

114 年地震前兆觀測作業與分析技術相關研究

子計畫一

地震電磁前兆現象的機器學習與資訊理論分析

陳建志 吳宗義

中華民國地球物理學會

摘要

基於全臺灣大地電磁場觀測系統與 GEMSTIP 和 MagTIP 演算法的整合基礎，本計畫探索將 TIP 概念結合當代機器學習方法的可能性，聚焦於解決測站中心導向預報的侷限，朝向網格單元中心導向的精細預報發展。先前研究顯示，引入資訊理論指標作為輸入特徵可有效增進預報表現；然而，特徵間的共線性問題與時間相依特性尚未被充分理解，限制了這些指標在 TIP 預報中的應用潛力。另一方面，原本的 GEMSTIP/MagTIP 演算法受限於傳統的最佳化方法，在發展精細空間預報上存在計算效能與技術障礙。本期計畫建立地電磁場統計與資訊理論指標的完整自相關性分析框架，透過樣本自相關函數與樣本偏自相關函數探索時間相依特性。研究揭示地磁場觀測中約 27 天的週期性特徵，可能與天體運行週期相關；地電場觀測則呈現突然平移現象，但普遍不具顯著的長期時間相依性。此外，透過 Ridge 迴歸與 PLS 迴歸建立以夏農熵預測費雪訊息的線性模型，成功分離獨立資訊的殘差特徵，為消除共線性提供可行方案。這些發現為未來基於決策樹演算法的特徵選擇與模型開發提供關鍵指引，並為發展高效能、精細化的 TIP 預報奠定基礎，進而促進基於 TIP 的地震電磁前兆研究與災害風險評估的進展。

關鍵字：地磁場異常、地震前兆、資訊分析

Abstract

This project integrates Taiwan's geoelectromagnetic observation system with GEM-STIP/MagTIP algorithms to explore combining TIP concepts with modern machine learning methods, aiming to transition from station-centric to grid cell-centric forecasting. While information theory indicators enhance prediction performance, their collinearity and tempo-ral dependencies require further investigation. Traditional optimization methods in current algorithms also impede fine-grained spatial forecasting development. This study establishes a comprehensive framework for autocorrelation analysis of geoelectromagnetic statistical and information-theoretic indicators, employing SACF and SPACF to examine temporal dependencies. Results reveal a 27-day periodicity in

geomagnetic observations, potentially linked to astronomical cycles, while geoelectric observations exhibit sudden shifts without significant long-term dependencies. Through Ridge and PLS regression, we successfully model Fisher Information from Shannon Entropy, separating independent residual features to address collinearity. These findings provide critical guidance for future decision tree-based feature selection and model development, establishing a foundation for refined TIP forecasting and advancing earthquake electromagnetic precursor research.

keyword : geomagnetic anomalies, earthquake precursor, informational analysis

壹、前言

(一)背景與研究概述

地震事件的「機率增加的時間」(Time of Increased Probability, TIP)概念，首先由 Keilis-Borok VI, Kossobokov VG [1] 於 1990 年代發表的「M8 演算法」(Algorithm M8) 中被提出。M8 演算法隨後被應用於環太平洋地震帶以至於全球範圍內的大型地震之中期預報分析研究，並使得 TIP 成為地震前兆研究中最重要之核心概念之一。TIP 的核心概念是當地震活動的相關參數開始展現出某些偏離歷史觀測常軌的行為，未來的某一段時期將進入地震「機率增加之時間」。在地震學中，TIP 演算法必須依據與地震相關且具有物理連結基礎之因素，透過量化異常現象定義特定時間窗口內的地震風險。

在本團隊先前的研究中，我們以統計及資訊理論指標對地電磁異常進行量化，根據與地震相關的電磁前兆異常計算在未來一段時間、目標範圍內 $ML \geq 5$ 的地震事件發生之機率，發展了 GEMS-MagTIP 演算法 [2-11]。演算法根據地電磁場時序觀測的偏度(skewness, S)、峰度(kurtosis, K)、費雪訊息(Fisher Information, FI)以及夏農熵 (Shannon Entropy, SE)在觀測時間窗內超過門檻的次數，定義未來一段時間是否為「地震機率增加的時間」。其中，預測時間窗(T_{pred})、先導時間(T_{lead})、觀測時間窗(T_{obs})、目標範圍(R_C ；又稱偵測半徑)等參數，是在模型訓練階段經由網格搜尋法(grid-search method)進行優化而得到。此外，GEMS-MagTIP 的運作邏輯為測站中心導向 (station-centric)，強調基於單一測站之觀測在測站周圍的 TIP 計算，並透過多個測站的 TIP 之疊加，實現涵蓋臺灣本島大部分地區之聯合測站機率預報。

然而，目前測站中心導向、使用網格搜尋法的 GEMS-MagTIP 存在著 TIP 在地理空間上的代表性假設問題、空間預報的不均勻性、解析度以及計算效能的限制。首先，測站中心為導向的邏輯假設單一測站偵測到的異常可代表整個 R_C 範圍內的地震活動特徵，此一代表性假設在許多情況下偏離現實，並使得預報結果高度依賴測站分布。此外，在測站偵測半徑範圍內宣告相同 TIP 數值的作法，在技術上難以產生在空間上具同質性¹的預報機率，並在測站覆蓋不足的區域產生預報上的盲區。最後，傳統網格搜尋法的參數最佳化過程是在朝著精細 TIP 預報發展上的關鍵技術性阻礙。考量到資料的規模與當代商用電腦的算力，使用網格搜尋法基於

¹這裡的同質性指的是空間中的機率值具有相同的性質與代表性，並非在空間上有相似的數值。

多測站資料為空間上單一網格單元計算 TIP 是不可行的作法。在機器學習領域中，決策樹不僅為發展成熟且高效能的演算法，更被廣泛應用在輸入特徵的重要性分析。相較於人工神經網路(Artificial Neural Network)，基於決策樹的模型能夠更好地處理包含不同資料種類的異質性資料集[12]，並在許多任務中達成頂級的表現[13]。由於決策樹演算法的效率、在應用上的彈性以及分析特徵重要性的能力，我們認為此方法特別適合我們從測站中心導向轉往以網格單元中心導向 (cell-centric) 發展 TIP 預報的需求。

而在本年度計畫中，我們首先進行一系列特徵工程(feature engineering)，主要目標在處理地電磁場的統計及資訊理論日指標之間的共線性問題，並探索指標的自相關特性以理解其在時間上的相依性，並為未來以決策樹為基礎繼續發展地震電磁前兆的關聯性研究做好準備。前期研究 [11] 顯示，在偏度、峰度之上引入夏農熵使得 TIP 預報對目標地震的擬合程度顯著提升，暗示地電磁場無序程度的增加可能與地震發生有著密切關聯。然而，同時引入夏農熵與費雪訊息卻導致擬合程度下降；我們認為這可歸因於兩指標之間的共線性關係，使得模型參數在訓練期間不穩定。在方法上，我們預計採用降維技術，例如主成分分析(Principal Component Analysis, PCA)、偏最小平方迴歸(Partial Least Squares, PLS) 等方法應對共線性問題。而在自相關分析方面，我們應用樣本自相關函數(Sample Autocorrelation Function, SACF)及樣本偏自相關函數(Sample Partial Autocorrelation Function, SPACF)等技術，探索地電磁指標的時間相依特性(temporal dependency)。由於地電磁觀測值及其導出量是反映著潛在地球物理過程的時間序列，往往具有時間上的相依性[14]，這意味著當前與過去之數值之間存在關聯。透過自相關分析，我們得以一窺地電磁統計與資訊理論指標時間序列的內部結構特徵，揭示其趨勢與週期性模式。這對於後續的特徵選擇與模型設計至關重要。

貳、資料與方法

(一) 地磁場與地電場測站、資料與地震目錄

本期計畫使用了中央氣象署地電場與地磁場觀測站紀錄之時間序列。其中，磁力觀測系統共 18 測站，地磁測站安裝精度 0.1nT 的全磁場強度磁力儀或精度為 0.01nT 的三分量磁力儀，取樣率皆為每秒記錄一個觀測值(1Hz)。地電場觀測系統(GEMS)取樣率為 15Hz，量測精度為 1 微伏特 (1 μ V)。原始資料記錄著電極間的電位差，單位為伏特(V)。由於各地電站因場址的區域限制，各長偶極的方向與電極展距或有差異。在陳宏嘉的 GEMSTIP 研究[15] 中，原始資料經過前處理，使得地電場資料呈現正南北與正東西方向的紀錄，並將資料降取樣(使用平均法)為 1Hz；關於地電測站與地電資料前處理的更詳細資訊請見 [15, 16]。

本研究之資料期間選定為 2014-01-01 至 2024-06-25 為資料期間；資料期間涵蓋全場磁力觀測、三分量磁力觀測以及二分量地電場觀測重疊的區間，以便對演算法在各種情況下之運作進行測試。地電、磁測站地理位置與名稱代號請見 表 1 與圖 1。資料期間內各測站所記錄之資料的缺值一覽如 圖 2 所示，其中每日紀錄到之資料點數占一日總秒數之比例以色階標示，而完全無資料的日子呈現白色。

(二) 濾波頻段

先前基於地磁場的地震前兆研究顯示，根據經帶通濾波 0.001-0.003Hz (代號: ULF_A) 或 0.001-0.01Hz (代號: ULF_B) 的地磁訊號計算得到的 TIP，相較於未將 $\geq 0.01\text{Hz}$ 的成分濾除的組別，與目標地震的擬合程度明顯較佳[6-8]。根據基於地電場的地震 TIP 前兆研究與相關文獻回顧，0.00032 至 0.0178 Hz (代號: BP_35) 以及 0.0001 至 0.00178 Hz (代號: BP_40) 的頻帶被認為是訊噪比高且與地震活動相關性高的頻段[15]。以上四個頻段的對照請參照圖 3。在本期計畫中，我們同樣基於這四個頻段，計算各地電磁站的峰度、偏度、費雪訊息以及夏農熵。

(三) 峰度與偏度

偏度(S)與峰度(K)作為統計量的 3 階與 4 階矩，分別量化地電磁觀測分布的不對稱性與集中程度，如式 1 與式 2 所示。

$$S(x) = \text{skewness}(x) = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^3}{\left(\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2} \right)^3} \quad (1)$$

$$K(x) = \text{kurtosis}(x) = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^4}{\left(\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2 \right)^2} \quad (2)$$

在統計學中， n 階矩反映變數與期望值間的距離 $n \times$ 機率密度，用於描述分布函數的結構特徵。作為地電磁異常的指標，高偏度可能反映著一日期間內地電磁場強度的突然整體偏移，而高峰度則可能意味著當日觀測中存在顯著離群值。這些特徵在 GEMS-MagTIP 演算法中作為識別潛在地震前兆的關鍵輸入變數。

(四) 費雪訊息及夏農熵

費雪訊息量度(Fisher Information Measure, FIM)與夏農熵(Shannon Entropy, SE)兩者應用的結合稱為費雪-夏農分析法，廣泛應用於時間序列的動態複雜度的研究上 [17-20]。費雪訊息的概念與數學基礎由費雪(Fisher)在 1920 至 30 年代基於最大概似估計(Maximum Likelihood Estimation)所建立[21-23]，之所以被稱為「訊息」是因為它量化了我們從觀測樣本中對參數不確定性的了解程度，而了解程度越高則意味著估計(estimate)本身的確定性越高。夏農熵由夏農(Shannon)於 1948 年提出[24]，其借用熱力學熵的概念量化訊號的隨機性。由於越高的隨機性需要更多位元描述，因此夏農熵被視為衡量「資訊量」的指標；當應用於物理系統狀態變量觀測時，其反映該變量的全域隨機程度。

在數學上，費雪訊息透過對機率密度函數變化率的積分為時間序列在局域(local)層面的組織或秩序水平(orderness)進行量化；夏農熵則透過對相當於「意外性」的機率密度倒數之積分，為全域的不確定性或混亂程度(randomness)進行量化。

兩者之數學定義如下：

$$\text{FIM} = \int_{-\infty}^{+\infty} \frac{1}{f(x)} \left(\frac{\partial f(x)}{\partial x} \right)^2 dx \quad (3)$$

$$\text{SE} = H_X = - \int_{-\infty}^{+\infty} f(x) \log f(x) dx \quad (4)$$

費雪訊息對機率密度函數 $f(x)$ 的局域變化極為敏感，這表現在式 3 中的 $(\frac{\partial f(x)}{\partial x})^2$ 項；除了高度集中的機率分布外，當 x 在局域的分布不均(機率密度沿 x 方向急遽起伏)基本上也會導致較高的 FIM 數值。

此外，夏農熵常被轉換為只允許正值的指數型式(Shannon Entropy Power, SEP)，用於等效表達系統動態特徵之不確定性並具更直觀的物理意義：

$$\text{SEP} = \frac{1}{2\pi e} e^{2H_X} \quad (5)$$

在地電磁場的異常量化上，高 SEP 值顯示地電磁日變化具有較高隨機性，分布更為平坦；而高 FIM 值則意味著觀測值高度集中或分布急遽變化。最後值得一提的是，對於鐘形分布而言，標準差越大，SEP 越高而 FIM 越低，反之亦然。這種兩面性導致我們常在地電磁的 SEP 與 FIM 之間觀察到某種程度的線性關係，這也是本系列研究中需要探究與克服的共線性問題。

(五) 時間序列自相關分析方法

在地電磁時間序列的分析中，自相關性 (Autocorrelation) 是一項數學上的關鍵特性，其揭示了時間序列中的內部結構和時間相依性，並能用於識別被雜訊掩蓋的模式。其中，自相關函數是量化時間序列與其自身延遲(lag)版本之間關係的基本工具。對於時間序列 Y_t 而言， k 階延遲 (lag k ，即延遲時間 $t_k = t_0 - k$) 的自相關函數(Auto-correlation Function, ACF) $\rho(k)$ 定義為 Y_t 與 Y_{t-k} 之間歸一化的共變異數 (covariance)：

$$\rho(k) = \frac{\text{Cov}(Y_t, Y_{t-k})}{\sqrt{\text{Var}(Y_t) \cdot \text{Var}(Y_{t-k})}} \quad (6)$$

其中，任意兩隨機變數的共變異數在數學上的定義為：

$$\text{Cov}(X, Y) = \mathbf{E}[(X - \mathbf{E}[X])(Y - \mathbf{E}[Y])] \quad (7)$$

其中， $\mathbf{E}[X]$ 為隨機變數 X 的期望值，又記做 \bar{X} ； $\mathbf{E}[Y]$ 、 \bar{Y} 亦然。對於平穩時間序列， $\text{Var}(Y_t) = \text{Var}(Y_{t-k})$ ，式6 可簡化為：

$$\rho(k) = \frac{\text{Cov}(Y_t, Y_{t-k})}{\text{Var}(Y_t)} \quad (8)$$

在數值計算上，式6 可寫成：

$$r(k) = \frac{\sum_{t=k+1}^T (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^T (Y_t - \bar{Y})^2} \quad (9)$$

這又稱為樣本自相關函數 (Sample Autocorrelation Function, SACF) 估計。其中 \bar{Y} 是時間序列的均值， T 是資料點的總數量。

ACF 無法區分直接和間接的相關性²，因此在分析時間序列的結構時，僅依賴自相關函數往往無法明確辨識各階延遲的獨立貢獻。為克服此一侷限，偏自相關函數 (Partial Autocorrelation Function, PACF) 應運而生，其可在排除中間延遲項 (如 $Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1}$) 的影響後，對 Y_t 與 Y_{t-k} 之間相關性進行量化。

在本研究中，我們使用 Yule-Walker 法[25, 26] 計算 PACF。Yule-Walker 法的核心假設是時間序列可用自回歸 (Auto-regressive, AR) 模型描述。AR(p) 模型如下：

$$Y_t = c + \phi_{p,1}Y_{t-1} + \phi_{p,2}Y_{t-2} + \dots + \phi_{p,p}Y_{t-p} + \epsilon_t \quad (10)$$

其中 $\phi_{p,1}, \dots, \phi_{p,p}$ 為 AR(p) 模型參數， ϵ_t 為白噪音項， c 為常數。與 AR(p) 對應之 Yule-Walker 的數學表示式如下：

$$\rho^{(p)}(k) = \sum_{j=1}^p \phi_{p,j} \rho^{(p)}(k-j) \quad \text{for } k > 0 \quad (11)$$

與式 6 不同的是，Yule-Walker 方程式假設 Y_t 為一可由 AR(p) 模型描述之隨機過程，而 lag k 的自相關係數可由延遲較小的版本 (lag $k-1, k-2, \dots, 1$) 及對應之 AR 模型係數遞迴求得 [25]。透過 Yule-Walker 方程組在自相關係數與 AR 參數之間建立起線性關係，我們將得以對各階參數項遞迴求解進而得到各階偏自相關係數。在實作[27] 上，各階延遲的 PACF 值可透過 Durbin-Levinson 遞迴公式計算，如下：

²以AR(1)， $Y_t = \phi_1 Y_{t-1} + \epsilon_t$ ，為例：當延遲為 1 時，ACF $\rho(1)$ 反映 Y_t 與 Y_{t-1} 之間直接相關性，對於平穩過程，其值等於 ϕ_1 。當延遲為 2 時， Y_t 與 Y_{t-2} 之間並無直接關聯，但因 Y_t 直接依賴 Y_{t-1} ，而 Y_{t-1} 又直接依賴 Y_{t-2} ，因此 Y_t 與 Y_{t-2} 之間存在間接相關性。 $\rho(2)$ 會捕捉到這種間接相關性，其值為 ϕ_1^2 。同理，延遲為 3 時， $\rho(3)$ 反映 Y_t 與 Y_{t-3} 之間透過 Y_{t-1} 與 Y_{t-2} 的間接相關性，其值為 ϕ_1^3 。因此，ACF 在高階延遲時同時包含了直接與間接相關性，僅依賴 ACF 難以辨識各階延遲的獨立貢獻。

$$\alpha(k) \equiv \phi_{k,k} = \begin{cases} \rho(1), & k = 1 \\ \frac{\rho(k) - \sum_{j=1}^{k-1} \alpha_{k-1}(j) \cdot \rho(k-j)}{1 - \sum_{j=1}^{k-1} \alpha_{k-1}(j) \cdot \rho(j)}, & \text{for } k > 1 \end{cases} \quad (12)$$

其中， $\alpha_{k-1}(j)$ 為 AR($k-1$) 模型的第 j 係數；即 $\phi_{k-1,j}$ 。式 12 即透過先前求得之 PACF 值遞迴求得 k 階延遲的偏自相關係數 $\alpha(k)$ 。PACF 反映隨機變數在 $t=0$ 與 $t-k$ 項的直接相關性。Yule-Walker 方程式建立了 AR 模型係數 (式 10) 與自相關函數數值 (式 9) 之間的聯結，使得 k 階延遲的直接相關性可透過 Durbin-Levinson 演算法，藉由移除中間延遲項 (1 至 $k-1$) 的線性效應，分離出 Y_t 與 Y_{t-k} 之間直接相關性。

對於完全可由 AR 模型描述的隨機過程，其各階 PACF ($\alpha(1), \dots, \alpha(p)$) 即對應到 AR 模型的係數 ($\phi_{1,1}, \dots, \phi_{p,p}$) 解；即 lag k 的 PACF $\alpha(k) = \phi_{k,k}$ 。在本研究中，我們將 ACF/PACF 的最大延遲期設定為 30 天 (對應至 lag 30)。為確保統計推論的穩健性，被用於自相關性分析的時間序列長度需至少大於最大延遲期的 4 倍 (即 120 天)；此一準則確保樣本數量足以提供可靠的自相關係數估計，特別是較高階延遲的係數。

地電與地磁場的觀測資料中通常存在大量缺失值 (在此研究中，缺失值的比例約為 0.401)，這是地球物理觀測資料的共同特徵。為應對此一狀況，我們為每個測站挑選最長的連續觀測序列進行自相關分析；若某測站的連續資料長度未達到前述 120 天的要求，則該測站不納入自相關分析的範圍。在離群值處理方面，我們採用四分位距 (IQR) 法則，將落在第一四分位數 (Q1) 減去 3 倍 IQR 以下或第三四分位數 (Q3) 加上 3 倍 IQR 以上的數值視為離群值，並以線性內插法進行取代。全部資料集 (地電磁場的日統計與資訊理論指標) 的整體離群值比例為 0.0386；對這個比例的資料進行線性內插，我們認為對 ACF/PACF 分析結果影響有限。值得注意的是，雖然我們在視覺化展示時將 FI 及 SE 以對數尺度呈現以利觀察，但在進行 ACF 與 PACF 分析時仍使用原始數值，以避免對數轉換可能引入的非線性效應影響時間相依結構的識別。

在機器學習的領域中，ACF 與 PACF 兩者的合併應用可以分析資料中的長期趨勢和週期性模式並透過比較觀測值與基於歷史自相關模式的預期值識別潛在的異常；這些特徵的分析結果可以提供模型輸入特徵選擇的方向指引，進而提升模型預測表現。本期研究將利用自相關性分析，識別地電磁指標中的時間相依模式與長期趨勢，並預計在未來的研究中根據分析結果進行模型的特徵篩選。

(六) 線性迴歸方法應對共線性特徵之應用

本期研究採用 Ridge 迴歸與偏最小平方迴歸 (Partial Least Squares, PLS) 方法處理 FI 與 SEP 在對數空間 (logspace) 中的共線性問題。兩種方法均透過不同的正則化策略，在共線性特徵存在時維持模型的穩定性與預測能力。

Ridge 迴歸透過在損失函數中加入懲罰項來限制迴歸係數的大小，從而緩解共線性問題 [28]。相較於最小平方線性迴歸法 (Ordinary Least Squares, OLS) 僅最小化殘差平方和 (式13)，Ridge 迴歸在目標函數中引入 L2 正則化項 (式14)：

$$\hat{\beta}_{OLS} = \arg \min_{\beta} \| \mathbf{Y} - \mathbf{X}\beta \|^2 \quad (13)$$

$$\hat{\beta}_{Ridge} = \arg \min_{\beta} \{ \| \mathbf{Y} - \mathbf{X}\beta \|^2 + \lambda \| \beta \|^2 \} \quad (14)$$

其中 \mathbf{Y} 為目標變數向量， \mathbf{X} 為輸入特徵矩陣， β 為迴歸係數向量， λ 為正則化參數。透過最小化此目標函數，Ridge 迴歸的解析解為：

$$\hat{\beta}_{Ridge} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{Y} \quad (15)$$

其中 \mathbf{I} 為單位矩陣。相較於 OLS，Ridge 迴歸的優勢在於當輸入特徵間存在高度共線性時， $\mathbf{X}^T \mathbf{X}$ 矩陣接近奇異(singular)，難以穩定逆推求解；而 Ridge 迴歸透過加入 $\lambda \mathbf{I}$ 確保矩陣的可逆性，從而避免數值不穩定與過度擬合。然而，對於簡單的線性關係問題，Ridge 迴歸會因為懲罰係數項而在估計中引入偏誤(bias)，使得估計值系統性地偏離真實值。此一偏誤－變異數權衡(bias-variance tradeoff)是 Ridge 迴歸的核心特性。

PLS 迴歸則採用降維策略應對共線性問題[29]。PLS 透過尋找輸入特徵 \mathbf{X} 與目標變數 \mathbf{Y} 的潛在變數 (latent factors/variables) 將原始高維度且可能共線性的特徵空間投影至低維度的潛在變數空間。PLS 的基本分解形式為：

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E}, \mathbf{Y} = \mathbf{U}\mathbf{Q}^T + \mathbf{F}, \text{ with } \max \text{cov}(\mathbf{t}_k, \mathbf{u}_k) \quad (16)$$

其中 \mathbf{T} 與 \mathbf{U} 分別為 \mathbf{X} 與 \mathbf{Y} 的分數矩陣 (score matrices)， \mathbf{P} 與 \mathbf{Q} 為載荷矩陣 (loading matrices)， \mathbf{E} 與 \mathbf{F} 為殘差項。PLS 的核心在於尋找能同時最大化 \mathbf{X} 與 \mathbf{Y} 共變異數的潛在成分。透過減少輸入特徵的維度，PLS 能有效簡化模型並避免因共線性導致的過擬合問題。相較於主成分分析 (PCA) 僅考慮 \mathbf{X} 的變異數，PLS 同時考慮 \mathbf{X} 與 \mathbf{Y} 的關聯性，因此在迴歸任務中往往具有更佳的預測表現。

本研究中，資料依據時間分為訓練期與測試期兩個區間。訓練期用於建立 SEP 與 FI 在對數空間中的線性關係模型，並透過交叉驗證(cross-validation)進行超參數 (hyperparameter) 調校；測試期則使用保留資料 (hold-out samples) 驗證訓練得到的模型，並計算預測殘差以評估模型的泛化能力。測試期的起始時間點為 2024-05-01。我們的目的是利用線性迴歸方法將 FI 可被 SEP 解釋的部分分離出來；透過比較 Ridge 與 PLS 迴歸對 SEP 與 FI 線性關係的擬合上的表現，我們得以評估不同正則化策略在後續資料處理與建模的適用性。

參、結果與討論

(一) 自相關性分析結果

各測站在不同濾波頻段、地電/磁場分量的統計與資訊理論指標時序圖展示於圖 4 至 圖 40；為了能夠在視覺上清楚呈現時間序列的特徵，我們在時序圖中對

FI 與 SEP 進行對數轉換。然而，實際進行 ACF/PACF 分析時仍使用原始數值，以保有資料的真實統計特性。在自相關性分析方面，各測站在不同濾波頻段、地電/磁場分量的統計與資訊理論指標之 ACF 及 PACF 函數圖展示於圖 44 至圖 80。

結果顯示，地電場與地磁場觀測資料中的時間相依結構與趨勢特徵，在一些統計或資訊理論量化指標上呈現顯著差異。首先，在多個地電測站(如 CHCH、DABA、ENAN、FENG、SHRL 等)的觀測資料中，FI 與 SE 時間序列可常常觀察到明顯的突然平移現象。這種特徵反映在 ACF 函數圖上的表現為極緩慢的衰減趨勢。為了驗證突然平移對自相關分析的影響，我們透過一階自回歸過程 (AR(1)) 的模擬範例 (圖 41 至圖 43)，說明當序列中出現平移時，即使隨機過程本身不具長期時間相依性，ACF 仍會呈現緩慢衰減的特徵。相較之下，在時間序列中段經歷相同平移的 AR(1) 過程之 PACF 則能正確反映其時間結構特徵。此一範例同時展示了 PACF 在辨識真實時間相依結構上的優勢。

即使前面提到的地電站之 FI 與 SE 的 ACF 呈現緩慢衰減現象，PACF 在除了 lag 0³ 以外的所有時間延遲版本之自相關係數皆維持在低水平 (多數低於 95% 信賴區間)，顯示地電場資料一般來說並無顯著的季節性變動模式。另外，我們發現 SHCH 測站的 SE 之 ACF 呈現較為特殊的特徵 (圖 58): ACF 緩慢衰減直到 lag 18 才降至信賴區間內。觀察對應的時間序列 (圖 18)，用於計算 ACF/PACF 的時序資料恰好涵蓋 SE 出現大量離群值的資料區間。由於離群值經線性內插，再加上對同樣時序所進行的 PACF 絕大多數仍落在信賴區間內，我們推測此種特徵是資料處理所導致的人為現象。

地磁測站的結果則展現出有別於地電測站的特性。大多數地磁測站的 ACF 與 PACF 均保持在較低水平 (低於或接近 95% 信賴區間)，僅在第 27 個延遲附近出現超越信賴區間的「丘狀」特徵。丘狀特徵代表連續數個延遲版本的自相關係數呈現較高的水平；相較於單一不連續的高自相關係數，丘狀特徵更不可能是隨機造成的結果。此特徵主要表現在 SE 與 FI 的 ACF/PACF 函數圖中，具備橫跨多個測站的一致性。綜合以上，這很有可能為真實的季節性訊號而非隨機結果。此 27 天週期性橫跨臺灣南北各測站，暗示其背景機制可能發生在地球或大氣系統尺度。

我們已知有兩種主要的天文活動具有約 27 天的週期性，分別是太陽自轉活動與月球軌道運動。太陽上具有可影響地磁觀測的來源，例如發射高速太陽風的冕洞 (coronal holes) 以及其他發生在太陽表面上的活躍事件，例如太陽閃焰 (solar flares) 和日冕物質的拋射 (coronal mass ejections, CMEs) 等；由於太陽自轉速度因緯度而異 [30]，這些活動從地球觀測的太陽自轉週期 (稱為會合週期, synodic period) 來看，通常約以 27 天再現，例如卡靈頓自轉週期 (Carrington rotation period, 27.2753 days) 或巴特爾斯自轉週期 (Bartels' rotation period, 27 days) [31, 32]。另一方面，月球軌道週期 (sidereal period, 即一般所指的公轉週期) 為 27.3 天，而月球的公轉運動透過潮汐效應影響著地球 [33]。

變化具有重要意義。值得注意的是，此延遲 27 天的自相關性在 2018 年前觀測的全場地磁場資料 (CS、LY、NC、TW、YH、TT 等測站) 中特別明顯，而

³ACF/PACF 的 lag 0 對應的自相關係數必為 1，反映著零延遲版本之間相關性必為 100% 的事實。

三分量地磁資料雖可觀察到類似特徵但不明顯。我們尚不清楚此特徵在不同時期的顯著性差異是否反映著某種現象的強弱或發生機率隨時間推移而改變，或僅為儀器及其他可能人為因素所導致的不同。

總體而言，ACF 與 PACF 分析不僅揭示了地電磁場指標的時間結構特徵，也幫助我們識別資料特性突然改變的現象，可為地電磁異常的定義與背景變化之區分提供重要參考與指引。

(二) 應對共線性問題的迴歸分析

為了應對資訊理論指標之間的共線性問題，我們提出利用線性迴歸模型分離相依變數中不可解釋的部分，進而降低輸入特徵的相依性。具體而言，我們的目標是利用 SEP 預測 FI，而預測值與實際值之間的殘差 ($r_{FI} = FI_{\text{預測}} - FI$) 即為 SEP 無法透過線性關係解釋的部分。本研究結果可作為後續建模的特徵工程之參考；例如，以 FI 的殘差取代原本的 FI 值作為輸入特徵，從而消除共線性並改善模型表現。

在實驗設計上，我們分別對地電場測站與地磁場測站建立迴歸模型。對於地電場測站，輸入特徵為南北與東西分量之夏農熵指數 ($\log_{10}(\text{SEPNS})$ 、 $\log_{10}(\text{SEPEW})$)，目標變數則為對應分量之費雪訊息對數值 ($\log_{10}(\text{FINS})$ 、 $\log_{10}(\text{FIEW})$)；對於三分量地磁場測站，輸入特徵為三個分量的夏農熵指數 ($\log_{10}(\text{SEP}x)$ 、 $\log_{10}(\text{SEP}y)$ 、 $\log_{10}(\text{SEP}z)$)，目標變數則為對應的費雪訊息對數值 ($\log_{10}(\text{FI}x)$ 、 $\log_{10}(\text{FI}y)$ 、 $\log_{10}(\text{FI}z)$)。我們採用 2024-05-01 作為訓練期與測試期的分界點，訓練期資料用於模型訓練與超參數調整，測試期資料則用於評估模型的泛化能力。模型的預測表現以平均絕對誤差 (Mean Absolute Error, MAE) 衡量，其計算基於交叉驗證 (Cross-validation, CV) 程序；交叉驗證的摺 (fold) 數根據樣本數量動態調整，最少為 3 摺，最多為 10 摺，以確保每摺皆有足夠的樣本數。在超參數的設定方面，對於 Ridge 迴歸，正則化參數 λ 的搜尋範圍設定為 0 至 100，共 50 個網格點；對於 PLS 迴歸，潛在成分數 (n) 的搜尋範圍為 1 至輸入特徵的數量 (對地電站為 2，對地磁站為 3)。

圖 81 與圖 82 展示了 Ridge 迴歸在不同正則化參數下的交叉驗證 MAE 變化。結果顯示，正則化參數的學習曲線因測站種類 (地電或地磁) 以及同一種類內不同測站而呈現不同的特徵。對於地電站的結果而言，在 $\lambda < 50$ 的範圍內學習曲線變化不大；可觀察到一些測站的學習曲線在此範圍內隨 λ 增加而微幅下降，一些測站則隨 λ 增加而上升；幾乎全部測站在 $\lambda > 50$ 之後 MAE 隨 λ 增加而上升。對於地磁站的結果而言，不同測站呈現不同特徵的學習曲線：同樣在 $\lambda < 50$ 的範圍內，可觀察到一些測站的學習曲線隨 λ 增加而微幅下降，一些測站 MAE 則隨 λ 增加而快速上升；同樣幾乎全部測站在 $\lambda > 50$ 之後 MAE 隨 λ 增加而有不同程度的上升。總結來說，在 $\lambda < 50$ 的範圍內進行超參數調整，透過交叉驗證得到適當的偏誤-變異數權衡，即可得到良好的以 SEP 預測 FI 的線性模型。圖 83 與圖 84 展示了 PLS 迴歸模型在訓練期之交叉驗證 MAE 隨潛在成分數變化的情形。對於三分量地磁測站與二分量地電站，幾乎所有測站都顯示較高的潛在成分數可以降低 MAE，反映 PLS 迴歸透過提取更多潛在成分能更有效捕捉輸入特

徵與目標變數之間的關係。

圖 85 與 圖 86 展示了在前述的超參數範圍中以交叉驗證得到的最佳模型之 MAE 比較。結果顯示，多數情況下，PLS 和 Ridge 迴歸的結果非常接近，難以辨別誰優誰劣，除了以下情況：對於 ENAN、WANL、FENL 地電站，在 BP-35 與 ULF-B 濾波頻段的結果中，PLS 的 MAE 較 Ridge 高；對於 LN 地磁站，PLS 的 MAE 亦較 Ridge 高。這些例外情況可能反映 PLS 迴歸在特定測站與頻段組合下的過度擬合現象。

圖 87 與 圖 88 展示了根據訓練期得到的最佳 Ridge 迴歸模型，對測試期資料進行預測的實際值與預測值散佈圖。結果顯示，Ridge 迴歸模型在很大的尺度範圍（從 1 至大於 10 個數量級）內，都能很好地從 SEP 預測 FI，驗證了模型的泛化能力與穩健性。殘差時序圖展示於 圖 89 與 圖 90。殘差代表 FI 中 SEP 無法透過線性迴歸模型解釋的部分，其包含了獨立於 SEP 的資訊。未來可以 FI 的殘差取代原本的 FI 值，作為 TIP 模型的輸入特徵，從而消除 FI 與 SEP 之間的共線性，並可能改善模型的預測表現。

綜合以上，本研究透過 Ridge 迴歸與 PLS 迴歸成功建立以 SEP 預測 FI 的線性模型，並驗證模型在測試期的泛化能力。研究結果揭示適當的正則化參數範圍（ $\lambda < 50$ ），並確認殘差作為特徵工程候選的可行性，為後續發展基於決策樹的 TIP 預報模型奠定基礎。

肆、結論

本期計畫建立了地電磁場統計與資訊理論指標的完整自相關性分析框架，透過 SACF 與 SPACF 系統性地探討費雪訊息與夏農熵的時間相依特性。研究結果揭示地磁場觀測中約 27 天的週期性特徵；特徵橫跨多個測站，可能與太陽自轉或月球軌道等天體運行週期相關，為判斷地磁異常的背景變化提供重要參考。對於地電場觀測，自相關分析識別出資訊理論指標的突然平移現象，PACF 分析則顯示地電場資料普遍不具顯著的長期時間相依性，反映其相對穩態的馬可夫性質。這些發現為理解地電磁場觀測的時間相依結構提供新見解，並為後續特徵工程與模型開發建立重要基礎。

在應對共線性問題方面，本研究透過 Ridge 迴歸與 PLS 迴歸成功建立以夏農熵預測費雪訊息的線性模型，並針對各測站與濾波頻段組合進行系統性的超參數調校與交叉驗證。研究確認在正則化參數 $\lambda < 50$ 的範圍內，Ridge 迴歸能有效達到偏誤—變異數權衡，並在測試期展現良好的泛化能力。透過分離費雪訊息中可被夏農熵線性解釋的部分，我們獲得包含獨立資訊的殘差特徵，可作為未來 TIP 模型輸入特徵的候選，從而消除共線性並可能改善模型表現。

此外，本期計畫亦完成嘉義縣中埔鄉自然電位接地點場勘工作，確認三層國小、汧水國小及仙草國小關嶺分校為適合設置地電觀測站的候選點位，為未來擴展觀測網路提供實地評估基礎。

基於本期研究成果，未來工作將朝向建立基於決策樹演算法的地震前兆特徵

重要性分析方法發展。我們預計將傳統 GEMS-MagTIP 演算法在現代機器學習框架下重現，並開發基於 TIP 概念的決策樹模型，實現特徵重要性的量化分析，進而發展提供精細風險評估的機率預報模式。地電磁場指標並非單純的隨機數列，而是潛在地球物理過程的體現。本研究透過自相關分析所建立的時間序列內部結構特徵之理解，以及透過迴歸分析所發展的共線性處理方法，為後續機器學習模型的開發奠定關鍵基礎。

參考文獻

- [1] Keilis-Borok VI, Kossobokov VG (1990) Premonitory Activation of Earthquake Flow: Algorithm M8. *Physics of the Earth and Planetary Interiors* 61:73–83. [https://doi.org/10.1016/0031-9201\(90\)90096-G](https://doi.org/10.1016/0031-9201(90)90096-G)
- [2] 陳建志, 陳宏嘉, 徐漢倫 (2014) 地震電磁前兆現象之研究 (計畫編號: MOTC-CWB-103-E-02)
- [3] 陳建志, 陳宏嘉, 徐漢倫 (2015) 地震電磁前兆現象之研究: TIP 預警模型之建立與顯著性測試 (計畫編號: MOTC-CWB-104-E-02)
- [4] 陳建志, 陳宏嘉, 孫郁勝 (2016) 地震電磁前兆現象之研究: 地震電磁前兆現象之研究 (計畫編號: MOTC-CWB-105-E-01)
- [5] 陳建志, 陳宏嘉, 孫郁勝 (2017) 地震電磁前兆現象之研究: 地電場的天然時間分析法與其地震前兆的初步研究 (計畫編號: MOTC-CWB-106-E-01)
- [6] 陳建志, 吳宗義 (2020) 台灣地區 109 年地震前兆監測資料彙整及分析—以 MagTIP 預報模型評估地磁場異常與地震的關連性 (計畫編號: MOTC-CWB-109-E-01)
- [7] 陳建志, 吳宗義 (2019) 地磁場異常與地震的關連性研究 (計畫編號: MOTC-CWB-108-E-01)
- [8] 陳建志, 吳宗義, 陳宏嘉 (2021) 台灣地區 110 年地震前兆監測資料彙整及分析—多變量 MagTIP 預報系統之研究 (計畫編號: MOTC-CWB-110-E-01)
- [9] 陳建志, 吳宗義, 陳宏嘉 (2022) 台灣地區 111 年地震前兆監測資料彙整及分析—資訊理論應用於地電磁場相關聯之地震前兆研究 (計畫編號: MOTC-CWB-111-E-01)
- [10] 陳建志, 吳宗義, 陳宏嘉 (2023) 台灣地區 112 年地震前兆監測資料彙整及分析—資訊理論應用於地電磁場相關聯之地震前兆研究 (計畫編號: MOTC-CWB-112-E-04)
- [11] 陳建志, 吳宗義, 陳宏嘉 (2024) 台灣地區 113 年地震前兆監測資料彙整及分析—資訊理論應用於地電磁場相關聯之地震前兆研究 (計畫編號: MOTC-CWA-113-E-04)
- [12] Kadra A, Lindauer M, Hutter F, Grabocka J (2021) Well-tuned Simple Nets Excel

on Tabular Datasets. In: Beygelzimer A, Dauphin Y, Liang P, Vaughan JW (eds) *Advances in Neural Information Processing Systems*

- [13] Bentéjac C, Csörgö A, Martínez-Muñoz G (2021) A Comparative Analysis of Gradient Boosting Algorithms. *Artificial Intelligence Review* 54:1937–1967. <https://doi.org/10.1007/s10462-020-09896-5>
- [14] Helman DS (2020) Seismic Electric Signals (SES) and Earthquakes: A Review of an Updated VAN Method and Competing Hypotheses for SES Generation and Earthquake Triggering. *Physics of the Earth and Planetary Interiors* 302:106484. <https://doi.org/10.1016/j.pepi.2020.106484>
- [15] Chen H-J (2018) 地電訊號異常與地震的關聯性研究
- [16] Hsu H-L (2013) 台灣天然電磁場觀測研究
- [17] Martin M, Pennini F, Plastino A (1999) Fisher's Information and the Analysis of Complex Signals. *Physics Letters A* 256:173–180. [https://doi.org/10.1016/S0375-9601\(99\)00211-X](https://doi.org/10.1016/S0375-9601(99)00211-X)
- [18] Lovallo M, Shaban A, Darwich T, Telesca L (2013) Investigating the Time Dynamics of Monthly Rainfall Time Series Observed in Northern Lebanon by Means of the Detrended Fluctuation Analysis and the Fisher-Shannon Method. *Acta Geophysica* 61:1538–1555. <https://doi.org/10.2478/s11600-012-0094-9>
- [19] Telesca L, Lovallo M, Chamoli A, et al (2013) Fisher-Shannon Analysis of Seismograms of Tsunamiogenic and Non-Tsunamiogenic Earthquakes. *Physica A: Statistical Mechanics and its Applications* 392:3424–3429. <https://doi.org/10.1016/j.physa.2013.03.049>
- [20] Telesca L, Lovallo M (2017) On the Performance of Fisher Information Measure and Shannon Entropy Estimators. *Physica A: Statistical Mechanics and its Applications* 484:569–576. <https://doi.org/10.1016/j.physa.2017.04.184>
- [21] Fisher RA (1925) Theory of Statistical Estimation. *Mathematical Proceedings of the Cambridge Philosophical Society* 22:700–725. <https://doi.org/10.1017/S0305004100009580>
- [22] Frieden BR (1990) Fisher Information, Disorder, and the Equilibrium Distributions of Physics. *Physical Review A* 41:4265–4276. <https://doi.org/10.1103/PhysRevA.41.4265>
- [23] Frieden BR (1998) *Physics from Fisher Information: A Unification*. Cambridge University Press, Cambridge
- [24] Shannon CE (1948) A Mathematical Theory of Communication. *Bell System Technical Journal* 27:379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- [25] Shumway RH, Stoffer DS (2025) ARIMA Models. *Time Series Analysis and Its Applications: With R Examples* 85–175
- [26] Brockwell PJ, Davis RA (2016) Modeling and Forecasting with ARMA Processes. *Introduction to Time Series and Forecasting* 121–155

- [27] Lin D, Milan Bouchet-Valat, Noack A, et al (2025) JuliaStats/StatsBase.Jl: V0.34.5
- [28] JuliaStats (2025) MultivariateStats.Jl: Julia Package for Multivariate Statistics and Data Analysis. <https://github.com/JuliaStats/MultivariateStats.jl>. Accessed 14 Nov 2025
- [29] Alvim L (2025) PartialLeastSquaresRegressor.Jl: Partial Least Squares Regressor for Julia. <https://github.com/lalvim/PartialLeastSquaresRegressor.jl>. Accessed 14 Nov 2025
- [30] Solar Rotation Varies by Latitude - NASA. <https://www.nasa.gov/image-article/solar-rotation-varies-by-latitude/>. Accessed 29 May 2025
- [31] Bartels J (1934) Twenty-Seven Day Recurrences in Terrestrial-Magnetic and Solar Activity, 1923–1933. *Terrestrial Magnetism and Atmospheric Electricity* 39:201–
- [32] 202. <https://doi.org/10.1029/TE039i003p00201>
- [33] Beck JG (2000) A Comparison of Differential Rotation Measurements - (Invited Review). *Solar Physics* 191:47–70. <https://doi.org/10.1023/A:1005226402796>
- [34] Guoqing L (2005) 27.3-Day and 13.6-Day Atmospheric Tide and Lunar Forcing on Atmospheric Circulation. *Advances in Atmospheric Sciences* 22:359–374. <https://doi.org/10.1007/BF02918750>

圖表

表1: 地電、磁測站地理位置與名稱代號。本報告未使用之測站資料以淺灰字標示。

代號	站名	種類	經度	緯度	觀測起始	觀測結束
MS	馬仕	地磁	120.63301	22.61089	2015-11-13	2018-10-12
TW	灣丘	地磁	120.528576	23.185017	2006-11-01	2019-09-04
TT	卑南	地磁	121.079922	22.817646	2006-11-01	
YL	玉里	地磁	121.297288	23.392399	2007-05-28	2019-09-22
HC	恆春	地磁	120.81367	21.940147	2007-05-09	2019-09-16
HL	加灣	地磁	121.604765	24.081107	2007-04-11	2019-04-11
PT	瑪家	地磁	120.65311	22.703066		2012-12-03
YH	爺亨	地磁	121.375767	24.669521	2007-11-02	2018-12-05
SL	雙龍	地磁	120.955803	23.788779	2007-04-09	2019-03-01
LY	鯉魚潭	地磁	120.780215	24.346513	2006-09-21	2019-01-17
NC	內城	地磁	121.682855	24.718023	2007-05-15	2019-08-31
KM	金門	地磁	118.45274	24.44934	2015-11-13	2019-01-22
CS	池上	地磁	121.2264	23.1112	2014-03-25	2019-01-09
MT	馬祖	地磁	119.923	26.1697	2020-09-10	
LN	蘭嶼	地磁	121.558	22.0375	2020-10-10	
ZB	知本	地磁	121.0648	22.7398	2020-10-06	
XC	新城	地磁	121.6095	24.0383	2020-10-06	
SM	日月潭	地磁	120.9076	23.881	2019-10-08	
CN	暨南	地磁	120.92817	23.95761	2022-12-29	
KUOL	過嶺	地電	121.142036	24.962872	2011-09-22	
HUAL	華陵	地電	121.367683	24.674475	2012-01-17	
TOCH	頭城	地電	121.805206	24.843529	2012-02-10	
ENAN	南澳	地電	121.78495	24.475822	2012-02-15	
SIHU	四湖	地電	120.229309	23.636976	2012-02-08	

HERM	和美	地電	120.501526	24.108841	2012-02-09	
CHCH	將軍	地電	120.161758	23.219656	2012-02-29	
DAHU	大湖	地電	120.902374	24.410578	2012-02-07	
KAOH	高雄	地電	120.289261	22.6577	2012-02-23	
PULI	埔里	地電	120.978783	23.920755	2012-03-01	
SHRL	士林	地電	121.561874	25.15589	2012-03-29	
SHCH	新城	地電	121.625008	24.11832	2012-04-02	
FENL	鳳林	地電	121.411156	23.715563	2012-04-03	
YULI	玉里	地電	121.318069	23.32472	2012-04-24	
RUEY	瑞源	地電	121.155685	22.973249	2012-04-23	
LIOQ	六龜	地電	120.663185	23.03211	2012-07-03	
LISH	梨山	地電	121.255091	24.253189	2012-11-07	
DABA	達邦	地電	120.748902	23.453463	2013-02-01	
WANL	萬巒	地電	120.593656	22.590934	2012-12-20	
FENG	楓港	地電	120.700691	22.204332	2012-12-20	
HUZS	磺嘴山	地電	121.600564	25.199053	2020-07-29	

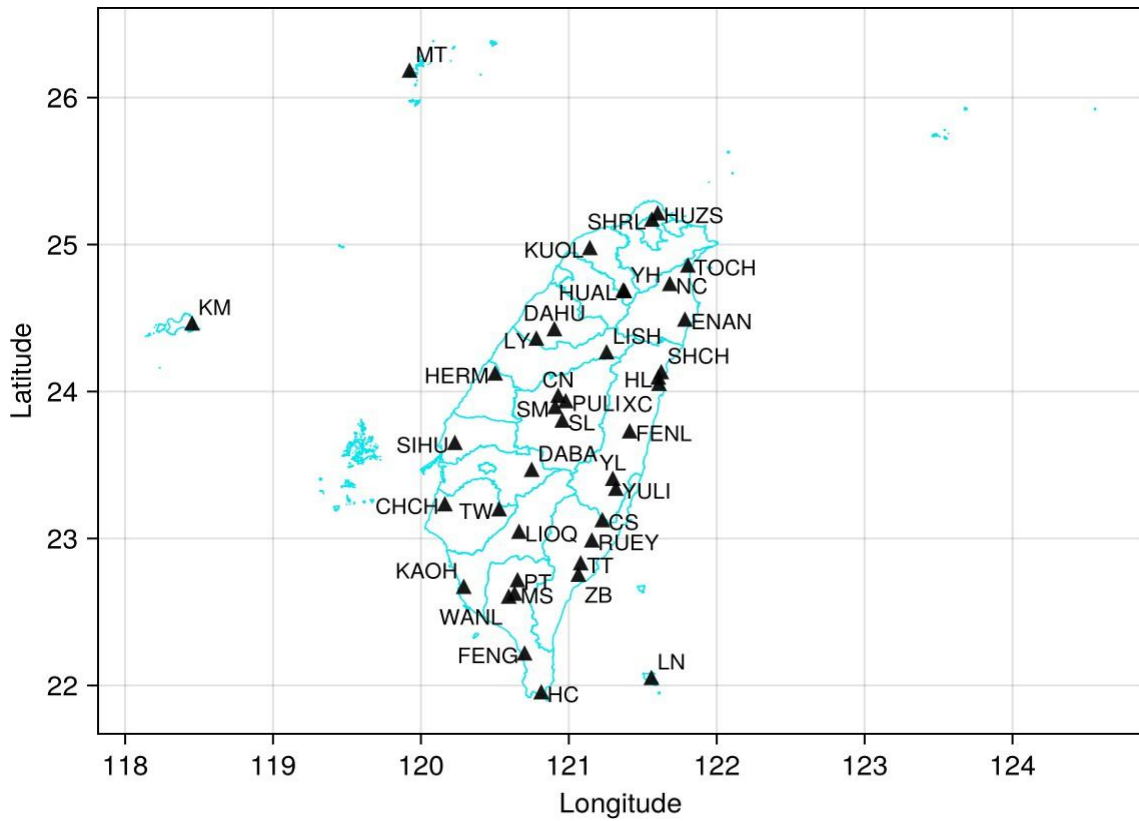


圖1: 中央氣象署地電與地磁觀測站一覽圖。



圖2: 使用之測站資料一覽。每日有效資料點數之比例以色階標示，完全無資料的日子呈現白色。測站軸之測站代碼依測站緯度由高至低(北至南)排列。

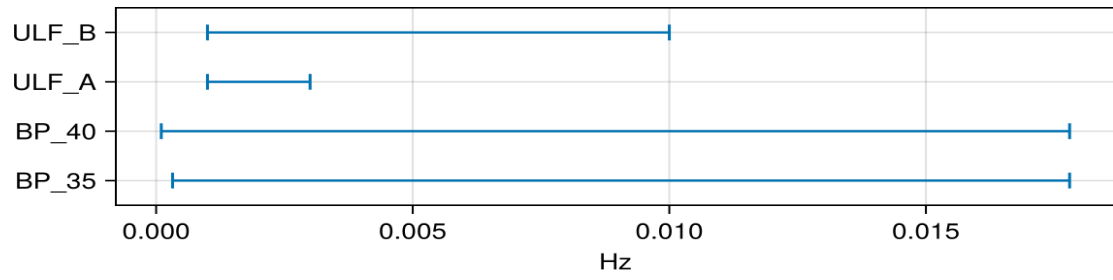


圖3: 本研究所涉之濾波頻段視覺化參考圖

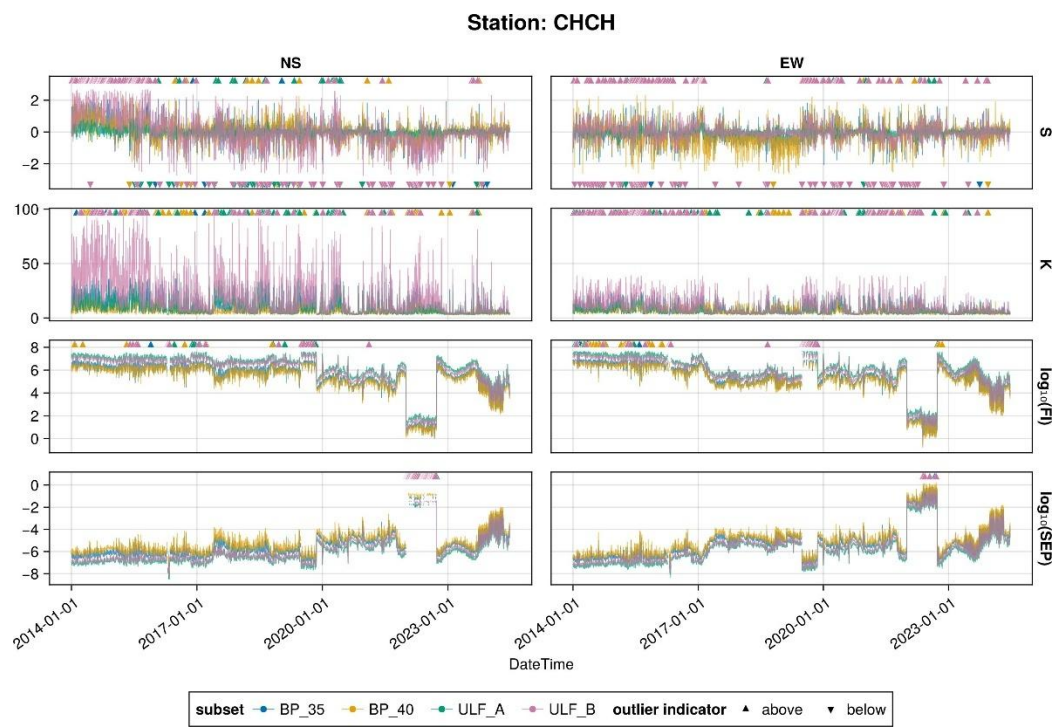


圖4: 將軍 (CHCH) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

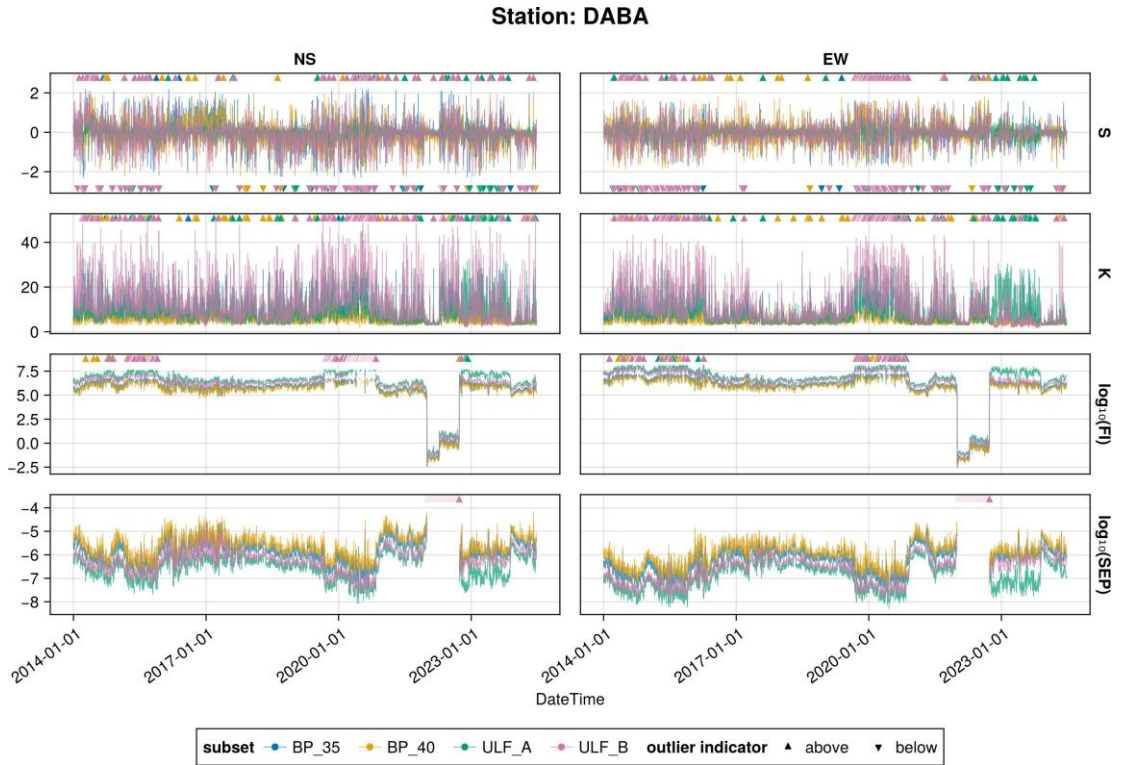


圖5: 達邦 (DABA) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

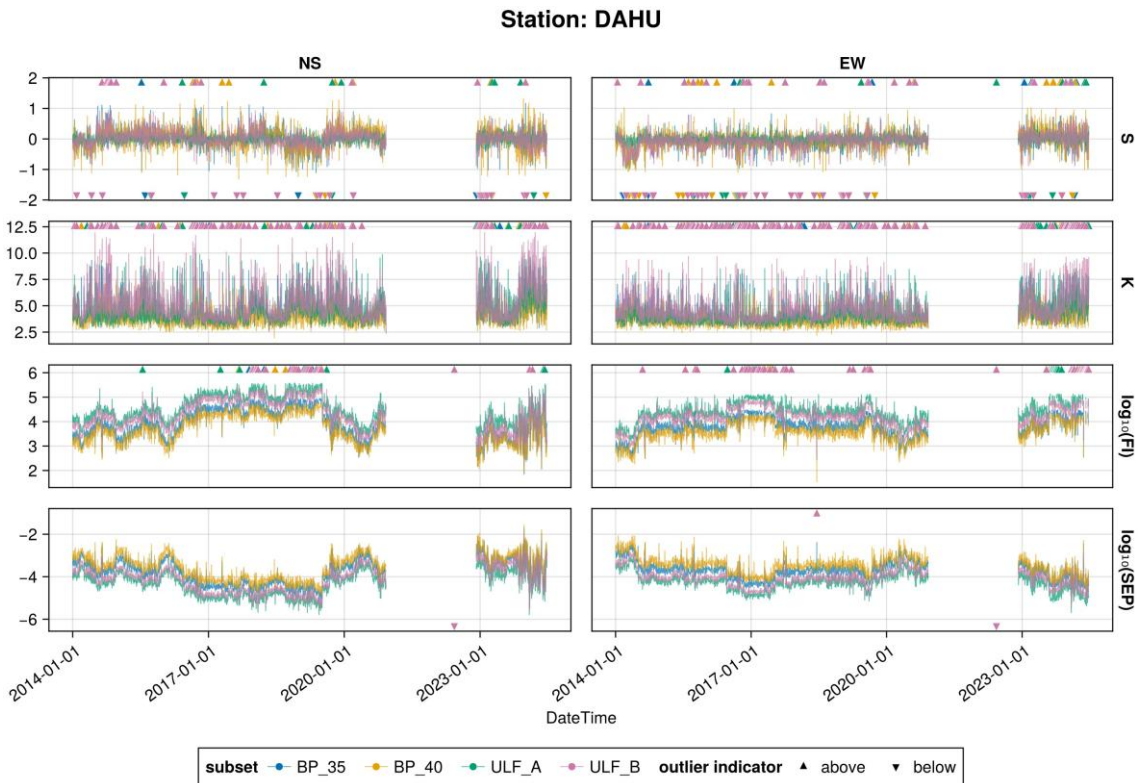


圖6: 大湖 (DAHU) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

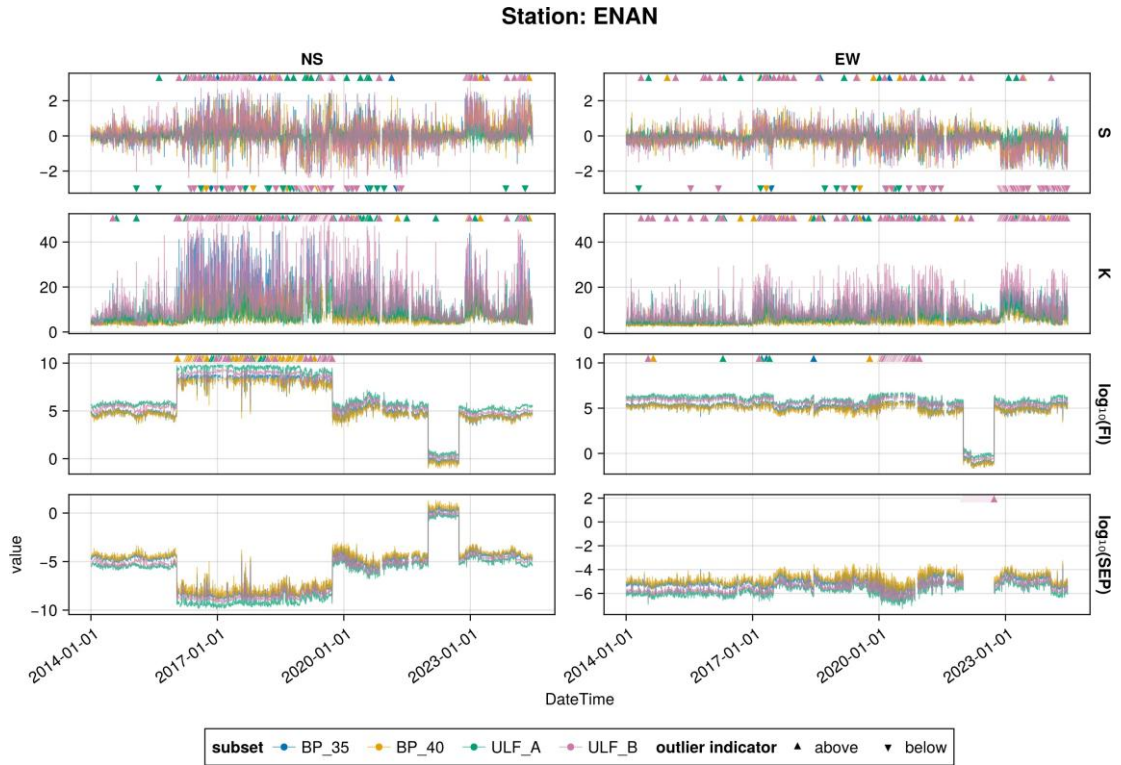


圖7: 南澳 (ENAN) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

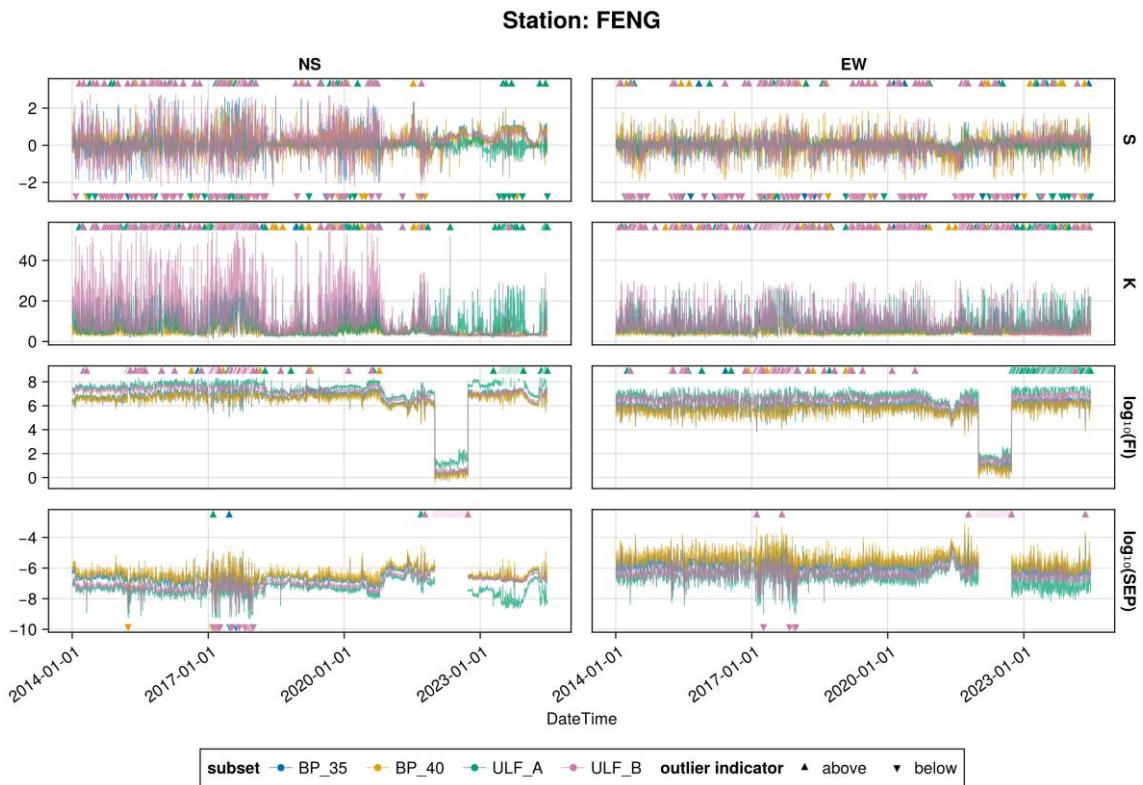


圖8: 楓港 (FENG) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

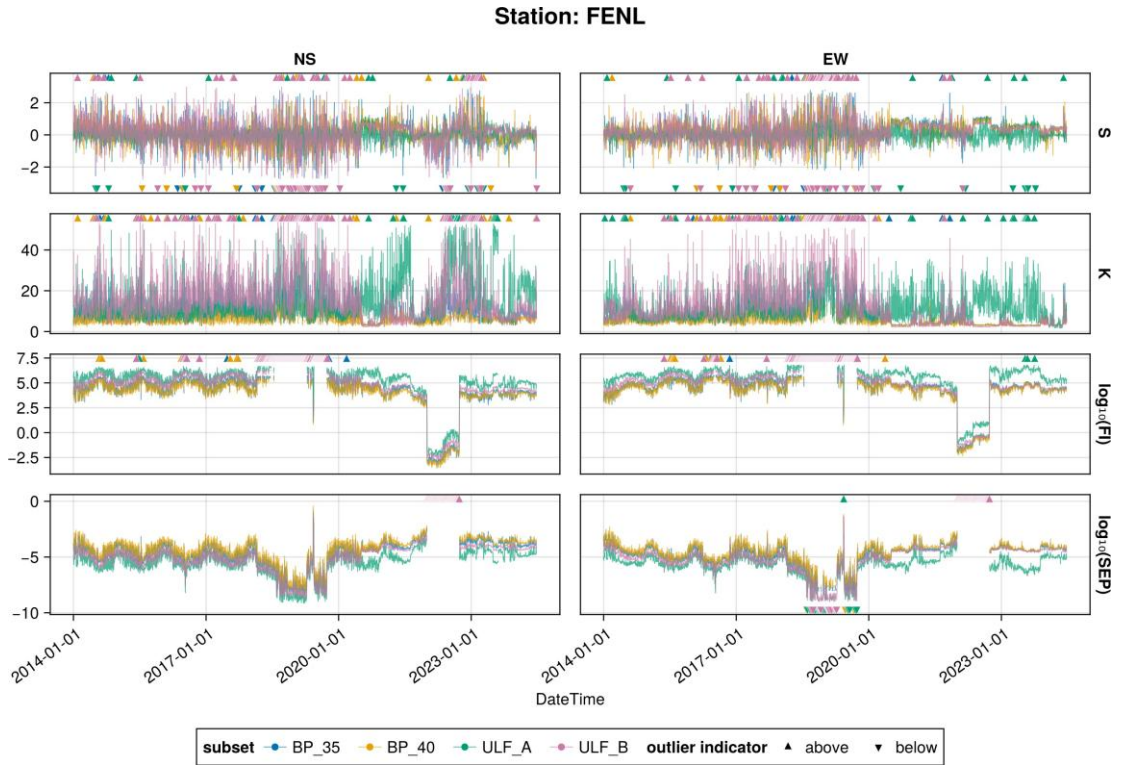


圖9: 鳳林 (FENL) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

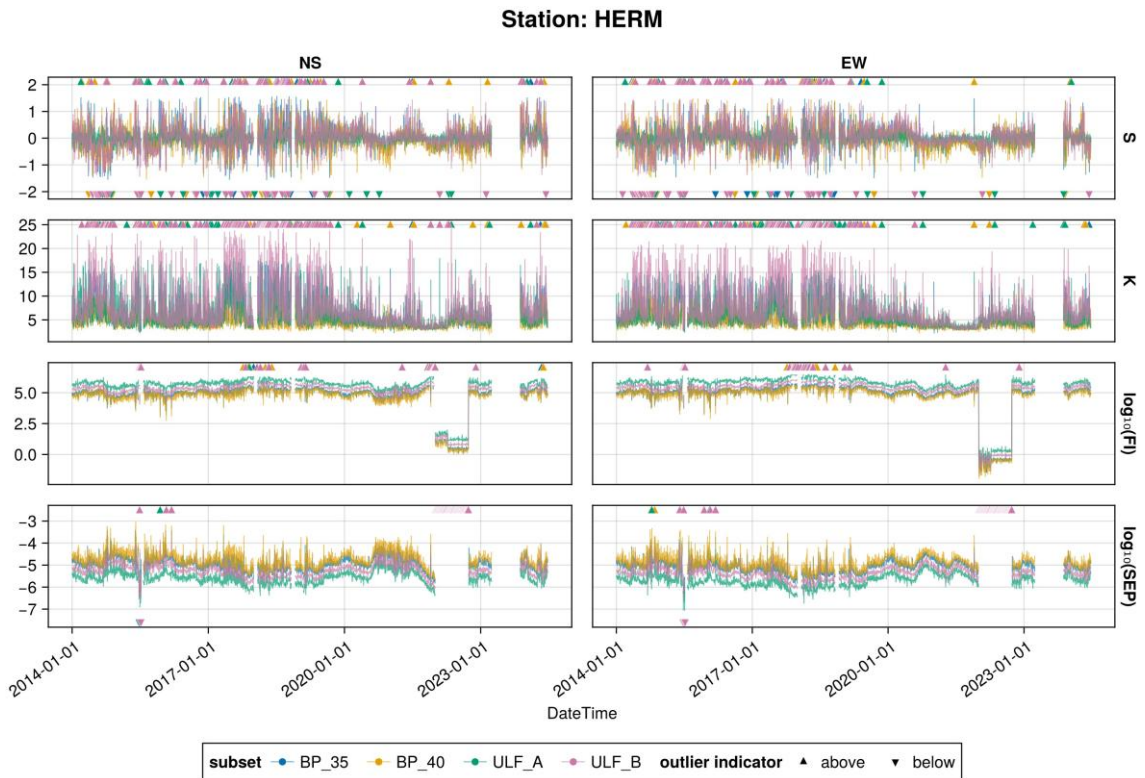


圖10: 和美 (HERM) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

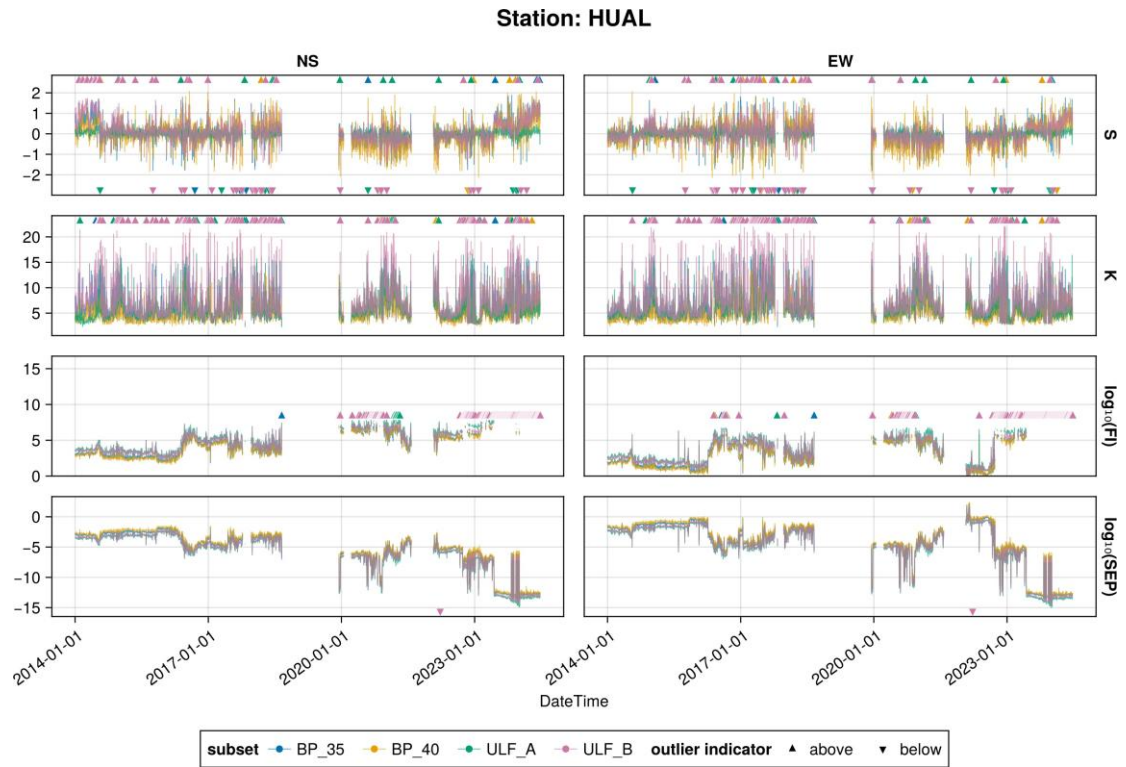


圖 11: 華陵 (HUAL) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

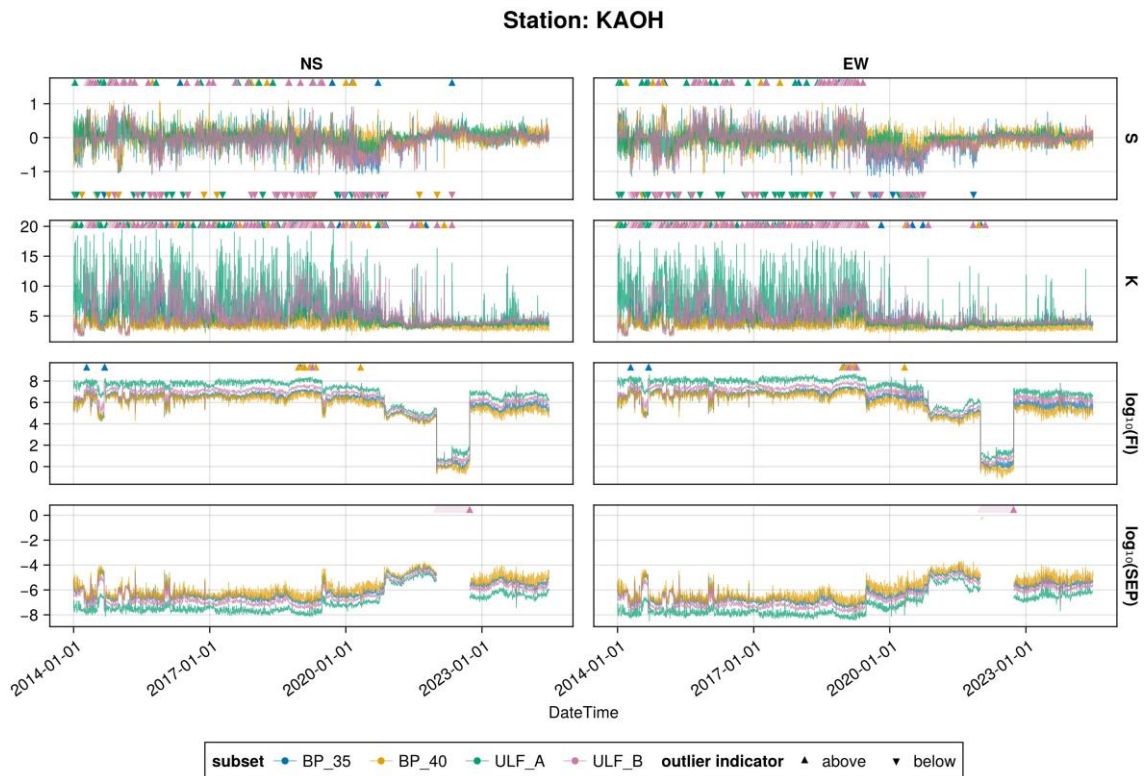


圖 12: 高雄 (KAOH) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

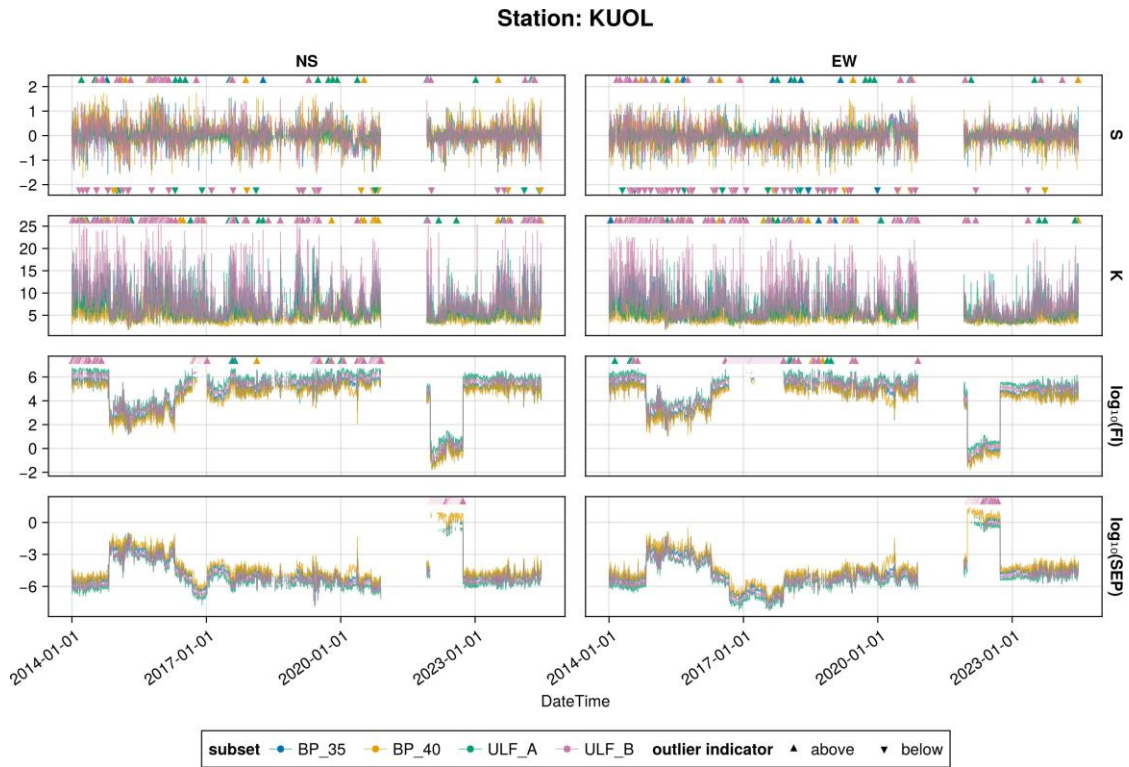


圖13: 過嶺 (KUOL) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

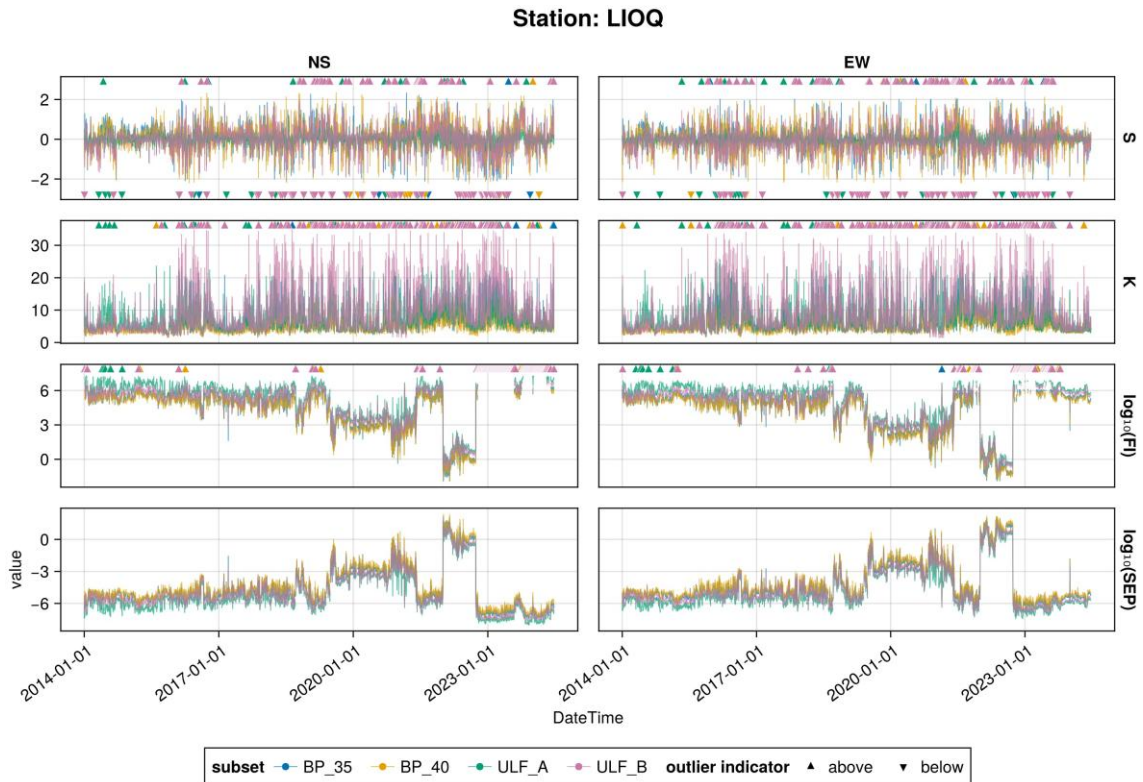


圖14: 六龜 (LIOQ) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

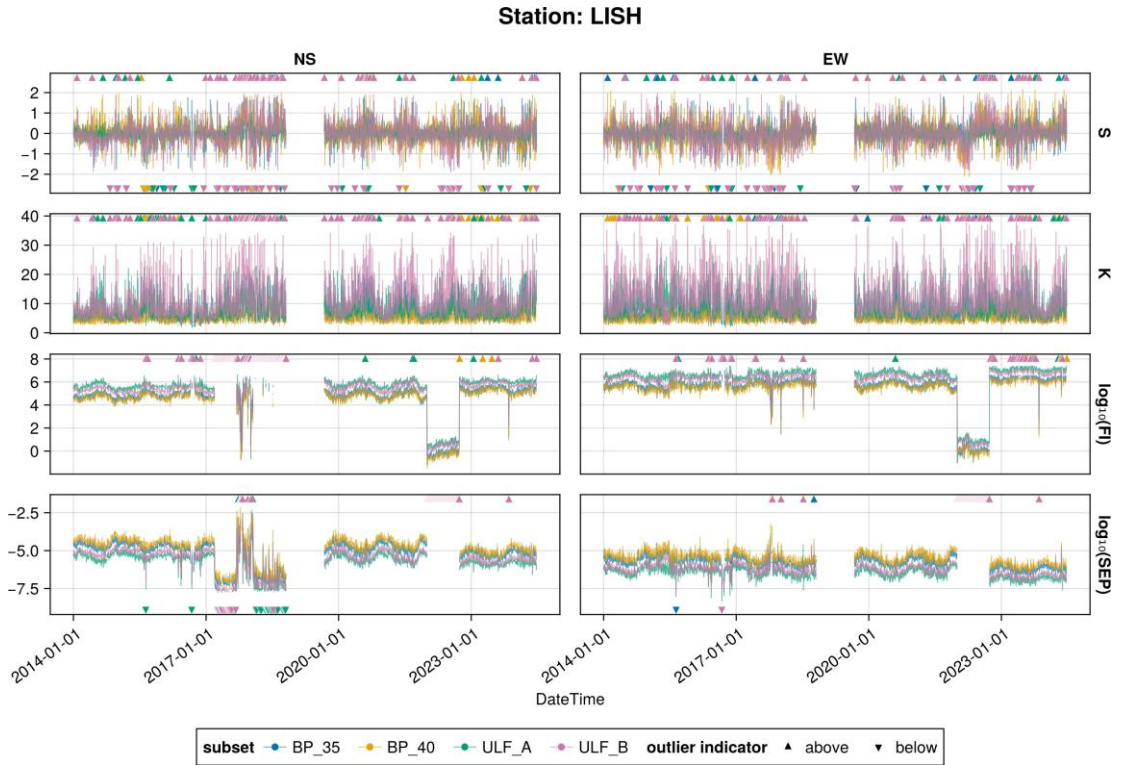


圖15: 梨山 (LISH) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

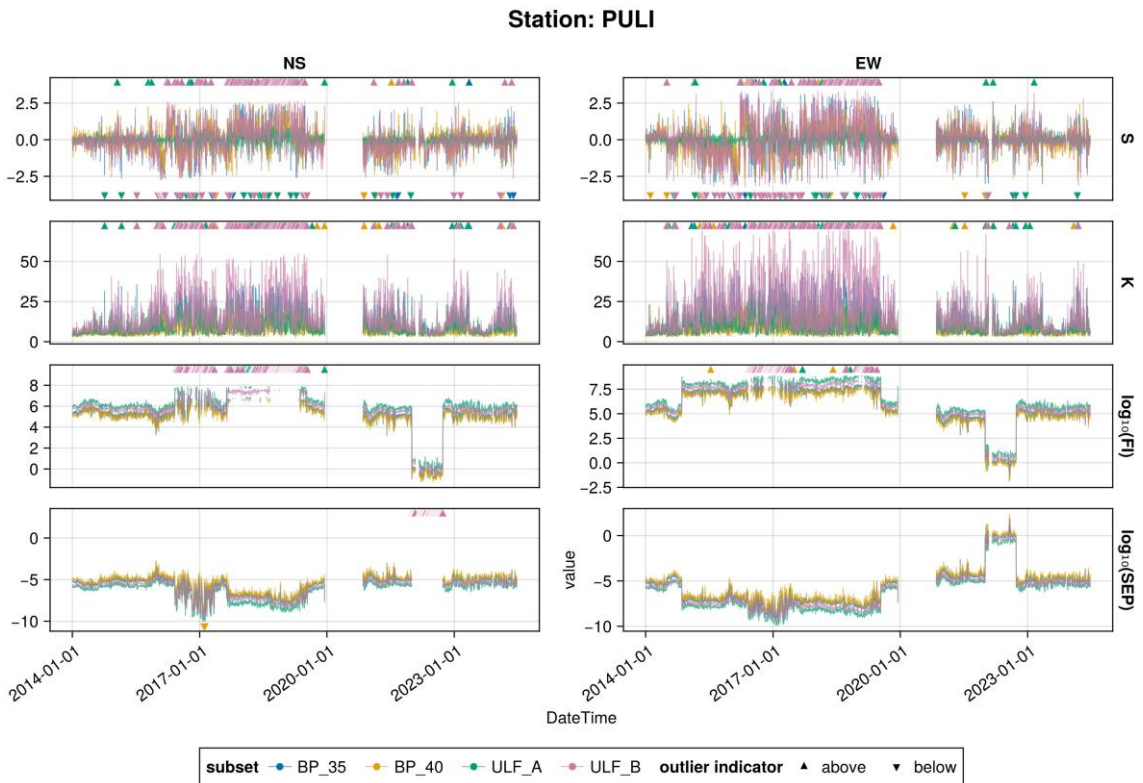


圖16: 埔里 (PULI) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

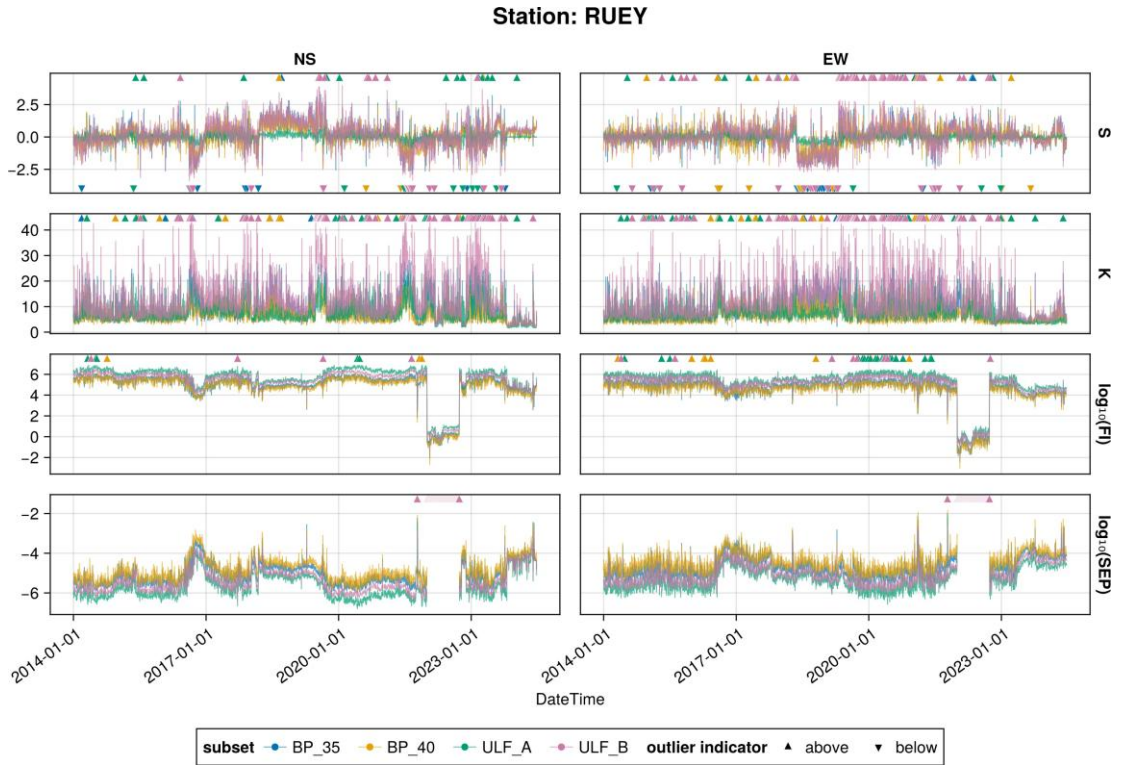


圖17: 瑞源 (RUEY) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

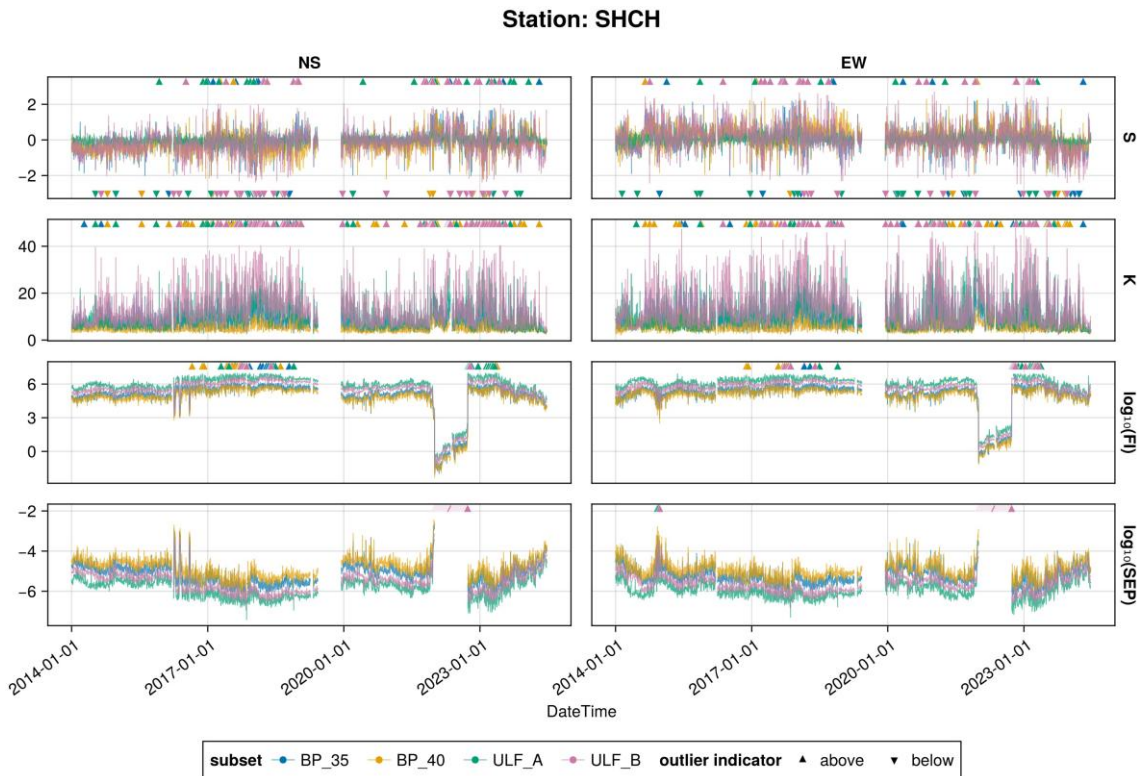


圖18: 新城 (SHCH) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

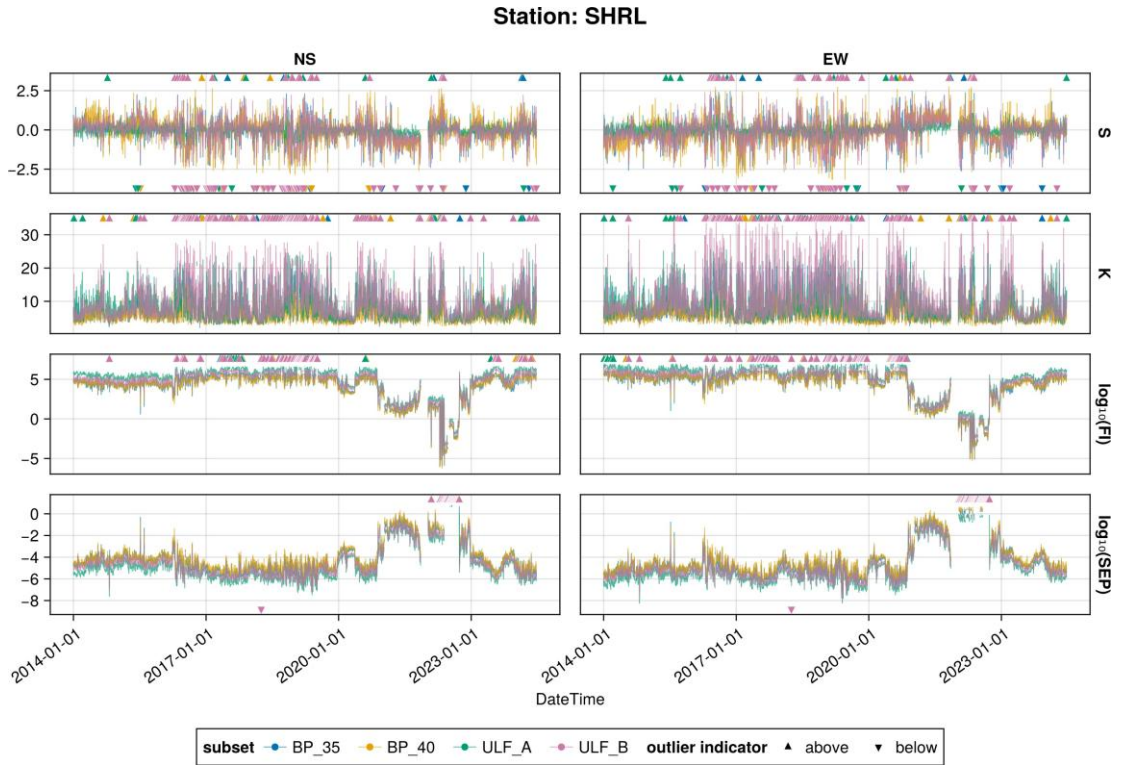


圖19: 士林 (SHRL) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

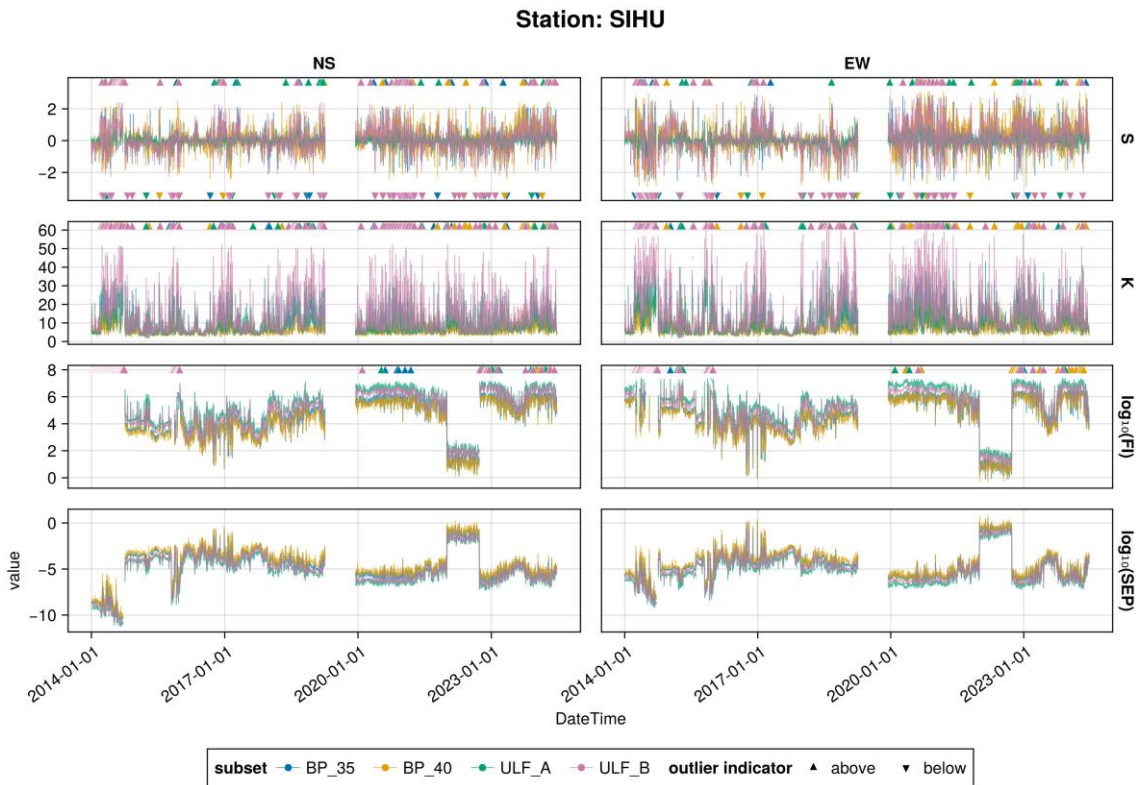


圖20: 四湖 (SIHU) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

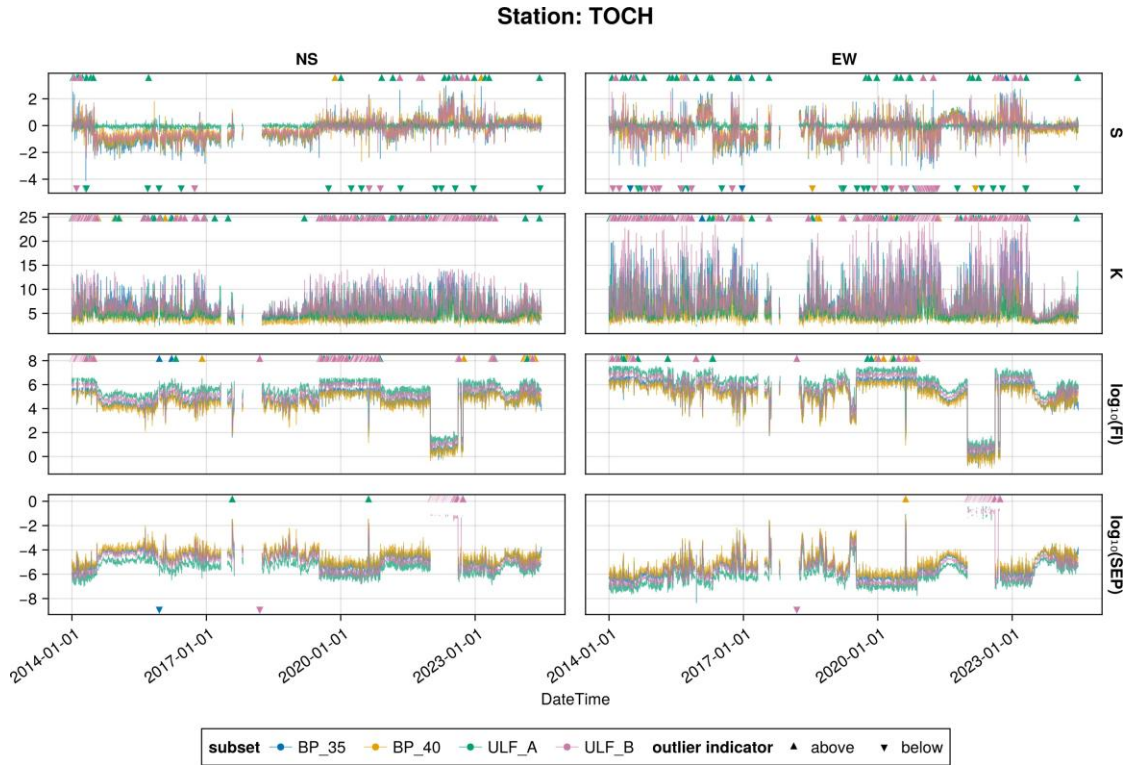


圖21: 頭城 (TOCH) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

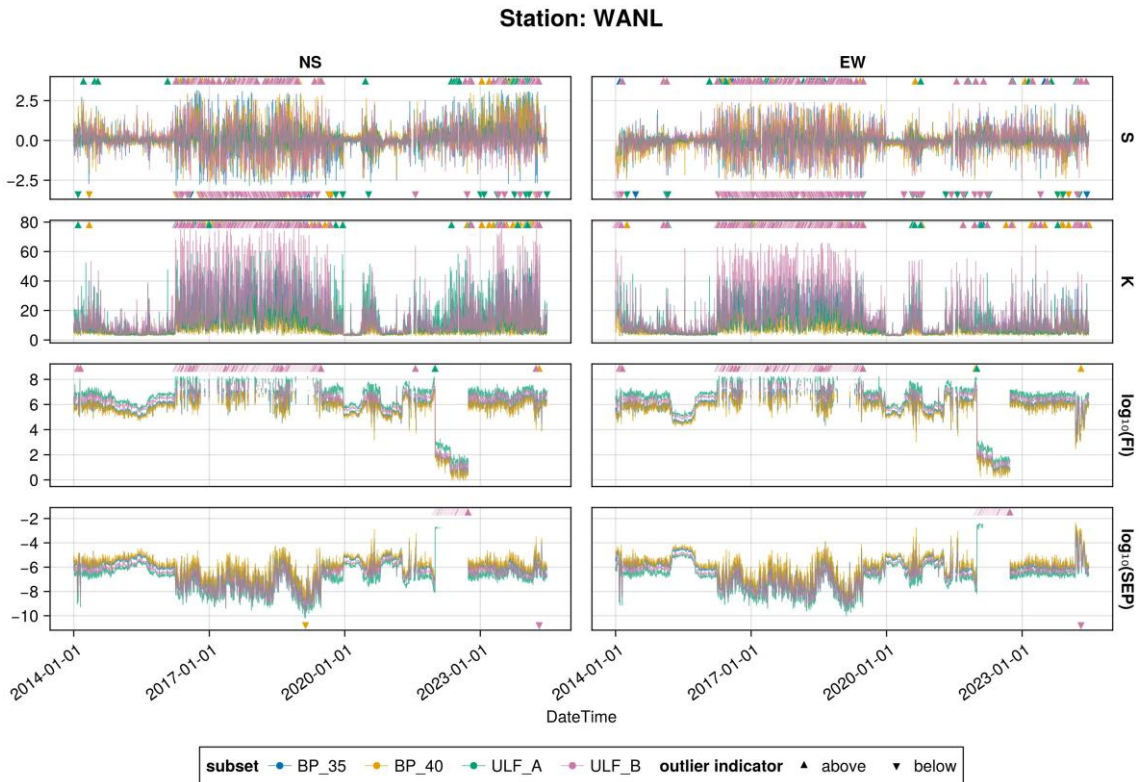


圖22: 萬巒 (WANL) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

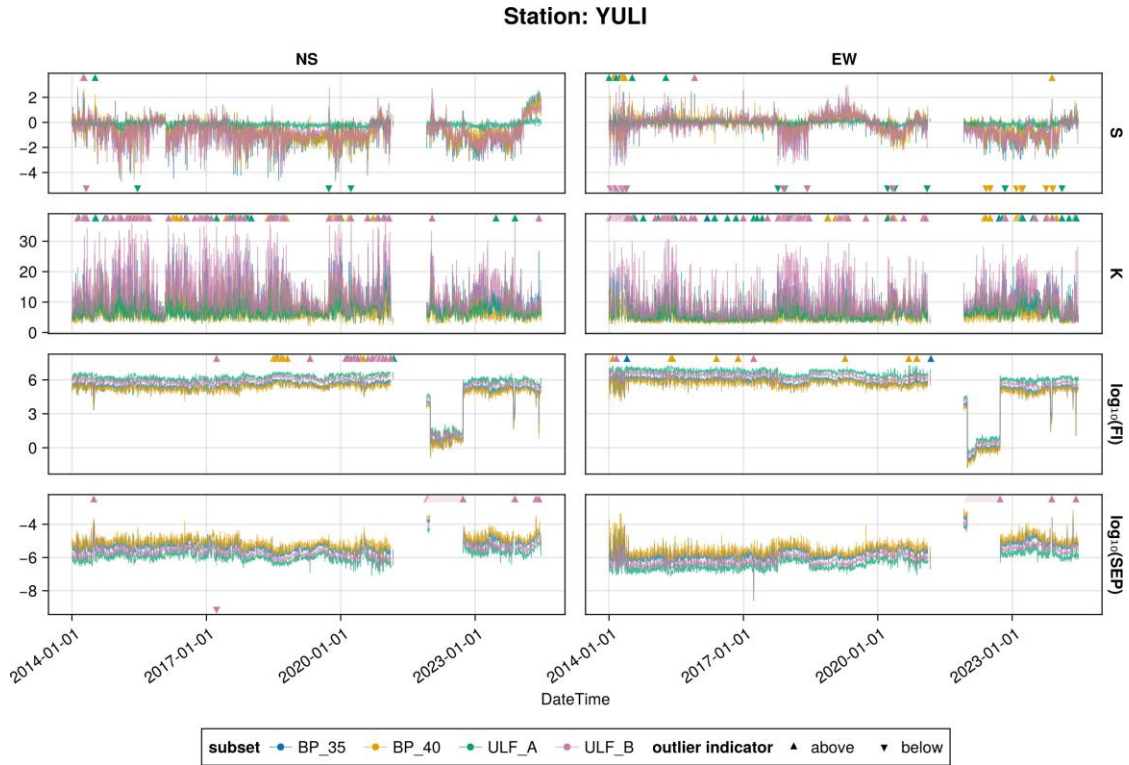


圖23: 玉里 (YULI) 地電站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

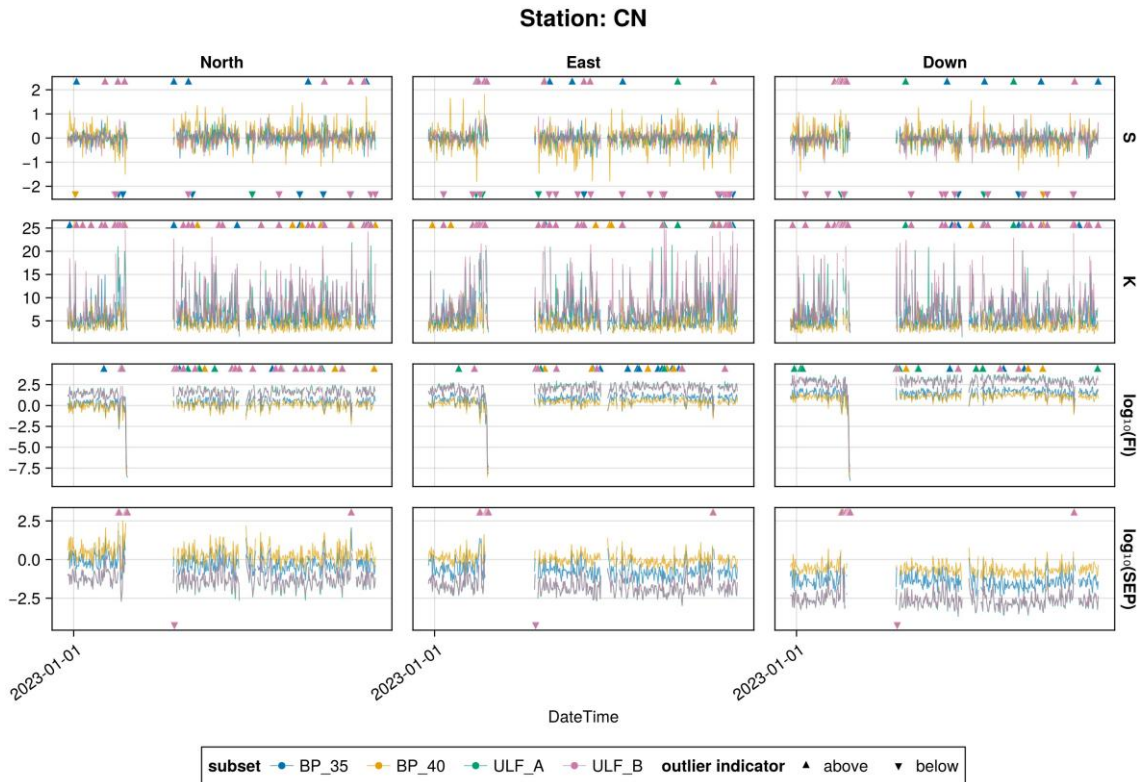


圖24: 暨南 (CN) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

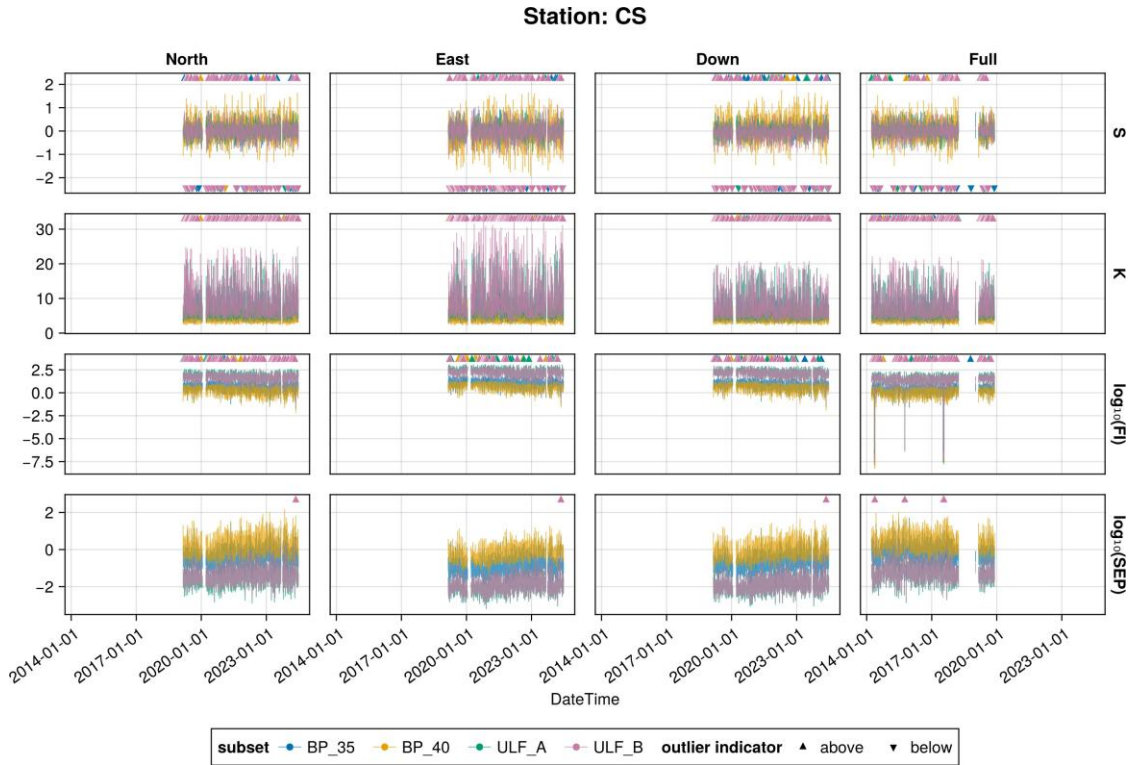


圖25: 池上 (CS) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

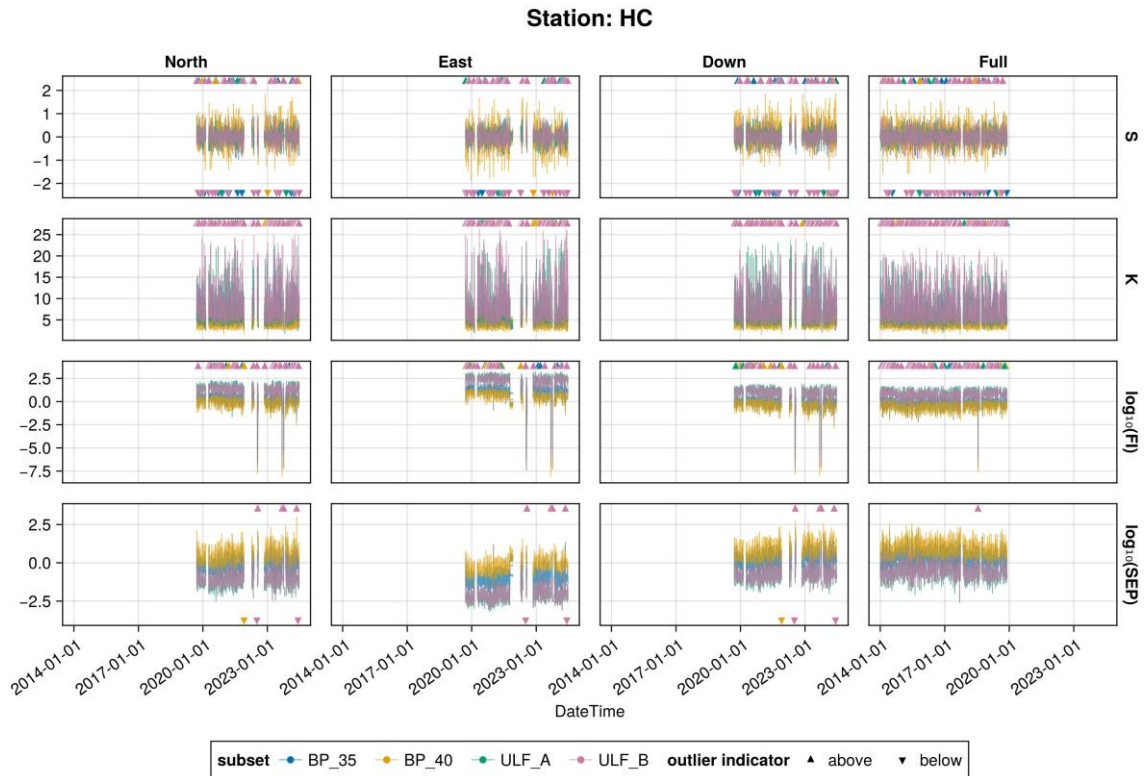


圖26: 恆春 (HC) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

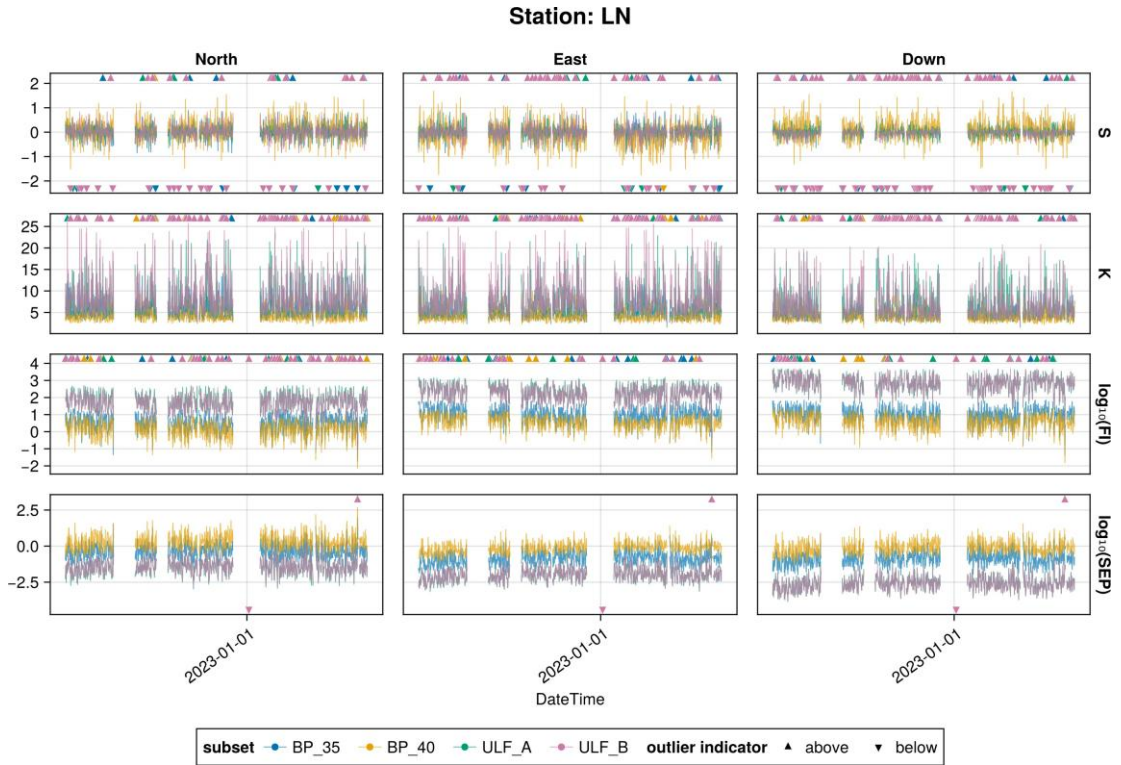


圖27: 蘭嶼 (LN) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

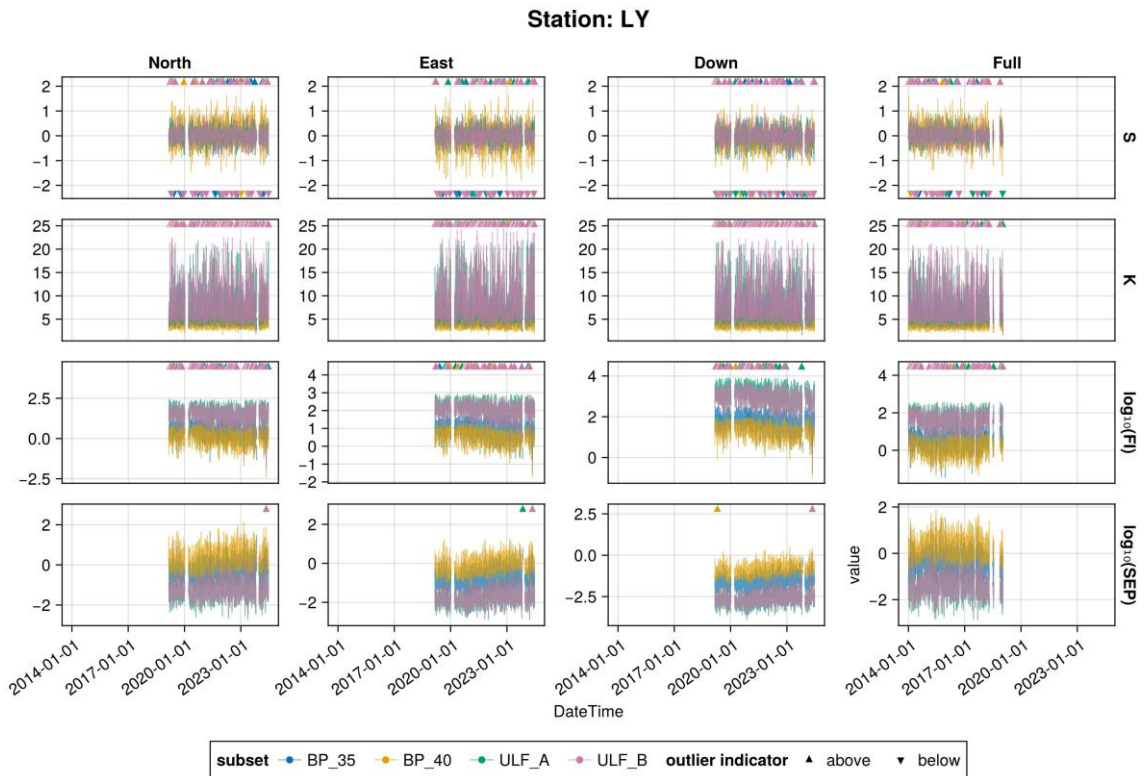


圖28: 鯉魚潭 (LY) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

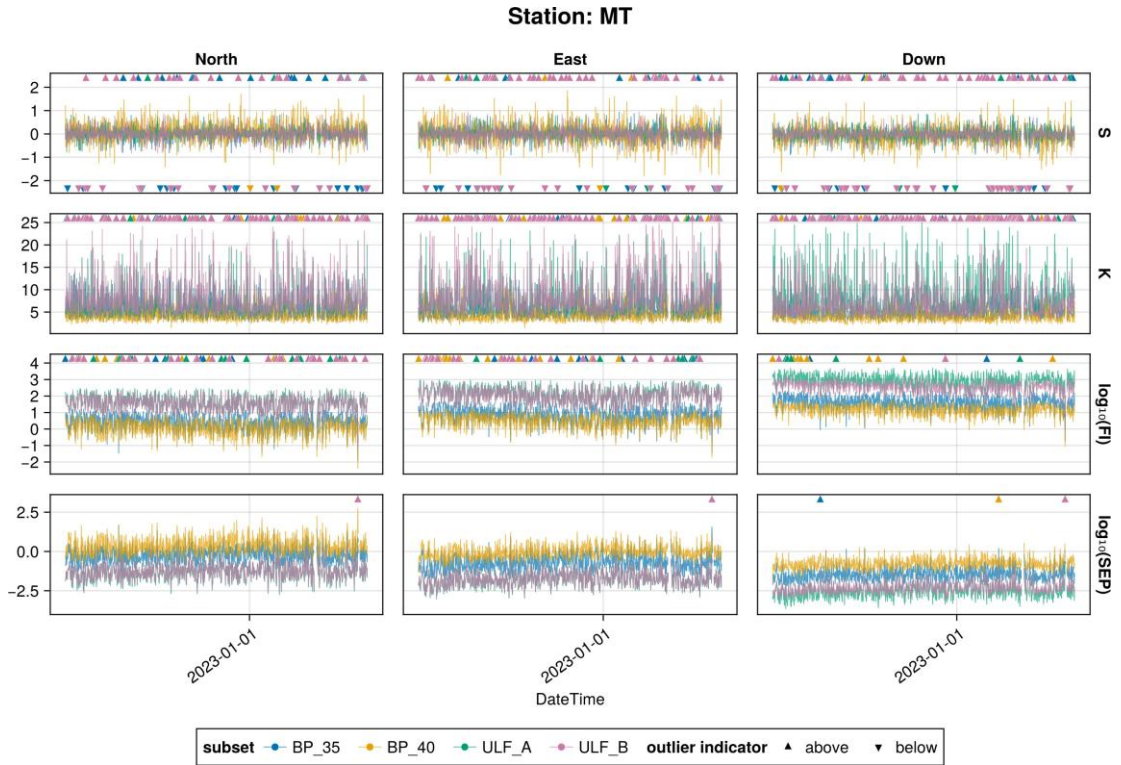


圖29: 馬祖 (MT) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

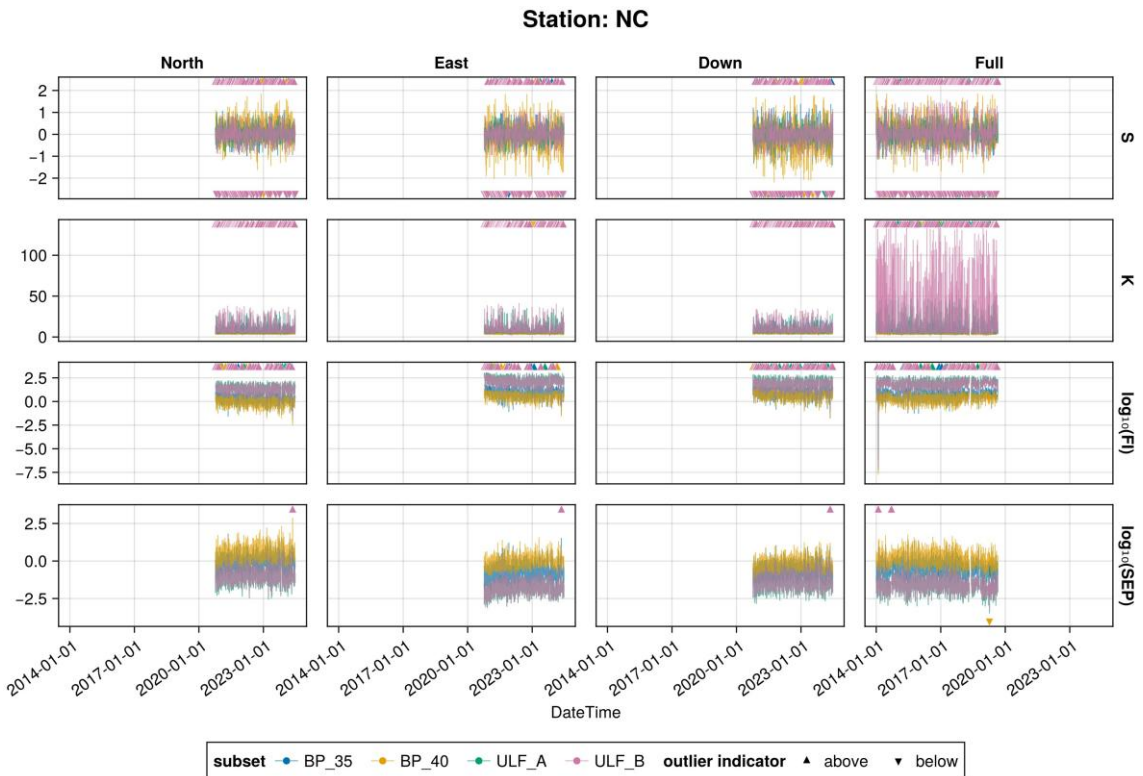


圖30: 內城 (NC) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

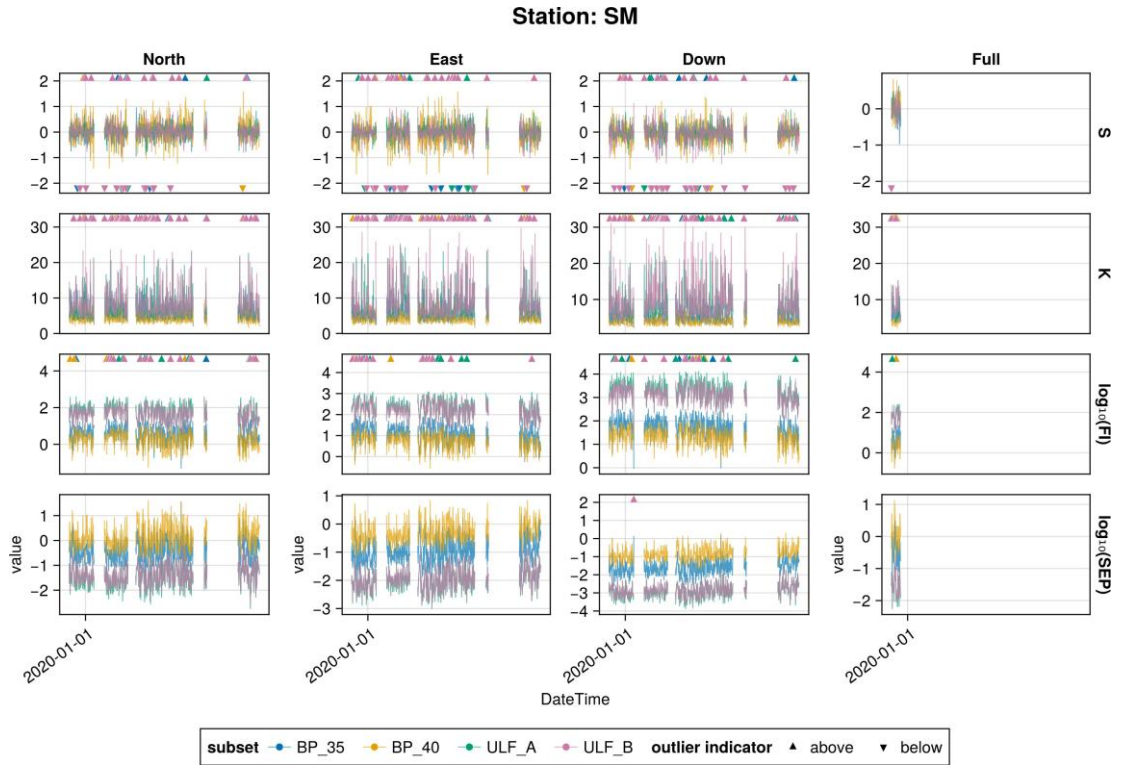


圖31: 日月潭 (SM) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

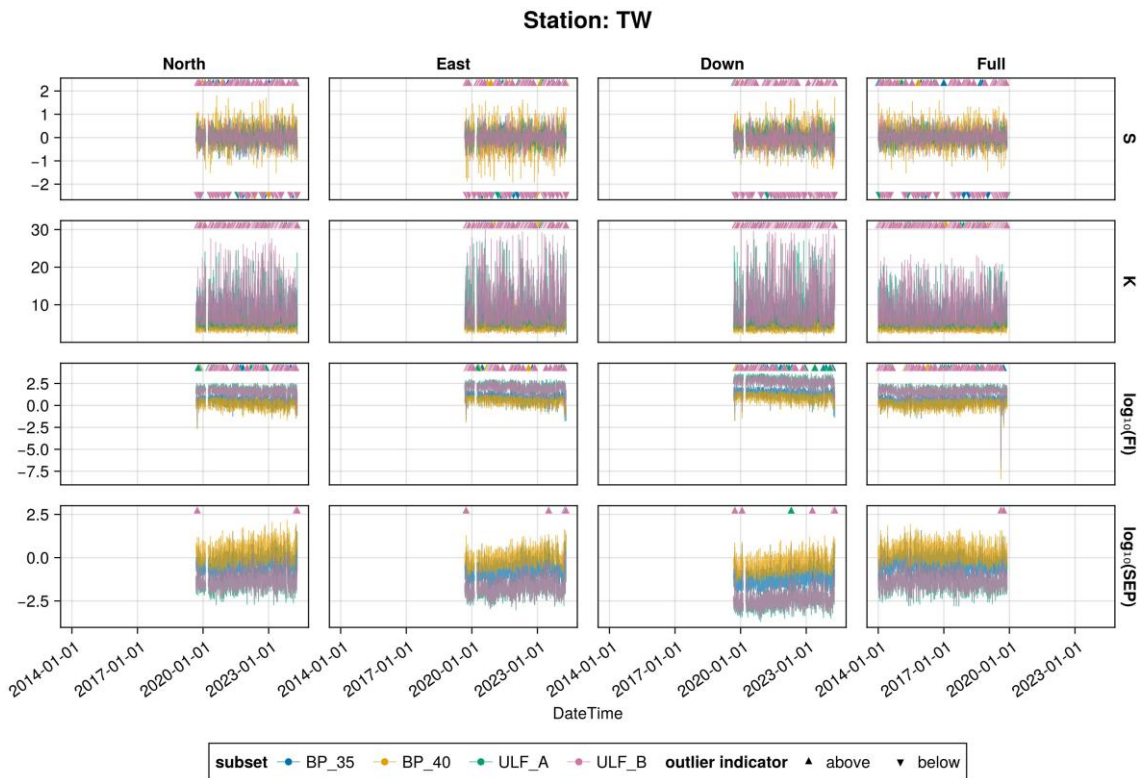


圖32: 灣丘 (TW) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

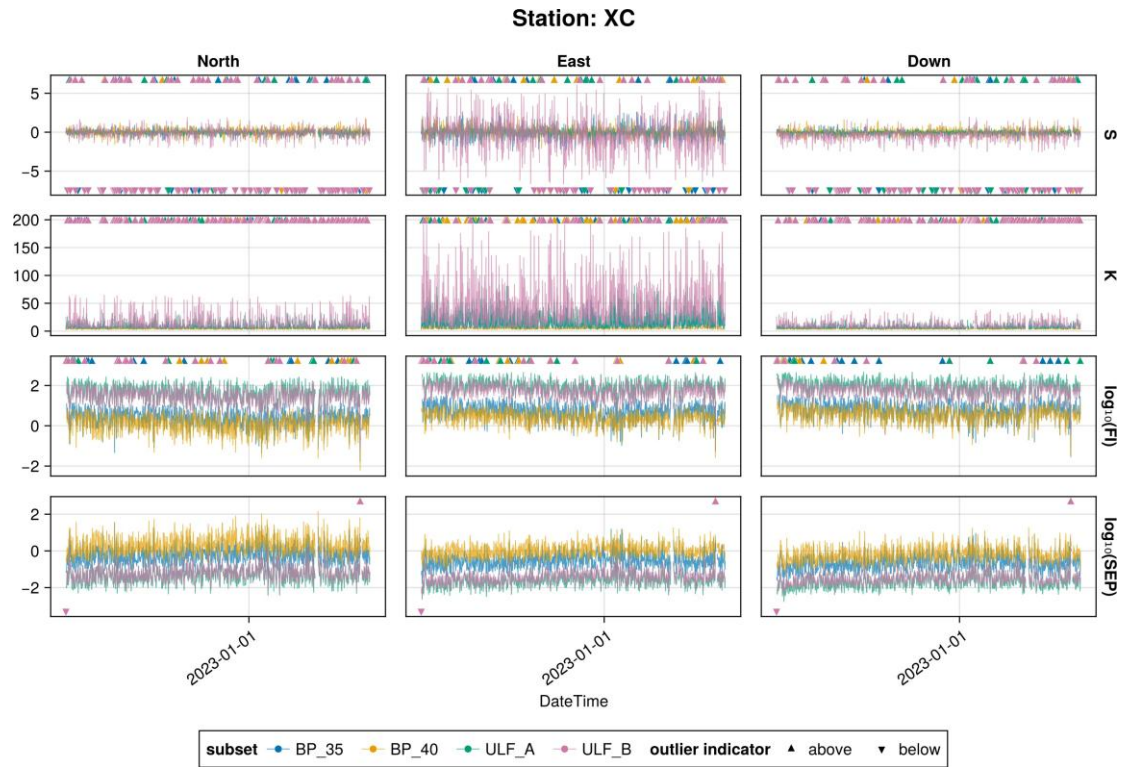


圖33: 新城 (XC) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

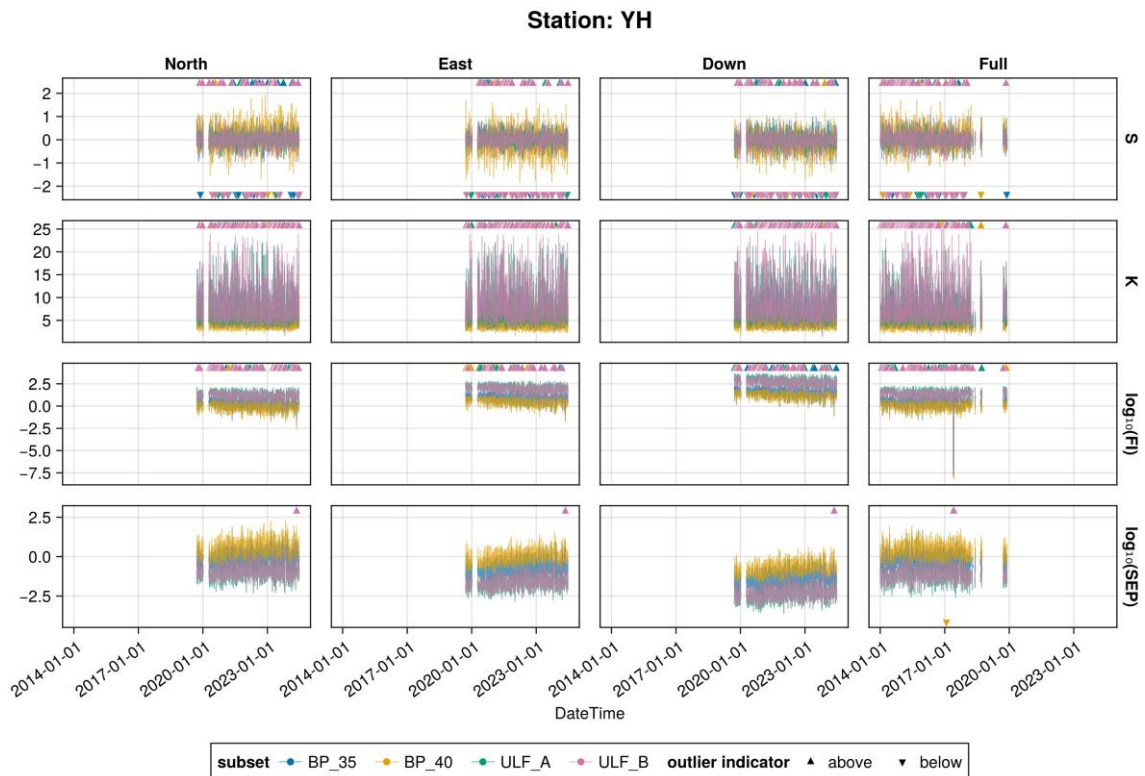


圖34: 爺亨 (YH) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

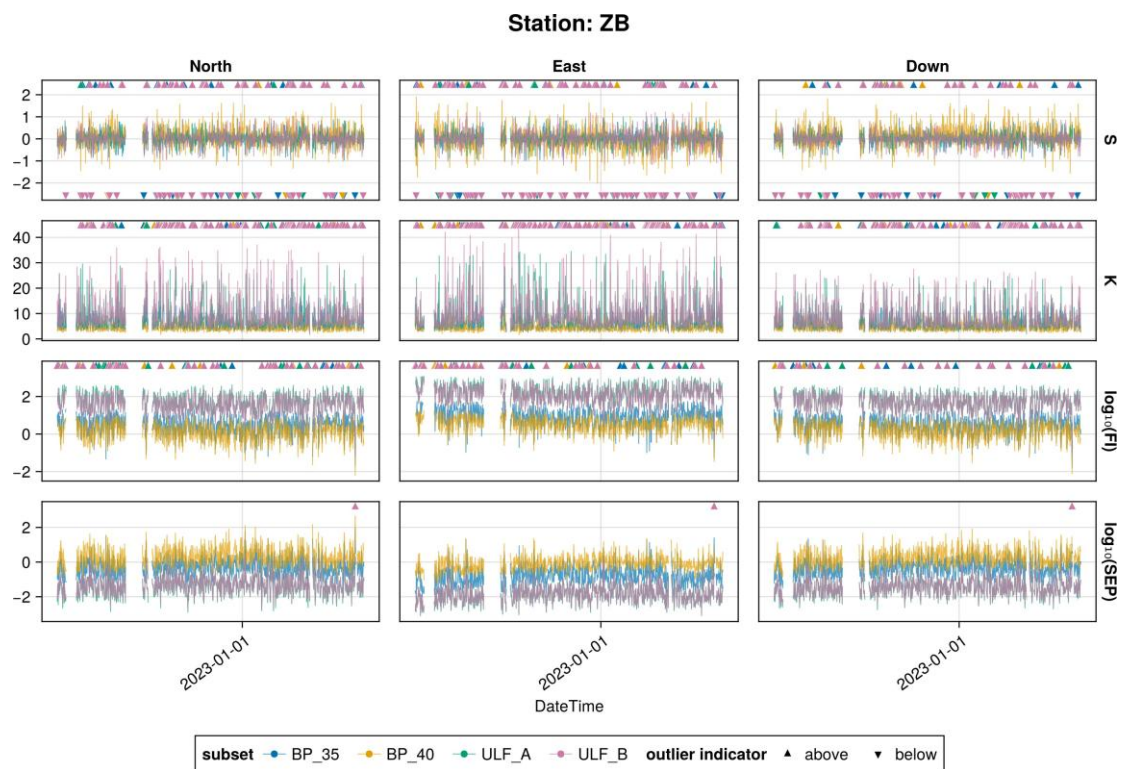


圖35: 知本 (ZB) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

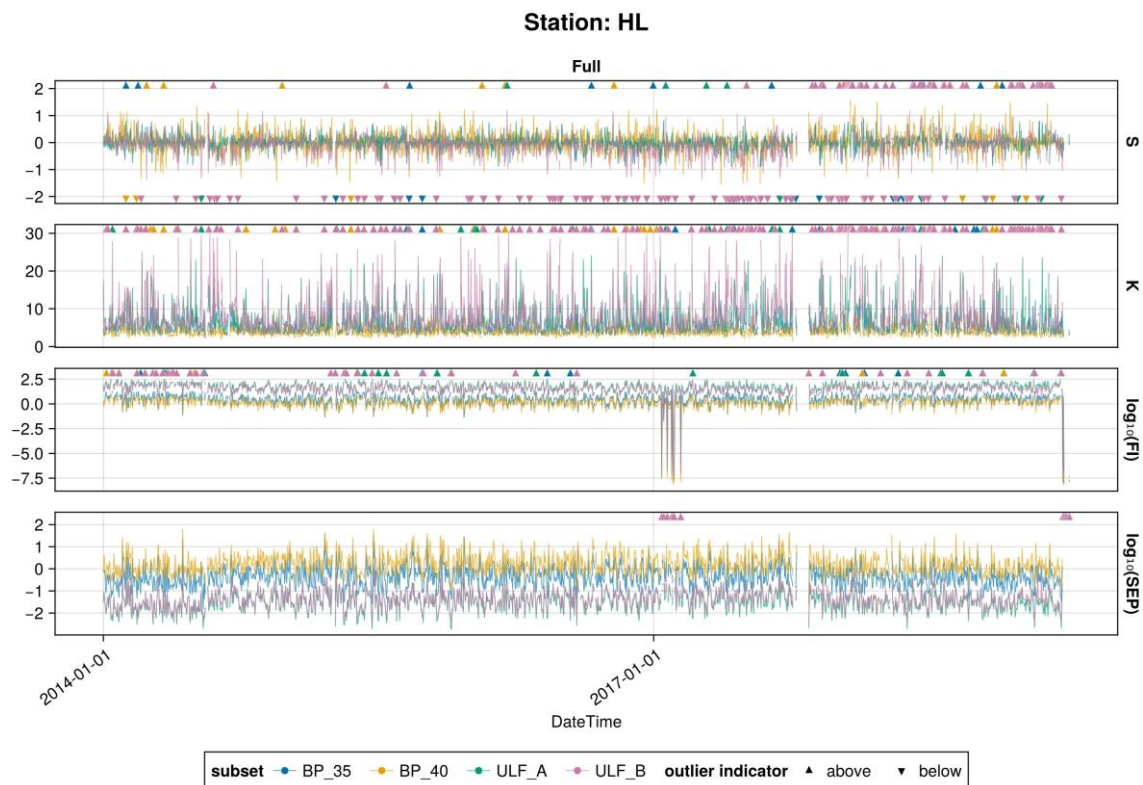


圖36: 加灣 (HL) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

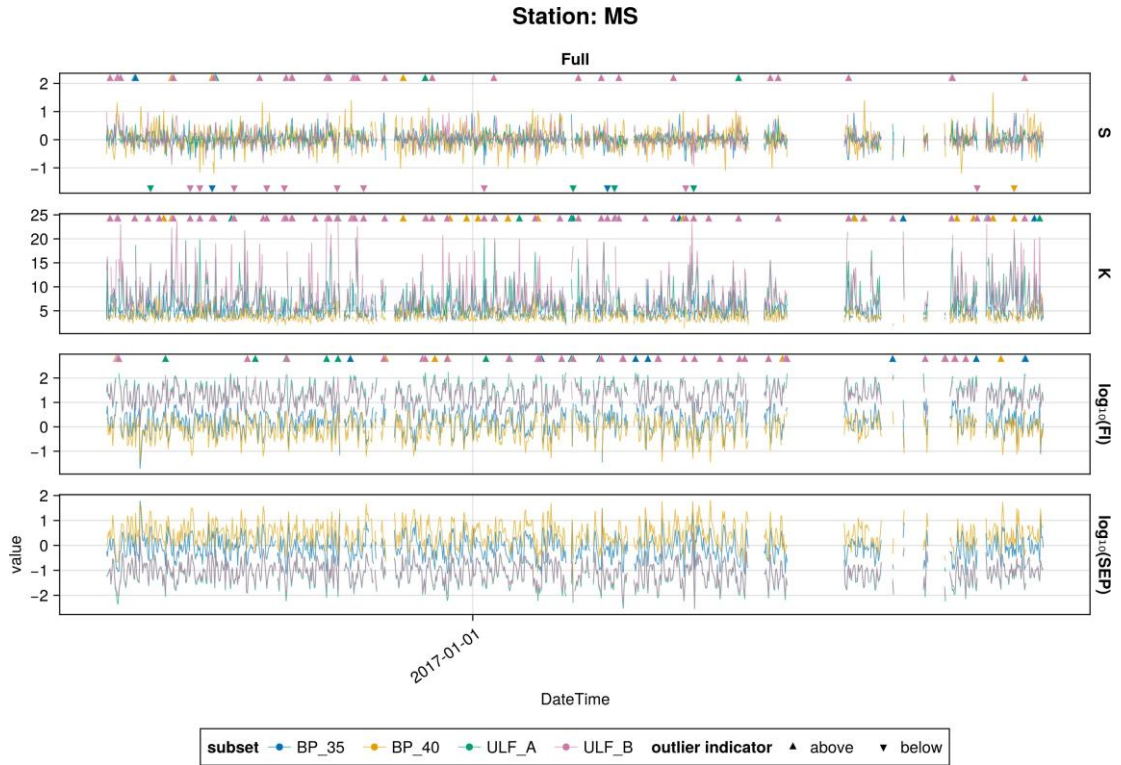


圖37: 馬仕 (MS) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

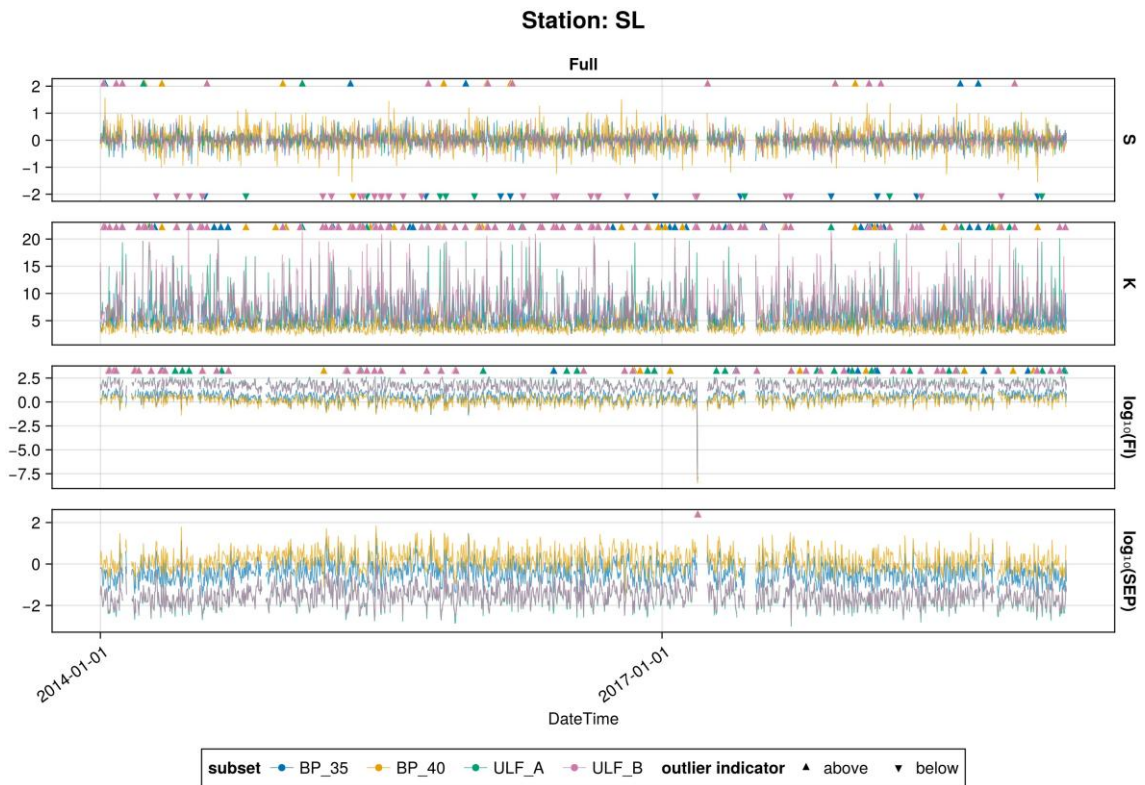


圖38: 雙龍 (SL) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

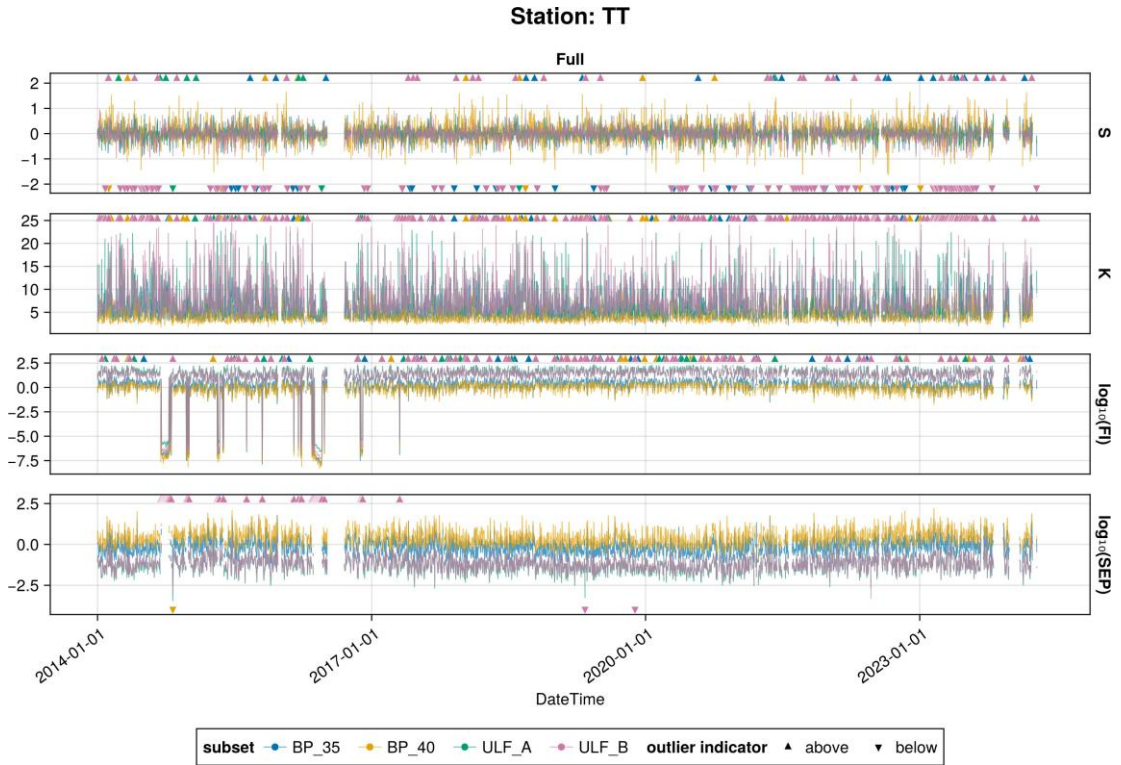


圖39: 卑南 (TT) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

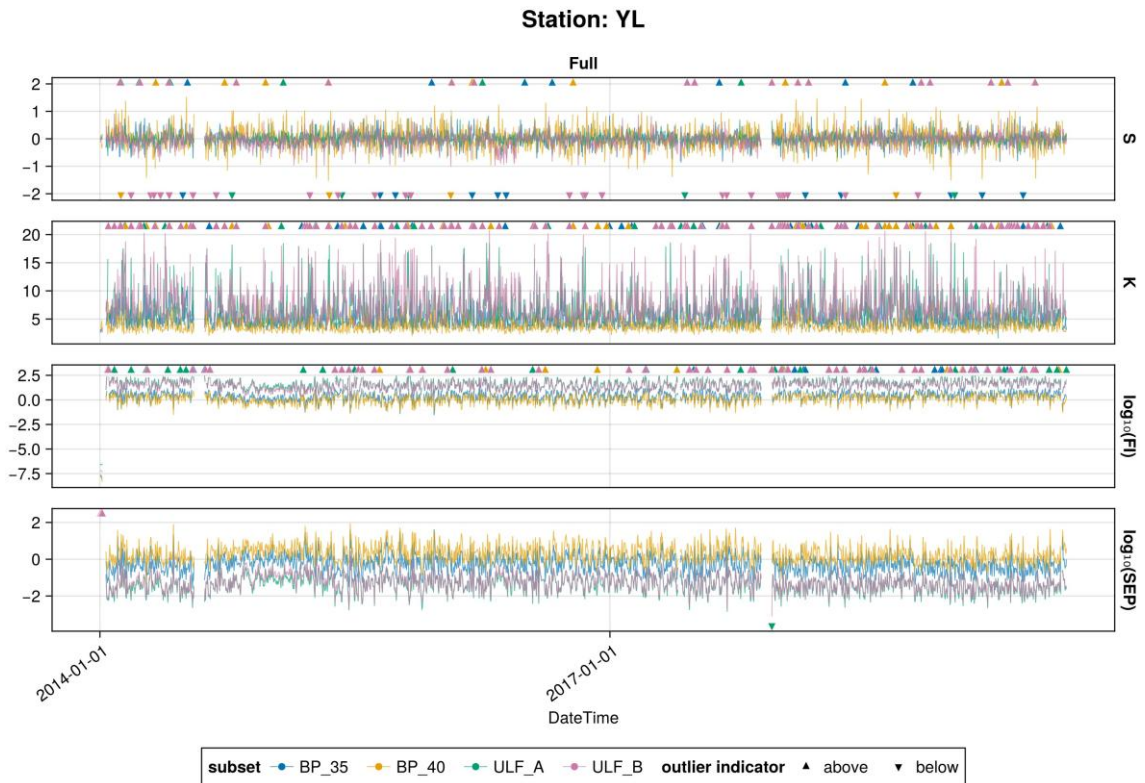


圖40: 玉里 (YL) 地磁站 各濾波頻段資料的統計指標與資訊理論指標時序圖。

Original ($Y_t = 0.7Y_{t-1} + \varepsilon_t$) vs its Shifted version

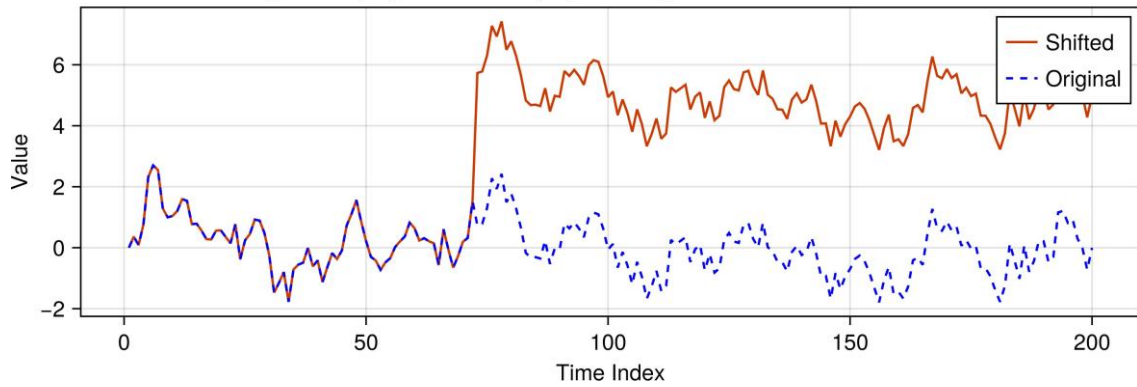


圖41: 範例 AR(1) 樣本路徑及其平移版本。

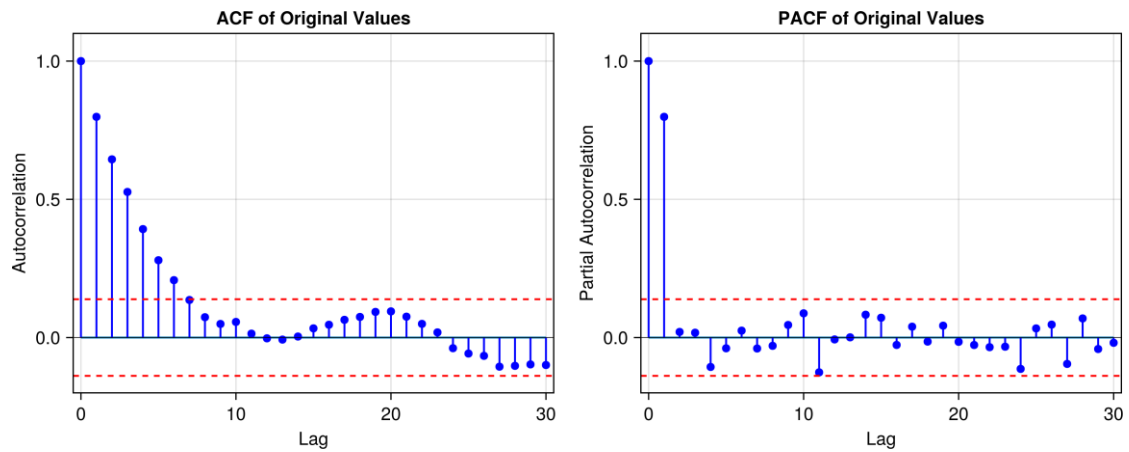


圖42: 圖41 中的原始 AR(1) 樣本路徑的樣本 ACF、PACF 函數圖。

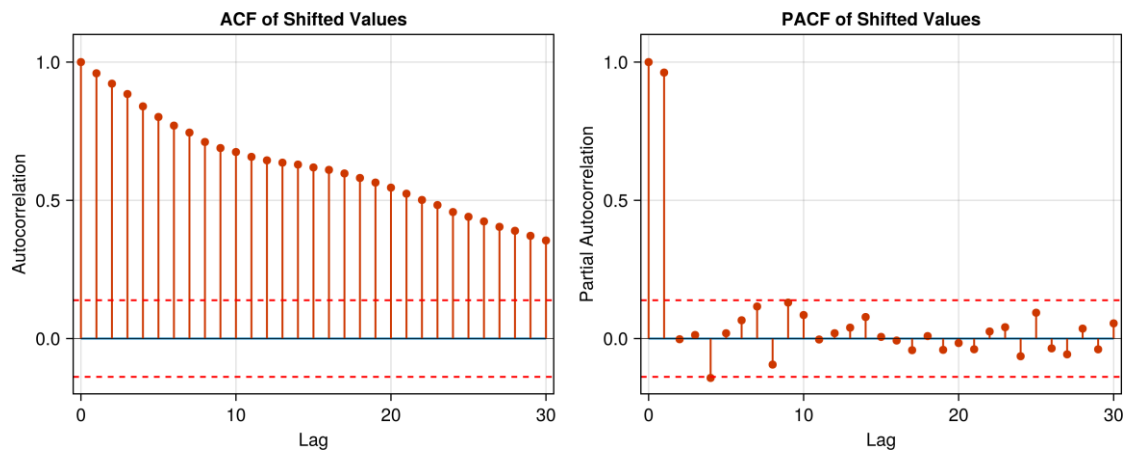
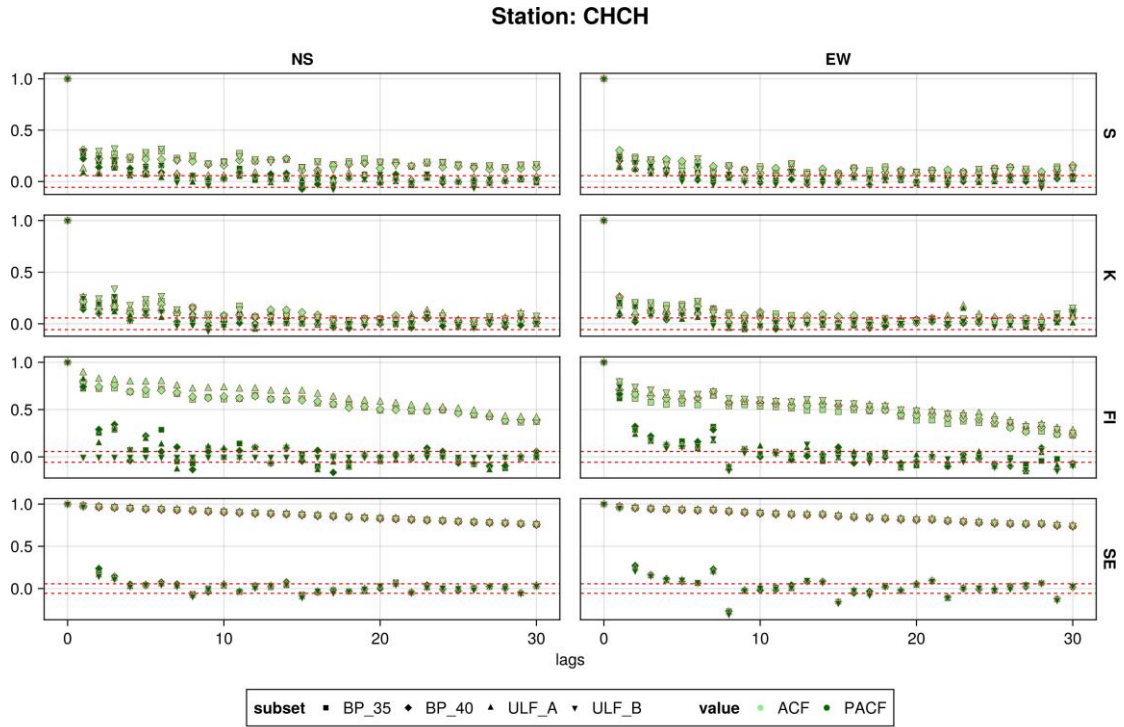
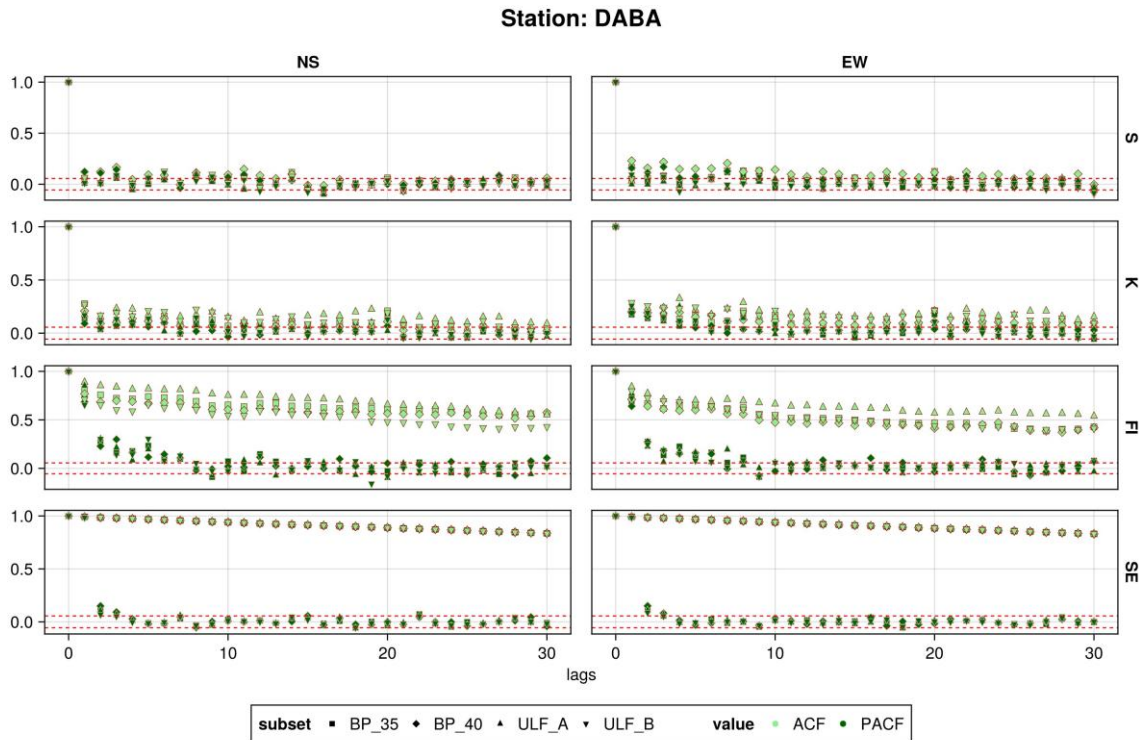


圖43: 圖41 中的平移版本 AR(1) 樣本路徑的樣本 ACF、PACF 函數圖。



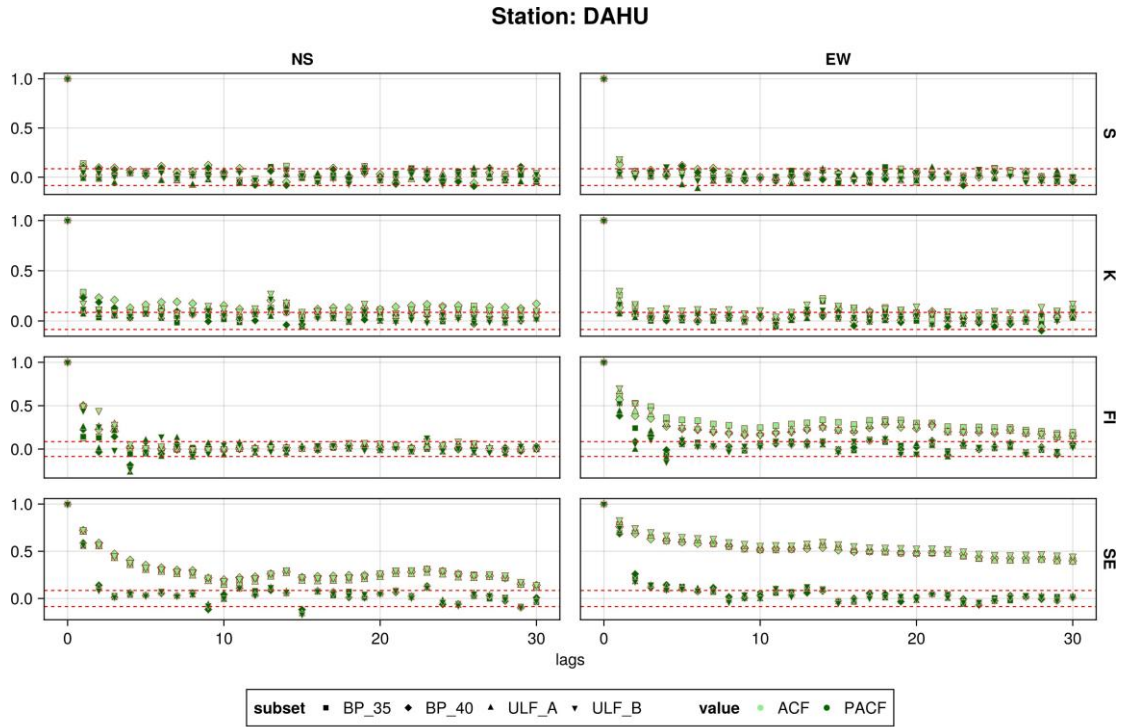
¹ACF/PACF is calculated on the continuous time-series from 2020-08-23 to 2023-12-02

圖44: 將軍 (CHCH) 地電站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



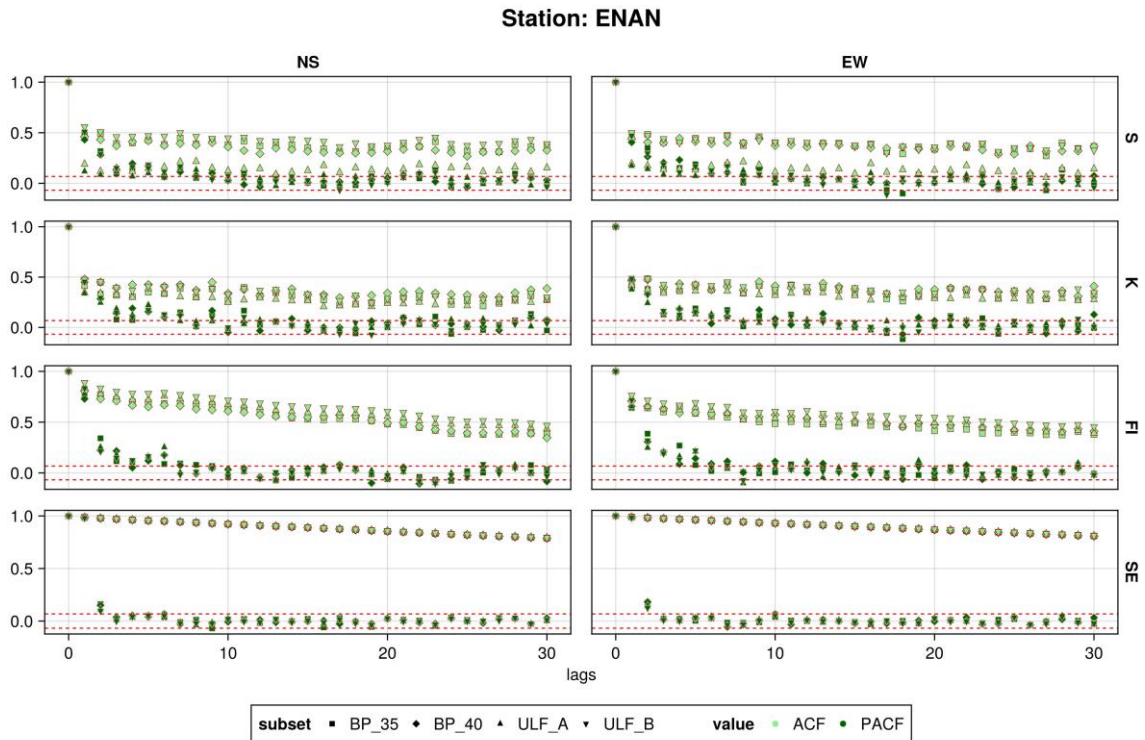
¹ACF/PACF is calculated on the continuous time-series from 2021-01-04 to 2024-05-13

圖45: 達邦 (DABA) 地電站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



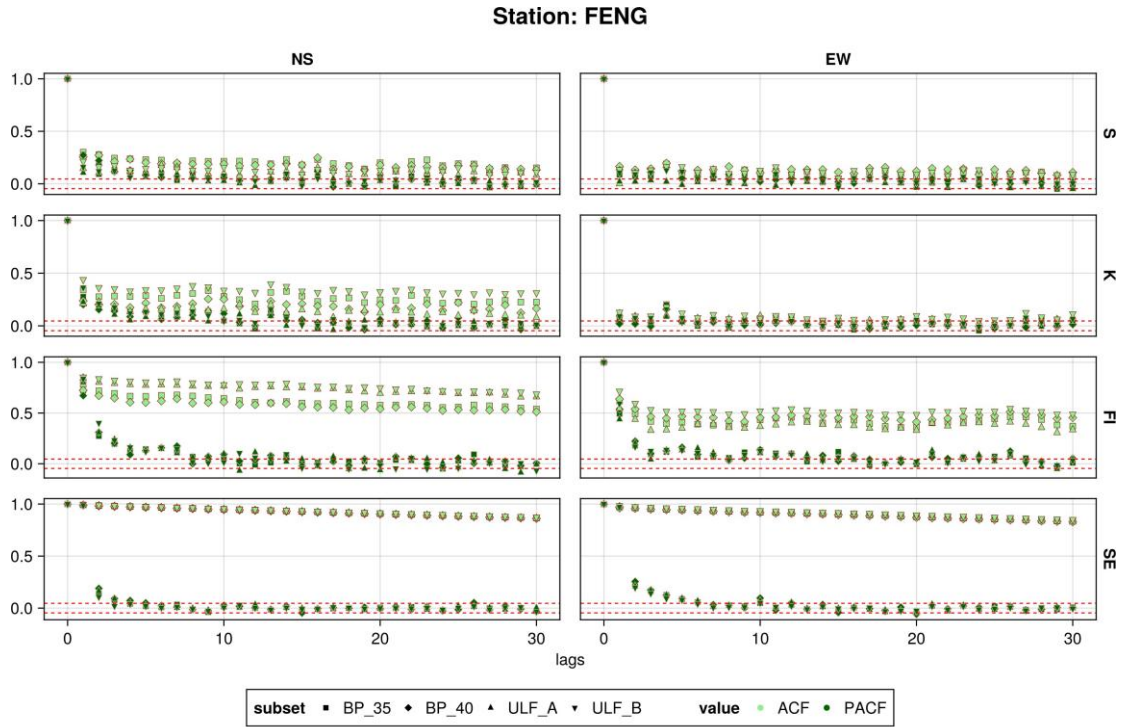
¹ACF/PACF is calculated on the continuous time-series from 2022-12-01 to 2024-05-13

圖46: 大湖 (DAHU) 地電站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



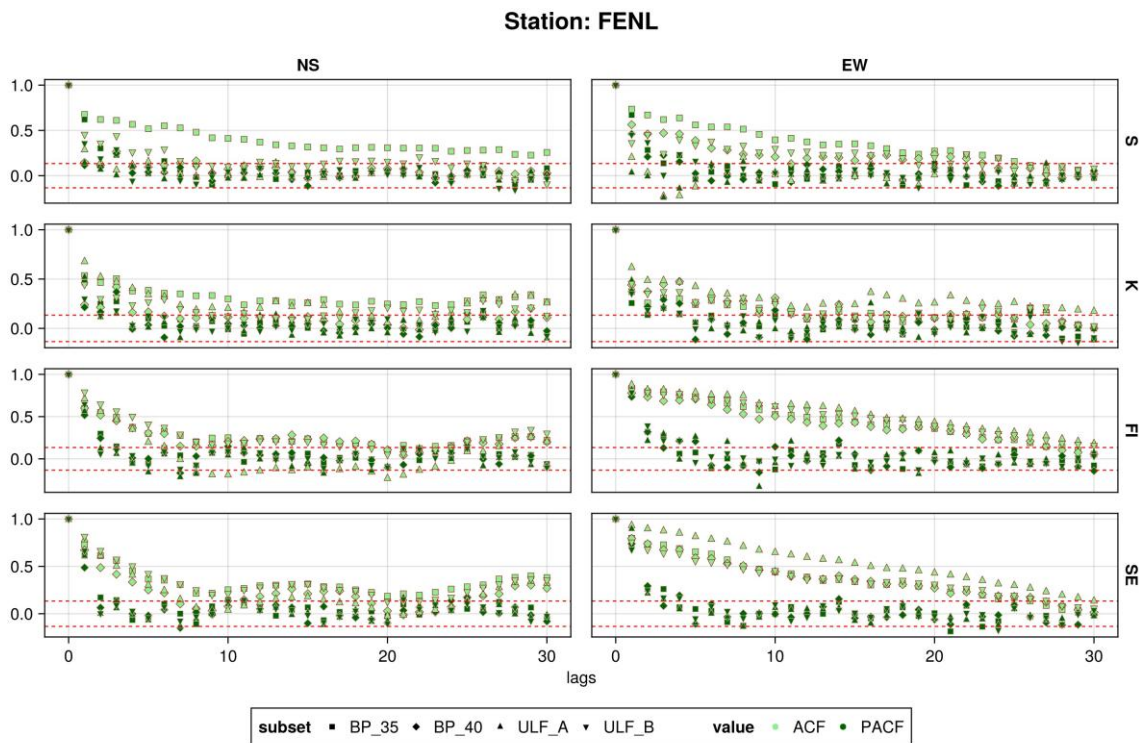
¹ACF/PACF is calculated on the continuous time-series from 2021-08-10 to 2023-11-25

圖47: 南澳 (ENAN) 地電站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



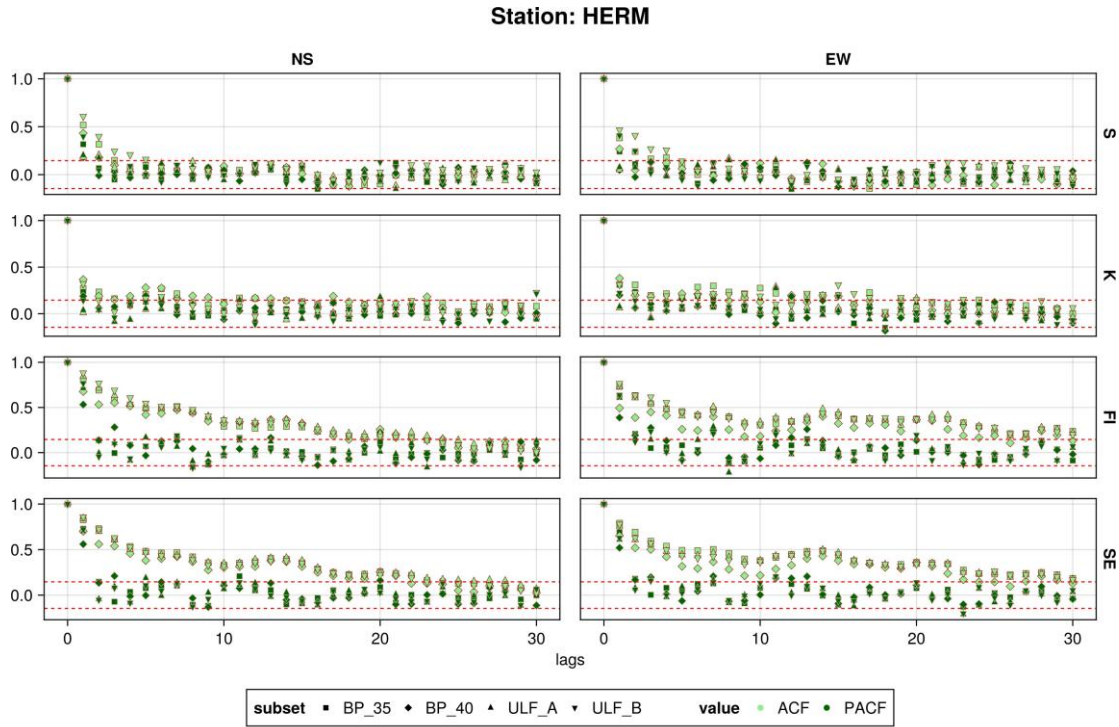
¹ACF/PACF is calculated on the continuous time-series from 2019-06-17 to 2024-05-13

圖 48: 楓港(FENG) 地電站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



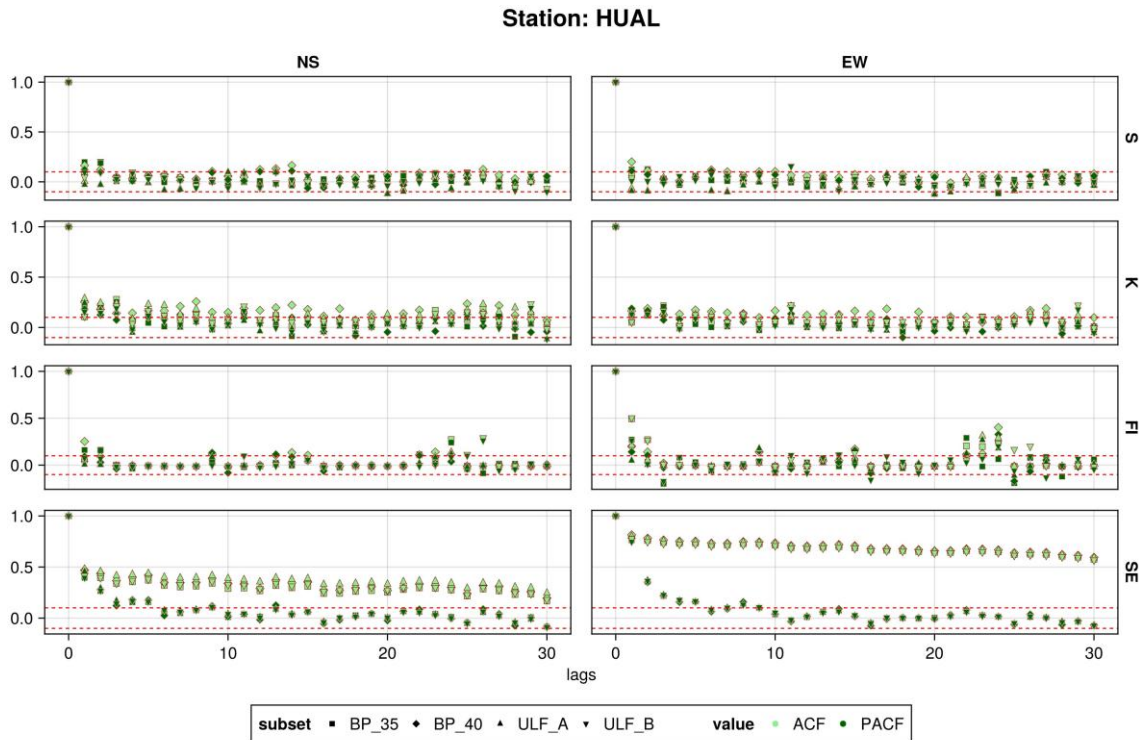
¹ACF/PACF is calculated on the continuous time-series from 2023-10-13 to 2024-05-13

圖49: 鳳林(FENL) 地電站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



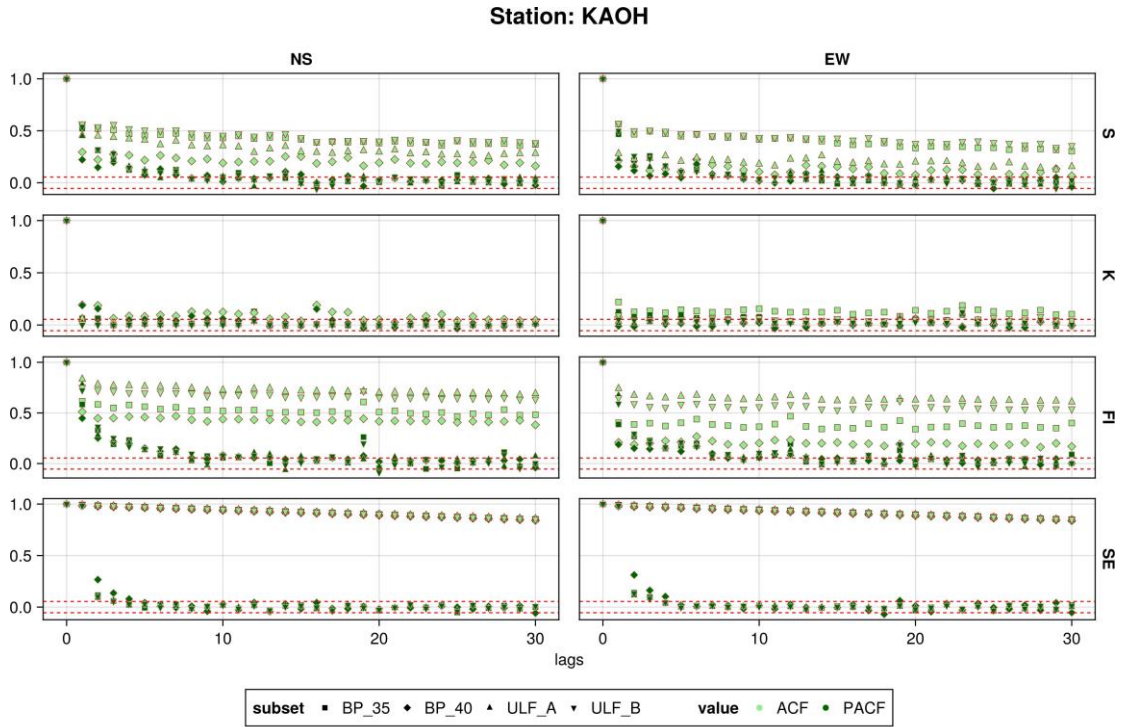
¹ACF/PACF is calculated on the continuous time-series from 2023-11-15 to 2024-05-13

圖 50: 和美 (HERM) 地電站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



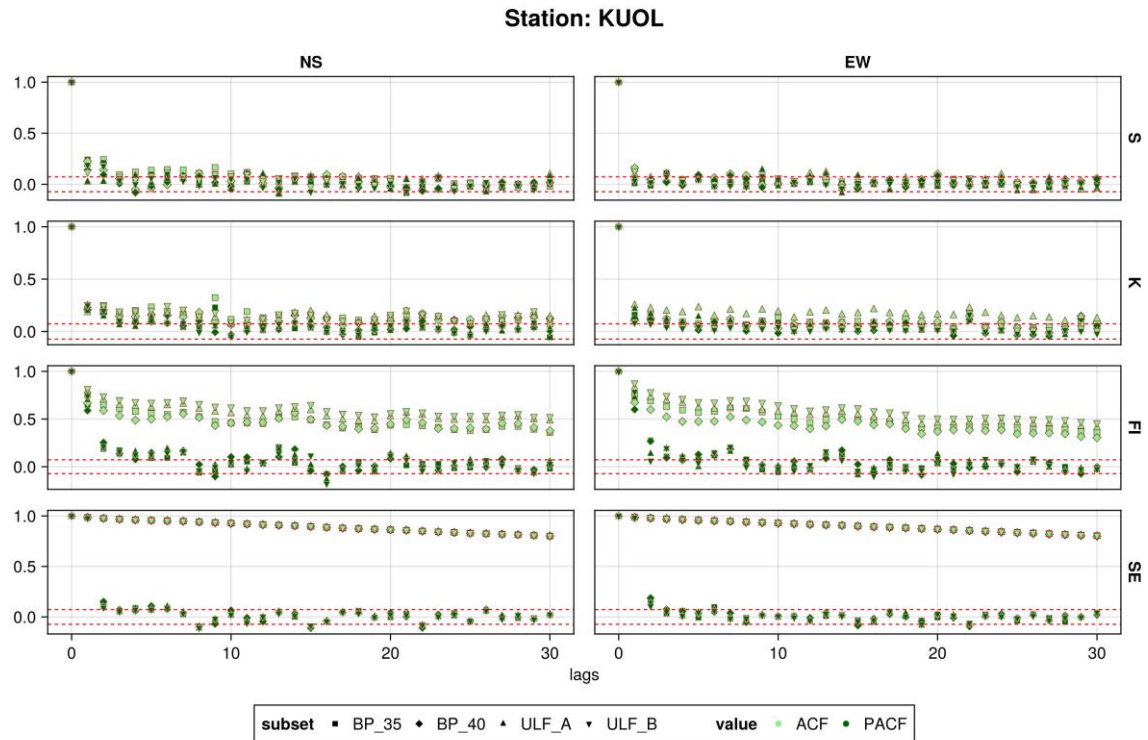
¹ACF/PACF is calculated on the continuous time-series from 2022-01-27 to 2023-02-12

圖 51: 華陵 (HUAL) 地電站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



¹ACF/PACF is calculated on the continuous time-series from 2020-11-03 to 2024-05-13

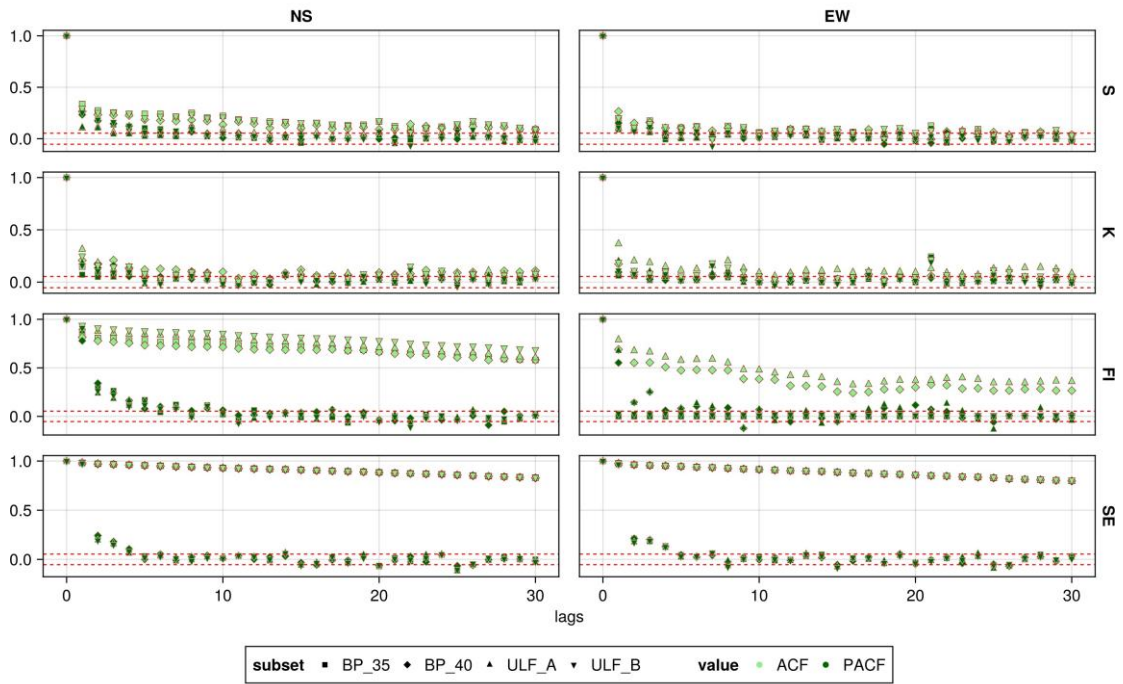
圖 52: 高雄 (KAOH) 地電站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



¹ACF/PACF is calculated on the continuous time-series from 2021-12-03 to 2023-11-23

圖 53: 過嶺 (KUOL) 地電站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。

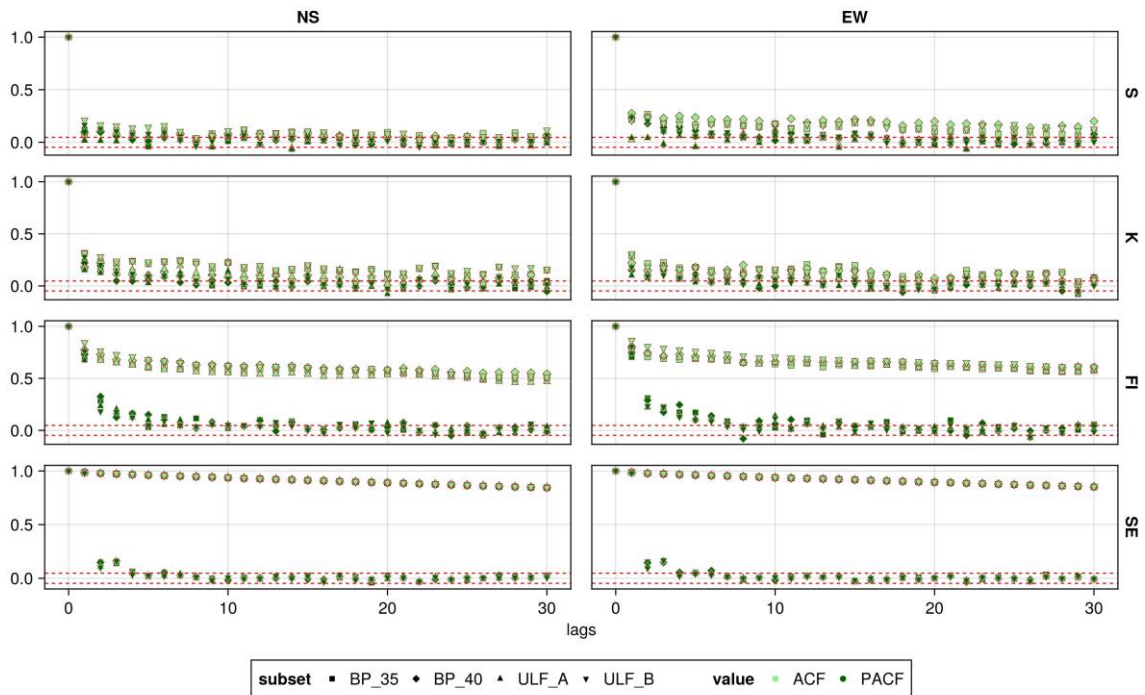
Station: LIOQ



¹ACF/PACF is calculated on the continuous time-series from 2020-09-21 to 2024-05-13

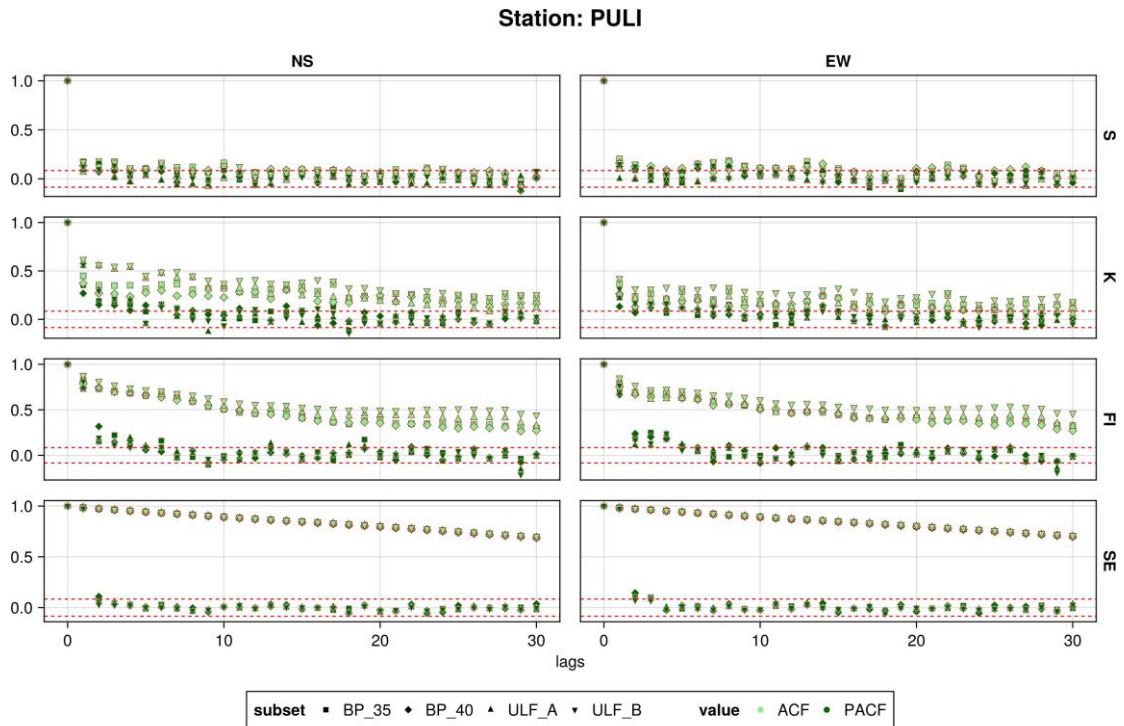
圖54: 六龜(LIOQ) 地電站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。

Station: LISH



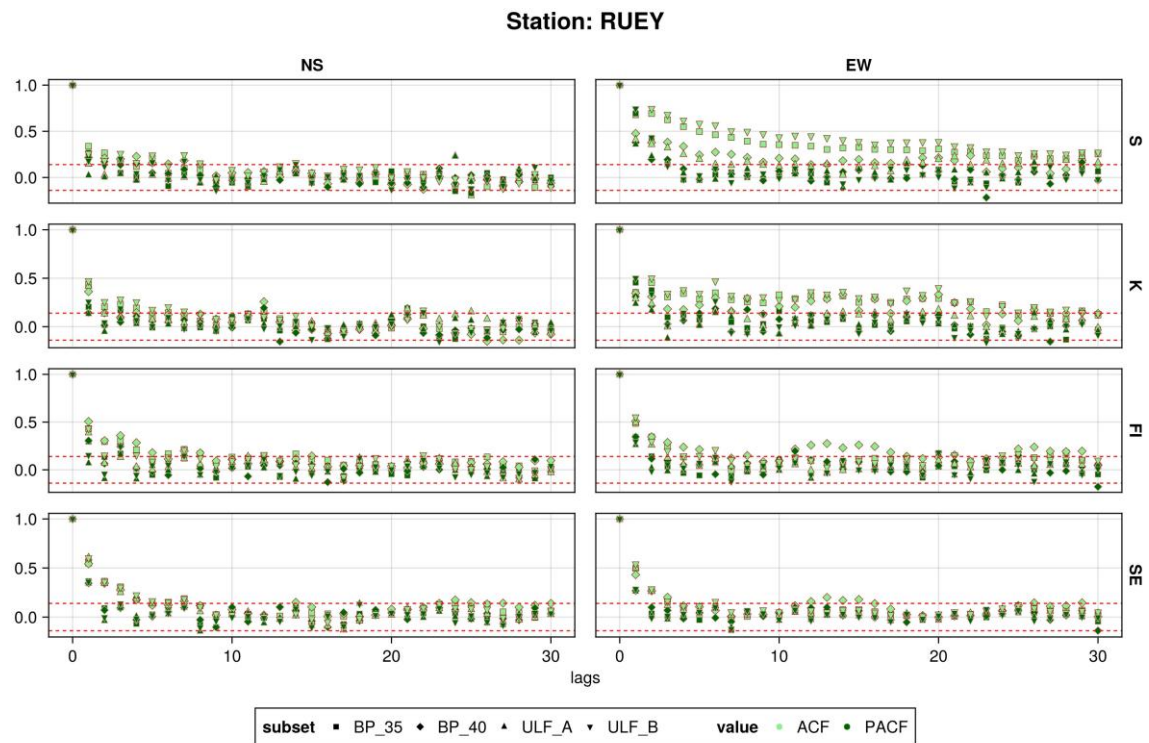
¹ACF/PACF is calculated on the continuous time-series from 2019-09-05 to 2024-05-13

圖55: 梨山(LISH) 地電站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



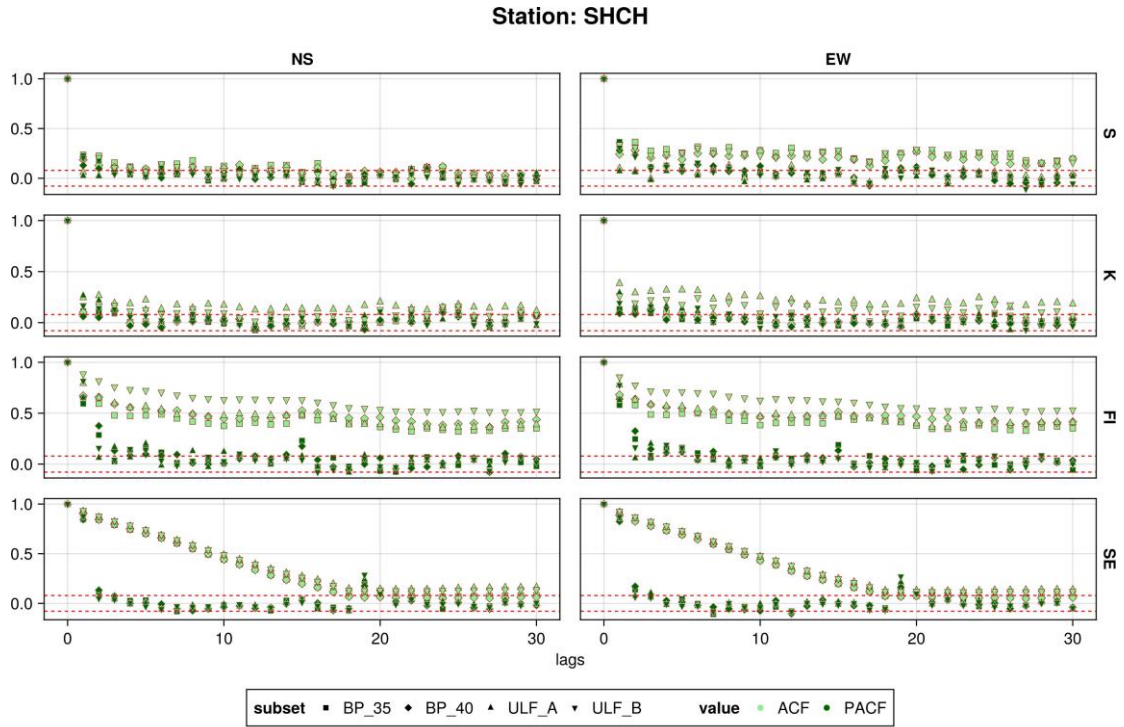
¹ACF/PACF is calculated on the continuous time-series from 2022-06-06 to 2023-11-24

圖 56: 埔里 (PULI) 地電站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



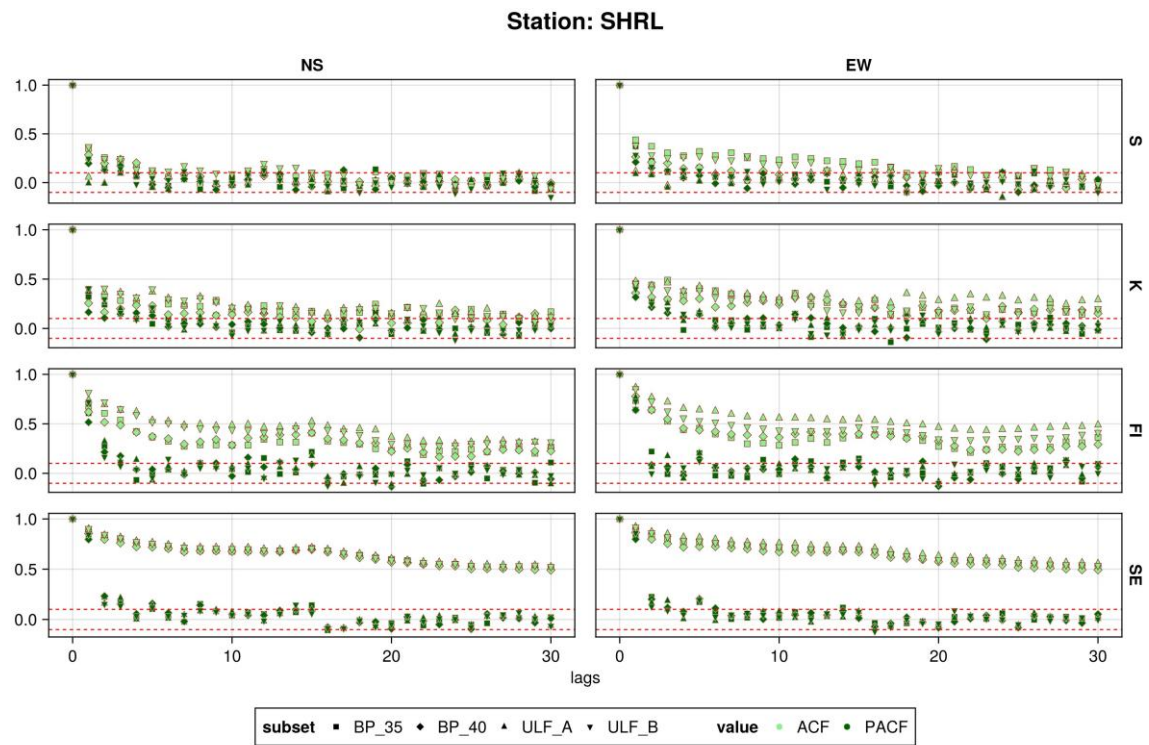
¹ACF/PACF is calculated on the continuous time-series from 2023-10-30 to 2024-05-13

圖 57: 瑞源 (RUEY) 地電站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



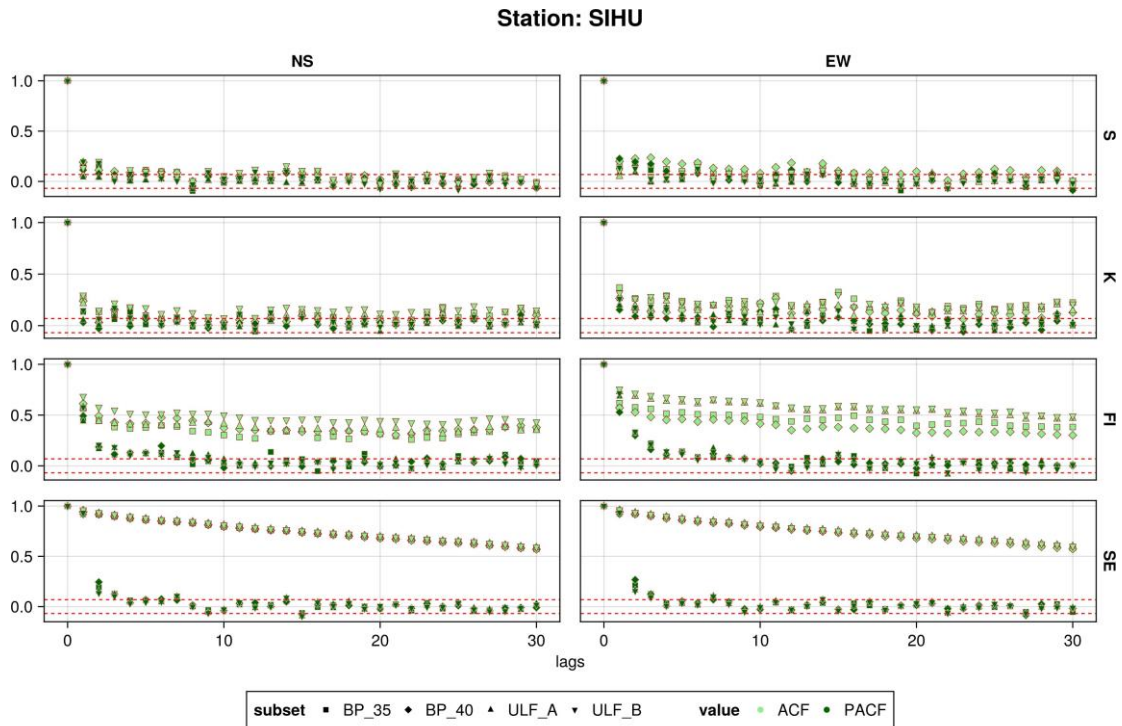
¹ACF/PACF is calculated on the continuous time-series from 2022-09-07 to 2024-05-13

圖58: 新城 (SHCH) 地電站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



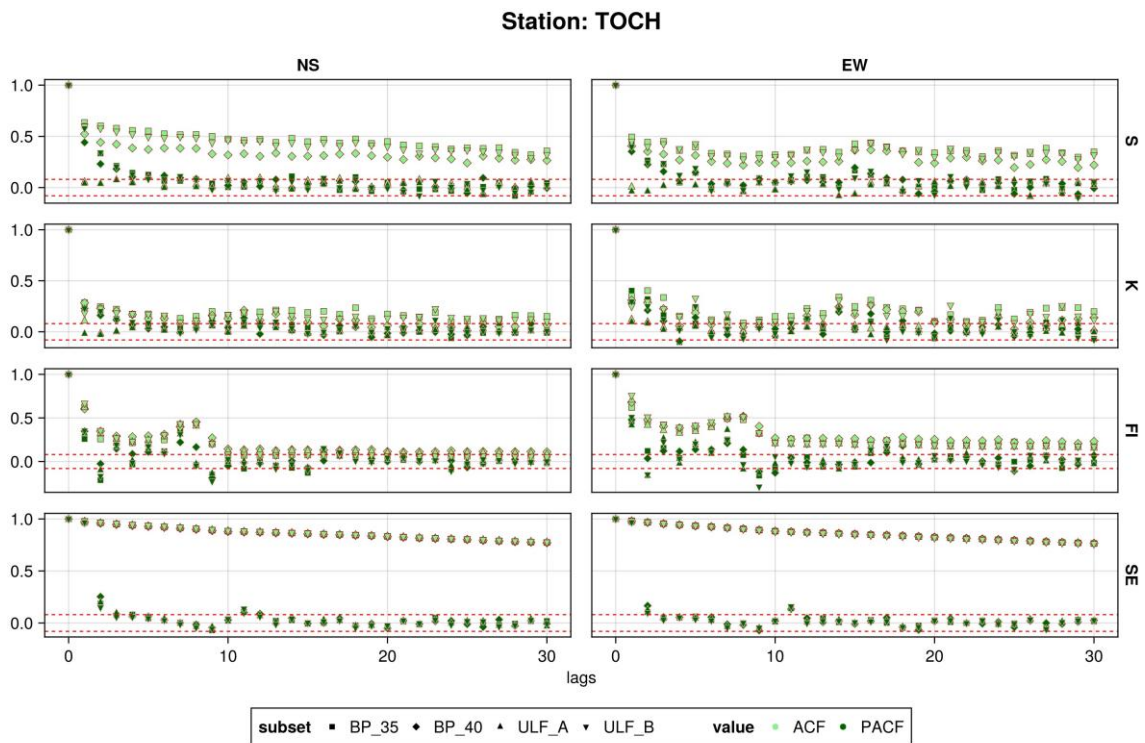
¹ACF/PACF is calculated on the continuous time-series from 2023-05-01 to 2024-05-13

圖59: 士林 (SHRL) 地電站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



¹ACF/PACF is calculated on the continuous time-series from 2019-12-11 to 2022-03-03

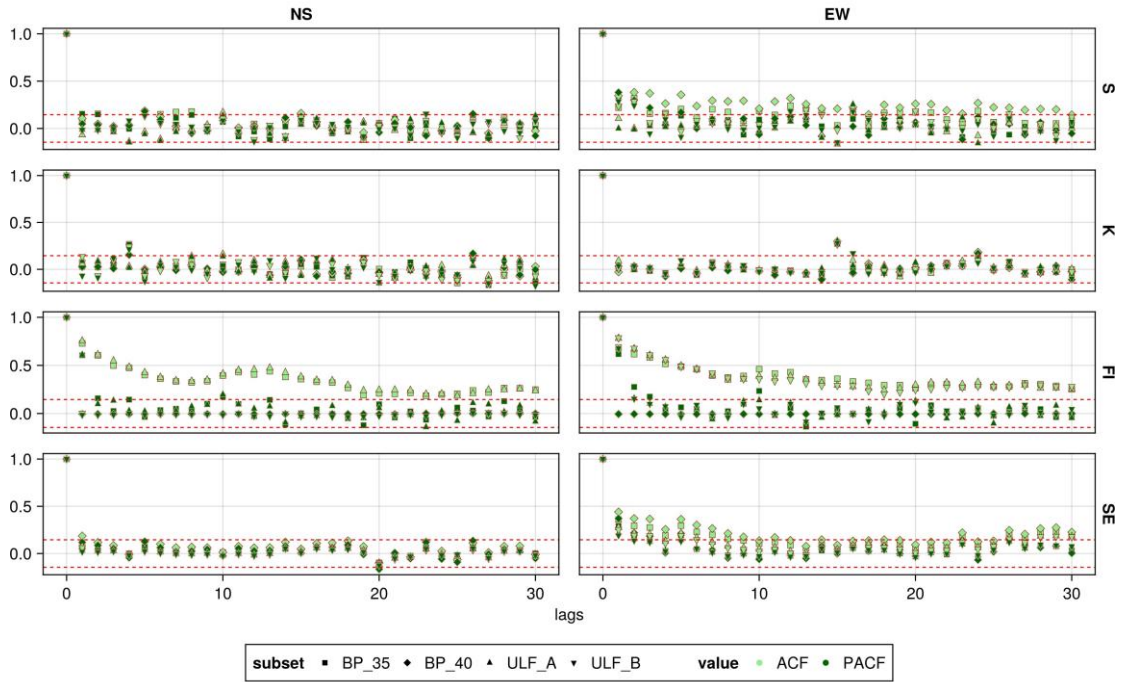
圖60: 四湖(SIHU) 地電站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



¹ACF/PACF is calculated on the continuous time-series from 2021-01-05 to 2022-08-24

圖61: 頭城(TOCH) 地電站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。

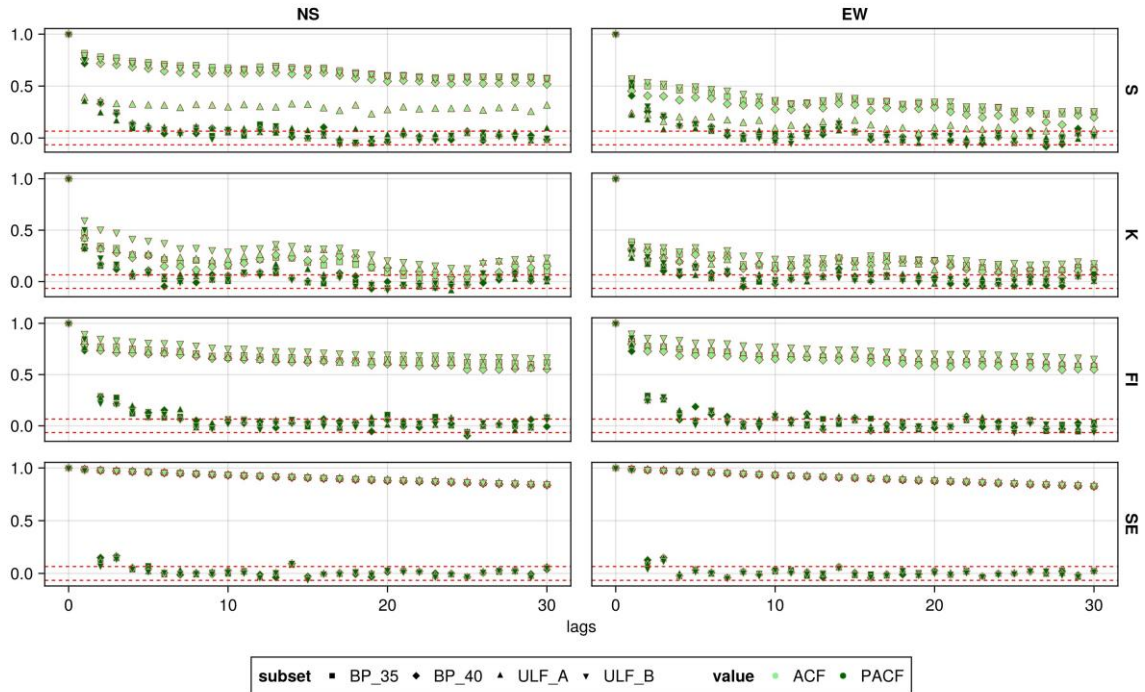
Station: WANL



¹ACF/PACF is calculated on the continuous time-series from 2023-11-14 to 2024-05-13

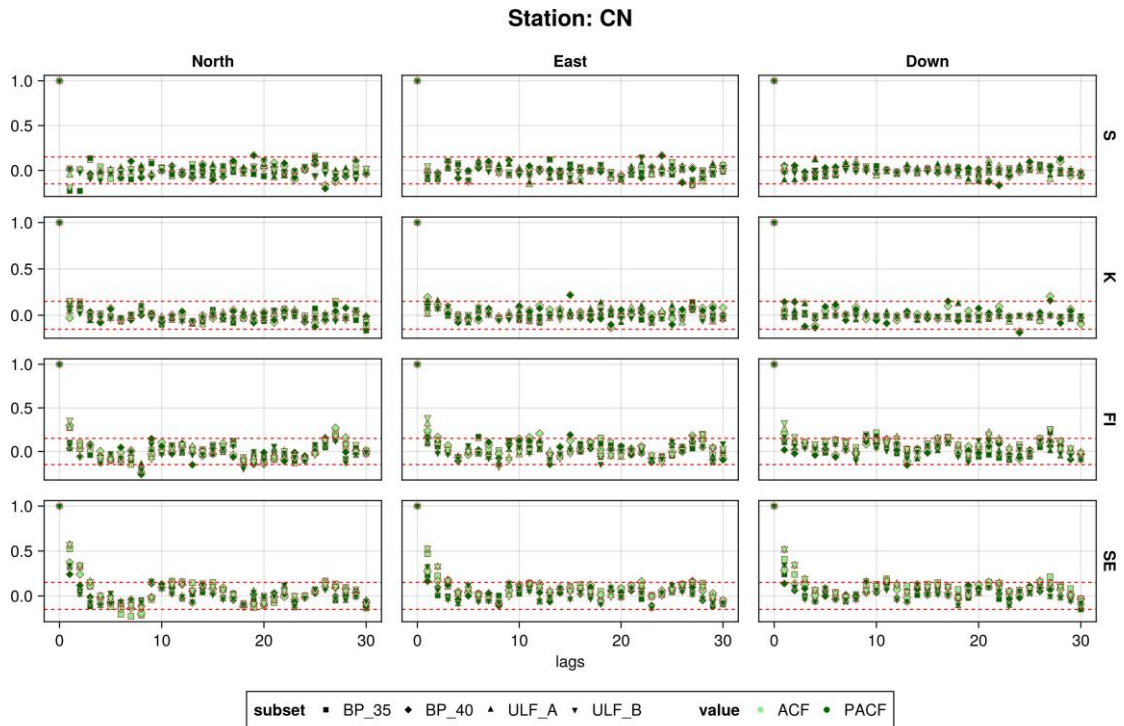
圖62: 萬巒 (WANL) 地電站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。

Station: YULI



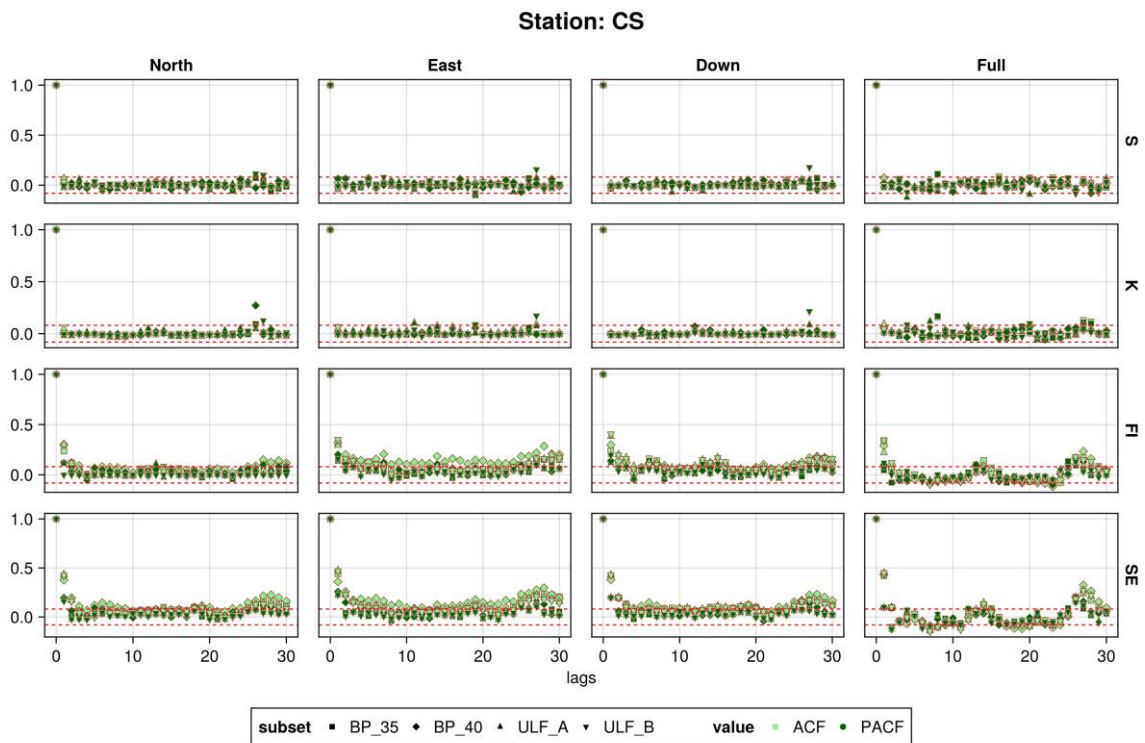
¹ACF/PACF is calculated on the continuous time-series from 2021-12-02 to 2024-05-13

圖63: 玉里 (YULI) 地電站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



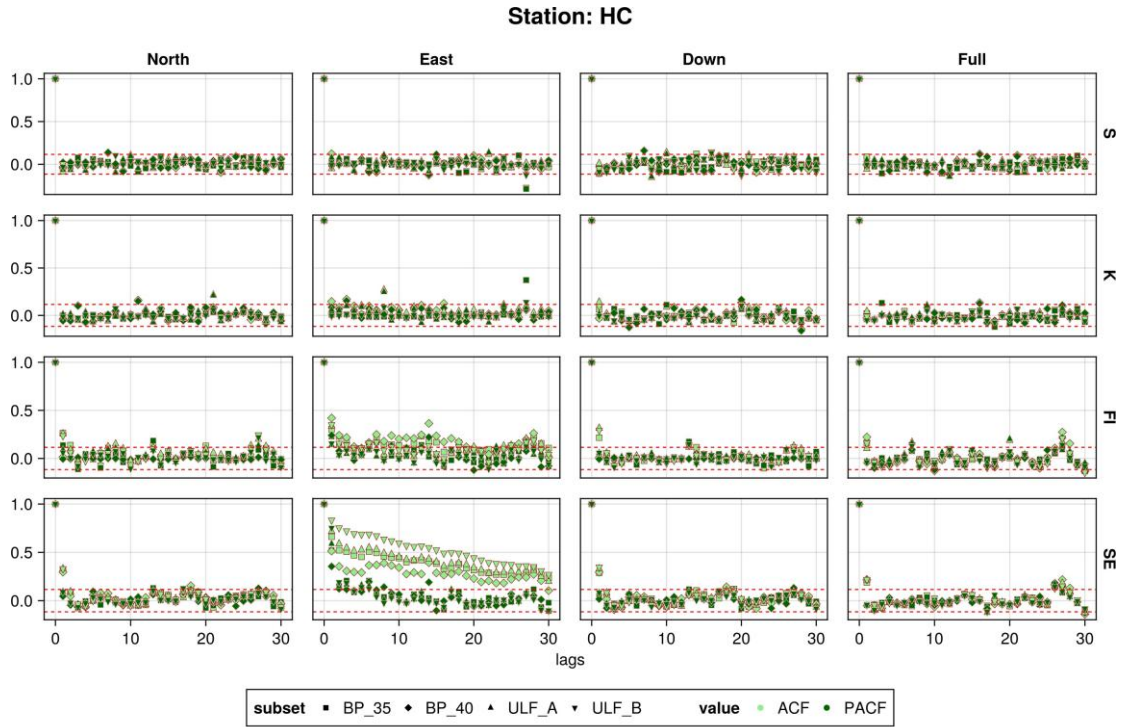
¹ACF/PACF is calculated on the continuous time-series from 2023-11-26 to 2024-05-13

圖64: 暨南(CN) 地磁站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



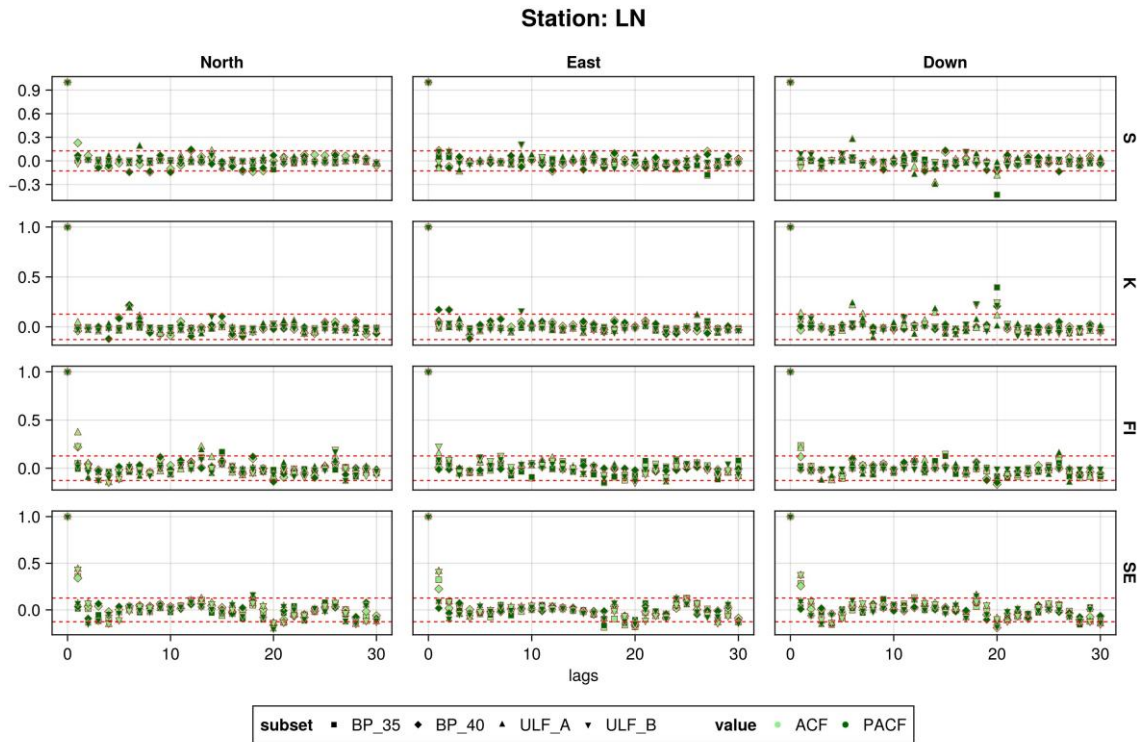
¹ACF/PACF is calculated on the continuous time-series from 2020-03-19 to 2023-09-03 and 2015-11-30 to 2017-07-08

圖65: 池上(CS) 地磁站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



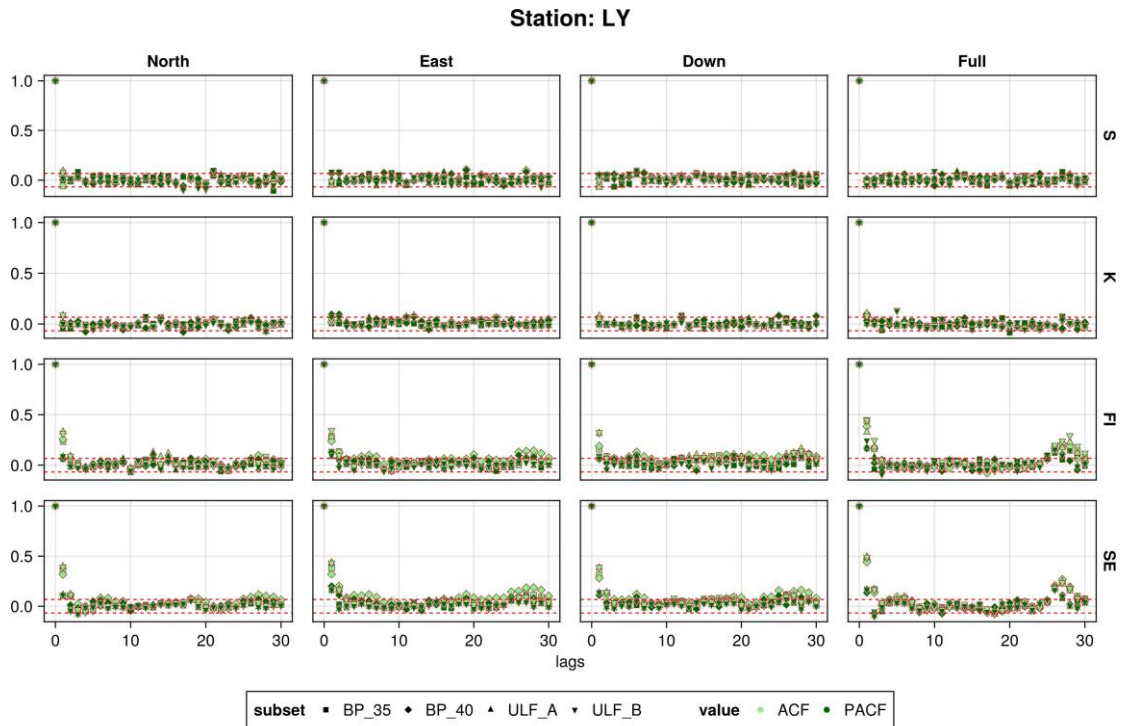
¹ACF/PACF is calculated on the continuous time-series from 2020-12-28 to 2021-11-30 and 2017-11-07 to 2018-08-20

圖66: 恆春(HC) 地磁站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



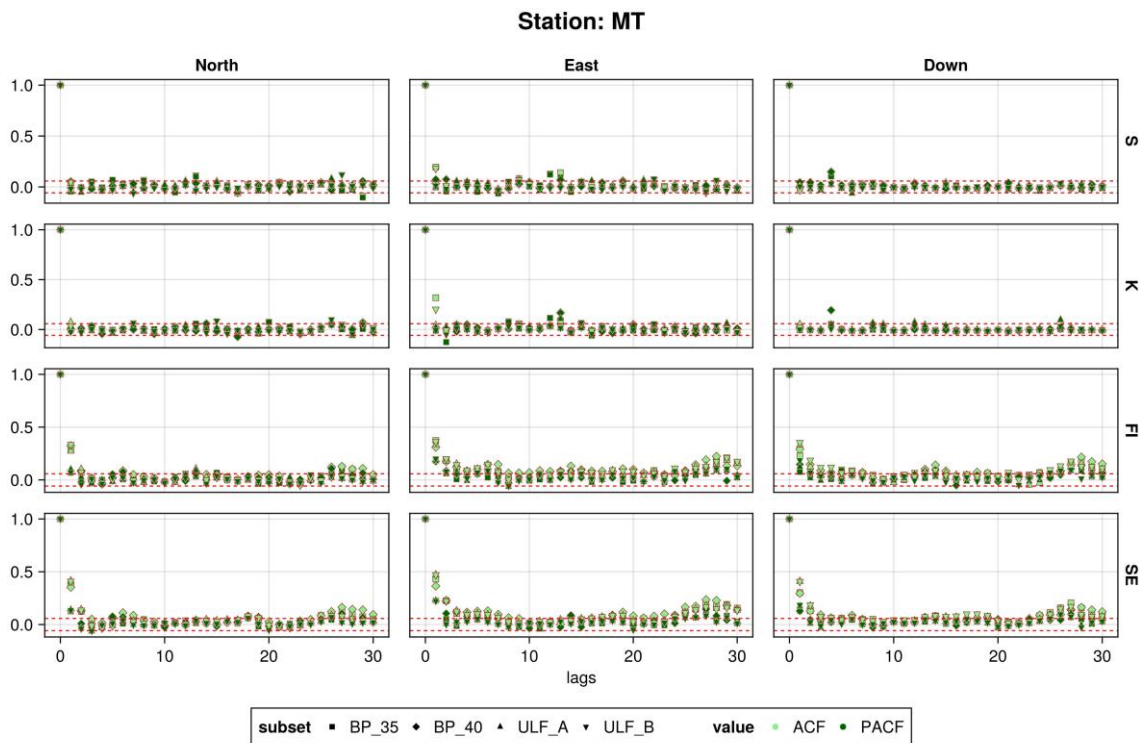
¹ACF/PACF is calculated on the continuous time-series from 2023-03-01 to 2023-10-23

圖67: 蘭嶼(LN) 地磁站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



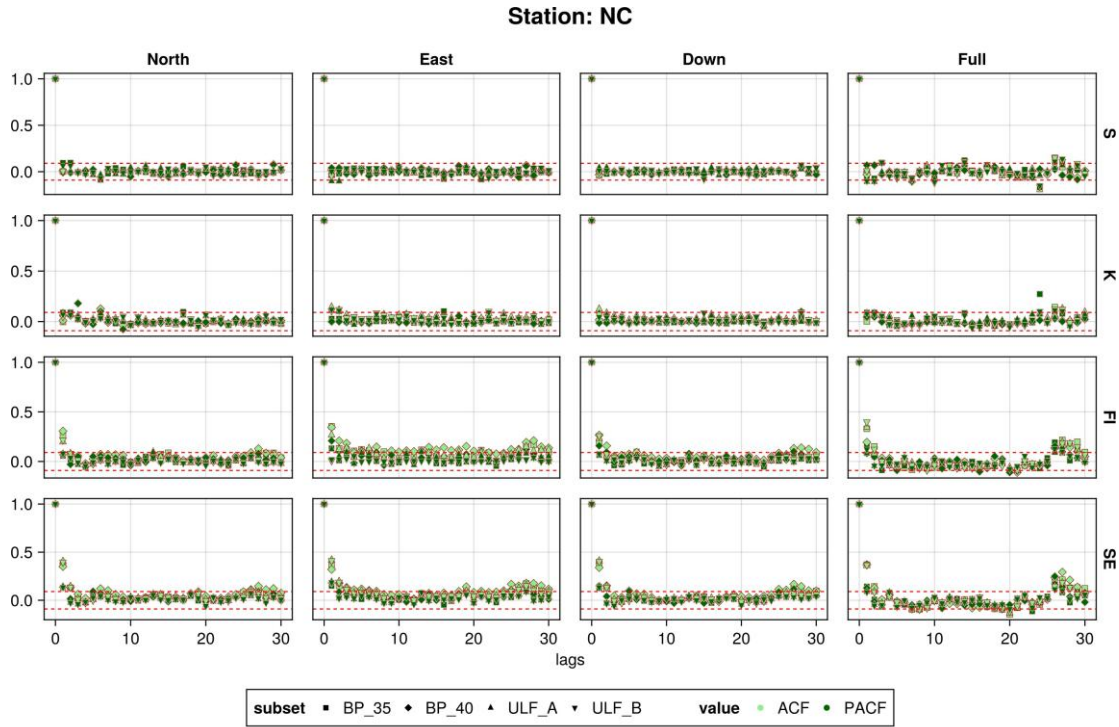
¹ACF/PACF is calculated on the continuous time-series from 2021-05-04 to 2023-10-23 and 2014-06-09 to 2016-09-27

圖68: 鯉魚潭 (LY) 地磁站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



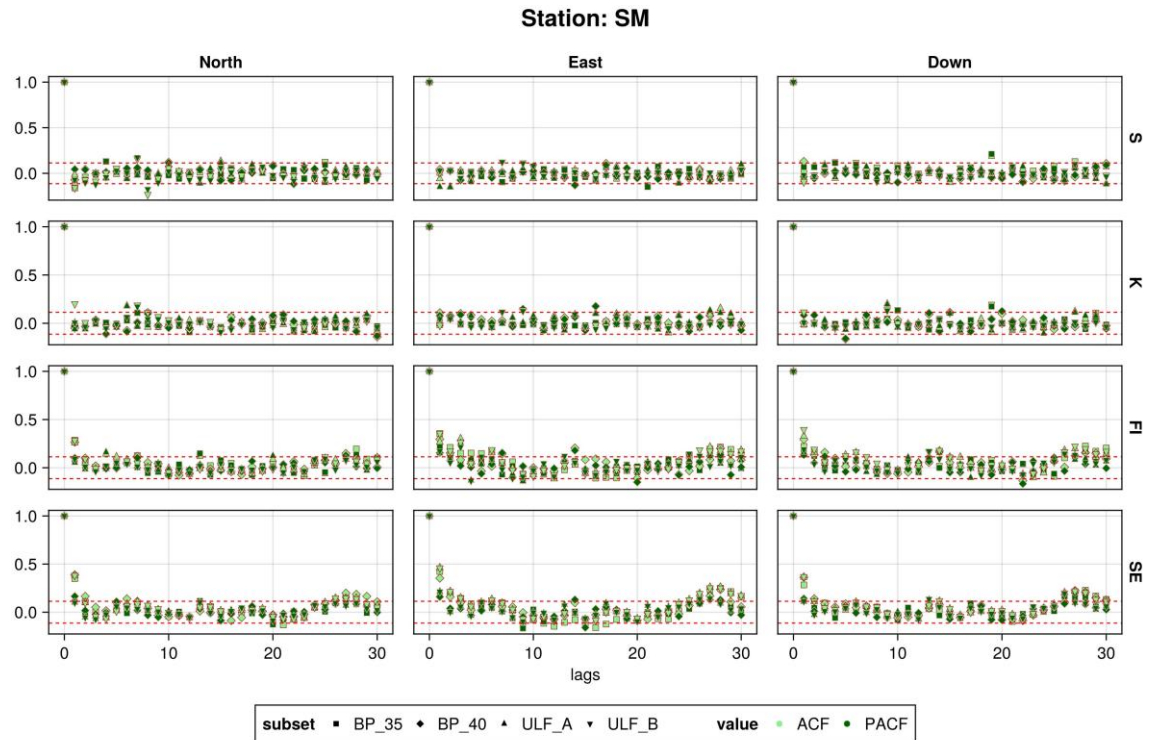
¹ACF/PACF is calculated on the continuous time-series from 2020-09-10 to 2023-10-23

圖69: 馬祖 (MT) 地磁站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



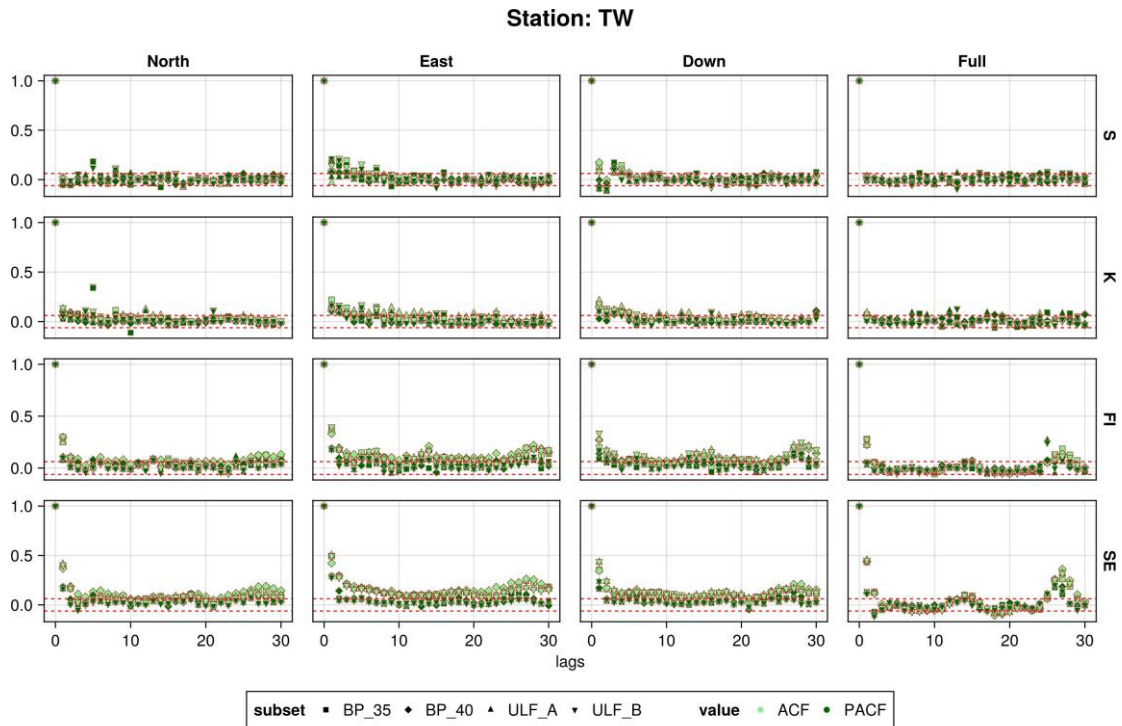
¹ACF/PACF is calculated on the continuous time-series from 2020-10-08 to 2023-10-23 and 2015-08-09 to 2016-11-19

圖70: 內城(NC) 地磁站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



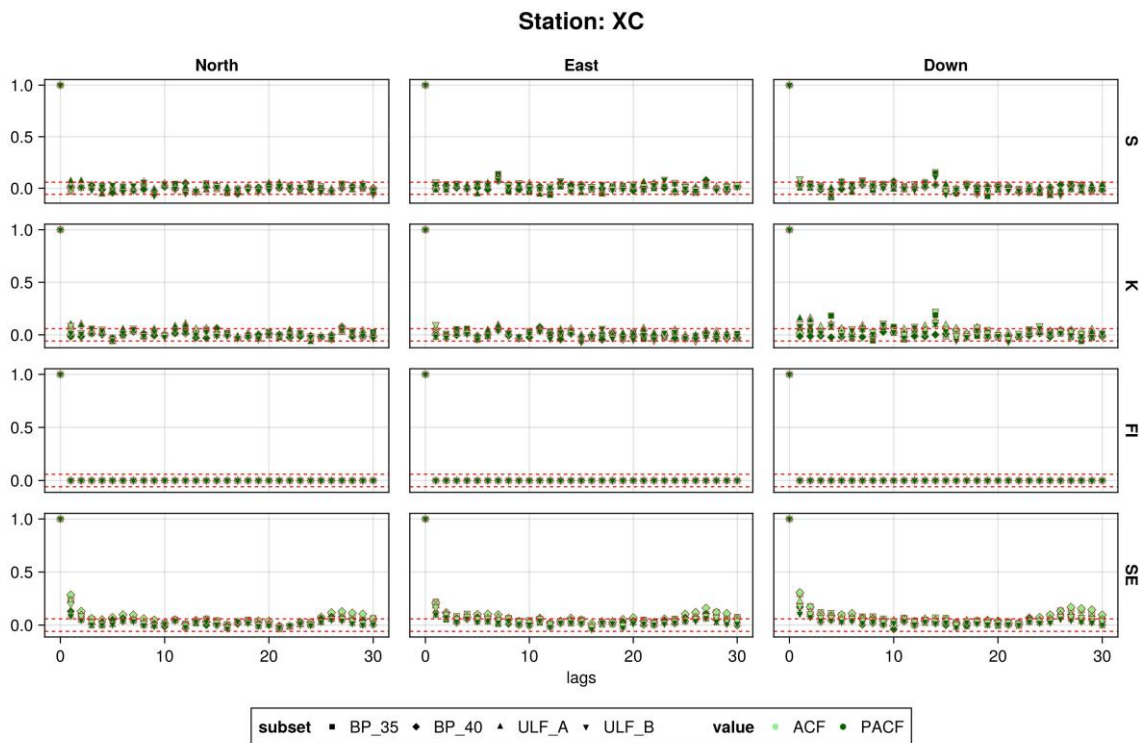
¹ACF/PACF is calculated on the continuous time-series from 2020-09-18 to 2021-07-13

圖71: 日月潭 (SM) 地磁站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



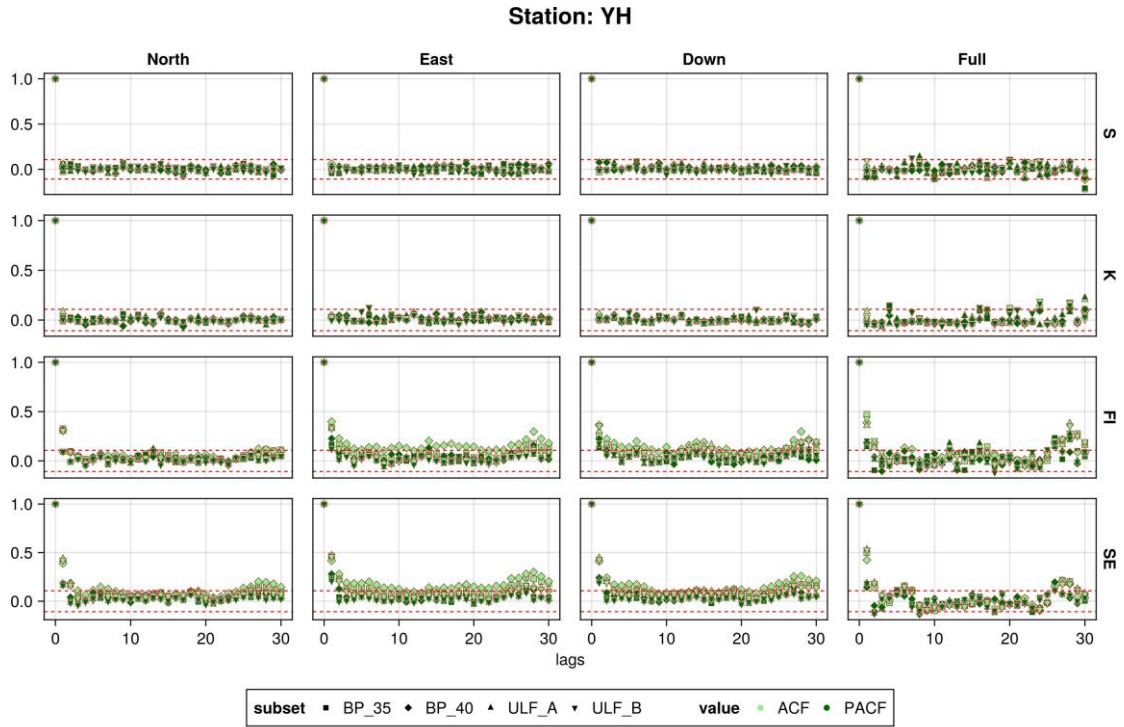
¹ACF/PACF is calculated on the continuous time-series from 2020-03-27 to 2023-10-23 and 2015-09-30 to 2018-06-26

圖72: 灣丘(TW) 地磁站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



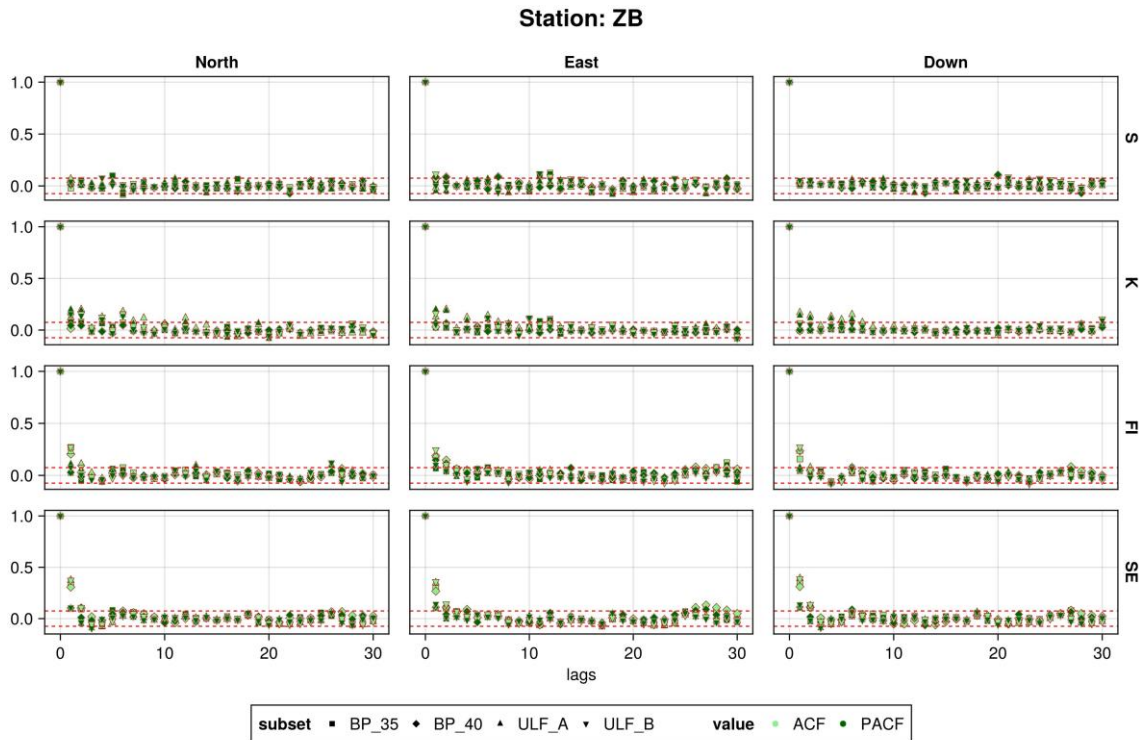
¹ACF/PACF is calculated on the continuous time-series from 2020-10-07 to 2023-10-23

圖73: 新城(XC) 地磁站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



¹ACF/PACF is calculated on the continuous time-series from 2020-04-10 to 2023-10-23 and 2014-09-10 to 2015-08-07

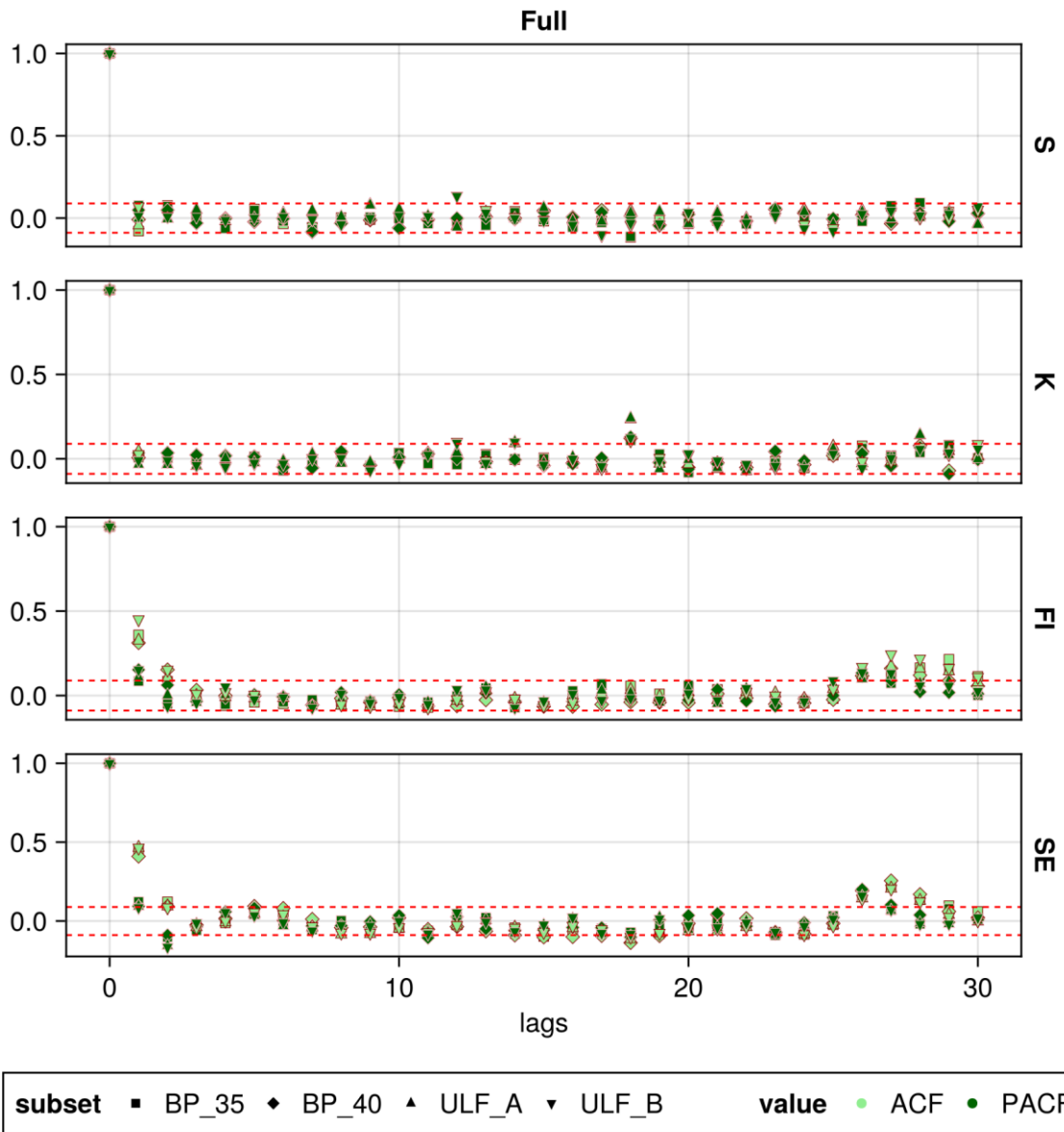
圖74: 爺亨(YH) 地磁站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。



¹ACF/PACF is calculated on the continuous time-series from 2021-12-01 to 2023-10-23

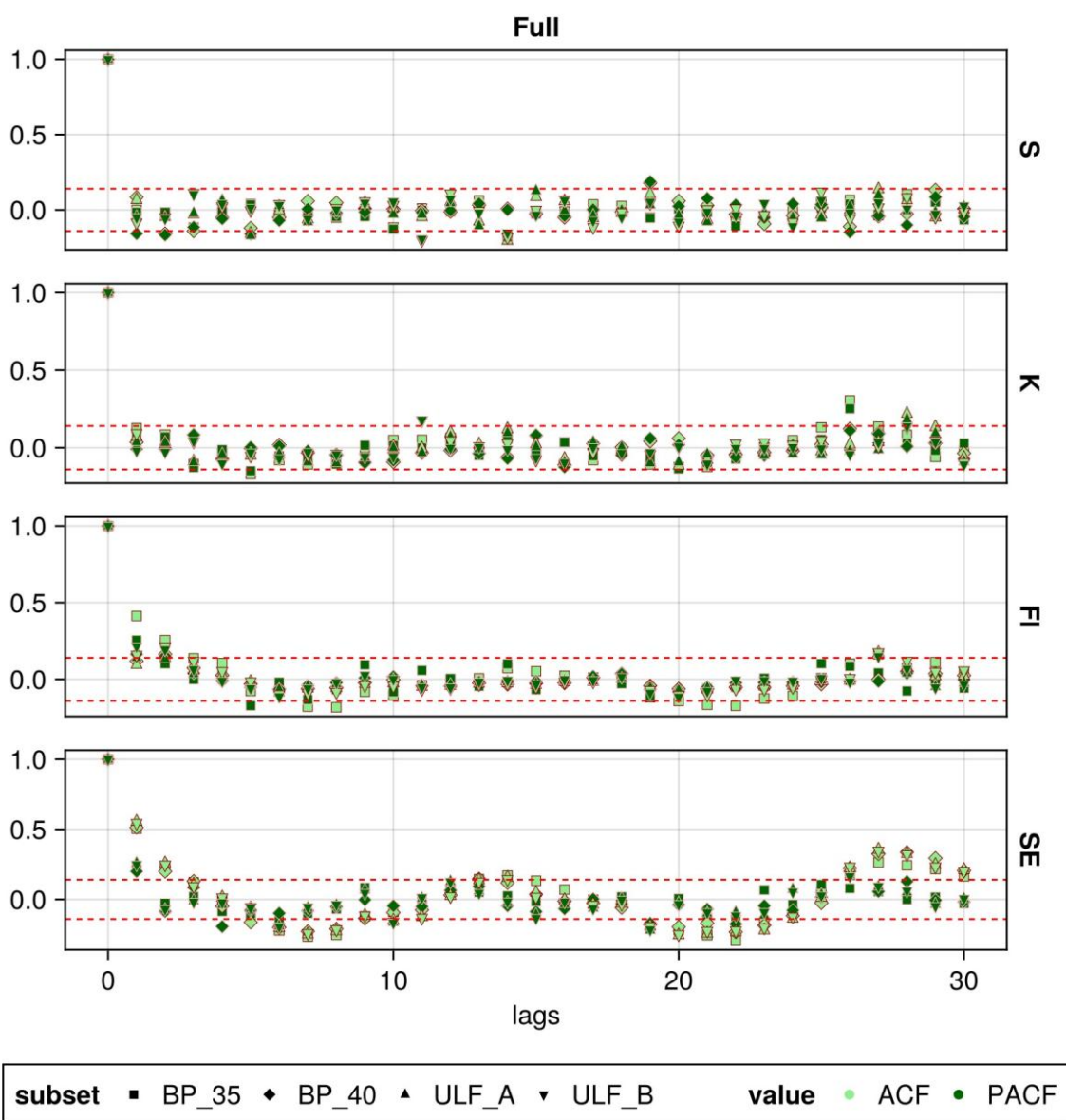
圖75: 知本(ZB) 地磁站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。

Station: HL



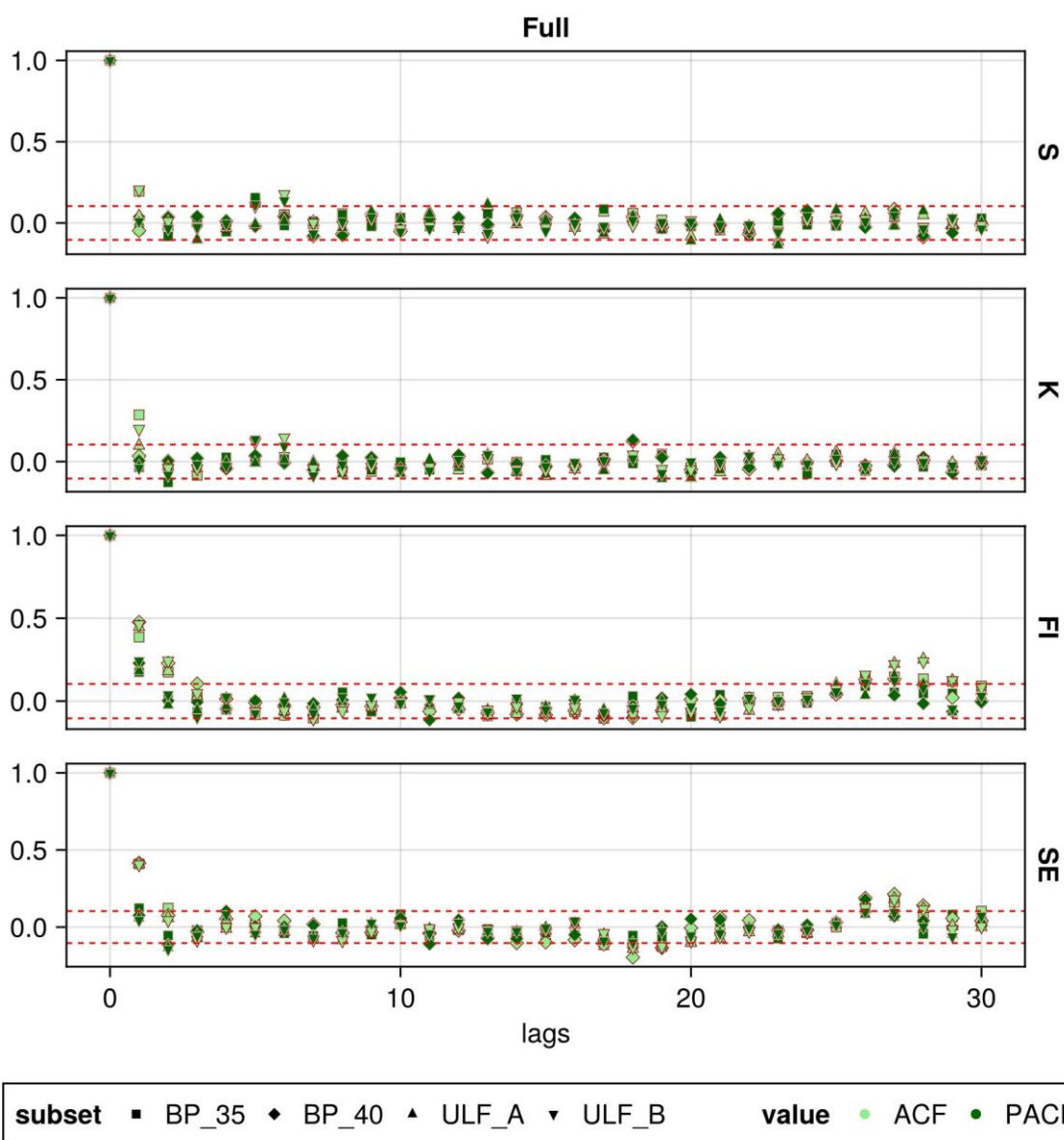
¹ACF/PACF is calculated on the continuous time-series from 2015-04-08 to 2016-08-05
 圖76: 加灣(HL) 地磁站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。

Station: MS



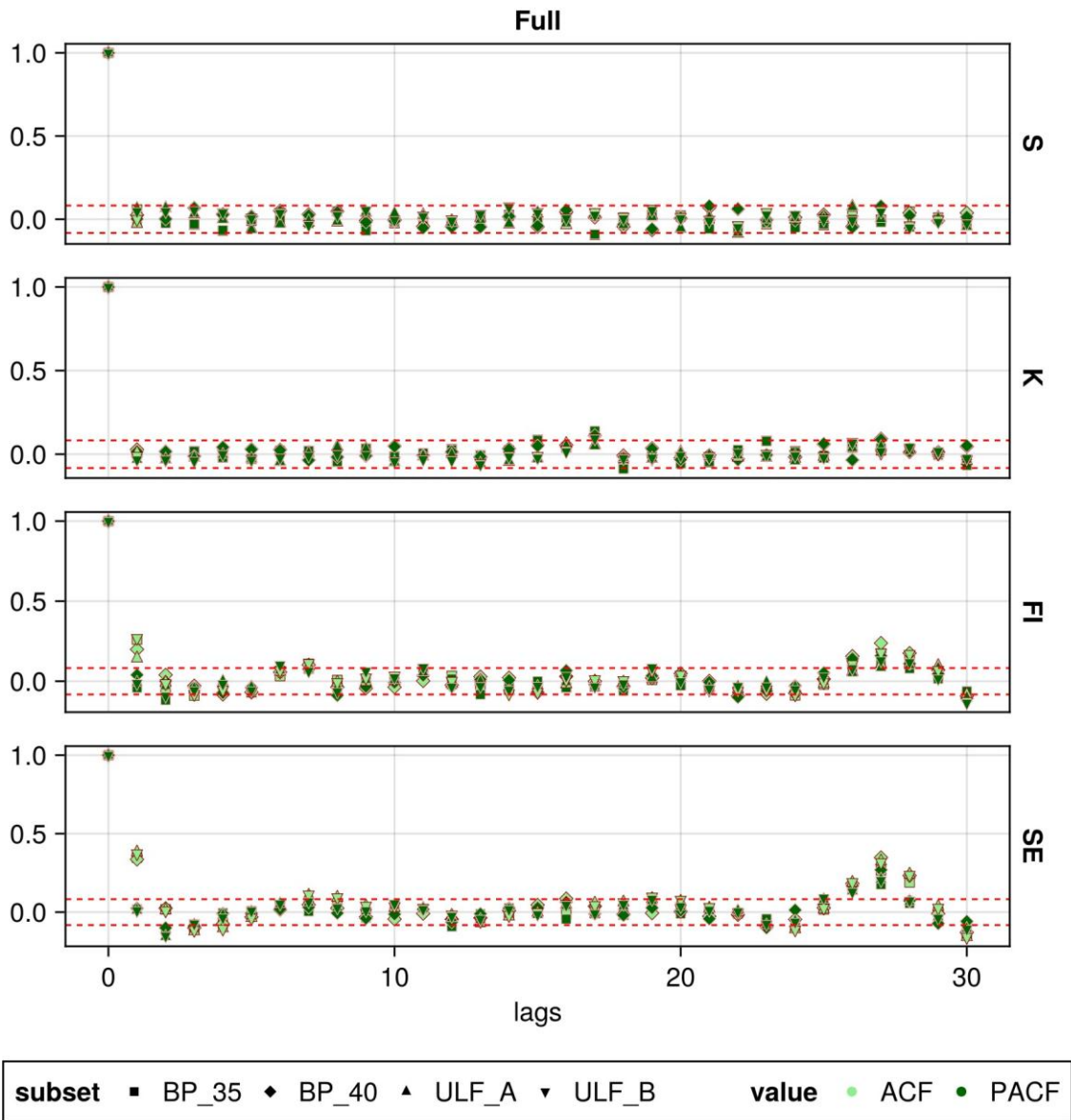
¹ACF/PACF is calculated on the continuous time-series from 2016-10-04 to 2017-04-16
 圖77: 馬仕(MS) 地磁站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。

Station: SL



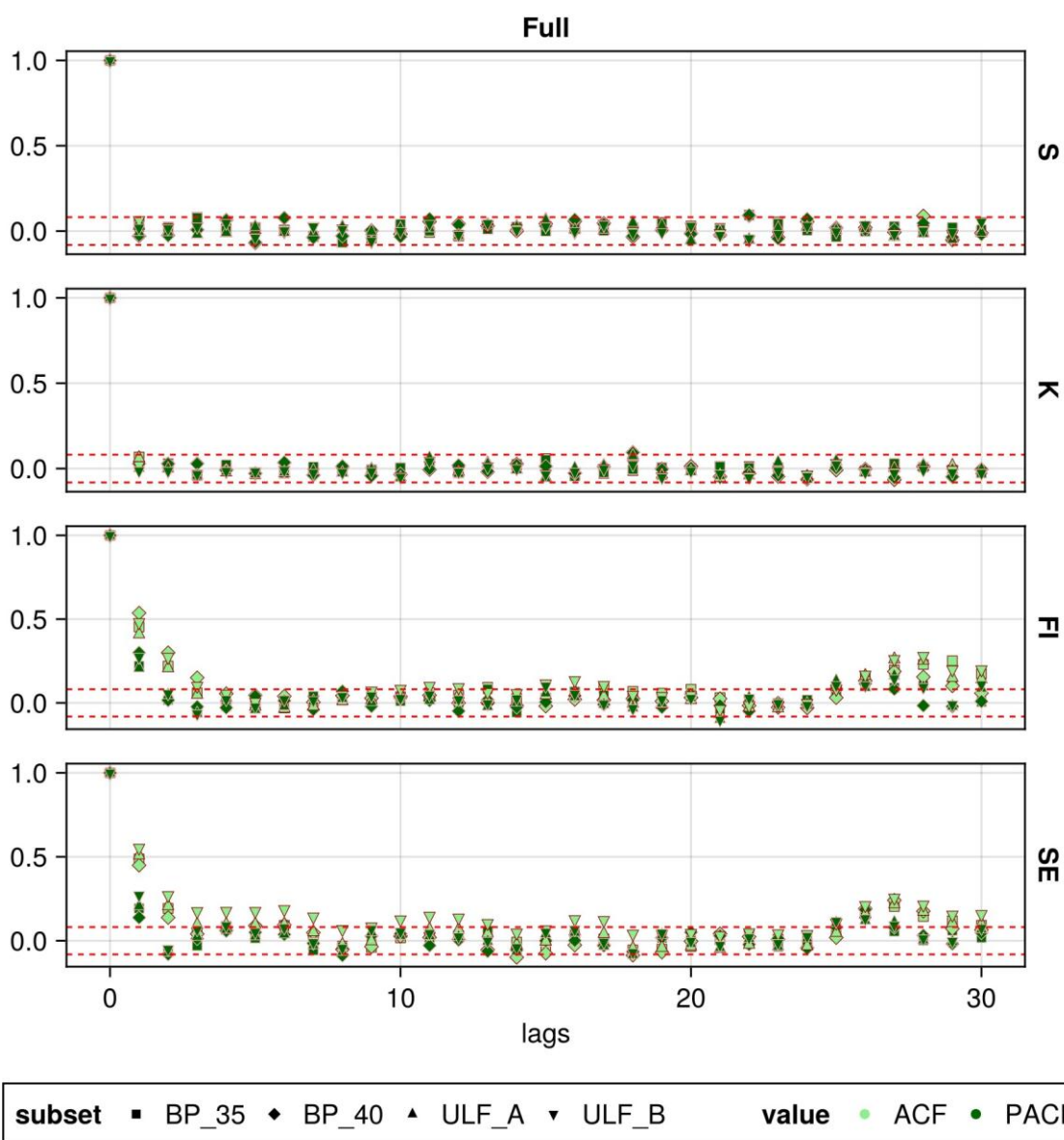
¹ACF/PACF is calculated on the continuous time-series from 2015-07-20 to 2016-07-10
 圖78: 雙龍 (SL) 地磁站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。

Station: TT



¹ACF/PACF is calculated on the continuous time-series from 2017-10-16 to 2019-05-03
 圖79: 卑南 (TT) 地磁站 各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。

Station: YL



¹ACF/PACF is calculated on the continuous time-series from 2014-08-14 to 2016-03-16
 圖80: 玉里(YL) 地磁站各濾波頻段資料的統計指標與資訊理論指標之樣本自相關函數 (SACF)、樣本偏自相關函數 (SPACF)圖。

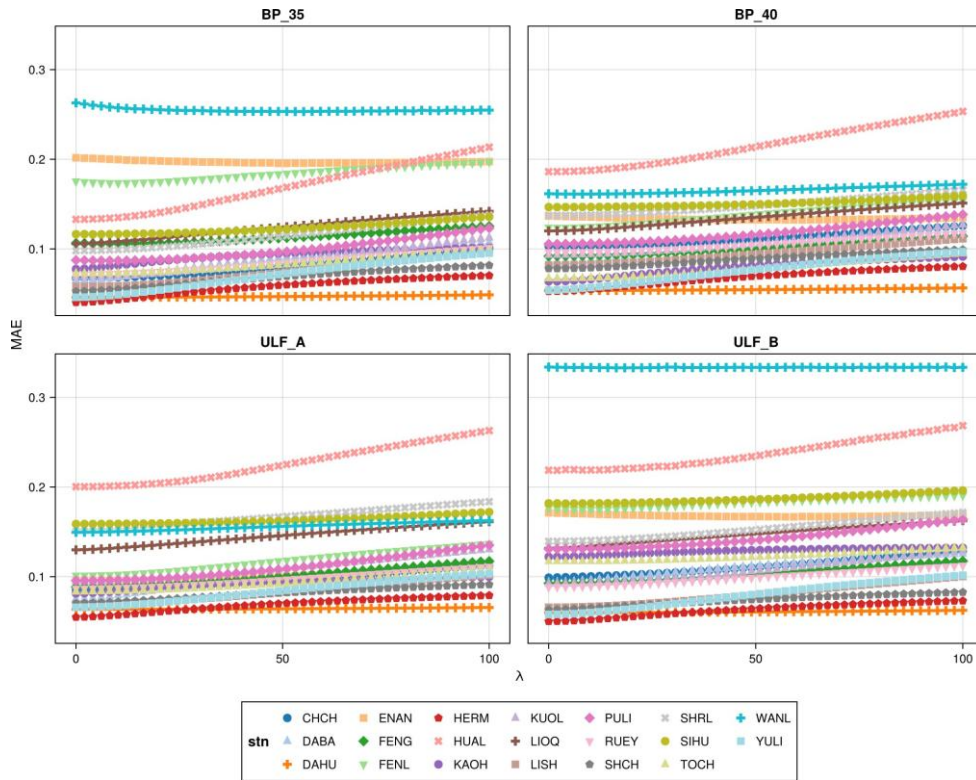


圖81: 地電場測站 Ridge 迴歸模型在訓練期之交叉驗證平均絕對誤差 (MAE) 隨正則化參數 λ 變化。

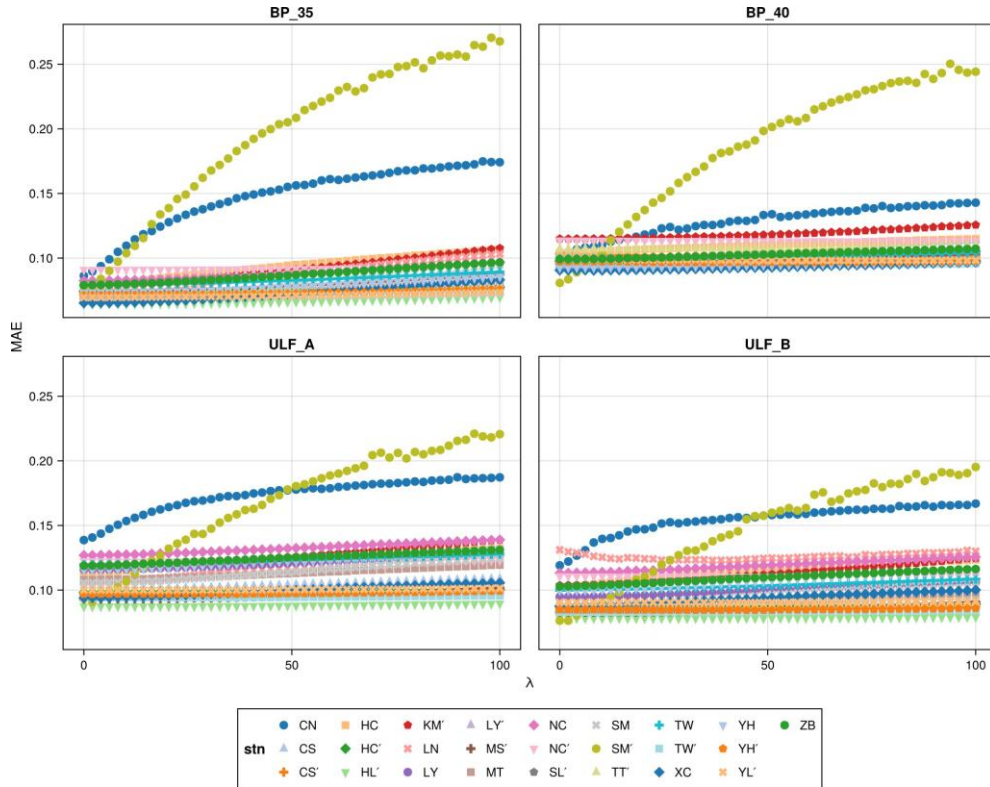


圖82: 地磁場測站 Ridge 迴歸模型在訓練期之交叉驗證平均絕對誤差 (MAE) 隨正則化參數 λ 變化。

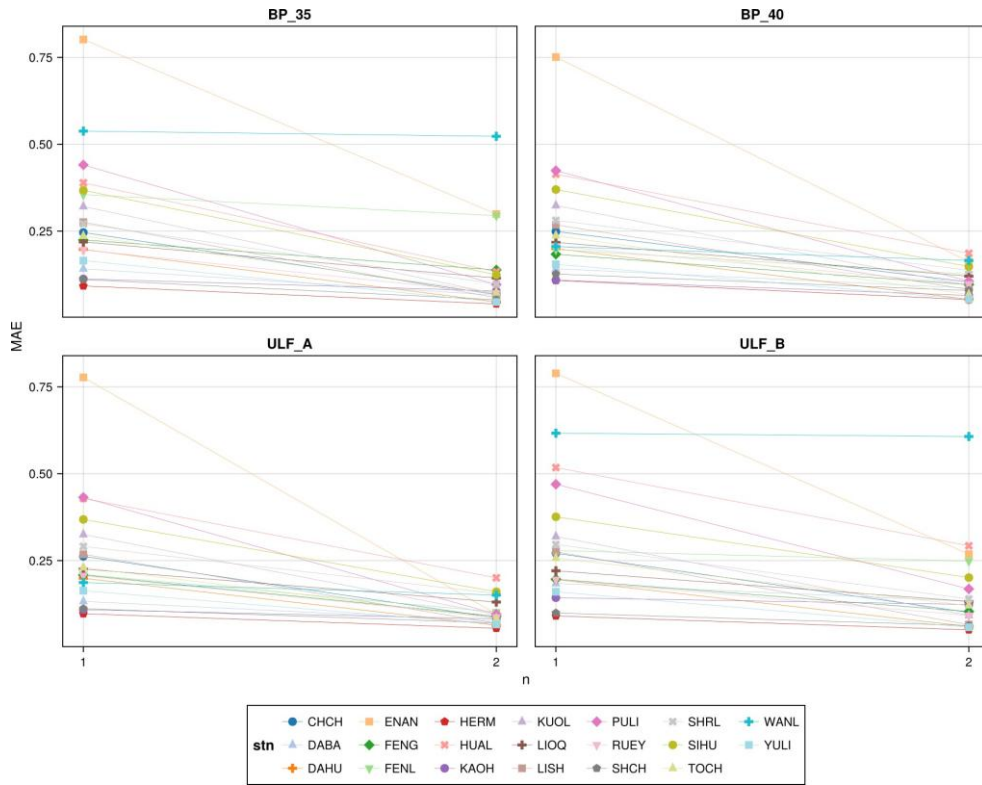


圖83: 地電場測站 PLS 迴歸模型在訓練期之交叉驗證平均絕對誤差 (MAE) 隨潛在成分數 n 變化。

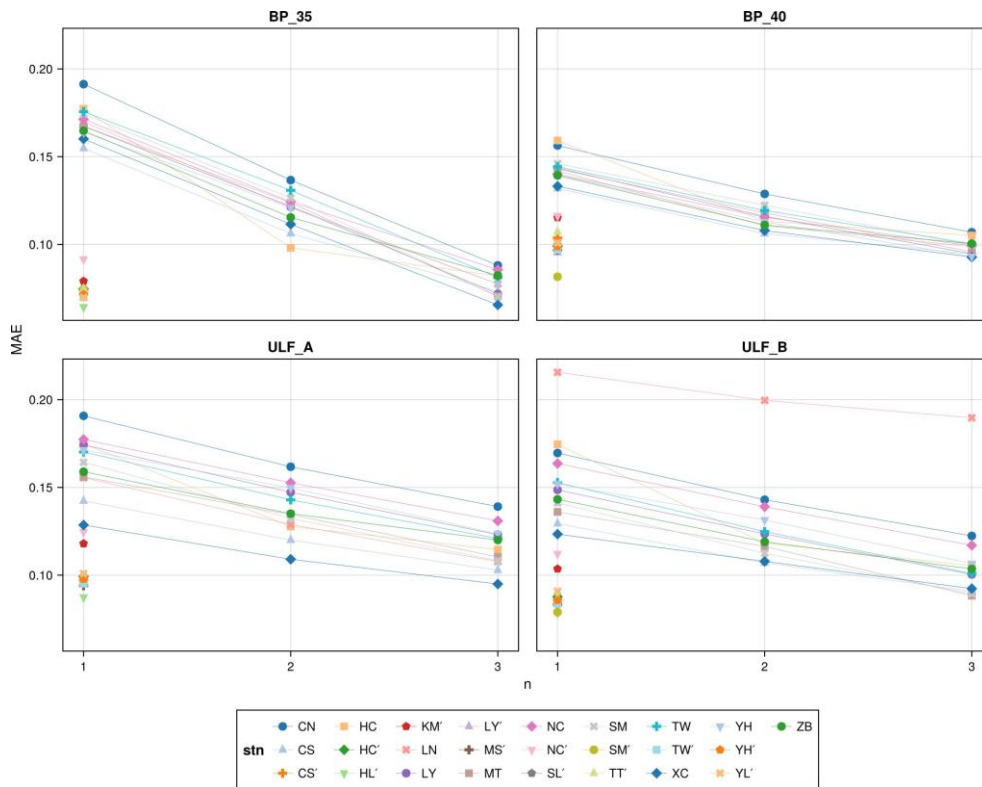


圖84: 地磁場測站 PLS 迴歸模型在訓練期之交叉驗證平均絕對誤差 (MAE) 隨潛在成分數 n 變化。其中，全磁場(單分量)測站分別顯示，以 ' 附註於測站代碼。

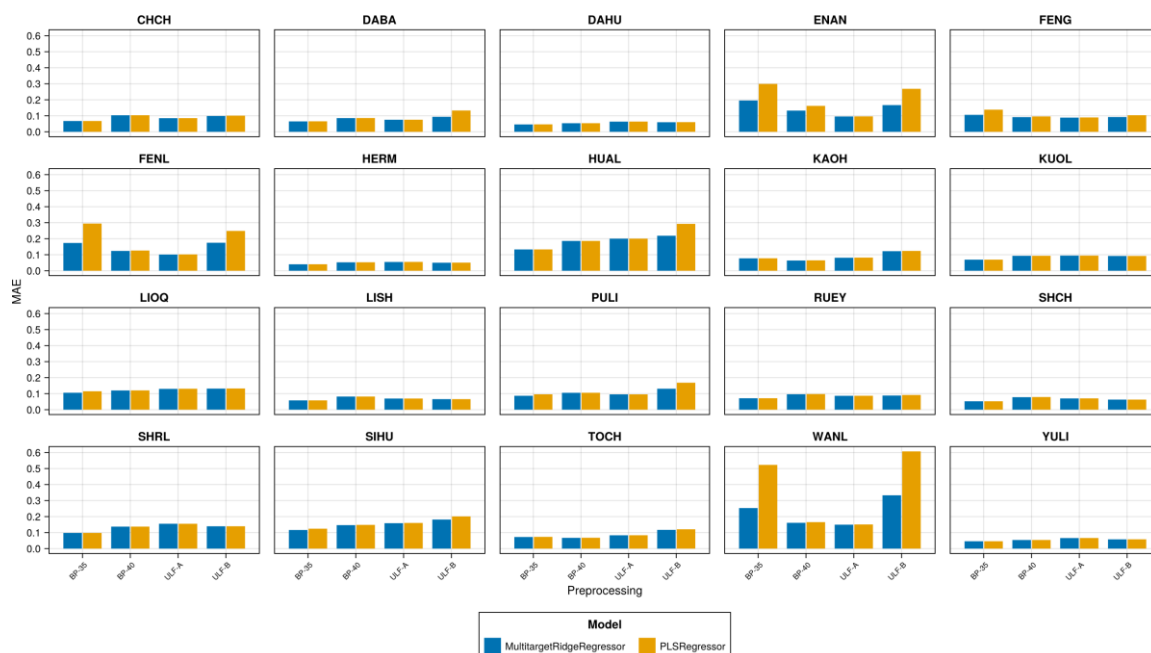


圖85: 各地電場測站在不同濾波頻段下，Ridge 與 PLS 迴歸模型之最佳平均絕對誤差 (MAE) 比較。

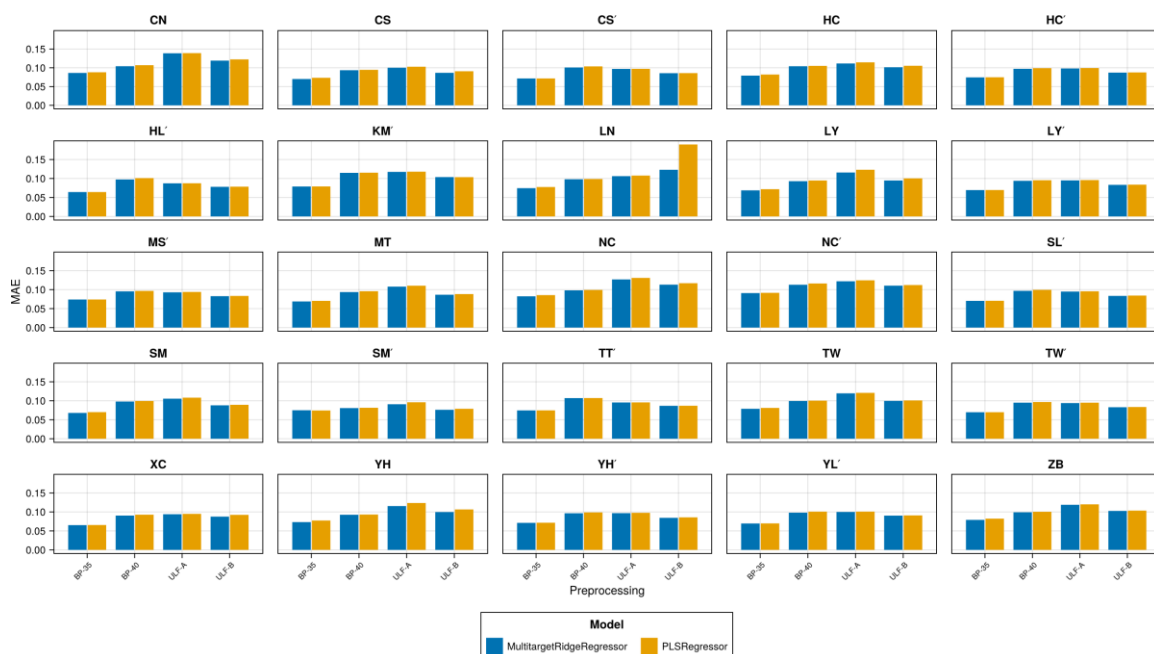


圖86: 各地磁場測站在不同濾波頻段下，Ridge 與 PLS 迴歸模型之最佳平均絕對誤差 (MAE) 比較。其中，全磁場(單分量)測站分別顯示，以 ' 附註於測站代碼。

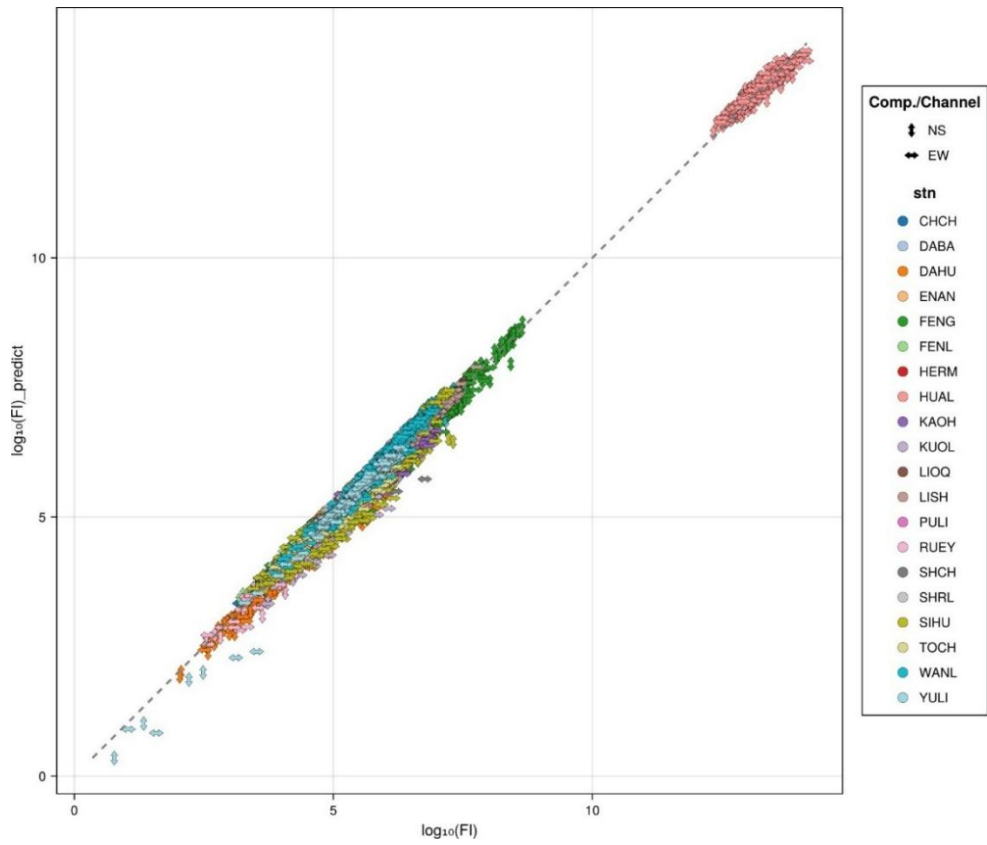


圖87: 地電場測站 Ridge 迴歸模型在測試期之實際值與預測值散佈圖。

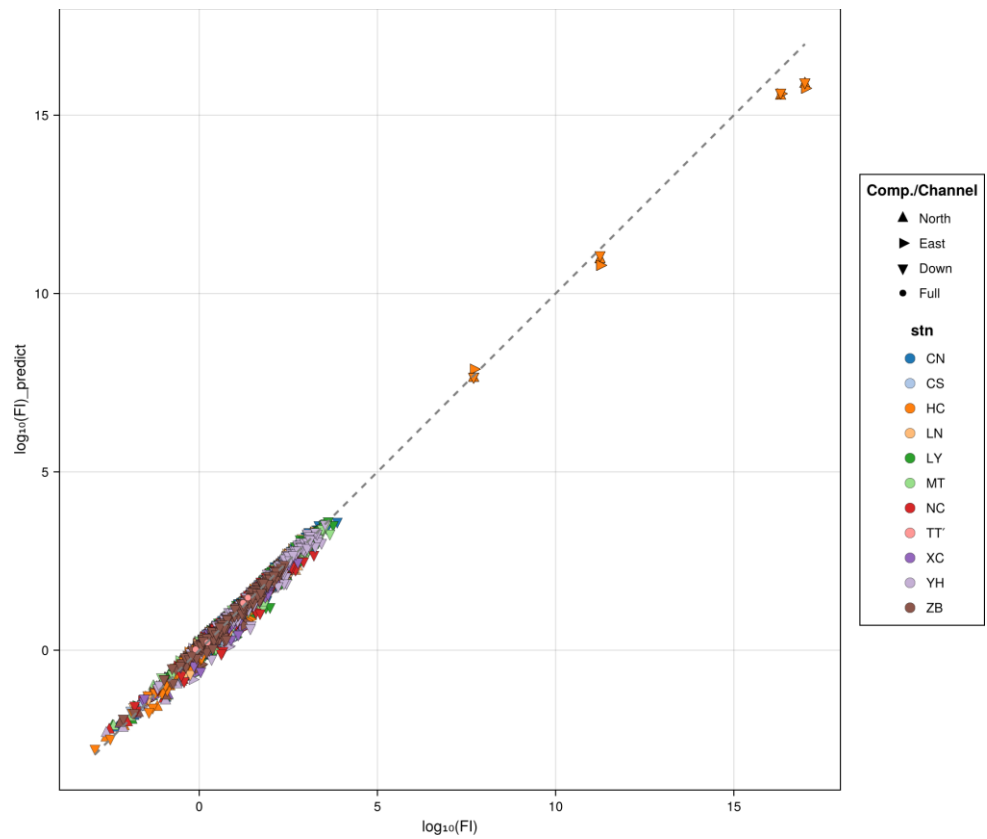


圖88: 地磁場測站 Ridge 迴歸模型在測試期之實際值與預測值散佈圖。

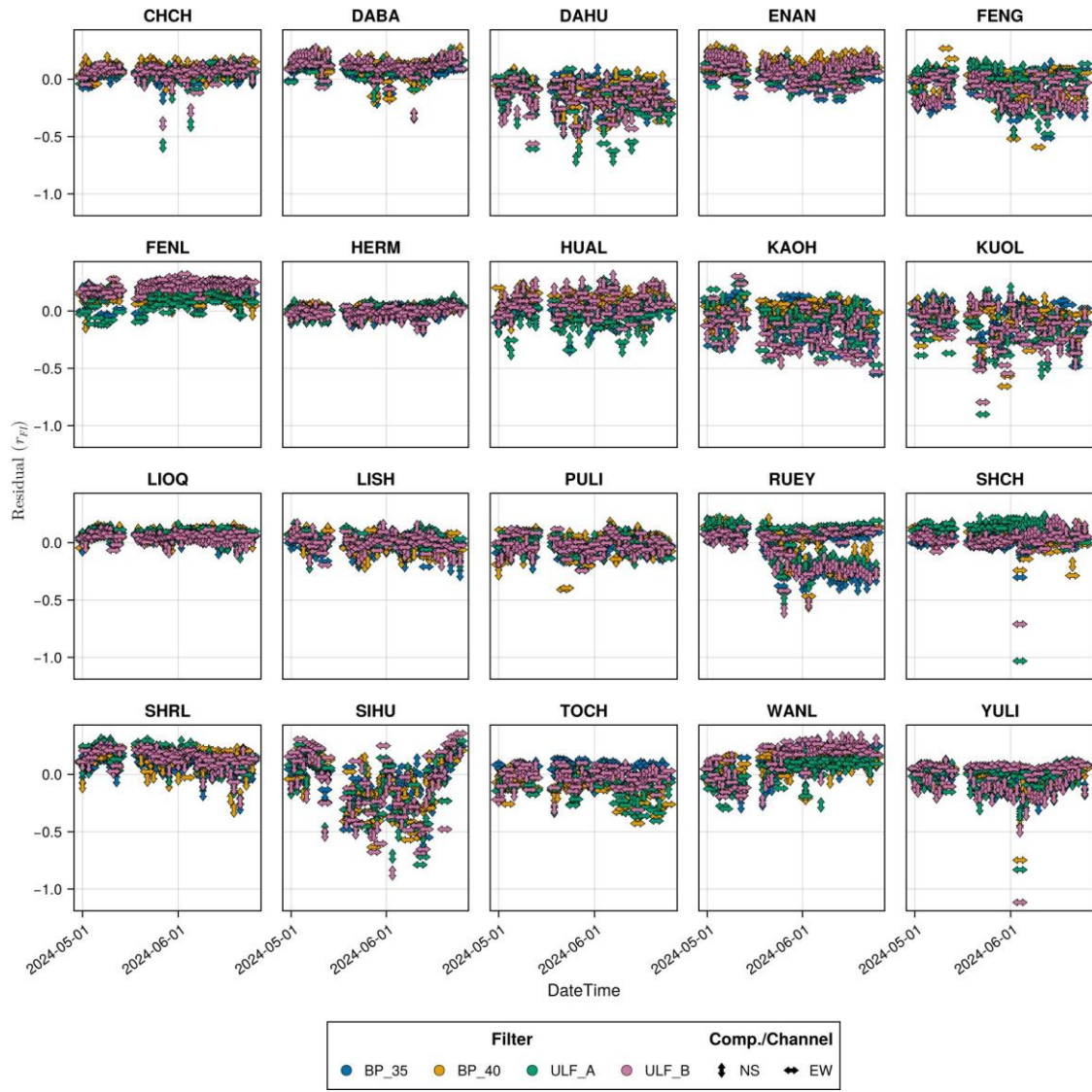


圖89: 地電場測站 Ridge 迴歸模型在測試期之預測殘差時序圖。

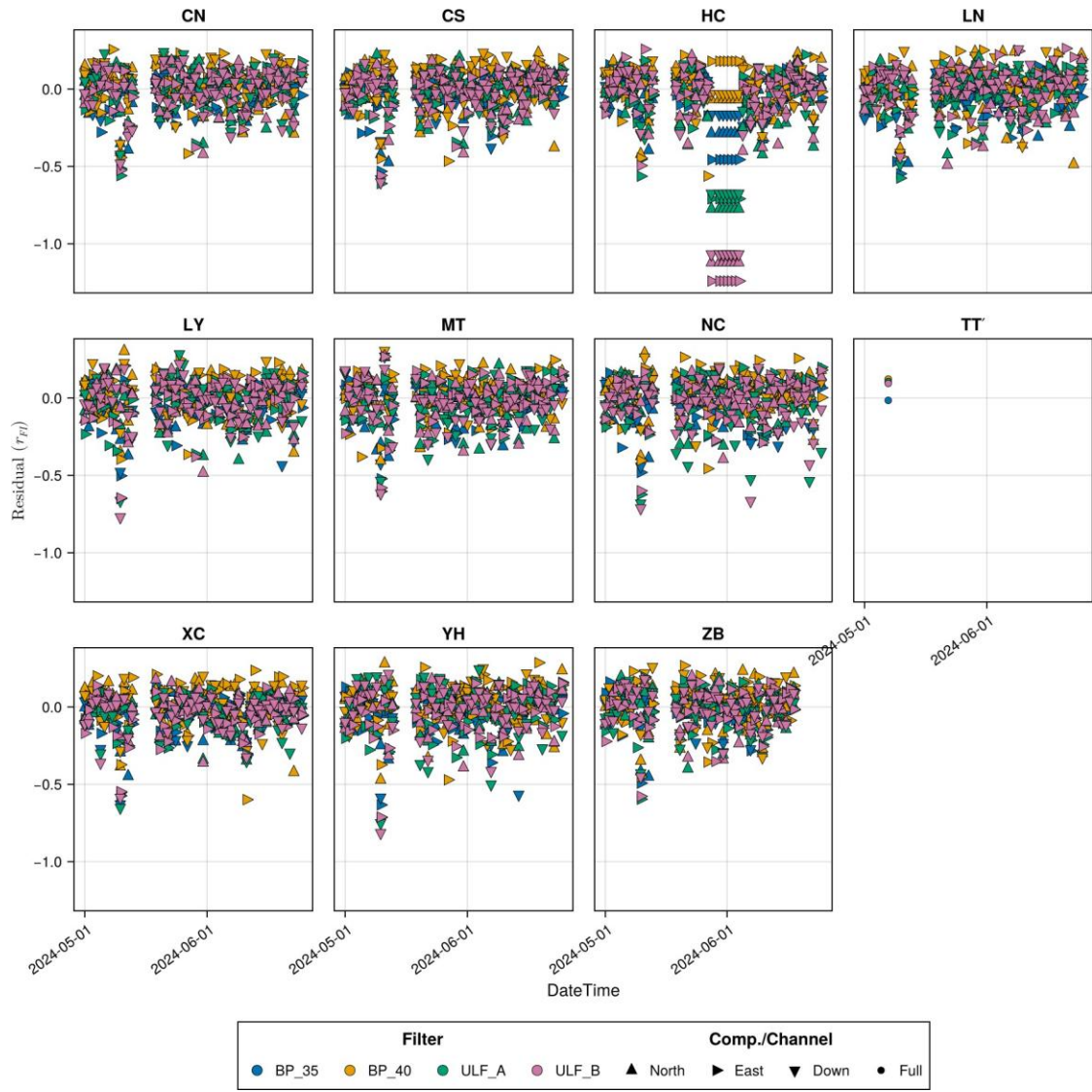


圖90: 地磁場測站 Ridge 迴歸模型在測試期之預測殘差時序圖。

114 年地震前兆觀測作業與分析技術相關研究

子計畫二

地震前兆監測—三分量地磁資料分析

顏宏元 羅祐宗 郭炫佑

中華民國地球物理學會

摘要

九二一大地震後，經檢視臺灣地區磁力連續觀測資料（全磁場強度）發現，位於車籠埔斷層兩端的磁力站—鯉魚潭站及曾文站，都觀測到磁力異常訊號，這些異常訊號與地震的發生有時間上的巧合。一般認為，當岩層受到應力作用而產生裂隙，地層內所含的帶磁礦物會散發在空間中，造成地球磁場的變化。計畫主要為檢視臺灣地區地磁連續觀測網之三分量地磁資料品質，同時也分析地磁連續觀測網資料，以了解地磁場強度改變與地震活動的關聯性。

關鍵詞：地磁、地震活動

Abstract

After 921Chi-Chi earthquake, numerous magnetic anomaly signals were observed in the northern and southern parts of the Chelungpu fault, and it is coincided with the time of Chi-Chi earthquake occurred. It is generally believed that the magnetic minerals contained in the formation will be dispersed in space, causing changes in the geomagnetic field. This project includes not only checking the quality of the magnetic observation network data, but also analyzing 3-components geomagnetic observations associated with seismicity and changed in tectonic stress.

Keywords: Geomagnetic, seismicity

壹、前言及研究目的

在 1999 年九二一集集大地震後，在距車籠埔斷層北端 8 公里的鯉魚潭磁力站，地球磁場總強度在地震發生前一個多月就記錄到擾動的訊號，振幅達到 180 nT，這些擾動訊號的振幅逐漸變小；在同年 10 月 22 日嘉義大地震發生的同時，這些擾動訊號就沒有被偵測到。這些地磁場總強度擾動的訊號應該和大地震的發

生有關 (Yen et al., 2004)；另外，位於臺南的曾文地磁觀測站，似乎也記錄到一些擾動的訊號，研判與嘉義大地震有關。

臺灣地磁連續觀測網，大致上可以分為三個時期：

(1) 1988~2001 年由中央研究院地球科學研究所建置的臺灣地區地磁連續觀測網，站內安裝 G-856 質子進動磁力儀 (Proton-Precession magnetometer, G856)，每 5 分鐘觀測地磁全磁場強度資料存於儀器的硬碟中，每兩個月由技術人員前往各站下載資料。

(2) 2002~2006 年中央大學改造中研院地球科學研究所地磁連續觀測網，共有 12 個觀測站，使用的儀器為加拿大 GEM 公司所生產之 GSM90F 型磁力儀，取樣率為 1Hz，記錄資料傳送至中央大學的記錄中心。

(3) 2007 年後，中央大學將所建置的「臺灣地區地磁場連續觀測網」移轉給中央氣象局。2014 年起為了加強對池上斷層的監測，於池上近斷層增設了池上(CS)站。

由於儀器老舊，地磁連續觀測網陸續有些測站出現了資料中斷的現象。自 2018 年開始進行地磁觀測儀器更新，將原量測全磁場的儀器汰換為三分量磁力儀，此儀器由日本 TIERRA TECNICA 公司所製造之儀器，取樣率為 1Hz，包含接收器、前置放大器、主機和 GPS，用網路將所觀測的資料即時傳輸到記錄中心。

本計畫主要為檢視地磁連續觀測網之地磁觀測網資料品質，同時也分析全臺三分量地磁連續觀測網資料，以了解地磁場強度改變與地震活動的關聯性。

貳、研究方法-磁力觀測作業

臺灣因受地震活動、電離層潮汐與洋流等多重因素影響，呈現豐富的地磁微脈動現象，加上密集的地震監測網，是全球研究地磁微擾與地震活動關聯性的重要地區。中央氣象署自 2018 年起全面更新觀測設備，以三分量磁力儀取代原先的全磁場量測儀器，重新建置全臺三分量地磁連續觀測網。本研究即在這套高品質觀測資料基礎上，深化地磁日變化分析、磁場異常偵測，以及其與地震過程之潛在相關性。

一、資料評估與日變化振幅分析

研究的第一步是全面檢視各地磁測站的資料品質，包括長期穩定度、噪訊來源、設備狀況與環境干擾，藉此確立能作為參考基準的最佳測站（本研究以馬祖站為主）。同時，針對可能增設的新測站，亦透過周遭磁環境特徵與背景訊號分析，確保新站點具備足夠的穩定性與研究價值。

地磁場日變化分析則採用 Liu et al. (2006) 的方法，計算每日磁力值的最大與最小值之差，以獲得日變化振幅。由於日變化主要受太陽風與地磁層交互作用控制，在臺灣這樣的區域尺度中，各測站理應呈現高度一致的日變化幅度。因此，將各站與馬祖站之振幅比值進行比較，可用來檢測區域性異常訊號。過去在 921

集集地震的研究案例中，地震前後 15 天資料顯示 LY 與 TW 兩站的比值明顯偏離背景變化，推測與孕震區內導電度的改變或震前電流產生有關。本研究延續此方法，以期在現代高解析度觀測網下檢測相似訊號。

二、多測站聯合分析與主成分分解 (PCA)

為提升異常特徵的辨識度，同時降低資料的高維度特性，本研究進一步採用主成分分析 (PCA) 整合多測站三分量資料。PCA 能有效分離跨測站的共模磁場變化，如太陽活動與磁暴效應，使殘留的主成分更能反映局部地殼導電性變化或震前可能的磁場異常。透過降維後的主成分空間分布，即可檢測不同區域之間是否存在具有一致時序的異常訊號，並比對地震發生位置與時間，強化對震前磁場反應的解讀能力。

然而 PCA 的物理解釋通常具有挑戰性，因此本研究將 PCA 的結果與原始地磁波形、日變化振幅、電離層資料以及地震目錄交叉比對，確保主成分的來源具有合理的地球物理意義。

三、資料處理挑戰與因應策略

由於地磁資料具有連續、取樣密度高、資料量龐大的特性，資料處理往往面臨運算量大與品質不一致的問題。本研究透過自動化資料品質檢測、跳變偵測、缺值處理與噪訊過濾，提升整體分析穩定度。對於資料品質較差的測站，透過重新校正設備、改善觀測環境，或以多測站比值與 PCA 降低其影響。此外，為因應資料處理運算量，研究中也採用高效率演算法與分段計算方式，確保長期資料的處理得以順利進行。對於 PCA 主成分不易解讀的問題，則透過空間分布與時間序列的比對，協助判定其與地震相關性，並以既有地磁物理模型與區域構造背景加以驗證。

參、結果與討論

(1) 檢視磁力觀測站的記錄

臺灣地區三分量地磁觀測站的取樣率是 1Hz，每一個觀測站一年總計有 3 千多萬筆記錄。以相鄰兩點的全磁場強度變化量差異超過 10nT 的個數及儀器傾斜狀態，來檢視因今年 (2025) 1 月至 10 月各測站觀測資料品質。臺灣地區三分量地磁觀測站的記錄及資料品質分析，依測站代碼英文順序如后：暨南 (CN)、池上 (CS)、恆春 (HC)、金門 (KM)、蘭嶼 (LN)、鯉魚潭 (LY)、馬祖 (MT)、內城 (NC)、曾文 (TW)、新城 (XC)、爺亨 (YH) 及知本 (ZB)。各測站分別繪製 2 張圖：先繪出地磁場三個分量的原始記錄及所計算出全磁場強度的結果；再呈現相鄰兩點的全磁場強度變化量差異超過 10nT 的個數及儀器傾斜狀態 (圖二至圖十三)。

有 8 個地磁觀測站的整體資料品質較佳，分別是位於臺灣西半部鯉魚潭 (LY)、曾文 (TW)、爺亨 (YH) 等 3 個站，位於臺灣南端的恆春 (HC)，位於臺灣東部的池上 (CS)、蘭嶼 (LN)、知本 (ZB) 等 3 個站及馬祖 (MT) 地磁觀測站。儀器傾斜部分會出現同震偏移現象，各觀測站的儀器姿態記錄都還穩定。不過在 7 月中下旬因為學校停電導致所有觀測站都有 2 天沒有資料。

暨南 (CN) 地磁站的記錄如圖二 (a) 和 (b) 所示，該站在今年 5 月底至 9 月中旬的資料有中斷的情況，不過有記錄時的資料品質相當的不錯，其儀器傾斜姿態記錄穩定。圖三 (a) 和 (b) 是池上 (CS) 的地磁觀測記錄，該站地磁場三分量記錄及傾斜記錄品質都非常的正常，不過在垂直分量 (HZ) 可以看到 10 個月偏移了 80 nT (圖三 (a))。圖四 (a) 和 (b) 是恆春 (HC) 的地磁觀測記錄，該站於 1 月中旬資料有短暫中斷的情形及 5 月 1 日地磁三個分量都有突跳的情形，突跳之後該站地磁觀測恢復良好的記錄資料品質，在垂直分量 (HZ) 也記錄到 6 個月間偏移了 80 nT (圖四 (a))。不過在儀器的姿態卻有不穩定的情況，如圖四 (b) 所示，原因待釐清。圖五 (a) 和 (b) 是金門 (KM) 的地磁觀測記錄，該站地磁場南北及東西分量的資料似乎有問題 (圖五 (a))，且 4 至 8 月底的資料是中斷的。9 月 20 日在金門東方發生規模 5.0 的地震，雖然 9 月份恢復觀測，但在儀器傾斜資料並沒有記錄到此一地震所造成的同震訊號。蘭嶼 (LN) 地磁觀測站的記錄如圖六 (a) 和 (b) 所示，該站的資料在 3 月初恢復正常觀測及資料傳輸，所記錄之地磁資料品質相當的不錯，儀器傾斜姿態正常且穩定。圖七 (a) 和 (b) 是鯉魚潭 (LY) 的地磁觀測記錄，該站的資料品質佳且沒有受到干擾。

圖八 (a) 和 (b) 是馬祖 (MT) 地磁站的觀測記錄，該站的資料品質佳而被選定為整個地磁觀測網的基準參考站。內城 (NC) 地磁觀測站的記錄如圖九 (a) 和 (b) 所示，該站的資料在 1 至 9 月所記錄之地磁資料品質相當的不錯，儀器傾斜姿態正常且穩定，不過從 10 月份開始資料就受到嚴重的干擾，儀器姿態卻沒有受到影響。曾文 (TW) 地磁觀測站的記錄品質一直都很不錯，不過在去 (2024) 年 4 月儀器發生故障送修，並在今年初重新安裝後恢復正常觀測及資料傳輸 (圖十 (a) 和 (b))。今 (2025) 年 1 月 21 日清晨在臺南市楠西區發生規模 6.4 的地震，在震央附近的曾文 (TW) 地磁觀測站受地震影響，各分量觀測到不等量的突跳情形 (圖十 (a))，儀器姿態也受到地震影響而發生偏移的現象，該站的資料在群鎮結束後恢復正常觀測及資料傳輸。架設在原大漢技術學院運動場旁的新城 (XC) 地磁觀測站，因附近常有人為活動而影響地磁觀測資料的品質 (圖十一 (a) 和 (b))。該站在 7 至 9 月資料中斷，於 9 月底重新啟動後，各分量及全磁場強度都有明顯的偏移，原因有待釐清。爺亨 (YH) 地磁觀測站一直都很不錯的資料品質 (圖十二 (a) 和 (b))，除了在 10 月 20 日可能是儀器發生傾斜而造成地磁資料的突跳事件，其他時間資料品質都非常穩定。知本 (ZB) 磁力站在 3 月下旬恢復觀測 (圖十三 (a) 和 (b))，所記錄到地磁資料及儀器姿態都相當不錯。

在今年度資料品質較佳的觀測站，都可以看到各測站的 Y 和 Z 分量有些偏移現象。經比對儀器姿態記錄，該偏移現象與儀器傾斜狀態並無直接關聯，可以進一步的探究是否與地磁場長期變化有關。

(2) 地磁場強度改變與地震活動的關聯性

在採用日變化比值分析方法來探討地磁場強度改變與地震活動的關聯性，過去都選擇資料品質佳且地震活動度低的馬祖 (MT) 觀測站做為參考站。本計畫為了解地磁觀測站日變化比值的背景，分別選擇本島資料品質較佳的地磁觀測站進行分析研究，分別是池上 (CS)、鯉魚潭 (LY)、曾文 (TW) 及爺亨 (YH) 等 4 個地磁觀測站資料。因應分析方法，以 1 月 14 日當天為 0 天，取前 10 天及後 290 天進行分析 1/1 至 10/31 的地磁訊號，15 天為一窗格計算，每次以 1 天移動窗格，進行日變化振幅的比值分析，標準差為 1/1 至 10/31 地磁訊號的統計結果。

2025 年 1 月 21 日清晨在臺南市楠西區發生規模 6.4 的地震。在震央附近的曾文 (TW) 地磁觀測站於 2024 年 4 月儀器發生故障送修，在今年 2 月初重新安裝後恢復正常觀測及資料傳輸 (圖十 (a) 和 (b))，因資料的中斷而無法對於發生在臺南楠西的群震進行分析。不過仍嘗試選擇離震央最近且位於西半部的暨南 (CN) 地磁觀測站進行分析，以馬祖 (MT) 地磁觀測站做為參考站。圖十四為暨南 (CN) 相對馬祖 (MT) 站全磁場日變化差值的比值結果。雖然在 3 月中下旬該站相對馬祖 (MT) 站全磁場日變化差值的比值皆有大於標準差的情形，但以其偏移的密度及振幅應該與 0121 楠西地震活動沒有相關性。

爺亨 (YH) 地磁觀測站相對其他 3 個地磁觀測站的地震活動度較低，通常做為本島的參考站。先選擇位於桃園市復興區的爺亨 (YH) 地磁觀測站與馬祖 (MT) 觀測站進行分析，判別兩站之間有無存在比值上的差異，此做法可確認異常來源為何。圖十五為爺亨 (YH) 站相對馬祖 (MT) 站全磁場日變化差值的比值結果。兩站全磁場強度的日變化差值比值中位數值為 0.9，標準偏差的上、下界在 ± 0.2 之間，全磁場強度日變化差值的比值在 3 月之後有部分比值大於一個標準偏差。位於臺灣苗栗的鯉魚潭 (LY) 地磁觀測站，一直是記錄穩定且資料品質佳的觀測站，將鯉魚潭 (LY) 站分別相對馬祖 (MT) 及爺亨 (YH) 兩個站，進行全磁場日變化差值比值的分析，結果如圖十六 (a) 和 (b) 所示。圖十六 (a) 為鯉魚潭 (LY) 站相對馬祖 (MT) 站之全磁場強度日變化差值的比值，中位數值為 0.95，標準偏差的上、下界在 ± 0.15 之間，全磁場強度日變化差值的比值也是在 3 月之後有部分比值大於一個標準偏差。鯉魚潭 (LY) 站相對爺亨 (YH) 站之全磁場強度日變化差值的比值 (圖十六 (b))，中位數值為 0.95，標準偏差的上、下界在 ± 0.1 之間，全磁場強度日變化差值的比值幾乎都小於一個標準偏差。

另選擇位於臺南的曾文 (TW) 地磁觀測站，該站也是資料品質不錯的觀測站，將曾文 (TW) 站分別相對馬祖 (MT) 及爺亨 (YH) 兩個站，進行全磁場日變化差值比值的分析，結果如圖十七 (a) 和 (b) 所示。曾文 (TW) 站相對馬祖 (MT) 站之全磁場強度日變化差值的比值 (圖十七 (a))，中位數值為 0.85，標準偏差的上、下界在 ± 0.25 之間，全磁場強度日變化差值的比值有不少比值大於一個標準偏差。曾文 (TW) 站相對爺亨 (YH) 站之全磁場強度日變化差值的比值 (圖十七 (b))，中位數值接近 1.0，標準偏差的上、下界約在 ± 0.2 之間，全磁場強度日變化差值的比值在 7、8 月間有部分比值大於一個標準偏差。在圖十七 (a) 和 (b) 中，於 0121 楠西地震群前後都沒有發現日變化差值的比值偏離標準差的情況。嘗試選擇資料品質不錯的池上 (CS) 地磁觀測站，該站位於地震頻仍的臺東地區，將池上 (CS) 站分別相對馬祖 (MT) 及爺亨 (YH) 兩個站，進行全磁場日變化差

值比值的分析，結果如圖十八 (a) 和 (b) 所示。池上 (CS) 站相對馬祖 (MT) 站之全磁場強度日變化差值的比值 (圖十八 (a))，中位數值為 0.8，標準偏差的上、下界在 ± 0.25 之間，全磁場強度日變化差值的比值大於一個標準偏差，其分佈較為零散。池上 (CS) 站相對爺亨 (YH) 站之全磁場強度日變化差值的比值 (圖十八 (b))，中位數值接近 0.9，標準偏差的上、下界約在 ± 0.15 之間，全磁場強度日變化差值的比值有少許大於一個標準偏差，但其零散分佈的情形沒有相對馬祖站發散。

地球磁場的變化 (微脈動) 是和地震活動、洋流及電離層潮汐等有關。如果兩個測站所記錄到的資料品質佳且沒有受前述因素所造成地磁場強度的改變，全磁場強度日變化差值的比值應該都在標準差範圍內。過去從資料品質檢視認為該站的資料品質佳而被選定為整個地磁觀測網的基準參考站 (圖八 (a) 和 (b))。依據上述的分析結果顯示，馬祖 (MT) 地磁站的觀測資料似乎有被觀測站附近的局部地磁場微小變化所影響。馬祖 (MT) 地磁站是否有季節因素或是因花崗岩地質所造成次生磁場干擾，將利用磁場的主成分分析 (Principal Component Analysis, PCA) 來探討可能的原因。

(3) 主成分分析 (Principal Component Analysis, PCA)

本研究針對臺灣多個穩定記錄的地磁測站 (CS、LY、MT、NC、TW、YH) 之三分量地磁資料 (HX、HY、HZ) 進行主成分分析 (PCA)，目的是區分區域性共變磁場與局部異常訊號，並探討其與地殼導電構造及潛在地震活動的關係。在資料前處理階段，本研究將每日地磁記錄串接為連續時間序列，並統一以五分鐘解析度進行重採樣。所有測站資料均依其磁偏角旋轉至地理坐標系，使 X 與 Y 分量對應地理北與東方向；短暫缺測則以保守線性插值補齊，以維持時間序列之連續性。

主成分分析分別應用於 X、Y、Z 三個分量，並在分析前對各測站資料進行標準化處理，以去除振幅差異並強化量測間的相關結構。在初步測試中 (以 05/01~05/15 資料為例)，PCA 已成功提取全臺測站共同的 PCA1 與 PCA2 趨勢 (圖十九)，提供一套穩定的區域背景場模型。分析結果顯示，PCA 的前兩個主成分 (PC1 與 PC2) 即可解釋大部分的資料變異 (Explained var (%): X=98.3+1.1, Y=99.2+0.4, Z=92.4+5.2)，累積貢獻率對 X、Y 分量皆超過 99%，Z 分量也達到約 97%。PC1 代表全區同步的共模磁場變化，主要反映外部電流系統 (磁暴、電離層電流等) 的控制；PC2 則呈現不同測站間之相位或振幅差異，可能與地下導電結構、局部感應效應或地形與地質差異有關。透過 PCA loading，可以進一步確認各測站在不同主成分中的貢獻程度 (類似多項式中的係數)，並反映其是否受地方性效應影響。

後續分析亦將採用「扣除 PC1 + PC2」的方法，透過殘餘場強調地方性異常訊號，其結果如圖二十至二十五。初步判斷，各站在剩餘場的振幅皆差不多在正負 10nT 內，說明該訊號確實為 local 所造成，而馬祖 (MT) 與內城 (NC) 站在垂直方向有較高頻的變化，爺亨 (YH) 站在 X、Y 方向有較高頻的變化。未來將從時間與空間兩個面向進一步分析這些殘餘場，此方法成功分離共模與非共模訊號，為未

來探討臺灣地磁微擾與地下構造或地震前兆之間的連結奠定重要基礎。

(4) 新建地磁站找站事宜

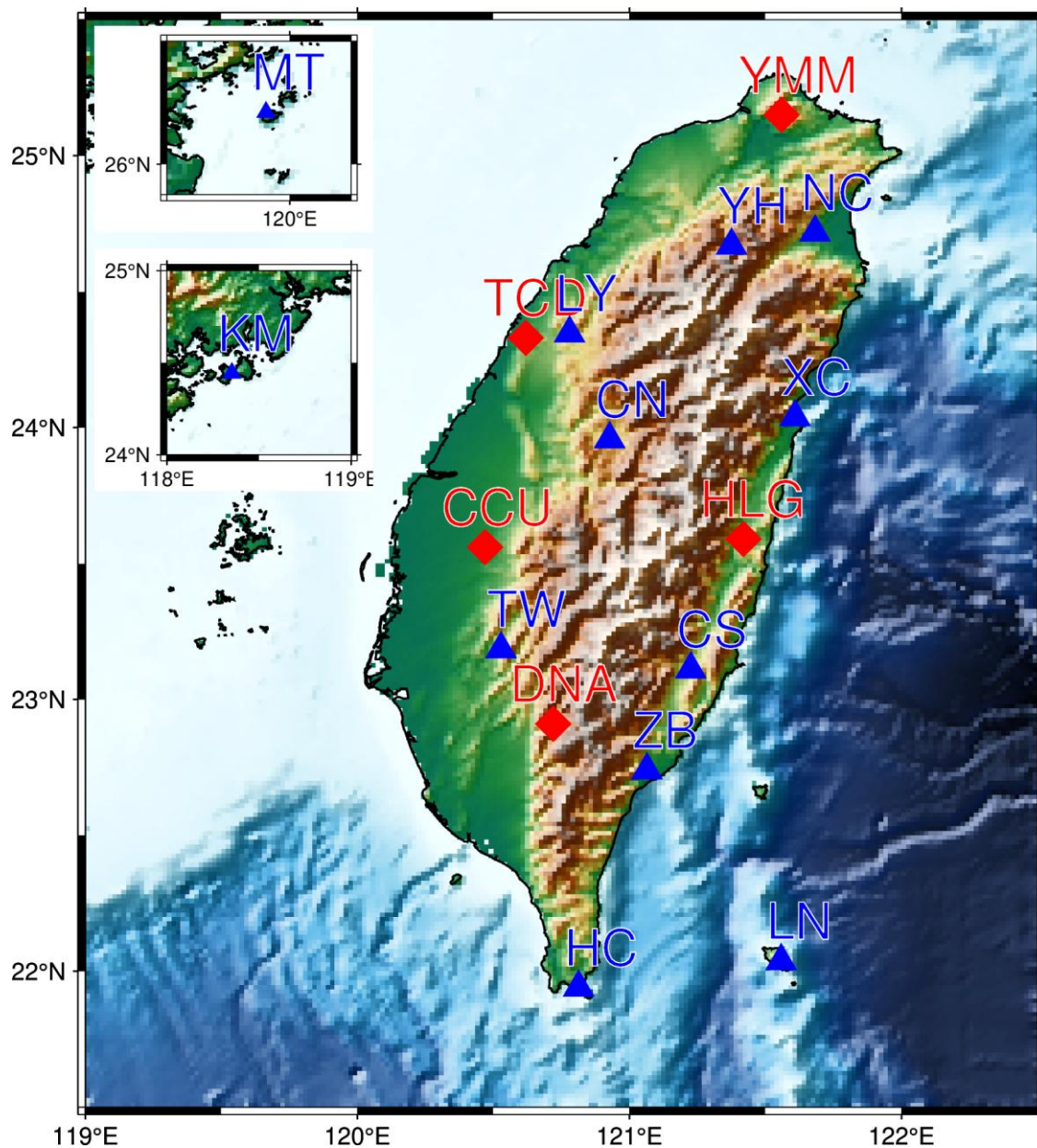
本年度已陸續進行新設之地磁站的找站工作。目前已完成三個站的勘查及整修工程，分別位於阿里山公路上之山美農場茶園一隅、中正大學校園內一隅及矯正署光復外役監獄的農場一隅。中正大學觀測站於 2005 年建站完成並裝設三分量地磁觀測儀器，觀測站因年久失修而無法裝設新的觀測儀器，目前已經完成內外部的整工程，整修前後的照片如圖二十六和二十七。

肆、誌謝

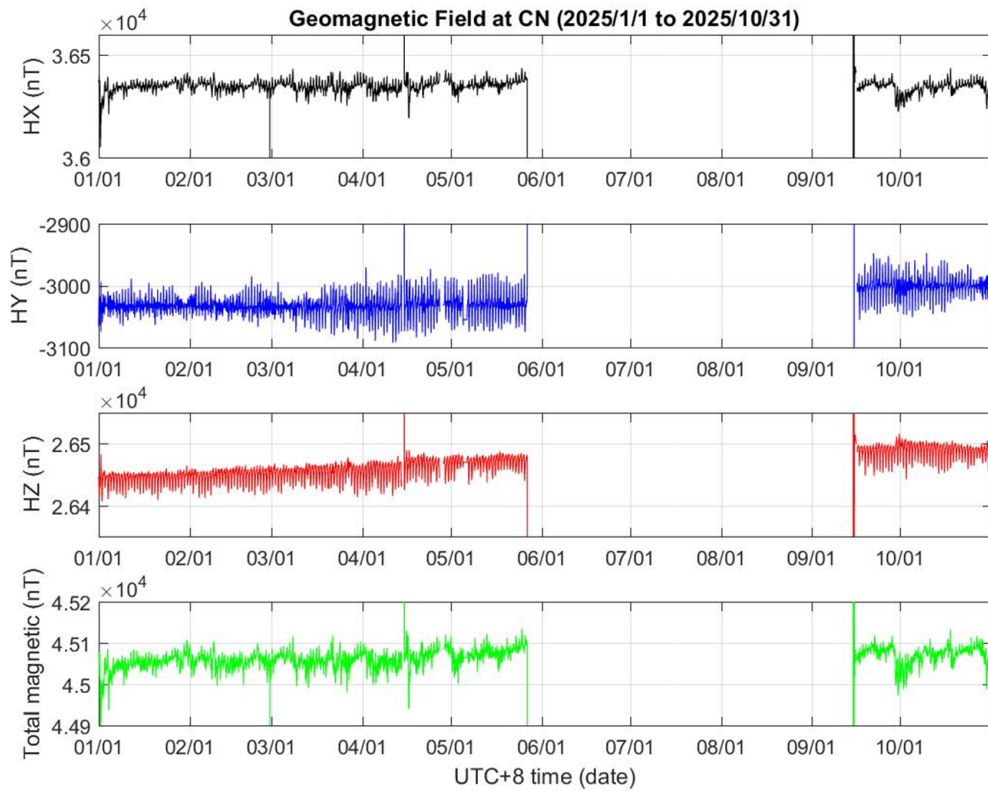
本計畫由中央氣象署提供經費，地震中心同仁協助測站維運及資料處理，在此深表感謝。

參考文獻

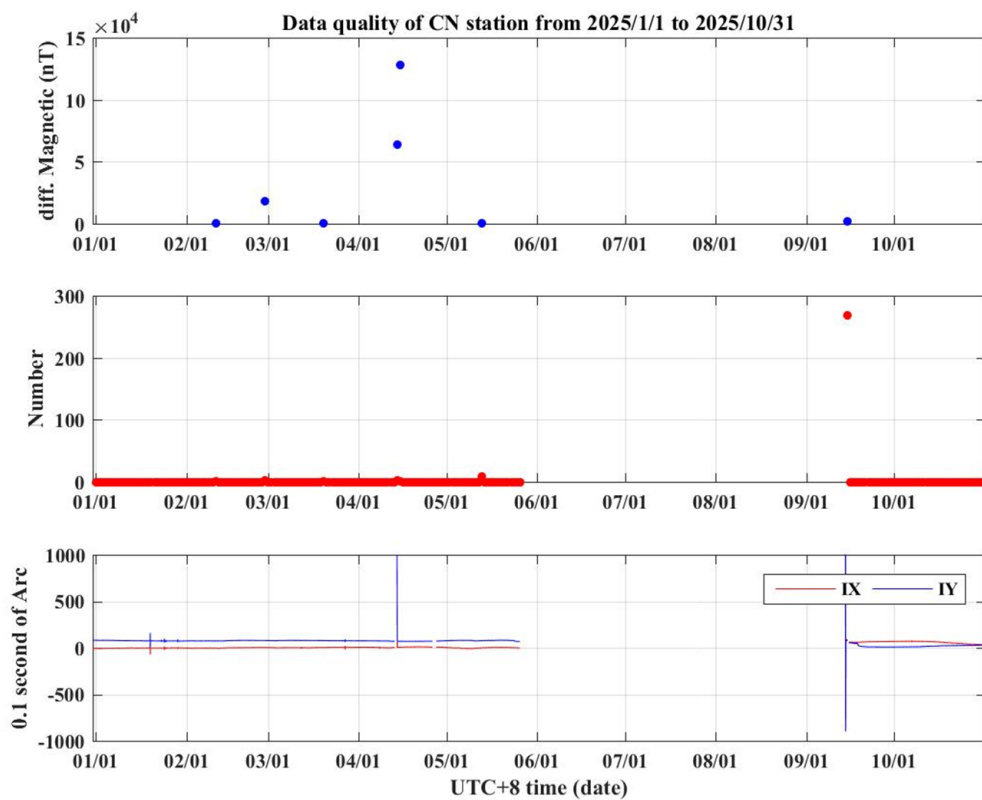
- Liu, J. Y., C. H. Chen, Y. I. Chen, H. Y. Yen, K. Hattori, K. Yumoto (2006). Seismo-geomagnetic anomalies and $M \geq 5.0$ earthquakes observed in Taiwan during 1988–2001. *Physics and Chemistry of the Earth*, 31, 215–222.
- Yen, H. Y., C. H. Chen, Y. H. Yeh, J. Y. Liu, C. J. Lin and Y. B. Tsai (2004). Geomagnetic fluctuations during the Chi-Chi, Taiwan earthquake, *Earth, Planets and Space*, 56, 39-45.



圖一、中央氣象署三分量地磁連續觀測網之測站位置圖。(紅色菱形為中研院地球所的三分量地磁站)。

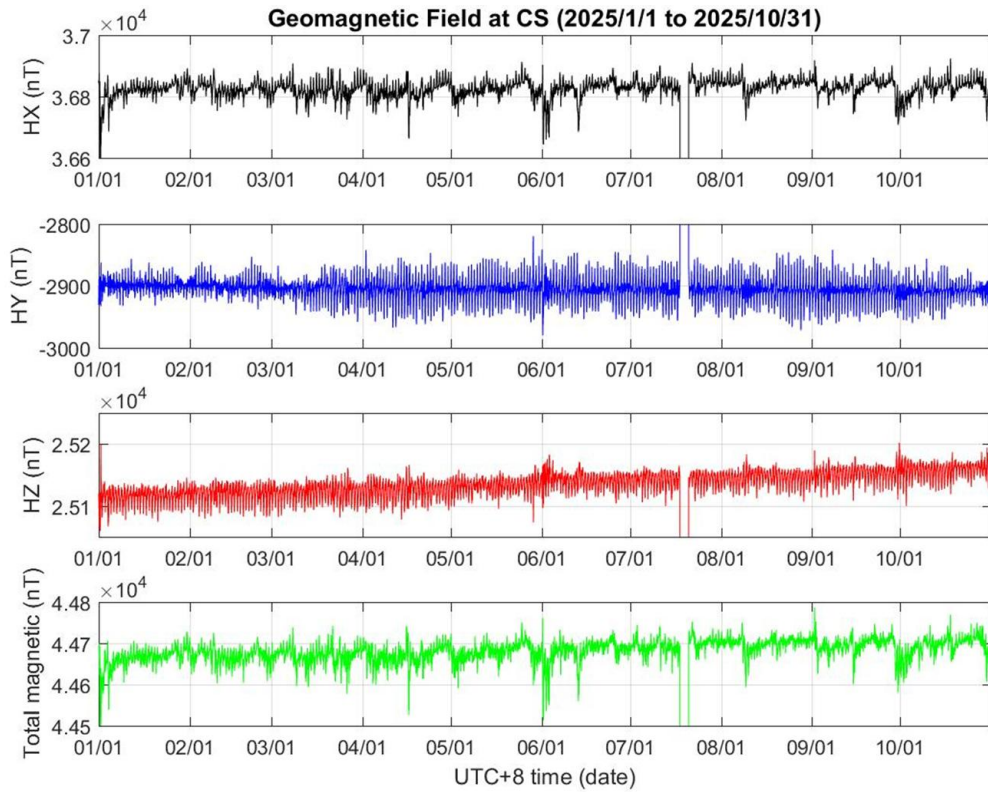


(a)

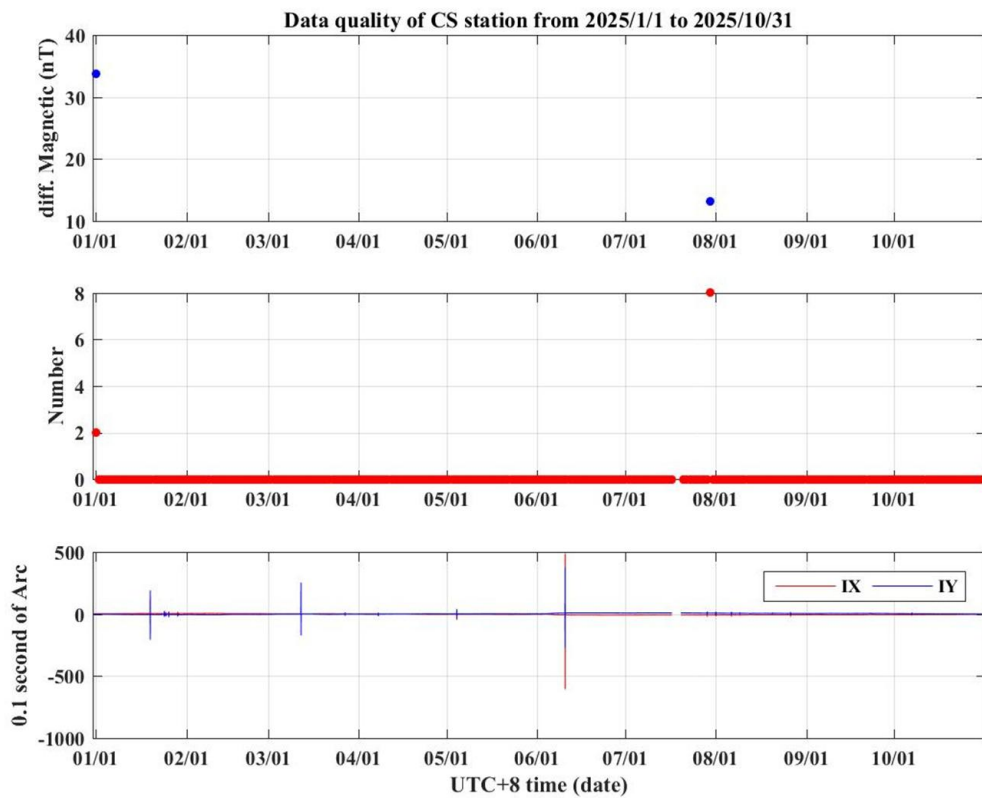


(b)

圖二、(a) 暨南 (CN) 觀測站地磁場三個分量的原始記錄及所計算出全磁場強度的結果；(b) 相鄰兩點的全磁場強度變化量差異量、差異超過 10nT 的個數及儀器傾斜狀態。

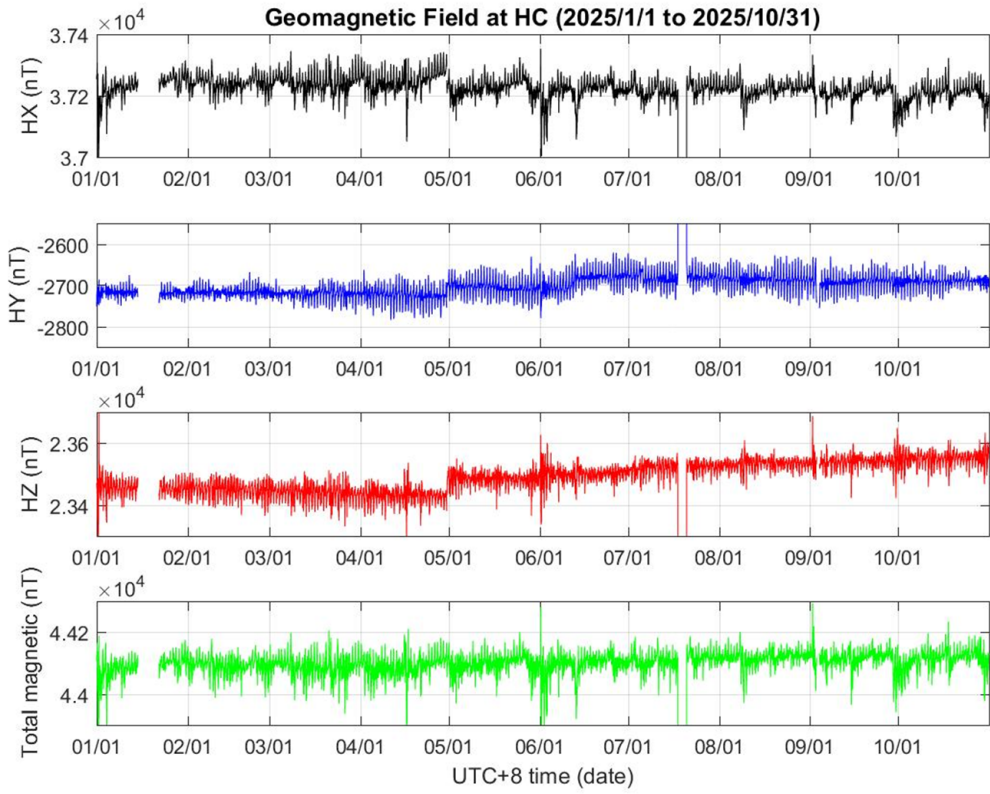


(a)

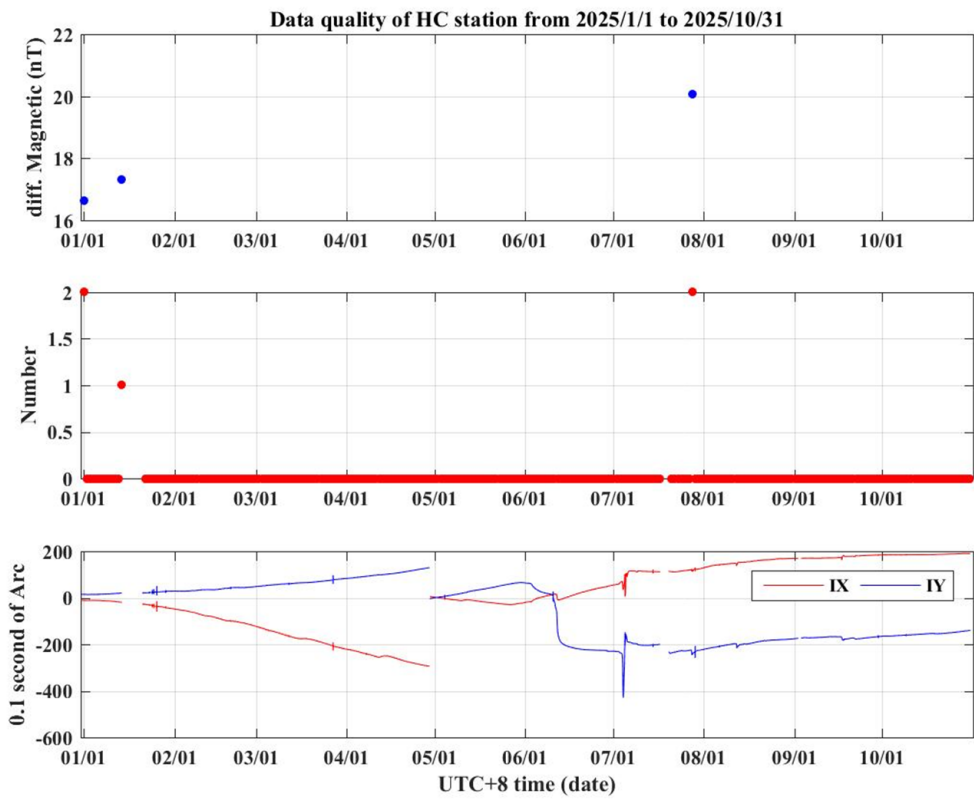


(b)

圖三、 (a) 池上 (CS) 觀測站地磁場三個分量的原始記錄及所計算出全磁場強度的結果; (b) 相鄰兩點的全磁場強度變化量差異量、差異超過 10nT 的個數及儀器傾斜狀態。

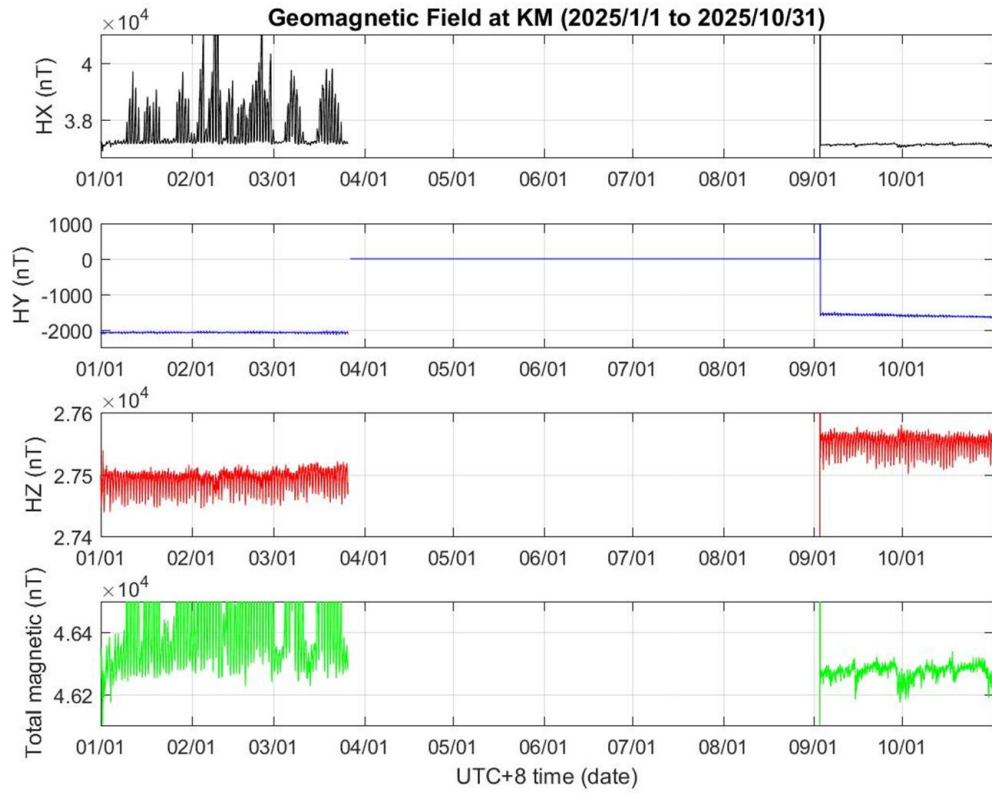


(a)

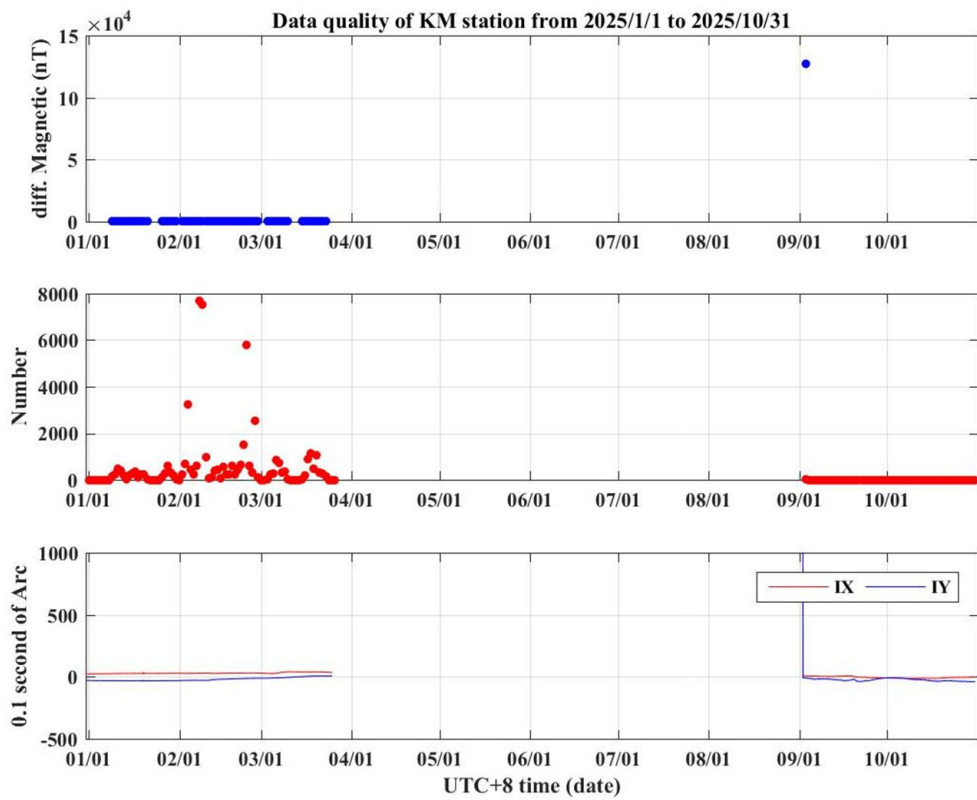


(b)

圖四、恆春 (HC) 觀測站三個分量及全磁場結果(上)、全磁場強度變化量差異量、差異超過 10nT 的個數及儀器傾斜狀態(下)。

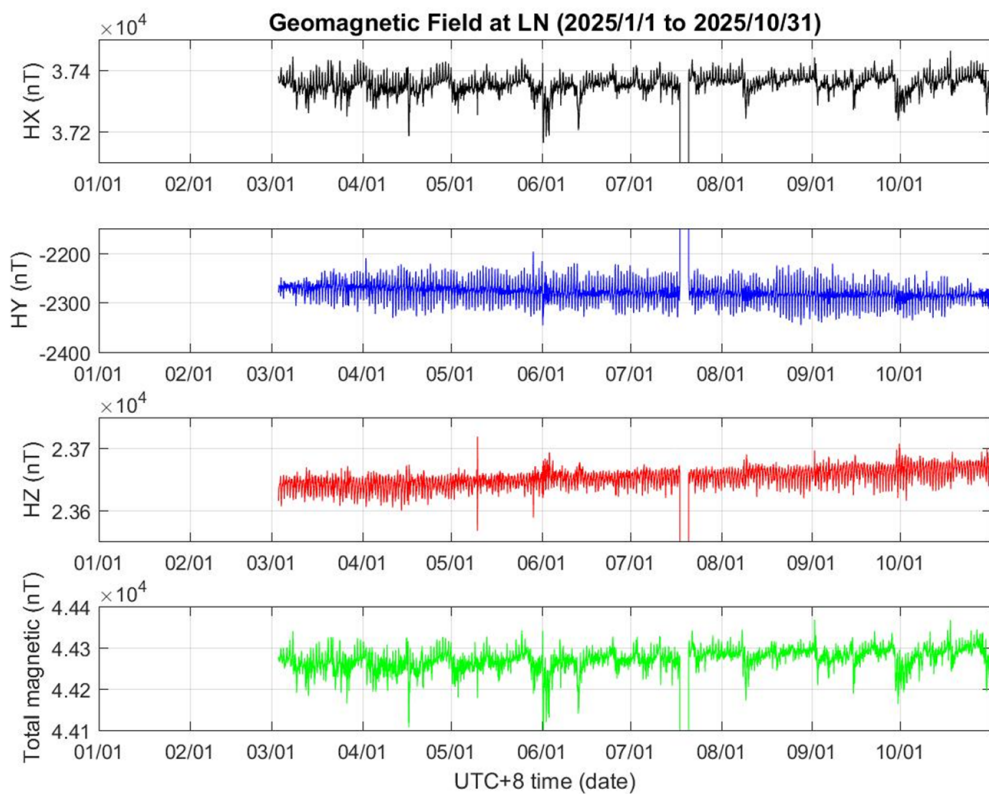


(a)

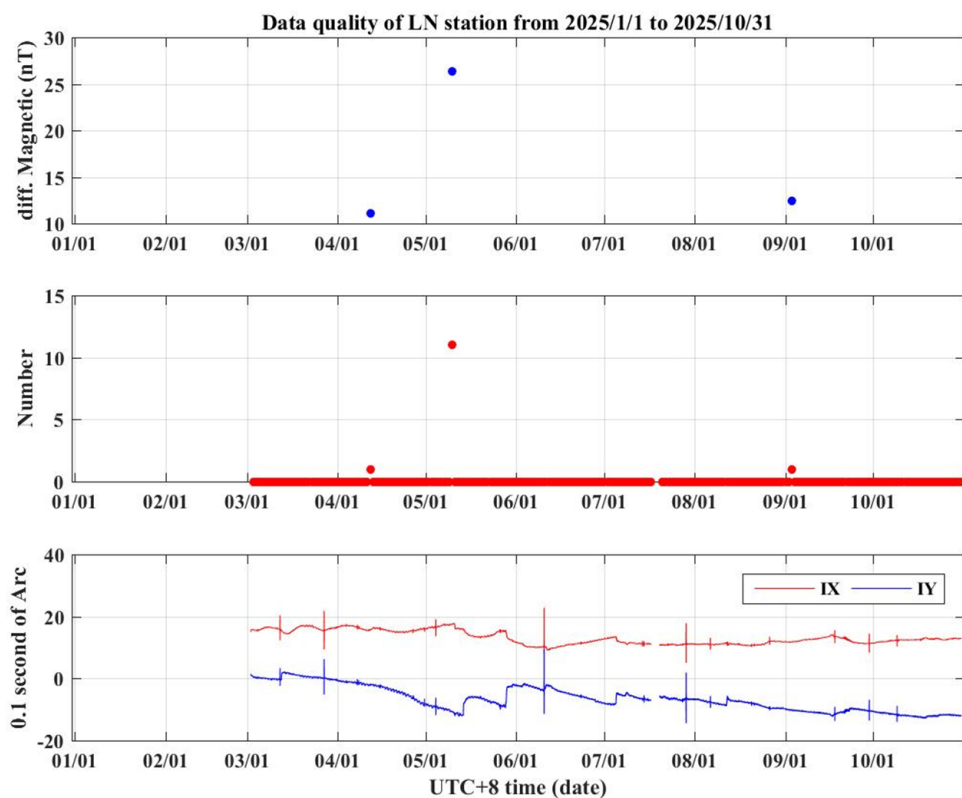


(b)

圖五、金門 (KM) 觀測站三個分量及全磁場結果(上)、全磁場強度變化量差異量、差異超過 10nT 的個數及儀器傾斜狀態(下)。

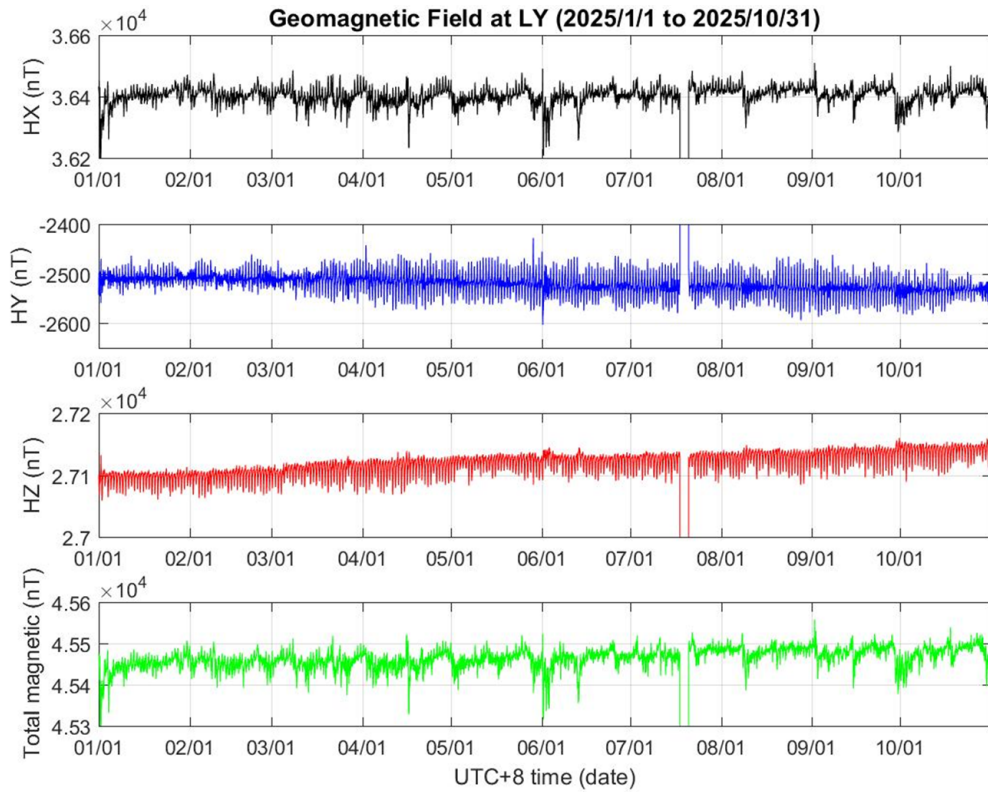


(a)

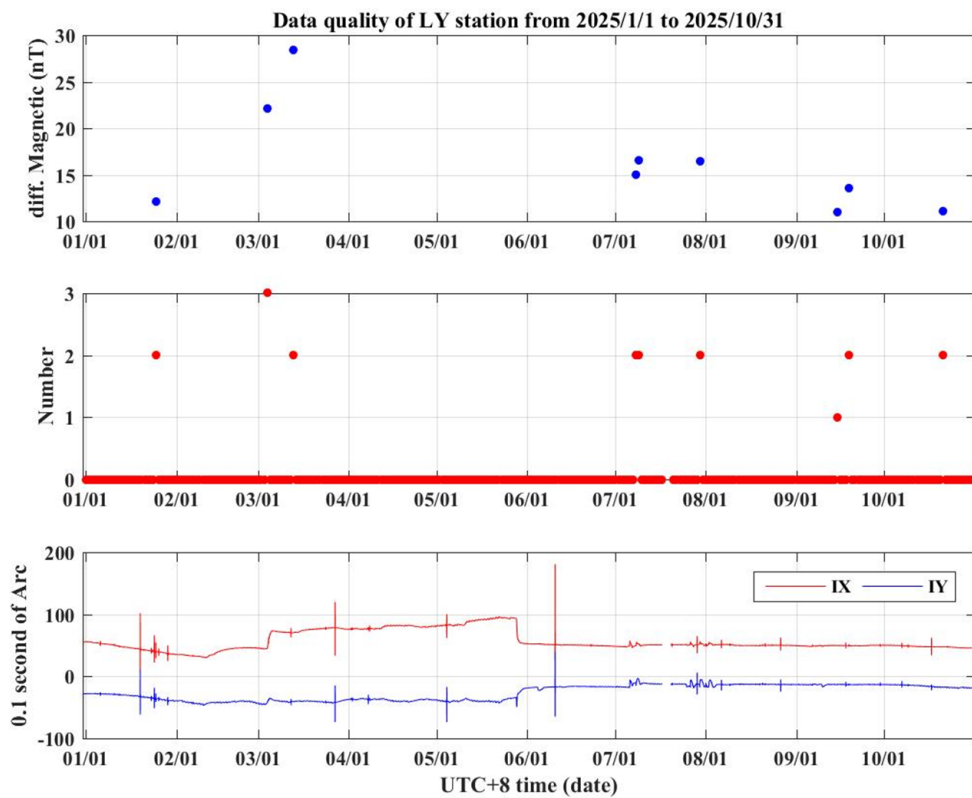


(b)

圖六、蘭嶼 (LN) 觀測站三個分量及全磁場結果(上)、全磁場強度變化量差異量、差異超過 10nT 的個數及儀器傾斜狀態(下)。

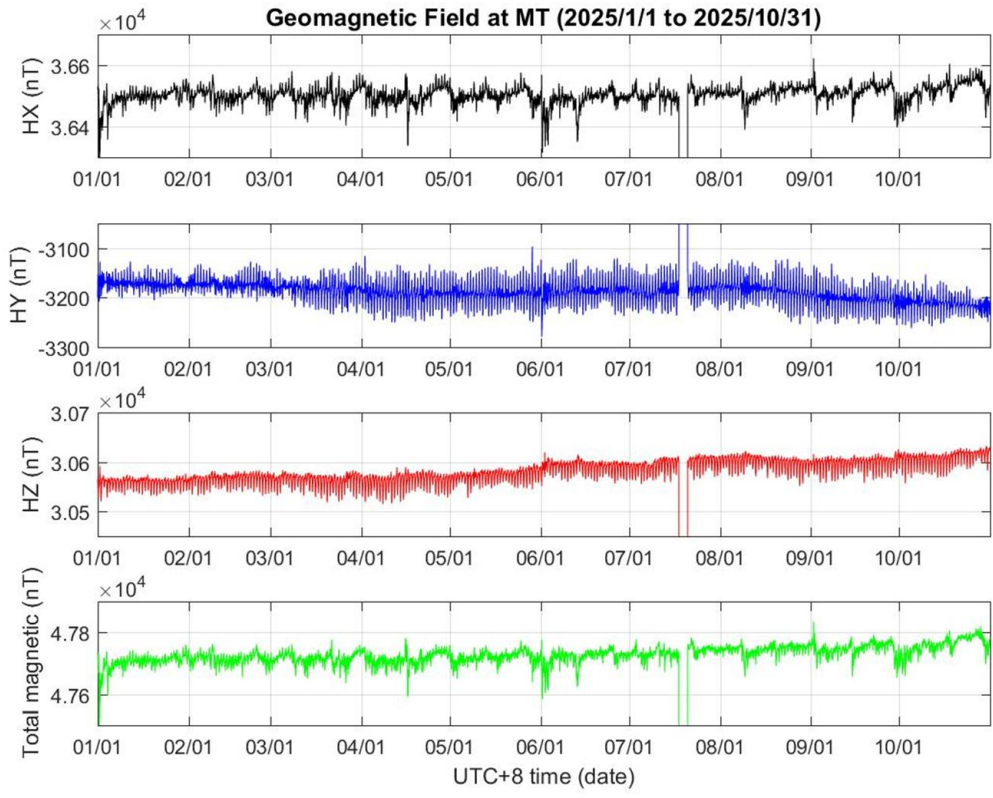


(a)

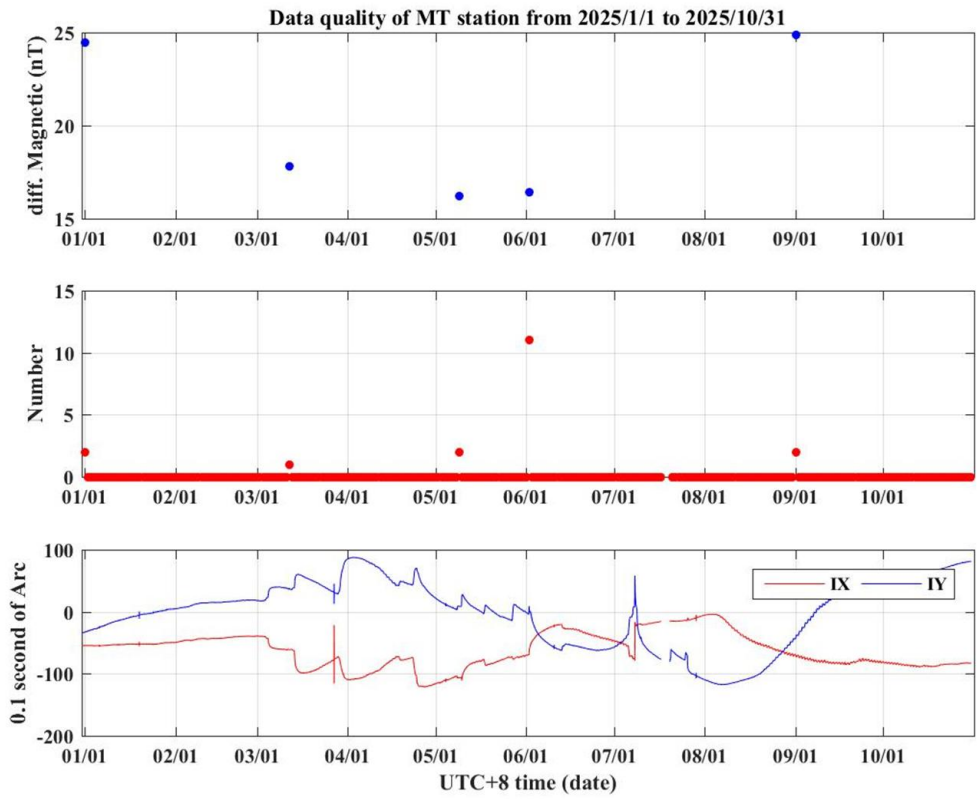


(b)

圖七、鯉魚潭 (LY) 觀測站三個分量及全磁場結果(上)、全磁場強度變化量差異量、差異超過 10nT 的個數及儀器傾斜狀態(下)。

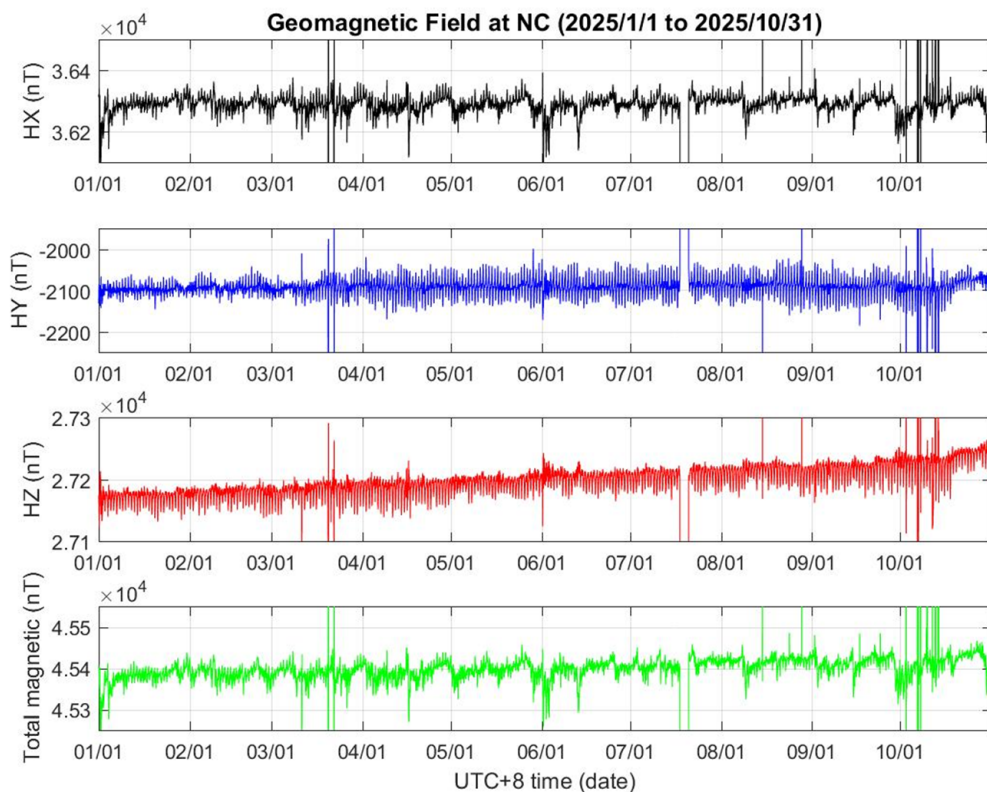


(a)

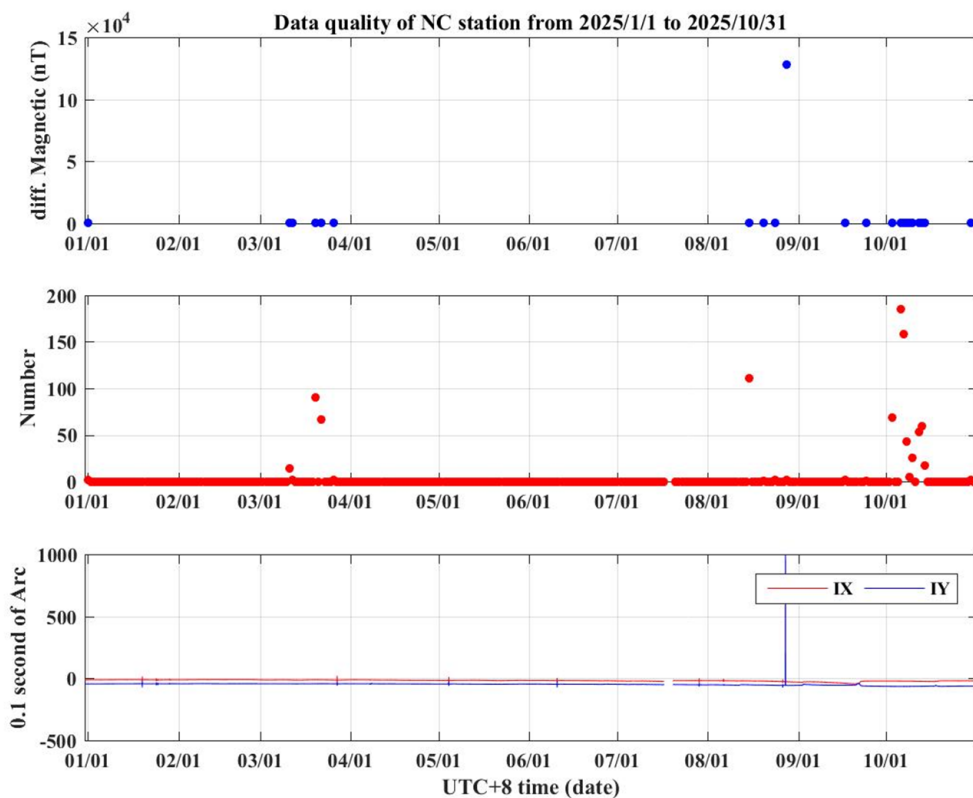


(b)

圖八、馬祖 (MT) 觀測站三個分量及全磁場結果(上)、全磁場強度變化量差異量、差異超過 10nT 的個數及儀器傾斜狀態(下)。

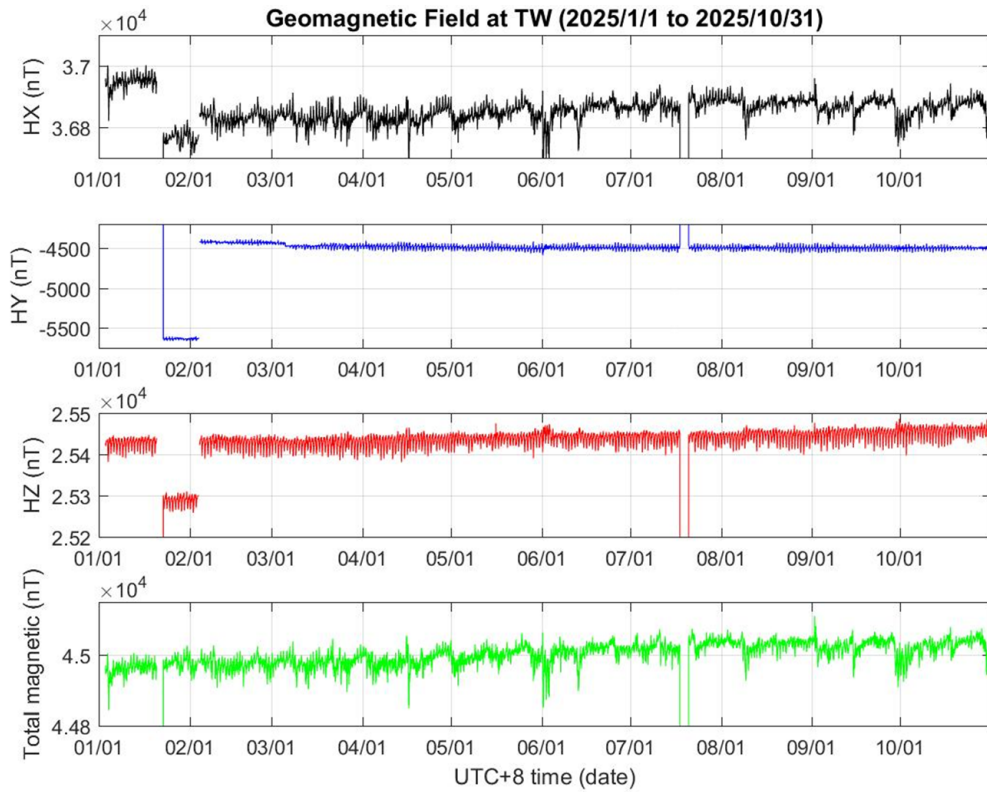


(a)

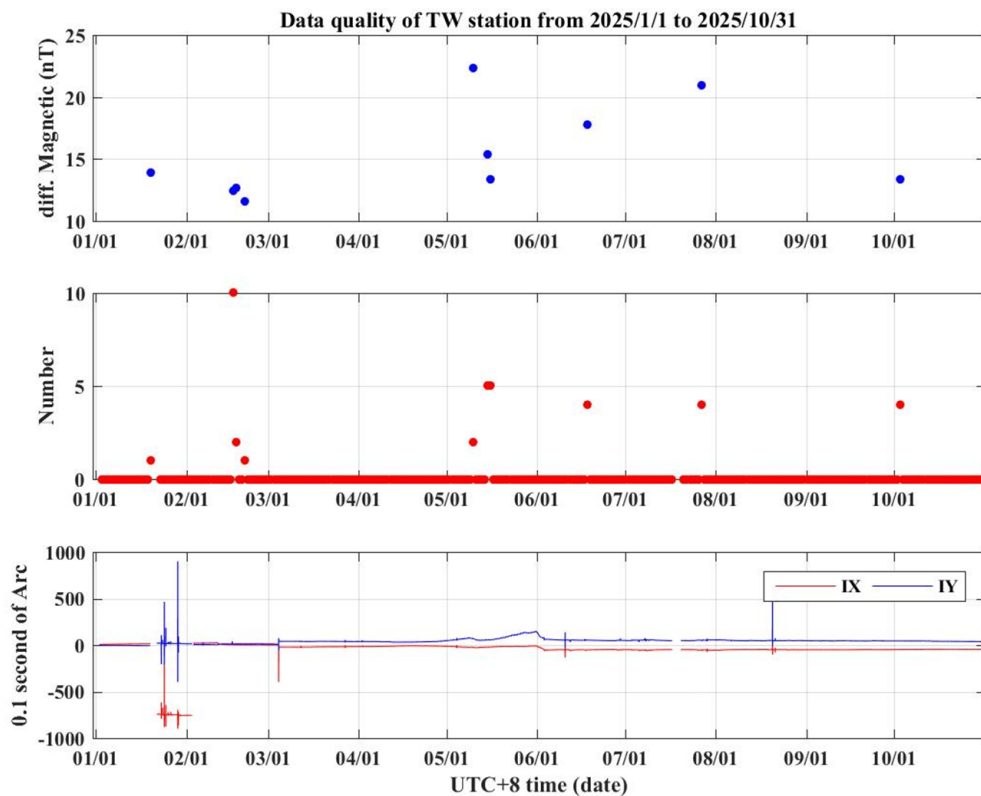


(b)

圖九、內城 (NC) 觀測站三個分量及全磁場結果(上)、全磁場強度變化量差異量、差異超過 10nT 的個數及儀器傾斜狀態(下)。

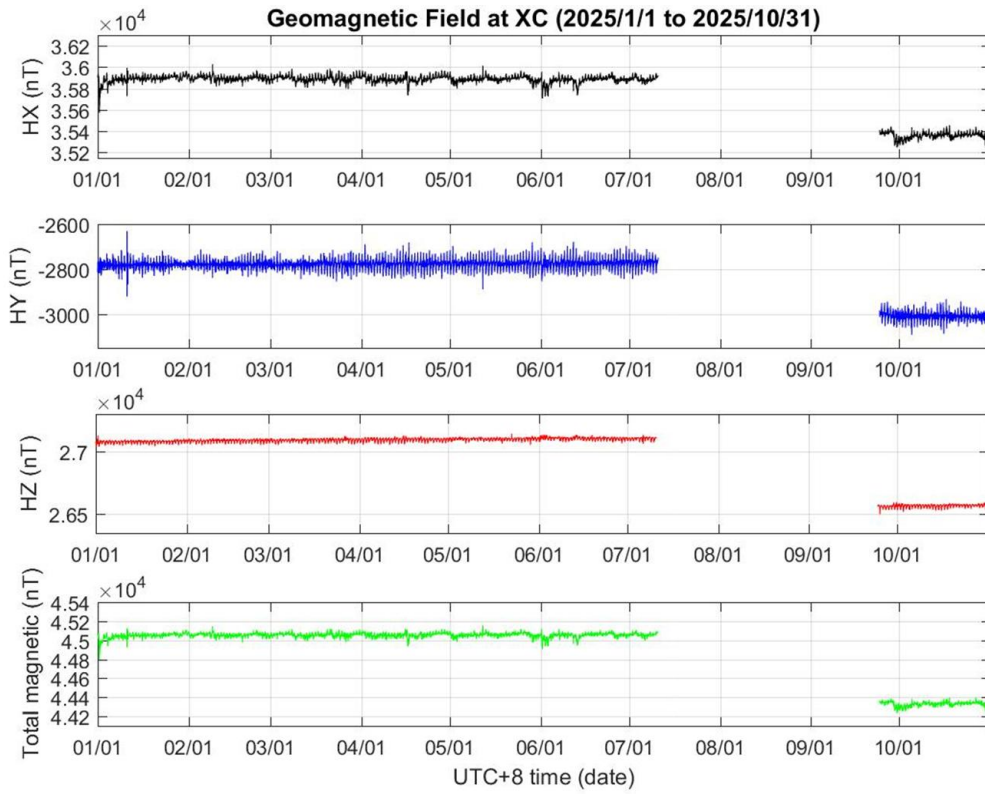


(a)

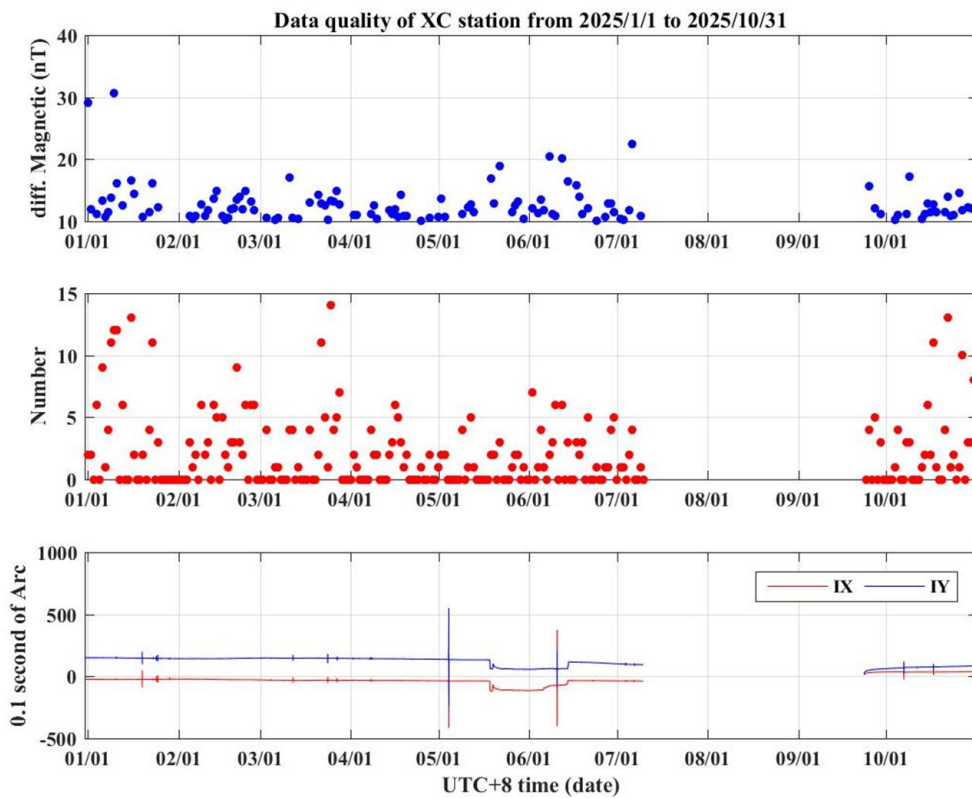


(b)

圖十、曾文 (TW) 觀測站三個分量及全磁場結果(上)、全磁場強度變化量差異量、差異超過 10nT 的個數及儀器傾斜狀態(下)。

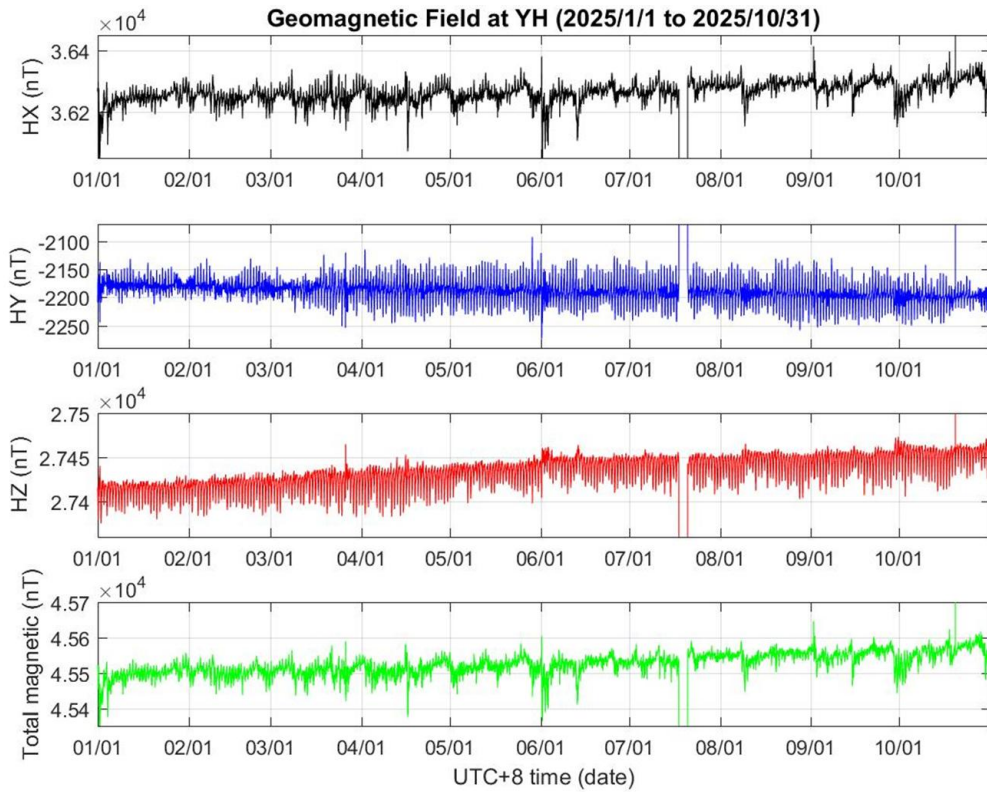


(a)

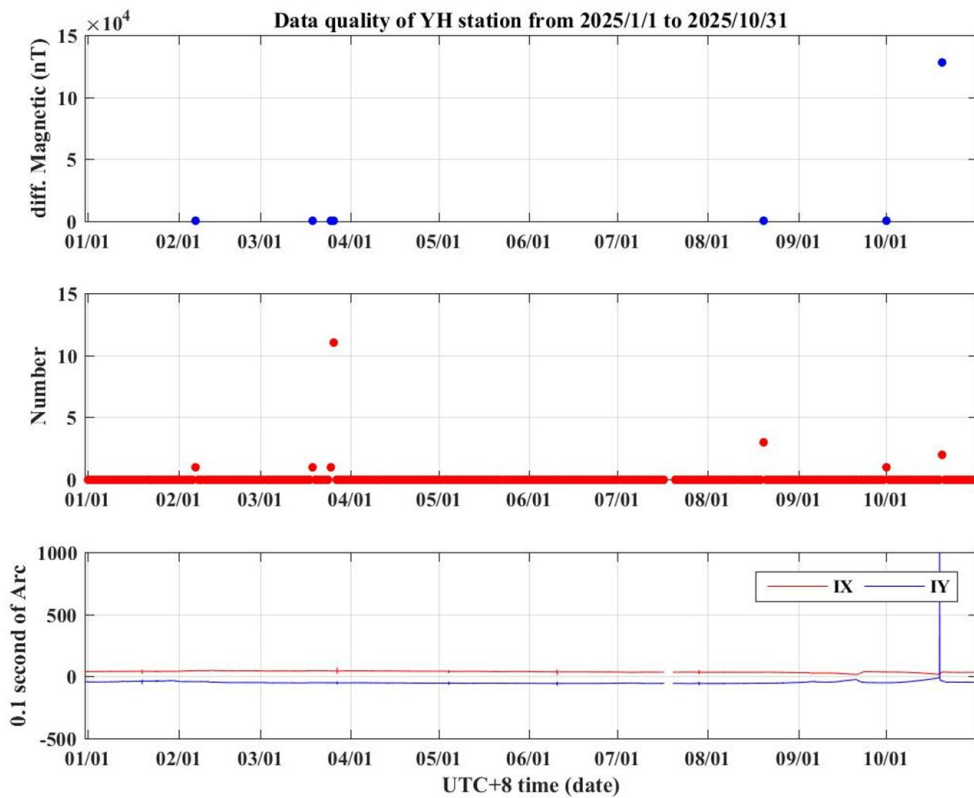


(b)

圖十一、新城 (XC) 觀測站三個分量及全磁場結果(上)、全磁場強度變化量差異量、差異超過 10nT 的個數及儀器傾斜狀態(下)。

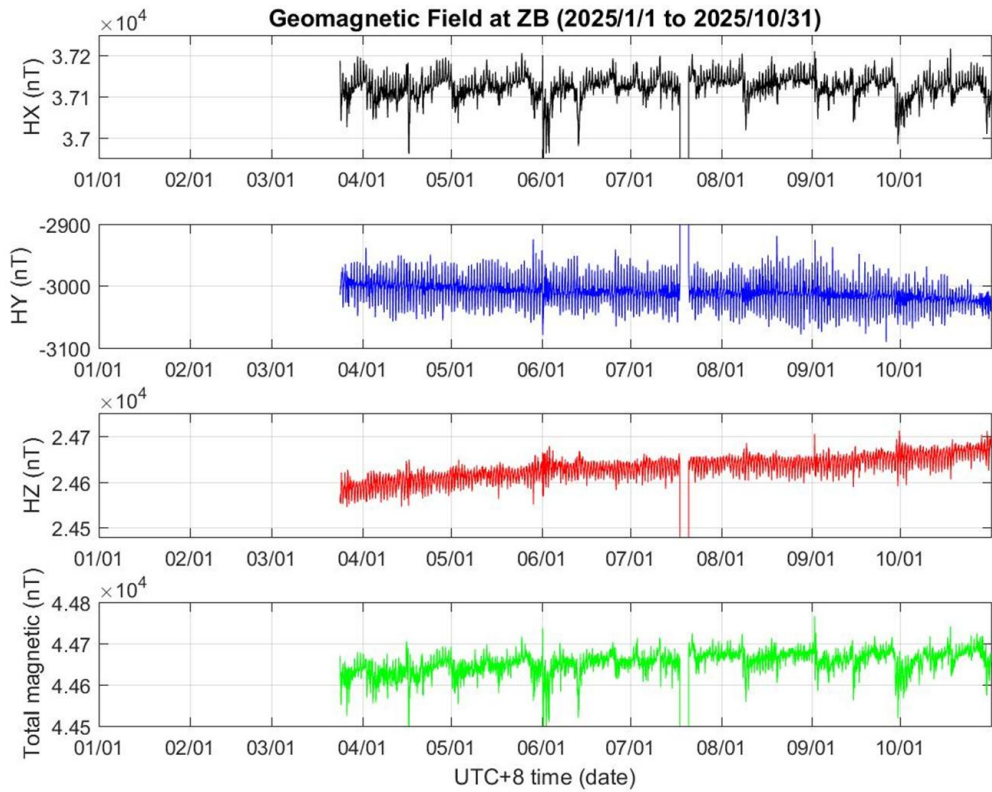


(a)

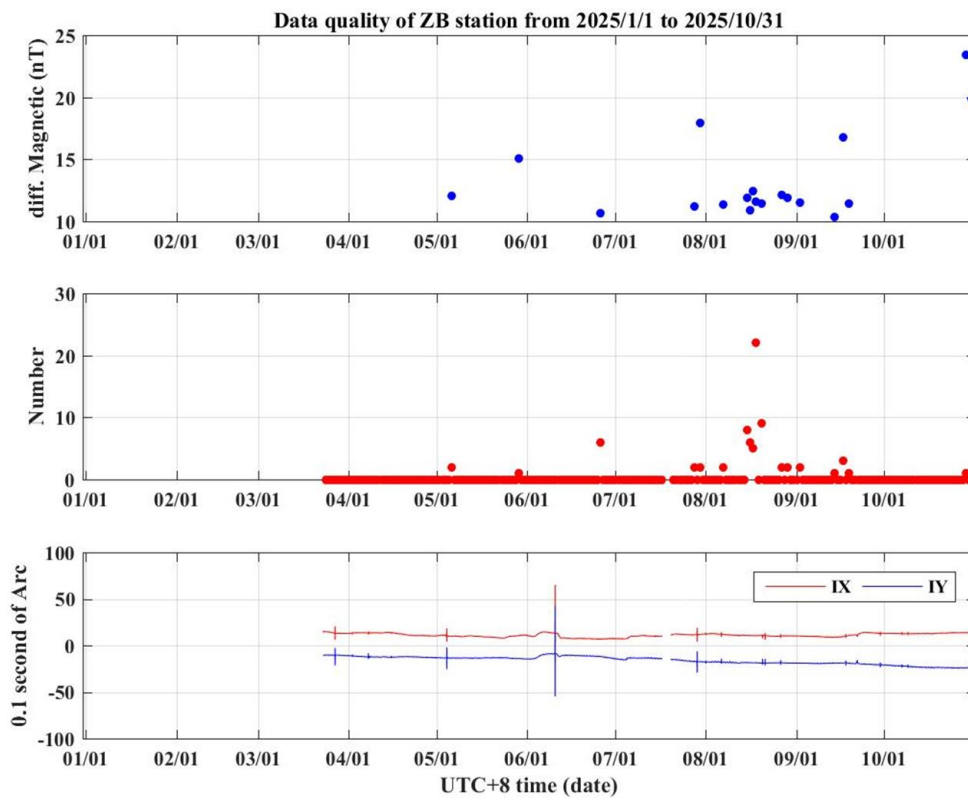


(b)

圖十二、爺亨 (YH) 觀測站三個分量及全磁場結果(上)、全磁場強度變化量差異量、差異超過 10nT 的個數及儀器傾斜狀態(下)。

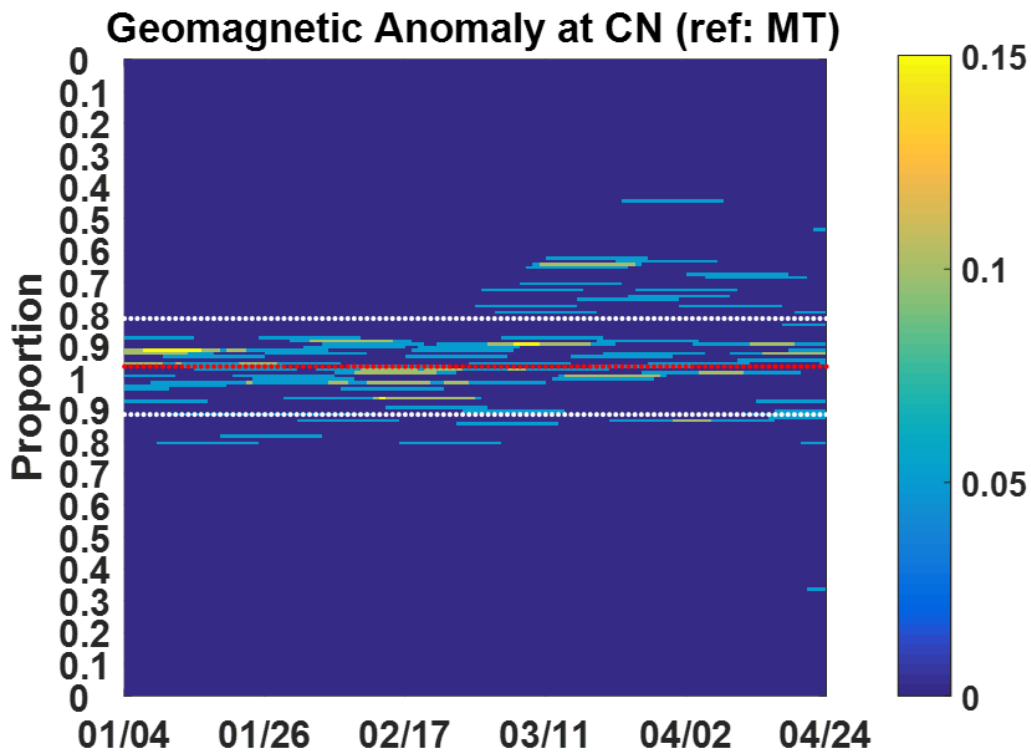


(a)

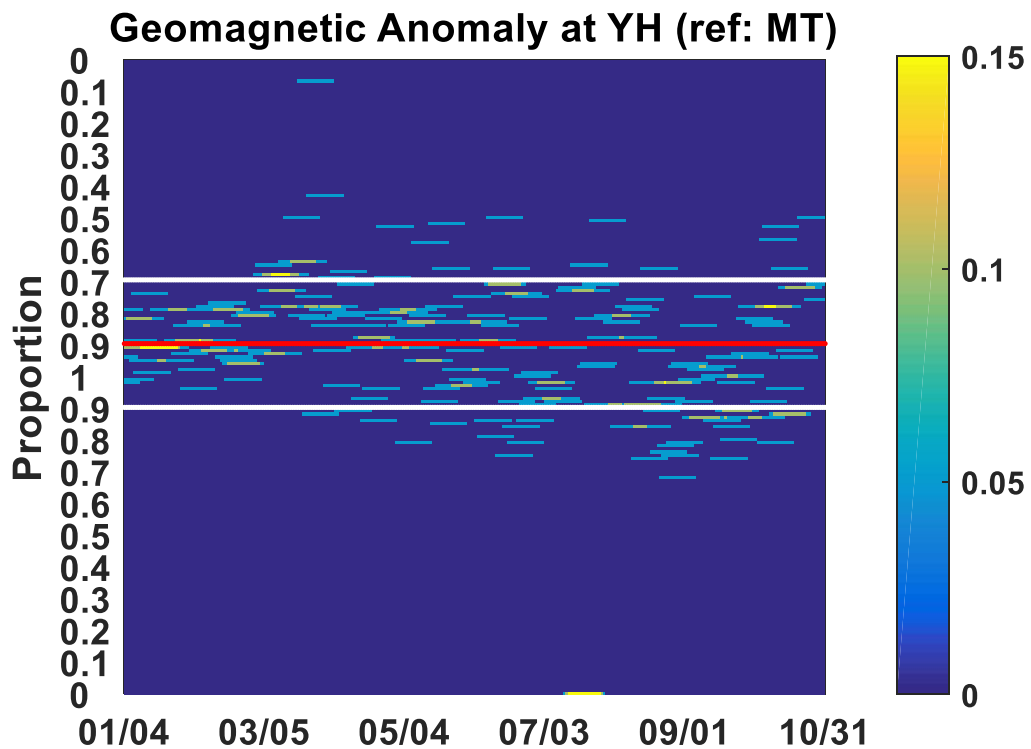


(b)

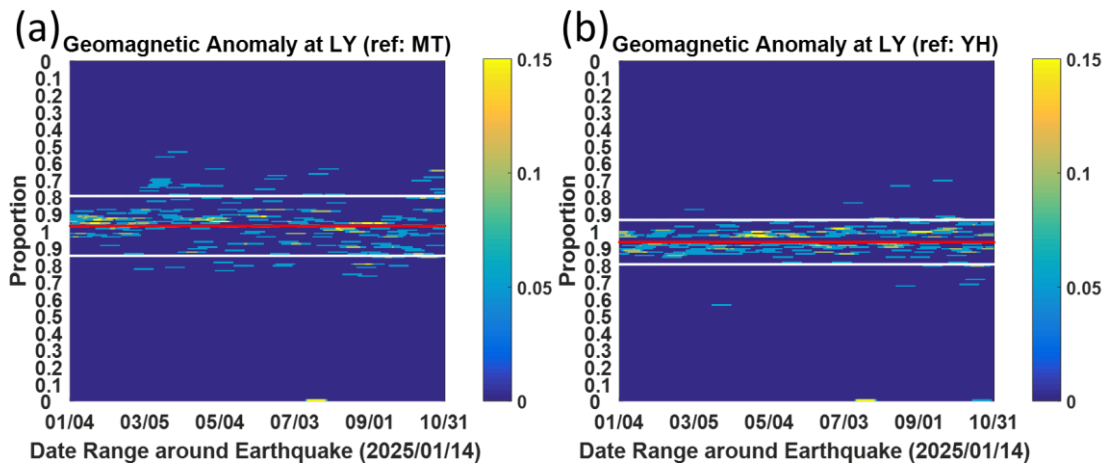
圖十三、知本 (ZB) 觀測站三個分量及全磁場結果(上)、全磁場強度變化量差異量、差異超過 10nT 的個數及儀器傾斜狀態(下)。



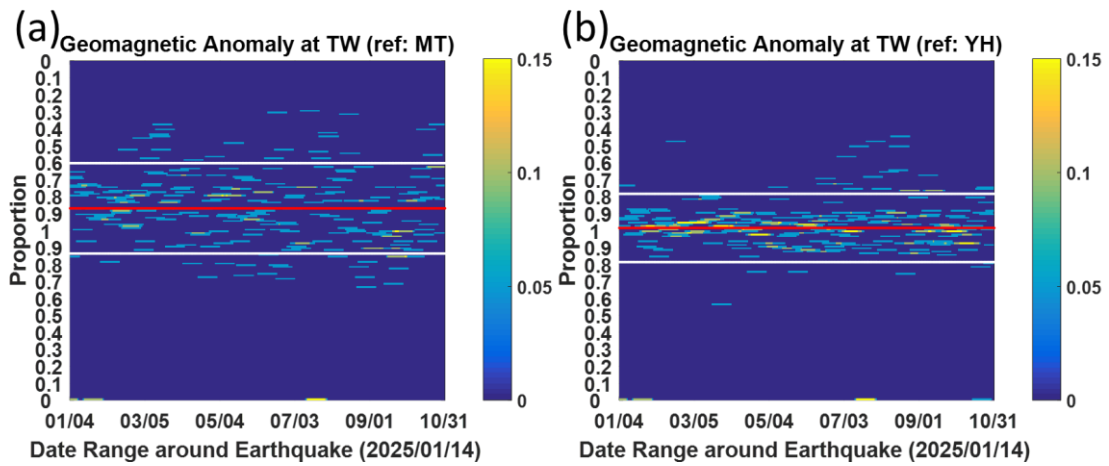
圖十四、分析近半年之日變化振幅比值，以暨南（CN）觀測站相對馬祖站(MT)的全磁場資料進行日變化差值的比值計算，以地震前半年的資料當作背景值，20天為一窗格計算，每次以1天移動窗格，設定4月10日當天為0天，觀察其前80天及後10天的比值分布。



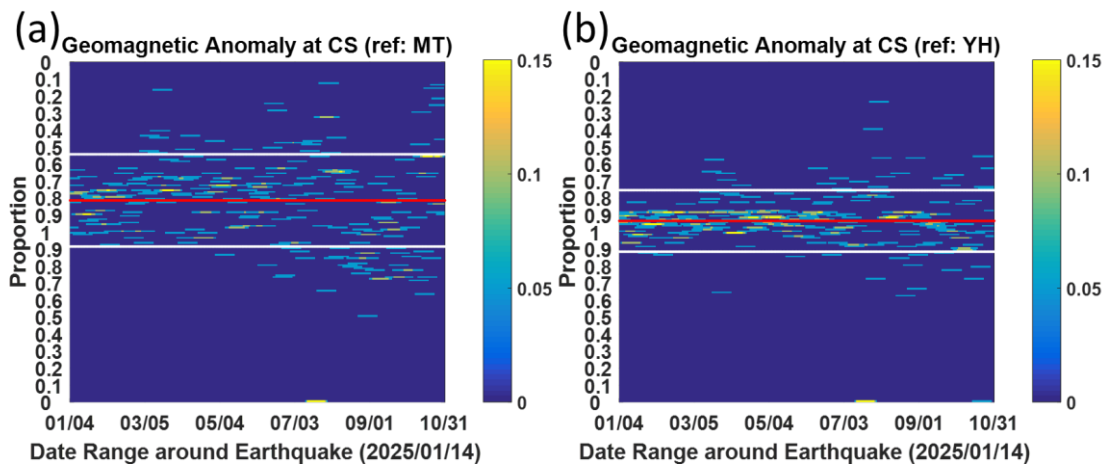
圖十五、為爺亨（YH）站相對馬祖（MT）站全磁場日變化差值的比值結果。



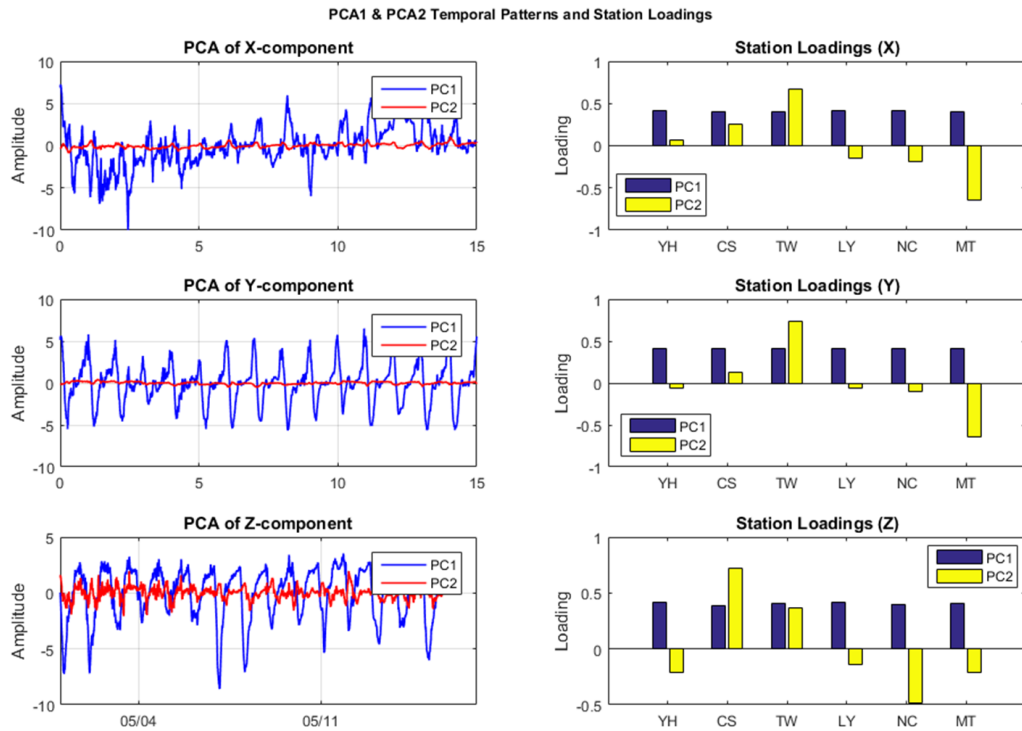
圖十六、鯉魚潭 (LY) 站分別相對馬祖 (MT) 及爺亨 (YH) 兩個站，進行全磁場日變化差值比值的分析



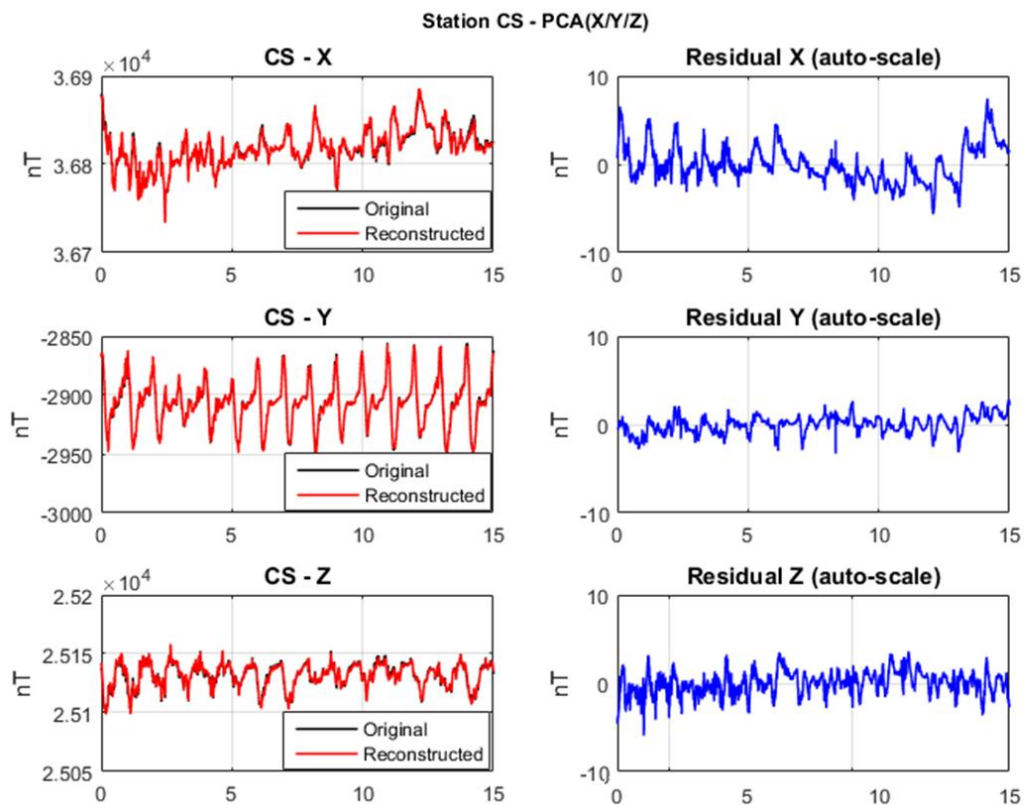
圖十七、曾文(TW) 站分別相對馬祖 (MT) 及爺亨 (YH) 兩個站，進行全磁場日變化差值比值的分析



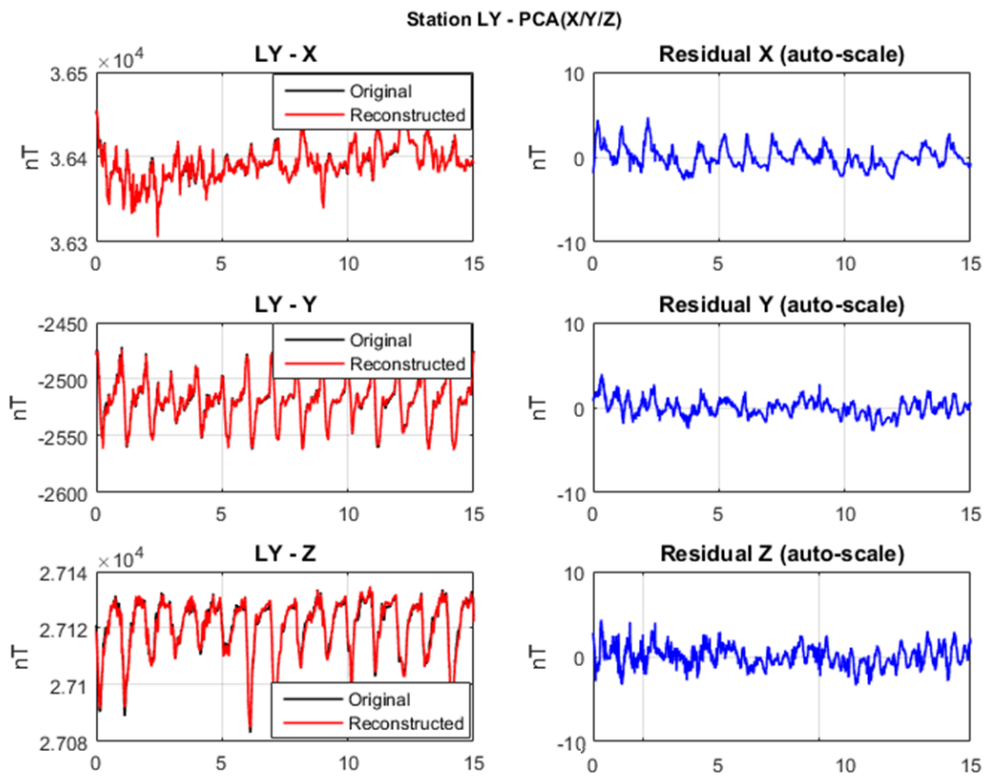
圖十八、池上 (CS) 站分別相對馬祖 (MT) 及爺亨 (YH) 兩個站，進行全磁場日變化差值比值的分析



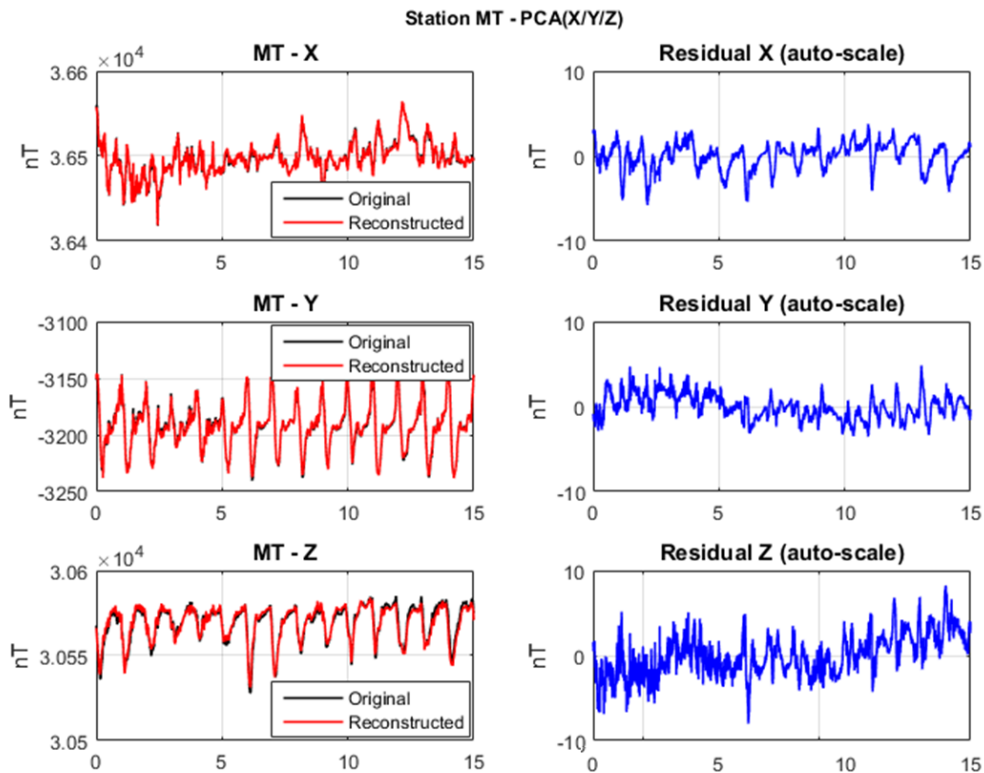
圖十九、計算今年度多個穩定記錄之地磁站共同之 PCA1 及 PCA2 的加成結果，之後藉由扣除主成分 1 和 2 的磁力值，得到每個站的剩餘結果進行異常訊號討論。



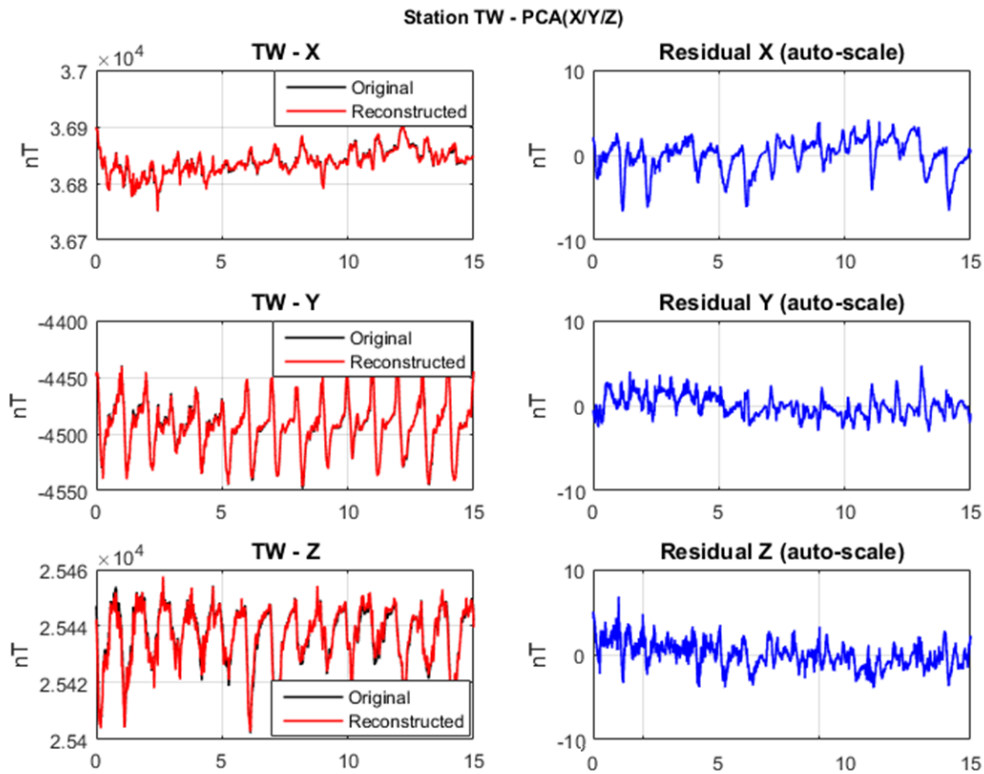
圖二十、池上 (CS) 站扣除 PC1 + PC2 後的殘餘場來強調地方性異常訊號。



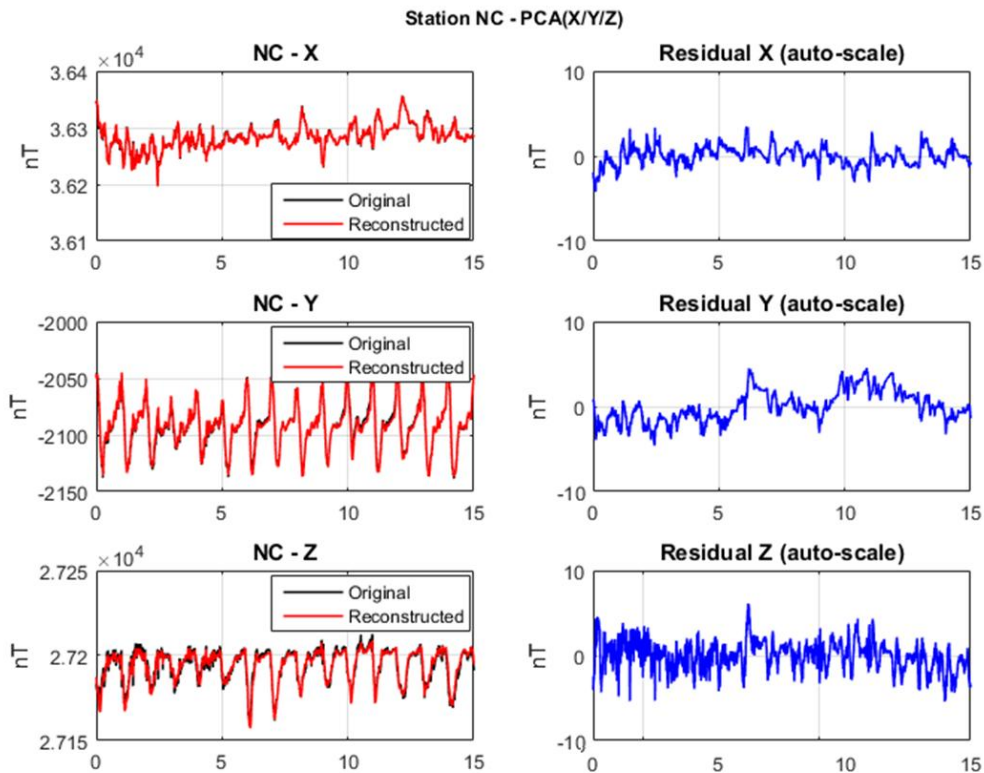
圖二十一、鯉魚潭 (LY) 站扣除 PC1 + PC2 後的殘餘場來強調地方性異常訊號。



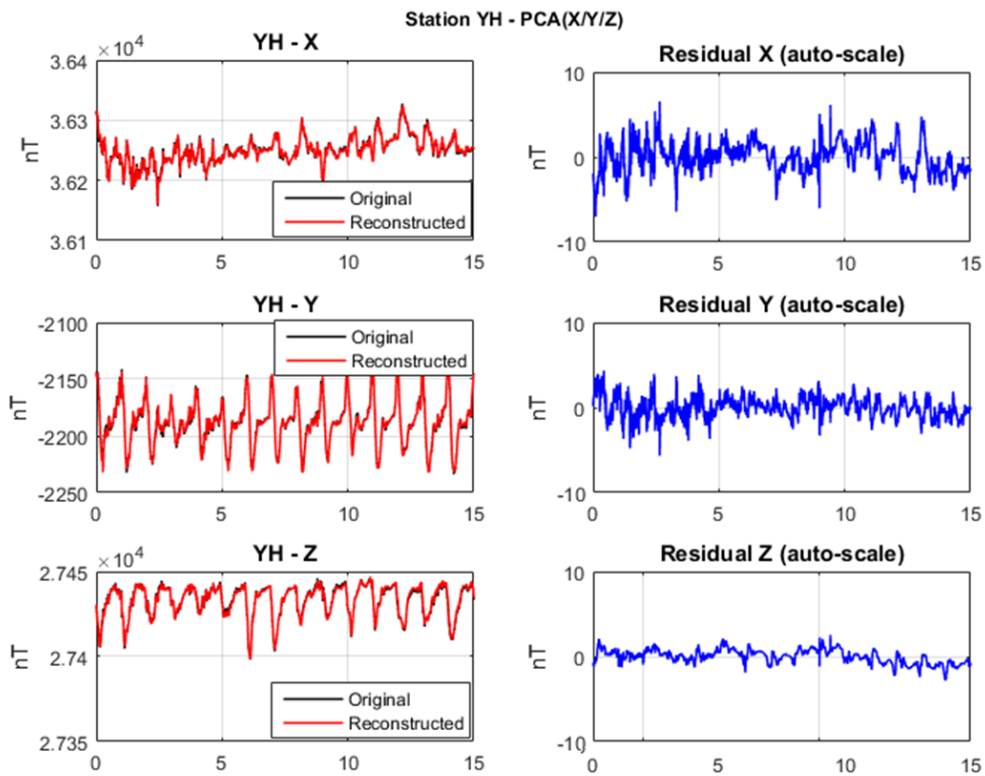
圖二十二、馬祖 (MT) 站扣除 PC1 + PC2 後的殘餘場來強調地方性異常訊號。



圖二十三、曾文 (TW) 站扣除 PC1 + PC2 後的殘餘場來強調地方性異常訊號。



圖二十四、內城 (NC) 站扣除 PC1 + PC2 後的殘餘場來強調地方性異常訊號。



圖二十五、爺亨 (YH) 站扣除 PC1 + PC2 後的殘餘場來強調地方性異常訊號。



圖二十六、中正大學地磁站整修前照片。



圖二十七、中正大學地磁站整修後照片。

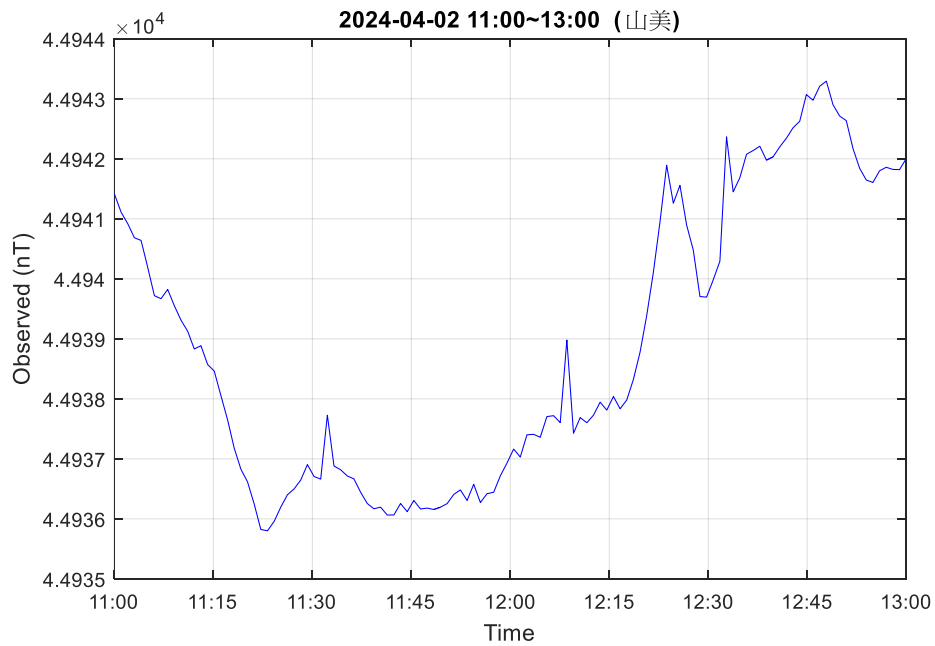
站名：山美(阿里山站)

聯絡人：部舞咖啡莊園－溫春梅 (0922-199-288)

位置及相關描述：山美休閒農場下方茶園一隅

觀測時間：2024/04/02 11:00~13:00

觀測資料(中心點)：東經 120.655010、北緯 23.406356

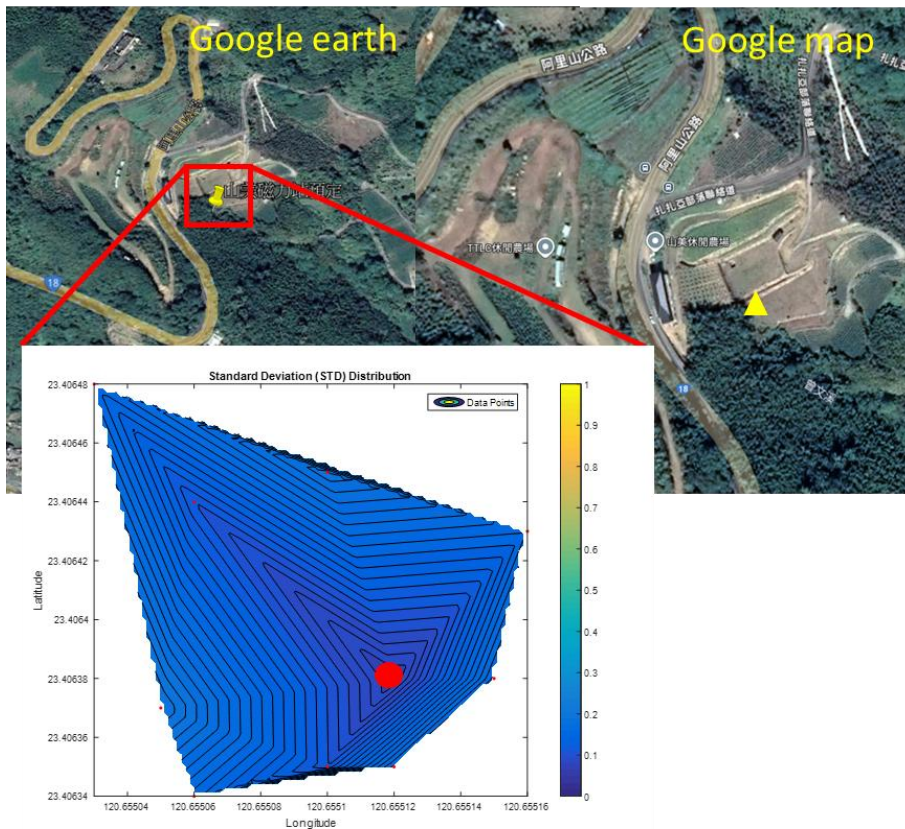


圖一、山美站預計站址之連續紀錄

觀測資料(周圍點)：

表一、山美磁力站觀測資料表

Num.	Lon.	Lat.	Elev.	Time	obsv.1.	obsv.2.	obsv.3.	obsv.4.	obsv.5.
1	120.66	23.406	1073		44939	44939.02	44938.96	44938.94	44938.99
					44938.65	44939.04	44939	44939.3	44939.13
2	120.66	23.406	1074		44937.26	44937.08	44937.2	44937.17	44937.11
					44937.59	44937.36	44937.4	44937.51	44937.56
3	120.66	23.406	1074		44933.41	44933.26	44933.28	44933.28	44933.22
					44933.21	44933.14	44933.16	44933.22	44933.29
4	120.66	23.406	1074		44934.8	44935.71	44935.21	44935.16	44935.13
					44935.12	44935.11	44935.26	44935.03	44934.92
5	120.66	23.406	1074		44939.84	44940.14	44939.78	44939.84	44939.84
					44939.94	44939.92	44940.03	44940	44940.08
6	120.66	23.406	1074		44940.01	44939.93	44939.98	44940.04	44940.09
					44940.08	44940.18	44940.26	44940.32	44940.32
7	120.66	23.406	1074		44930.29	44930	44929.99	44930.02	44929.98
					44929.89	44929.96	44930.01	44929.93	44929.85
8	120.66	23.406	1074		44913.62	44913.31	44913.39	44913.14	44913.44
					44913.52	44913.7	44913.52	44913.58	44913.38
9	120.66	23.406	1075		44928.5	44928.37	44928.55	44928.49	44928.33
					44928.57	44928.1	44928.53	44928.89	44928.62
10	120.66	23.406	1075		44940.27	44939.82	44939.78	44939.96	44939.87
					44940.04	44940.03	44940.05	44940.15	44940.23



圖二、山美磁力站測量位置及各個觀測點的標準差作圖

現場照片：



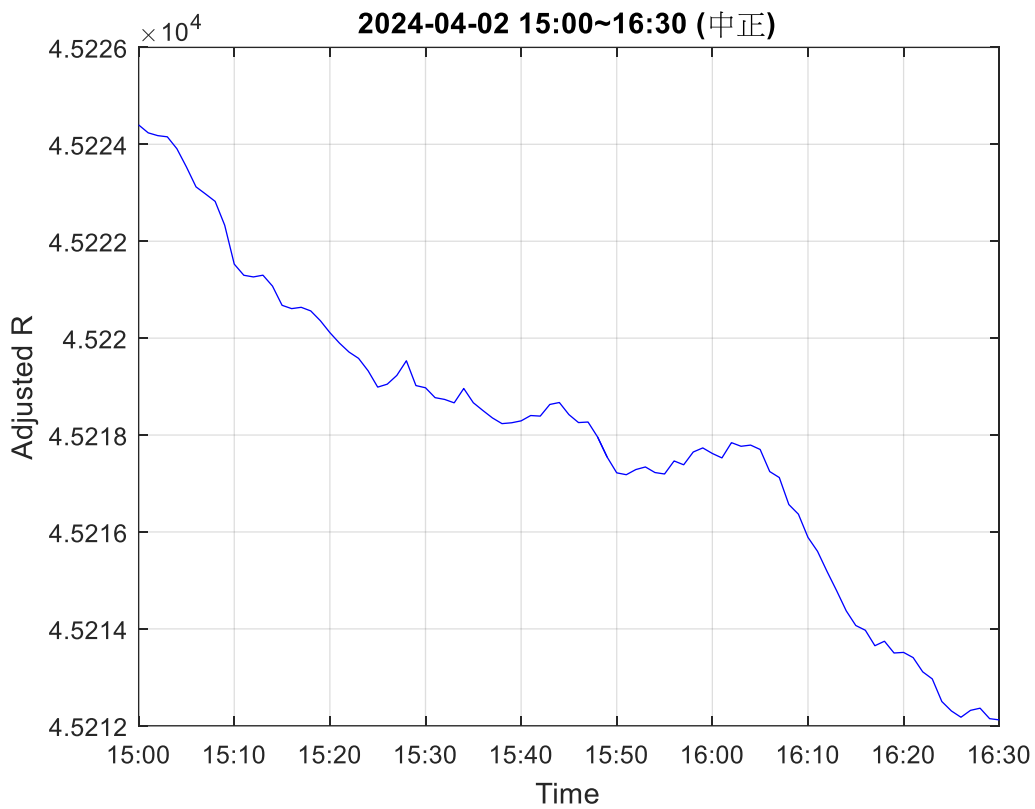
站名：中正大學

聯絡人：顏宏元教授(中央大學)、溫士忠副教授(中正大學)

位置及相關描述：原 2005 年中正大學三分量磁力站站址

觀測時間：2024/04/02 15:00~16:30

觀測資料(中心點)：東經 120.47878、北緯 23.56633 (約略、因觀測點在站房內)

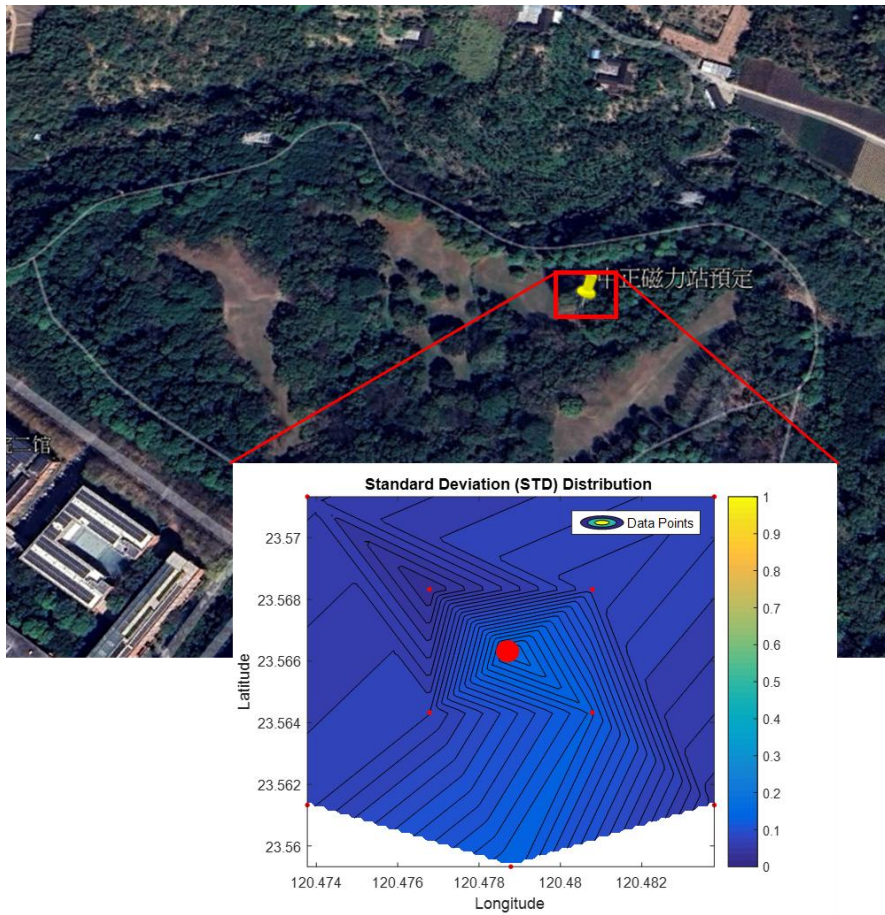


圖三、中正大學站預計站址之連續紀錄

觀測資料(周圍點)

表二、中正大學磁力站觀測資料表

Num.	Lon.	Lat.	Elev.	Time	obsv.1.	obsv.2.	obsv.3.	obsv.4.	obsv.5.
1	120.48	23.566	101		45218.76	45219.31	45219.3	45219.26	45219.24
					45219.15	45219.25	45219.26	45219.27	45219.31
2	120.48	23.568			45224.08	45224.08	45224.06	45224.03	45224.04
					45223.99	45224.04	45224.02	45224.06	45224.08
3	120.48	23.568			45223.68	45223.72	45223.8	45223.88	45223.87
					45223.88	45223.92	45223.93	45223.91	45223.82
4	120.48	23.564			45212.1	45212.06	45212.05	45212.05	45212.07
					45212.05	45212.18	45212.21	45212.17	45212.24
5	120.48	23.564			45214.96	45214.95	45214.9	45214.83	45214.7
					45214.75	45214.7	45214.7	45214.63	45214.58
6	120.48	23.559			45205.37	45205.48	45205.51	45205.46	45205.54
					45205.4	45205.36	45205.63	45205.22	45205.19
7	120.47	23.571			45226.7	45226.75	45226.74	45226.67	45226.64
					45226.66	45226.63	45226.72	45226.74	45226.8
8	120.47	23.561			45212.4	45212.51	45212.51	45212.46	45212.52
					45212.53	45212.55	45212.43	45212.39	45212.33
9	120.48	23.571			45225.15	45225.08	45225	45225.05	45225.09
					45225.15	45225.25	45225.19	45225.2	45225.21
10	120.48	23.561			45218.3	45218.38	45218.41	45218.43	45218.4
					45218.39	45218.45	45218.47	45218.42	45218.42



圖四、中正大學磁力站測量位置及各個觀測點的標準差作圖

現場照片：





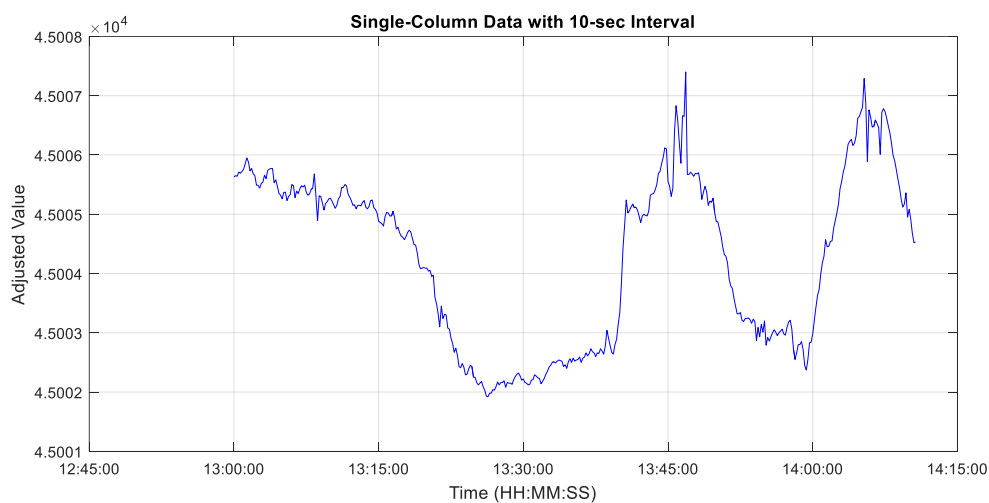
站名：光復站

聯絡人：光復外役監獄 (有聯繫人的 line)

位置及相關描述：光復外役監獄南邊農場放牧區

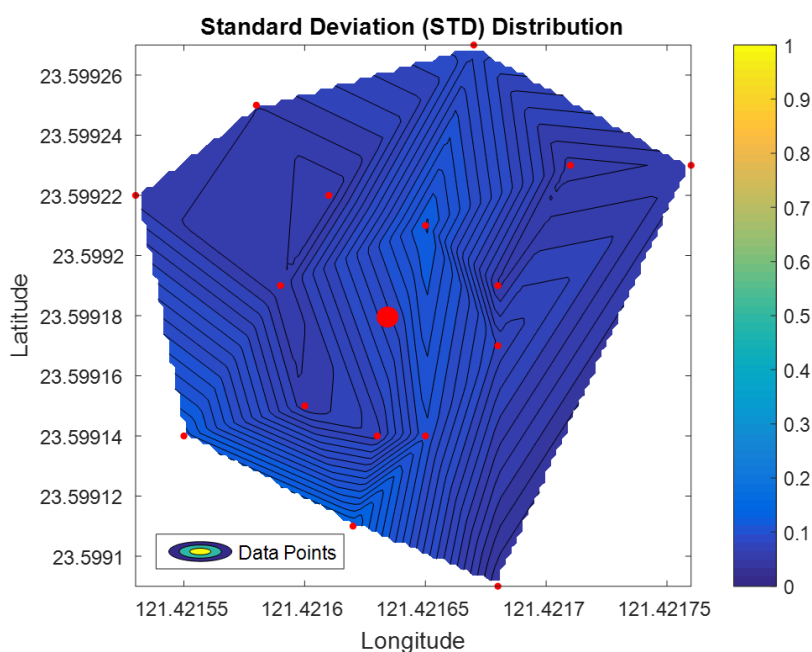
觀測時間：2025/06/25 13:00 ~ 14:13

觀測資料(中心點)：TWD97: (293030.323, 2610698.315) (121.421637, 23.599103),
H=311.692m



圖五、光復站預計站址之連續紀錄

觀測資料(周圍點)：



圖六、中正大學磁力站測量位置及各個觀測點的標準差作圖

現場照片：





114 年地震前兆觀測作業與分析技術相關研究

子計畫三

以深度學習影像處理為基礎之大地變形地震前兆研究

江振宇 林彥廷 彭敏鳳

國立臺北大學

摘要

在上半年的計劃中，我們 GNSS 測站為中心劃定同心圓區域，旨在識別與地震預測相關的測站組合。模型輸入包含中心站及鄰近最多三站之 GNSS 位移 與目標區域過去地震統計，利用 MinGRU 模型預測隔日地震統計或規模 ≥ 4 且深度 < 30 公里淺層地震機率。針對花蓮地區訓練大量模型後發現，僅部分測站（單獨或特定組合）對預測特定小規模極淺層地震（如規模 1-2、深度 5-30 公里）具有限能力，且加入兩個鄰近站能提升精確度。然而，所有模型均無法有效預測規模四以上大地震。研究結論確認 GNSS 數據對特定小型地震具有部分預測的潛力，且對預測大規模地震的具有侷限性。因此在下半年的計劃，我們針對規模 > 5.5 及深度 < 30 公里的地震，以花蓮地區所有的測站及週邊地震事件設計了兩個實驗，再次確認輸入的參數及序列長度、預測目標區間對模型影響。結果顯示，在預估區間為一天時，即使是訓練集，模型也未能有預測能力。而同時使用兩種輸入參數則只在 90-365 天的區間有影響。放寬輸入長度和預估的確對模型訓練有幫助，但學習到的參數，並未具有普遍性，因此在測試集上的仍未有預估能力。

中文關鍵詞：大地變形、地震規模、深度學習

Abstract

In the first half of the year's plan, we designated concentric circular regions centered on GNSS stations to identify combinations of stations relevant to earthquake prediction. The model inputs included GNSS displacements from the central station and up to three nearby stations, along with historical earthquake statistics for the target region. A MinGRU model was used to predict the next day's earthquake statistics or the probability of shallow earthquakes with magnitude ≥ 4 and depth < 30 km. After training numerous models for the Hualien region, we found that only certain stations (either individually or in specific combinations) exhibited limited predictive capability for particular small-scale, extremely shallow earthquakes (e.g., magnitude 1–2, depth 5–30 km), and incorporating

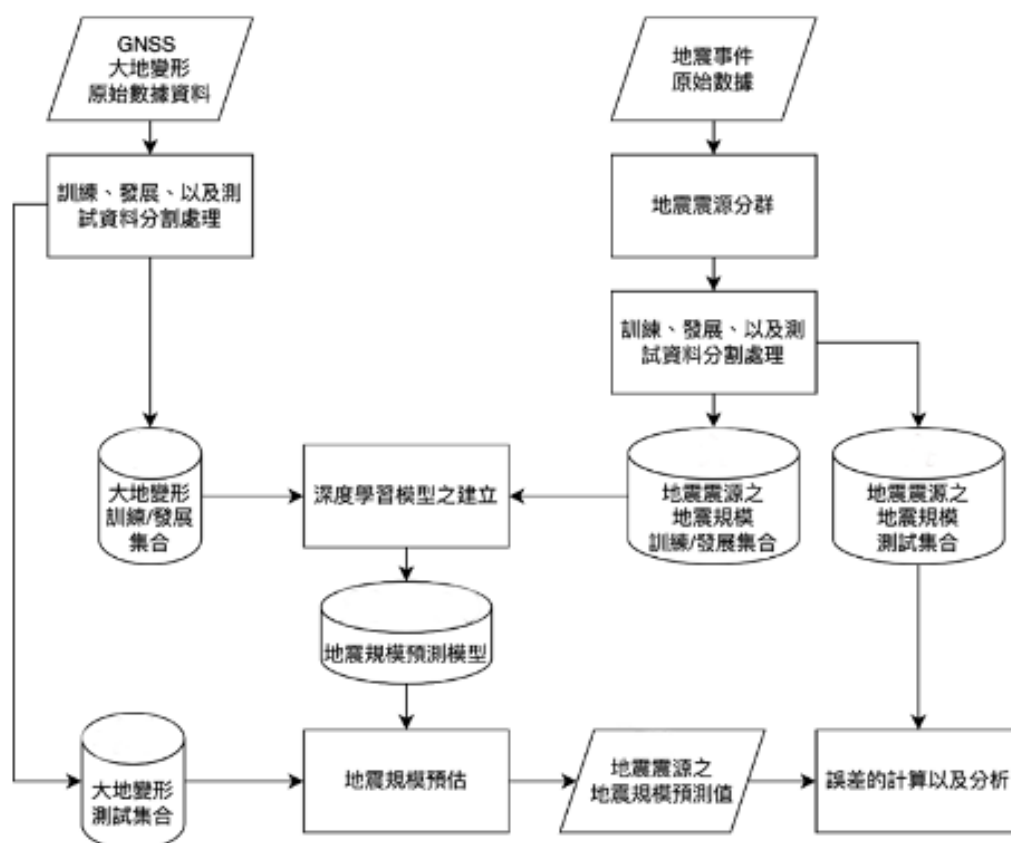
two nearby stations improved accuracy. However, all models were unable to effectively predict large earthquakes with magnitude ≥ 4 . The research conclusion confirmed that GNSS data has some predictive potential for specific small earthquakes but has limitations in predicting large-scale earthquakes. Therefore, in the second half of the year's plan, focusing on earthquakes with magnitude > 5.5 and depth < 30 km, we designed two experiments using all GNSS stations in the Hualien region and surrounding seismic events to further examine the impact of input parameters, sequence length, and prediction time window on the model. The results showed that when the prediction window was one day, the model lacked predictive ability even on the training set. Simultaneously using both types of input parameters only had an effect in the 90–365-day window. Relaxing the input length and prediction window did help model training, but the learned parameters lacked generalizability, so the model still showed no predictive ability on the test set.

Keywords: deformation, earthquake magnitude, deep learning

壹、前言

大型地震事件可視為發生在時間序列上的離散的事件，以機率統計的角度來講，若將大型地震發生與否當作二元決策的問題，在時間序列上，大型地震的機率很低，而沒有大型地震機率很高，因此大型地震是非常稀疏的事件，所以如果使用機器學習裡面的分類問題來解題，會有目標（大型地震）數量稀疏的問題。因此本研究團隊於先前研究「臺灣地區 113 年地震前兆監測資料彙整及分析 - 以深度學習及訊號處理為基礎的地震前兆研究」將原本的分類問題轉化為回歸問題，由 GNSS 所提供之「大地變形」參數來預估包含震源地理資訊之地震規模數值，採取深度學習 (deep learning) 方法，儘量避免人工設計特徵參數，直接建立一以深度學習模型為基礎之地震規模預測模型，以資料驅動 (data-driven) 方法，讓模型由輸入的原始數據 (raw data) 來進行特徵參數抽取 (feature extraction)，並預估目標震源之地震規模數值。

該研究的架構如圖一所示，我們將「GNSS 大地變形原始數據資料」以及「地震事件原始數據」都分成訓練、發展以及測試集合，由於地震事件在每日可能不只發生一次且會在不同經度、緯度、深度發生，為了要能夠讓以「日」為單位紀錄的「GNSS 大地變形原始數據資料」能與地震事件以「日」為單位同步處理，我們將依據地震事件的經度以及緯度將臺灣地區分群成 16 個區域，並且記錄落於這些 16 個區域內地震規模以及次數，接下來我們建立一「深度學習模型」由以「日」為單位的 630 個 GNSS 測站提供之大地變形資料來預估 16 個區域的地震規模或是次數。



圖一：研究架構圖

該研究將地震事件的預估使用以下數學式表示：

$$z_t = f(x_{t-T:t-1}, y_{t-T:t-1}) \quad (1.1)$$

其中 z_t 代表第 t 日待預估的地震事件， $x_{t-T:t-1}$ 代表地震事件發生前 T 日（第 $t-T$ 日）至前一日（第 $t-1$ 日）的 GNSS 觀測資料， $y_{t-T:t-1}$ 代表地震事件發生前 T 日（第 $t-T$ 日）至前一日（第 $t-1$ 日）已經發生的地震事件觀測資料，我們為了模擬已發生地震事件對於未來地震事件以及 GNSS 量測參數的影響，所以將 $y_{t-T:t-1}$ 作為輸入參數，而 $f(\cdot)$ 代表一個由觀測資料以及已發生地震事件對應待預估地震事件的映射函數 (mapping function)，此映射函數於本研究中將採用深度學習模型來實現，比如使用以 Transformer (Vaswani et al., 2023) 或是 MinGRU (Feng et al., 2024) 的 RNN 為基礎的類神經網路架構。

基於數學式 (1.1) 的想法，該研究分別在執行中的上半年以及後半年提出了兩種研究方法：1) 研究方法-1：端至端資料驅動法、以及 2) 研究方法-2：引入知識之機器學習法。以下為此兩種方法的概述：

1) 研究方法-1：端至端資料驅動法

- 考量臺灣全部的 GNSS 資料對應到臺灣 16 個量化區域的最大地震規模。

- 輸出：16 個量化區域未來一天、一週、兩週、三週、以及一個月內最大的地震規模，此 16 個量化區域是由 1999 年至 2023 年之有感地震事件序列之經度以及緯度做 *k*-means clustering (Lloyd, 1982) 得到。
- 輸入：包含過去 10、30、60、或 90 天的 GNSS 測站之 dE、dN、以及 dU 資料，以及 16 個量化區域已發生過的地震，GNSS 測站共有 630 個，而過去地震每日都有 16 個區域的最大地震規模值。
- 輸入-輸出對應模型：使用 Transformer 為基礎的回歸模型。
- 優點：僅需要將輸入以及輸出資料對應好，直接採用成熟的 deep learning 模型進行輸入-輸出的對應，Transformer 先天在 deep learning 模型的參數裡可以表達輸入序列對於輸出序列的對齊關係。
- 缺點：Transformer 參數量過大，容易造成過度擬和或是參數學習失敗，且此方法未加入任何地理資訊知識，地震包含非臺灣本島以及較不會造成傷害的深層地震，在輸入中也包含了非本島的 GNSS 測站，此外，所預估的地震僅留下每日最大的地震規模，因此會損失部分資訊。而實驗結果地震規模預估之均方根誤差過大，且難以對應到地震是否會發生的機率。

2) 研究方法-2：引入知識之機器學習法

- 將臺灣本島各縣市拆分成面積相似的 25 個區域，盡可能保留原有的縣市界、各區域面積相近、山區和平原分開、且少地震區獨立分區。目前先僅討論地震次數最多的「花東 1」地區討論。
- 輸出：「花東 1」地區每日的地震事件，此事件以 50 個維度表示，由 5 類的地震深度以及 10 種規模組合而成，其中深度分成五類：<5 公里、<10 公里、<30 公里、<70 公里、<300 公里；而地震規模則是將原本浮點數的規模以向下取整的方式分成 0~9 十類。每一個維度就是記錄某一個深度（五類）以及某一種地震規模（十類）在當天發生的次數
- 輸入：「花東 1」區域內地震當日之過去 364 天的 GNSS 測站之 dE、dN、以及 dU 資料，以及「花東 1」區域內地震當日之過去 364 天已發生過的地震事件。
- 輸入-輸出對應模型：使用 MinGRU 為基礎的回歸模型。
- 優點：加入地理資訊知識，移除較無關係的深層地震以及外島資料。MinGRU 模型的參數量不會過大，且可以對抗模型更新時可能遇到的梯度消失問題，另外，所預估的地震目標包含不同深度以及規模對應的數目，充分保留地震事件的資訊，所預估的地震事件，可以轉換成發生某種地震規模的精確度、召回率、以及準確率。
- 缺點：此方法必須在不同區域（25 個區域）使用不同資料集合進行實驗，且目前僅考慮「花東 1」區域內的地震事件以及 GNSS 資料。

本次研究主要基於前項研究方法-2，將針對一些考慮不足的部分稍做改進，並調整實驗的方向。

貳、研究目的及意義

根據先前計畫「研究方法-2：引入知識之機器學習法」的結果，在模型的準確率方面，比純粹使用歷史統計資訊猜測地震發生的準確率可以增加 5-10% 左右的正確性，代表有部分地震是可以由 GNSS 輸入參數來預估，雖然改善的比率不高，但至少是一個使用機器學習或是深度學習技術預測地震事件的開端。

先前計劃是將臺灣本島分成 25 個區域，以區域內的測站和地震事件進行分析。然而，一次使用所有測站進行預測，無法確定哪些測站或其組合對地震事件有相關性。這種不規則且複雜的區域劃分方式也需要改善。本研究將提出改進資料選擇和訓練方法，以解決上述問題。

參、研究方法-3

3.1 問題描述

在之前方法 (研究方法-1/2) 中，我們是使用一個區域中全部的測站和地震資料來進行預估，在數學式上，未特別引入輸入測站組合和目標地震區域這兩個變量，因此我們需要將地震的預估函數稍做修改：

$$\hat{z}_t^A = f_\theta(x_{t-T:t-1}^J, y_{t-T:t-1}^A) \quad (3.1)$$

其中 J 代表某一個 GNSS 測站組合， A 代表地震事件的區域，我們以下標來表示時間或是時間段，單獨的 t 表示某一特定的日期， $t-T:t-1$ 代表的 $t-T$ 到 $t-1$ (含) 日的時間段； x 代表 GNSS 測站的觀測資料、 y 代表已發生過地震事件、 \hat{z}_t^A 代表模型預估的地震事件， $f_\theta(\cdot)$ 代表一個由 GNSS 測站觀測資料以及已發生地震事件對應待預估地震事件的映射函數 (mapping function)，此映射函數於本研究中將採用參數為 θ 的深度學習模型來實現。

根據以上的數學符號定義，以下對於式 (3.1) 裡面的各項變數做更詳細的說明：

- $x_{t-T:t-1}^J \in \mathcal{R}^{T \times |J| \times 3}$ ：是對應到此 GNSS 測站組合 J (其中有 $|J|$ 個測站)，第 $t-1$ 至 $t-T$ 日的 GNSS 觀測資料，其中每一個測站貢獻 dN 、 dE 、 dU 的三個變量，它們分別代表地殼每一日南北向、東西向、以及高度與參考點座標的位移數值，且共有 T 日的資料，所以 $x_{t-T:t-1}^J \in \mathcal{R}^{T \times |J| \times 3}$ 。
- $y_{t-T:t-1}^A \in \mathbb{N}^{T \times 5 \times 10}$ ：從第 $t-T$ 到 $t-1$ 日間某區域 A 內的已觀測的地震事件。地震事件使用 5×10 的矩陣表示。將深度分成五類：0: 0~5 公里、1: 5~10 公里、2: 10~30 公里、3: 30~70 公里、4: 70~300 公里；而地震規模則是將原本浮點數的規模以向下取整的方式分成 0~9 十類，兩者形成 50 類事件，

每一個事件就是記錄某一個深度（五類）以及某一種地震規模（十類）在當天發生的次數。

- \hat{z}_t^A ：「預測地震事件」，也就是函數 $f_\theta(\cdot)$ 所預估的第 t 日的某區域 A 內的地震事件，依「目標地震事件」 z_t^A 不同有兩種定義。

目標地震事件 z_t^A 是一個向量，本研究將實驗以下兩種地震事件的向量表示方法：

1. $z_t^A = y_t^A \in \mathbb{N}^{5 \times 10}$ ：代表第 t 日在區域 A 內所發生的地震事件統計，與輸入的過去地震事件相同的表示方式。
2. $z_t^A = y_t^A \in \{0,1\}$ ：代表第 t 日有沒有發生大於等於規模 4 且深度 < 30 公里的地震事件，也就是 $z_t^A = 1$ 代表有發生，而 $z_t^A = 0$ 代表沒有發生。

如果 $z_t^A \in \mathbb{N}^{10 \times 5}$ ，則我們採用 MMSE 作為模型函數 $f_\theta(\cdot)$ 的訓練條件，每一日（第 t 日）所貢獻的損失含數值為各個類別的地震數量的均方差：

$$MSE_t(x_{t-T:t-1}^J, y_{t-T:t-1}^A, z_t^A | \theta) = \frac{1}{MD} \sum_{d=0}^{D-1} \sum_{m=0}^{M-1} \|z_t^A(d, m) - \hat{z}_t^A(d, m)\|^2 \quad (3.2)$$

我們將此類模型稱為迴歸模型，其中 d 代表不同地震深度分類， m 代表不同地震規模。

但若 $z_t^A \in \{0,1\}$ ，則我們採用 cross entropy (CE) 作為模型函數 $f_\theta(\cdot)$ 其參數的訓練條件，每一日（第 t 日）所貢獻的損失含數值為是否發生規模 ≥ 4 且深度 < 30 公里地震事件的 binary cross entropy loss：

$$CE_t(x_{t-T:t-1}^J, y_{t-T:t-1}^A, z_t^A | \theta) = -z_t^A \log(\hat{z}_t^A) - (1 - z_t^A) \log(1 - \hat{z}_t^A) \quad (3.3)$$

我們又將此類模型稱為機率模型。

在定義好要訓練的模型類型後，下一小節我們將說明如何去設定模型所需測站組合 J 及目標地震事件區域 A 。

3.2 輸入測站組合及目標地震區域的選取

本研究旨在找到能用來進行有效地震預估的不同的 GNSS 測站組合。但由於測站總數龐大，要完整地探討所有的組合是不太可行的。因此，我們先假設地震事件與其鄰近的測站具有較大的相關性，劃定地震周圍的一個區域，選用其中的測站的不同組合去預估地震事件。

不同於之前使用縣市界來劃定一個特定的區域，本次使用較為簡便處理的方式。對於所有的測站 Ω ，我們會以某一 GNSS 測站 $c \in \Omega$ 為做為中心測站，取一個半徑 20 公里及半徑 40 公里的圓 C_1 及 C_2 ，再取 C_1 內的地震事件做為預測

3.3 實驗設計

本次實驗的首要目標是找出那一種測站的組合對地震預測較有幫助。基於前次的結果有部分的地震是可以被預測的，因此我們這次仍以花蓮地區的測站做為主要的實驗對象，同時去排除目前已停止使用的測站資料，我們將這個集合設為 H 。即使已排除了一些測站，測站組合數量 $|2^{S_c}|$ 仍非常龐大，因此我們只選擇 2^{S_c} 中基數不大於 3 的組合，而每個模型的整體建立流程可簡略表示成以下的虛擬碼：

```

For all  $c \in H \subset \Omega$  // 對於每一個做為中心測站 GNSS 測站  $c$ 
    // 準備一些範圍
    Prepare  $C_1, C_2, S^c, 2^{S_c}$ 
    For all  $I \in 2^{S_c}, |I| \leq 3$  // 對於所有鄰近測站組合  $I$ 
        // 設定輸入測站
        Set  $J = \{c\} \cup I$ 
        // 預測規模深度之次數
         $\theta_{c,J}^* = \arg \min_{\theta} \sum_t MSE_t(x_{t-T:t-1}^J, y_{t-T:t-1}^{C_1}, z_t^{C_1} | \theta)$ 
        or
        // 預測是否發生規模 4 以上淺層地震
         $\theta_{c,J}^* = \arg \min_{\theta} \sum_t BCE_t(x_{t-T:t-1}^J, y_{t-T:t-1}^{A_1}, z_t^{A_1} | \theta)$ 
    end
end

```

由於訓練資料量並沒有非常大，因此在模型架構方面，不使用 Transformer 而是使用結構較為簡單的 MinGRU 架構 [1]，且總層數較少。模型的架構如下 表格 1，其輸出為長度為 365 的時間序列，每一個時間步都含有 50 類地震的預測數量，而我們只取最後一個時間步 ($t=364$) 做為預測結果。

表格 1：模型架構，其中 B 是 batch size、K 是區域測站數量

層	模組名稱	輸入形狀	輸出形狀

1	Linear	(B, 365, $K*3+5*10$)	(B, 365, 128)
2	LeakyReLU	(B, 365, 128)	(B, 365, 128)
3	LayerNorm	(B, 365, 128)	(B, 365, 128)
4	MinGRU	(B, 365, 128)	(B, 365, 128)
5	LayerNorm	(B, 365, 128)	(B, 365, 128)
6	MinGRU	(B, 365, 128)	(B, 365, 128)
7	LayerNorm	(B, 365, 128)	(B, 365, 128)
8	MinGRU	(B, 365, 128)	(B, 365, 128)
9	Linear	(B, 365, 128)	(B, 365, 50) or (B, 365, 1)
10	Reshape	(B, 365, 50) or (B, 365,1)	(B, 365, 5, 10) or (B, 365, 1)
11	SoftPlus or Sigmoid	(B, 365, 5, 10) or (B, 365,1)	(B, 365, 5, 10) or (B, 365, 1)

中心測站集合 H 中，包含了花蓮縣目前仍存活的 25 測站，我們並以此中心測站集合總共訓練了 26,335 個模型。而這 25 個測站的簡單資訊如表格 2。

表格 2：中心測站資訊

測站名	鄰近測站數	地震事件	規模 4 以上極淺層地震	模型數量
DCHU	28	11039	56	3683
FUDN	26	11678	63	2952
JPEI	17	12171	64	834
SPAO	13	14710	77	378
KNKO	17	15481	66	834
YUL1	27	8201	48	3304
JSU2	17	12682	53	834
SOFN	14	75337	386	470
SHUL	15	41678	227	576
YENL	14	82369	412	470
BLOW	13	75883	417	378

WARO	14	33161	172	470
JPIN	25	10620	55	2626
NDHU	15	72224	358	576
SLIN	14	33294	172	470
HUAP	9	40983	235	130
SCHN	12	92695	484	299
HUAL	15	97804	501	576
SICH	12	92698	484	299
FONB	17	21097	110	834
CHUN	18	11215	48	988
JULI	25	7906	44	2626
DNFU	15	24172	140	576
FLNM	15	31389	177	576
TUNM	15	67668	339	576

訓練時，所有的資料集都會分成兩個子集：民國 89-110 年的訓練資料集和民國 111-112 年的測試資料集。對於每個模型均在訓練集上，以 batch size =128 訓練完整的 5 個 epochs。而測試資料集則用於模型的評估。

模型的名稱均以輸入的測站組命名，測站組合的名稱均以均以中心測站做為開頭用底線「_」來連接鄰近測站，如中心測站 HUAL 加上鄰近測站 YENL 及 WARO，其名稱會命名為 HUAL_YENL_WARO。

模型評估方面，除了訓練時所用的均方差之外，還會另外計算各類別預測地震發生的精確度 (precision)、召回率 (recall) 及準確率 (accuracy)。這部分是以數值 1 為有無地震發生的門檻值，並以此計算下面的四組統計量：

$$N_{TP}(d, m) = \sum_t [z_t(d, m) \geq 1] [\hat{z}_t(d, m) \geq 1] \quad (11)$$

$$N_{FN}(d, m) = \sum_t [z_t(d, m) \geq 1] [\hat{z}_t(d, m) < 1] \quad (12)$$

$$N_{FP}(d, m) = \sum_t [z_t(d, m) < 1] [\hat{z}_t(d, m) \geq 1] \quad (13)$$

$$N_{TN}(d, m) = \sum_t [z_t(d, m) < 1] [\hat{z}_t(d, m) < 1] \quad (14)$$

分別代表正確預估有地震 (TP, true positive)、漏報(FN, false negative)、誤報(FP, false positive)、正確預估無地震 (TN, true negative)的數量，其中[.] 為 Iverson bracket。從上面四組統計量，可以分別計算精確度 P、召回率 R、準確率 A：

$$P(d, m) = \frac{N_{TP}(d, m)}{N_{TP}(d, m) + N_{FP}(d, m)} \quad (15)$$

$$R(d, m) = \frac{N_{TP}(d, m)}{N_{TP}(d, m) + N_{FN}(d, m)} \quad (16)$$

$$A(d, m) = \frac{N_{TP}(d, m) + N_{TN}(d, m)}{N_{TP}(d, m) + N_{FN}(d, m) + N_{FP}(d, m) + N_{TN}(d, m)} \quad (17)$$

用以評估模型偵測地震的性能（這三項指標均為越高越好）。

3.4 具體成果

觀察訓練完的 26,335 個回歸模型的結果，我們發現大部分的模型是不具有預測能力的，即所有的輸出均預測為無地震，僅有少量的模型能夠產生有地震的預測。由於模型的數據眾多，我們僅能在此就一些觀察結果嘗試去回答下面幾個問題。

3.4.1 單一測站是否有預測地震的能力？

在全部 25 個測站中，只有 13 個測站有預測周邊地震的能力。能夠預測的地震類型只分布在 $[d, m] = [1, 1]$ (深度 5-10 公里，規模 1)、 $[2, 1]$ (深度 10-30 公里，規模 1)、 $[2, 2]$ (10-30 公里，規模 2) 三組。其結果列在表格 3。

表格 3：單一測站的預測能力

中心測站	地震事件類別 [深度類別 d , 規模 m]	對應精確度 P
WARO	[2, 1]	0.45
NDHU	[1, 1], [2, 1], [2, 2]	0.45, 0.32, 0.55
FLNM	[2, 1]	0.44
TUNM	[1, 1], [2, 1], [2, 2]	0.39, 0.31, 0.52
BLOW	[1, 1], [2, 1], [2, 2]	0.38, 0.34, 0.55
SHUL	[1, 1], [2, 1]	0.00, 0.48
SICH	[1, 1], [2, 1], [2, 2]	0.38, 0.38, 0.60
SCHN	[1, 1], [2, 1], [2, 2]	0.39, 0.37, 0.60
YENL	[1, 1], [2, 1], [2, 2]	0.39, 0.36, 0.57
HUAL	[1, 1], [2, 1], [2, 2]	0.39, 0.40, 0.63

SOFN	[1, 1], [2, 1], [2, 2]	0.46, 0.33, 0.57
HUAP	[1, 1], [2, 1]	0.33, 0.41
SLIN	[2, 1]	0.46

3.4.2 加入鄰近測站後，對精確度 P、準確率 A 是否一定有幫助？

我們檢查了前項問題中所列的 13 個中心測站的類別 [2,1] 的精確度數值，並列出精確度下降和上升的組合數。在精確度上，以加入兩個鄰近測站的組合最能夠提升精確度，提升幅度約為 2-7 個百分點。加入三個測站似乎並無明顯的數值提升甚至還會下降。結果放在表格 4。

表格 4：鄰近測站對精確度和準確度的影響

中心測站	提升	降低	單測站(P, A)	表現前三名的模型 (P, A, name)
WARO	85	385	0.45, 0.78	(0.50, 0.79, WARO_FLNM_FONB), (0.48, 0.79, WARO_YENL_LSAN), (0.48, 0.79, WARO_YENL_JPEI)
NDHU	105	471	0.32, 0.36	(0.38, 0.55, NDHU_FLNM_SLIN), (0.38, 0.54, NDHU_FLNM_WARO), (0.38, 0.54, NDHU_FLNM_TUNM)
FLNM	111	465	0.44, 0.80	(0.58, 0.82, FLNM_FONB_KNKO), (0.56, 0.82, FLNM_SHUL_FONB), (0.56, 0.82, FLNM_NDHU_FONB)
TUNM	100	476	0.31, 0.35	(0.36, 0.50, TUNM_FLNM_SLIN), (0.35, 0.49, TUNM_FLNM_WARO), (0.35, 0.47, TUNM_NDHU_SLIN)
BLOW	39	339	0.34, 0.35	(0.36, 0.42, BLOW_NDHU_HNSN), (0.36, 0.41, BLOW_YENL_HNSN), (0.36, 0.41, BLOW_SCHN_HNSN)
SHUL	61	515	0.48, 0.75	(0.54, 0.78, SHUL_SICH_TUNM), (0.54, 0.78, SHUL_SICH_SLIN),

				(0.54, 0.78, SHUL_SCHN_NDHU)
SICH	46	253	0.38, 0.38	(0.40, 0.44, SICH_NDHU_TUNM), (0.39, 0.43, SICH_SHUL_TUNM), (0.39, 0.43, SICH_NDHU_HUAN)
SCHN	47	252	0.37, 0.38	(0.40, 0.44, SCHN_NDHU_HUAN), (0.39, 0.44, SCHN_NDHU_TUNM), (0.39, 0.43, SCHN_SOFN_TUNM)
YENL	74	396	0.36, 0.38	(0.40, 0.50, YENL_FLNM_WARO), (0.40, 0.50, YENL_FLNM_SLIN), (0.40, 0.49, YENL_NDHU_SLIN)
HUAL	105	471	0.40, 0.41	(0.42, 0.47, HUAL_WARO_SLIN), (0.42, 0.47, HUAL_NDHU_TUNM), (0.42, 0.47, HUAL_FLNM_WARO)
SOFN	82	388	0.33, 0.37	(0.39, 0.53, SOFN_NDHU_SLIN), (0.39, 0.53, SOFN_FLNM_TUNM), (0.38, 0.53, SOFN_WARO_SLIN)
HUAP	16	114	0.41, 0.72	(0.48, 0.76, HUAP_SCHN_SUA2), (0.47, 0.76, HUAP_SICH_SUA2), (0.47, 0.76, HUAP_HUAL_SUA2)
SLIN	62	408	0.46, 0.78	(0.48, 0.79, SLIN_YENL_WARO), (0.48, 0.79, SLIN_YENL_TUNM), (0.48, 0.79, SLIN_YENL_JPEI)

3.4.3 中心測站本身無預測地震能力，但加入其他測站後具有預測能力的組合。

在所有組合中只有 FUDN_WULU_DCHU_TAPE 一組會將類別 [0, 4] 預估為有地震，但其精確度為 0。其他的組合均為落以 DNFU 中心的組合，共有 66 組，其中精確度最高的為 DNFU_JSU2_CHUN 這一組，其在類別 [2, 1] 的精確度為 0.29。

3.4.4 模型是否能夠預測規模 4 以上地震？

在所有組合中只有一個模型會去預測規模 3 以上的地震，但其在測試集上的精確度為 0.0，如表格 5。

表格 5：能夠預估規模 4 以上地震的模型

測站組合	地震事件類別	精確度
FUDN_WULU_DCHU_TAPE	[0, 4]	0.0

3.4.5 規模 4 以上、極淺層機率模型所遇到的狀況？

我們同樣也訓練了 26335 個使用 CE loss 的機率模型，但是與回歸模型的狀況不同，在這個條件下，所有的模型均不會輸出 0 以外的數值。由於回歸模型的也無法預測規模 3 以上的地震，從這個結果來推論，可能 GNSS 的資料和過去地震分布對於規模較大的地震預測並無幫助。

3.5 小結

本研究延續先前的研究方法-2，在資料的篩選上做了一些變動，改使用同心圓的方式設定目標區域。並且用大量不同的測站組合大量訓練模型，目的是找出對預測較有幫助的測站組合。下面列出一些觀察和討論：

- 無法保證模型一定能夠學到某種與地震相關的特徵。沒學到特徵的模型無法確定和地震事件完全無關。
- 區域劃分上，我們取半徑 40 公里和 20 公里的圓作為輸入和輸出的區域條件。我們無法確定將地震事件限縮在這樣一個區域中，是否是最好的做法。
- 模型只能夠預測規模較小、深度較淺的地震事件，且準確度和精確度並不高。
- 為了避免組合數過多，我們以一次 4 個測站的資料為上限，尚未使用更多測站的狀況進行實驗。
- 對於地震事件較少的區域，因資料偏壓，模型會選擇預估成無地震，有待使用其他方法去解決，比如使用 Balanced Cross Entropy (Cui, Y. Cui, Y., 2019) 或是 Focal Loss (Lin, T. Y., 2017)。

肆、研究方法-4

4.1 問題描述

觀測以及預估之時間區間尺度：上述的「研究方法-3」存在最大的問題是對物理量的定義範圍，如所觀測的 GNSS 資訊 $x_{t-T:t-1}^j$ 應考量更長的時間，因為大地變形的影響應該是至少 1 年以上 ($T \geq 365$) 的長期觀察；而函數 f_θ 所預估的地震事件，由於大地變形的能量累積是屬於長期影響，所以所造的大規模地震應該也是在某一個長時間區間內發生，也就是說不應該只預估未來一天，應該是預估未來的幾個月會發生強震的機率。

聚焦具傷害性的地震事件預測：在研究方法-3 之中，為了要能夠充分利用數據，所以我們將預測目標設定為下一天 50 個事件出現的次數，也就是 5 種不同深度範圍以及 10 種量化地震規模 (0 至 9)，然而這樣對於真正需要被關注的規模和深度無法被模型學習到其關係，因此本研究方法將針對規模為 5.5 以上且地震深度為 30 公里以內的淺層地震作為待預估事件。

以釋放能量觀點觀察：地震所造成的傷害會與單位時間內釋放的總能量有關，由目前的地震事件預測結果來看，模型比較能預測到規模小於 2 的地震事件，在一天內可以有多次小規模的地震，然而這些小規模的地震如果經過累積計算能量，應該也具有一些物理意義，如果能讓模型預估出未來某個時間區間的地震釋放量

量，可以和發生大規模傷害之地震事件做一些對應分析。如果在短時間內有很大的總能量，代表就越有可能或有高功率的大規模地震發生。

根據以上的問題討論，我們將地震預測模型的數學表示重新說明：

$$\hat{z}_{t,\tau}^A = f_{\theta}(x_{t-T:t-1}^J, y_{t-T:t-1}^A) \quad (4.1)$$

- $x_{t-T:t-1}^J \in \mathcal{R}^{T \times |J| \times 3}$ ：是對應到 GNSS 測站組合 J (其中有 $|J|$ 個測站)，第 $t-1$ 至 $t-T$ 日的 GNSS 觀測資料，其定義如同「研究方法-3」。而本方法考量的觀察時間長度為 1、2、3、5 年的四種設定，所以 $T=356, 712, 1068, 1780$ 。
- $y_{t-T:t-1}^A \in \mathbb{N}^{T \times 5 \times 10}$ ：從第 $t-T$ 到 $t-1$ 日間某區域 A 內的已觀測的地震事件，其定義如同「研究方法-3」， $T=356, 712, 1068, 1780$ 。
- $\hat{z}_{t,\tau}^A$ ：「預測地震事件」，也就是函數 $f_{\theta}(\cdot)$ 所預估的第 t 日至第 $t+\tau$ 日的某區域 A 內的地震事件，本研究將考量第 t 日後 3、6、12、24 個月內的地震事件，也就是 $\tau = 90, 180, 365, 730$ 稱為「預測地震事件時間區間」。依「目標地震事件」 $z_{t,\tau}^A$ 不同有三種定義：
 1. $\hat{z}_{t,\tau}^A \in \mathbb{N}^1$ ：代表於區域 A 內，第 t 日至第 $t+\tau$ 日之間出現規模大於 5.5 且深度淺於 30 公里的地震次數。
 2. $\hat{z}_{t,\tau}^A \in \{0,1\}$ ：代表於區域 A 內，第 t 日至第 $t+\tau$ 日之間，有沒有發生大於等於規模 5.5 且深度 < 30 公里的地震事件，也就是 $\hat{z}_{t,\tau}^A = 1$ 代表有發生，而 $\hat{z}_{t,\tau}^A = 0$ 代表沒有發生。
 3. $\hat{z}_{t,\tau}^A \in \mathbb{R}^1$ ：代表於區域 A 內，第 t 日至第 $t+\tau$ 日之間，地震所釋放的總能量。也就是：

$$\hat{z}_{t,\tau}^A = \log_{-}\{1 \sum_{i=1}^K 10^{(11.8 + 1.5M_i)}\} \quad (4.2)$$

4.2 模型訓練

如果 $\hat{z}_{t,\tau}^A \in \mathbb{N}^1$ 或是 $\hat{z}_{t,\tau}^A \in \mathbb{R}^1$ ，則我們採用如數學式 (3.2) 的 MMSE 作為模型函數 $f_{\theta}(\cdot)$ 的訓練條件：

$$MSE_t(x_{t-T:t-1}^J, y_{t-T:t-1}^A, z_{t,\tau}^A | \theta) = \frac{1}{MD} \sum_{d=0}^{D-1} \sum_{m=0}^{M-1} \|z_{t,\tau}^A(d, m) - \hat{z}_{t,\tau}^A(d, m)\|^2 \quad (4.3)$$

但若 $z_{t,\tau}^A \in \{0,1\}$ ，則我們採用如數學式 (3.3) 的 binary cross entropy (CE) 作為模型函數 $f_{\theta}(\cdot)$ 其參數的訓練條件：

$$CE_t(x_{t-T:t-1}^J, y_{t-T:t-1}^A, z_{t,\tau}^A | \theta) = -z_{t,\tau}^A \log(\hat{z}_{t,\tau}^A) - (1 - z_{t,\tau}^A) \log(1 - \hat{z}_{t,\tau}^A) \quad (4.4)$$

由於規模大於 5.5 且深度淺於 30 公里的地震事件 (正樣本) 出現機率很少，很容易造成 $\Pr(z_{t,\tau}^A = 0)$ 會大於 $\Pr(z_{t,\tau}^A = 1)$ 許多，會造成很明顯的偏壓，以至於模型只要讓 $f_{\theta}(\cdot)$ 學習輸出成 0 (負樣本，代表沒有目標地震) 就可以達到很低

的 BCE 值，會很容易讓函式只要預估沒有大規模地震即可，為了解決這個問題，本研究將引入 balanced cross entropy:

$$BCE_t(x_{t-T:t-1}^J, y_{t-T:t-1}^A, z_{t,\tau}^A | \theta) = -\alpha z_{t,\tau}^A \log(\hat{z}_{t,\tau}^A) - (1 - \alpha)(1 - z_{t,\tau}^A) \log(1 - \hat{z}_{t,\tau}^A) \quad (4.5)$$

其中 $\alpha \in (0,1)$ 為發生「規模大於 5.5 且深度小於 30 公里的地震事件」(以下稱目標事件)的權重，可以設定為「沒有發生目標事件」之於所有觀測地震事件區間次數的比例，因此數學式 (4.5) 會將「有發生目標事件」所貢獻的損失值放大，讓模型更專注於辨認出「有發生目標事件」的可能性。

雖然 balanced cross entropy 能降低類別不平衡的影響，但當負樣本數量非常大、而且有大量「一看就知道是 0」的簡單負樣本時，模型仍然會把大部分的學習力氣浪費在這些容易判斷的樣本上，真正困難、真正重要的「少數正樣本」和「接近要發生的負樣本」反而得不到足夠的梯度。為了讓模型自動把注意力集中在難樣本上，本研究也考慮採用 focal loss(原用於目標偵測的不平衡問題)。標準的 focal loss 形式為：

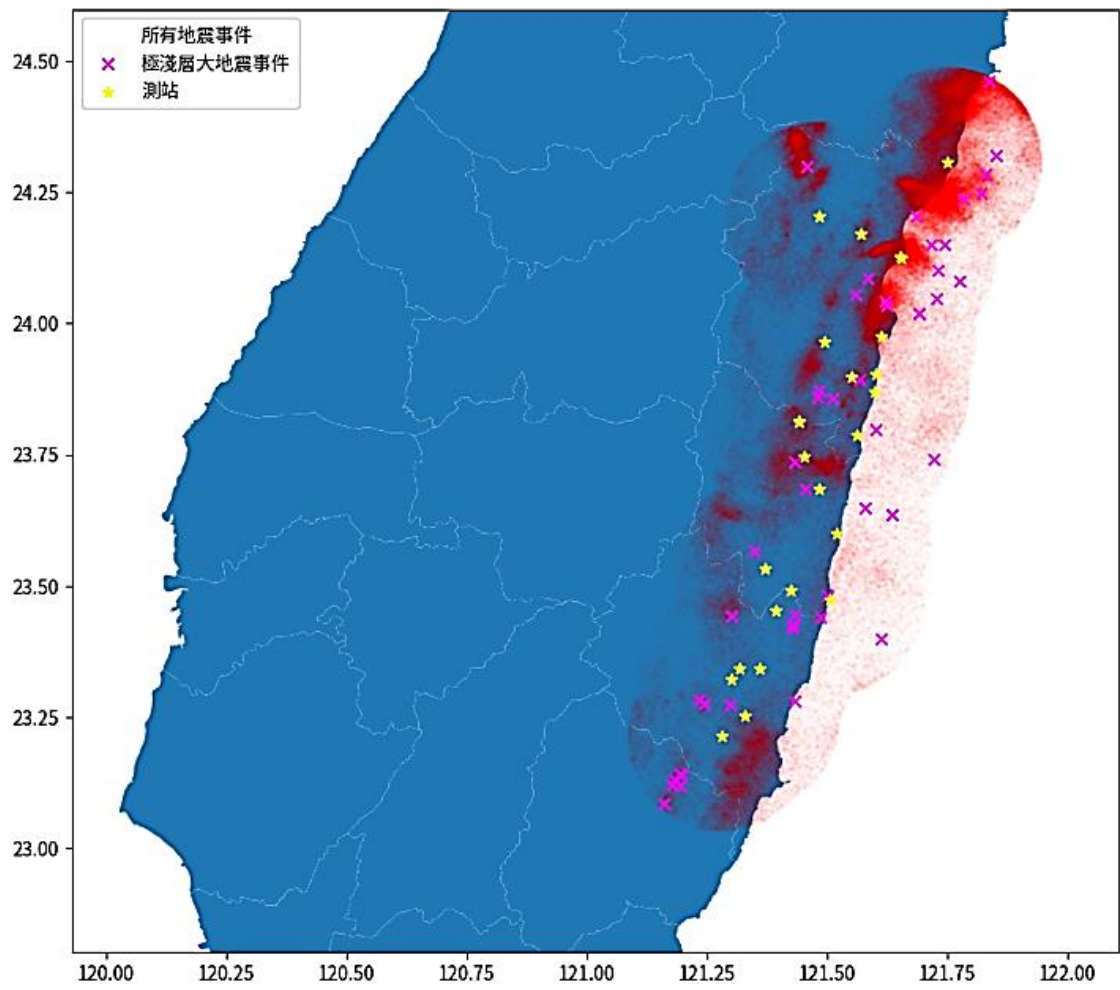
$$FLoss_t(x_{t-T:t-1}^J, y_{t-T:t-1}^A, z_{t,\tau}^A | \theta) = -\alpha(1 - \hat{z}_{t,\tau}^A)^\gamma z_{t,\tau}^A \log(\hat{z}_{t,\tau}^A) - (1 - \alpha)(\hat{z}_{t,\tau}^A)^\gamma (1 - z_{t,\tau}^A) \log(1 - \hat{z}_{t,\tau}^A) \quad (4.6)$$

其中：

α 與前述 balanced cross entropy 相同，用來做正負類的靜態重權； $\gamma \geq 0$ 稱為 focusing parameter，控制「要多強烈地壓低簡單樣本的損失」。 $\gamma = 0$ 時，就退化回 balanced cross entropy；當 $\gamma = 2$ (實務上常用)時，模型對「已經預測得很對的樣本」會給很小的梯度，反而把梯度留給「還預測不好的樣本」。大部分時間都沒有 $M \geq 5.5$ 的淺層地震，以致於這些樣本模型一下就學會「預測成 0」，在 focal loss 的權重項裡就會被 $(\hat{z}_{t,\tau}^A)^\gamma (1 - \hat{z}_{t,\tau}^A)^\gamma$ 壓得很小，而那些「模型已經覺得這某時間點有可能會發生」的樣本，會因為 $(1 - \hat{z}_{t,\tau}^A)^\gamma$ 得到較大的梯度，模型就會繼續去微調這些真正有資訊的區段。這正好符合我們要在大量 0 當中抓出「少數可能有前兆」的需求。

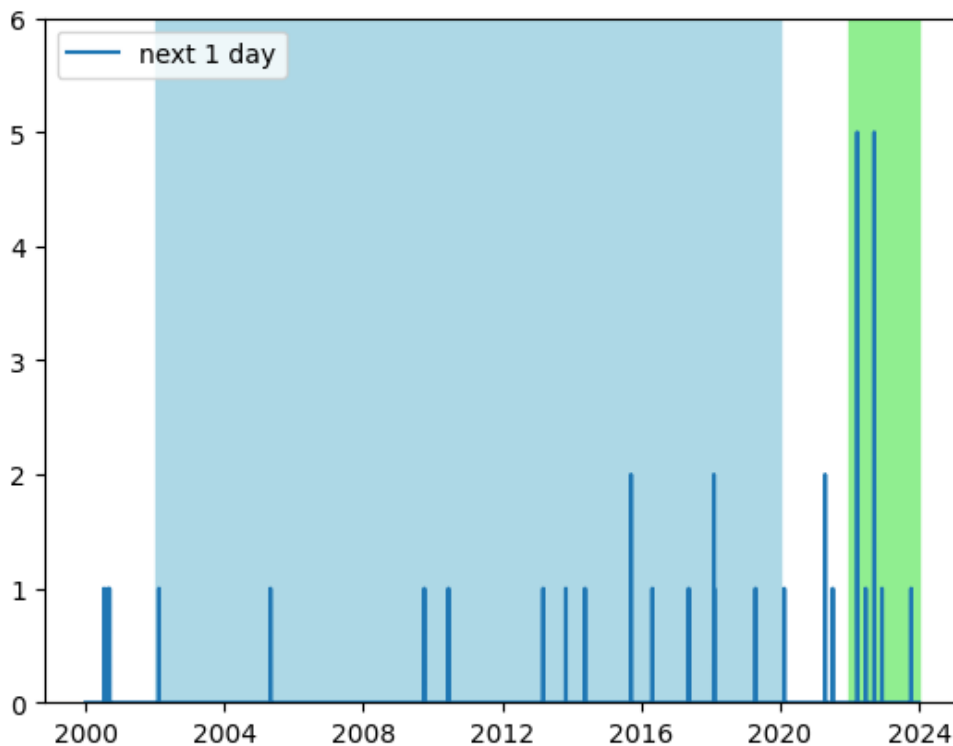
4.3 實驗設計

根據上一小節的討論，本次我們著重在確認模型是否具有預估目標地震(深度 $< 30\text{km}$, $M < 5.5$)的能力。在資料上，我們將目標限定在花蓮地區。GNSS 測站只使用目前仍存活的 25 個測站資料。而在地震資料上，則是從用這 25 個測站半徑 20km 內的地震事件。在資料劃分上，我們以 89-108 年的資料做為訓練集，並以 109-112 年的資料做為測試集。



圖三：目標資料在地圖上的分布狀況

在地震的時間分布上見圖四，我們可以明顯地觀察到 109 年後的目標地震次數明顯較密集，我們可以預期訓練出的模型在測試集上效果應較差。



圖四：目標地震的時間分布狀況。藍色和綠色區域分別為訓練集及測試集的時間帶

在模型輸入上，我們使用了 GNSS 測站資料，以及地震統計資訊。測站資料只選取 dX, dY, dZ 三種，並將缺失數值補為零。地震統計資訊則包含有每日極淺層地震次數、極淺層地震累積能量、淺層地震次數、淺層地震累積能量、中層地震次數、中層地震累積能量，共六種統計值。模型的預估的目標，則是去學習未來的幾日內「是否會發生目標地震」或是「發生目標地震的次數」。

我們以這樣的輸入輸出設計了兩個實驗：

一、固定輸入的時間長度為兩年，但輸入的參數分為：「只使用 GNSS 測站資料」、「只使用過去地震統計資」、「兩者一起使用」三種。輸出的目標區間則參考委員的建議，選擇了 5 種區間： $\tau=1, 90, 180, 365, 730$ 。目的是為了確認是否兩者要一起使用才有預測的效果。

二、固定輸入為「兩者一起使用」。輸入的長度改為 $T=30, 90, 180, 365, 730$ 五種，預估的區間 $\tau=1, 90, 180, 365, 730$ ，總共有 25 種組合。目的是用於判斷更長的輸入長度或是更長的預估區間，是否真的對模型的有正面的影響

在模型訓練上，我們會從訓練集中隨機拆分出 1/10 的資料做為開發集。而目標函數方面，如果是預估未來是否會發生目標地震，則使用 focal loss ($\gamma = 3$)，如果是預估未來目標地震的發生次數，則使用 MSE。每個模型均以 batch size=128 訓練完整的 5 個 epochs，並選取在開發集上 loss 最低者做為最終的模型。

在模型的評估上，我們會通過調整門檻值，計算出每個模型的 ROC 曲線 (Receiver Operating Characteristic curve) (ROC) 及其曲線下面積，做為效能評估指標。

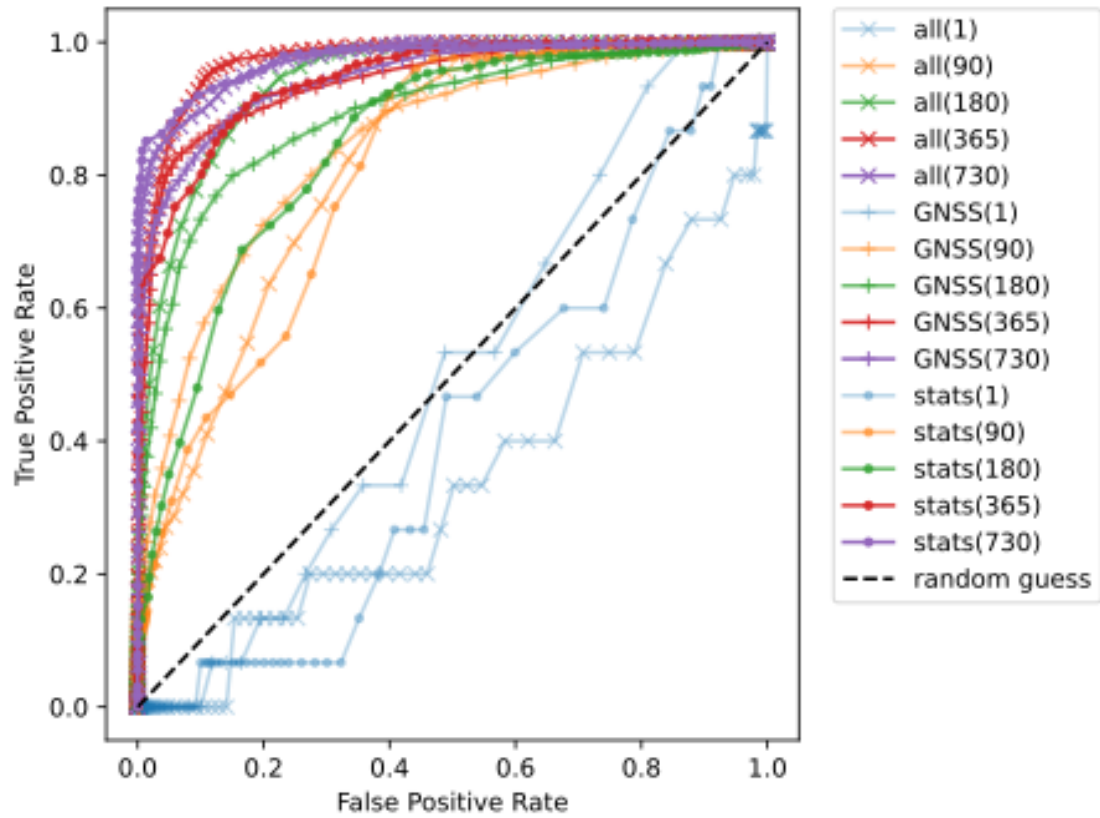
4.4 具體成果

表格 6 用來驗證是否因為預測地震事件時間區間 τ 增加，而正樣本的比例有所增加，因此可以降低預估地震事件的難度。

表格 6: 訓練集合及測試集合中目標地震事件的數量。 τ : 預測地震事件區間 (天)。
N: 總樣本數。H: 規模大於 5.5 且深度淺於 30 公里的事件數目 (正樣本)。

τ (天)	H/N (%) 訓練集合	H/N (%) 測試集合
1	15/6576 (0.23%)	8/732 (1.09%)
90	1083/6576 (16.47%)	433/732 (59.15%)
180	2163/6576 (32.89%)	523/732 (71.45%)
365	3678/6576 (55.93%)	650/732 (88.80%)
730	4979/6576 (75.71%)	650/732 (88.80%)

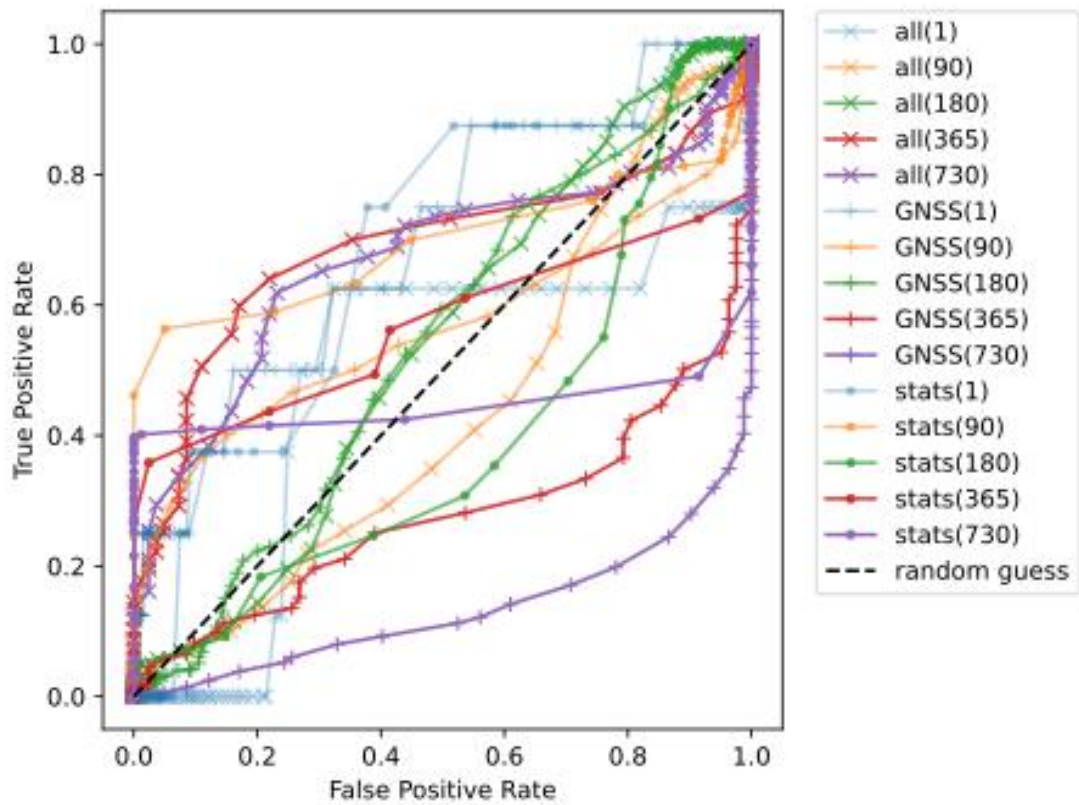
圖五是訓練集下的 ROC 曲線。從圖中可以看出，除了預估未來 1 天的模型，全部的模型都比隨機猜測要好，且有著極高的準確率，代表模型將訓練資料完全學起來。從表格 7 中我們也可以確認到隨著預估區間的增加，模型的準確度有明顯的提升。然而在測試集的結果中(見表格 8 及圖六)，有約半數的模型表現至比隨機猜測要差。在只使用 GNSS 測站資料下，隨著預估天數的增加，模型的表現呈現下降的趨勢，代表模型在訓練集學到的特徵與測試集非常不同，並不具有普遍性。在測試集中，只有同時使用 GNSS 及過去地震統計的模型，呈現隨預測天數的增加而上升的趨勢。



圖五：訓練集各模型的 ROC 曲線。

表格 7：訓練集各模型的曲線下面積

τ 天數	all	GNSS	stats
1	0.3427	0.5002	0.4144
90	0.8176	0.8451	0.8088
180	0.9418	0.8921	0.8464
365	0.9742	0.9423	0.9461
730	0.9729	0.9461	0.9761



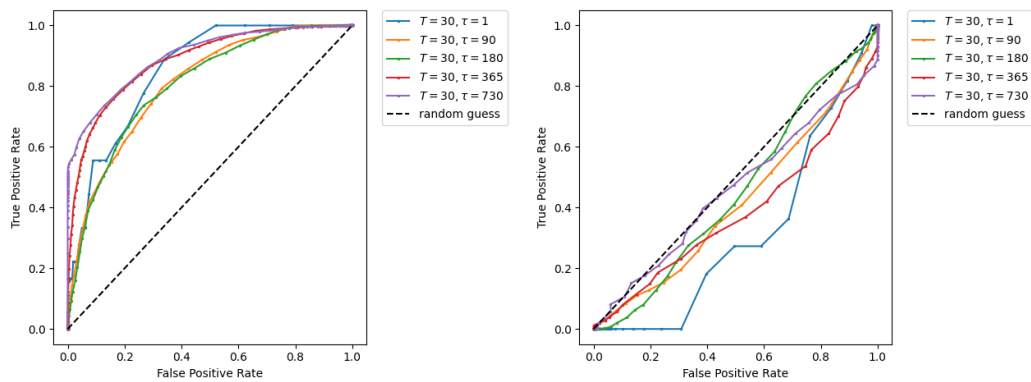
圖六：測試集的 ROC 曲線。

表格 8：實驗一測試集各模型的曲線下面積

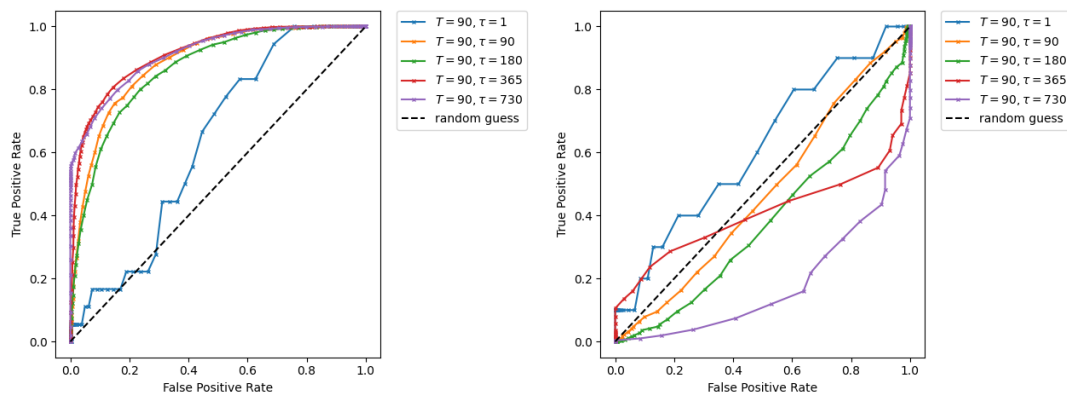
τ 天數	all	GNSS	Stats
1	0.4859	0.7007	0.6838
90	0.4403	0.5645	0.6953
180	0.5381	0.5392	0.4038
365	0.6893	0.2772	0.5672
730	0.6722	0.1326	0.4469

在實驗二中，可以觀察模型在輸入長度 T 和地震區間 τ 的表現。首先是從地震區間的部分（圖七到圖十一），對訓練集來說，隨著 τ 的增加，曲線也越往左上角彎曲，代表我們的目標的確對模型來說更容易學了。當中唯一的例外是在 $\tau=1$ 天的時候，模型在訓練集上的表現非常不好，ROC 曲線呈現有好有壞的狀況。再來是從輸入長度 T 來看（圖十二到圖十六），我們發現，除了在 $\tau=1$ 時，增加輸入長度的確對模型的效能有幫助，當 τ 越長，所需要的 T 也要越長。然而從測試集的結果也會觀察到，當 τ 越長，模型過擬合的狀況也更嚴重，ROC 曲線也往右下彎曲。從曲線下面積來看（表格 9 及表格 10），所有的以模型在測試集的表現上

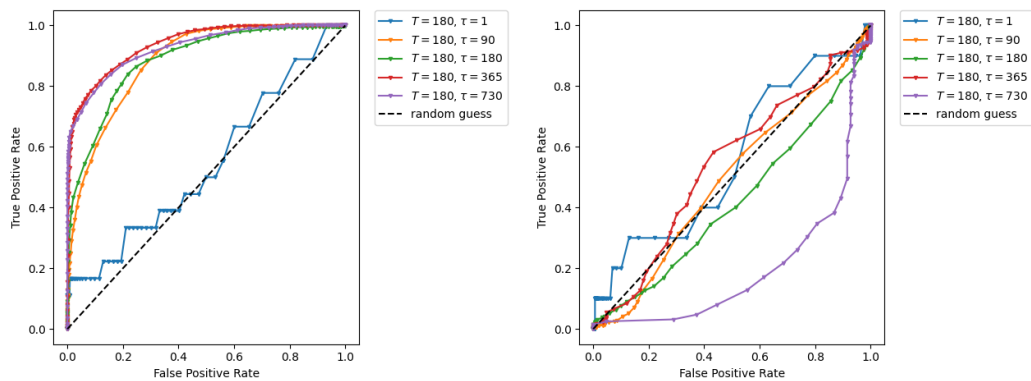
均不佳，與隨機猜測無異，甚至更差。這代表利用 89-108 年的資料學到的特徵並不適合用在 109-112 年。



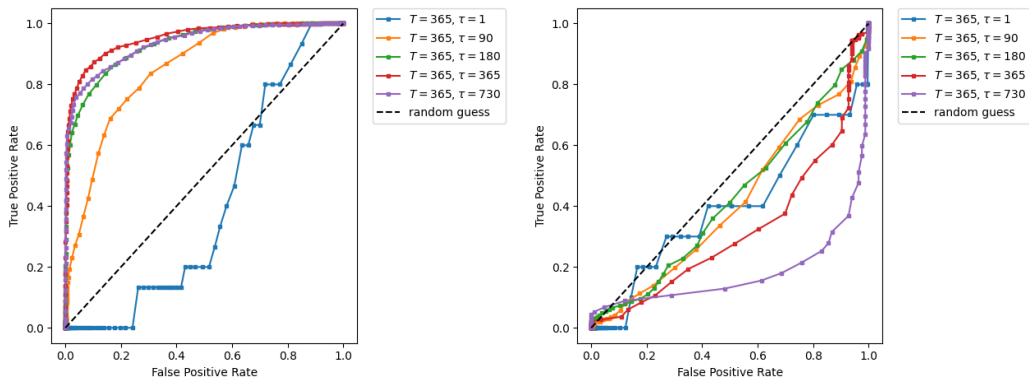
圖七：輸入長度為 30 天的 ROC 曲線，左圖為訓練集，右圖為測試集。



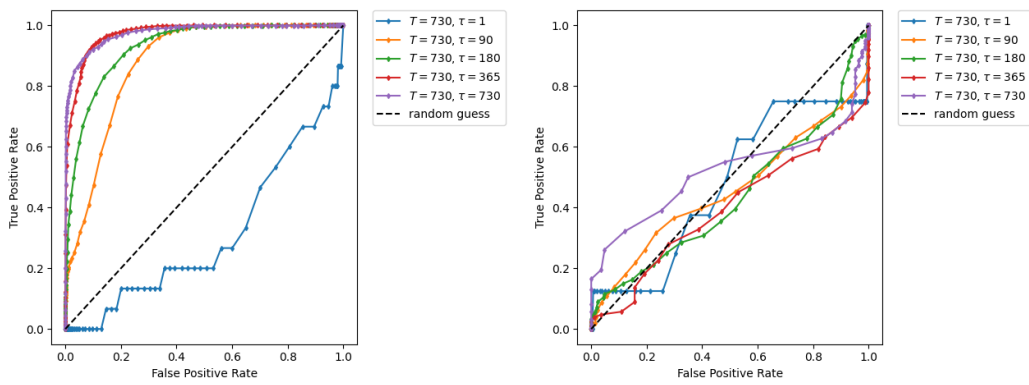
圖八：輸入長度為 90 天的 ROC 曲線，左圖為訓練集，右圖為測試集。



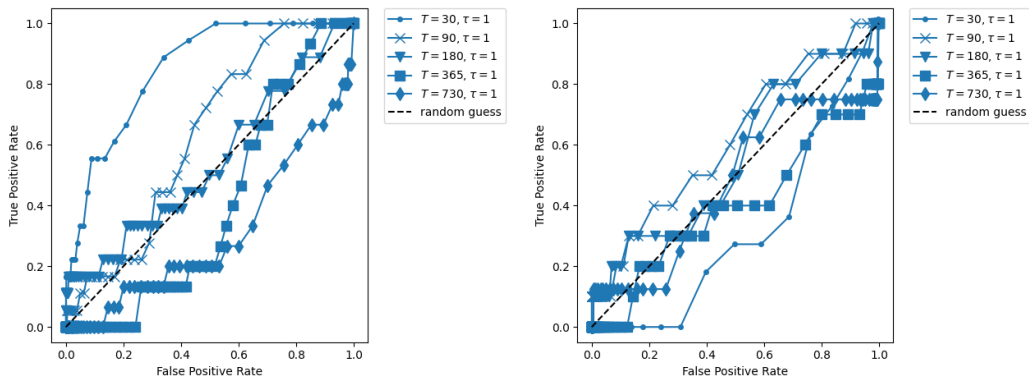
圖九：輸入長度為 180 天的 ROC 曲線，左圖為訓練集，右圖為測試集。



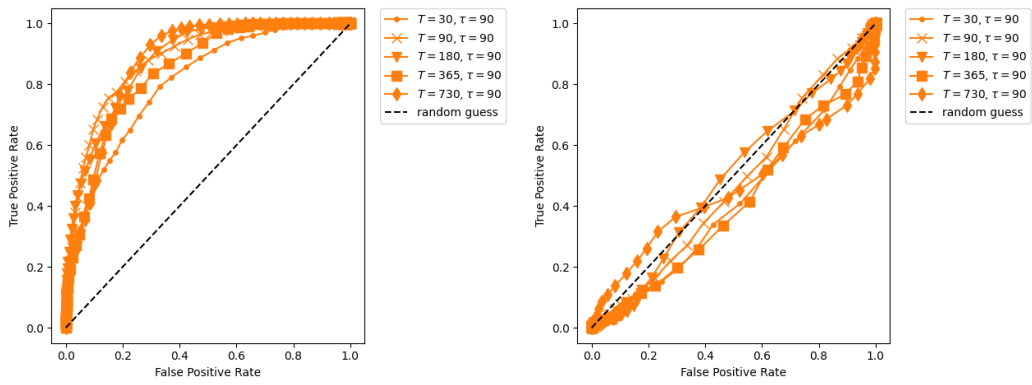
圖十：輸入長度為 365 天的 ROC 曲線，左圖為訓練集，右圖為測試集。



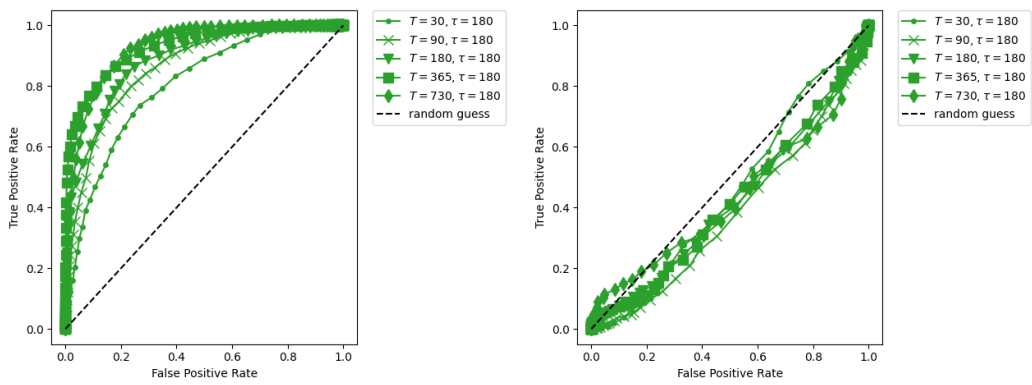
圖十一：輸入長度為 730 天的 ROC 曲線，左圖為訓練集，右圖為測試集。



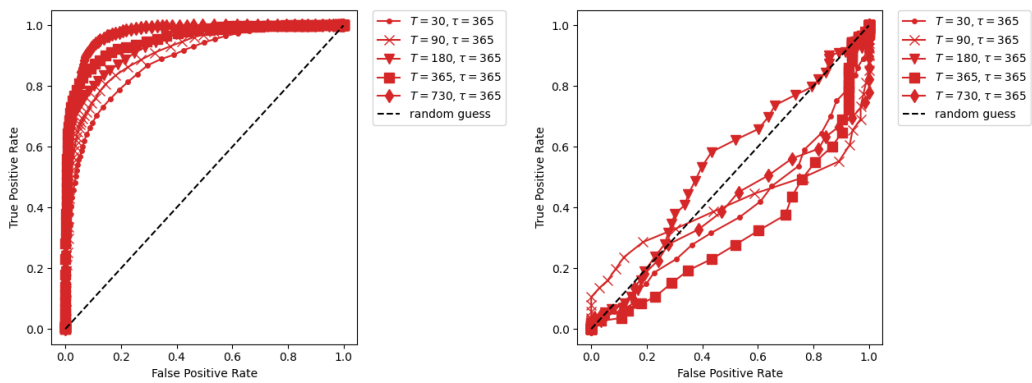
圖十二：預估期間為 1 天的 ROC 曲線，左圖為訓練集，右圖為測試集。



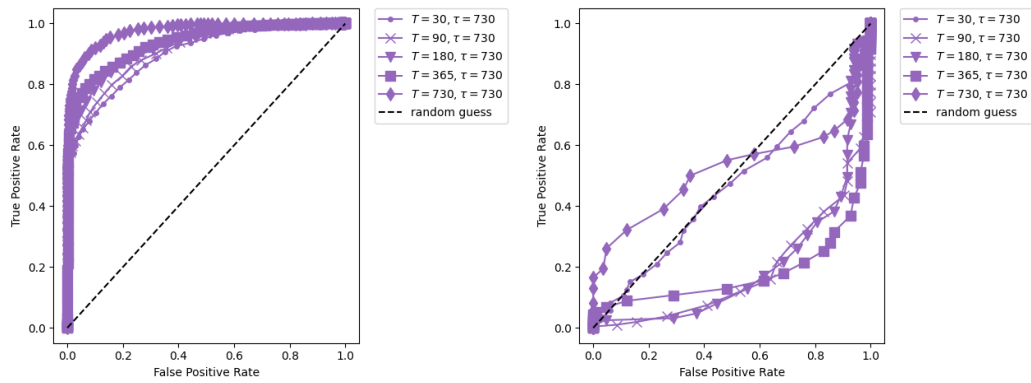
圖十三：預估期間為 90 天的 ROC 曲線，左圖為訓練集，右圖為測試集。



圖十四：預估期間為 180 天的 ROC 曲線，左圖為訓練集，右圖為測試集。



圖十五：預估期間為 365 天的 ROC 曲線，左圖為訓練集，右圖為測試集。



圖十六：預估期間為 730 天的 ROC 曲線，左圖為訓練集，右圖為測試集。

表格 9：實驗二各模型在訓練集上的曲線下面積。

$\tau \backslash T$	30	90	180	365	730
1	0.85	0.62	0.54	0.40	0.30
90	0.80	0.88	0.88	0.84	0.87
180	0.80	0.86	0.88	0.93	0.92
365	0.88	0.91	0.88	0.95	0.97
730	0.89	0.91	0.92	0.93	0.97

表格 10：實驗二各模型在測試集上的曲線下面積。

$\tau \backslash T$	30	90	180	365	730
1	0.31	0.61	0.55	0.39	0.47
90	0.41	0.47	0.49	0.41	0.45
180	0.45	0.38	0.41	0.42	0.43
365	0.38	0.40	0.53	0.32	0.39
730	0.40	0.18	0.20	0.18	0.51

伍、成果的價值與貢獻

使用 GNSS 資料和過去地震事件統計上，我們是有確認到模型的確可以學到一些特徵用以判別未來是否會發生地震，且放寬預估的區間的確能夠讓模型學得更好。在較長天數的預估區間，才需要同時使用 GNSS 及過去地震的統計。預估區間也並非越長越好，過長的預估區間，在測試集上會明顯比隨機猜測要差。

目前但最大問題是資料量實在是太少，使得模型學到的特徵並不一定具備普遍性，訓練出來的模型只適用於訓練集上。我們可能要有一個更好的方式，去從隨機學到的特徵中，再去篩選出真正有效的特徵。

陸、落實應用情形

- 研究方法-1(先前研究)：端至端資料驅動法所使用的數據、資料處理程式、以及地震規模預估程式，皆放置於 https://github.com/pliaoang/cwa_project，方便後續人員使用以及研究之延續。
- 研究方法-2(先前研究)：引入知識之機器學習法究所使用的一些處理好的數據、訓練腳本、模型套件，皆放置於 [https://github.com:orz-orz-orz-orz-cwa_project2.git](https://github.com:orz-orz-orz-orz/cwa_project2.git)，方便後續人員使用以及研究之延續。
- 研究方法-3(先前研究)：引入知識之機器學習法究所使用的一些處理好的數據、訓練腳本、模型套件，皆放置於 https://github.com:orz-orz-orz-orz-cwa_project3.git，方便後續人員使用以及研究之延續。
- 研究方法-4(本研究)：引入知識之機器學習法究所使用的一些處理好的數據、訓練腳本、模型套件，皆放置於 <https://github.com:orz-orz-orz-orz-cwa4.git>，方便後續人員使用以及研究之延續。

參考文獻

- Kitagawa, Y., Koizumi, N., Takahashi, M., Matsumoto, N., & Sato, T. (2006). Changes in groundwater levels or pressures associated with the 2004 earthquake off the west coast of northern Sumatra (M9.0). *Earth, Planets and Space*, 58(2), 173–179. <https://doi.org/10.1186/BF03353375>
- Liu, J. Y., Chen, C. H., Chen, Y. I., Yen, H. Y., Hattori, K., & Yumoto, K. (2006). Seismo-geomagnetic anomalies and $M \geq 5.0$ earthquakes observed in Taiwan during 1988–2001. *Physics and Chemistry of the Earth, Parts A/B/C*, 31(4–9), 215–222. <https://doi.org/10.1016/j.pce.2006.02.009>
- Liu, J. Y., Chen, Y. I., Pulinets, S. A., Tsai, Y. B., & Chuo, Y. J. (2000). Seismo-ionospheric signatures prior to $M \geq 6.0$ Taiwan earthquakes. *Geophysical Research Letters*, 27(19), 3113–3116. <https://doi.org/10.1029/2000GL011395>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*. <https://doi.org/10.48550/ARXIV.1706.03762>
- Wen, S., Chen, C.-H., Yen, H.-Y., Yeh, T.-K., Liu, J.-Y., Hattori, K., Peng, H., Wang, C.-H., & Shin, T.-C. (2012). Magnetic storm free ULF analysis in relation with earthquakes in Taiwan. *Natural Hazards and Earth System Sciences*, 12(5), 1747–1754. <https://doi.org/10.5194/nhess-12-1747-2012>
- Yu, S.-B. (2004). Preseismic Deformation and Coseismic Displacements Associated with the 1999 Chi-Chi, Taiwan, Earthquake. *Bulletin of the Seismological Society of America*, 91(5), 995–1012. <https://doi.org/10.1785/0120000722>
- Feng, L., Tung, F., Ahmed, M. O., Bengio, Y., & Hajimirsadegh, H. (2024). Were RNNs All We Needed? (arXiv:2410.01201). arXiv. <https://doi.org/10.48550/arXiv.2410.01201>
- 最新鄉鎮市區界線(TWD97 經緯度 EPSG:3824). 國土測繪圖資服務雲. (2023, March 23). https://maps.nlsc.gov.tw/MbIndex_qryPage.action?fun=8
- Cui, Y., Jia, M., Lin, T. Y., Song, Y., & Belongie, S. (2019). Class-balanced loss based on effective number of samples. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 9268–9277).
- Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 2980–2988).

114 年地震前兆觀測作業與分析技術相關研究

子計畫四

區域性與全球性地震前電離層異常統計分析

陳佳宏 劉正彥

國立成功大學

摘要

中央氣象署近年建構臺灣地區區域性高解析度電離層監測系統，提供臺灣上空電離層資訊，用以偵測地震前電離層異常情形。過去研究發現在大地震來臨前 1-5 天內出現電離層異常的機率最高，然而近年發現過去的這種異常趨勢有點不同了，除了異常類型(正異常、負異常)之外，異常預測有效時間(前 1-5 天)也有所不同，因此進一步的研究與修改是必需的。本研究將針對 2024 年 4 月 3 日 M7.2 花蓮地震與 2022 年 9 月 18 日 M6.8 臺東地震前電離層異常為對象，尋找地震前時間與空間前兆，並利用過去 M6.0 以上地震進行統計分析研究，測試最佳電離層預測有效時間。除此之外，建構全球電離層時間與空間異常解算程式，提供做為臺灣地區電離層異常驗證與判定。

關鍵字：電離層，全電子含量，地震預警。

Abstract

In recent years, the CWA has established a regional high-resolution ionospheric monitoring system for Taiwan, providing information on the ionosphere above Taiwan to detect ionospheric anomalies before earthquakes. Past research has found that the probability of ionospheric anomalies appearing 1-5 days before a major earthquake is highest. However, in recent years, it has been discovered that this past anomaly trend has changed somewhat. In addition to the types of anomalies (positive anomalies, negative anomalies), the effective prediction time (1-5 days prior) has also differed, necessitating further research and modifications.

This study will focus on the ionospheric anomalies before the M7.2 Hualien earthquake on April 3, 2024, and the M6.8 Taitung earthquake on September 18, 2022, to search for temporal and spatial precursors before earthquakes. It will also conduct statistical analysis research using past earthquakes of M6.0 and above to test the optimal effective prediction time for ionospheric anomalies. Furthermore, a global ionospheric

temporal and spatial anomaly resolution program will be constructed to provide verification and determination of ionospheric anomalies in the Taiwan region.

Keywords : Ionosphere , Total Electron Content (TEC) , Earthquake Early Warning.

壹、前言，研究目的及意義

地震預測是地球科學領域持續面臨的巨大挑戰。儘管科學家們在理解地震成因和機制方面取得了長足進展，但準確預測地震的發生時間、地點和強度仍然是遙不可及的目標。傳統的地震預測方法主要依賴於地震活動性監測、地殼形變觀測以及地磁場變化等手段，但這些方法的預測能力和時效性仍然有限，常常無法提供足夠提前的預警時間。

近年來，利用電離層異常作為地震前兆的研究受到廣泛關注，為地震預測提供了一種新的思路。電離層是地球大氣層中的一個電離區域，其電子密度受各種因素影響，包括太陽活動、地磁活動以及地球本身的物理過程。一些研究表明，在地震發生前，由於地殼應力累積和釋放過程中產生的電磁效應，電離層的電子密度可能會出現異常變化，從而形成可觀測的地震電離層前兆。這些異常變化可以通過全球導航衛星系統(GNSS)測量到的全電子含量(TEC)來檢測。國際上關於電離層地震前兆的研究已持續數十年，許多研究利用全球 GNSS 網絡的數據來探測地震前電離層 TEC 的異常變化。研究方法涵蓋了統計分析、機器學習和物理建模等多種手段。然而，由於地震的複雜性和電離層本身的變動性，現有的預測模型的準確性和可靠性仍然有限，預測的成功率不高，預報時間也存在不確定性。許多研究集中在特定區域或特定類型的地震，缺乏普適性。

中央氣象署近年來建構的臺灣地區區域性高解析度電離層監測系統，為研究臺灣地區的地震電離層前兆提供了寶貴的數據資源。該系統提供高時間和空間解析度的 TEC 數據，可以更精細地捕捉電離層的微小變化，從而提高地震前兆檢測的精度和可靠性。過去的研究已經證實，在大地震發生前 1-5 天內，電離層異常出現的機率相對較高，這為利用電離層異常進行地震預測提供了初步的依據。

然而，近年來的研究也發現，過去建立的電離層異常預測模型的有效性有所下降，地震前電離層異常的趨勢出現了變化。這不僅表現在異常的類型(正異常或負異常)上存在更多樣性，而且異常出現的有效時間(1-5 天)也存在不確定性，有些地震可能在更早或更晚的時間出現明顯的電離層異常，甚至有些地震在傳統的 1-5 天時間窗口內沒有明顯的異常。這種現象可能與地震的規模、震源深度、斷層特性以及其他地球物理因素的複雜交互作用有關。因此，需要進一步的研究來探討電離層異常的物理機制，改進現有的預測模型，並提高地震預測的準確性。

本研究將重點關注近年來臺灣地區發生的幾次較大規模地震，例如 2024 年 4 月 3 日花蓮 M7.2 地震和 2022 年 9 月 18 日臺東 M6.8 地震，分析這些地震發生前電離層 TEC 的變化特徵，並與過去 M6.0 以上地震的電離層數據進行比較分析，以期建立更可靠和更具普適性的電離層地震前兆預測模型。研究目的具體在以下三個方面：

1. 精確刻畫地震電離層前兆:透過分析 2024 年花蓮 M7.2 地震和 2022 年臺東 M6.8 地震, 以及其他 M6.0 以上地震的電離層數據, 精確刻畫不同規模地震發生前電離層 TEC 的時間與空間異常特徵, 包括異常的類型、強度、持續時間和空間範圍等。特別關注近年來觀察到的電離層異常趨勢變化, 例如異常類型和預測有效時間的改變。
2. 開發全球電離層異常解算程式: 開發一套自動化處理全球 GIM 數據的程式, 用於識別和分析全球範圍內的電離層異常, 並驗證臺灣地區高解析度電離層數據中檢測到的異常的區域性和全球背景影響。此程式將為未來研究提供重要的數據分析工具。
3. 建立更可靠的地震電離層異常判斷準則: 基於高解析度區域性電離層數據和全球 GIM 數據, 建立更可靠和更具普適性的電離層地震異常判定準則。

貳、研究方法與進行步驟

本研究將採用多種研究方法, 以期全球、深入地分析臺灣地區地震與電離層異常的關係, 並改進地震電離層異常判定準則。研究方法主要包含:

1. 數據收集與處理:
 - a. 高解析度電離層數據: 收集中央氣象署提供的臺灣地區高解析度電離層 TEC 數據, 時間範圍涵蓋 2024 年 4 月 3 日花蓮 M7.2 地震和 2022 年 9 月 18 日臺東 M6.8 地震發生前數週至數月, 以及地震發生後數天。數據的空間解析度需高於或等於 $0.25^{\circ} \times 0.25^{\circ}$, 時間解析度需達到 15 分鐘。
 - b. 全球電離層數據: 收集全球範圍的 GIM TEC 數據, 這些數據將用於驗證臺灣地區高解析度電離層數據的異常情況, 並探討區域性異常與全球電離層背景狀態的關係。數據來源可從 IGS 或其他相關數據中心(例如 CODE)獲取。
 - c. 地震目錄數據: 收集臺灣地區 M6.0 以上地震的目錄數據, 包括地震發生的時間、地點、規模和震源深度等訊息。這些數據將用於統計分析, 研究電離層異常與地震參數之間的關係。
 - d. 數據預處理: 對收集到的數據進行預處理, 包括去除異常值、填充缺失值、去除多徑效應和衛星噪聲等。採用適當的插值方法填充缺失數據, 例如克里金插值(Kriging)或線性插值(Linear)。去除多徑效應和衛星噪聲可以採用濾波技術或基於模型的校正方法。
 - e. 數據去趨勢: 利用多項式擬合或其他時間序列去趨勢方法, 去除 TEC 數據中的長期趨勢, 例如季節變化和太陽活動的影響。確保數據分析僅關注地震相關的異常變化。
2. 電離層異常識別與分析:

- a. 基線建立: 針對每個選定的地震事件, 選擇地震發生前一段相對平靜的時期作為基線, 利用移動平均法或多項式擬合法建立基線模型, 消除長期趨勢的影響。
 - b. 異常檢測: 基於基線模型, 採用統計方法(例如, 標準差法、閾值法)檢測 TEC 數據中的異常變化。需仔細設定閾值, 並評估不同方法的檢測效果。
 - c. 空間-時間分析: 分析 TEC 異常的空間分佈和時間演變, 確定異常區域的範圍、強度和持續時間。利用空間插值方法繪製 TEC 異常圖, 並結合地震的空間位置進行分析。
 - d. 異常分類: 根據 TEC 異常的類型(正異常或負異常), 以及異常出現的時間和空間範圍, 對電離層異常進行分類。
3. 全球電離層時間與空間異常解算程式開發:
- a. 程式設計: 開發一套全球電離層時間與空間異常解算程式, 該程式將能自動化處理全球 GIM TEC 數據, 並識別和分析全球範圍內的電離層異常。此程式將使用高效的算法, 以便處理大量數據。
 - b. 程式驗證: 使用已知地震事件的數據驗證程式, 確保程式能準確識別和分析電離層異常。評估程式的準確性和效率。
 - c. 臺灣地區異常驗證: 利用開發的程式處理全球 GIM 數據, 對臺灣地區高解析度電離層數據檢測到的異常進行驗證, 判斷其是否為區域性現象或受到全球電離層背景狀態的影響。

參、具體成果

數據收集與處理

為了分析全球電離層異常區域, 本研究採用 CODE 所製作的全球電離圖(Global Ionosphere Map, GIM)產品, 下載位置為 <ftp://gdc.cddis.eosdis.nasa.gov>。CODE 提供的 GIM 資料在時間上會有 2 天左右的延遲, 為了能夠儘早得知電離層的狀態, 因此在 CODE 最終產品發佈前會先使用 CODE 的即時 GIM 資料。兩者數值會有些許的差異, 但整體的趨勢不會有太大的改變。即時 GIM 資料不會儲存下來, 因此待最終 GIM 資料發佈後, 程式會下載最終 GIM 資料並進行後續的分析研究。

GIM 資料的空間解析度為經度 5 度與緯度 2.5 度, 時間解析度為 1 小時(圖 1)。為了與臺灣區域性 GNSSTEC 資料進行比對, 因此先將資料利用空間線性內插在目標位置(例如 24°N, 120°E)上, 然後再利用 cubic spline 內插成 15 分鐘資料, 最後再計算上下界值與異常天。由圖 2 兩者比較圖可以看到, 基本上 GNSSTEC 與 GIMTEC(紅色線)的每天變化趨勢非常相似, 並且異常天(紅色點)的日期也很接近, 說明兩者資料雖來源不同但具有高度一致性。兩者皆利用 GNSS 進行 TEC 的解算, 但由於資料量的不同以及解算方法不同, 因此有些許差異, 但可以看出兩

者資料展現出明顯的相關性，顯示出不同資料與方法在描述電離層活動時的互補性(GIMTEC 提供全球資料，GNSSTEC 提供高時空解析度資料)。目前自動下載資料、解算 GIMTEC 與繪製圖 2b 的程式已在氣象署遠端電腦上安裝、設置與測試完成，目前尚待分析方法(例如電離層異常來測基線、空間異常判定等)確定以及每日排程解算設定，即可提供除了臺灣區域性 GNSSTEC 資料之外的全球電離層資訊，可做為研究與分析地震電離層前兆的交叉驗證資料使用。

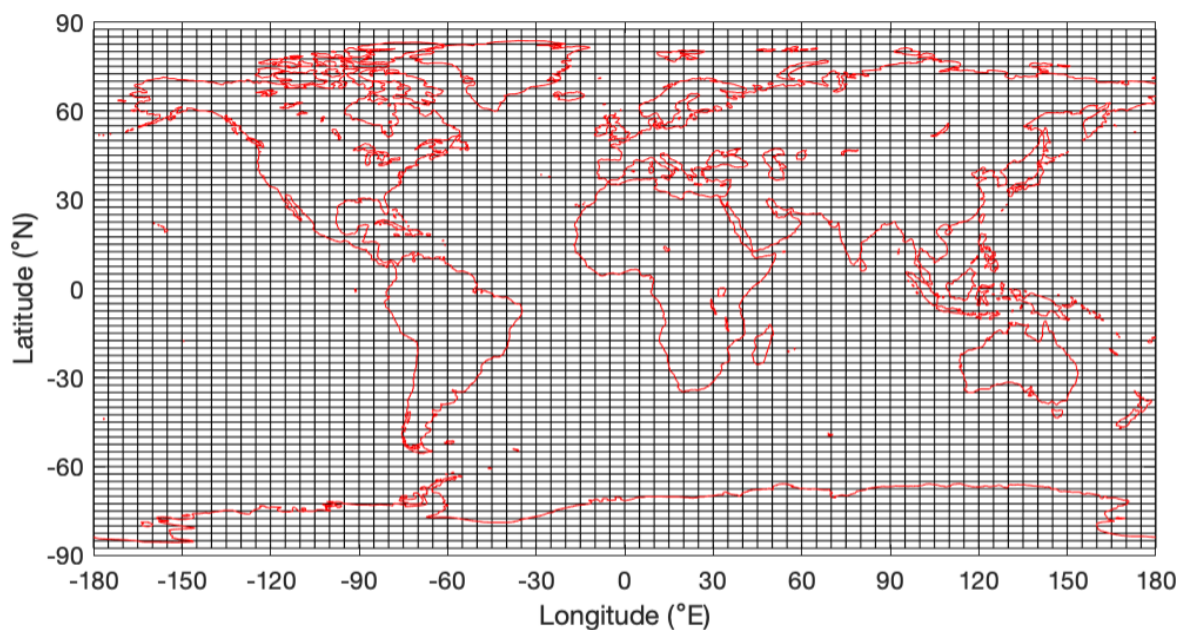


圖 1、全球電離圖空間解析度。全球共 73*71=5,183 個資料點。

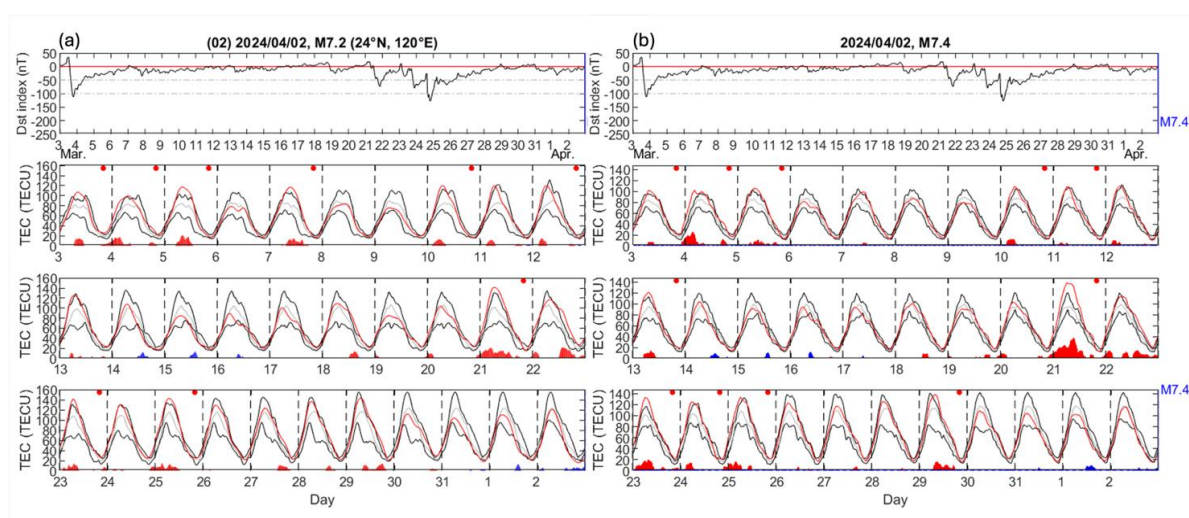


圖 2、臺灣區域 GNSSTEC(a)與全球電離圖(b)的比較。最上層子圖中的黑線為地磁 Dst 指數，第 2 到第 4 子圖中的紅線為當天 TEC 觀測資料、灰線為前 15 天之 TEC 中位數值、兩條黑線分別為上界與下界 TEC 區間值。子圖底部紅色與藍色區域分別代表當天 TEC 觀測值超過 TEC 上界與低於 TEC 下界的差值。藍色垂直線為地震發生時間與規模。程式判斷為異常天會用紅色點與藍色點(本例中無藍色點)代表

TEC 正異常與負異常。

個案地震 TEC 資料分析

本研究先針對 2024 年 4 月 2 日(UT 時間)的花蓮地震進行分析，然後再擴及多筆地震資料進行統計研究，找尋判斷地震前電離層時空異常之最佳方法。下圖 3a 為利用 GIM 資料所內插出在震央上空之 TEC 變化情形，可以看到在 3/30 與 3/31 兩天，紅色線的 TEC 值有明顯低於 TEC 下界的情況，因此程式將這兩天判斷為電離層下界異常天。圖 3b 為 3/31 0800UT 時全球 TEC 值與 TEC 下界值之間的差異，顏色越藍色代表差異越大。可以看到在臺灣有明顯的負值存在，代表 TEC 非常低於下界值，並且可以看到負值位於理論地震孕震的範圍之內，也是全球負值最多最集中的區域。地震孕震的半徑範圍是參考 Dobrovolsky et al. (1979)所提出的公式：

$$R = 10^{0.43M} \quad (1)$$

其中 R 為半徑，單位為公里；M 為地震規模。花蓮地震所計算出來的地震孕震半徑範圍約為 1520.55 公里。

根據過去地震前電離層異常的研究經驗以及電離層理論物理，地震前所造成的電離層異常通常不會一下子就消失了，而是會持續一段時間。因此，本研究除了找出電離層異常的時間與空間位置之外，另外找出異常連續出現 9 小時以上的空間位置。圖 3b 中紅色線所圈選的範圍即為符合上述條件的地區，其中一部分位於花蓮地震的孕震半徑範圍之內，南半球印度與印度洋的異常區域為北半球異常區域共軛的結果。另外，就全球異常區域分布來看，雖然非洲南部與南大西洋地區的 TEC 異常值不大，但因為連續出現 9 個小時以上，因此為偵測為 TEC 的空間異常區域。

然而，這裡值得注意的是，圖 3a 與圖 2b 是使用完全相同的 GIM 資料，但是圖 2b 中的異常天為 3/29，並且為正異常，這與圖 3a 的結果不同。造成這個差異主要是兩張圖有以下 2 點不同：(1)參考的背景天數不同，圖 2b 使用地震前 15 天資料做為背景值計算 TEC 上下界，而圖 3a 則是使用地震前 7 天資料。(2)TEC 上下界範圍不同，圖 2b 使用 1.9 倍的 IQR 範圍去計算 TEC 上下界值，而圖 3a 則是使用 3 倍的 IQR 去計算。目前氣象署所使用的 TEC 異常天計算資料是使用 15 天前觀測資料做為參考天，並且使用 1.9 倍的 IQR 範圍(與圖 2b 相同設定)，若是使用這個設定的話，將得到 3/29 為 TEC 正異常天。圖 3c 為 3/29 0800UT 的全球正異常分布情形，可以看到在臺灣上空有非常顯的正異常，並且其連續出現超過 9 個小時以上，南半球同樣也可以看到電離層共軛的結果。除此之外，雖然 TEC 異常值沒有臺灣地區那麼大，但在臺灣緯度以北的幾乎整個北半球都可以看到 TEC 正異常的分布，比照圖 3a 上圖的地磁活動來看，也沒有特別的地磁擾動發生，因此無法解釋其來源為何。

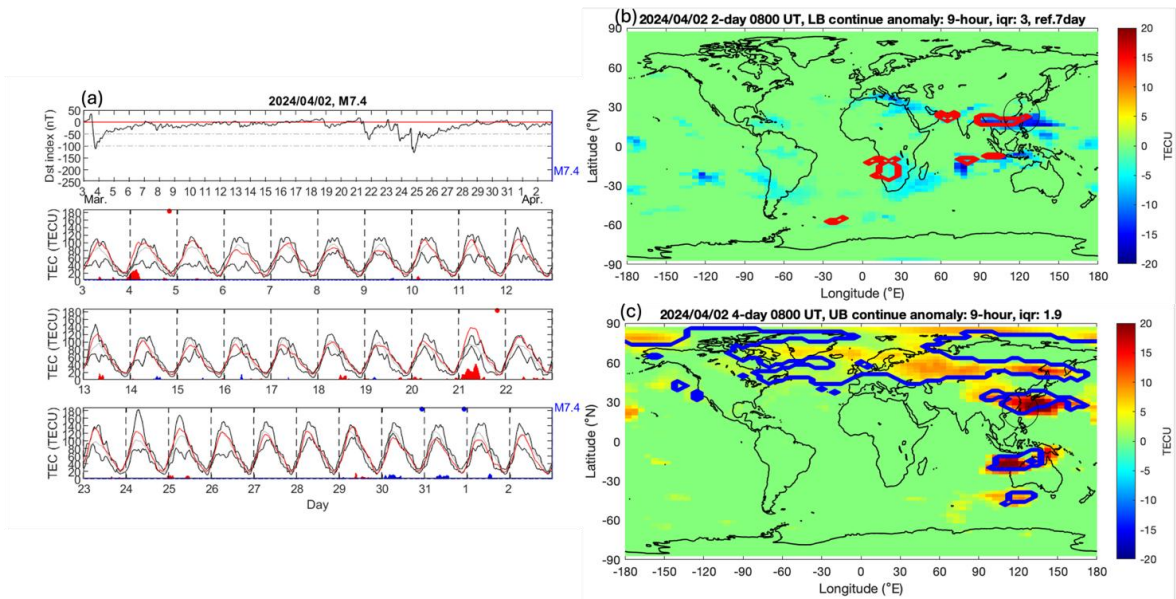


圖 3、2024 年 4 月 2 日花蓮地震震央上空 GIMTEC 時間(a)與空間(b、c)變化情形。上下界 TEC 值採用 3 倍 IQR 值。(b、c)圖中紅色線範圍為 TEC 下界異常累積 9 小時以上之異常區域，顏色為低於下界之 TEC 差值。震央附近的圓為地震孕震過程的可能範圍。(b)與(c)的差異為(b)是 TEC 下界異常，(c)是 TEC 上界異常，此外兩者的設定也不同，詳細說明請參考本文。

以上結果說明，不同的參考天數以及 IQR 值，將會得到不同的結果。過去氣象署所使用的設定是根據過去臺灣區域發生之地震所建立起來的統計分析最佳設定，但是同一個 2024 年的花蓮地震事件來說，不同的設定將導致 TEC 正異常與負異常不同的結果，並且日期也都不相同，特別是過去研究也有指出，不同地區的電離層地震前兆的特性皆有所不同，因此需要謹慎評估目前使用的異常計算參數是否適合所有地震事件。

全球地震統計分析設定

為了能夠消除以上的疑慮，本研究選取了 2020 年到 2025 年 10 月初的期間內，全球發生規模大於等於 7 以上的 385 筆地震進行統計分析研究，圖 4 為本研究統計分析地震的全球分布地理位置。圖 5 為本研究進行全球地震前兆統計分析所使用的地震樣本背景資料，呈現自 2000 年 1 月 1 日至 2025 年 10 月 10 日期間，全球規模大於等於 7.0 之強震的年度分布與累積趨勢。根據統計，此時間區間內全球共記錄到 385 筆規模 7.0 以上的地震事件，從藍色長條圖的年度分布可見，全球強震的發生頻率在某些年份有顯著增加(如 2010 年前後)，而某些年份則相對平靜。然而，從黑色實線所呈現的累積趨勢觀察，儘管年度計數有所起伏，但長期而言，全球強震的發生率呈現一個相對穩定的增長態勢。

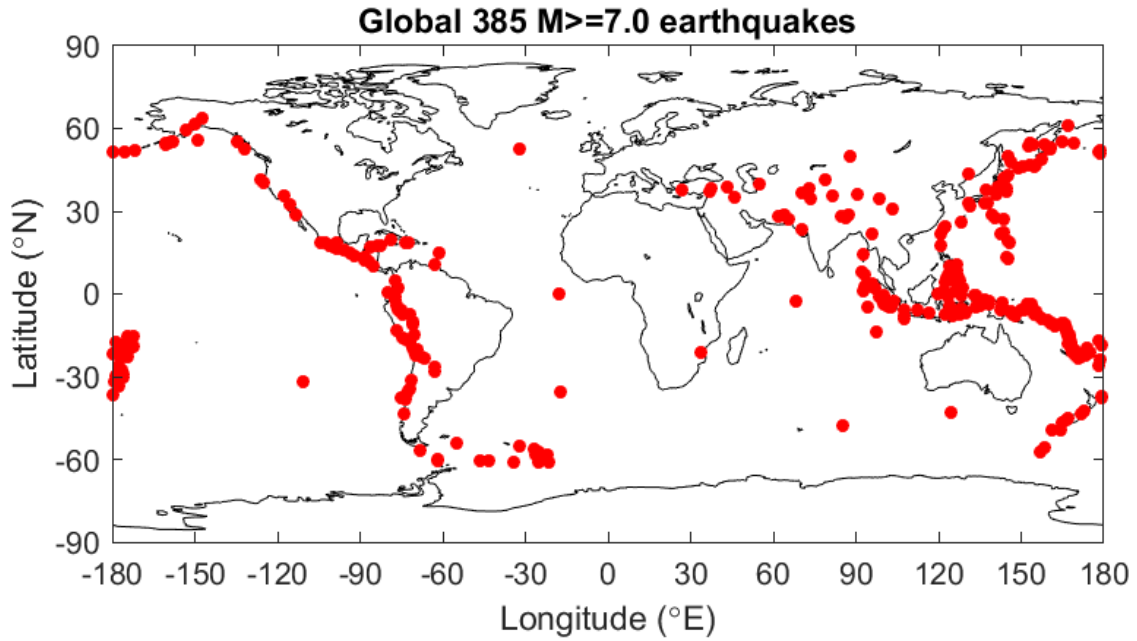


圖 4、2020 年到 2025 年 10 月初地震規模 7.0(含)以上的震央分布。

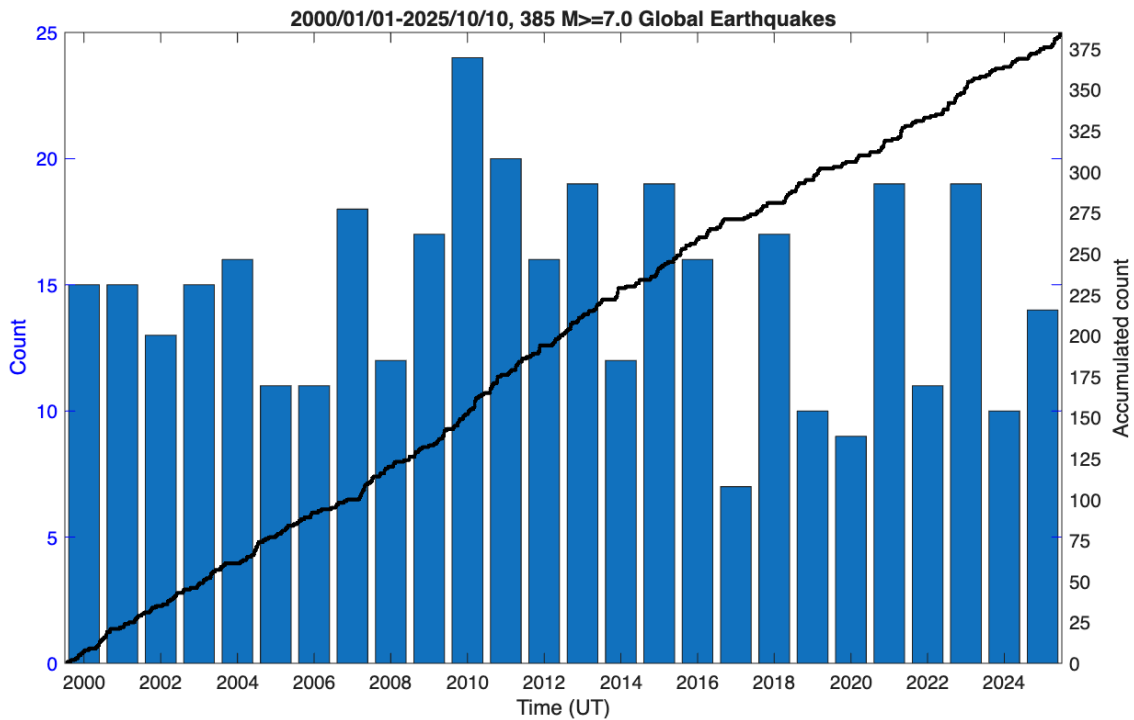


圖 5、2000 年至 2025 年 10 月期間，全球規模大於等於 7.0 之地震年度分布與累積計數統計圖。圖中藍色長條圖(對應左側縱軸)為每年的地震發生次數，黑色實線(對應右側縱軸)則為自 2000 年起算的累積地震總數。此統計區間內共記錄 385 筆地震事件。

為進一步探討地震發生的時間特性，本研究分析了此 385 筆全球 7.0 以上地震

樣本，在其震央當地時間(local time, LT)上的分布情形，如圖 6 所示。從分布直方圖可見，儘管在某些特定時段(如當地時間傍晚 18 點)的計數略為突出，但整體而言，強震的發生次數在一天之內並未呈現出顯著的集中趨勢或明確的日夜週期性。其分布狀態接近隨機，並無任何時段顯示出統計上的顯著高發生率或低發生率。此結果說明，對於全球尺度下的強震活動而言，其觸發機制與由太陽照射所主導的當地日夜循環之間，並無明顯的物理關聯。基於此項分析，本研究在後續探討電離層地震前兆時，將不把地震發生的當地時間列為一個主要控制變因或分類依據，從而更能聚焦於分析震前天數等關鍵時間尺度。

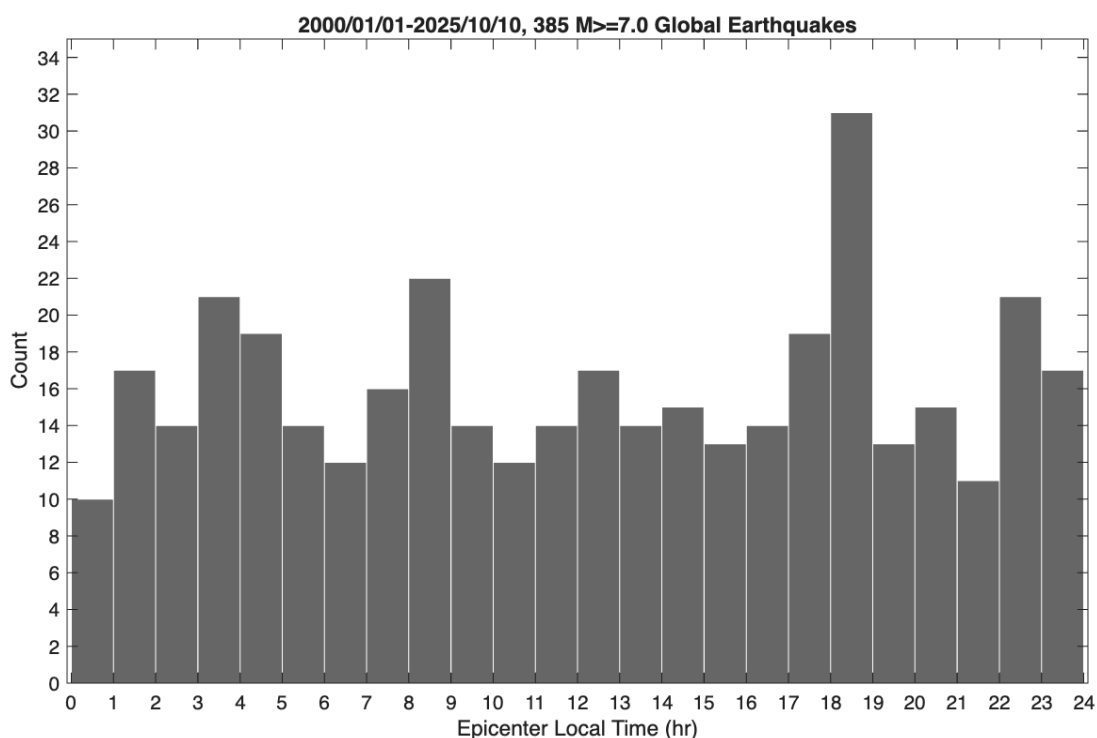


圖 6、2000 年至 2025 年 10 月期間，全球規模大於等於 7.0 之地震，其發生時間對應於震央當地時間的統計分布圖。橫座標為一天 24 小時，縱座標為地震計數。

接著，為了建立統一的電離層異常值標準，因此重新定義電離層背景值的設定：參考背景天設為 16 天，上下界值利用標準差值來表示。其中，因為電離層電漿變化有所謂的準 16 天週期(quasi 16-day period)，因此參考背景天設定為 16 天，可包含整個變化週期。此外，為了不要讓 TEC 上下界根據不同地震事件而有所改變，因此採用標準差值的計算，看 TEC 值到達幾倍的標準差值即可知道依據統計學標準差的概念而得知其異常程度：68% 數值位於 1 倍標準差範圍內、95% 數值位於 2 倍標準差範圍內、99.7% 數值位於 3 倍標準差範圍內。將上述的標準與設定進行資料分析處理後，可以得到圖 7 的結果。由圖 7 可以明顯看到紅色的 TEC 值每天位於幾倍標準差的區間內，即可知道其值的異常程度。圖中顯示，3/29 的 TEC 達到 2 倍標準差區間內，若將超過 2 倍標準差定義為異常天的話，則 3/29 可定義為正異常天，這與圖 2b 的判斷結果一致。而圖 3a 所判斷 3/30 以及 3/31 為負異常天在圖 7 中則顯示，這兩天的 TEC 值雖然較低，但仍然位於 1 倍標準差範圍之內，

屬於正常區間內的 TEC 值。

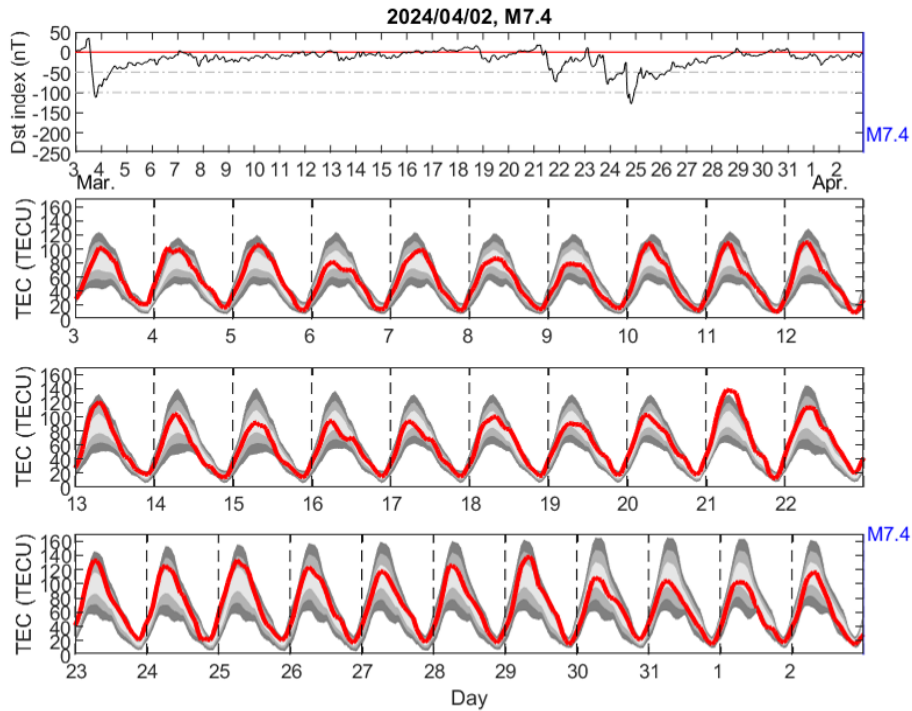


圖 7、2024 年 4 月 2 日花蓮地震電離層 TEC 變化情形。圖的格式與圖 2 相同，灰色區域為不同標準差的區域，顏色最深的為 3 倍標準差，顏色次深的為 2 倍標準差，顏色最淺的為 1 倍標準差。

將圖 7 中 3/29 的上界異常空間分布繪製出來，如下圖 8 所示。可以看到在花蓮地震孕震範圍內的 TEC 值有明顯超過 2 倍標準差的區域，南半球也同樣存在有共軛效應。除此之外，北半球較高緯度地區同樣存在有超過 2 倍標準差的區域，跟圖 3c 的分布很相似，但是範圍就沒有那麼大，特別是在東歐地區，新的設定讓 TEC 小於 2 倍標準差(圖 8)。從連續發生異常的區域來看，跟圖 3c 比起來範圍縮小許多，但花蓮地震孕震範圍內仍然可以看到連續的異常區域，而南半球的 TEC 異常共軛區域也不再連續超過 9 小時，整體連續異常區域被侷限在特定地區。

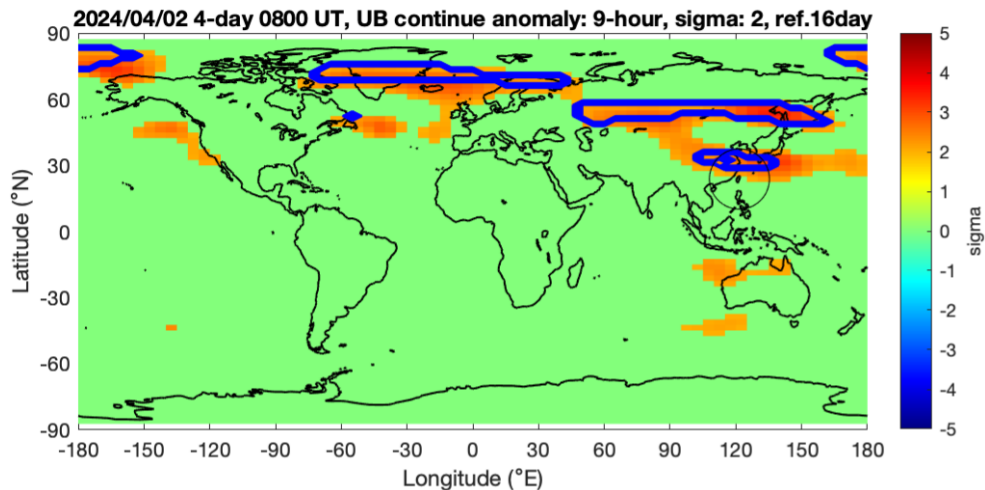


圖 8、2024 年 4 月 2 日花蓮地震電離層 TEC 標準差值空間分布情形。圖中顯示大於 2 倍標準差的數值。藍色線圍起來的區域代表連續超過 9 小時以上的異常值區域。

全球地震時間異常分析

利用上述的設定標準，首先進行研究區間中 385 筆規模 7.0(含)以上的地震前電離層「時間異常」統計分析。為了從大量的地震案例中萃取出具普適性的前兆特徵，本研究利用前述建立的標準化異常判斷準則(即以 16 天為背景參考期，將超出正負 2 倍標準差範圍定義為異常)，對全球 385 筆規模 7.0 以上地震進行了疊加統計分析。分析的時間範圍涵蓋了地震發生前的 30 天，結果如圖 9 所示，呈現了在地震前的每一天出現電離層正異常(TEC > +2 倍標準差)與負異常(TEC < -2 倍標準差)的地震案例所佔的百分比。可以明顯看到，在整個震前 30 天的觀測窗口內，正異常的發生率(灰色虛線)幾乎全面性的高於負異常的發生率(黑色實線)。為了更清晰的量化此趨勢，圖中長條圖表示每日正、負異常發生率之間的差值，灰色長條代表正異常大於負異常，黑色長條代表正異常小於負異常。由圖可見，長條圖始終為正值(除了第 29 天之外)，說明了對於全球強震而言，其電離層前兆以正異常為主導現象。此一不對稱性在特定時間點尤為突出，其中在地震前第 12 天附近，正異常的主導性(即長條圖高度)有相對高峰，顯示這個時段可能是電離層前兆訊號較為集中的關鍵窗口。此統計結果確立了正異常作為主要觀測目標的重要性，也提供未來預測模型有效時間設定的重要數據支持。

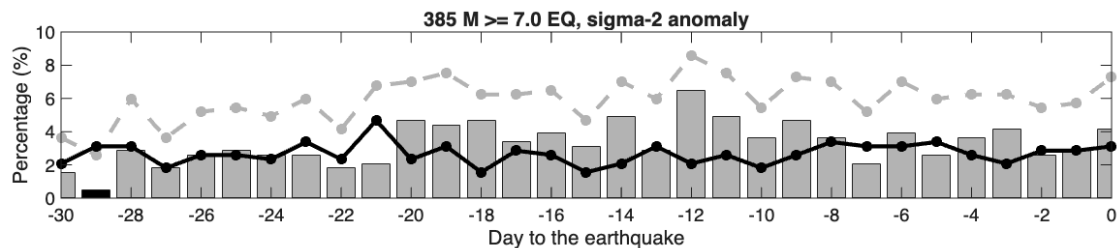


圖 9、全球 385 筆規模 7.0 以上地震，其震前 30 天電離層 TEC 異常發生率之統計結果。圖中灰色虛線與黑色實線分別表示每日 TEC 超出 2 倍標準差上界(正異常)與低於 2 位標準差下界(負異常)的地震案例百分比。長條圖則為正異常與負異常百分比之差值，正值為灰色長條圖，代表正異常較多，負值為黑色長條圖，代表負異常較多。

接著進一步探究地震規模是否影響電離層前兆的特徵，本研究將 385 筆地震樣本依據規模大小進行分類，並分別進行疊加統計分析，結果如圖 10 所示。此圖由 6 個子圖(a)至(f)組成，每個子圖代表一個不同的最低規模門檻，圖中的符號定義與前述圖 9 相同。圖 10a 為所有規模 7.0 以上地震統計結果，與圖 9 相同，作為比較的基準。圖 10b 篩選出規模 7.2 以上的 249 筆地震，其統計樣貌與圖 10a 類似，正異常在震前第 12 天附近依然顯著。隨著規模門檻的提升，圖 10c(170 筆)、10d(111 筆)、10e(59 筆)至 10f(28 筆)，樣本數雖然逐漸減少，但一個清晰的趨勢也隨之浮

現。那就是，在地震發生前第 3 天的正異常趨勢，隨著地震規模的增加而越來越顯著。在規模 7.0-7.4 的統計中(圖 10a-c)，第 3 天前的正異常訊號並不突出。然而，從規模 7.6(圖 10d)開始，第 3 天前的正異常發生率顯著提升，並在規模 7.8 以上(圖 10e-f)的地震中，成為最為明確且強烈的統計特徵之一，顯示出其作為極強地震前兆的高度潛力。此外，值得注意的是，對於規模較大的地震群體(圖 10c-f)，在震前約三週(第 21 天附近)反而出現了顯著的負異常訊號，此現象在較低規模的統計中並不清晰。這暗示著巨型地震的孕震過程可能更為複雜，其電離層擾動可能呈現出一個「先負後正」的時序特徵。

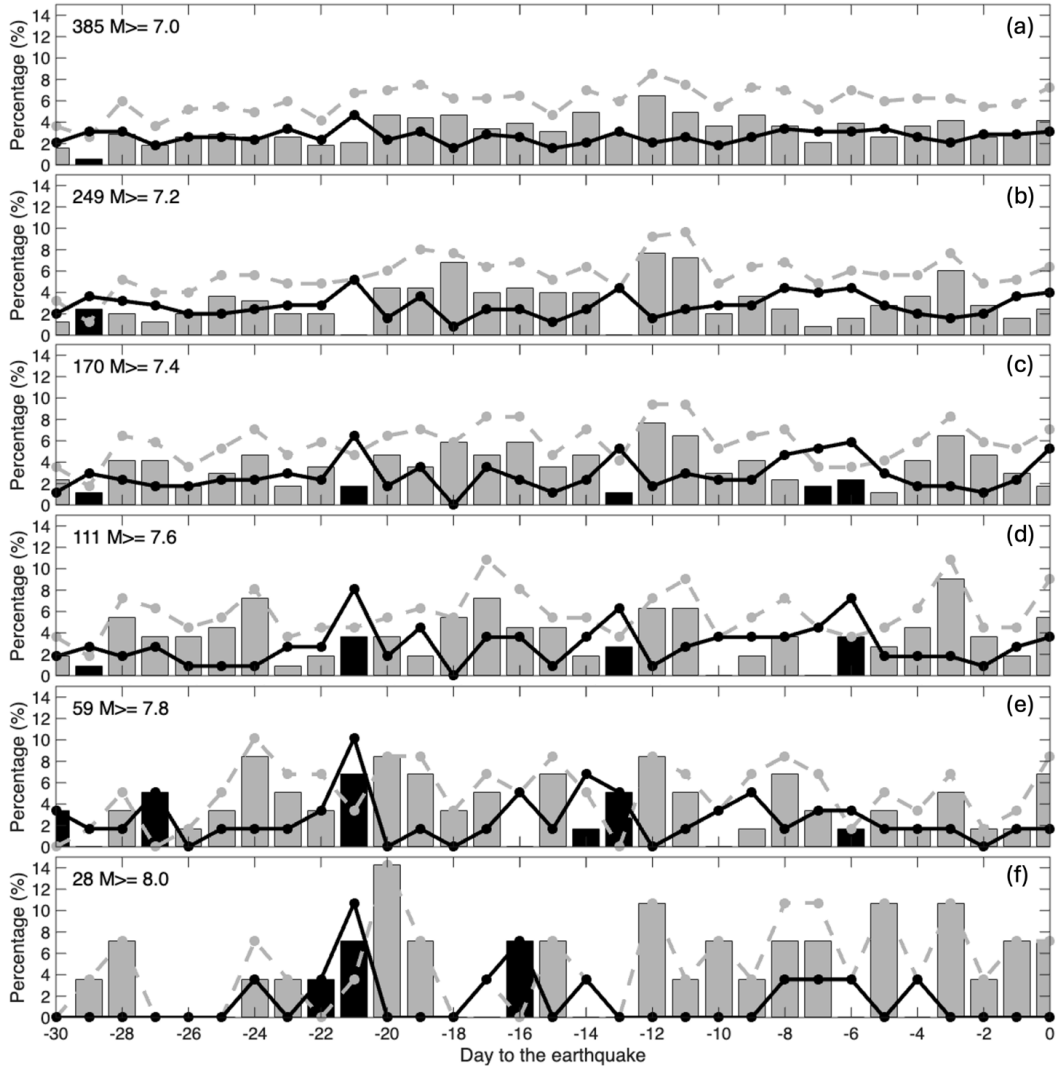


圖 10、全球地震依不同規模門檻分類後，其震前 30 天電離層 TEC 異常發生率之統計分析。(a)至(f)分別對應規模 7.0 以上(385 筆)、7.2 以上(249 筆)、7.4 以上(170 筆)、7.6 以上(111 筆)、7.8 以上(59 筆)及 8.0 以上(28 筆)的地震群體。圖中符號定義與圖 9 相同。

有鑑於電離層的物理特性，特別是電漿的分布與傳輸過程深受地球磁場的控制，因此，震央所在的地磁緯度被認為是影響電離層地震前兆形態的可能因素之一。為了驗證此一假設，本研究首先分析了 385 筆地震樣本在震央地磁緯度上的

分布情形，如圖 11 的統計分布所示。本研究將地震事件劃分為三個區域：(1)低緯度區域：地磁緯度介於正負 30 度內，共計 277 筆地震；(2)中緯度區域：地磁緯度介於正負 30 度至正負 60 度之間，共計 94 筆地震；(3)高緯度區域：地磁緯度大於 60 度，共計 14 筆地震。由分布可見，全球絕大多數的強震(約 96%)均發生於中、低緯度地區。

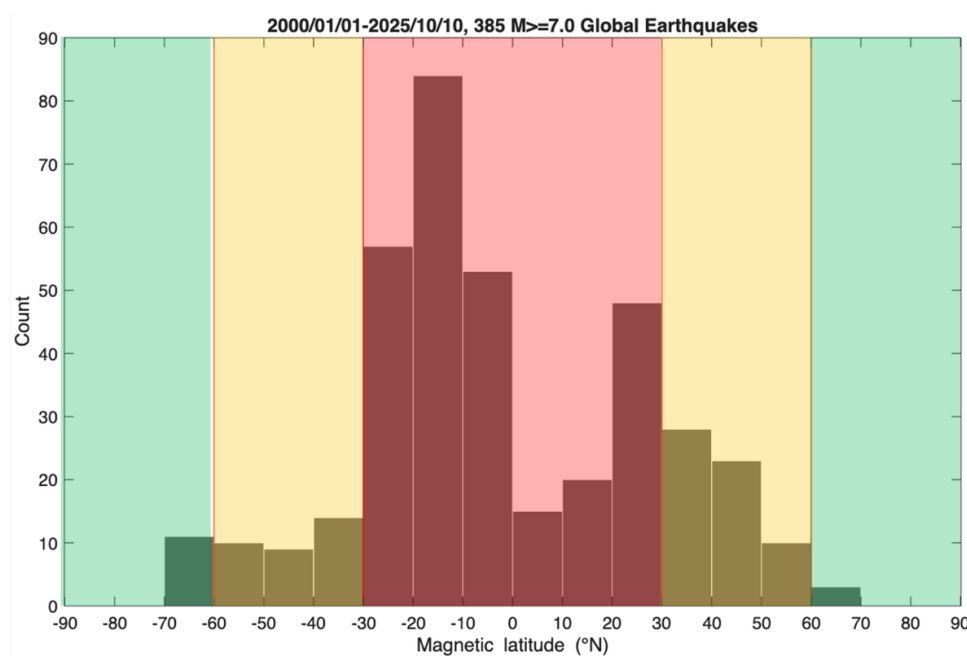


圖 11、全球 385 筆規模 7.0 以上地震之震央地磁緯度分布直方圖。圖中紅色垂直線標示出低緯度(正負 30 度內)、中緯度(正負 30 度至 60 度)與高緯度(大於 60 度)的分界線。紅色區域為低緯度地區、黃色區域為中緯度地區、綠色區域為高緯度地區。

基於此一分類，本研究進一步將電離層異常發生率依不同緯度區域進行獨立統計，以釐清地磁緯度的調控作用，結果如圖 12 所示。分析結果顯示，不同地磁緯度區域的地震前兆特徵存在顯著差異。在低緯度區域中有最多地震筆數，其統計特徵與全球總體結果(圖 9)最為相似，一個顯著的正異常高峰出現在地震前第 12 天附近，顯示這是低緯度地區電離層前兆最主要的特徵時間。在中緯度地區呈現出截然不同的前兆時間特徵，正異常的高峰期門移移至 2 個主要時段：震前第 8 至 11 天，以及在震前第 3e 天出現的另一個高峰，這與低緯度地區的單一高峰模式形成鮮明對比。最後，由於高緯度區域樣本數過少，其統計結果的波動性極大，訊號缺乏穩定性與代表性，因此雖然圖中呈現某些劇烈變化，但難以從中得出具有統計意義的結論。綜合而言，此項分析明確指出，地震前兆的時序特徵具有顯著的緯度依賴性，低緯度地區的前兆主要集中在震前 2 週，而中緯度地區的前兆則可能出現在震前 1 週附近及震前數天。此研究分析結果對於發展分區、乃至客製化的地震前兆監測模型具有重要的指導意義。

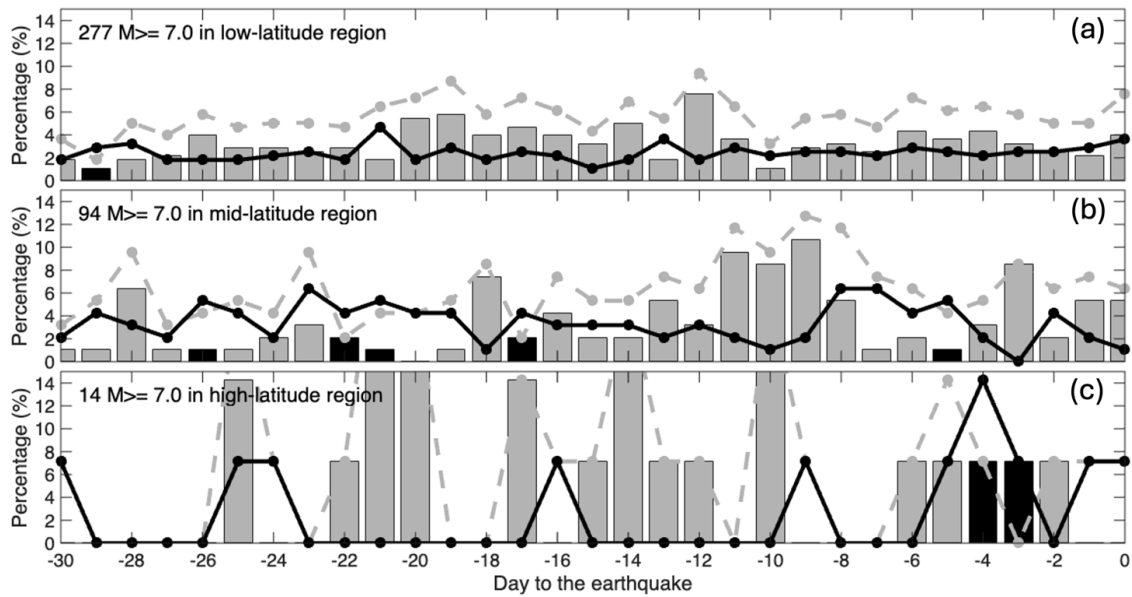


圖 12、依地磁緯度分類之電離層 TEC 異常發生率統計結果。(a)為低緯度區域，有 277 筆地震；(b)為中緯度區域，有 94 筆地震；(c)為高緯度區域，有 14 筆地震。圖中符號定義同圖 9。

電離層的狀態除受地球內部活動影響外，更主要受來自太陽的輻射與粒子活動所主導。為了釐清太陽活動性對地震前兆觀測的影響，本研究利用太陽 10.7 公分波長幅射通量指數(F10.7)作為太陽活動強弱的指標。F10.7 指數與太陽黑子數高度相關，是評估電離層背景狀態的常用參數。本研究的觀測區間(2000-2025 年)完整涵蓋了第 23、24 太陽活動週期，經歷了活動高峰與低谷，因此適合進行此項分類分析。本研究以 $F10.7=120$ sfu (solar flux unit) 為界，將 385 筆地震事件劃分為兩組：(1)發生於低太陽活動期($F10.7<120$)的 238 筆地震，以及(2)發生於高太陽活動期($F10.7\geq 120$)的 147 筆地震。兩組的疊加統計結果如圖 13 所示，其分析揭示了顯著的差異。在太陽背景相對平靜的時期，地震前兆的統計特徵為正異常的發生率高於負異常，並在震前第 20 天與第 12 天附近形成兩個明確的高峰。此結果與全球總體統計(圖 9)的趨勢高度吻合，說明在低背景干擾下，電離層正異常前兆較顯著。而當太陽活動劇烈時，電離層本身的變異性大幅增加，使得地震前兆的特徵變得不同。圖 13b 可看到，原先平靜天的正異常高峰被抑制或模糊化，整體趨勢變得與全球總體統計(圖 9)大不相同。更重要的是，負異常的發生率(黑色實線)也有顯著的提升，在震前約三週(第 20、22 天)甚至出現了負異常主導(黑色長條圖)的情況。這表明在高太陽活動期間，電離層的整體不穩定性可能同時誘發正、負兩種類型的擾動，從而掩蓋了來自地球內部的特定前兆訊號。

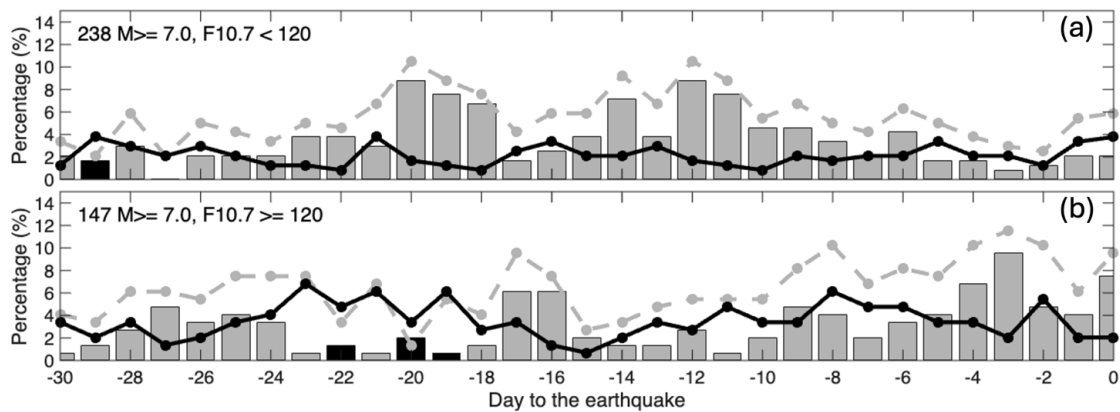


圖 13、依太陽活動性分類之電離層 TEC 異常發生率統計結果。(a)為低太陽活動期 (F10.7<120，238 筆)之結果，(b)為高太陽活動期(F10.7≥120，147 筆)之結果。圖中符號定義同圖 9。

全球地震空間異常分析

前述的疊加分析主要關注每日的異常發生率，為了進一步評估在一段時間內偵測到前兆的總體機率，本研究引入了「連續空間異常」的累積統計方法。此方法的計算步驟如下：首先，依據式(1)計算每筆地震的理論孕震半徑範圍。其次，檢視此範圍內是否存在持續 9 小時以上，且強度超過正負 2 倍標準差的電離層 TEC 異常。若在指定的時間窗口內滿足此條件，則該地震事件被記錄為「偵測到異常」。

下圖 14 展示了基於此方法，對全球 385 筆規模 7.0 以上地震進行的累積分析結果。圖中橫座標為自地震發生當日向前回溯的累積天數(例如 Days=5 代表分析地震前 0-5 天的時段)，縱座標則為在該累積時段內偵測到異常的地震案例所佔的百分比。結果顯示，在所有回溯天數下，正異常的累積偵測率均顯著高於負異常，再次證實了正異常是全球大地震最主要的電離層前兆形態。另外，隨著回溯分析的時間窗口擴大，偵測到異常的地震案例百分比也隨之穩定增長。圖 14a 顯示，當累積分析窗口達到 8 天(即震前 0-8 天)，已有超過 50% 的地震事件在其孕震區內可觀測到持續性的正異常。圖 14b 則統計了在累積時段內偵測到任一類型(正或負)異常的總體偵測率。若同一天出現正、負異常，僅計為 1 次，此結果量化了本方法在一個相對較短的時間窗口內的偵測效能，其重要發現為：當分析窗口僅回溯至震前 4 天時，便已有超過 50% 的地震在其孕震區內出現了可識別的持續性電漿異常現象。

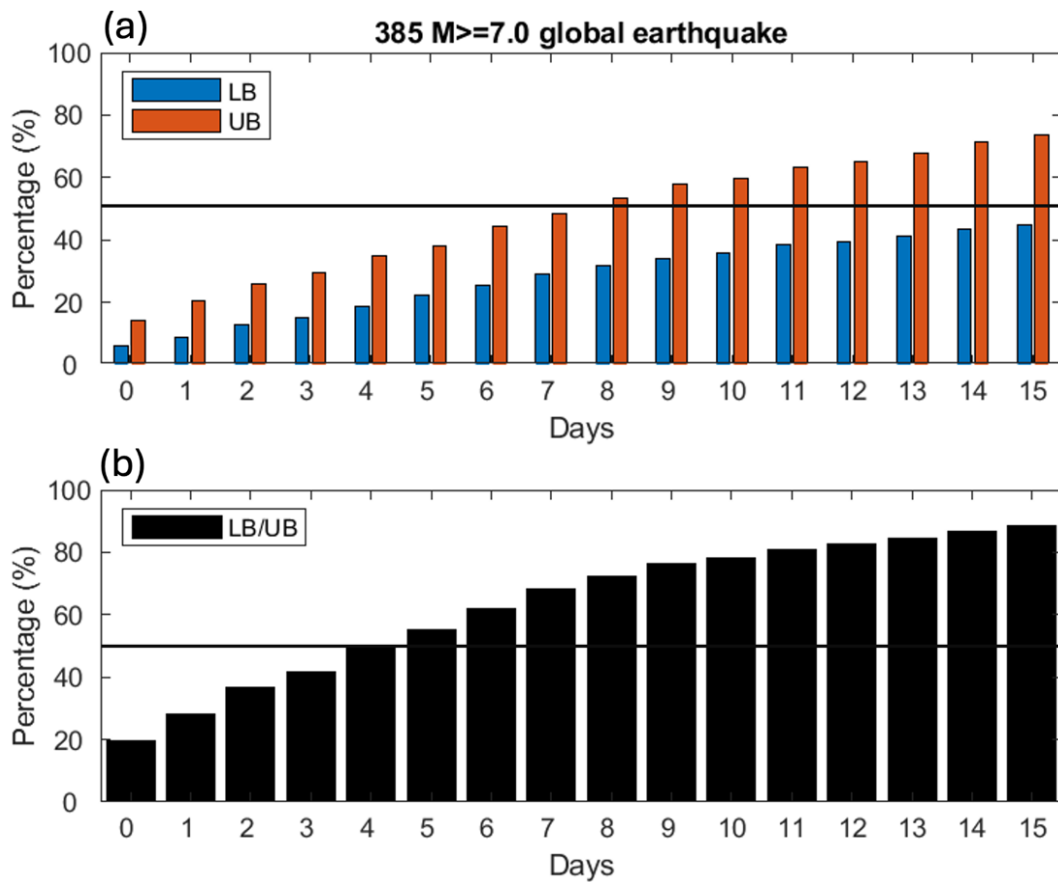


圖 14、全球 385 筆規模 7.0 以上地震，其震前孕震區內連續空間異常之累積偵測率統計。橫座標為自震前回溯的累積天數。異常定差為在孕震區內出現持續 9 小時以上、強度超過 2 倍標準差的 TEC 擾動。(a)分別統計正異常(UB)與負異常(LB)的累積偵測率。(b)統計任一類型異常(LB/UB)的總體累積偵測率。黑色橫線表示 50%。

為了系統性的建立地震規模、前兆觀測時間窗口與偵測率之間的量化關係，本研究進一步將前述的連續空間異常累積分析結果，依地震規模進行分類，並將其視覺化呈現，如圖 15 所示。圖中的橫座標為自地震前回溯的累積天數，縱座標為地震規模，而顏色則代表在對應規模與累積天數條件下，偵測到任一類型(正或負)連續空間異常的地震案例百分比。而圖中黑色等值線標示出了特定的偵測率：50%、75%及 90%。圖 15 呈現了一個趨勢：要達到相同的偵測率，地震規模越大，所需要的回溯觀測時間則越短。由 50%等值線的走向可以看到，對於規模 7.0 以上的地震(共 385 筆)來說，需要回溯約 4 天才能達到 50%的偵測率，但對於規模 7.6 以上的地震(共 111 筆)則僅需回溯約 2 天即可達到相同的偵測率。同樣的，由 75%偵測率等值線可以看到，對於規模 7.2 以上的地震(共 249 筆)來說，需要長達 7-8 天的回溯窗口才能涵蓋 75%的案例，而對於規模 8.0 以上的地震(共 28 筆)則僅需約 4-5 天的回溯期。上述結果表示，地震規模越大，其電離層前兆不僅更為顯著，且在震前更短的時間窗口內即可達到較高的偵測率。也就是說，越是具有破壞性的強震，其前兆訊號在孕震區內的可觀測性與即時性也越高。

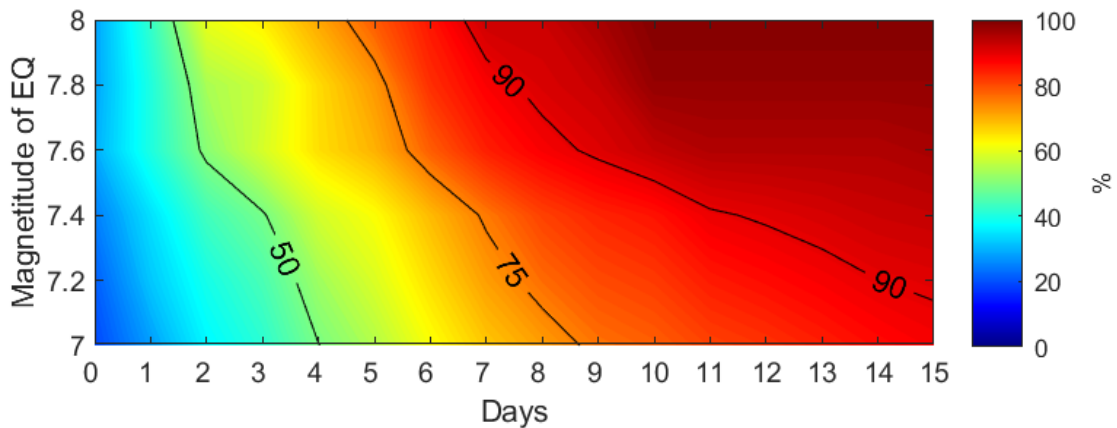


圖 15、地震規模與累積天數對連續空間異常偵測率之影響分析圖。橫座標為地震前回溯累積天數，縱座標為地震規模，顏色表示在對應條件下的累積偵測率(%)。圖中黑色等值線標示出特定的偵測率(50%、75%、90%)。

接著進一步探討地磁緯度對連續空間異常累積偵測率之間的關係，本研究依據討論「全球地震時間異常分析」時的緯度區域定義(圖 11)進行地震樣本分類，分別進行累積統計分析。結果如下圖 16 所示。圖中分別繪製了低地磁緯度(紅色實線)、中地磁緯度(藍色實線)與高地磁緯度(黑色實線)地區的累積偵測率曲線。可以看到，中、低地磁緯度地區的地震(紅色與藍色實線)均呈現出一個共同趨勢，隨著回溯累積天數的增加，偵測到連續空間異常的地震案例百分比也穩定上升。然而，兩者之間也存在著差異，若以 50% 的偵測率(圖中黑色虛線)作為一個評估效能的基準點來看，低地磁緯度地區地震的累積偵測率曲線(紅色)的上升斜率較為明顯，大約在回溯 3-4 天的窗口即可達到 50% 的偵測率。而位於中地磁緯度地區的地震而言，其偵測率曲線(藍色)則相對平緩，需要約 5 天的回溯窗口才能達到相同 50% 的偵測率。高地磁緯度地區的地震由於樣本數過少，其曲線呈現階梯狀，不具備足夠的統計穩定性，故在此不作深入討論。藉由上述的比較可以看到，對於低地磁緯度地區的地震，其連續空間異常前兆達到一個固定偵測率所需要的時間窗口較短，這可能意味著低地磁緯度地區的地震前電離層耦合機制更為直接或強烈，使得可觀測到的異常訊號更集中出現在臨震前的數日內。

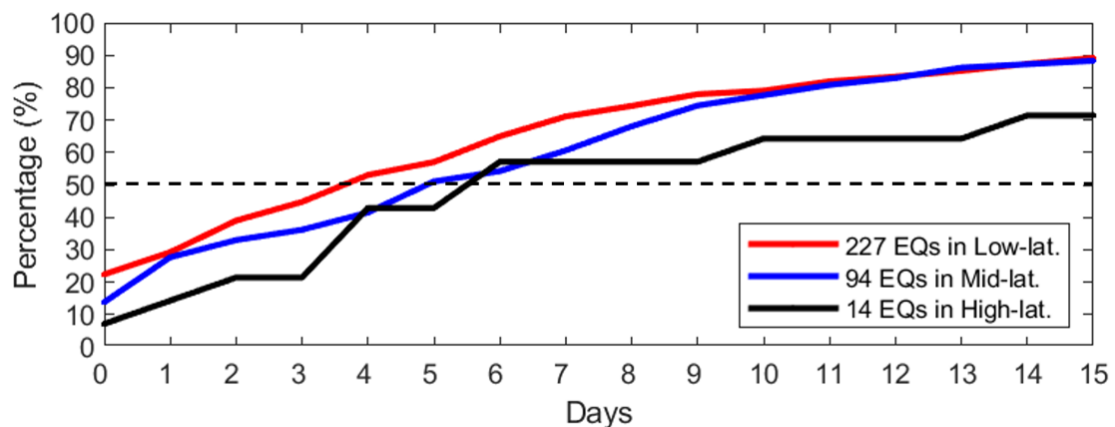


圖 16、依地磁緯度分類之連續空間異常累積偵測率比較圖。圖中紅色、藍色與黑

色實線分別代表位於低地磁緯度、中地磁緯度與高地磁緯度地區的地震，其在不同回溯累積天數下的偵測率。圖中黑色虛線標示 50% 偵測率。

最後，本研究將連續空間異常的累積偵測率與太陽活動性進行交叉分析，評估在不同背景條件下的觀測效能。如下圖 17 所示，圖中紅色與藍色實線分別代表發生於高太陽活動期($F_{10.7} \geq 120$)與低太陽活動期($F_{10.7} < 120$)的地震，其連續空間異常的累積偵測率，黑色虛線為 50% 的標示線。結果呈現出在高太陽活動期間，累積偵測率(紅色實線)在整個觀測窗口內均顯著高於低太陽活動期間(藍色實線)。若以 50% 偵測率為基準，高太陽活動期的地震僅需回溯約 2-3 天即可達到，而低太陽活動期的地震則需要約 5-6 天。這個結果並不意味著高太陽活動期更有利於前兆觀測，相反的極可能反映了電離層背景雜訊的增加。在高太陽活動期，由於太陽閃焰、地磁擾動等引發的電離層劇烈變化更為頻繁，這些非地震源的強烈擾動同樣會被本研究的正負 2 倍標準差的異常區間被視為異常，從而導致一個較高的偵測率。因此，雖然偵測率數值較高，但其中訊號的訊噪比卻可能更低，即訊號的可靠性與獨特性反而下降了。

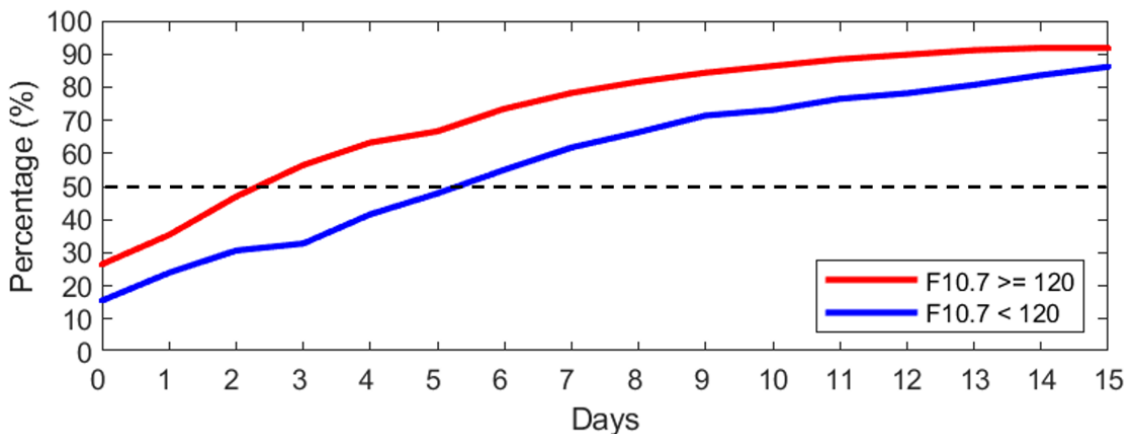


圖 17、依太陽活動性分類之連續空間異常累積偵測率比較圖。圖中紅色實線代表高太陽活動期($F_{10.7} \geq 120$)，藍色實線代表低太陽活動期($F_{10.7} < 120$)，黑色虛線標示 50% 偵測率。

肆、結論與建議

目前的研究成果顯示，在數據處理與分析工具的建置上，已經完成一套能夠自動化處理全球 GIM TEC 資料並偵測異常區域的程式，並在氣象署的遠端電腦上完成安裝與測試，即將進入排程觀測階段。這套程式能處理來自 CODE 等資料中心的 GIM 資料，並能將全球資料內插至特定位置進行比對分析。本研究透過比對臺灣區域性 GNSS TEC 資料與全球 GIM TEC 資料，證實兩者在描述電離層 TEC 的變化與異常日期上的一致性與互補性。全球 GIM 資料提供了廣闊的空間背景，而區域性 GNSS 資料則提供了精確的時間與空間解析度資料。

然而，在電離層異常的具體判斷上，本研究透過 2024 年花蓮地震的個案分析，

發現傳統判斷方法對參數設定(如背景參考天數或異常閾值)具有高度的敏感性，僅是調整設定就可能導致對同一事件的異常判斷結果截然不同。這強烈提示過去依賴特定統計經驗設定的方法，可能無法普適於所有地震事件和不同地區的電離層特性。為了建立更具普適性的判斷標準，本研究進一步建立了以 16 天為參考週期(對應電離層準 16 天週期)、以 2 倍標準差作為異常邊界的新標準，並對全球 385 筆規模 7.0 以上的地震進行了大規模統計分析。

在時間異常的分析上，本研究針對全球 385 筆強震前每日異常發生率進行了疊加統計，結果顯示，在整個震前 30 天的觀測窗口內，正異常的發生率幾乎全面的高於負異常，確立了正異常為全球強震最主要的電離層前兆形態，並且在地震前第 12 天與第 20 天附近最為集中。更深入的分類統計顯示，此時間特徵並非不變，而是受到多重物理因素的影響。地震規模越大，前兆訊號越趨向於臨震前出現，例如規模 7.6 以上地震在地震前第 3 天有顯著正異常。地震發生所在的地磁緯度也會影響電離層異常時序，低緯度地區集中在地震前第 12 天，中緯度地區則轉移至地震前第 8-11 天與第 3 天。此外，太陽活動的影響也十分關鍵，在低太陽活動期，電離層前兆訊號清晰可辨，但在高太陽活動期則受到顯著干擾與抑制。

在空間異常的分析上，本研究採用連續空間異常的累積統計方法，評估在一段時間內於孕震區內偵測到持續性異常的總體機率。分析結果再次證實了正異常的主導性，並量化了偵測效能，在回溯 8 天的窗口內，有超過 50% 的地震可偵測到持續性正異常。此空間異常的偵測率同樣展現出強烈的依賴性。地震規模越大，達到相同偵測率(如 50%)所需的回溯時間越短，顯示強震的電離層前兆在空間上的可觀測性與即時性更高。地磁緯度的效應也十分明顯，低緯度地區達到 50% 偵測率所需的時間窗口(約 4-5 天)顯著短於中緯度地區(約 6 天)。而在太陽活動性的影響上，高太陽活動期的偵測率雖然較高，但這極可能反映了太空天氣事件造成的背景雜訊增加，反而降低了電離層異常訊號的訊噪比與可靠性。

總結來說，本研究在電離層地震前兆的資料處理工具開發、全球與區域資料的整合比對、以及建立新標準化分析方法上取得了階段性成果。透過大規模統計，不僅顯示了電離層前兆訊號的總體特徵，更說明了其受地震規模、地磁緯度與太陽活動性等因素的影響。基於目前的進度與分析結果，本研究顯示電離層地震前兆在「時間」與「空間」維度上各自的特徵與規律。然而，將兩者獨立分析仍有其局限性，因此結合電離層前兆的時間與空間特徵，並透過兩者之間的相互驗證與疊合比對，將有機會更全面、更可靠的識別出潛在的地震電離層前兆區域與時間窗口，並有效降低單一指標可能造成的誤判。

為探討如何進行兩者的結合，並量化分析兩者之間的複雜關係，本研究接下來預計導入人工智慧與機器學習模型。此模型將不再依賴單一統計閾值，而是能夠學習時間異常(如地震前特定天數的發生率)與空間異常(如孕震區內異常的持續時間、範圍與強度)之間的關聯模式。模式將整合地震規模、地磁緯度、太陽活動性等多維度參數作為輸入特徵，透過訓練大量的歷史數據，自主尋找出高可信度前兆的時空資訊。最終目標是建立一個能夠客觀評估潛在異常事件是否為地震前兆的機率性預警框架，期望藉此將電離層前兆的監測從現象的統計描述，推進至更具實用價值的預測性科學研究階段。

伍、成果的價值與貢獻、落實應用情形

全球 GIM TEC 解算程式已於氣象署遠端電腦上進行安裝、設置與測試完成，接下來將進行排程觀測與產出全球電離層異常區域偵測資料。

參考文獻

Dobrovolsky, I. R., S. I. Zubkov, and V. I. Myachkin (1979), Estimation of the size of earthquake preparation zones, *Pure Appl. Geophys.*, 117, 1025-1044.