警之閾值，並以 F1 分數來確定最適合的閾值。

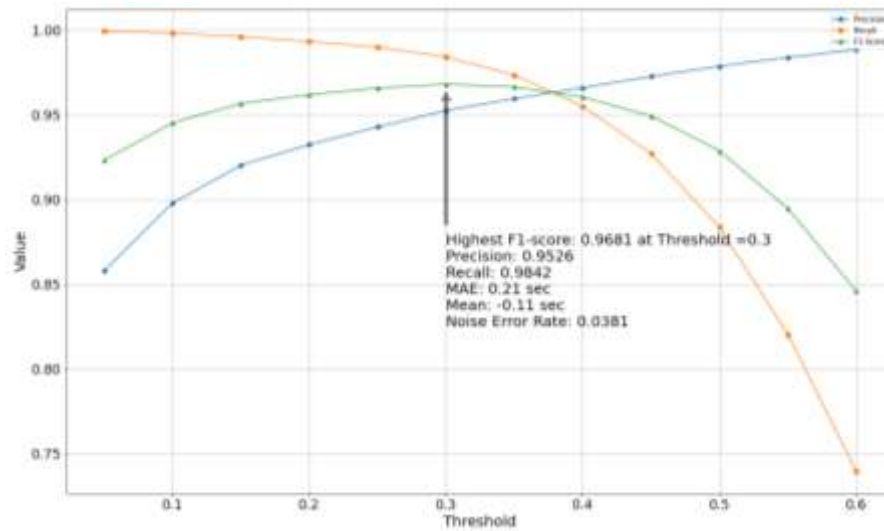


圖 3-26 DAPPER 在不同閾值下地震相位表現

圖 3-26 顯示了 DAPPER 在不同閾值下的地震相位選取表現，包括精確率 (Precision)、召回率 (Recall) 和 F1 分數等指標。最佳表現在閾值設定為 0.3 時達成，此時精確率為 95.26%、召回率為 98.42%、F1 分數為 96.81%、MAE 為 0.21 秒、平均值(Mean)為 -0.11 秒，且雜訊錯誤率(Noise Error Rate)為 0.038%。

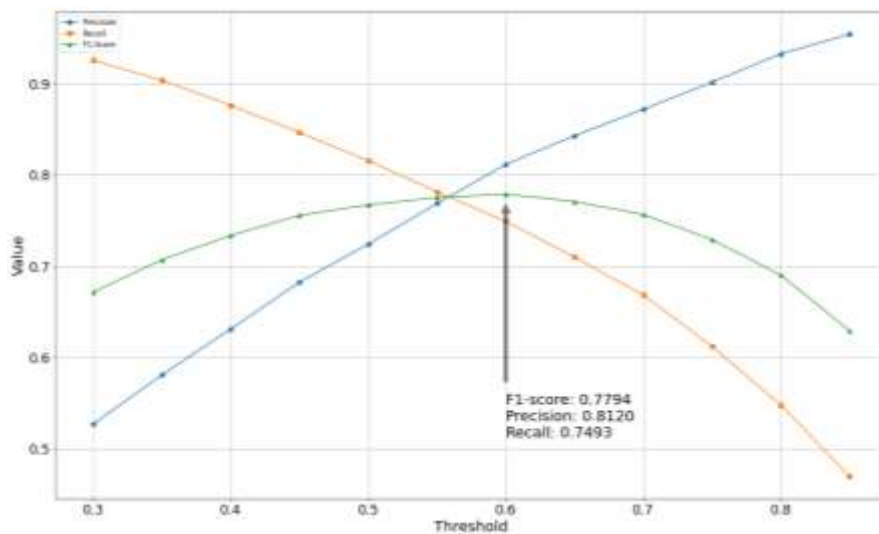


圖 3-27 DAPPER 在不同閾值下預測 PGA 是否大於 25gal 表現(在 P 波相位後 1 秒、2 秒、3 秒)

圖 3-27 顯示了 DAPPER 在 P 波相位後 1 秒、2 秒、3 秒測試集上針對 PGA 預測在不同閾值下的表現。最佳表現在閾值為 0.6 時，此時精確率為 0.81、召回率為 0.74，且 F1 分數為 0.77。

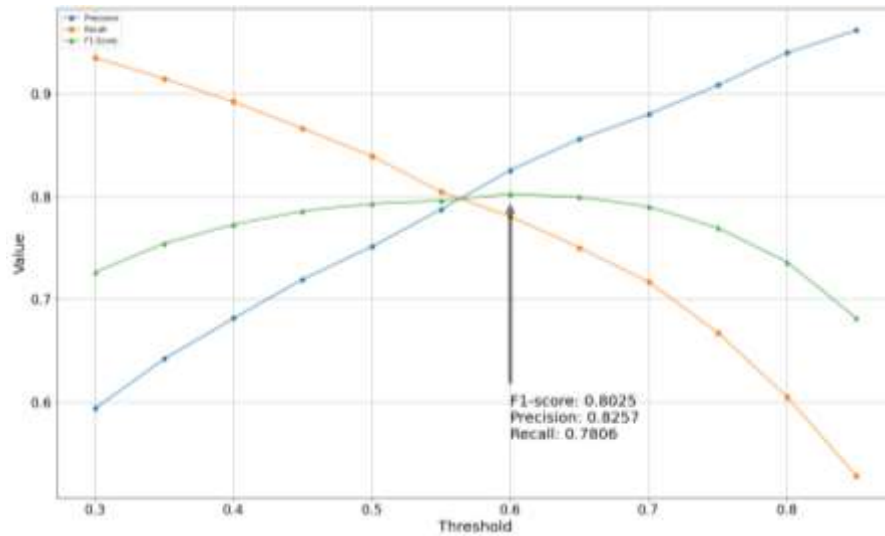


圖 3-28 DAPPER 在不同閾值下預測 PGA 是否大於 25gal 表現(在 P 波相位後 2 秒、3 秒)

圖 3-28 顯示了 DAPPER 在 P 波相位後 2 秒、3 秒上針對 PGA 預測在不同閾值下的表現。最佳表現在閾值設定為 0.6 時，此時精確率為 0.82、召回率為 0.78，且 F1 分數為 0.80。

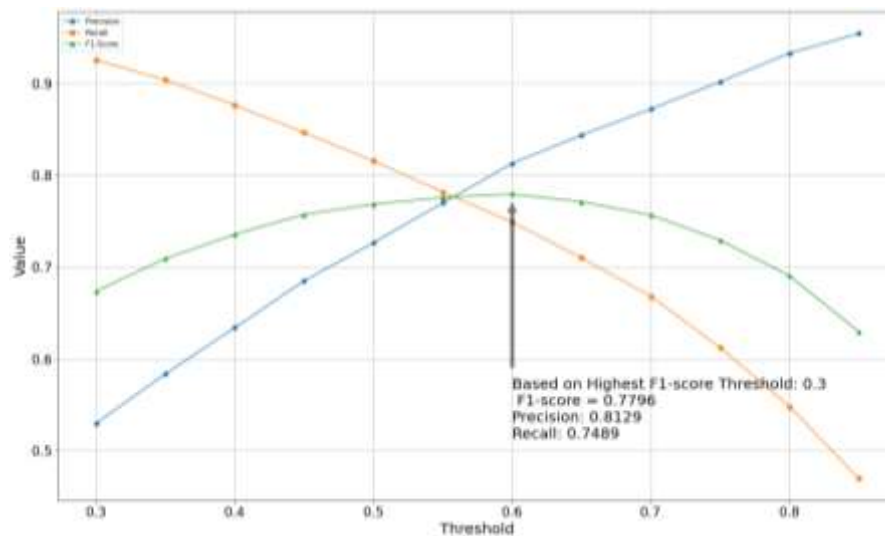


圖 3-29 DAPPER 在不同閾值下現地型單站預警表現(在 P 波相位後 1 秒、2 秒、3 秒)

圖 3-29 顯示了 DAPPER 作為單站即時地震預警系統 (EEWS) 在 P 波相位後 1 秒、2 秒、3 秒上的表現。地震相位選取閾值設定為 0.3，並在不同 PGA 預測閾值下評估精確率、召回率和 F1 分數。最佳表現在閾值設定為 0.6 時，此時精確率為 0.8129、召回率為 0.7489，且 F1 分數為 0.7796。

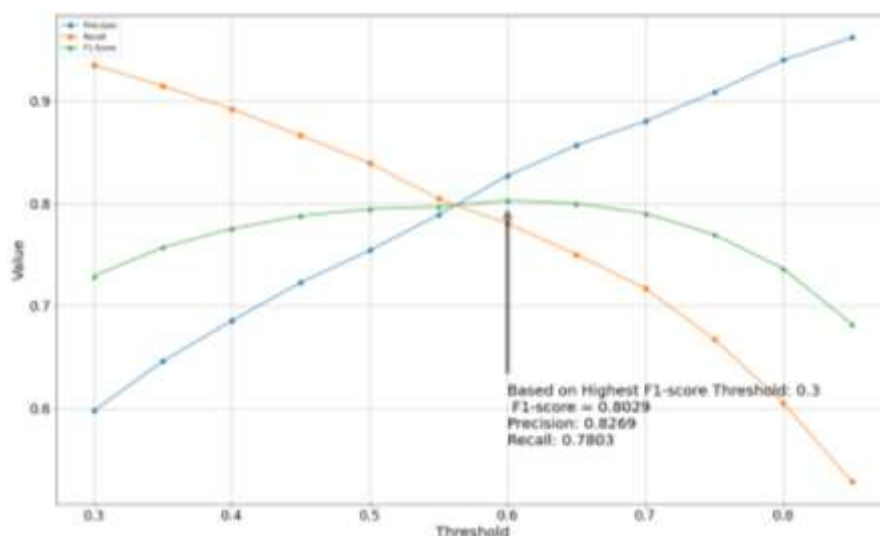


圖 3-30 DAPPER 在不同閾值下現地型單站預警表現(在 P 波相位後 2 秒、3 秒)

圖 3-30 顯示了 DAPPER 作為單站即時地震預警系統 (EEWS) 在在 P 波相位後 2 秒、3 秒的表現。地震相位選取閾值設定為 0.3。最佳表現在閾值設定為 0.6 時達成，此時精確率為 0.82、召回率為 0.78，且 F1 分數為 0.80。

3.7.2 與特定任務模型比較

地震相位標記的比較將使用 Liao 等人 [2] 提供的數據集。該數據集包含 2019 年 CWSN 的 122,356 條波形資料，每條持續時間為 30 秒。在 30 秒的信號中包含了 P 波和 S 波的到達時間，取樣頻率為 100Hz。性能將與兩個模型進行比較：RED-PAN [2] 和 Z-score&2.5gal [3]。

表 3-11 顯示了比較結果，其中粗體表示 DAPPER 的表現，底線表示每個類別中的最佳表現。DAPPER 模型在召回率 (Recall) 上達到最佳，但與其他兩個模型相比，其精確率 (Precision) 較低，這表明當前的 DAPPER 在相位選取上與真實值之間的差異較大。這一現象在平均值 (Mean) 和平均絕對誤差 (MAE) 的統計數據中也顯而易見，DAPPER 與 ground truth 的偏差比其他模型更大。

表 3-11 與不同地震相位標記之模型比較

Model	F1-score(%)	Precision(%)	Recall(%)	Mean(s)	MAE(s)
RED-PAN	98.89	98.77	98.83	0.03	0.033
Z-score&2.5gal	97.28	97.16	97.22	0.05	0.088
DAPPER	96.07	93.65	99.84	-0.13	0.24

PGA 預測的比較將使用 施柏鈞 [3] 提供的數據集，該數據集包含來自 TSMIP 的 4,056 條波形記錄。每條波形記錄包括 P 波到達後的 2 秒和 3 秒區間，總長度為 30 秒，取樣率為 100Hz。性能將與另外兩個模型進行比較：基於 CNN 的模型和 Z-score & 2.5gal。

表 3-12 和表 3-13 分別顯示了在 P 波相位後 1 秒、2 秒、3 秒和在 P 波相位後 2 秒、3 秒的結果。粗體數值代表 DAPPER 模型的表現，底線數值表示每個類別中的最佳表現。DAPPER 模型在 F1 分數和召回率 (Recall) 上達到最佳表現，並且相比 Z-score & 2.5gal 模型，增加頻率域資訊改善了低召回率的問題。

表 3-12 跟不同預測 PGA 模型比較(在 P 波相位後 1 秒、2 秒、3 秒)

Model	F1-score	Precision	Recall
CNN-based	78.27	78.24	78.30
Z-score&2.5gal	16.13	<u>97.42</u>	8.82
DAPPER	<u>83.65</u>	84.61	<u>82.71</u>

表 3-13 跟不同預測 PGA 模型比較(在 P 波相位後 2 秒、3 秒)

Model	F1-score	Precision	Recall
CNN-based	77.50	<u>82.74</u>	72.90
Z-score&2.5gal	55.84	77.98	43.49
DAPPER	<u>81.69</u>	81.99	<u>81.38</u>

3.8 20240403 花蓮事件之事後分析

本章將以 2024 年 4 月 3 日發生的花蓮地震作為案例研究。將使用 CWASN 網絡中 140 個站點的波形數據，模擬模型在 P 波到達後 1、2、3、5、7 和 10 秒的預測結果。該事件發生於 2024 年 4 月 3 日上午 7:58:09.8 (GMT+8)。震央位於北緯 23.86 度、東經 121.58 度，深度為 22.5 公里，規模為 7.2 ML。根據中央氣象署的地震報告，臺灣本島所有縣市的震度均達到 4 級以上。因此，臺灣本島所有縣市應接收到警報資訊。

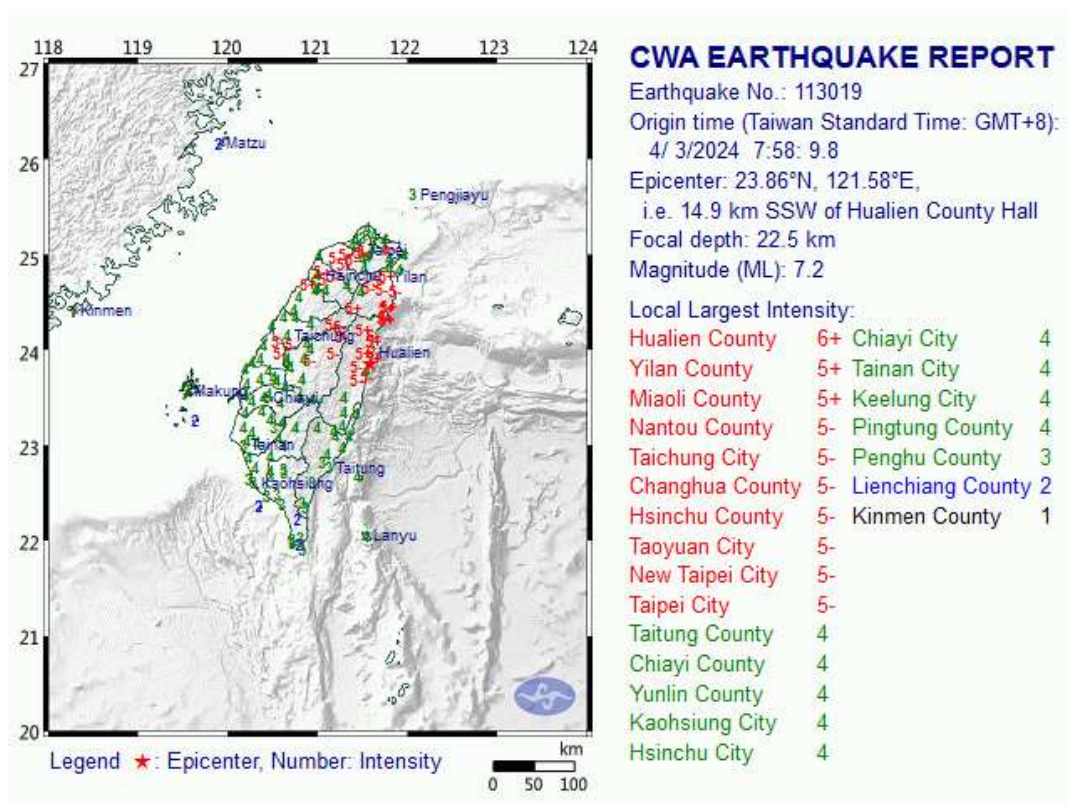


圖 3-31 20240403 花蓮地震事件報告

表 3-14 顯示了 DAPPER 在 0403 地震事件中相位選取的表現。F1 分數為 87.05，精確率 (Precision) 為 89.36，召回率 (Recall) 為 84.85，平均值 (Mean) 為 -0.24 秒，平均絕對誤差 (MAE) 為 0.54 秒。

表 3-14 DAPPER 在 0403 事件之地震相位標記表現

DAPPER				
F1-score	Precision	Recall	Mean(s)	MAE(s)
87.05	89.36	84.85	-0.24	0.54

以 CHK 站為例，其 PGA 為 111.41 gal。圖 3-32、圖 3-33 和圖 3-34 分別展示 P 波到達後 1 秒、2 秒和 7 秒的預測結果。根據第一秒的結果，DAPPER 正確地選取了相位，僅與真實值相差 0.05 秒。然而，在第二秒時，DAPPER 選取的相位出現了延遲，而在第七秒之後，機率值不再超過閾值。這表示 DAPPER 目前傾向於在時間窗口的右側選取相位，當 P 波相位到時間窗口中間就不容易標記 P 波相位。

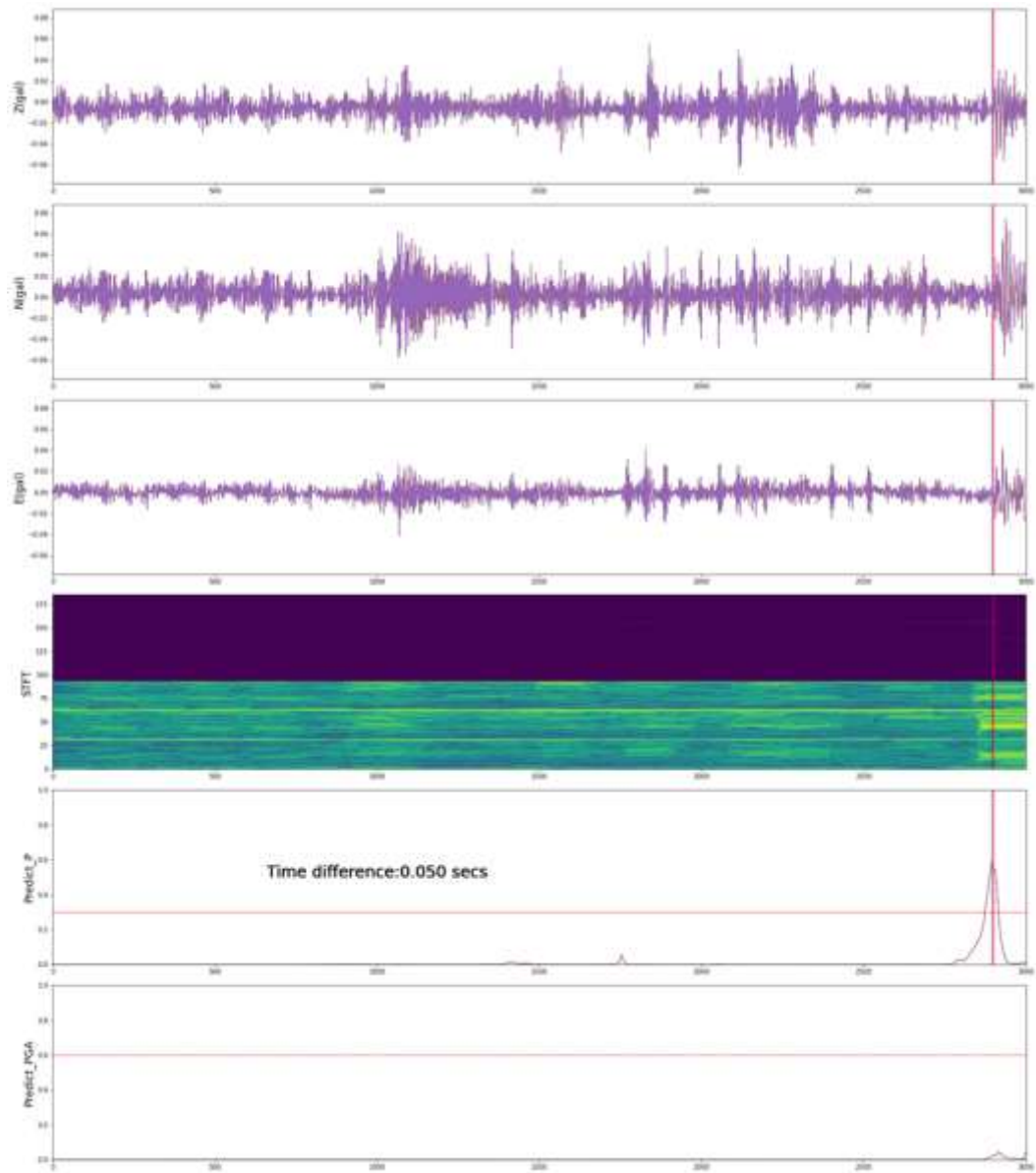


圖 3-32 DAPPER 標記測站 CHK 在 P 波相位後 1 秒之情況

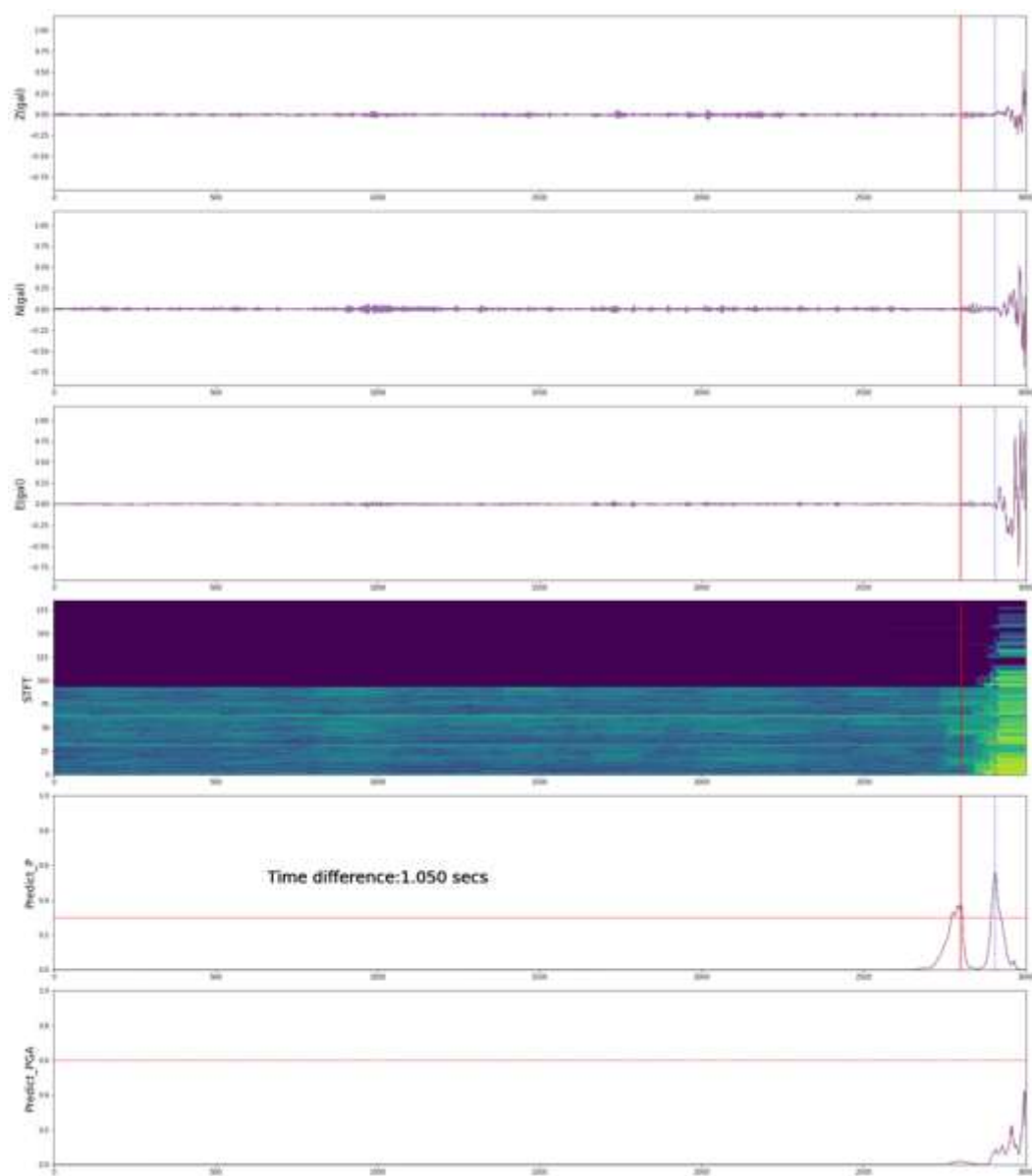


圖 3-33 DAPPER 標記測站 CHK 在 P 波相位後 2 秒之情況

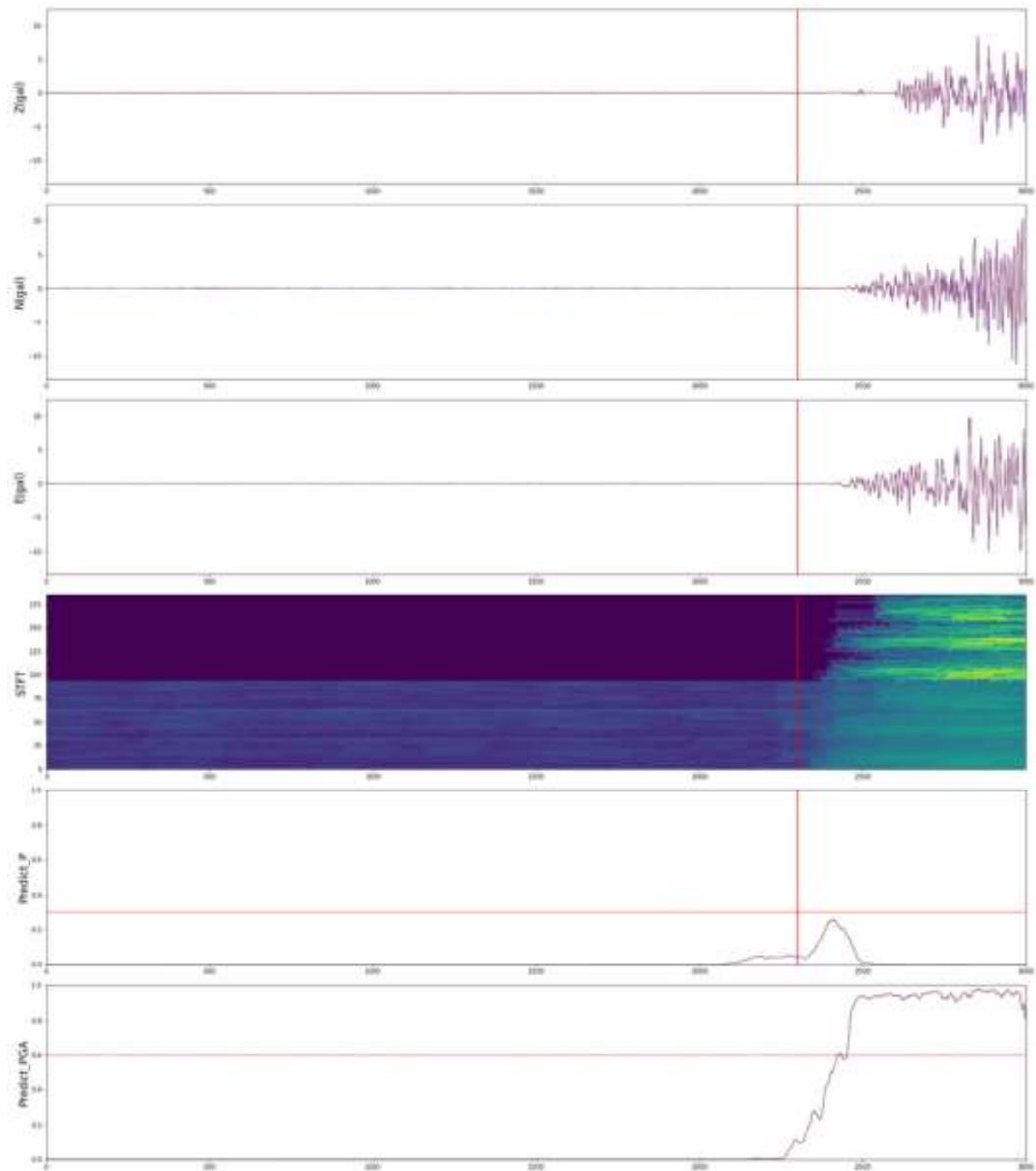


圖 3-34 DAPPER 標記測站 CHK 在 P 波相位後 7 秒之情況

再以 NOU 站為例，其 PGA 為 51.28 gal。圖 3-35、圖 3-36、圖 3-37 和圖 3-38 分別展示了 P 波到達後 1 秒、2 秒、5 秒和 7 秒的預測結果。在第一秒中，儘管模型顯示機率值上升，但機率值沒有超過閾值。在第二秒中，雖然模型選取了相位，但與 ground truth 相比存在明顯的時間差異。到第五秒時，概率值降至閾值以下，而在第七秒時進一步下降，這再次顯示 DAPPER 傾向於在時間窗口的右側標記 P 波相位。這可能是因為訓練數據中有相當一部分的 P 波相位於時間窗口的右側，導致模型在學習後傾向於關注時間窗口的右側。雖然這對於實時選取有幫助，但這也意味著如果模型未能在 P 波到達後的前幾秒內標記相位，後續被標記到可能性將降低。

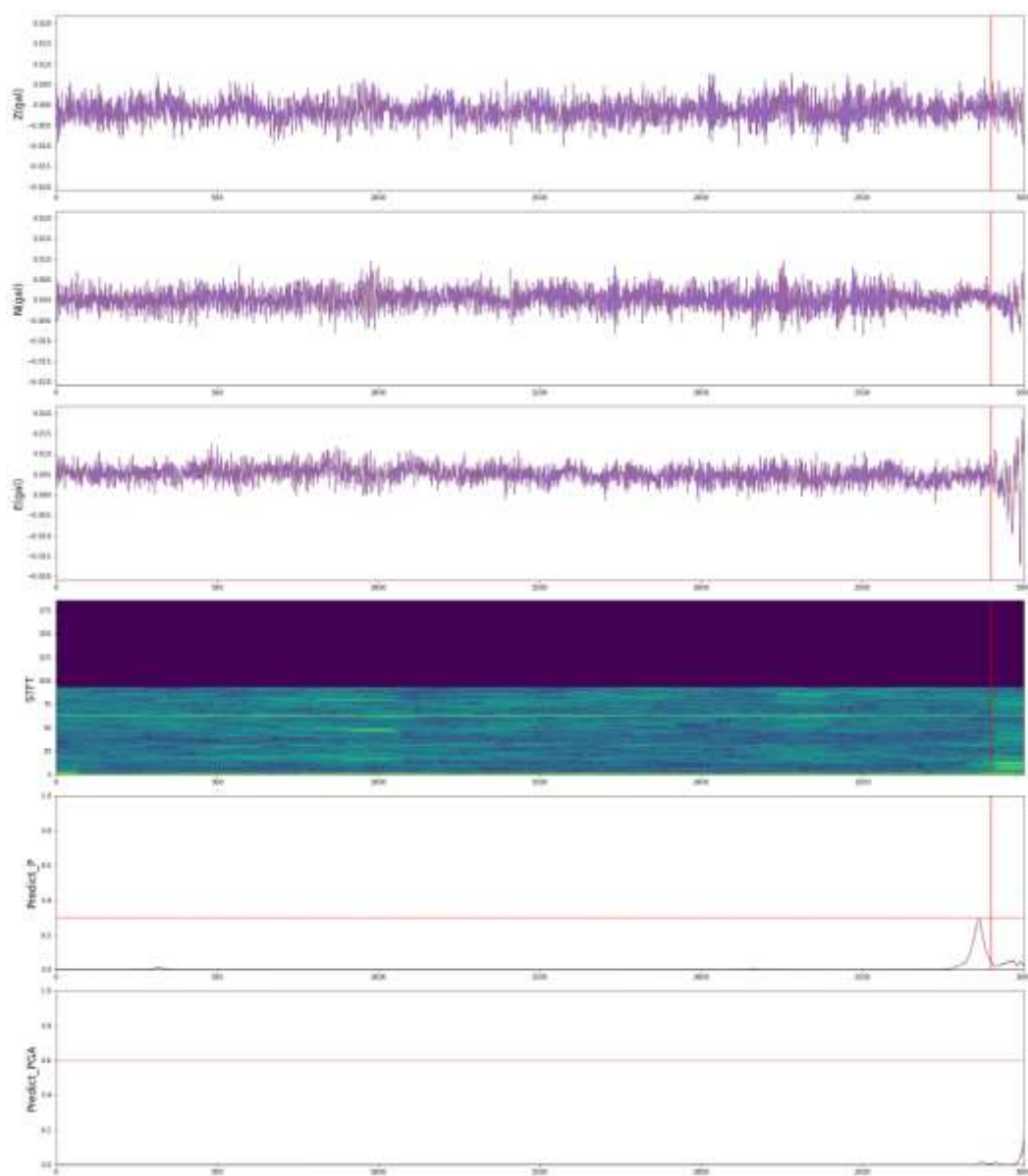


圖 3-35 DAPPER 標記測站 NOU 在 P 波相位後 1 秒之情況

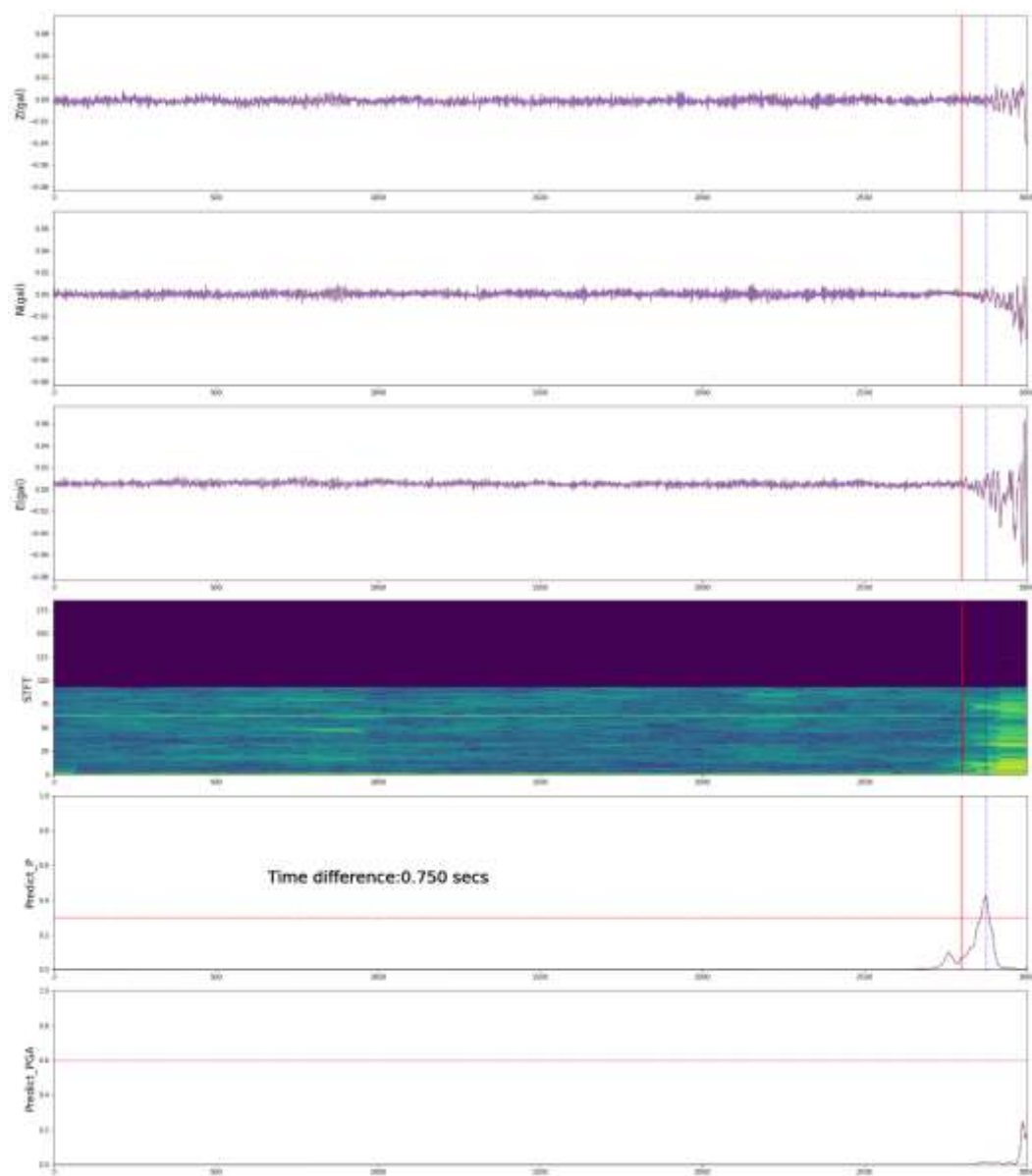


圖 3-36 DAPPER 標記測站 NOU 在 P 波相位後 2 秒之情況

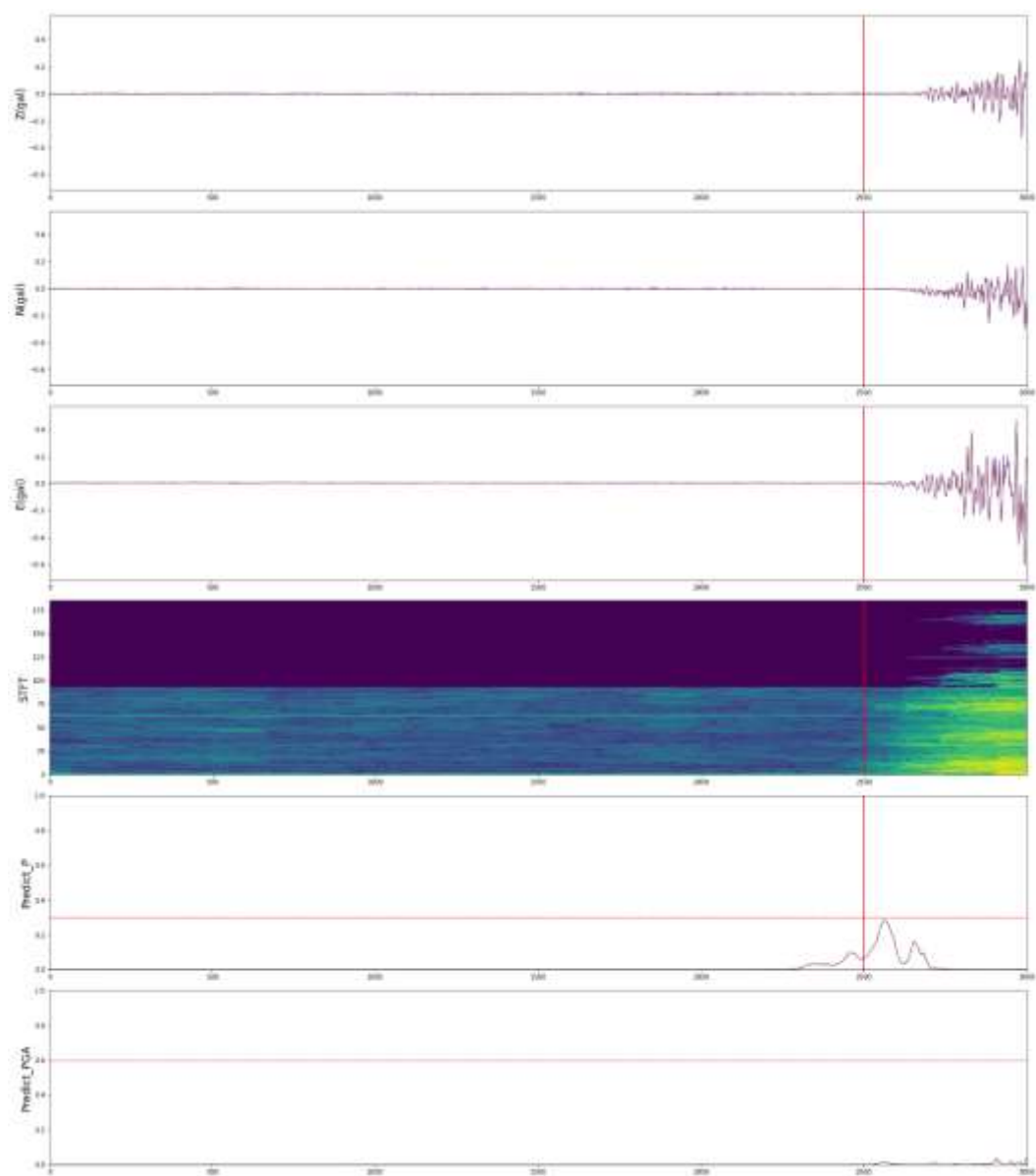


圖 3-37 DAPPER 標記測站 NOU 在 P 波相位後 5 秒之情況

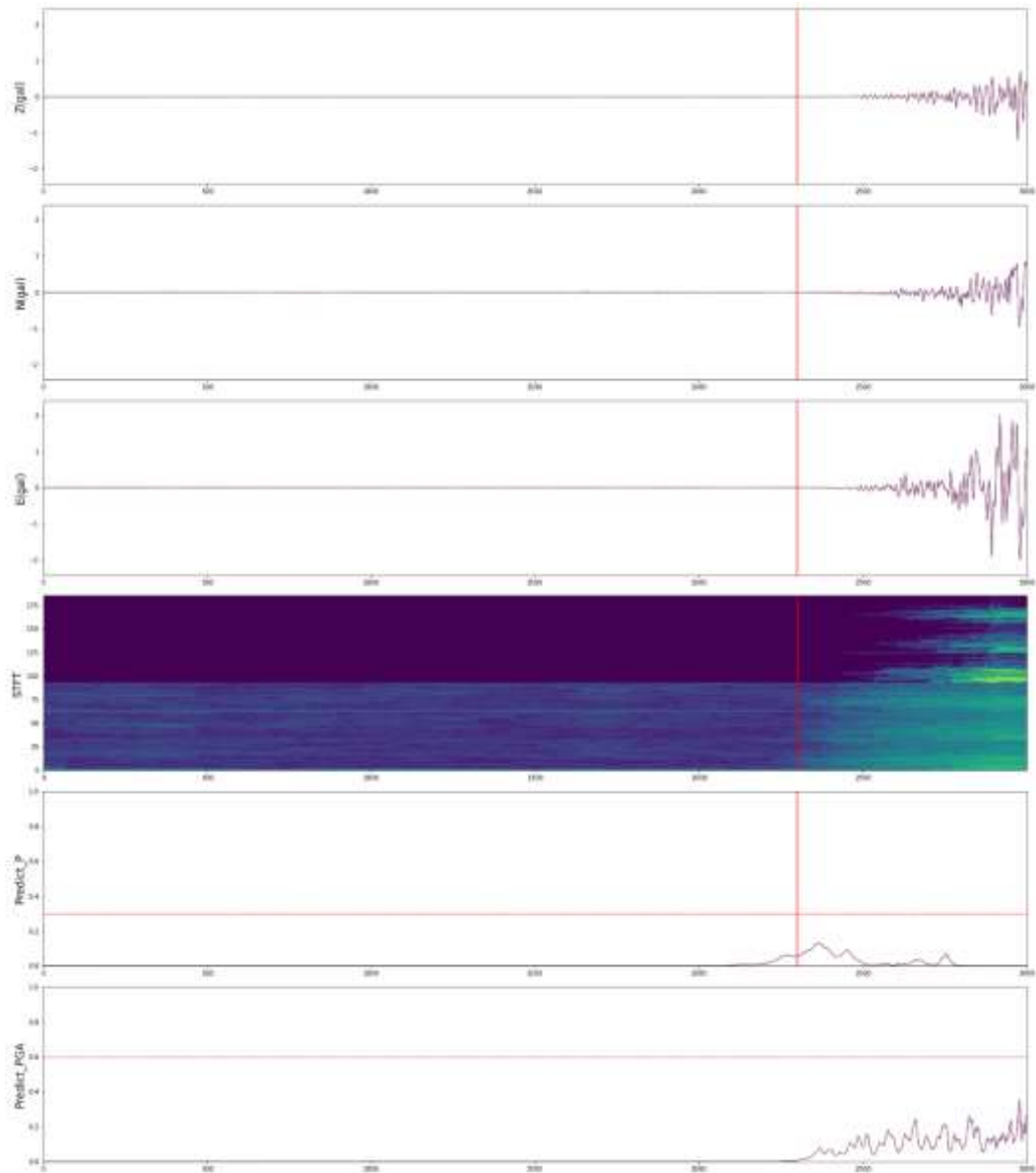


圖 3-38 DAPPER 標記測站 NOU 在 P 波相位後 7 秒之情況

此外，圖 3-39、圖 3-40 和圖 3-41 分別展示了 EAS 站在 P 波到達後 1 秒、3 秒和 5 秒的預測結果。該站的 PGA 為 17.42 gal。從圖中可以觀察到，與 ground truth 相比，DAPPER 在實際相位到達前的初始波動期間選取了相位，導致了顯著的時間差異。直到第 5 秒時，相位才被選取得更接近真實值，但仍然與 ground truth 存在明顯的差距。

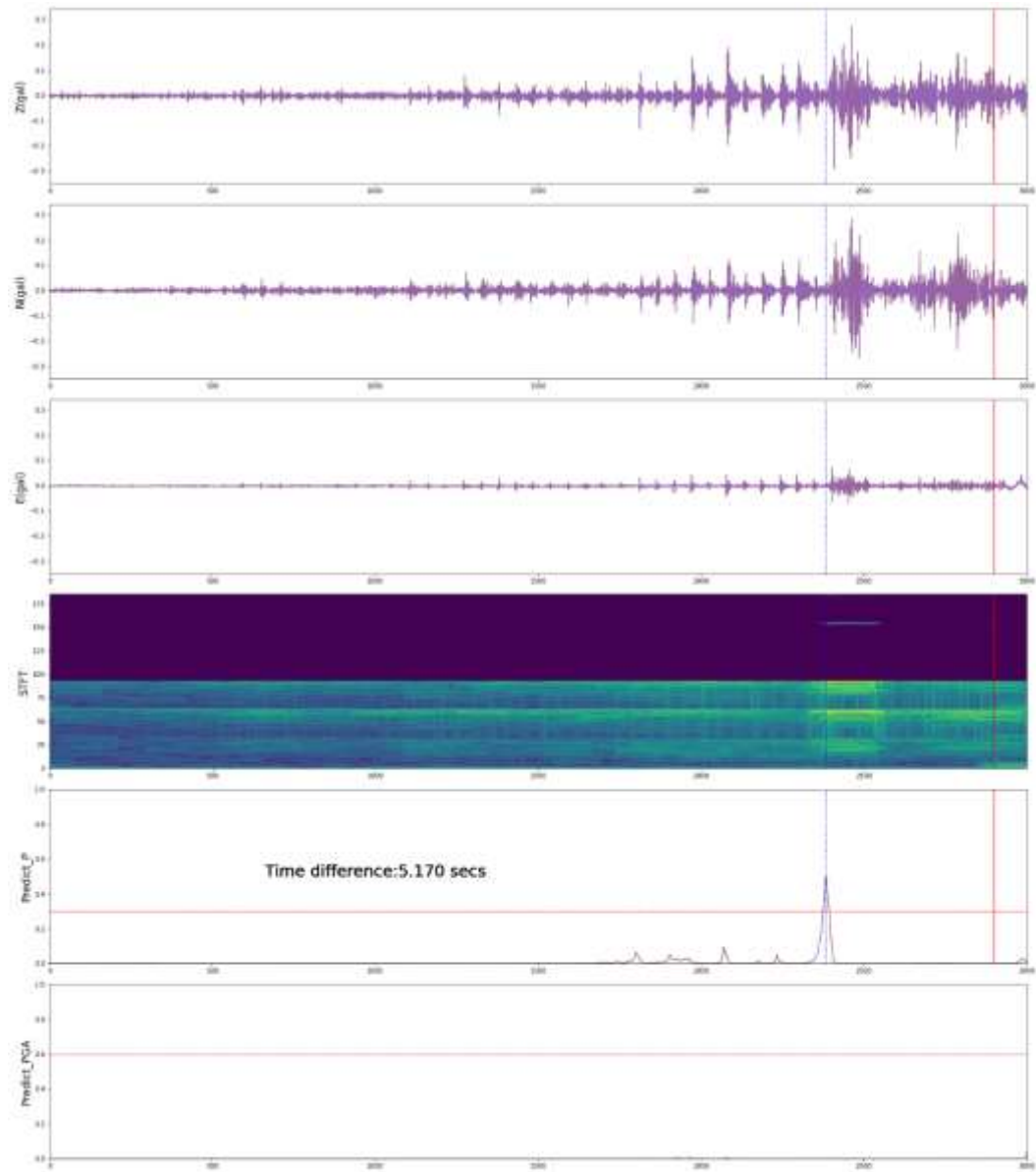


圖 3-39 DAPPER 標記測站 EAS 在 P 波相位後 1 秒之情況

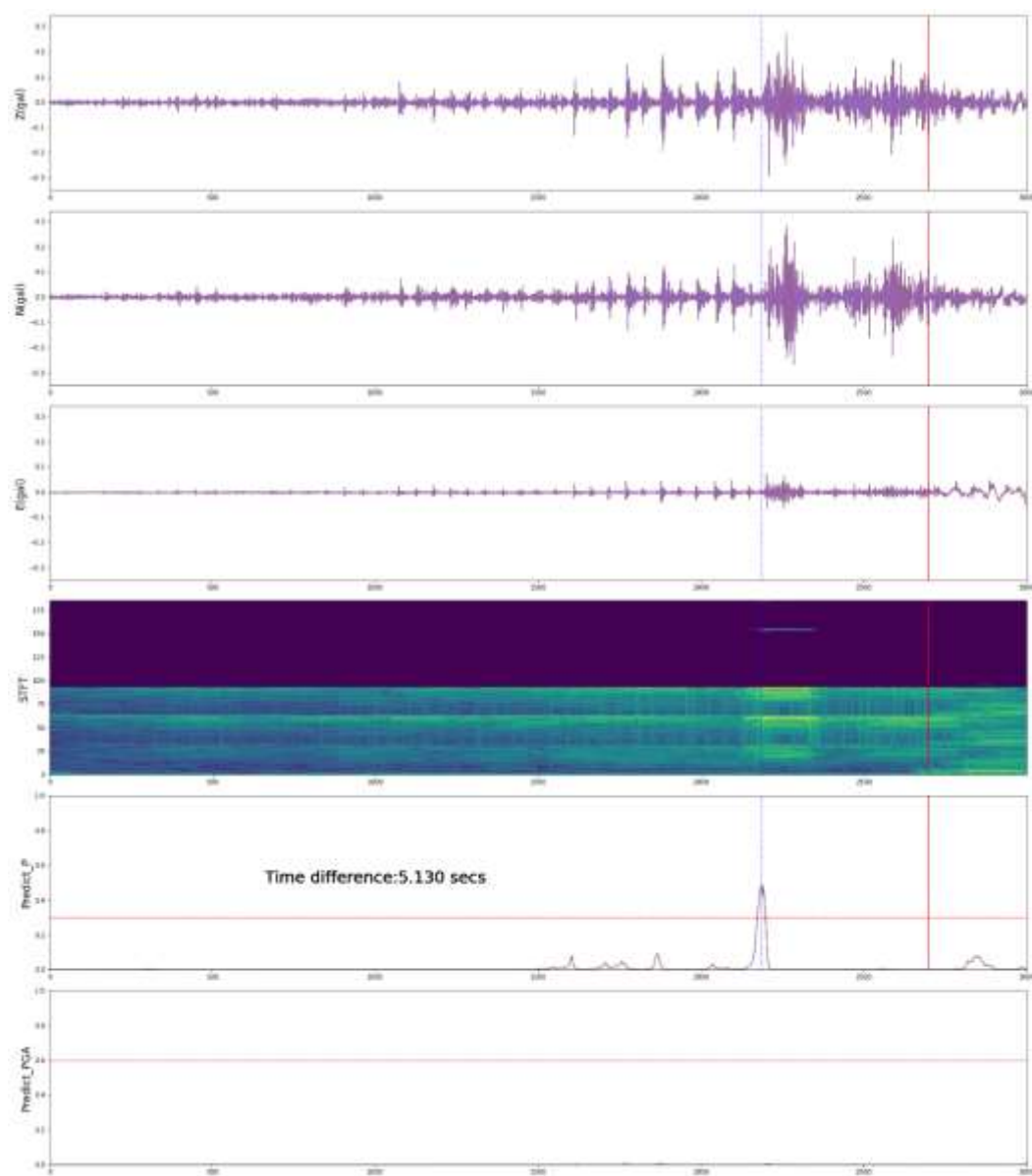


圖 3-40 DAPPER 標記測站 EAS 在 P 波相位後 3 秒之情況

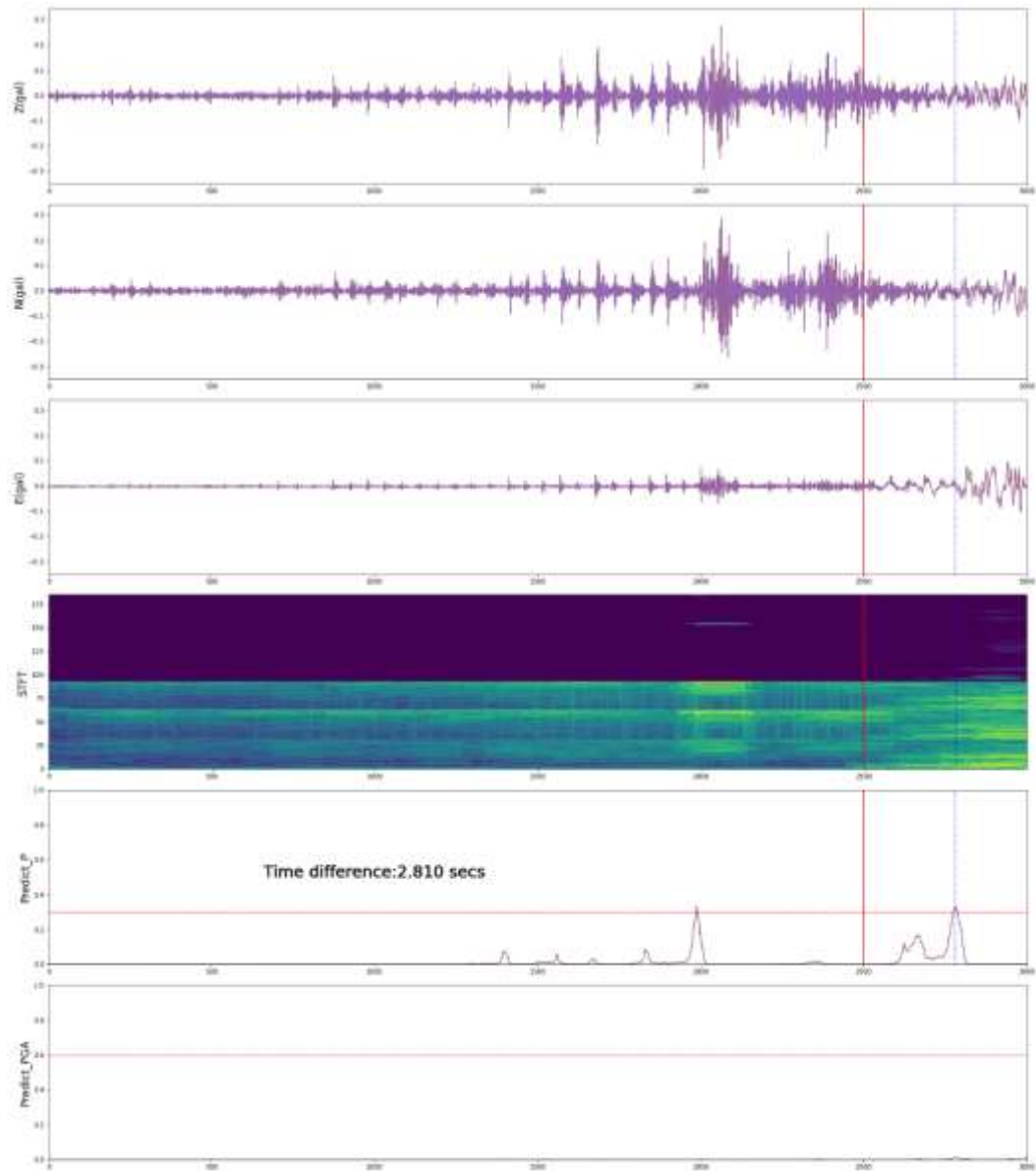


圖 3-41 DAPPER 標記測站 EAS 在 P 波相位後 5 秒之情況

為了模擬模型的實際警報性能，必須考慮到一旦發出預警，便無法撤回。通常會依賴 P 波到達後幾秒內的預測結果來發布警報。本次模擬將評估 P 波到達後 1、2、3、5、7 和 10 秒的預測和警報性能。P 波到時為氣象署提供。

圖 3-42 顯示了 PGA 超過 25 gal 的測站，主要集中在臺灣南部，僅有少數站點未超過 25 gal。圖 3-43 展示了 DAPPER 在 P 波到達後不同秒數的預測表現，包括召回率 (Recall)、精確率 (Precision) 和 F1 分數。對於此次事件，P 波到達後 1、2 和 3 秒的 PGA 預測顯示出較高的漏報率，表現不如理想。整體表現在 5 秒後顯著改善，這與測試數據集中觀察到的結果有所不同。

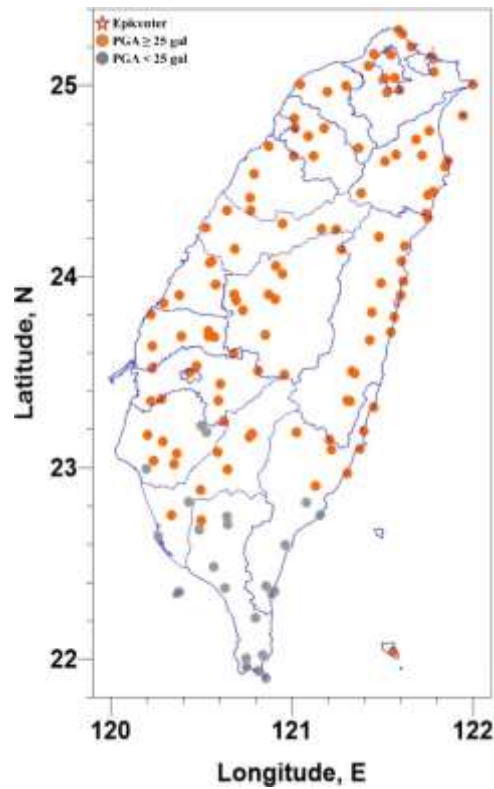


圖 3-42 實際 PGA 大於和小於 25gal 之測站

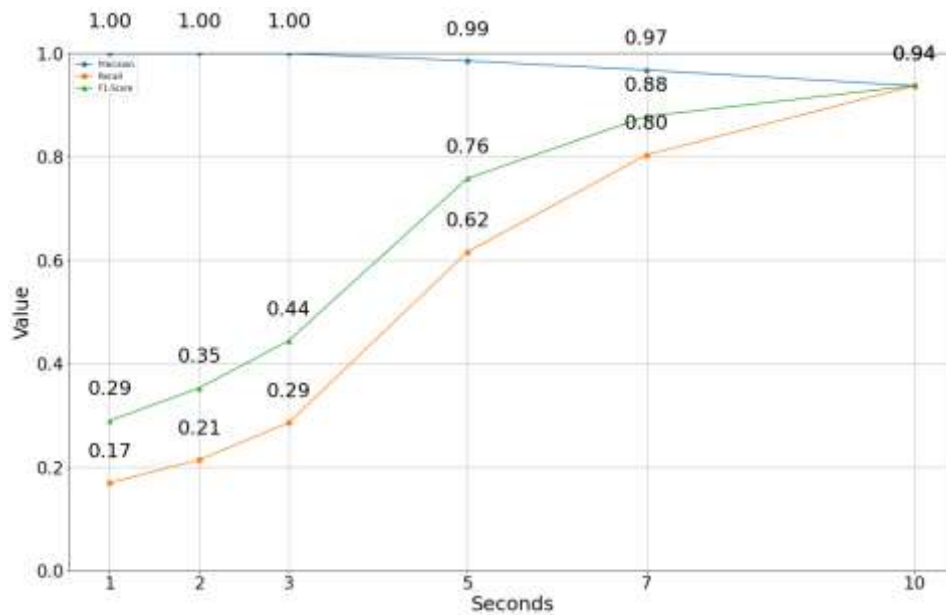


圖 3-43 DAPPER 在 P 波相位後 10 內之預測表現

在此次事件中，發現許多站點的 P 波到達時間與 PGA 超過 25 gal 之間存在顯著的時間差。例如，如圖 3-44 所示，CHK 站測得的 PGA 為 111.41 gal，從 P

波到達到 PGA 超過 25 gal 的時間差為 15.3 秒。附錄 C 提供了所有站點的 P 波到達時間與 PGA 超過 25 gal 的時間之詳細統計摘要。根據統計，此事件中 P 波到達與 PGA 超過 25 gal 的平均時間差為 13.98 秒，中位數為 11.35 秒，標準差為 9.96 秒。

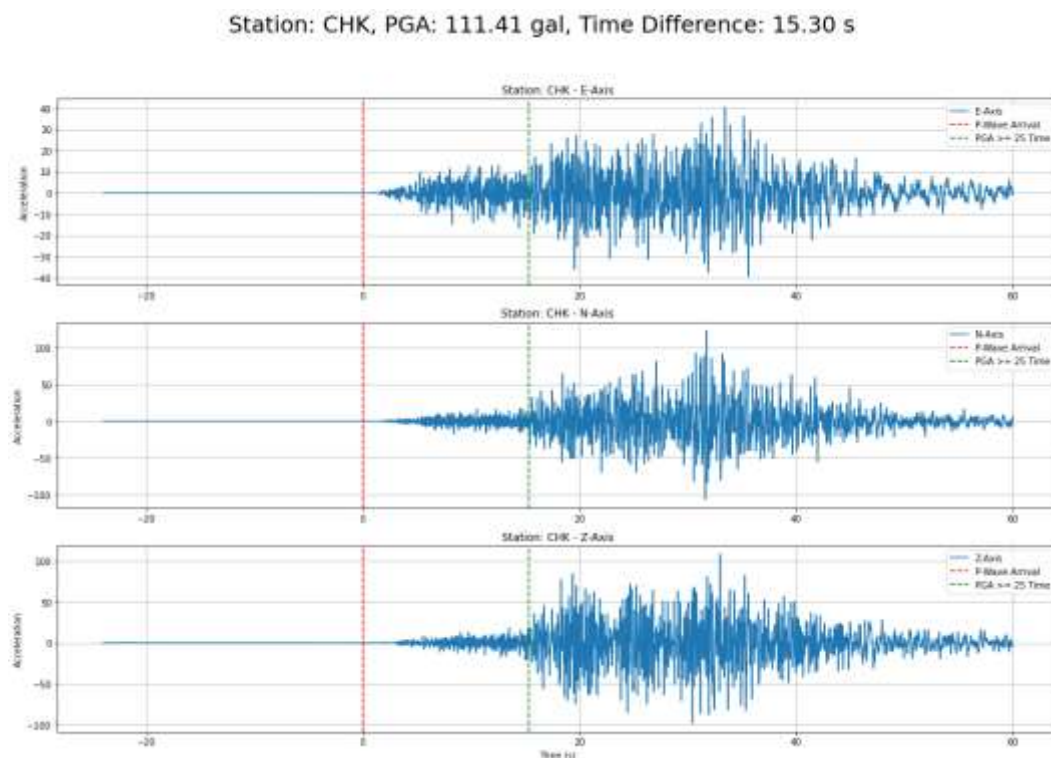


圖 3-44 測站 CHK 之三軸向加速度 60 秒訊號

為此，我們將 PGA 超過 25 gal 的時間作為標準。如果預測的 PGA 超過 25 gal，且該站點的 PGA 已經超過此數值，則將其額外標記為「Late alert」。圖 3-32 至圖 3-41 顯示了不同秒數單站的警報結果，分類為 Alert in time、Late alert、Miss alert、False alert 和 No alert needed。Alert in time 是某測站 PGA 大於 25 的時刻還沒到但模型此刻認為 PGA 大於 25。

Late alert 是某測站 PGA 已經大於 25 的時刻已經過了，但模型此刻才認為 PGA 大於 25。

Miss alert 是某測站 PGA 在未來某個時候才大於 25，而模型在此刻不認為 PGA 會大於 25。

False alert 是某測站 PGA 在此刻和未來都不會大於 25gal，但模型在此刻認為 PGA 大於 25gal。

No alert needed 是 PGA 在此刻和未來都不會大於 25gal 的測站。

可以觀察到，在 P 波到達後的第一秒內，已有三個站點的 PGA 值超過了 25 gal。隨後，直到 P 波到達後的第七秒，才有另一個站點出現 Late alert。預測結果

表明，雖然在 P 波到達後前三秒的表現不如預期，但大多數測站能夠在當地 PGA 超過 25 gal 之前發出警報。然而，也可以觀察到在第 5、7 和 10 秒時出現了一些誤報情況。

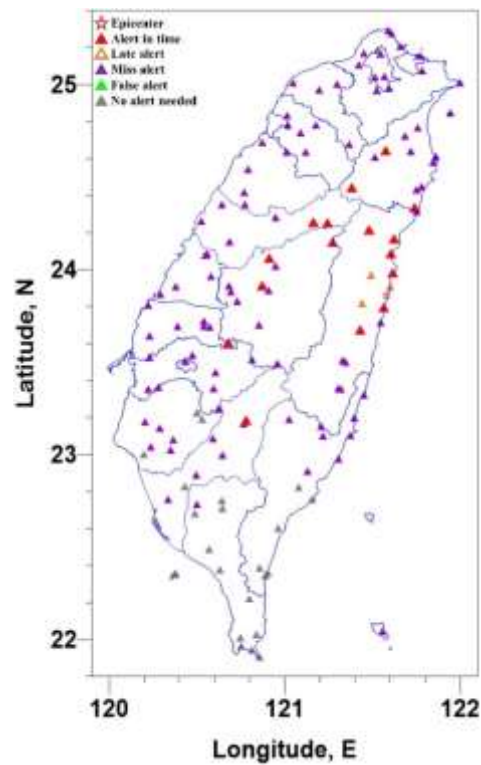


圖 3-45 各測站在 P 波相位後 1 秒發報情況

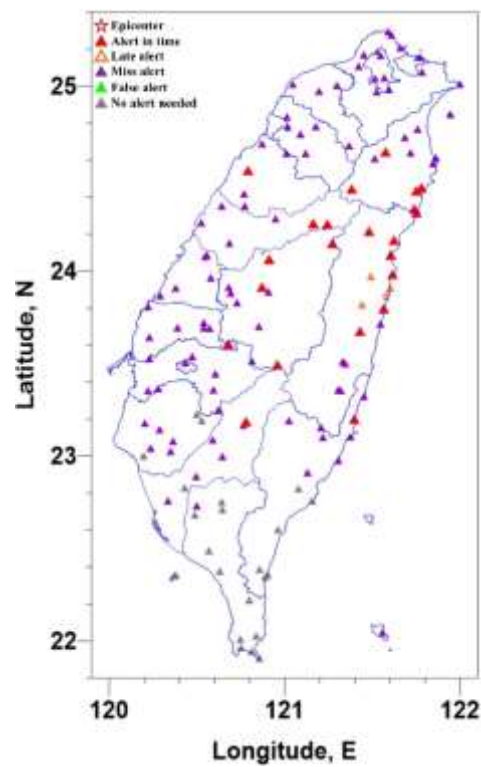


圖 3-46 各測站在 P 波相位後 2 秒發報情況

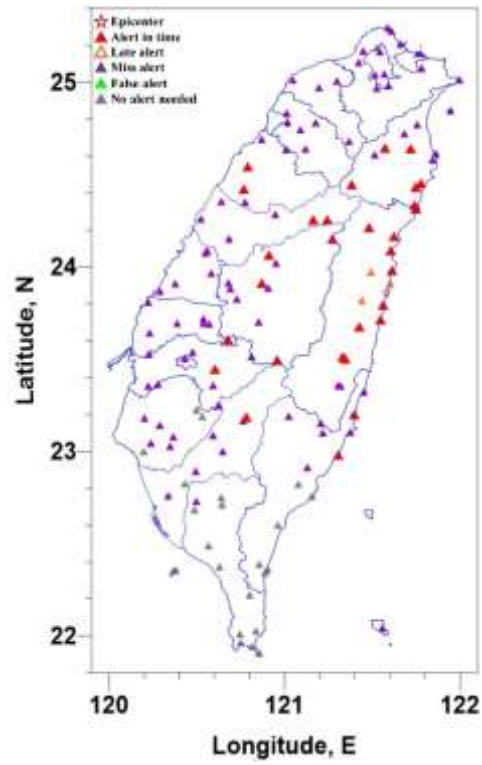


圖 3-47 各測站在 P 波相位後 3 秒發報情況

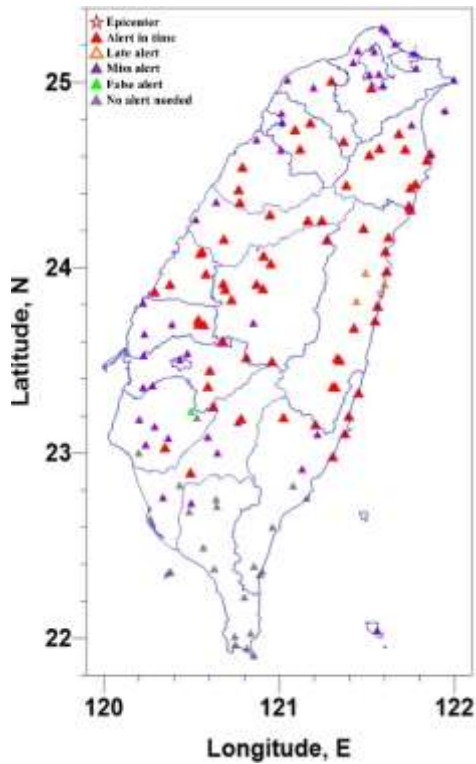


圖 3-48 各測站在 P 波相位後 5 秒發報情況

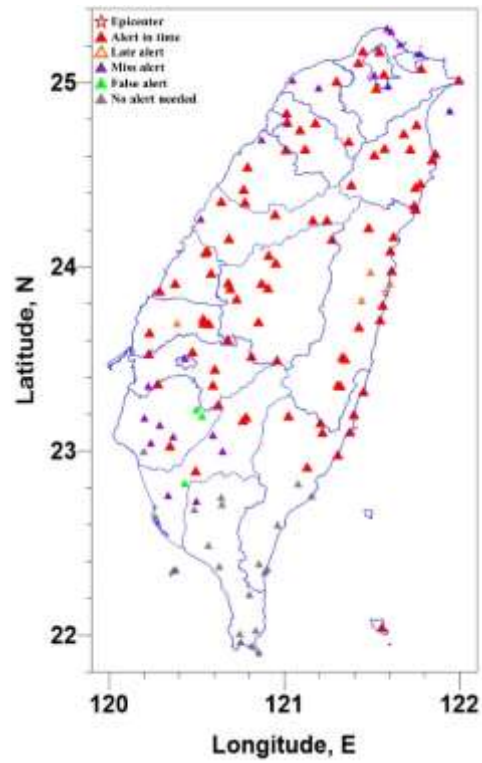


圖 3-49 各測站在 P 波相位後 7 秒發報情況

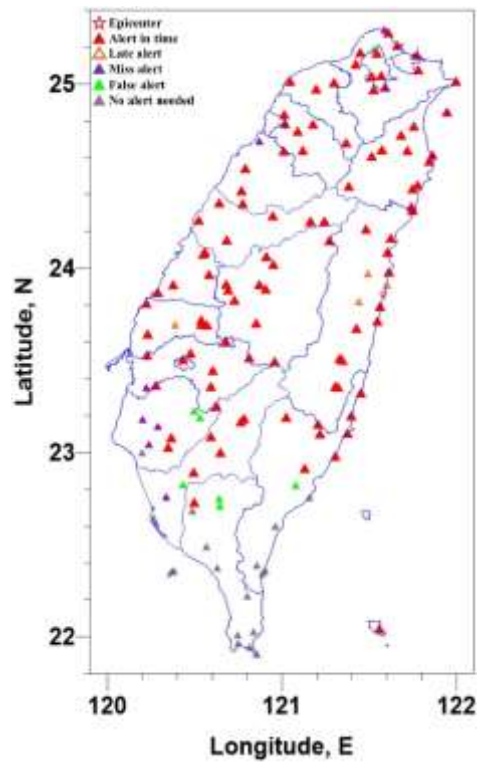


圖 3-50 各測站在 P 波相位後 10 秒發報情況

用 DAPPER 提供的 P 波相位後兩秒之波形資料對各縣市模擬警報發布情況。

顯示了警報發布的情況，如圖 3-51。圖中時間是以地震發生時間(上午 7:58:09.8，GMT+8) 為基準，不考慮當地警報發布時間，各縣市在震央發震後多少秒收到警報。根據圖中結果，除了基隆外，臺灣所有地區均收到了警報。雖然基隆的站點預測 PGA 超過 25 gal，但未選取相位，導致漏報。在此次事件中，模型達到了 100% 的精確率 (Precision)、94.11% 的召回率 (Recall) 以及 96.87% 的 F1 分數。

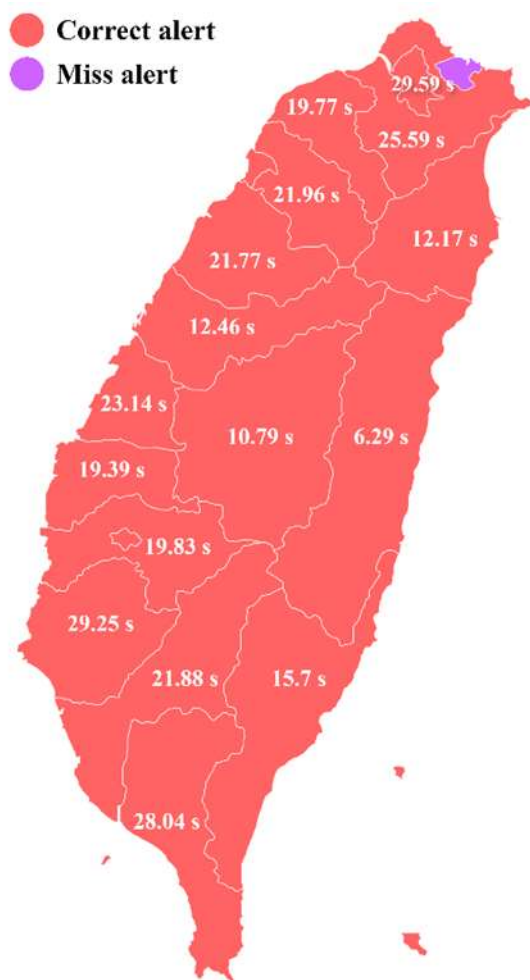


圖 3-51 各縣市發震時間至最早出現 25gal 測站之時間差

圖 3-52 為各縣市之 leading time。圖中的數字為各縣市中最早出現 25gal 之測站，模型對此測站預測 PGA 大於 25gal 之時間-此測站出現 25gal 的時間，所以負的值表示提早幾秒發出警報。故此圖中之縣市除了南投和花蓮外，其餘縣市都有預警時間。

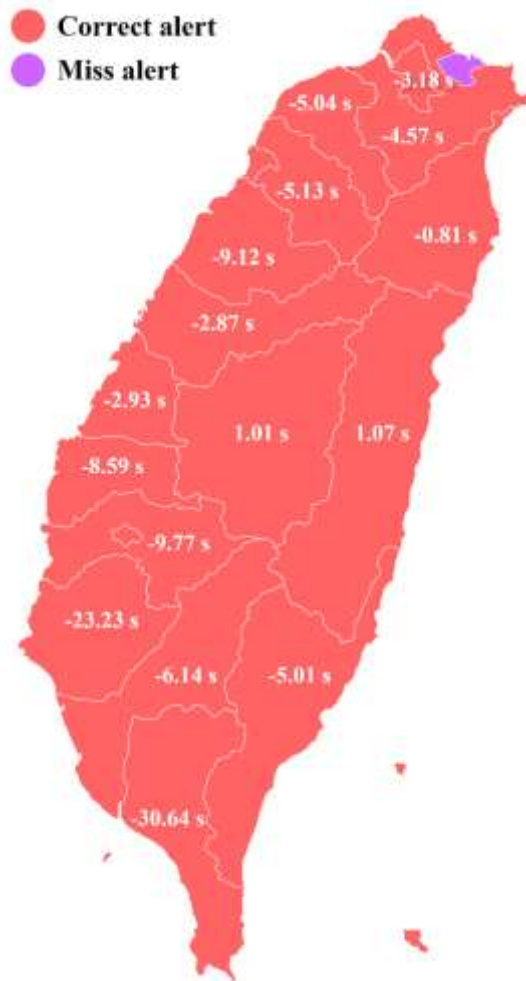


圖 3-52 各縣市 leading time

4 結論

在 TSMIP 測站現地型 CNN 強震預警模型改善方面，本計畫依照氣象署之建議，在研究室架設 MQTT 通訊架構，並接收即時 P-alert 之波形資料，以接近氣象署線上環境。除了架設 MQTT 通訊架構外，本計畫並將 Triton Inference Server 部屬在研究室之 MQTT 通訊架構，期能縮短模型預測時間。根據所有觸發事件的統計結果顯示，確實可以縮短資料前處理與模型預測所需時間，但對於接收相位資訊所需的時間並沒有縮短，預計將於 2024 年 12 月將此一新的 MQTT 通訊架構部屬在氣象署。

在 CNN 模型於氣象署即時運行結果方面，本研究統計 2024 年 1 至 9 月的 9 筆震度達 5 級以上的地震事件單站預測結果，其 Precision 為 78.9%，Recall 為 67.6%，但各事件中均存在漏報和誤報的情況。漏報多由地震事件特性所致，部分事件初期波形較小，使得在 P 波到達後 2 秒內的預測結果偏低；此外，某些事件在主震前發生較小地震導致觸發位置並非主震，亦使得預測結果偏低。而誤報的主要因為某些事件中程式記錄的 P 波起始點位因更新延遲而導致延後，使得預測值偏大，

此部分程式之缺點已改善完成，等待之後的地震事件加以驗證。

針對地震預警系統的網格化設計，本研究探討依據面積、測站數量及人口數量劃分網格的可行性。結果顯示，以區域面積和測站數量為劃分基準能有效提升預測準確性，即使有少數漏報，整體誤差仍在可接受範圍內。然而，若網格過小或過多，可能出現狹長且不規則的網格形狀，影響系統的預警覆蓋效果。相對之下，使用人口數為劃分依據因人口分布不均，造成網格大小差異過大，較不適合地震預警系統。綜合考量後，建議以區域面積和測站數量為主要劃分依據，以提升預警覆蓋範圍的準確性。

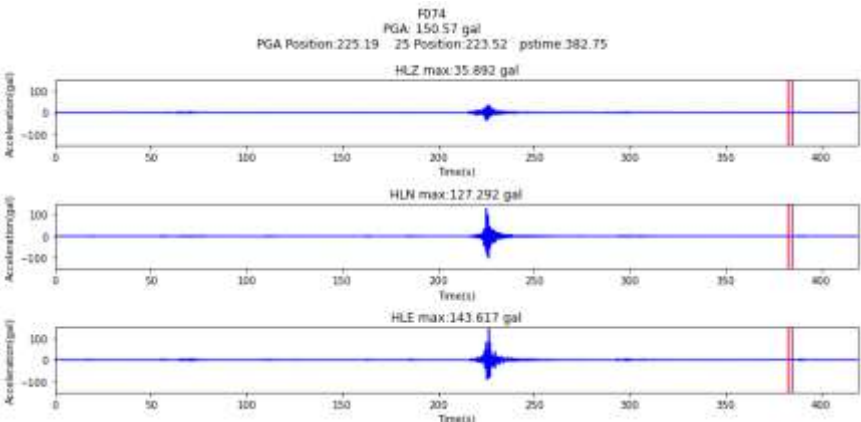
本計畫另一方面提出一可同時標注 P 到時且預測此地震 PGA 是否大於 25gal 之模型 DAPPER。相較於去年計畫之模型，本年度新增時頻域的波形特徵作為輸入，而在增加時頻域特徵後，對於預測 PGA 是否超過 25gal 之任務有顯著改善漏報之情況，但對地震相位標記之任務則無顯著改善。DAPPER 在 20240403 花蓮事件之表現如下，在地震事件中相位選取的表現方面，F1 分數為 87.05%，Precision 為 89.36%，Recall 為 84.85%；在對縣市預警發報方面，其 F1 分數為 96.87%、Precision 為 100%、Recall 為 94.11%。

參考文獻

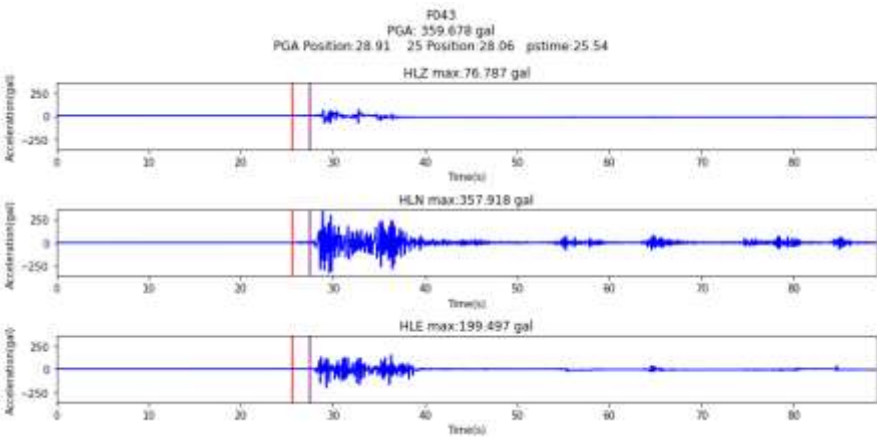
- [1] 黃昭文, "卷積神經網路應用於初達波預估當站最大地表加速度之研究," 碩士, 營建工程系, 國立臺灣科技大學, 臺北市, 2020.
- [2] W.-Y. Liao, E.-J. Lee, D.-Y. Chen, P. Chen, D. Mu, and Y.-M. Wu, "RED-PAN: Real-Time Earthquake Detection and Phase-Picking With Multitask Attention Network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-11, 2022, doi: 10.1109/tgrs.2022.3205558.
- [3] 陳威樑, "單站法即時地震波到時標記及強震預警預測模型開發," 碩士, 營建工程系, 國立臺灣科技大學, 臺北市, 2023.
- [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.

附錄

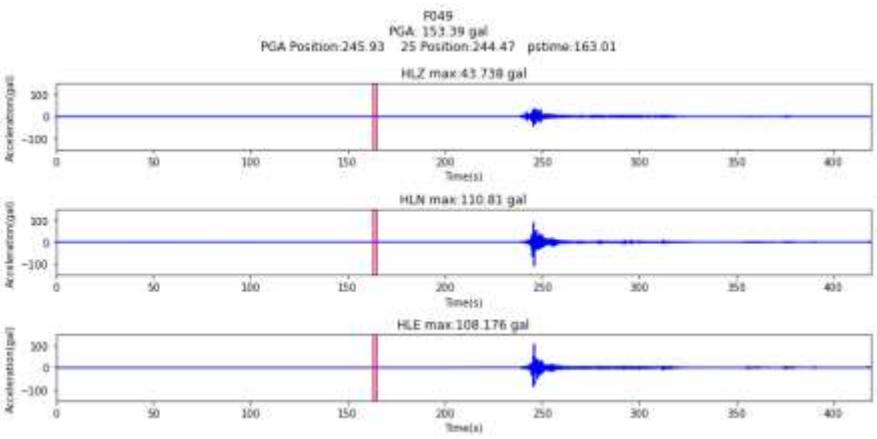
附錄 A：2024 年 1-9 月漏報之測站波形圖



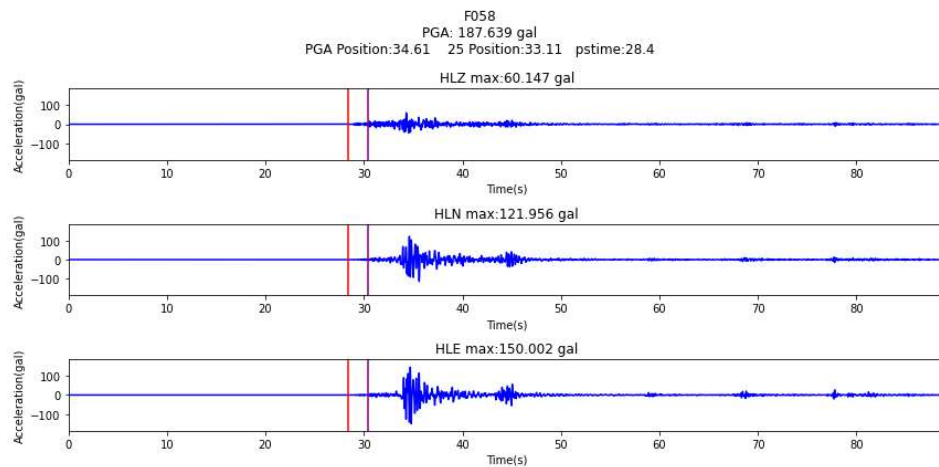
20240403(編號 021)F074 測站波形圖



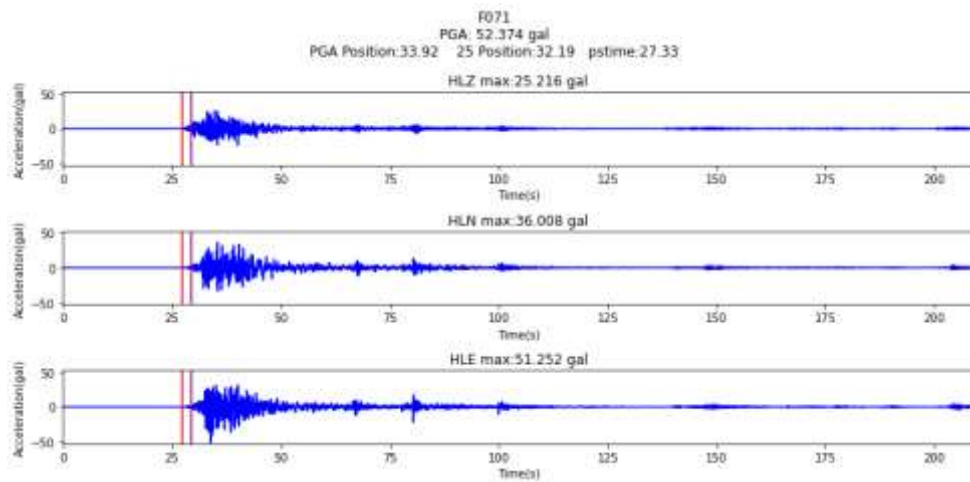
20240423(編號 266)F043 測站波形圖



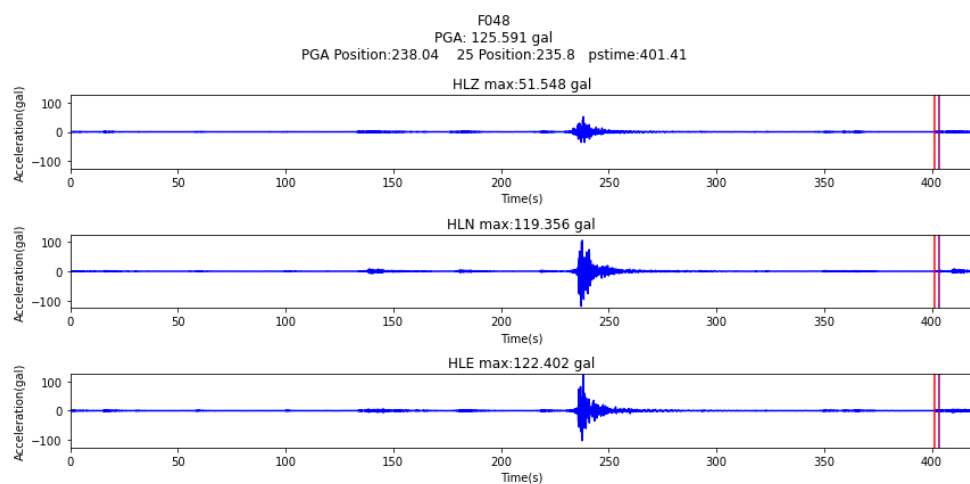
20240423(編號 266)F049 測站波形圖



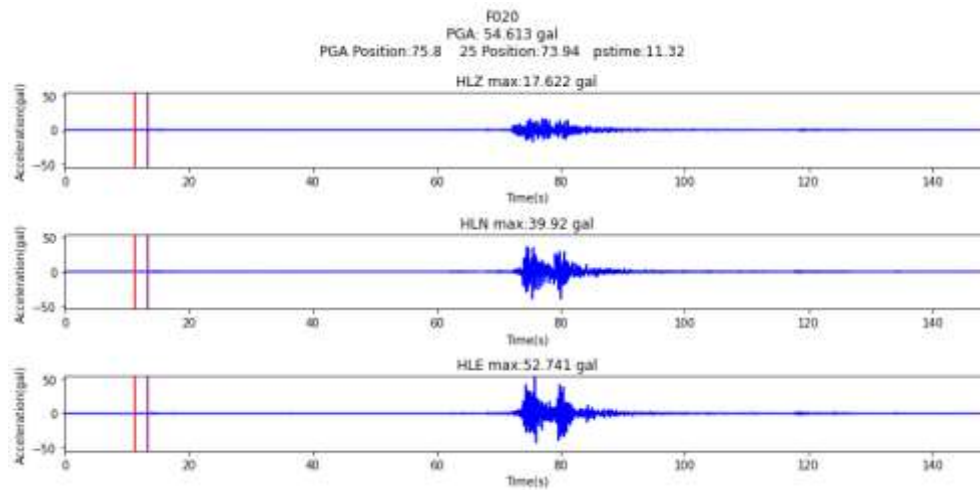
20240423(編號 266)F058 測站波形圖



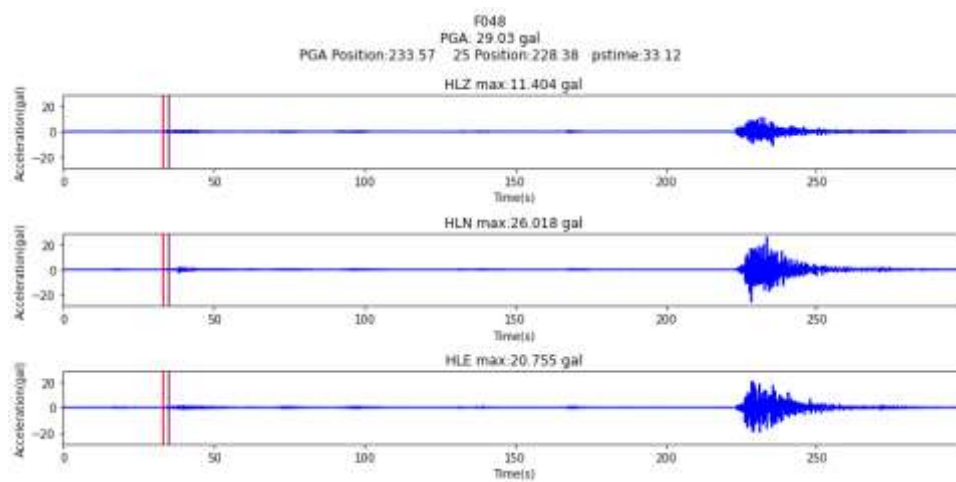
20240423(編號 266)F071 測站波形圖



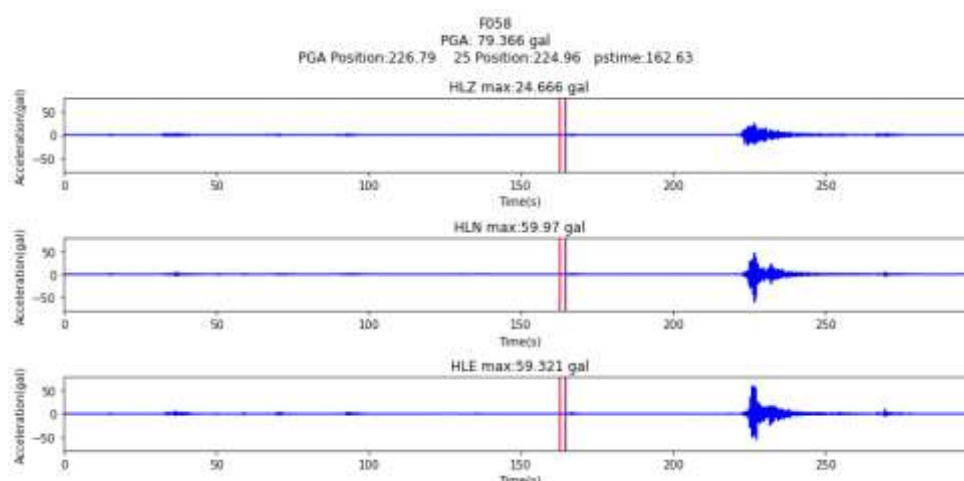
20240423(編號 267)F048 測站波形圖



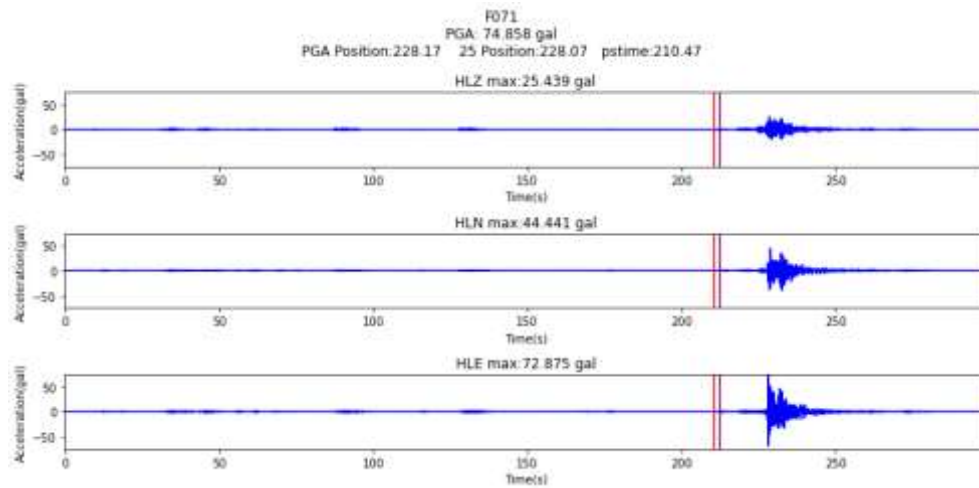
20240423(編號 300)F020 測站波形圖



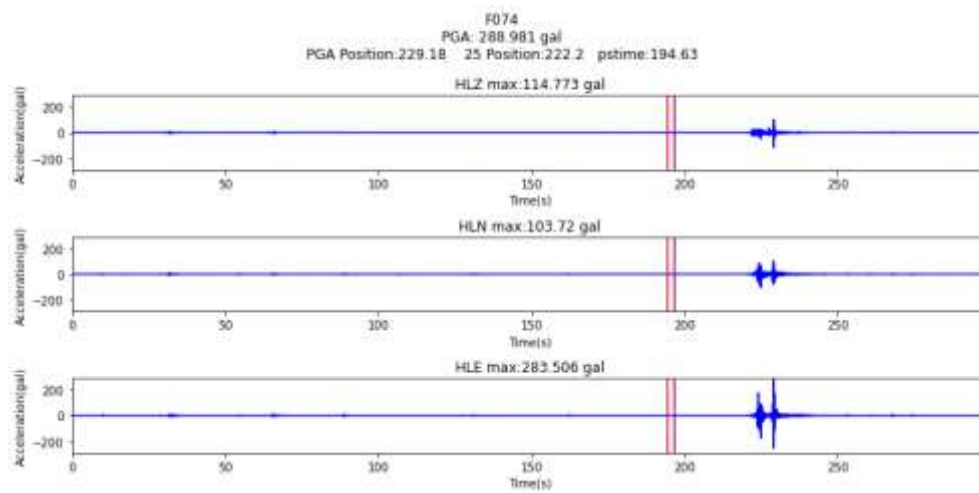
20240423(編號 300)F048 測站波形圖



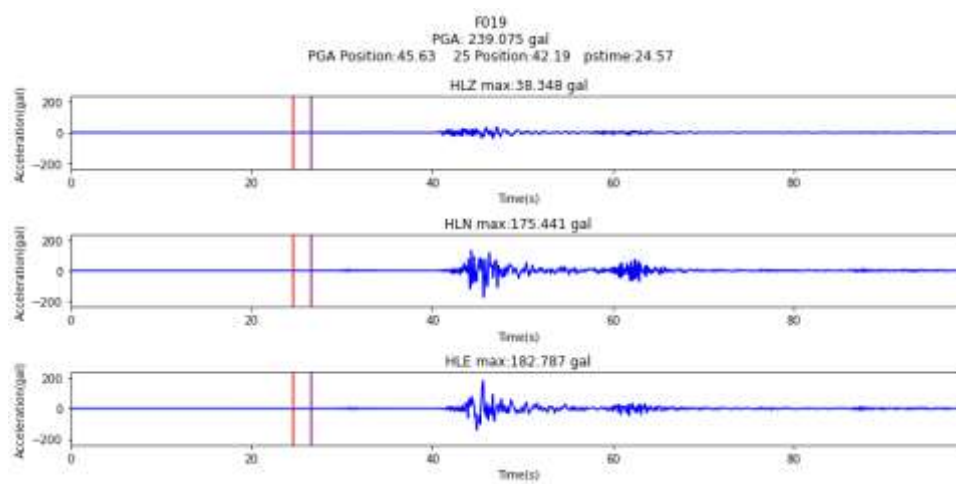
20240423(編號 300)F058 測站波形圖



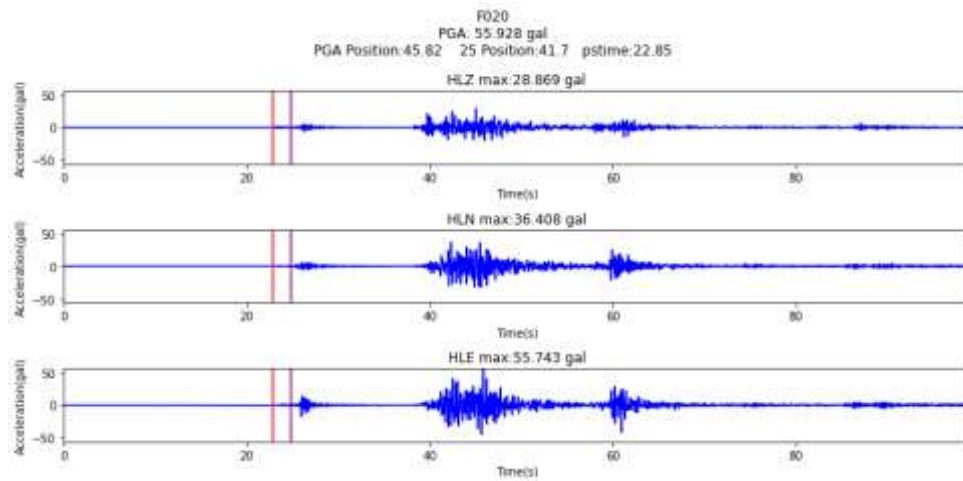
20240423(編號 300)F071 測站波形圖



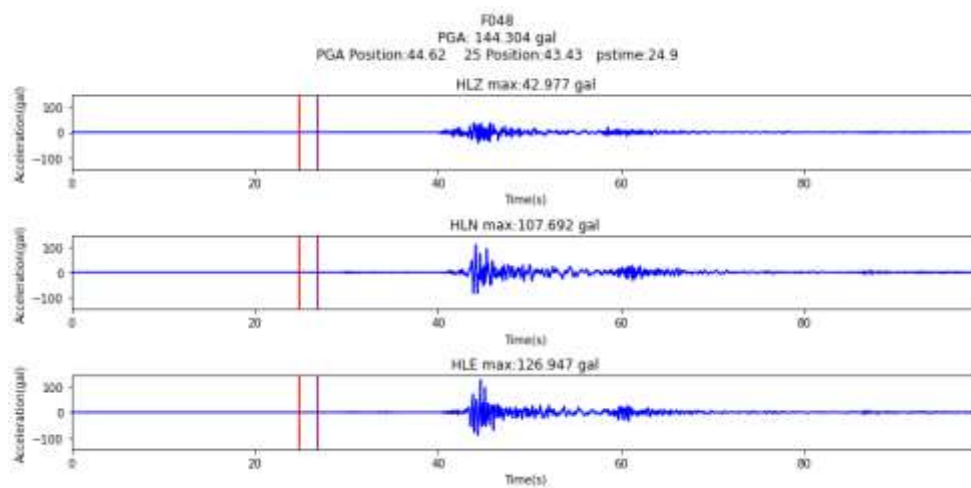
20240423(編號 300)F074 測站波形圖



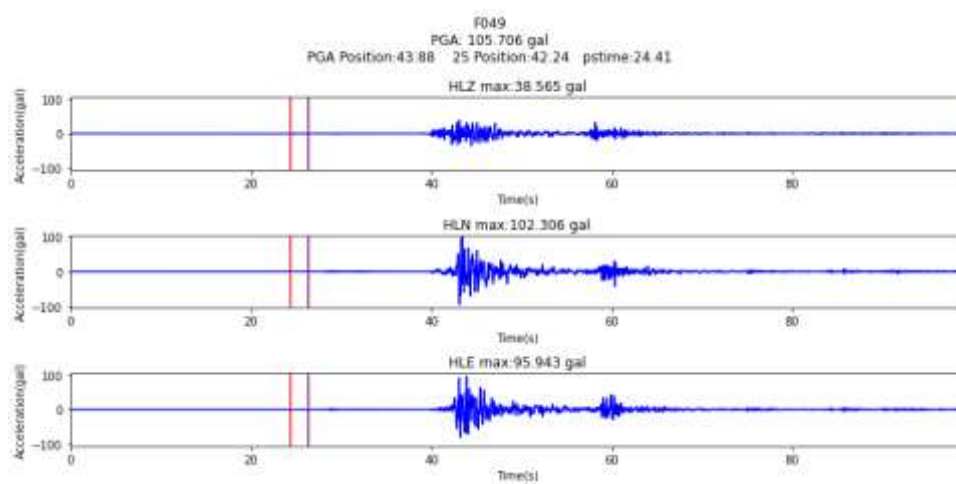
20240423(編號 307)F019 測站波形圖



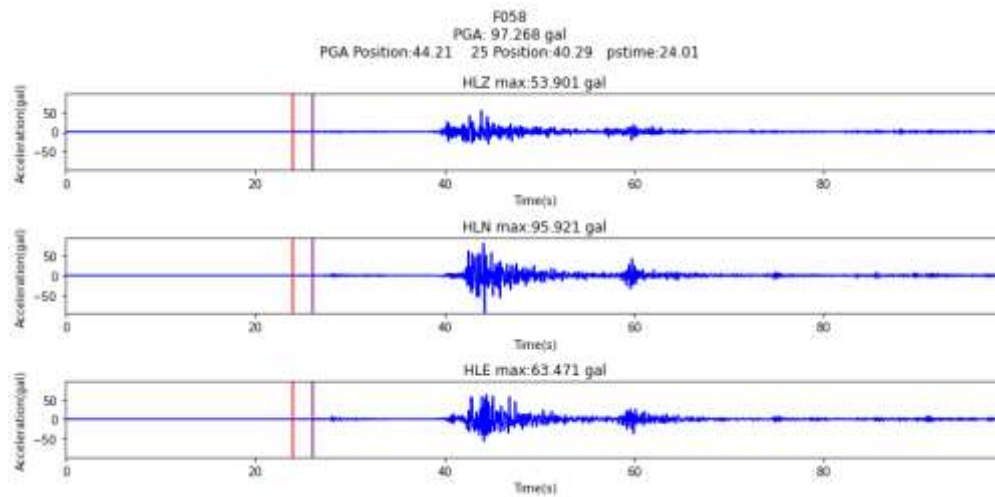
20240423(編號 307)F020 測站波形圖



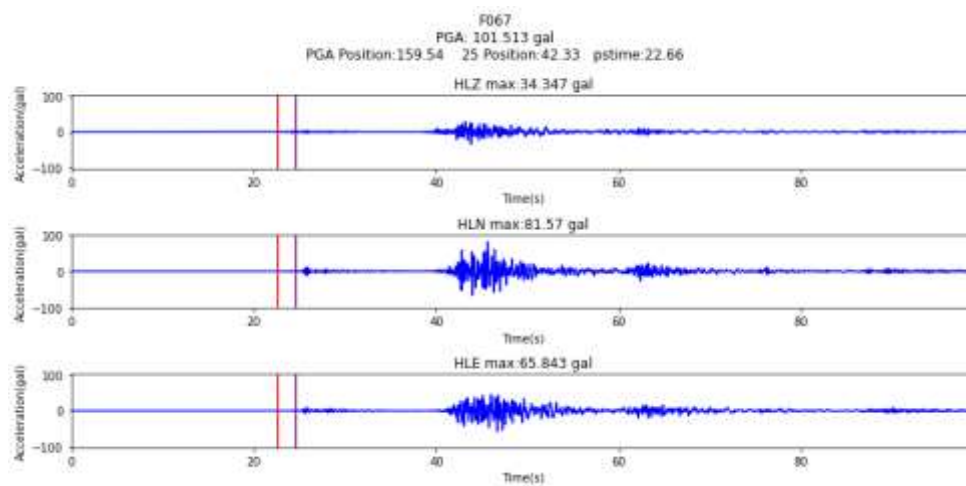
20240423(編號 307)F048 測站波形圖



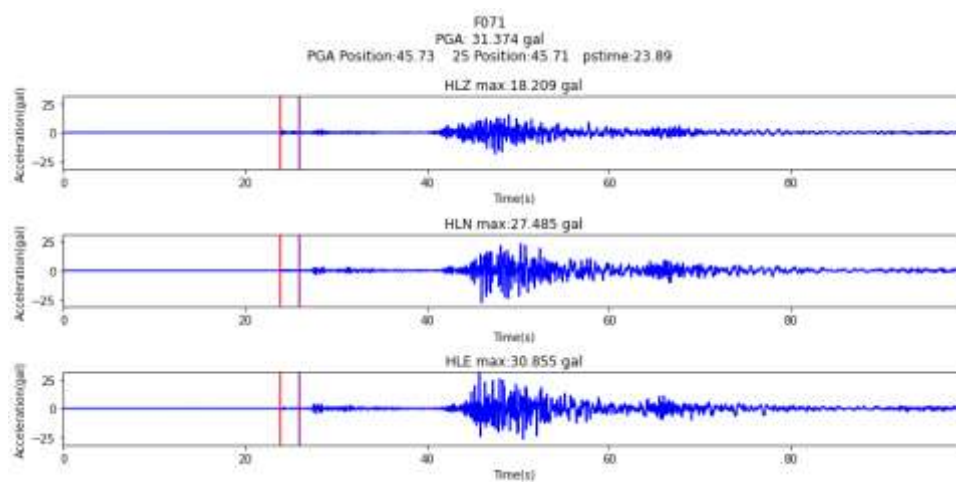
20240423(編號 307)F049 測站波形圖



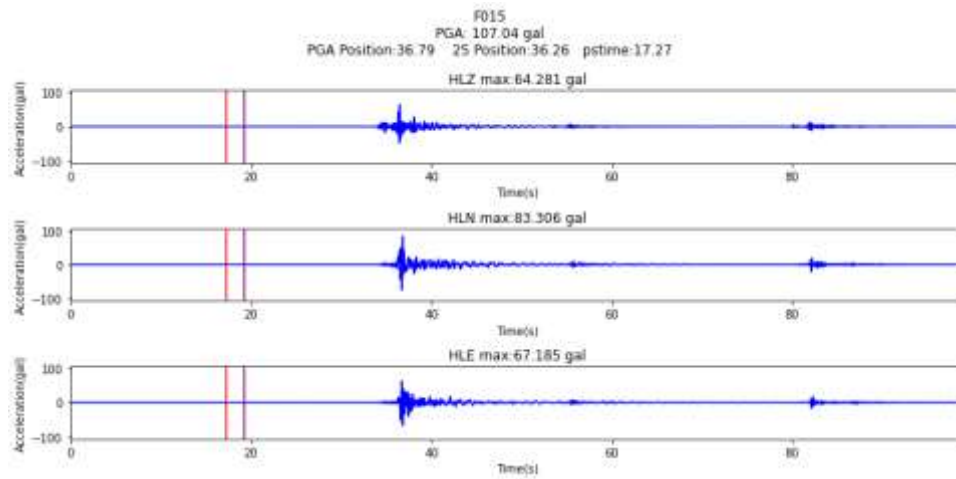
20240423(編號 307)F058 測站波形圖



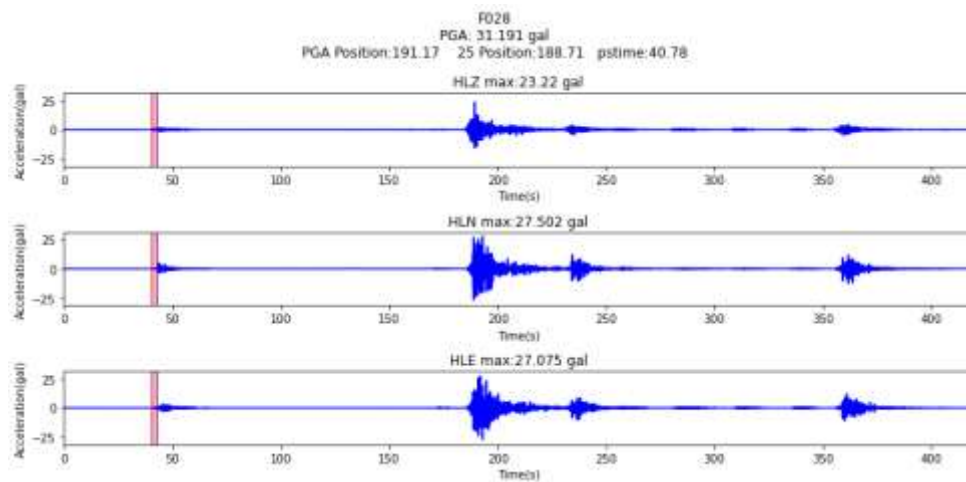
20240423(編號 307)F067 測站波形圖



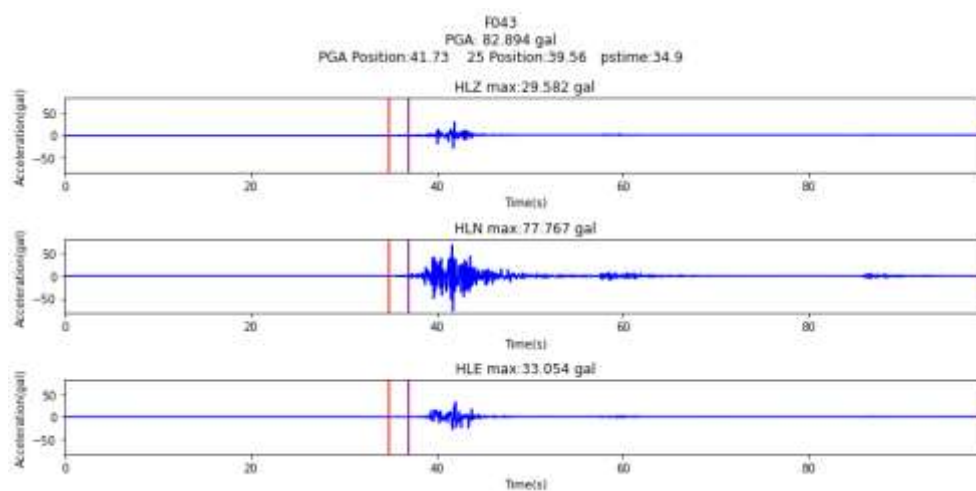
20240423(編號 307)F071 測站波形圖



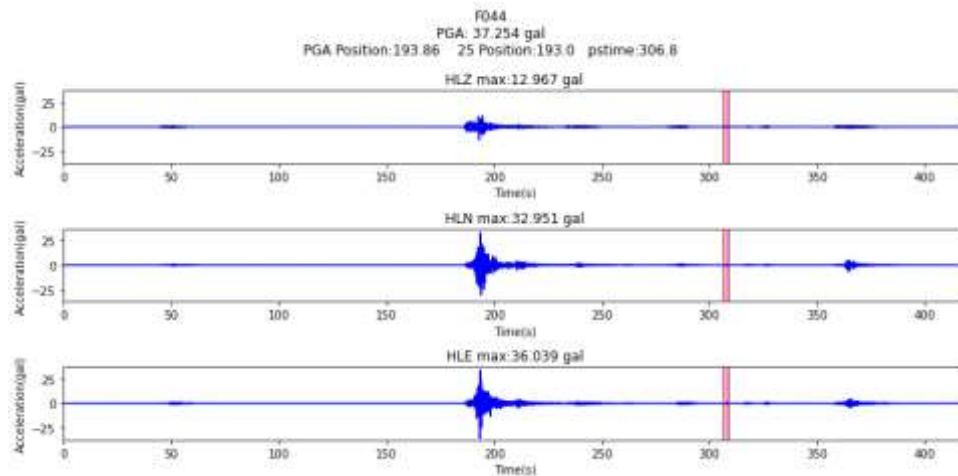
20240423(編號 318)F015 測站波形圖



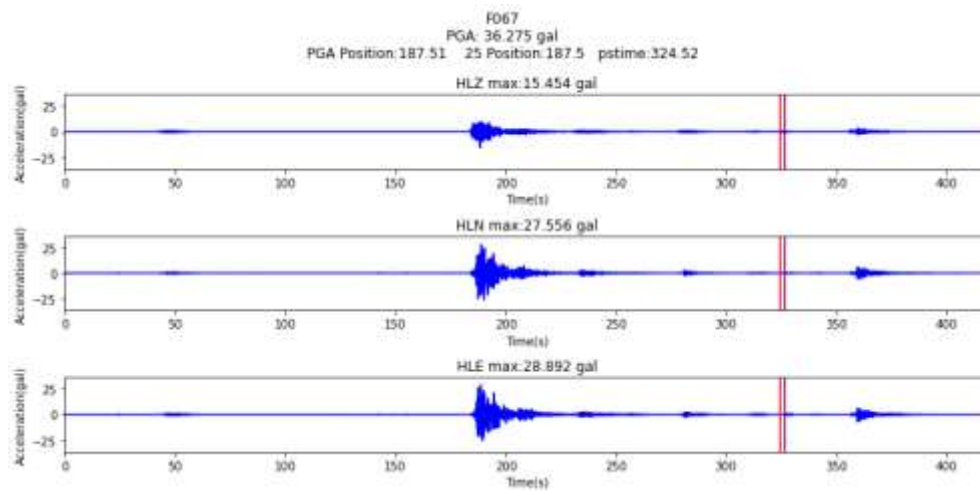
20240423(編號 318)F028 測站波形圖



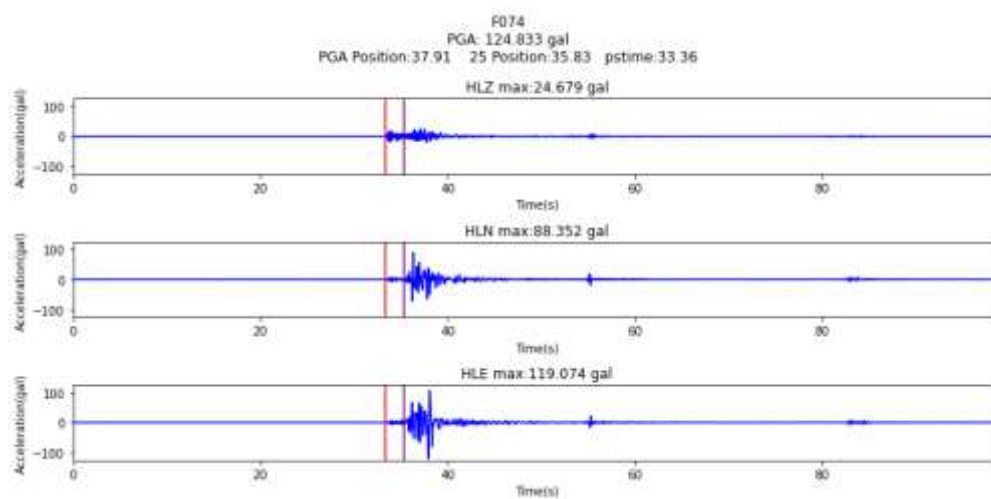
20240423(編號 318)F043 測站波形圖



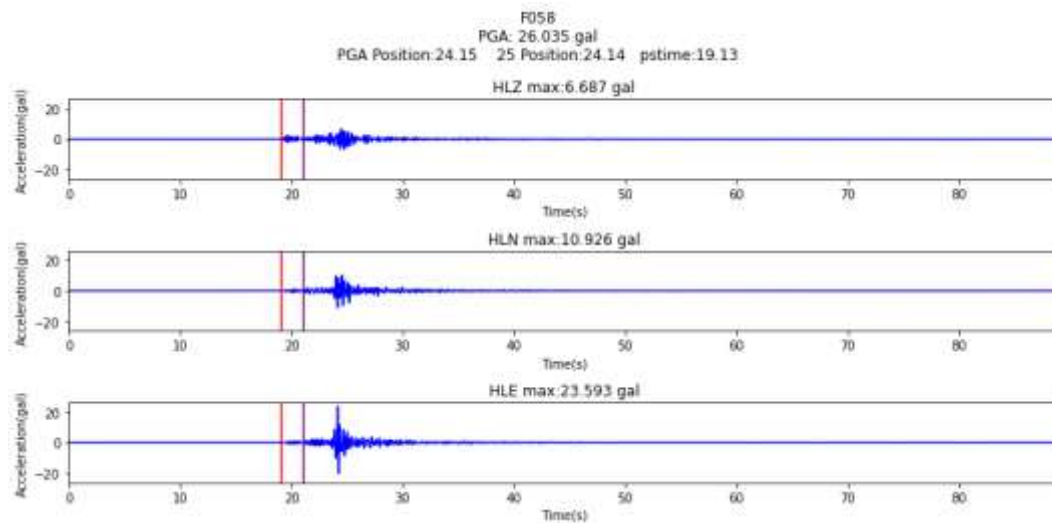
20240423(編號 318)F044 測站波形圖



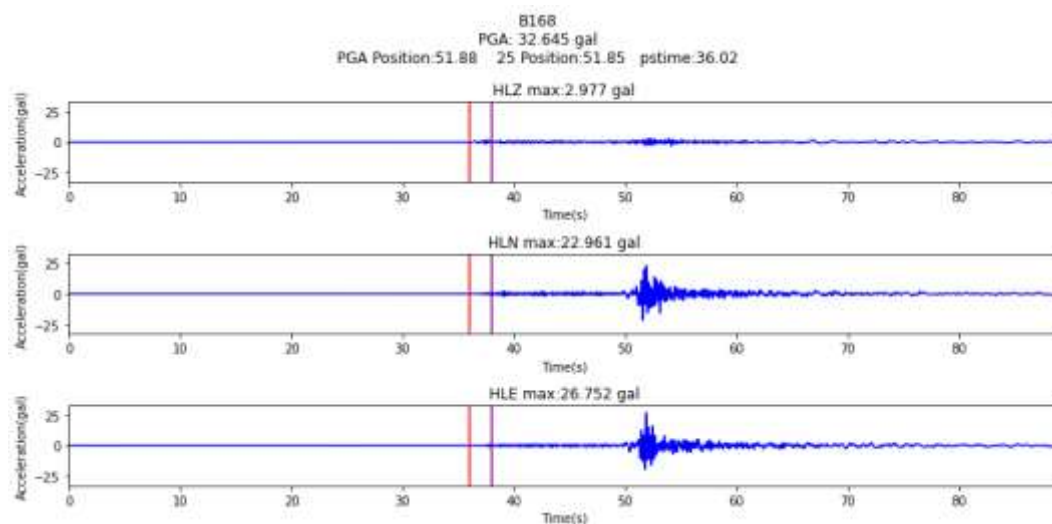
20240423(編號 318)F067 測站波形圖



20240423(編號 318)F074 測站波形圖

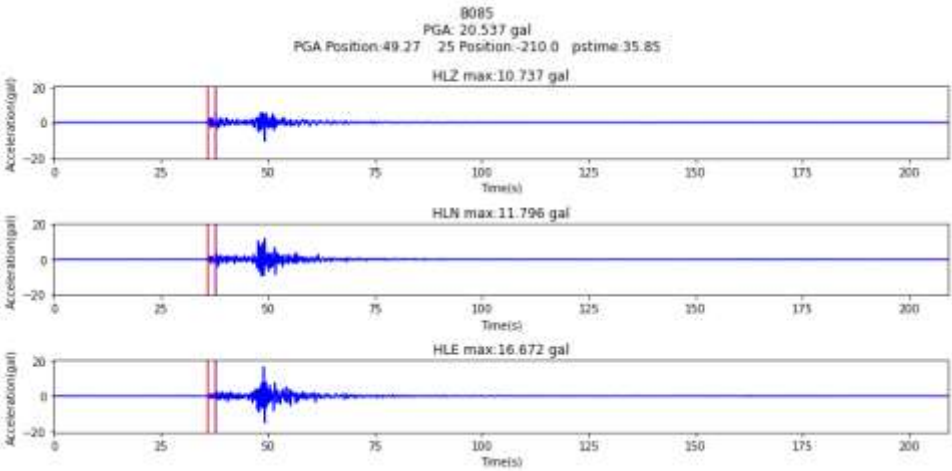


20240508(編號 377)F058 測站波形圖

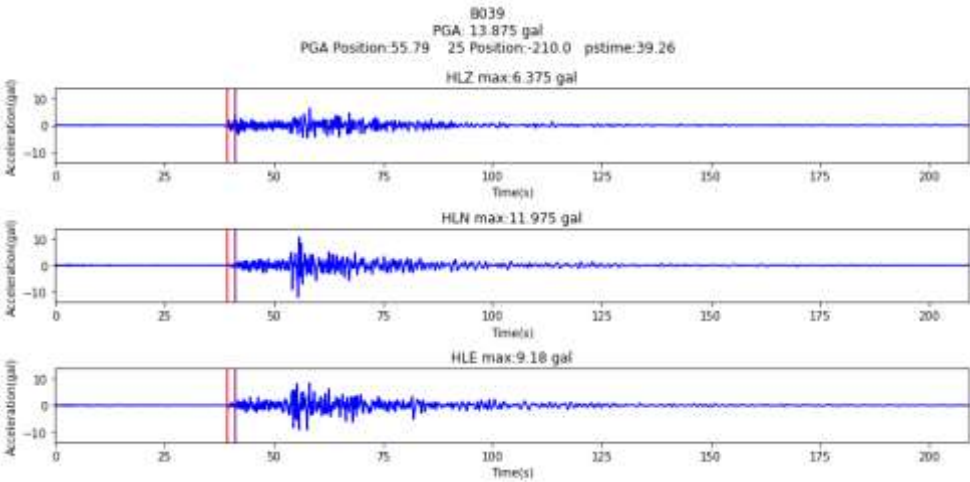


20240815(編號 442)B168 測站波形圖

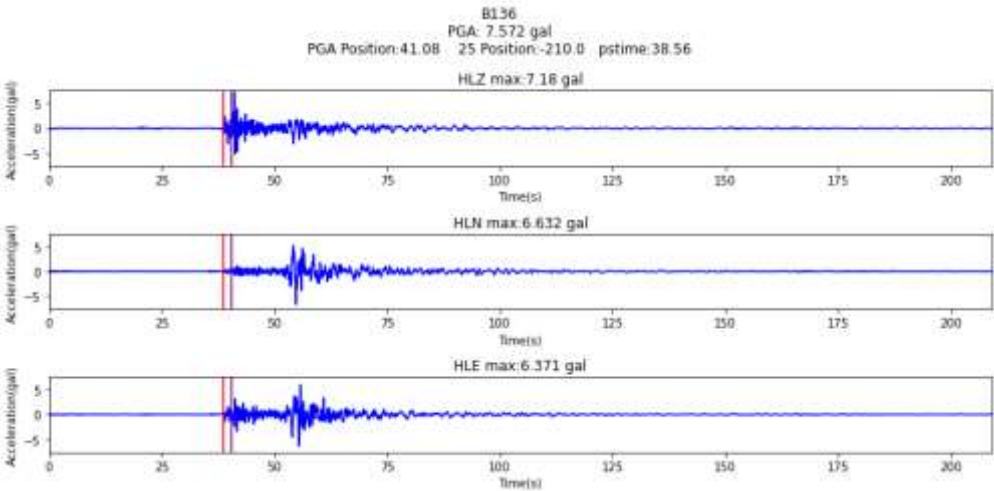
附錄 B：2024 年 1-9 月誤報之測站波形圖



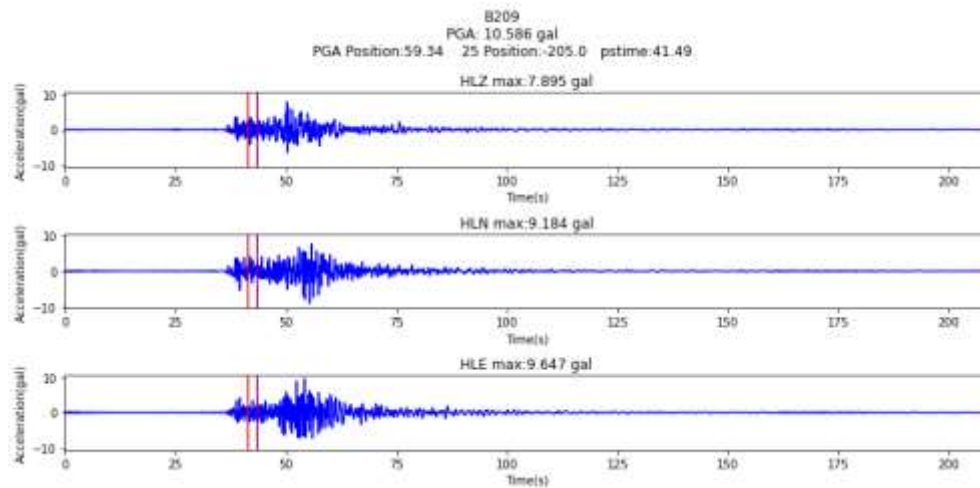
20240423(編號 267)B085 測站波形圖



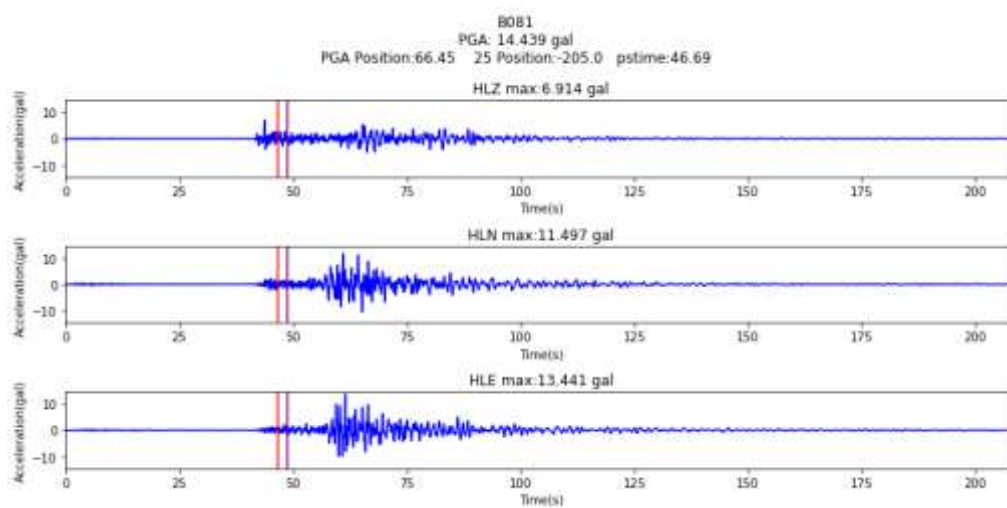
20240423(編號 267)B039 測站波形圖



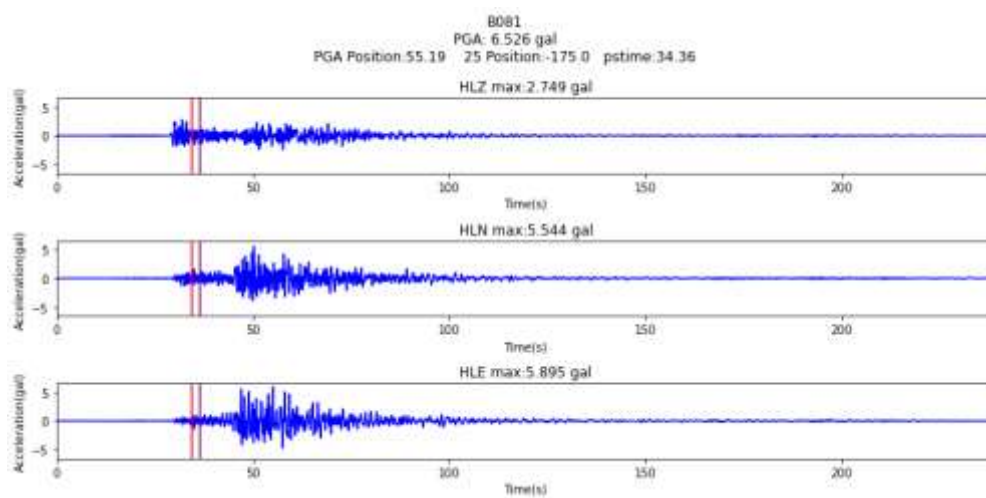
20240423(編號 267)B136 測站波形圖



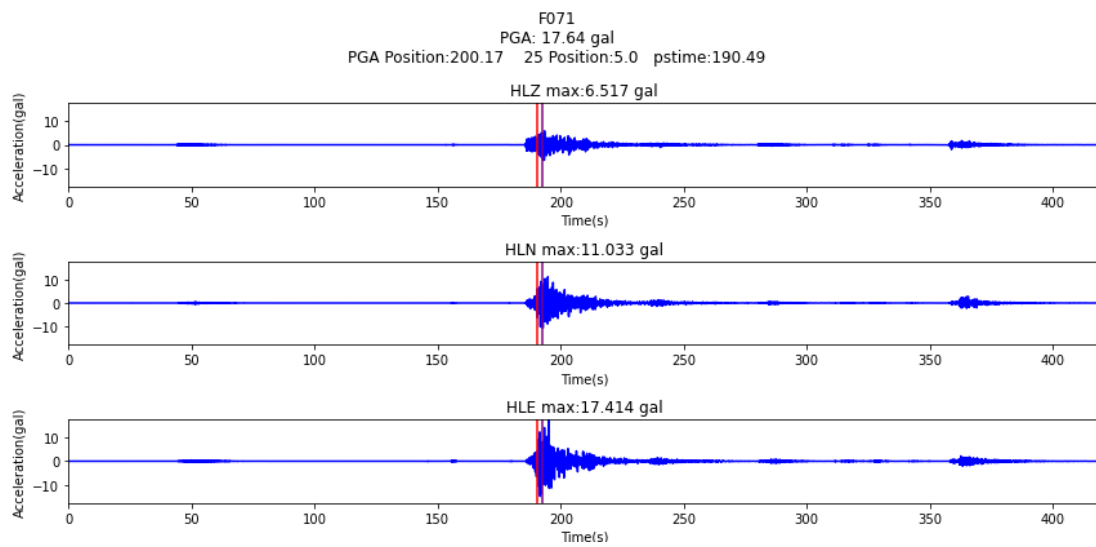
20240423(編號 267)B209 測站波形圖



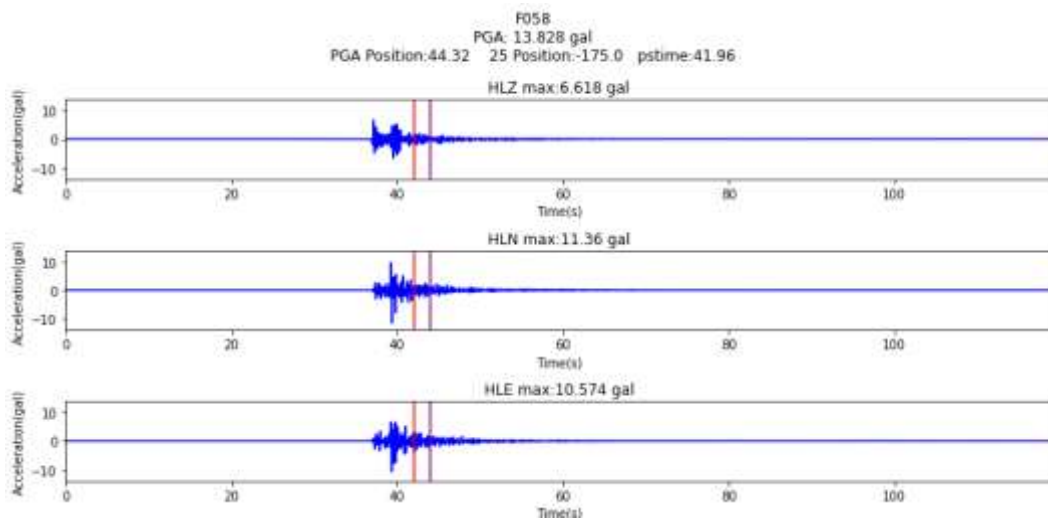
20240423(編號 267)B081 測站波形圖



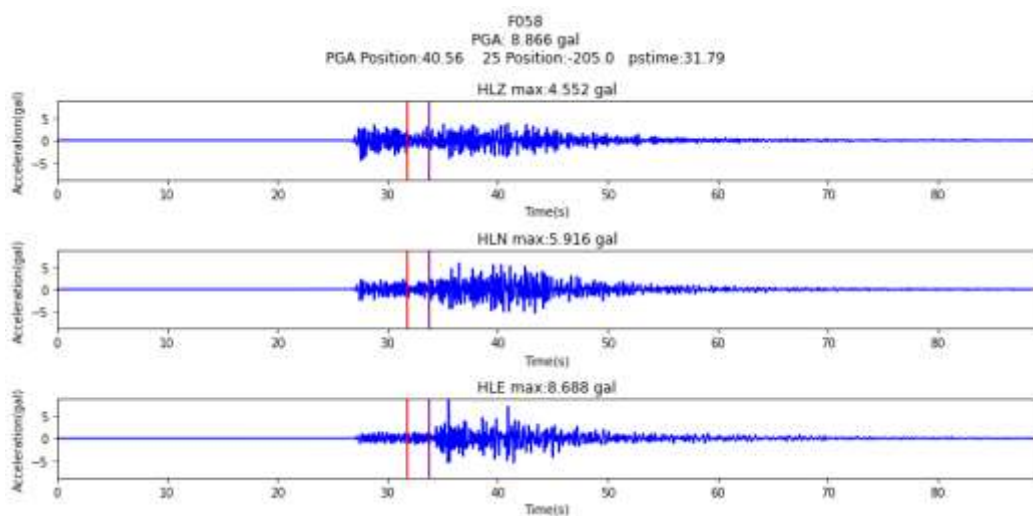
20240423(編號 307)B081 測站波形圖



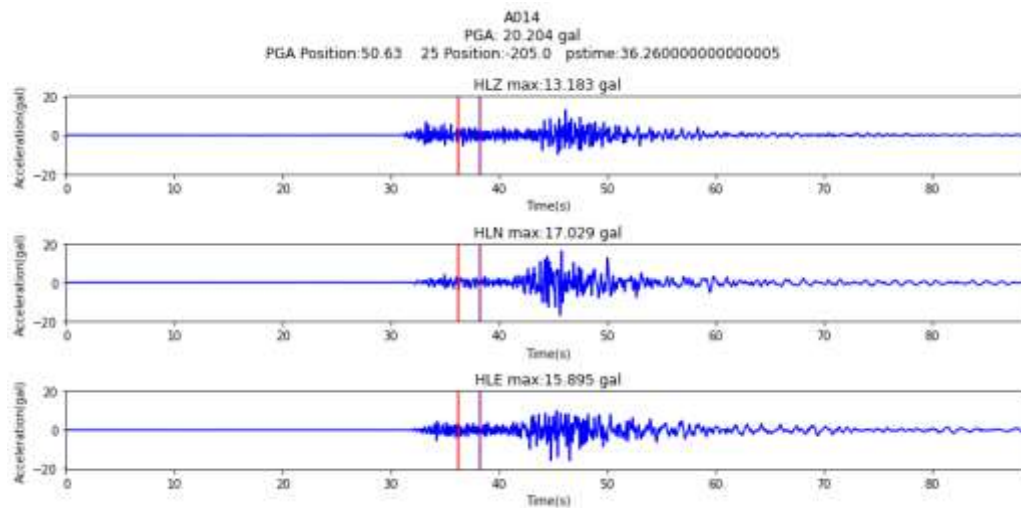
20240423(編號 318)F071 測站波形圖



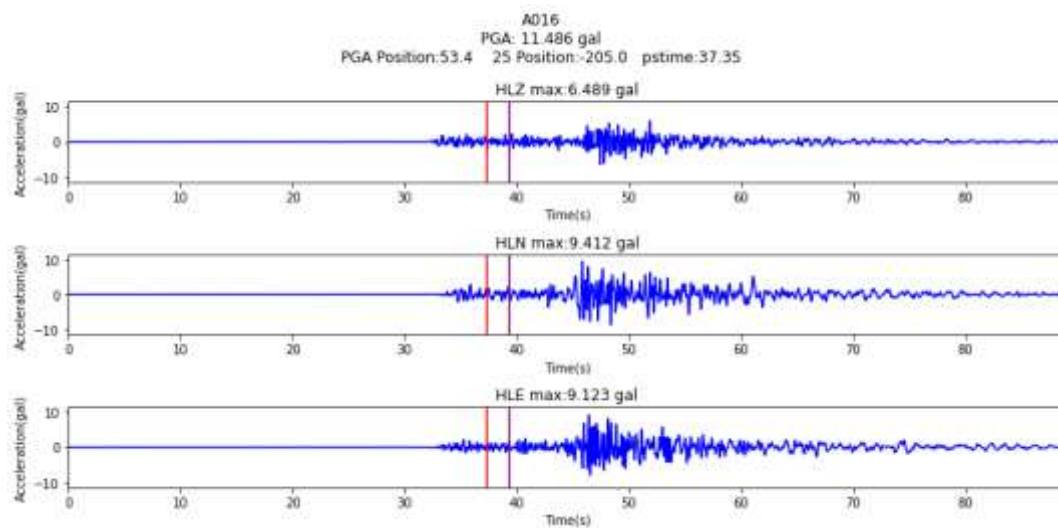
20240425(編號 341)F058 測站波形圖



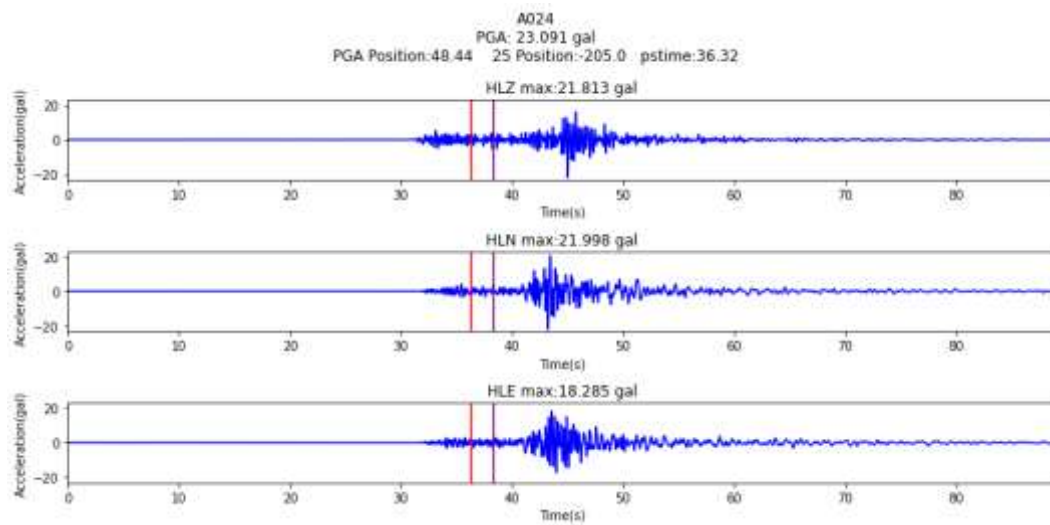
20240815(編號 442)F058 測站波形圖



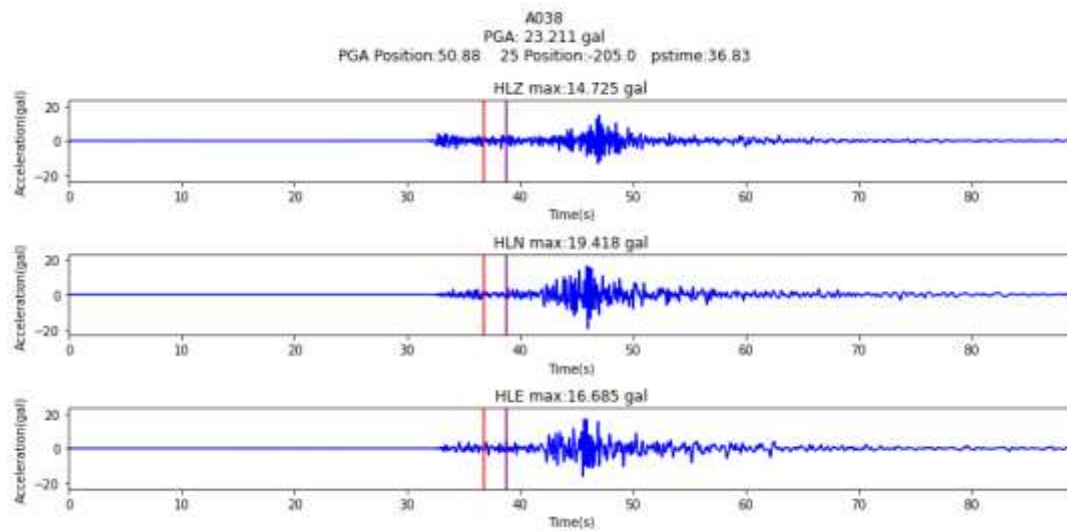
20240815(編號 442)A014 測站波形圖



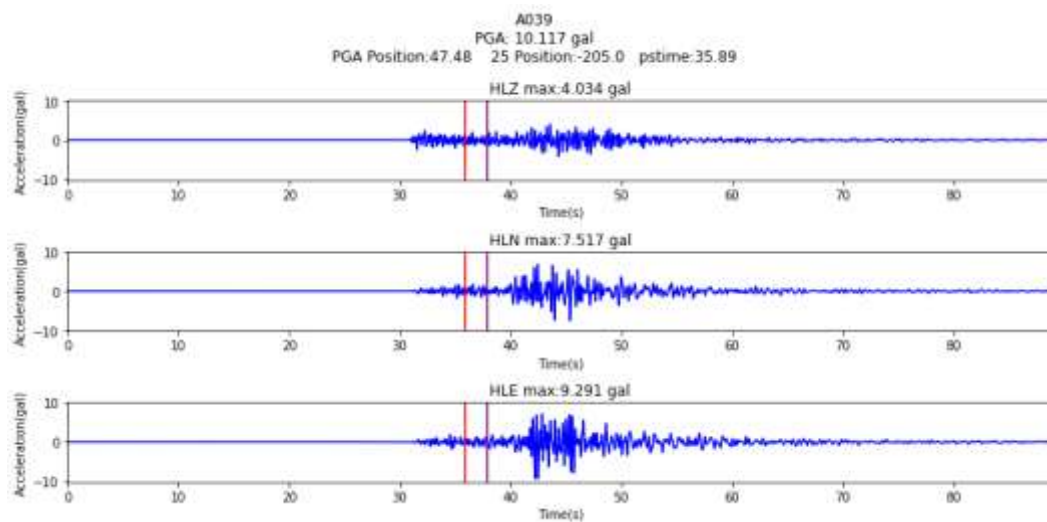
20240815(編號 442)A016 測站波形圖



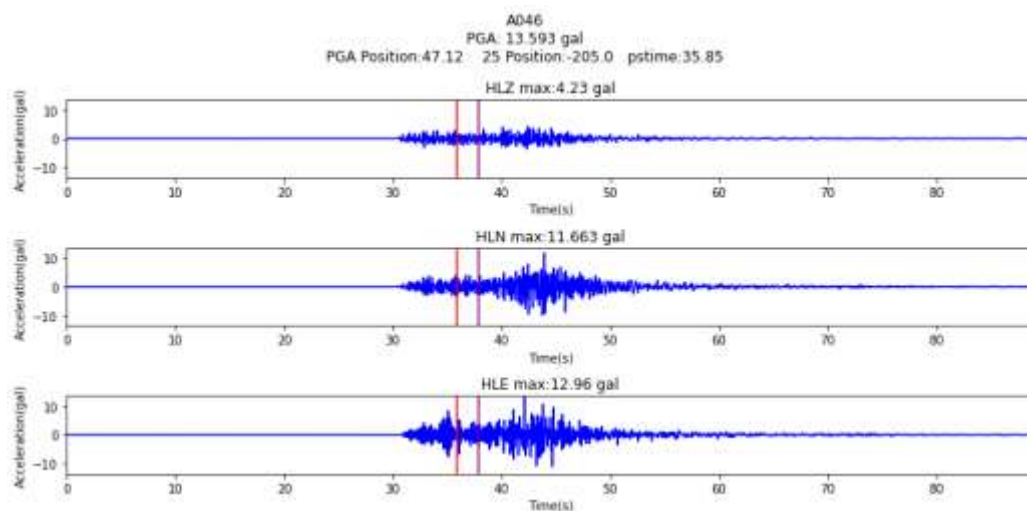
20240815(編號 442)A024 測站波形圖



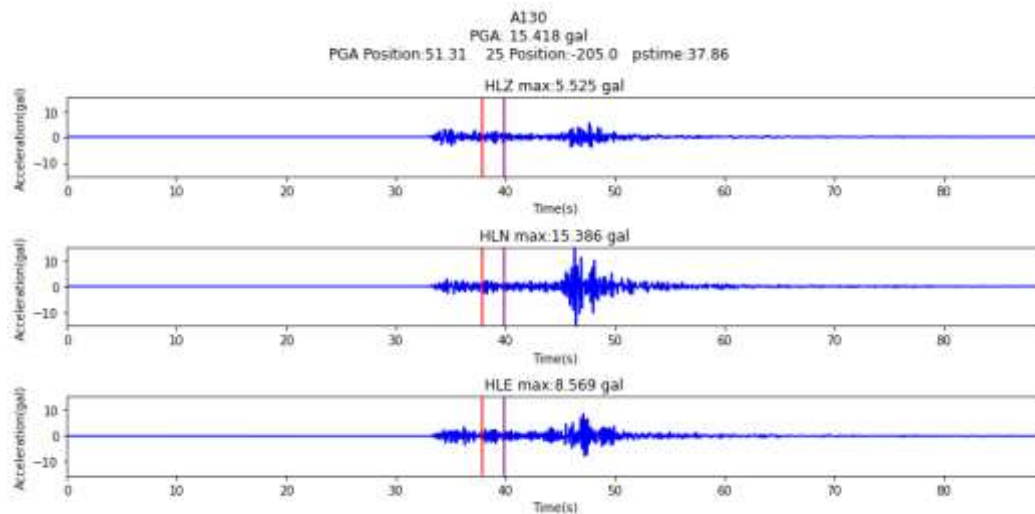
20240815(編號 442)A038 測站波形圖



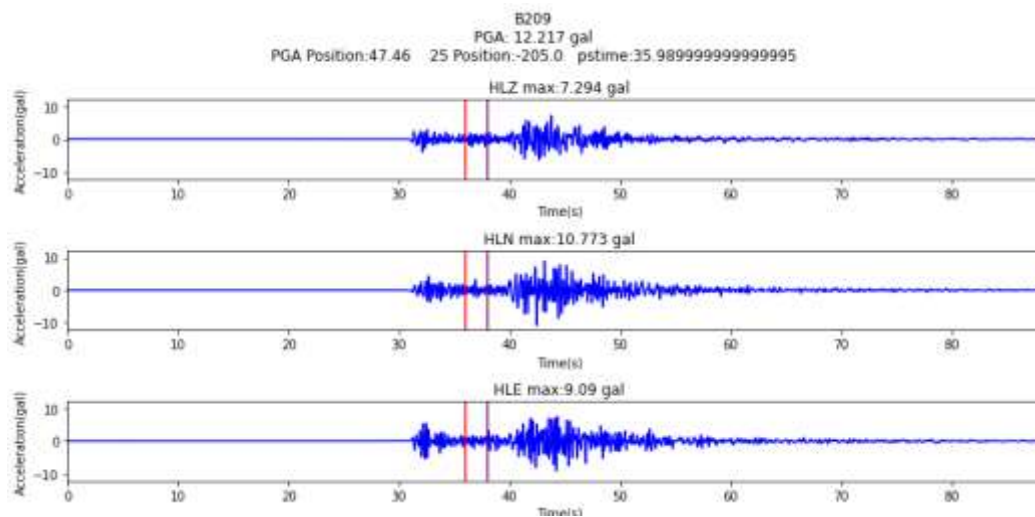
20240815(編號 442)A039 測站波形圖



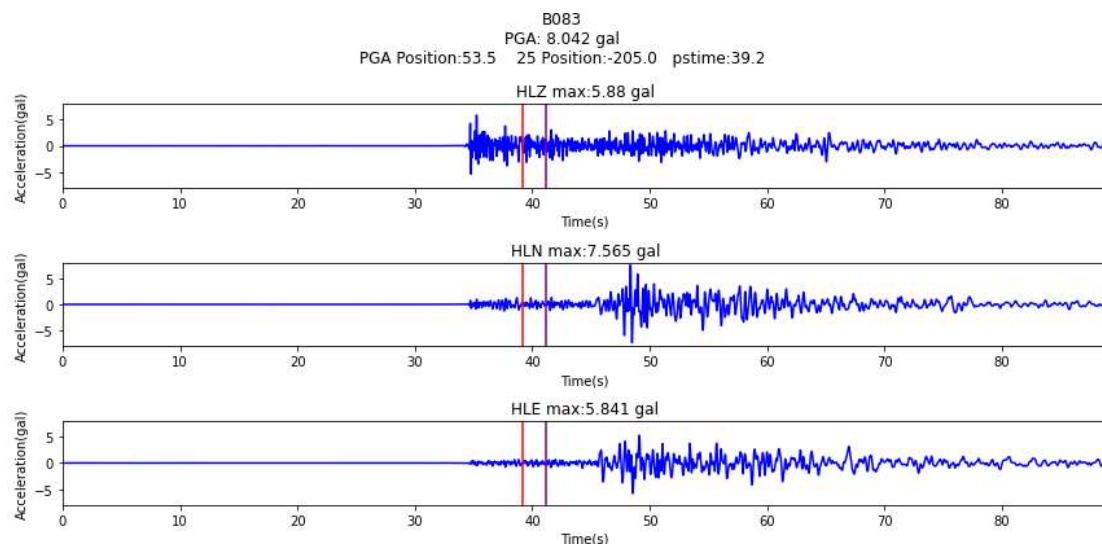
20240815(編號 442)A046 測站波形圖



20240815(編號 442)A130 測站波形圖



20240815(編號 442)B209 測站波形圖



20240815(編號 442)B083 測站波形圖

附錄 C：所有測站 P 波到達時間與 PGA 超過 25 gal 時刻表

Station	Longitude	Latitude	PGA	P-Wave Arrival Time	PGA >= 25 Time	Time Difference (s)
EYL	121.6	23.9	595.79	07:58:14.09	07:58:15.05	0.96
SHUL	121.56	23.79	271.65	07:58:14.42	07:58:18.09	3.67
ETM	121.49	23.97	493.85	07:58:14.61	07:58:15.02	0.41
ESL	121.44	23.81	227.7	07:58:14.65	07:58:15.46	0.81
HWA	121.61	23.98	428.76	07:58:14.70	07:58:16.05	1.35
EGC	121.55	23.71	547.2	07:58:15.31	07:58:20.55	5.24
TWD	121.6	24.08	275.45	07:58:15.67	07:58:17.85	2.18
EGFH	121.43	23.67	219.64	07:58:15.94	07:58:18.97	3.03
ETL	121.62	24.16	1049.9	07:58:16.76	07:58:18.32	1.56
ETLH	121.48	24.21	472.82	07:58:17.54	07:58:18.81	1.27
EHY	121.33	23.5	86.54	07:58:18.34	07:58:26.49	8.15
EHYH	121.35	23.49	188.71	07:58:18.37	07:58:25.31	6.94
WHF	121.27	24.14	344.15	07:58:18.50	07:58:21.51	3.01
EHP	121.75	24.31	360.71	07:58:19.68	07:58:23.51	3.83
EAH	121.74	24.33	697.85	07:58:19.97	07:58:22.78	2.81
FUSS	121.24	24.25	123.88	07:58:20.26	07:58:25.13	4.87
TWF1	121.31	23.35	26.25	07:58:20.46	07:58:45.06	24.6
EYUL	121.32	23.35	126.93	07:58:20.80	07:58:27.62	6.82
TWT	121.16	24.25	154.31	07:58:21.26	07:58:26.91	5.65
ENA	121.75	24.43	401.3	07:58:21.48	07:58:24.46	2.98
ECB	121.45	23.32	138.54	07:58:21.50	07:58:32.99	11.49
WPL	120.95	24.01	59.09	07:58:21.74	07:58:33.48	11.74
NNS	121.38	24.44	306.35	07:58:21.94	07:58:28.56	6.62
EWT	121.78	24.45	676.77	07:58:21.94	07:58:25.23	3.29
SML	120.91	23.88	159.19	07:58:22.47	07:58:32.79	10.32
WCS	120.91	24.06	115.29	07:58:22.78	07:58:32.49	9.71
TYC	120.87	23.91	27.95	07:58:23.01	07:58:47.99	24.98
YUS	120.96	23.49	182.86	07:58:23.45	07:58:31.09	7.64

Station	Longitude	Latitude	PGA	P-Wave Arrival Time	PGA >= 25 Time	Time Difference (s)
WHY	120.85	23.7	80.86	07:58:23.50	07:58:37.31	13.81
CHKH	121.4	23.19	162.53	07:58:23.50	07:58:30.51	7.01
EHD	121.21	23.15	107.4	07:58:23.52	07:58:31.13	7.61
CHK	121.37	23.1	111.41	07:58:23.97	07:58:39.27	15.3
WHP	120.95	24.28	60.5	07:58:24.04	07:58:38.38	14.34
NDT	121.51	24.6	98.62	07:58:24.08	07:58:35.40	11.32
ESA	121.84	24.58	205.97	07:58:24.19	07:58:31.53	7.34
NDS	121.72	24.63	317.52	07:58:24.61	07:58:29.64	5.03
ENT	121.57	24.64	185.8	07:58:24.63	07:58:30.13	5.5
ECS	121.22	23.1	84.13	07:58:24.71	07:58:44.62	19.91
ELD	121.03	23.19	48.16	07:58:24.85	07:58:43.19	18.34
TWC	121.86	24.61	148.66	07:58:24.86	07:58:35.27	10.41
ALS	120.81	23.51	67.57	07:58:25.63	07:58:39.40	13.77
NSK	121.37	24.67	94.85	07:58:25.74	07:58:34.61	8.87
TWE	121.68	24.72	115.18	07:58:25.87	07:58:36.34	10.47
WJS	120.73	23.82	121.74	07:58:26.01	07:58:39.14	13.13
EDH	121.3	22.97	190.43	07:58:26.42	07:58:34.11	7.69
WNT	120.69	23.88	54.97	07:58:26.75	07:58:42.08	15.33
NFF	121.12	24.63	127.8	07:58:26.76	07:58:36.89	10.13
WNT1	120.68	23.91	61.16	07:58:26.95	07:58:39.84	12.89
ILA	121.76	24.76	126.55	07:58:27.00	07:58:37.56	10.56
CHN5	120.68	23.6	90.04	07:58:27.19	07:58:38.45	11.26
LONT	121.13	22.91	74.94	07:58:27.55	07:58:45.50	17.95
TWQ1	120.77	24.35	84.27	07:58:27.57	07:58:41.65	14.08
TCU	120.68	24.15	56.88	07:58:27.84	07:58:40.93	13.09
NST	121.01	24.63	82.09	07:58:27.98	07:58:40.69	12.71
WYL	120.58	23.96	92.89	07:58:28.38	07:58:40.25	11.87
TWG	121.08	22.82	23.02	07:58:28.57	-	-
STYH	120.78	23.18	68	07:58:28.68	07:58:41.43	12.75

Station	Longitude	Latitude	PGA	P-Wave Arrival Time	PGA >= 25 Time	Time Difference (s)
WGK	120.57	23.68	88.5	07:58:28.76	07:58:42.21	13.45
NSY	120.77	24.41	86.85	07:58:28.91	07:58:41.79	12.88
WDLH	120.54	23.69	181.59	07:58:28.93	07:58:38.98	10.05
WCHH	120.56	24.08	302.34	07:58:28.94	07:58:35.87	6.93
KSHI	121.18	24.78	129.85	07:58:28.95	07:58:38.38	9.43
WDL	120.54	23.71	184.67	07:58:28.97	07:58:39.03	10.06
NJD	121.09	24.74	177.68	07:58:28.98	07:58:36.97	7.99
STY	120.77	23.16	28.67	07:58:29.13	07:59:01.09	31.96
WCH1	120.55	24.07	216.31	07:58:29.19	07:58:37.96	8.77
WCKO	120.6	23.44	262.32	07:58:29.28	07:58:39.55	10.27
EGS	121.94	24.84	53.59	07:58:29.37	07:58:50.02	20.65
WWC	120.52	24.26	59.78	07:58:29.57	07:58:43.79	14.22
NMLH	120.79	24.54	391.14	07:58:29.58	07:58:36.64	7.06
WDJ	120.64	24.35	42.51	07:58:29.61	07:58:54.76	25.15
TTN	121.15	22.75	21.83	07:58:30.07	-	-
CHN4	120.59	23.35	73.29	07:58:30.08	07:58:47.88	17.8
TWA	121.59	24.98	67.9	07:58:30.39	07:58:51.47	21.08
NHDH	121.53	24.96	122.75	07:58:30.39	07:58:39.96	9.57
HSN1	121.02	24.78	76.96	07:58:30.43	07:58:41.81	11.38
WTP	120.62	23.24	29.04	07:58:30.44	07:58:59.85	29.41
NJN	120.87	24.68	114.85	07:58:30.59	07:58:42.50	11.91
CHN2	120.47	23.53	99.69	07:58:30.69	07:58:45.33	14.64
WTK	120.39	23.69	156.96	07:58:30.82	07:58:37.78	6.96
WRL	120.38	23.9	179.88	07:58:31.05	07:58:37.44	6.39
CHY	120.43	23.5	67.96	07:58:31.27	07:58:50.71	19.44
SNS	120.5	23.22	22.65	07:58:31.35	-	-
HSN	121.01	24.83	106.06	07:58:31.43	07:58:43.94	12.51
TAP	121.51	25.04	84.45	07:58:31.65	07:58:43.78	12.13
NCU	121.19	24.97	94.16	07:58:31.70	07:58:42.15	10.45

Station	Longitude	Latitude	PGA	P-Wave Arrival Time	PGA >= 25 Time	Time Difference (s)
NHY	121.57	25.04	100.95	07:58:31.76	07:58:42.58	10.82
ECL	120.96	22.6	13.35	07:58:31.83	-	-
NTY	121.3	25	114.85	07:58:31.86	07:58:40.85	8.99
TWB1	122	25.01	73.41	07:58:31.86	07:58:53.87	22.01
CHN1	120.53	23.18	24.16	07:58:32.05	-	-
WTC	120.29	23.86	80.17	07:58:32.08	07:58:43.05	10.97
NWF	121.78	25.07	177.34	07:58:32.18	07:58:42.94	10.76
SGS	120.59	23.08	30.15	07:58:32.38	07:59:08.86	36.48
NHW	121.05	25.01	83.86	07:58:32.43	07:58:54.41	21.98
SLG	120.65	22.99	29.35	07:58:32.50	07:59:02.56	30.06
ZUZH	121.54	25.16	76.68	07:58:32.63	07:58:54.10	21.47
WML	120.22	23.8	86.52	07:58:32.85	07:58:43.88	11.03
NOU	121.77	25.15	51.28	07:58:32.89	07:58:56.10	23.21
WSF	120.23	23.64	81.71	07:58:32.95	07:58:49.71	16.76
TWS1	121.42	25.1	58.13	07:58:33.00	07:58:54.34	21.34
ANP	121.53	25.18	23.94	07:58:33.22	-	-
WSL	120.23	23.52	121.12	07:58:33.59	07:58:44.60	11.01
ICHU	120.28	23.36	80.69	07:58:33.70	07:58:45.37	11.67
NTS	121.45	25.16	152.1	07:58:34.10	07:58:43.10	9
SSD	120.64	22.74	18.37	07:58:34.23	-	-
NWR	121.66	25.2	40.91	07:58:34.28	07:58:57.38	23.1
NSM	121.59	25.29	26.95	07:58:34.68	07:59:07.13	32.45
SMG	120.64	22.71	22.42	07:58:34.79	-	-
CHN8	120.22	23.35	48.75	07:58:34.80	07:59:00.10	25.3
TWY	121.61	25.27	25.66	07:58:35.15	07:59:05.35	30.2
TAW	120.9	22.36	7.75	07:58:35.59	-	-
EAS	120.86	22.38	17.42	07:58:35.66	-	-
CHN3	120.36	23.08	43.71	07:58:35.69	07:59:10.00	34.31
SCL	120.2	23.17	45	07:58:35.86	07:59:03.05	27.19

Station	Longitude	Latitude	PGA	P-Wave Arrival Time	PGA >= 25 Time	Time Difference (s)
SSH	120.29	23.14	38.2	07:58:35.86	07:59:02.28	26.42
TAWH	120.89	22.34	19.08	07:58:35.91	-	-
SCS	120.49	22.89	106.74	07:58:35.92	07:58:49.01	13.09
SHH	120.35	23.02	84.87	07:58:36.27	07:58:56.62	20.35
TAI1	120.24	23.04	27.81	07:58:37.18	07:59:21.39	44.21
TAI	120.2	23	16.99	07:58:37.47	-	-
TWM1	120.43	22.82	19.32	07:58:37.82	-	-
LAY	121.56	22.04	33.44	07:58:37.96	07:59:07.71	29.75
SLIU	120.8	22.22	3.99	07:58:38.15	-	-
SGL	120.5	22.72	27.96	07:58:38.62	07:59:28.84	50.22
SCZ	120.63	22.37	8.37	07:58:38.65	-	-
SSP	120.57	22.48	23.29	07:58:38.69	-	-
SPT	120.49	22.68	22.32	07:58:38.74	-	-
SNJ	120.34	22.75	34.35	07:58:38.98	07:59:29.57	50.59
WSS	120.26	22.64	8.38	07:58:40.53	-	-
PCY	122.08	25.63	16.44	07:58:40.98	-	-
WDG	119.67	23.26	7.64	07:58:40.99	-	-
TWK1	120.81	21.94	4.77	07:58:41.71	-	-
SMS	120.84	22.02	21.84	07:58:41.97	-	-
WLC	120.37	22.35	6.5	07:58:43.21	-	-
WLCH	120.38	22.35	6.41	07:58:43.33	-	-
SEB	120.86	21.9	11.07	07:58:43.57	-	-
HEN	120.75	22	19.66	07:58:43.89	-	-
SNW	120.75	21.96	16.67	07:58:45.51	-	-
MSU	119.92	26.17	4.46	07:58:52.38	-	-

機器學習技術建立現地型地震預警系統相關研究(III)

子計畫四

優化地震預警系統：以機器學習技術進行地震規模快速估算

李恩瑞 吳秉芬 蔡文淮 廖勿渝

國立成功大學

摘要

本研究目前將即時地震波形處理流程和已經開發的即時地震波相挑選模型 RED-PAN (Real-Time Earthquake Detection and Phase-Picking With Multitask Attention Network) 架設於中央氣象署新架構的伺服器並進行即時資料處理。目前測試結果流程能近乎即時 (0.6 秒以內) 對 500 個以上測站的即時波形資料進行處理並用 RED-PAN 對地震波相到時更新。在利用多個測站即時地震波形資料，對地震規模快速評估的機器學習模型上，本計畫開發了基於 Fourier Neural Operator (FNO) 的深度學習模型。該模型結合多測站波形特徵與震源距離校正參數，通過時頻域特徵提取，實現地震規模的快速估算。該機器學習模型以 2012 年至 2023 年規模大於 3.5 的地震事件作為模型訓練、驗證和測試。結果顯示，當最佳模型在使用第五個測站收到 P 波後 3 秒的波形進行規模評估時，約 95% 的地震事件評估誤差小於 0.4。對於規模大於 5 的地震事件，模型的預測結果穩定且誤差大多小於 0.5，表現出對大規模地震事件的高精度評估能力。由於模型演算法設計相對簡潔，運算效率極高。測試中，模型的單次運算時間可小於 0.05 秒，證明其具有接近即時的地震規模評估能力，適合於地震預警系統的實時應用。

關鍵詞：地震規模、地震預警系統、地震災害

Abstract

This study integrates the real-time seismic waveform processing workflow and the previously developed real-time seismic phase-picking model, RED-PAN (Real-Time Earthquake Detection and Phase-Picking With Multitask Attention Network), into the new EMQX server architecture of the Central Weather Administration (CWA) for real-time seismic data processing. The testing results demonstrate that the workflow can process real-time waveform data from more than 500 stations and update seismic phase arrivals using RED-PAN within approximately 0.6 seconds, achieving near real-time performance. This study also developed a deep learning model based on the Fourier Neural Operator (FNO) for rapid earthquake magnitude evaluation using real-time seismic waveform data from multiple stations. The model integrates waveform features from multiple stations with source distance correction parameters and utilizes time-frequency domain feature extraction to achieve rapid magnitude estimation. The machine learning model was trained, validated, and tested using seismic recordings of events with magnitudes greater than 3.5 from 2012 to 2023 in the CWA catalog. The results indicate that when using waveform recordings available 3 seconds after the P-wave arrival at the fifth station, the model achieves magnitude errors of less than 0.4 for approximately 95% of the events. For seismic events with magnitudes greater than 5, the model demonstrates stable predictions with most errors below 0.5, highlighting its high accuracy in assessing large-magnitude earthquakes. The simplicity of the model's algorithm design ensures high computational efficiency. Testing shows that the model's single inference time is less than 0.05 seconds, confirming its capability for near real-time earthquake magnitude evaluation, making it well-suited for practical applications in earthquake early warning systems.

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

一、前言

臺灣位於環太平洋地震帶，是全球地震活動最頻繁的地區之一。由於菲律賓海板塊和歐亞板塊的隱沒與弧陸碰撞造山作用(Teng, 1990; Wu et al., 1997)，臺灣擁有複雜的地質結構，並時常發生地震。在臺灣的歷史紀錄中，多次大規模地震造成了重大人員傷亡，且對經濟和基礎設施帶來嚴重損失(Kao, 2000; Chi et al., 2001)。地震災害對都市化人口密集地區的影響特別大，而地震預警系統的建立成為減少地震影響的重要手段。透過地震預警系統，利用地震發生後鄰近測站的即時觀測資料，可迅速估算震源參數，並在預測到潛在災害性地表震動到達前發出警報(Wu et al., 2007; Wu and Kanamori, 2008; Hsiao et al., 2009; Chen et al., 2019)。這些警報能夠為民眾提供幾秒到幾十秒的反應時間，使人們有機會在地震波抵達前採取避險措施，有效減少災害的影響(Allen et al., 2009; Allen and Melgar, 2019)。除了保護個人安全，地震預警系統還可以防護關鍵基礎設施，如橋梁、高速公路、發電廠及鐵路系統等(Gasparini et al., 2007)。此外，地震預警系統能提高公眾對地震風險的認識，增強備災意識。

在地震預警中，地震規模是評估地震能量釋放的重要指標，不僅影響震度估算，亦是決定是否發佈預警的重要參數(Wu and Zhao, 2006; Lin and Wu, 2010; Chen et al., 2017)。地震預警系統通常依賴地震的最早波相（P 波）來估計震源參數與地震規模，並進一步預測破壞性更強的 S 波到達時間及強度，以便快速發出警報，減少地震帶來的損失(Wu and Zhao, 2006; Lin and Wu, 2010; Chen et al., 2017)。目前，許多地震災害頻繁的國家，如臺灣、日本、墨西哥、美國與義大利等，已建立或正在開發各自的地震預警系統(Gasparini et al., 2007)。臺灣的地震預警系統由中央氣象署負責，透過遍布全島的地震測站網絡進行監測，並向公眾提供即時的地震預警服務。儘管測站分布密集，且具有即時訊號回傳能力，但臺灣複雜的斷層系統及部分斷層接近人口密集區，對現有地震預警系統提出了更高的精確性需求。為了進一步降低地震災害對人員及基礎設施的威脅，需要更高效且準確的地震規模快速評估方法。

在地震規模的快速評估中，主要有兩種方法：

1. 基於統計的觀測資料分析：

傳統的統計方法利用早期 P 波數據，如波形振幅、頻率、週期參數（ τ_c ）及位移振幅參數（Pd）等特徵，結合統計回歸建立這些參數與地震規模之間的關係，進行快速規模估算(Wu and Zhao, 2006; Shieh et al., 2008; Lin and Wu, 2010; Lin et al., 2011; Chen et al., 2017; Nazeri et al., 2017)。然而，傳統方法的應用存在一定局限性。首先，這些方法通常針對單個測站的觀測數據進行地震規模的計算，導致不同測站評估的規模數值可能存在差異。此外，傳統方法未充分利用多個測站的觀測數據來進行地震規模的整體評估，而僅局限於單測站的數據特徵。這樣無法有效整合多測站的震幅信息進行全局規模的綜合判斷，可能影響地震規模評估的準確性。

2. 機器學習方法：

隨著深度學習技術的發展，機器學習方法開始在地震規模評估中展現出優勢(Mousavi and Beroza, 2020; Kuang et al., 2021; Zhang et al., 2021)。機器學習模型（如卷積神經網絡 CNN 和支持向量機 SVM）可以自動從地震波形數據中提取特徵，尋找其與地震規模的潛在關係，並結合多測站數據進行綜合分析，提供更快速且準確的規模估計。

本研究主要在運用機器學習技術，進一步提升地震規模評估的速度與準確性。我們將結合已開發的深度學習模型 RED-PAN (Real-Time Earthquake Detection and Phase-Picking With Multitask Attention Network) (Liao et al., 2022)，利用即時地震波形數據進行 P 波挑選，並結合多測站的 P 波到時後的波形資料，進行快速且

精確的地震規模評估。本研究期望能有效增強地震預警系統的地震規模評估的速度與準確性，以提高地震預警系統的整體效能，減少地震災害的影響。

二、前期開發機器學習模型部署

中央氣象署原有的地震預警系統使用開源軟體 Earthworm，託管全自動化的地震演算法流程。然而，隨著近年來通訊協定和傳輸系統的快速發展，中央氣象署於 2024 年中旬更新了資料通訊軟體架構，改用 **Message Queuing Telemetry Transport (MQTT)** 協定進行資料傳輸。MQTT 是一種輕量級訊息傳輸協議，特別適合於網路環境受限的設備通訊，其採用資料發佈者-訂閱者模式進行解耦操作。訊息代理程式(broker)作為第三方元件，負責篩選和分發來自發佈者的訊息給訂閱者。在此架構下，將地震數據傳輸至 MQTT broker 伺服器，而微型服務(microservices)作為訂閱者，透過 MQTT 的 API（如 Python 或 C 語言）訂閱特定測站的資料。此外，這些微型服務還可作為資料發佈者，將產製的資訊回傳至 MQTT broker，供其他微型服務共享使用。這種架構不僅提升了不同演算法之間的資訊共享效率，還能實時交換地震波形、波相到時及震源位置等數據，支援即時地震處理。

本研究所設計的地震預警系統以微型服務架構為核心，將不同的地震演算法模組部署為獨立的微型服務（圖 1）。每個微型服務的輸出會暫時儲存於共享的網路通訊系統（如 MQTT 或 Apache Kafka）中，供其他服務使用，達成高效的資訊共享。系統運行包含以下主要步驟：

1. 從中央氣象署的 MQTT broker 接收所有測站的即時波形資料封包。
2. 將波形資料依測站合併為連續的波形紀錄。
3. 利用基於機器學習的地震波相挑選模型(RED-PAN)，挑選出各測站的 P 波。
4. 根據挑選的 P 波進行地震震源的初步定位。

中央氣象署委託中保防災科技股份有限公司建置的資訊傳輸系統基於 EMQX 公司的 MQTT broker，專注於物聯網架構中的訊息傳遞與處理。同時，本系統的微型服務之間的主要通訊則採用 Apache Kafka。相較於 MQTT，Kafka 更適合於大規模即時數據流的處理，具備分散式持久化數據儲存與緩存的功能，有利於系統的擴展和維護。

本系統使用 NVIDIA Triton Server 部署深度學習模型，提供統一的模型 API 支援（圖 1）。不論是基於 TensorFlow 或 PyTorch 開發的模型，都可以轉換並壓縮至 ONNX 格式，以便使用 NVIDIA Triton Server 自動調度 GPU 叢集資源進行高效推論。GPU 的運算效率遠高於 CPU，特別是在地震預警系統中，能夠顯著縮短

處理時間。氣象署目前擁有兩台伺服器，其中一台配備 4 張 NVIDIA A100 GPU，另一台配備 2 張入門級推論 GPU NVIDIA A2。此外，氣象署規劃將外部人員開發的地震預警演算法部署在一台僅有 CPU 的伺服器上，並透過 HTTP/gRPC 協定與 GPU 伺服器進行聯繫，完成模型推論工作。為了實現系統的可擴展性與長期維護，本系統的所有微型服務均已容器化，並納入中央氣象署內部搭建的 Kubernetes 叢集管理。這種設計允許未來開發的新演算法以容器化形式部署，直接訂閱 Kubernetes 中的 MQTT broker 或 Kafka broker 的共享資訊即可運作，無需停止其他已部署的微型服務，確保系統運行的穩定性與靈活性。

本系統於 2024 年 9 月 18 日 18:00 至 9 月 19 日 09:20 期間，利用中央氣象署的硬體設施進行運算測試。在資料採集階段，共接收到來自 506 個地震測站的資料封包，以 0.5 秒的時間間隔進行波相挑選。測試中共進行了近 40 萬次模型推論，推論階段的時間中位數為 0.3 秒，後處理階段的時間中位數為 0.16 秒，整體符合即時運算的需求（圖 2）。然而，在部分情況下，推論和後處理的總時間超過了 0.5 秒，產生的延遲可能與網路傳輸問題或伺服器其他程式造成的 CPU 滿載有關，需進一步分析與優化（圖 2）。

三、資料和資料處理流程

在進行近即時地震規模評估時，本研究聯合多個地震測站的波形數據，包含中央氣象署(CWA)地震觀測網(CWASN)、臺灣強地動觀測網(TSMIP)和臺灣寬頻地震網(BATS)提供的加速度資料（圖 3）。所選地震事件的規模以中央氣象署公佈的地震目錄為基準，範圍涵蓋 2012 年至 2023 年 7 月的地震事件，且規模大於 3.5。為確保地震事件具足夠的觀測覆蓋，僅納入波形資料來自 10 個以上測站的地震事件，共篩選出約 5,700 筆地震事件（圖 4），用於模型訓練與測試中規模以上地震規模的評估能力。

在數據處理方面，我們採用深度學習模型 RED-PAN (Real-Time Earthquake Detection and Phase-Picking With Multitask Attention Network) (Liao)對地震波形進行 P 波挑選，並將地震儀原始振幅資料(counts)轉換為物理量單位。為模擬即時地震預警情境，選定第 5 個測站的 P 波被偵測時作為規模評估的起始時間。波形處理包括去除平均值、去趨勢，並在 P 波到時的波形上應用單邊 Hanning taper 將 P 波前的波形資料設為 0，避免因前一地震或背景噪音的過大振幅干擾導致規模評估誤差。根據研究(Mousavi and Beroza, 2020)，地震波形的振幅大小對規模評估至關重要，因此本研究保留波形的原始物理量振幅，未進行正規化處理。

為準備機器學習模型的訓練和驗證資料，我們將同一地震的多個測站資料組合，提取測站與震源距離最近前 10 個測站的波形，以第 5 個測站的 P 波到時作為起始預測點。為模擬即時輸入，每隔 1 秒將波形向前位移，並生成 10 組時間間隔位移 1 秒波形切片資料。如果測站尚未偵測到 P 波，對應的輸入數據設為 0，以避免噪音干擾模型判斷（圖 5）。

為了平衡資料集並提高模型的測試可信度，地震事件依照規模分為 3 至 8 共 5 個範圍，再將每組資料隨機分成訓練、驗證與測試資料集，比例為 8:1:1。最後，合併三種資料集以構成規模平衡的最終資料集，用於模型的訓練、驗證與測試。

四、方法

本研究針對聯合多測站地震波形進行近即時地震規模評估，設計結合 Fourier Neural Operator (FNO)與長短期記憶模型(Long Short-Term Memory, LSTM)的機器學習架構。此模型旨在有效提取地震波形的多維特徵，並結合震源距離校正參數，透過學習波形的時序特性，達成快速且準確的地震規模評估（圖 6）。

模型的輸入為一個三維張量，包含 3 個波形分量（東西向、南北向、垂直向），來自 10 個測站的波形數據，以及每個測站 30 秒的波形（對應 3000 個數據點，採樣率 100 Hz）。多測站波形首先進入三層 FNO1d 模塊(Sun *et al.*, 2023)進行特徵提取。FNO 模塊設計了兩條並行路徑：一是透過傅立葉轉換將波形轉換至頻率域進行特徵學習，學習完成後再回到時間域；二是透過 1×1 卷積核的卷積層提取時間域特徵。這兩條路徑的結果最終相加，使模型能夠同時學習波形的頻率域與時間域特徵，進一步提升對地震波形的特徵的提取能力。

震源距離資料作為模型的校正參數，經過單層單神經元的線性轉換後生成校正項。此設計參考傳統地震規模計算方法中的距離校正公式，用以修正不同測站距離對波形特徵的影響。經 FNO 處理的波形特徵與距離校正參數相乘，使最終輸入的波形同時具備時序性與規模相關的物理信息（圖 6）。

經過校正的波形特徵接著輸入雙向 LSTM 模型進行處理。LSTM 能夠捕捉波形前後的時序特徵，進一步學習波形中隱含的地震規模特徵，並在保留關鍵時序信息的基礎上提取重要的規模特性。經 LSTM 處理後的數據進入全連接層(Fully Connected Layer)，最終輸出單一的地震規模評估值（圖 6）。

整體而言，此模型架構有效結合了 FNO 與 LSTM 的優勢，充分考慮地震波形的多維特徵提取，包括頻率域與時間域的學習，以及震源距離對波形特徵的影響。

五、結果

在本研究中，我們將地震波形數據按照 8:1:1 的比例分為訓練集、驗證集和測試集，並採用 L1 Loss 作為模型的損失函數。訓練過程中，批次大小設為 256，優化器選擇 Adam，學習率設定為 0.001。這些參數的設計旨在提高訓練效率，並確保模型具有良好的收斂性。在訓練過程中，自動挑選驗證集損失(loss)最低的模型儲存，作為最終的地震規模即時評估模型。

為測試不同特徵提取方法在地震規模評估中的性能，我們對模型進行了多種架構比較，包括 Fourier Neural Operator (FNO)、卷積神經網絡(CNN)和全連接網絡(FC)。在模型參數數量上，CNN 和 FNO 模型均約為 25 萬個參數，而 FC 模型由於架構特性，包含約 300 萬個參數。當模型對測試集的地震規模評估標準差(STD)

越接近 0，則模型的評估結果越準確。測試結果表明（圖 7），FNO 在所有測試中表現最佳，因此被選為快速地震規模評估的最優模型。

模型訓練完成後，在不同時間長度的地震波形資料下，進行地震規模預測的測試顯示，隨著使用的波形長度增加，模型的預測性能逐漸提升。具體表現在地震規模預測的均方根誤差(Root-Mean-Square Error, RMSE)、平均絕對誤差(Mean Absolute Error, MAE)以及誤差的標準差(Standard Deviation, σ)均呈下降趨勢（表 1）。結果顯示，機器學習模型能有效利用多測站波形隨時間變化的特徵，提升地震規模的預測精度。

圖 8 展示了不同波形時間長度下的地震規模評估分布圖及柱狀圖。當模型使用從第五個測站 P 波到達後的 3 秒內波形資料進行規模評估時，約 95%的地震規模評估誤差小於 0.4。此外，隨著波形時間增加，對於規模大於 5.5 的地震事件，模型的評估結果更加穩定且誤差更小。這表明，延長波形時間有助於提升對較大規模地震的評估精度，這對地震預警系統具有重要的實際應用意義。

在計算效能方面，由於模型的架構設計相對簡潔，運行效率顯著。在使用 CPU (AMD Ryzen 9 3950X 16-Core)的測試中，模型對 6000 組波形數據進行規模評估的平均運行時間為 0.0198 秒，標準差為 0.00143 秒。這意味著，該模型能夠在即時波形資料上實現接近即時的地震規模評估，對實際地震預警應用具有高度的速度優勢和可行性。

六、結論

本年度的研究工作主要聚焦於兩個方面：第一，將前期開發的機器學習模型進行應用，部署於中央氣象署新建構的伺服器上，以實現近乎即時的資料處理與地震波相判斷；第二，利用機器學習方法，針對多測站的即時波形資料進行快速地震規模估算。

在第一部分的測試中，我們成功將系統部署於氣象署伺服器上，針對超過 500 個測站的即時波形資料進行處理，並利用機器學習模型進行波相判斷。結果顯示，在大多數情況下，系統能於 0.6 秒內完成每次更新，達到接近即時的處理效率，表現出穩定的即時數據處理能力。

在第二部分的地震規模評估研究中，我們採用了基於 Fourier Neural Operator (FNO) 的深度學習架構，結合震源距離校正與長短期記憶網絡(Long Short-Term Memory, LSTM)，對地震波形的時序特徵進行提取與分析。測試結果表明，該方法在即時地震規模評估中具有良好的性能，特別是對於大規模地震事件，模型能穩定且準確地進行估算。同時，由於模型設計簡潔，其單次推論時間低於 0.05 秒，展現出極高的計算效率，極適合應用於地震預警系統的實時運行。

未來，我們計畫將本研究開發的機器學習模型進一步整合並部署於中央氣象署系統，進行實際應用測試，驗證模型的穩定性與性能表現。同時，嘗試建構基於機器學習的地震預警系統，以提升地震災害預警的準確性與即時性，減少地震對社會的影響。

表 1 地震規模誤差統計資料。近即時地震規模評估機器學習模型在第 5 個 P 波到達後不同秒數(t)所預測地震規模的均方根誤差 (RMSE)、平均絕對誤差 (MAE) 以及誤差的標準差 (SD)。

T(s)	RMSE	MAE	SD
1	0.247	0.182	0.246
2	0.211	0.158	0.211
3	0.191	0.145	0.191
4	0.181	0.139	0.181
5	0.179	0.138	0.179
6	0.169	0.13	0.169
7	0.17	0.129	0.17
8	0.167	0.129	0.167
9	0.16	0.124	0.16
10	0.173	0.132	0.173

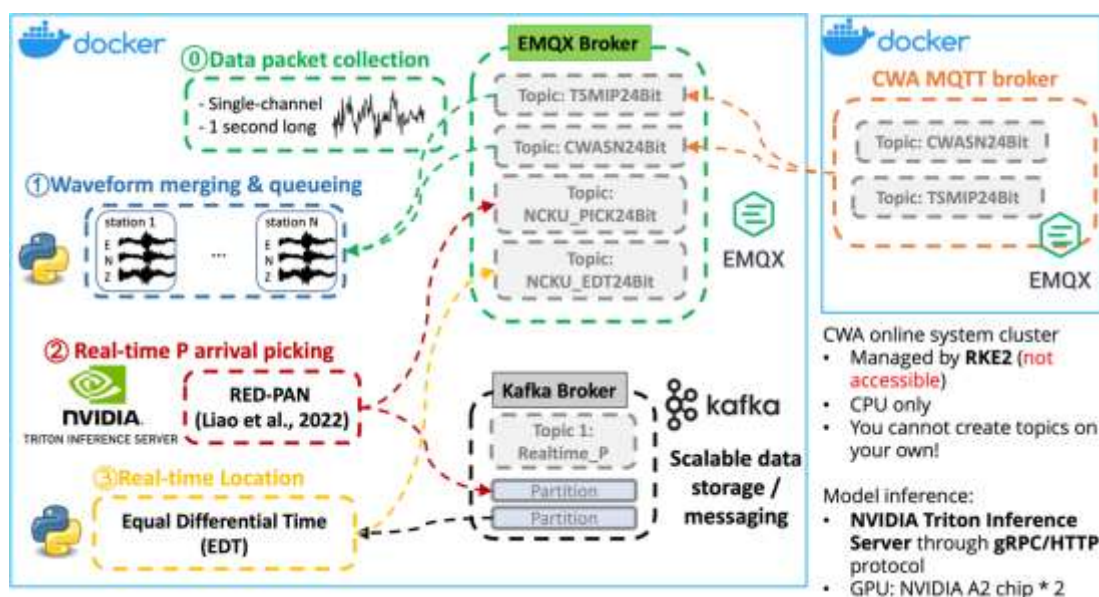


圖 1 地震資料處理流程架構圖。本計畫於氣象署部署氣象署的 EMQX 服務器傳輸架構，以及初期開發的機器學習模型和即時震源定位的資料處理流程。

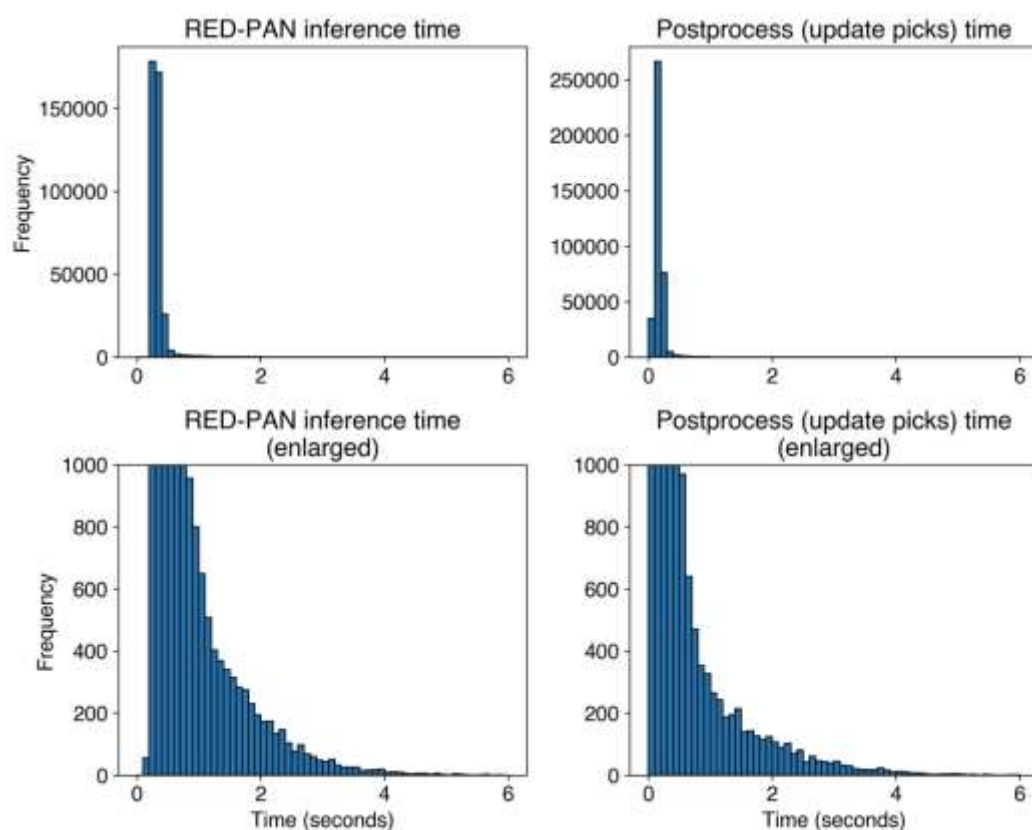


圖 2 RED-PAN 模型波相判識與波相更新處理時間分布圖。RED-PAN 模型在中央氣象署伺服器上對即時波形的波相判識（左圖）及波相更新時間（右圖）的分布情況。上方的柱狀圖為整體處理時間的分布，顯示大多數能在極短時間內完成，表現出良好的效能。下方的柱狀圖為處理時間的放大版，但在少數情況下，處理效能略低，導致部分時間分布出現延遲。

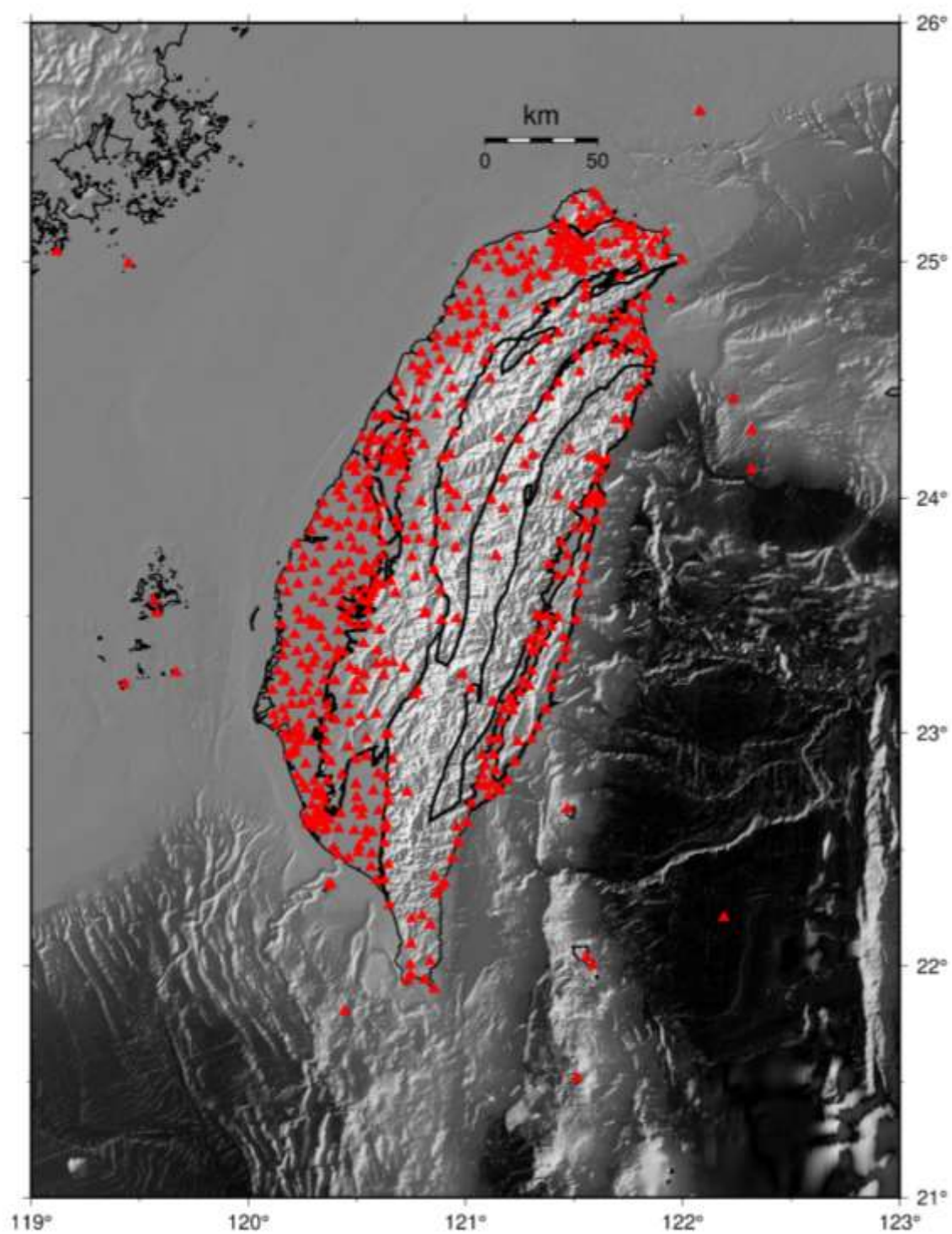


圖 3 近即時地震規模評估使用的地震資料地震測站分布圖。中央氣象署地震觀測網(CWASN)、臺灣強地動觀測網(TSMIP)和臺灣寬頻地震網(BATS)地震測站分布圖。

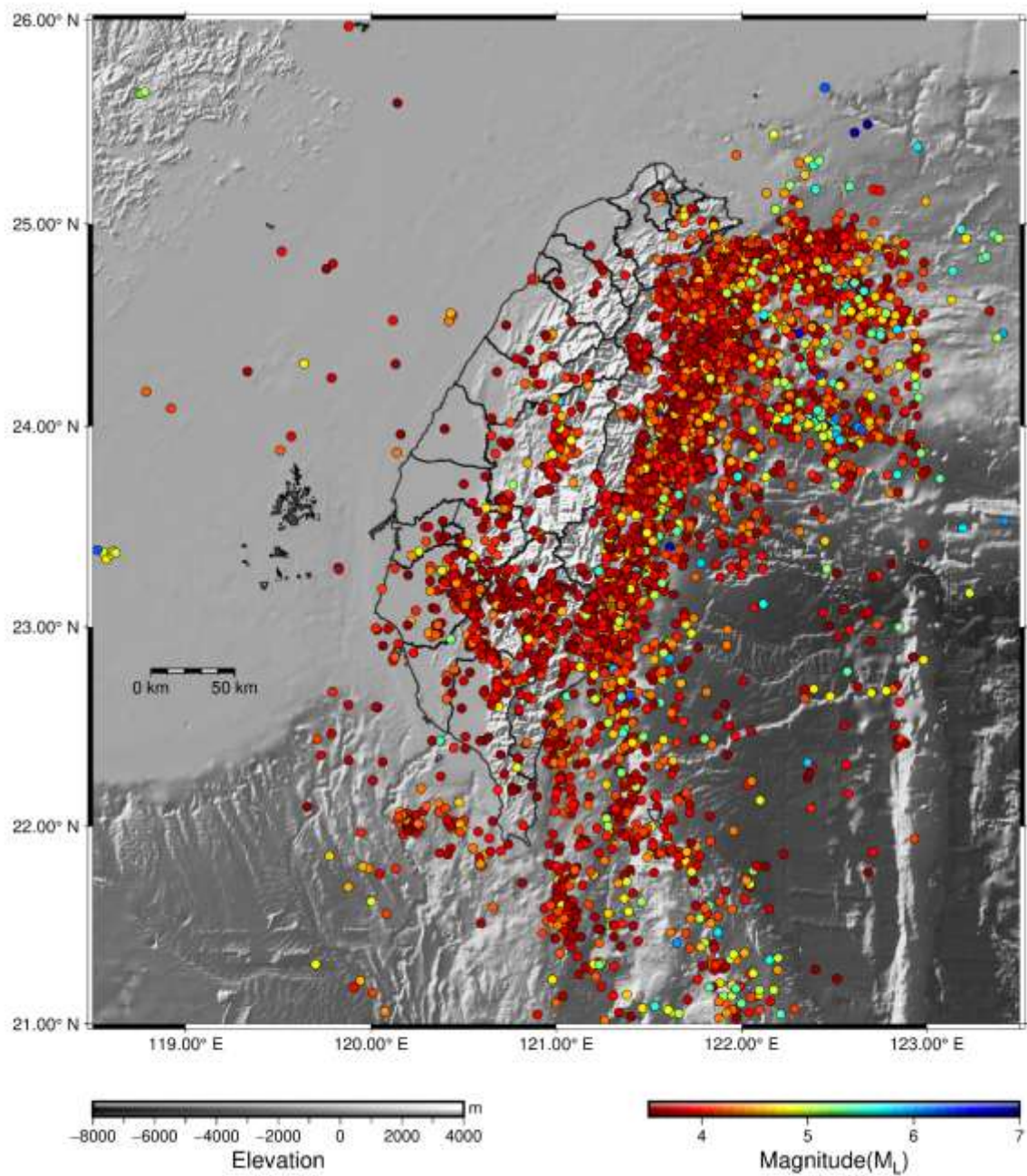


圖 4 近即時地震規模評估使用的地震事件分布圖。圖中顯示 2012 年至 2023 年 7 月間，規模大於 3.5 的地震事件分布，點的顏色表示地震震源深度，展示的地震事件分布涵蓋臺灣主要地震活動區域。



圖 5 資料處理與時間窗切片示意圖。多測站地震波形資料的處理流程及時間窗切片的方式，每個時間窗包括 30 秒的波形資料，並依時間每 1 秒逐步滑動，以捕捉不同時間段的波形特徵。

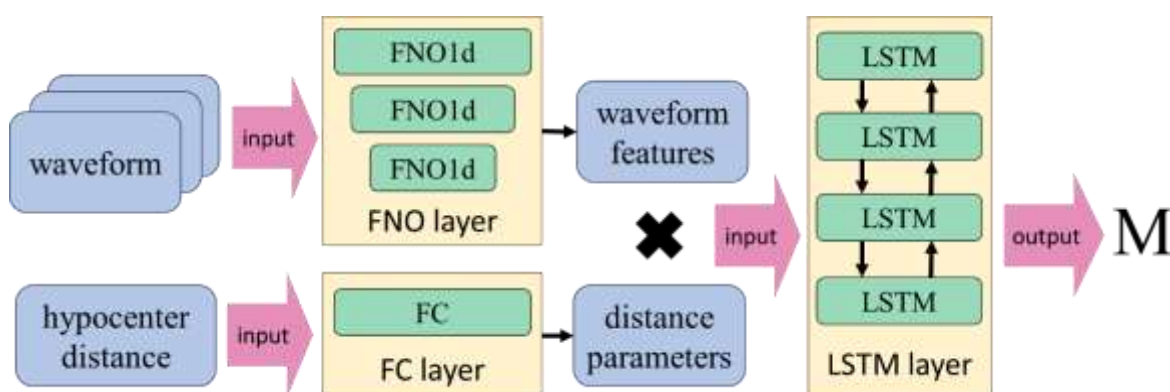


圖 6 地震規模評估模型架構示意圖。模型的輸入包括多測站的地震波形資料和震源距離數據。波形資料經過三層 Fourier Neural Operator (FNO) 模塊提取時頻域特徵，生成波形特徵；震源距離數據則通過全連接層 (FC layer) 轉換為距離校正參數。兩者的輸出相乘後，進入長短期記憶網絡 (Long Short-Term Memory, LSTM)，用於保存並學習地震波形中的時序性規模特徵，最終輸出地震規模 M 。

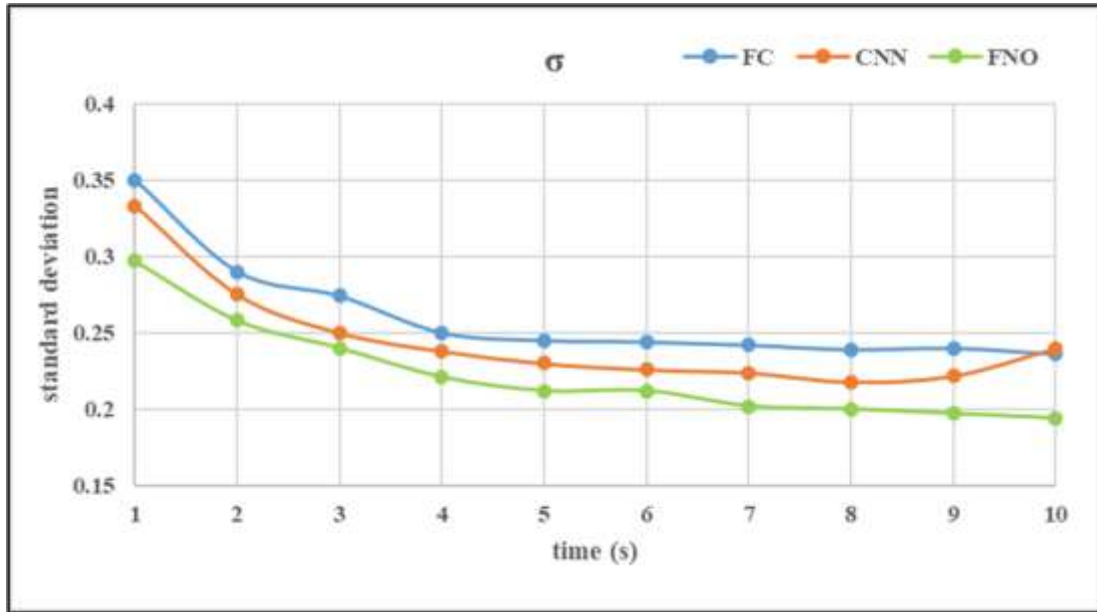


圖 7 不同機器學習模型架構的模型表現比較。此圖展示了全連接網絡(FC)、卷積神經網絡(CNN)和 Fourier Neural Operator (FNO)三種模型在地震規模評估中的性能。橫軸表示從第 5 個測站 P 波到達後的 1 至 10 秒，縱軸為模型的標準差。結果顯示，FNO 模型的標準差始終低於 CNN 和 FC，表現出最佳的穩定性與準確性。

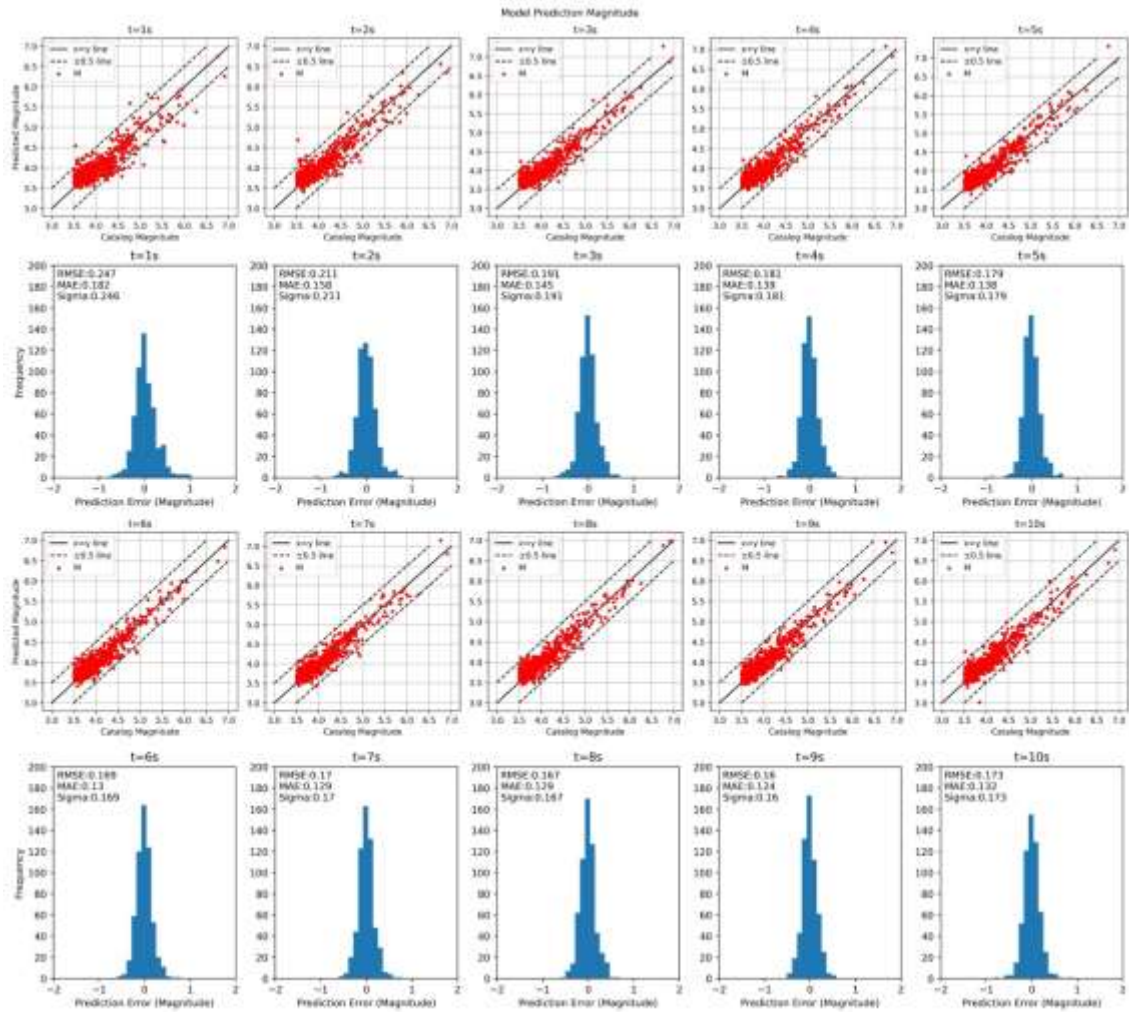


圖 8 地震規模評估結果與誤差分布圖和柱狀圖。此圖展示機器學習模型對地震規模進行評估的結果及誤差分布。上方的散布圖以中央氣象署觀測的地震規模為橫軸（X 軸），模型預測的地震規模為縱軸（Y 軸）。圖中的黑色實線表示理想的一致性（預測值等於觀測值），虛線則表示 0.5 的誤差範圍。下方的柱狀圖展示了模型預測與觀測值之間的誤差分布，橫軸為誤差值，縱軸為頻率。

七、系統部署於中央氣象署伺服器使用說明

本系統目前部署於中央氣象署的一台伺服器中，該伺服器配備 4 片 A100 GPU，系統路徑為 `/home/rtd/NCKU/rpflow_mqtt_docker`。系統相關部署與操作說明文檔位於 `/home/rtd/NCKU/rpflow_mqtt_docker/docs/README.md`，文檔內包含以下內容：

- 系統簡介
- 微型服務的組成架構
- 參數調整方法

系統所需的所有參數由 `/home/rtd/NCKU/rpflow_mqtt_docker/config.py` 文件進行管理與設定。例如，連接 MQTT broker 所需的參數（如 `host`、`port`、`username`、`password` 等）均已記錄在該檔案中。

氣象署內部將即時波形數據傳輸至 MQTT broker 下的以下兩個主題 (topic)：TSMIP24Bit 和 CWASN24Bit。本系統接收這兩個主題的數據，並在即時狀態下完成以下任務：

1. 將挑選出的 P 波結果回傳至 MQTT broker 的主題：NCKU_PICK24Bit。
2. 將快速定位的結果回傳至 MQTT broker 的主題：NCKU_EDT24Bit。

系統啟動說明

按照 `/home/rtd/NCKU/rpflow_mqtt_docker/docs/README.md` 中的指示，進入 `/home/rtd/NCKU/rpflow_mqtt_docker` 目錄，執行以下指令即可啟動服務：

```
$ docker compose up
```

完成上述步驟後，系統服務將成功部署並開始運行。

參考文獻

- Allen, R. M., P. Gasparini, O. Kamigaichi, and M. Böse (2009). The Status of Earthquake Early Warning around the World: An Introductory Overview, *Seismological Research Letters* **80**, no. 5, 682–693, doi: 10.1785/gssrl.80.5.682.
- 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.-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.
- Chen, D.-Y., Y.-M. Wu, and T.-L. Chin (2017). An empirical evolutionary magnitude estimation for early warning of earthquakes, *Journal of Asian Earth Sciences* **135**, 190–197, doi: 10.1016/j.jseaes.2016.12.028.
- Chi, W.-C., D. Dreger, and A. Kaverina (2001). Finite-Source Modeling of the 1999 Taiwan (Chi-Chi) Earthquake Derived from a Dense Strong-Motion Network, *Bulletin of the Seismological Society of America* **91**, no. 5, 1144–1157, doi: 10.1785/0120000732.
- 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.
- Kao, H. (2000). The Chi-Chi Earthquake Sequence: Active, Out-of-Sequence Thrust Faulting in Taiwan, *Science* **288**, no. 5475, 2346–2349, doi: 10.1126/science.288.5475.2346.
- Kuang, W., C. Yuan, and J. Zhang (2021). Network-Based Earthquake Magnitude Determination via Deep Learning, *Seismological Research Letters* **92**, no. 4, 2245–2254, doi: 10.1785/0220200317.
- Lee, E.-J., W.-Y. Liao, G.-W. Lin, P. Chen, D. Mu, and C.-W. Lin (2019). Towards Automated Real-Time Detection and Location of Large-Scale Landslides through Seismic Waveform Back Projection, *Geofluids* **2019**, 1–14, doi: 10.1155/2019/1426019.
- 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.
- Lin, T.-L., and Y.-M. Wu (2010). Magnitude estimation using the covered areas of strong ground motion in earthquake early warning: PGA MAGNITUDE EARLY WARNING, *Geophysical Research Letters* **37**, no. 9, n/a-n/a, doi: 10.1029/2010GL042797.

- Lin, T.-L., Y.-M. Wu, and D.-Y. Chen (2011). Magnitude estimation using initial P-wave amplitude and its spatial distribution in earthquake early warning in Taiwan, *Geophysical Research Letters* **38**, no. 9, n/a-n/a, doi: 10.1029/2011GL047461.
- Mousavi, S. M., and G. C. Beroza (2020). A Machine-Learning Approach for Earthquake Magnitude Estimation, *Geophysical Research Letters* **47**, no. 1, doi: 10.1029/2019GL085976.
- Nazeri, S., Z. H. Shomali, S. Colombelli, L. Elia, and A. Zollo (2017). Magnitude Estimation Based on Integrated Amplitude and Frequency Content of the Initial P Wave in Earthquake Early Warning Applied to Tehran, Iran, *Bulletin of the Seismological Society of America* **107**, no. 3, 1432–1438, doi: 10.1785/0120160380.
- 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.
- Shieh, J.-T., Y.-M. Wu, and R. M. Allen (2008). A comparison of τ_c and $\tau_p^{\max}$ for magnitude estimation in earthquake early warning, *Geophysical Research Letters* **35**, no. 20, doi: 10.1029/2008GL035611.
- Sun, H., Z. E. Ross, W. Zhu, and K. Azizzadenesheli (2023). Phase Neural Operator for Multi-Station Picking of Seismic Arrivals, *Geophysical Research Letters* **50**, no. 24, e2023GL106434, doi: 10.1029/2023GL106434.
- Teng, L. S. (1990). Geotectonic evolution of late Cenozoic arc-continent collision in Taiwan, *Tectonophysics* **183**, no. 1, 57–76.
- Wu, Y.-M., C.-H. Chang, L. Zhao, J. B. H. Shyu, Y.-G. Chen, K. Sieh, and J.-P. Avouac (2007). Seismic tomography of Taiwan: Improved constraints from a dense network of strong motion stations, *Journal of Geophysical Research* **112**, no. B8, doi: 10.1029/2007JB004983.
- 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.
- Wu, F. T., R.-J. Rau, and D. Salzberg (1997). Taiwan orogeny: thin-skinned or lithospheric collision?, *Tectonophysics* **274**, no. 1, 191–220.
- Wu, Y.-M., and L. Zhao (2006). Magnitude estimation using the first three seconds P-wave amplitude in earthquake early warning, *Geophysical Research Letters* **33**, no. 16, doi: 10.1029/2006GL026871.
- Zhang, X., M. Zhang, and X. Tian (2021). Real-Time Earthquake Early Warning With Deep Learning: Application to the 2016 M 6.0 Central Apennines, Italy Earthquake, *Geophys Res Lett* **48**, no. 5, doi: 10.1029/2020GL089394.