

高精度衛星遙測土壤含水量系統建立

Establishment of a Highly Precise Satellite Remote
Measurement System for Soil Moisture Content

創新研究計畫
成果報告書



主辦機關：行政院農業委員會水土保持局

執行機關：國立臺灣大學

中華民國 111 年 12 月

高精度衛星遙測土壤含水量系統建立

Establishment of a Highly Precise Satellite Remote
Measurement System for Soil Moisture Content

創新研究計畫

成果報告書

主辦機關：行政院農業委員會水土保持局

執行機關：國立臺灣大學

目 錄

表目錄	表-1
圖目錄	圖-1
摘要	摘-1
第一章 計畫背景	1-1
一、計畫動機	1-1
二、計畫背景	1-2
三、計畫目標	1-5
第二章 計畫架構	2-1
一、研究開發項目	2-1
二、研究方法	2-2
三、研究執行步驟	2-3
第三章 研究進度及成果	3-1
一、衛星融合技術驗證	3-1
二、土壤含水量推估驗證	3-7
第四章 研究貢獻與侷限性	4-1
一、討論與建議	4-1
二、成果貢獻	4-2
三、未來展望	4-4
參考文獻	參-1
附錄一 審查意見回覆	附 1-1

表 目 錄

2-1	使用測站之地文資訊	2-31
2-2	各土石流測站所使用之(左)雨量與(右)濕度的自動測站 .	2-31
2-3	土石流監測站基本資訊	2-34
2-4	土壤含水量測量的時空間差異	2-34
3-1	同質區域的二度分帶座標邊界框	3-1
3-2	同質區域的模型表現性衡量指標	3-4
3-3	目標土石流觀測站的關注區	3-8
3-4	特徵選擇結果	3-9

圖 目 錄

1-1	每日土壤含水量與每日 CWSI 峰值相關性	1-5
2-1	計畫流程架構圖	2-3
2-2	衛星影像融合推估算法架構	2-10
2-3	MODIS 衛星遙測演算產品資料庫服務畫面	2-17
2-4	機器學習模型概念圖	2-23
2-5	機器學習模式建構示意圖	2-27
2-6	k 折交叉驗證示意圖 (k=4)	2-27
2-7	使用土石流測站分布圖	2-28
2-8	使用土石流測站之含水量觀測時間序列	2-29
2-9	中台灣坡度空間分布	2-29
2-10	中台灣坡向空間分布	2-30
2-11	中台灣高程空間分布	2-30
2-12	小時累積降雨與對應測站之土壤含水量比較圖	2-32
2-13	12 小時前平均濕度與對應測站之土壤含水量比較圖	2-32
2-14	土石流觀測站空間分布圖	2-35
3-1	農田 (a) 以及邊坡 (b) 的預測地表反射值與實際 Landsat 觀測值之散布圖	3-3
3-2	丘陵地 (a)、山區 (b) 和草原 (c) 的預測地表反射值與實際 Landsat 觀測值之散布圖	3-3
3-3	農田的 HISTARFM 推估結果	3-4
3-4	邊坡的 HISTARFM 推估結果	3-5
3-5	高山的 HISTARFM 推估結果	3-5
3-6	丘陵地的 HISTARFM 推估結果	3-6
3-7	草原的 HISTARFM 推估結果	3-6
3-8	所有觀測區域的誤差協方差	3-7
3-9	本計畫建構機器學習模型流程圖	3-9
3-10	以 CWSI 與高程為輸入特徵之模型驗證結果	3-9

4-1	實際與卡曼濾波器最佳即時推估產製之預測地表反射值 (NDVI) 之散布圖	4-2
-----	---	-----

摘要

前言

土石流的發生與臨前土壤含水量息息相關。透過土壤含水量的觀測儀，安裝儀器之敏感山坡地的土壤含水量可被即時追蹤，但是卻無法提供更廣泛的空間資訊。衛星遙測的應用，可以提供同時具有時間和空間特性的土壤含水量指標，不過同時也受限於雲遮蔽和時空間解析度不足的問題，在臺灣山坡地形變化較為劇烈的地區尚不能達到應用的標準。

Highly scalable temporal adaptive reflectance fusion model (HISTARFM) 是一種衛星影像融合技術，可將不同空間與時間解析度的衛星影像利用統計方法進行融合，產出同時具有高空間解析度以及高觀測頻率的衛星影像，特別針對遙測影像因為雲遮蔽影響的平滑化與補遺。此融合影像的高時空間解析度，對於土壤含水量的推估以至於土石流的災害防範預警，能夠有顯著的增進。

為了達到偵測與監控的目的，自動化土壤含水量的模式需要被開發。作物水分逆境 (Crop Water Stress Index, CWSI) 的方法建立於植物根部與土壤含水量的互動，可以透過植被的蒸發散量表現，判斷土壤的水分含量，該指數以地面熱平衡方程式 (Surface Energy Balance Algorithm for Land, SEBAL) 推估衍生而至。在過去 SEBAL 模型需要人工選取錨點像素才可進行運算，而 PySEBAL 模式使其能自動化選取，在達成自動化土石流預警又更近一步。

本計畫目標將開發衛星融合的技術以及自動化土壤含水量推估方法。結合兩者，基於高時空間解析度遙測影像，臺灣山坡地的土壤含水量能夠有更高頻率和更高解析度的空間觀測資料。

研究方法

本計畫系統架構主要兩個主要模型的開發，資料蒐集，以及兩個模型的驗證。模型一為衛星融合技術，以 Landsat 衛星和 MODIS 衛星兩大開源衛星作為遙測資料輸入，以 HISTARFM 技術作為核

心，開發在臺灣地區的衛星融合技術。驗證方法則採用同質小區域的人工隨機挖空後補遺的方法，驗證評估該方法的精確性和準確性。模型二為土壤含水量推估系統，一樣以 Landsat 衛星和 MODIS 衛星作為輸入，加入由臺灣氣候變遷推估資訊與調適知識平台 (Taiwan Climate Change Projection Information and Adaptation Knowledge Platform, TCCIP) 的氣象資料進行 PySEBAL 自動地表蒸發散量的推估，後衍生推導 CWSI 的數值圖層，作為土壤含水量的指標。驗證方法為選取目標土石流測站的該點為的 CWSI 點資料以及該測站的土壤含水量資料，以 2019 年全年，評估兩者的相關性和趨勢。

資料的蒐集包括 Landsat 8-9 以及 MODIS 的地表反射率、返照率以及地表溫度，來自 TCCIP 的氣象資料面資料 (氣溫、風速、濕度以及日射量)，和來自水保局土石流觀測站的土壤含水量資料。

研究成果與討論

衛星融合資料推估選定了五個同質小區域，分別是高山 (玉山群峰)、丘陵地 (鹿谷鄉)、邊坡 (古坑鄉)、草原 (墾丁龍磐草原)、農地 (嘉南平原)。透過 HISTARFM 的卡曼濾波器，產出最佳即時推估的結果，再透過具有時序性的偏差修正取得最終推估成果。高山區域使用卡曼濾波器作為最佳即時推估的結果最佳， $RMSE=0.0659$ 、 $MAE=0.0501$ 、 $MBE=0.0423$ 、 $MSE=0.00434$ 以及 $r=0.955$ 。相反地，農田在目前模型中的推估效果最差 $RMSE=0.209$ 、 $MAE=0.165$ 、 $MBE=0.0722$ 、 $MSE=0.0437$ 以及 $r=0.0949$ 。此外，針對補遺區域的數值推估，本研究擷取最後更新的誤差協方差 W_k ，作為判定不確定性的依據。當中農田區域中有較多不確定性較高的區域，在其他地區則有少數區域有高不確定性。

在農地的補遺精準度和準確性皆欠佳，該原因為衛星融合技術中尚無法偵測過於頻繁變動的區域。為了若要將此衛星融合技術應用於高變動的區域，貝氏模型平均的優化、模型中 γ 的動態調整以及偏差修正的優化仍然需要進行進一步的研究。

土壤含水量推估中，本研究分別採用 MODIS 以及 Landsat 的衛星資料，再輔以 TCCIP 的氣象資料，推估出地表的各蒸散量的面資

料。藉由蒸發散量資訊，本研究先推估出個別衛星的作物水分逆境指數 CWSI，而後使用 HISTARFM 進行衛星融合產出每日高精度的 CWSI 圖層。選定 7 座土石流測站進行最後推估的驗證，分別為上安、神木、豐丘、九份二山、坪頂、南豐以及來義。推估結果顯示，測站土壤含水量變化和 CWSI 指標並無顯著的相關性。

不顯著的原因，可能由兩個因素造成。一是 HISTARFM 的結果中，針對每個像素皆有推估的不確定性。本研究則進一步將資料不確定性作為權重放入相關性的討論之中，發現加入像素點不確定性之後，兩者的相關性有趨向理論上呈現的相關性，在豐丘、來義、南豐皆由原本的無相關或正相關修正至負相關。二是可能來自 CWSI 指標和測站土壤含水量儀偵測的含水量深度所導致。

結論

本計畫開發了衛星融合的演算法以及自動化的植被蒸發散推估。在融合演算法中，本研究計畫完善了個別衛星資料的資料庫建置，和各項地理空間資料的存取及運算功能。在土壤含水量推估中，CWSI 雖然表現上與土壤含水量儀在相同地點監測的數據相關性結果欠佳，但 CWSI 仍在淺層土壤含水量推估上具有潛力。往後可以與前人研究，例如合成孔徑雷達 (Synthetic aperture radar, SAR)、觀測站附近的地貌以及土壤質地等因素共同來預測現地的土壤含水量，也可透過增設不同深度的土壤含水量儀，更精確的指出 CWSI 適合監測的土壤含水量深度。

第一章 計畫背景

一、計畫動機

土壤含水量，是土壤中所含水分重與土粒重之比。測定土壤含水量可以掌握作物對水的需求狀況，對農業生產是重要的指標之一。除了農業之外，該指標亦有其他的用途，過去的研究指出此指標對於山坡地的水土保持上有重大的意義，可以藉由此指標來判斷土石流發生的可能性。形成土石流的三大基本條件為豐富的堆積物、陡峻的坡度與充足的水源。前兩個誘發條件較不易受到環境的變動而產生劇烈的變化，然而土壤水源的變動是快速且劇烈的，因此也成為土石流災害監測的關鍵。過去，土石流預警研究多利用雨量資料做為大範圍區域的警戒基準參數，目前最常使用的為降雨強度與有效累積降雨量，該指標雖可以提供土石流發生的預警，但是與真正土石流發生的關聯性尚隔了一層關係。

一定強度的降雨雖為造成土石流的最大原因，但強降雨若配合不同的臨前土壤含水量情形，將導致不一樣的結果，臨前土壤含水量的多寡，作為先決條件，將一定影響土石流發生的機率在強降雨來臨之前。過去的研究指出，若臨前含水量過高（>5%），較不易導致土石流的發生，由於多於的水分不易進入土層而造成岩床與堆積物之間滑動；若臨前含水量較低時，降雨來臨時入滲速度比臨前含水量較高時更快，更有機會造成岩床與堆積物之間滑動；若土壤趨於乾燥狀態，降水滲入的速度與量也會較低，情況和臨前含水量高時雖然狀況不同但卻導致相同的結果。

退一步來說，地表逕流的產生和土壤含水量息息相關。國內學者專家也提到，土石流的引發會依據各種不同環境條件而決定，若土體顆粒細、疏鬆、含水量較高且具有較陡地形時，僅需少量的降水即可能引發土石流。由於土質和地形短時間不會有劇烈變動，故除了降雨量之外，臨前土壤含水量也同為土石流啟動機制之關鍵因子，若能即時掌握各地區土壤含水量情形，將大幅提高對於土石流預警的準確度。

二、計畫背景

(一) 衛星遙測影像應用於土壤含水量推估

目前對於土壤含水量的監測相當有限，考慮到土壤含水量觀測儀的設備及設置的人力成本，該儀器大多佈設於已知的潛勢溪流集水區內，對於土壤含水量觀測之空間區域相對侷限，無法對於所有敏感山坡地區域之土壤含水量進行有效之監控。為解決此一問題，近年許多國內外研究開始在探討透過衛星遙測資料對於土壤含水量進行評估，如能應用遙測大面積之影像資訊進行適當的處理，即可獲得大範圍且即時之資訊，除節省觀測儀設備建置及維護的人力成本，也能對大範圍的土壤含水量進行更全面的監控及評估，提升土石流預警效果。例如，陳錕山在 2009 年即利用 MODIS 衛星資料演算台灣全區之土壤含水量 (交通部中央氣象局, 2009)。彭文飛在 2010 年融合不同解析度之 SMAP、MODIS、SPOT、Sentinel-1 衛星觀測資料與降雨資料，對土壤含水量進行高精度預測模式建構 (行政院農業委員會水土保持局, 2010)。2012 年陳繼藩利用 MODIS 與 AMSR-E 遙測資料透過地表溫度與常態植生標指推估地表土壤含水量，結果顯示依照高度分區作法產生之成果比不分區之方法有更佳的精確度 (行政院國家科學委員會, 2012)。

地表植被的蒸發散量結合地表溫度和植生指數，被認為可獲得更加完整的土壤含水量資訊。水保局於 2016 年的時空巨量資料推估技術於坡地災害風險預警之發展應用報告中，即是使用地表植被蒸發散量作為土壤含水量的推估依據，藉由植物潛勢蒸發散量計算出作物水分逆境指數 (Crop water stress index, CWSI)，該指數被證實與土壤含水量的多寡呈現一負線性相關性，透過土石流觀測站的土壤含水量做驗證，可以推估取得空間上各地點的土壤含水量資訊。不過此報告中，為求土石流的災害預警，用於推估土壤含水量的衛星使用 MODIS 的產品，該衛星最大空間解析度達到 500 公尺，對於變化較小的區域較無顯著影響，但由於臺灣山坡地形變化較為劇烈，該衛星所提供的空間解析度尚不能達到應用的標準，結果也顯示該遙測推估數值無法監測到劇烈變

化的山坡地形的土壤含水量變化。

(二) 衛星融合技術對於遙測資料的增進

以上的推估方法尚無法建立良好的土石流災害預警系統，原因並非土壤含水量遙測推估的方法，而是受限於單一衛星遙測衛星影像的不足的時空間解析度。

土壤含水量的觀測儀在收取資料的時間尺度可以達到每五分鐘一筆，該方法雖然可以即時的追蹤土壤含水量的變化，但是卻無法提供更廣泛的空間資訊。衛星遙測上，MODIS 衛星提供每日資料，空間解析度達到 500 公尺；日本的向日葵衛星在臺灣地區具有每十分鐘一筆資料，空間解析度為 1000 公尺；Sentinel 衛星具有五天一張的衛星影像，空間解析度達到 10 公尺；Landsat 衛星則是八天一張，達到的空間解析度為 30 公尺；SPOT 衛星 26 天才會完成一次繞行，該空間解析度達到六公尺。不同衛星在時間和空間有所權衡，低空間解析度使山坡地的土壤含水量推估較為失準，而低時間解析度則沒有辦法即時提供土石流的災害預警。

此外，遙測資料的缺漏也是一個問題，由於衛星影像普遍受氣象雲的干擾，在陰雨天時，地表反射將受到雲的阻擋而導致遙測資料的缺漏，而無法推算地面上土壤含水量，若該觀測點位在有衛星影像觀測時皆受到雲遮蔽，則此處可以用於判斷土石流發生的資訊需要追溯到上次具有影像的時間點，對於推估的結果顯而易見地將有較大的偏差產生。

為了能遙測上時空間解析度以及補強資料缺漏的問題，國外發展了許多不同的衛星遙測影像融合技術，由於遙測影像中的空間網格解析度和時間觀測頻率互為權衡關係，擁有細緻的空間解析度衛星影像，其時間的觀測頻率便較長，反之，觀測頻率高的衛星則有著較低的空間解析度。自適應時空反射率融合模式 (Spatial and temporal adaptive reflectance fusion model, STARFM) 可將不同空間與時間解析度的衛星影像利用統計方法進行融合，產出同時具有高空間解析度以及高觀測頻率的衛星影像，此理論發展的也相當成熟，同時透過高頻率的衛星影像，也可以大大降低

雲遮蔽所帶來觀測值缺失的情況 (Gao et al., 2006; Wang & Huang, 2017)。

Moreno-Martinez (Moreno-Martínez et al., 2020) 的研究中，提出了較 STARFM 模式來得可靠的資料融合方法，在衛星影像的資料融合上，最大的困難點是需要對於因為雲層遮蔽所造成的缺失值進行補遺，而補遺的不確定性也就成為了資料融合後遙測影像品質好壞的關鍵點。Moreno-Martínez 等人提出了較過去多數人使用的自適應時空反射率融合模式來得可靠的資料融合方法，稱為 Highly scalable temporal adaptive reflectance fusion model (HISTARFM)，特別針對遙測影像因為雲遮蔽影響的平滑化與補遺 (Moreno-Martínez et al., 2020)。該方法運用衛星充足歷史資料的特性，對於低空間精度的衛星資料降尺度上有良好的表現，也具有較好的雲補遺能力。藉此可以同時提升衛星資料的時空間頻率，也能產出沒有雲遮蔽干擾的衛星影像，對於土壤含水量的推估以至於土石流的災害防範預警，能夠有顯著的增進。

(三) 其他環境參數對土壤含水量預測之影響

CWSI 的方法建立於植物根部與土壤含水量的互動，前人研究指出 CWSI 和 0-40cm 的土壤含水量有可靠的線性關係，且 CWSI 在土壤含水量較低時較佳，由於植株在水份稀缺時對水份變化是較為敏感的，圖 1-1 顯示 82% 的土壤含水量變異可被 CWSI 解釋 ($R=0.82$) 在 CWSI(0.6-1) 之間 (Mangus et al., 2016; Ru et al., 2020)

除了 CWSI，許多環境因子也和土壤含水量有一定的相關性。降雨量的多寡亦會使植生指數、地表溫度產生改變，其直接影響的因素即是降雨的入滲量，影響入滲的相關因子為土壤性質 (土壤種類、土壤本體含水量、土壤臨前水文條件)、地表覆蓋、地形以及降雨特徵 (降雨強度) 相關 (吳怡瑩，劉哲欣，張志新, 2013)。坡度、坡向以及高程也會影響降雨入滲量，則間接影響地表穩度與植生指數，進而使土壤含水量推估產生差異。除此之外，地貌、地溫以及濕度對於土壤含水量亦有影響，對於預測土壤含水量有一定的貢獻 (行政院農業委員會水土保持局, 2010)。

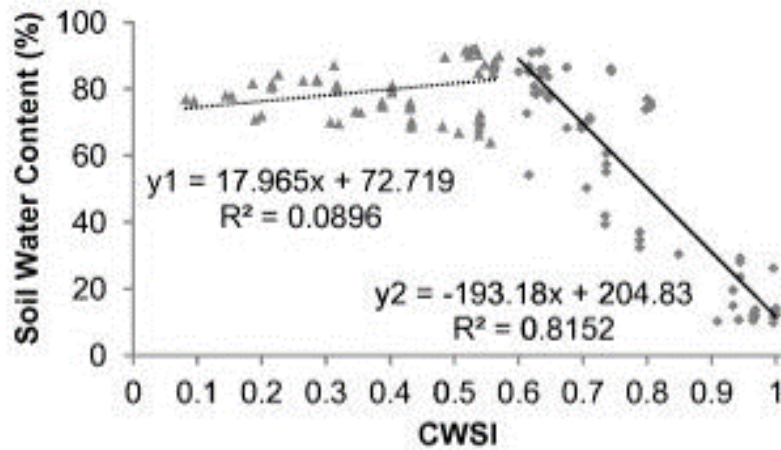


圖 1-1: 每日土壤含水量與每日 CWSI 峰值相關性

三、計畫目標

坡地的土石流預警，需要降雨資料作為即時的判斷依據，同時也需要該地的臨前土壤含水量資訊來達成更加精準的決策。過去使用土壤含水量觀測儀做土壤含水量的收集，但無法遍及所有敏感山坡地，應用遙測影像推估方法因此發展。此外，雖然許多環境因子與臨前土壤含水量皆有關聯，但是開發透過植被狀態來動態推估土壤含水量，有助於往後加入其他修正因子已更高時空間精度的方式來設計自然山坡地的土壤含水量推估。

本創新計畫將著重於精進提升單一指標對土壤含水量推估的時空間精度，以植被的實際蒸散量來計算作物水分逆境指數 (Crop water stress index, CWSI)，藉此推估土壤含水量。提升時空間精度方面，本創新計畫將應用最新方法的衛星影像融合技術-HISTARFM，並且將其進行修改。該技術被證實具有高精度且沒有嚴重偏差，不過該方法在地表資訊的推估上假設地表反射數值遵守高斯分佈，本計畫將修改此假設，並且求出地表反射數值的機率密度函數 (Probability density Function, PDF)，根據分佈特性可以考慮更多資訊取得較佳的推估值。

將衛星融合的技術與土壤含水量推估方法結合，基於高時空間解析度遙測影像，可以針對山坡地的土壤含水量進行更高頻率和更高解析度的觀測。此外，若再加入其他環境參數模擬現地情況，以

此修正土壤含水量使之因地制宜，更進一步提昇預測的精準度。該計畫願景是希望增進臺灣山坡地的監控與偵測，進而達到即時防範土石流災害，減輕天然災害發生對民生的影響。

第二章 計畫架構

一、研究開發項目

(一) 資料蒐集

分別蒐集研發所需要的各項資料，本計畫主要由三處開源資料庫取得資料。(1) 衛星資料：由於本計畫使用衛星融合演算法，故需要分別從 MODIS 以及 Landsat 兩衛星之的開源資料庫取得資料；(2) 氣象資料：氣象資料將臺灣氣候變遷推估與資訊平台計畫 (Taiwan Climate Change Projection and Information Platform, TCCIP) 取得歷史氣象資料，同時也從臺灣之中央氣象局觀測資料平台蒐集所需要的氣象資料；(3) 土石流監測站：土石流監測站提供土壤含水量資料將作為研究計畫驗證之用。

(二) 異元衛星融合算法

模組研發的第二步驟是衛星遙測影像資料融合模組研發，本計畫會應用目前最新的衛星影像融合算法-HISTARFM，並且對其推估器進行再現與修改，以貝氏最大熵法推估 (Bayesian maximum entropy, BME) 取代原本的卡曼濾波器 (Kalman filter)，旨提升估計之準確性，並且對於二者在臺灣各類型地貌的推估進行評估比較。

(三) 土壤含水量推估模型

第三步會開發土壤含水量估算法，含水量推估的方法會以植被生理與土壤含水量的相互關係著手，透過地表植被蒸發散量的變化去推估土壤水分含量。此方法將參照 105 年水土保持局之時空巨量資料推估技術於坡地災害風險預警之發展應用中的遙測土壤含水量推估方法，但本次的遙測資料來源將取代原本單一衛星資料，改以融合衛星資料作為遙測之料的輸入。

(四) 算法及模型驗證方法

1、異元衛星融合驗證

本計畫驗證方法參考 Moreno-Martinez (2020) 中的驗證方式並加以修正。透過將已知資料挖空，利用融合算法補遺，並且分析補遺結果的不確定性 (Moreno-Martínez et al., 2020)

歸因於本計畫的目標以及臺灣地表變動的特性，本計畫篩選五種地形地貌，農地、邊坡、丘陵、高山以及草原以評估融合的結果，以 $RMSE$ 、 MAE 、 R^2 對融合結果進行評估。

2、土壤含水量驗證

本計畫採用行政院農業委員會水土保持局土石流防災資訊網所提供的土壤含水量資料來進行推估結果的率定。本計畫將選擇至少四個測站，蒐集該處的土壤含水量歷史觀測數據，利用該歷史數據的時間和空間資訊，來擷取至本計畫推估取得之該點位的 CWSI，繪製推估之 CWSI 以及土壤含水量的散布圖並且計算皮爾遜相關係數 (Pearson correlation coefficient)，以此驗證推估結果趨勢是否吻合與儀器監測結果。

二、研究方法

計畫將區分為兩個大項目接續完成，分別為系統模組開發以及系統測試與驗證。圖 2-1 為本計畫的計畫架構圖。

本計畫系統架構含八項工作子項目，包含兩個主要模型的開發，資料蒐集，以及兩個模型的驗證。

1. 融合衛星最佳先驗推估：貝氏模型平均
2. 即時最佳推估：卡曼濾波器
3. 即時精準最佳推估：貝氏最大熵法
4. 蒸發散量推估：地面能量平衡方程式

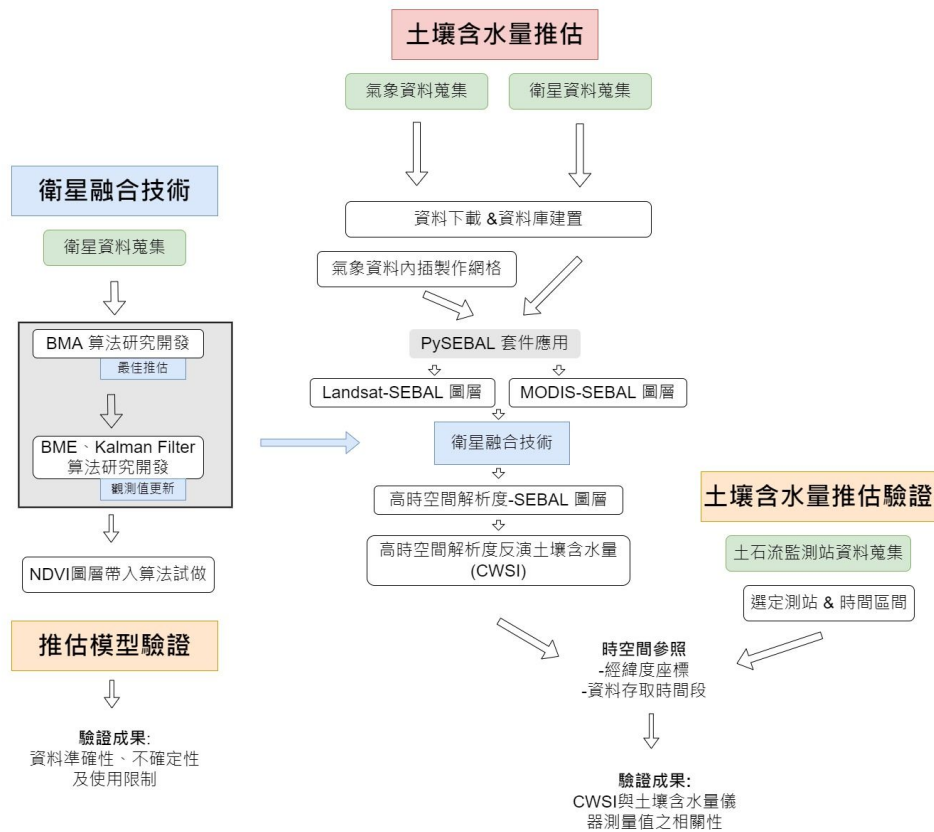


圖 2-1: 計畫流程架構圖

5. 土壤含水量數值推估：作物水分逆境指數
6. 資料參數蒐集
7. 衛星融合推估模型驗證
8. 土壤含水量推估模型驗證

三、研究執行步驟

(一) 融合衛星最佳先驗推估：貝式模型平均

貝氏模型平均 (Bayesian model averaging, BMA) 考慮模型的不確定性，藉由透過一個以上的模型來進行補遺及預測。它使用多個資訊來源來降低不穩定度，以提供更強健的估計值 (Hoeting et al., 1999)。在衛星觀測上，BMA 整合兩個或兩個以上的資訊，降低地表反射數值的不確定性，藉此提供更為強健的推估。該模型分為兩個部份，最佳推估以及誤差修正，以此來適應不同地表

地貌的特性。

1、最佳推估

當衛星資料在當前時刻 k 的觀測資料缺失時，其最佳的推估值為先驗估計值 x_k^- ，該值存在即時衛星觀測值於該給定時刻不存在。藉由結合 Landsat 衛星反射率的氣候資料（2013 至當前時刻 k 之歷史資料的平均和變方和 MODIS 衛星的反射率，利用貝氏模型平均（Bayesian model averaging, BMA），可以求得先驗估計值 x_k^- 。

首先為了能將 MODIS 反射率與 Landsat 衛星反射率的氣候資料結合，MODIS 資料需要先降尺度至 Landsat 的空間解析度，而空間分解（Spatial disaggregation）的方法使用線性映射（Linear mapping）。本計畫使用普通最小平方法（Ordinary least square, OLS）來做線性映射，單一像素透過給定時間往回推一年期間所提供的資料點為線性迴歸的觀測值，MODIS 觀測值在此是先透過最近鄰居法（nearest neighbor algorithm）重新採樣至 Landsat 的尺度 (30m)，以此來取得像素對像素的資料結構，再透過 OLS 取得各像素之相關係數 β_i ， i 在這邊代表衛星資料的像素點，式 2.1 為 OLS 的矩陣表達式。

$$\beta_i = (U_{i,MOD}^T U_{i,MOD})^{-1} U_{i,MOD}^T z_i \quad (2.1)$$

$U_{i,MOD}$ 為增廣矩陣 $[1, u_{i,MOD}]$ ，1 是全為數值 1 的向量， $u_{i,MOD}$ 是一個向量，其值為像素 i 在有 Landsat 資料的時間點時的無缺失 MODIS 反射率，其時間區間為給定時刻 k 往前推一年。 z_i 是一個向量，代表 Landsat 資料的時間區間中各像素的在不同時間的反射率，而向量 β_i 是一對迴歸係數（斜率以及截距），每個像素 i 的迴歸係數是獨立的。

當求出迴歸係數 β_i 後，空間分解後的 MODIS 在時刻 k 經由被線性映射得到

$$u_{i,k,MOD}^{30} = U_{i,k,MOD} \beta_i \quad (2.2)$$

$u_{i,k,MOD}^{30}$ 在這邊為 MODIS 反射率在像素點 i 和時刻 k 空間分解後的結果。線性模型被假設是無偏估，表示空間分解後的 MODIS 其平均值與 Landsat 的觀測平均值相同，由此可以推算分解後 MODIS 反射率的誤差協方差藉由式2.3

$$P_{k,MOD} = E[(z_i - U_{i,k,MOD} \cdot \beta_i)(z_i - U_{i,k,MOD} \cdot \beta_i)^T] \quad (2.3)$$

分解後的 MODIS 反射率和 Landsat 的氣候資料在此分別代表兩個模型，兩個資訊來源，藉 BMA 可以運算出先驗估計值 x_k^- 的平均值和變方，如式2.4及式2.5所示。

$$x_k^- = \bar{z} \frac{P_{k,MOD}}{\bar{P}_{k,LS} + P_{k,MOD}} + u_{i,k,MOD}^{30} \frac{\bar{P}_{k,LS}}{\bar{P}_{k,LS} + P_{k,MOD}} \quad (2.4)$$

$$P_k^- = (1 - \gamma) \left(\frac{1}{\bar{P}_{k,LS}} + \frac{1}{P_{k,MOD}} \right)^{-1} \quad (2.5)$$

其中 $\bar{z}_k$ 是 Landsat 氣候資料在 2013 年至當前時刻 k 的 Landsat 平均值， $\bar{P}_{k,LS}$ 是 Landsat 氣候資料在 2013 至當前時刻 k 的變方， γ 為一個分數，代表由偏差貢獻到總體的誤差協方差的比率，而 $P_{k,MOD}$ 是分解後 MODIS 反射率的誤差協方差，見式2.3。

2、偏差修正

偏差修正的需求來自觀測值和最佳推估值的相異。在貝氏模型平均之中，最佳推估值是來自於歷史資料且假設為無偏估，但實際是衛星融合補遺上，每個像素的估算受到前一

個觀測值的影響較大。基於此觀念，最終融合影像的推估結果除了貝氏模型平均的結果之外還需要考慮前一次的觀測值，以前一次觀測值來修正得到最終推估。

基於假設，最終推估的誤差協方差 W_k 是誤差協方差 P_k 及偏差協方差 T_k 的總合

$$W_k = P_k + T_k \quad (2.6)$$

在此有兩個不同的推估，一個來自貝氏模型平均的結果稱為狀態推估 (state model)，另一個來自偏差推估 (bias estimate)。偏差推估在當前時刻 k 是由 $k-1$ 修正而來，而最先的偏差由實際反射率觀測值和先驗估計值相減而得，在另外偏差預報誤差協方差 T_k 是一部分的先驗反射率估計誤差，修正來自式2.7：

$$b_k^- = b_{k-1} T_k^- = \gamma \left(\frac{1}{\bar{P}_{k,LS}} + \frac{1}{P_{k,MOD}} \right)^{-1} \quad (2.7)$$

γ 代表用於偏差修正的觀測值資訊，當 $\gamma = 0$ 表示資訊皆來自於貝氏模型平均的無偏估推估，當 $\gamma = 1$ 表示所有誤差接來自於偏差。 γ 控制了過濾器修正至實際偏差的速度和敏感性，但不完全決定最終推估值。在本計畫會設定 $\gamma = 0.6$ ，它較能掌握地面覆蓋的權衡，使地表變化較大的農田和變化小的天然森林都能有較佳的融合推估結果。

(二) 即時最佳推估：卡曼濾波器

貝式模型平均 (Bayesian Model Averaging, BMA) 的最佳推估提供了各個像素點位的先驗推估值 (priori estimate) 在時刻 k 。在卡爾曼濾波器的假設之中，所有的變數為常態分布，而在此由於在時刻 k 的先驗分布是由外部來源 BMA 的最佳推估取得，故傳統動態的卡爾曼濾波被靜態所取代。少了動態更新，以下的方

程式變成單純的一次性更新。

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} \quad (2.8)$$

$$x_k = x_k^- + K_k (z_k - H x_k^-) \quad (2.9)$$

$$P_k = (I - K_k H) P_k^- \quad (2.10)$$

k 代表時刻， x_k^- 是先驗推估值來自 BMA 的最佳推估值， z 是當下的觀測資料。 K 是卡爾曼獲益，代表 x^- 和 z 相對權重， P_k^- 為先驗推估值的誤差協方差， H 為觀測矩陣，用來描述模型和實際觀測數據的相關。 H 為 $m \times n$ 的矩陣，而 m 小於 n ， m 代表時當下觀測資料 z 實際有被觀測的像素點， n 為 x^- 最佳推估的像素點。 R 是 Landsat 的誤差協方差。最後 P_k 為後驗估計的誤差斜方差。

此外，考量地表變化並非全然是無偏估的。事實上，資料的時序性對於資料推估也會有所影響，故在卡曼濾波器之外，也會加入持續性模型 (persistence model)，該模型假定時間序列中未來的數值，與當前時間點的數值是相同的。式2.13 表示透過此模型修正即時最佳推估的結果。而式2.14和式2.15用於修正最終推估的誤差協方差 W_k ，以代表模型的不確性。

$$L_k = T_k^- H^T (H T_k^- H^T + H P_k^- H^T + R)^{-1} \quad (2.11)$$

$$b_k = b_k^- + L_k (z_k - H(x_k^- - b_k^-)) \quad (2.12)$$

$$\tilde{x}_k = x_k^- - (I - K_k H) b_k \quad (2.13)$$

$$T_k = (I - K_k H) T_k^- \quad (2.14)$$

$$W_k = P_k + T_k \quad (2.15)$$

其中 T_k 和 b_k 來自式2.14至式2.15。 x_k^- 、 x_k 、 P_k 、 R 來自式2.12至式2.13。最終， $\tilde{x}_k$ 為最佳即時推估， W_k 為其誤差協方差。

(三) 即時精準最佳推估：貝氏最大熵法

貝氏最大熵法 (Bayesian maximum entropy, BME) 是一種近代時空模式與繪圖 (Modern spatio temporal modeling and mapping) 的方法。可自動整合非線性的推估技巧與非高斯的機率密度函數理論，也可以處理不同種類與尺度的資訊與不確定性的資料型態。

以一張衛星遙測影像來說，在貝氏最大熵法架構中，可將每一個像素點上的反射率組成視為一個時空隨機場 (Spatial-temporal random field)。令 $p = (s, t)$ ，表示在時空間隨機場的一項反射率特性，而其中向量是時空間中的一個位置 (s 為空間中地理座標位置， t 為時間)。隨機場模式提供了普遍和數學嚴謹架構去調查大自然現象。這樣可以提高在複合空間-時間域的預測能力。在機率場模式的機率密度函數，場中某個空間時間點 $p = (s, t)$ 的反射值可令為 $X_p = X_{s,t}$ 。而對於場中每個點的反射率 X_p ， f_{KB} 可表示如式2.16所示：

$$\begin{aligned} P_{KB}[x_1 \leq X_{p1} \leq x + 1 + dx_1, \dots, x_k \leq X_{pk} \leq x_k + dx_k] \\ = f_{KB}(p_1, \dots, p_k) dX_{p1} \dots dX_{pk} \end{aligned} \quad (2.16)$$

下標 KB 代表用來建構的此機率密度函數的“知識根據” (Knowledge base)。

對於知識根據，貝氏最大熵法將其分為兩大類，即 G-KB 與

S-KB。G-KB 是指全面性的知識根據 (General KB)，記為 G ，包含一些物理、生物的定律跟一些與空間時間相關的時空理論模式 (例如：平均值、共變異數、半變異數、高階動差等等)，適用於整個觀察區域。S-KB 是指特定點的知識根據 (Site specific KB)，記為 S ，包含確定性資料 (Hard data) 與不確定性資料 (Soft data) 這兩種資料型態。確定性資料就是觀測點的值為一確切的值，不確定性資料則是使用機率密度函數來描述觀測點的值。而總知識根據 K-KB(Total KB)，記為 K ，即是將全面性與特定點的知識根據結合，也就是 $K = G \cup S$ 。

貝氏最大熵法的基礎理論方程式如式2.17所示

$$\begin{cases} \int dx (g - \bar{g} e^{\mu^T g}) = 0 \\ \int dx \varepsilon_s e^{\mu^T g} - A f_k(X_p) = 0 \end{cases} \quad (2.17)$$

g 為一向量，代表所有由全面性知識根據而來的函式 $g_a (\alpha = 1, 2, \dots)$ 。 $\bar{g}$ 表示 g 的統計期望值。 μ 也為一向量，代表所有與空間時間位置、 g 相關的所有係數 μ_α ，可以將其視為對全面性知識根據的函式之間的相對重要程度。 ε_s 是特定點知識根據的的運算子 (Operator)。 A 是標準化參數。 f_K 是每個空間時間中所有每一個點的機率密度函數，下標 K 表示此機率密度函數式基於總知識根據而來。式中 g 與 ε_s 是需輸入的部分， μ 與 f_K 則是未知待定的變數。

因為 f_K 是空間時間中所有每一個點的機率密度函數，所以欲推估的點 p_k 也包含在 f_K 之中。只要得到欲推估點的機率密度函數，便可以依研究上的需要，推估在該點上的某些統計特性，例如：平均值，變異數、高階動差等等。

經由替換貝氏最大熵法，將可不再侷限在高斯分布的假設下，可假設非線性系統，考量推估的不確定性，進而建構出合理的衛星融合算法。貝氏模型平均的結果，先驗估計值 x_k^- 以及誤差協方差 P_k^- 將和偏差誤差的結果，偏差推估值 b_k^- 和偏差預報誤差協方差 T_k^- 皆將放入貝氏最大熵法作為不確定性資料 (Soft data) 而當天的實際觀測值 (空間分解後的 MODIS 或是 Landsat)

則作為確定性資料 (Hard data) 使用，而整體的資料融合架構即如圖 2-2所示。

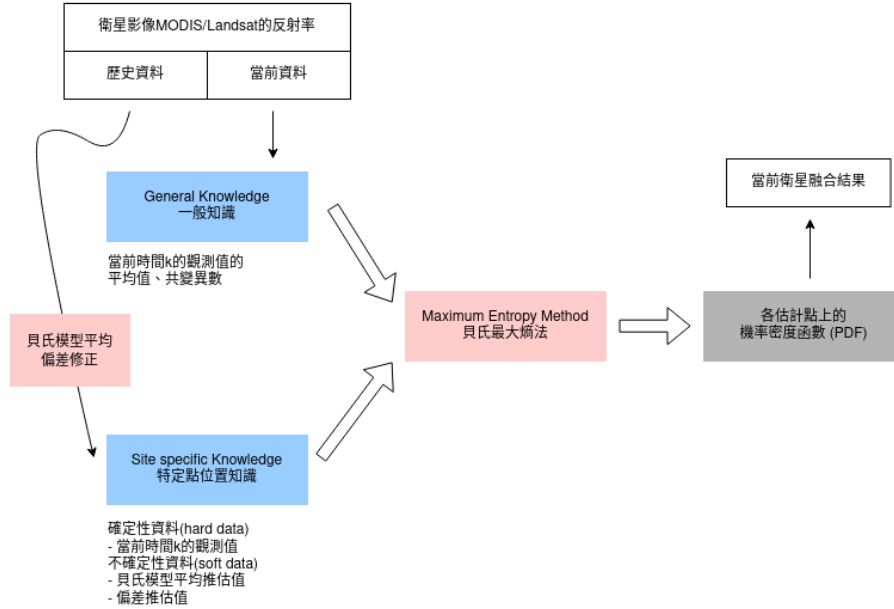


圖 2-2: 衛星影像融合推估算法架構

(四) 蒸發散量推估：地面能量平衡方程式

1、淨輻射能量 (Net Radiation, R_n) 推估

淨輻射能量之物理方程式，淨輻射量即為目標系統所接收之總輻射能量值。式中可知計算方法為將短波即長波輻射依據特性及地表反射率做運算，各參數詳細說明如式2.18

$$R_n = (1 - \alpha)K^\downarrow + L^\downarrow + L^\uparrow - (1 - \epsilon_0)L^\downarrow \quad (2.18)$$

其中 α 為地表反射率， ϵ_0 為「寬頻」地表放射率， $K^\downarrow$ 為短波輻射量 (W/m^2)， $L^\downarrow$ 為入射長波輻射量 (W/m^2)， $L^\uparrow$ 為出射長波為出射長波輻射量 (W/m^2)，如式2.18所示。

$$K^\downarrow = S_c \cos \theta_s d_r \tau_{sw} \quad (2.19)$$

其中 S_c 為太陽常數， $1367 (W/m^2)$ ； θ_s 為太陽入射天頂角 (solar incidence angle)； d_r 為反平方相對地日距離； τ_{sw} 為短波波段大氣穿透率。

$$L^\downarrow = 0.85 \times (-\ln \tau_{sw}) \times \sigma T_a^4 \quad (2.20)$$

其中 τ_{sw} 為短波波段大氣穿透率； σ 為史帝夫波茲曼常數 ($5.67 \times 10^{-8} W/(m^2 K^4)$)； T_a 為氣溫 (K)。

$$L^\uparrow = \epsilon_0 \sigma T_s^4 \quad (2.21)$$

其中 ϵ_0 為「寬頻」地面熱發射率 (surface emissivities)； σ 為史帝夫波茲曼常數 ($5.67 \times 10^{-8} W/(m^2 K^4)$)； T_s 為地表溫度 (K)。

2、土壤熱通量 (Soil heat flux, G) 推估

土壤熱通量 (Soil heat flux, G) 為目標系統未接收到之輻射能量，轉而由土壤接收。土壤熱通量與 T_s (地表溫度)、 α (地表反射率)、常態化差值植生指數 (Normalized difference vegetation index, NDVI) 及淨輻射量 (R_n) 有關，各參數詳細說明如式2.22所示。

$$G = T_s(0.0038 + 0.0062\alpha)(1 - 0.978NDVI^4)R_n \quad (2.22)$$

R_n 為淨輻射量 (W/m^2)； T_s ：瞬時地表溫度 (K)； α 為地表反射率 (Surface albedo)；NDVI 為常態化差值植生指標。

3、可感熱通量 (Sensible heat flux, H) 推估

可感熱通量 (Sensible heat flux, H) 為使目標系統產生溫度改變之輻射能量值。根據地表熱平衡方程式可以求得。

$$R_n = G + H + LE \quad (2.23)$$

$$H = R_n - G - LE \quad (2.24)$$

R_n 為淨輻射能量， G 為土壤熱通量， LE 為潛熱通量。

其計算與空氣密度、空氣比熱、地表溫度與氣溫之差值及熱傳空氣動力阻抗等參數相關，各參數詳細說明如式2.25所示。

$$H = \frac{\rho C_p}{r_{ah}} dT \quad (2.25)$$

其中 ρ 為空氣密度 (kg/m^3)， dT 為溫度差異在兩個高度之間； C_p 為空氣比熱 ($J/kg \cdot K$)； r_{ah} 為熱傳空氣動力阻抗 (Aerodynamic resistance to heat transport)，如式2.26所示。

$$r_{ah} = \frac{1}{k^2 u(Z)} \ln\left(\frac{Z_{ref}}{Z_{oh}}\right) \ln\frac{Z}{Z_{om}} \quad (2.26)$$

其中 k 為 Von Karman constant； Z_{ref} 為氣溫計之高度 (m)； Z 為動量傳遞粗糙長度 (m)； $Z_{oh} = 0.1Z_{om}$ (m)； Z 為氣象站風速儀高度 (m)； $u(Z)$ 為在 Z 高度下之風速 (m/s)。

在此， dT 為 T_s 瞬時地表溫度 (K) 和近地面氣溫的差異，地表溫度的取得不易，取而代之使用 dT 和 T_s 存在的線性相依性來求得 dT 。透過衛星影像中冷熱像素的選擇，透過各自的假設來推算 dT ，再利用多次的遞迴調整 r_{ah} 進而使 $dT = aT_s + b$ 達到穩定。

dT_{cold} 的假設

dT_{cold} 為冷像素點上近地面氣溫和地表溫度的差異，該像素位於濕度高的位置但非水體或雲，一般來說會尋找充分灌溉的田區，上面覆蓋植被，作物為主。在該處的表面溫度和近地表氣溫是假設為相同。由此推斷，該點位的 dT 為 0，並且根據式2.25， H_{cold} 也為 0。

dT_{hot} 的假設

dT_{cold} 為熱像素點上近地面氣溫和地表溫度的差異，該像素位於乾燥且裸露的農田土壤，該處的蒸發散 (Evapotranspiration, ET) 被假設為 0。此外，式 2.19 的反導，可以由 H_{hot} 求得 dT ，見公式。

$$H_{hot} = R_n - G - \lambda ET, \lambda ET = 0 \quad (2.27)$$

$$dT_{hot} = \frac{H_{hot} \times r_{ah}}{\rho C_p} \quad (2.28)$$

至此， dT_{hot} 可以求得，但由於大氣的不穩定性， r_{ah} 於式的計算方法存在著不穩定因素，該不穩定因素可以熱傳導的穩定校正 ψ_{z1} 以及 ψ_{z2} 表示。透過迭代的運算可以得到持續大氣修正下的 T_a 、 ρ 、 H ，每次迭代的 H_{hot} 會利用 Monin-Obukhov length(L) 來定義是否為穩定狀態，進而產生一個熱傳導的穩定校正參數。最終，迭代運算會進行直到 dT_{hot} 和 r_{ah} 在熱像素中達到穩定狀態。

自動熱/冷像素篩選

自動化冷/熱像素的選取需透過兩個階段：(1) 植被指數 NDVI 高低篩選和 (2) 地表溫度篩選。

- 植被指數 NDVI 高低篩選

NDVI 門檻，作為熱/冷像素的篩選門檻，該數值介於 0 到 1，代表影像中所有像素數值進行排列的百分位數，本次實驗根據前人研究選擇 0.1 作為冷/熱像素的篩選門檻。

- 地表溫度篩選

在透過 NDVI 篩選過後的像素，會先去除雲遮蔽和水體後，依據前人研究會在篩選當中前百分之五的像素放入熱/冷像素池，藉此算出冷/熱像素池的最小值、最大值、標準差以

及平均數 (Dhungel & Barber, 2018; Feng, 2015)。

4、潛熱通量 (Latent Heat flux, LE) 推估

穩定的 dT_{hot} 以及假設為 0 的 dT_{cold} ，搭配上對應像素的地表溫度 T_s ，可以求得 $dT = aT_s + b$ ，該線性關係套用於各個像素即可取得各像素的 H 。至此，潛熱通量 λET 可被推導。

5、文獻回顧-SEBAL 各項常數定義

針對氣象因子，本研究依據前人設定一些定值。氣溫垂直遞減率 (Temperature lapse rate) 依據國際民航組織 ICAO 定義國際標準大氣 ISA，從海平面至海拔 11km 的溫度遞減率微 6.5 攝氏度/公里 (Organization, 1993)。太陽常數 (Solar constant) 在大氣層之外，以地球與太陽之平均距離為準，在垂直於太陽輻射之單位面積上所受到之太陽輻射率，其值約為 1367 W/m^2 (Johnson, 1954)。玻滋曼常數根據科學技術數據委員會 CODATA 定義為 $5.6703 \times 10^{-8} \text{ W/m}^2 \text{ K}^4$ 。

對於尚未進行大氣校正的 Landsat 產品 (Level 1)，本研究根據 Wukelic, G., et al. (1989) 的研究定義各項對於地表溫度大氣校正的常數。路徑輻射率 (Path radiance) 在地表溫度 (Land surface temperature, LST) 波段定為 $0.91 (\text{W/m}^2/\text{sr}/\mu\text{m})$ (Wukelic et al., 1989)。空氣窄頻帶透射率 (Narrow band transmissivity of air) 定為 0.866 (Wukelic et al., 1989)。此外，水對於地表溫度抵銷值設為 3、陰影對於地表溫度抵銷值設定為 -1、雲對於地表溫度抵銷值設定為 -3，以及最大陰影以及雲的反照率設定為 0.1 和 0.4，此些數值用於進一步增進雲及水體遮蔽的正確性 (UNESCO-IHE, 2018)。

(五) 土壤含水量數值推估：作物水分逆境指數

地面熱平衡方程式 SEBAL 以淨輻射能量 (Net Radiation, R_n , W/m^2)、土壤熱通量 (Soil Heat flux, G , W/m^2)、可感熱通量 (Sensible Heat flux, H , W/m^2) 及潛熱通量 (Latent Heat flux, LE , W/m^2) 來代表地表上能量的變化情形，並以潛熱通量與淨輻射量減去土壤熱通量之比值來代表蒸發散比率。淨輻射能量表示接收於太陽

輻射之總輻射能量；土壤熱通量表示在淨輻射能量中傳遞於土壤之中的能量；可感熱通量表示因為地表溫度與氣溫之間的差值，所造成的能量交換項；潛熱通量即表示進行蒸發散所需之能量。CWSI 的計算公式亦可由實際蒸發散量及潛勢蒸發散量比值推估，以此比值可以去除氣象因素對於植物蒸散的影響，以此方法產生的代表性指標對於偵測單純土壤濕度較具有說服力，而該數值代表公式推導如式2.29所示：

$$CWSI = 1 - \frac{LE}{R_n - G} \quad (2.29)$$

E_a 代表每日實際蒸發散通量 ($kgm^{-2}day^{-1}$)，而 E_p 代表每日潛勢蒸發散量 ($kgm^{-2}day^{-1}$)(Baranowski et al., 2001)

(六) 資料參數蒐集

1、氣象資料庫蒐集及建置

本研究使用臺灣氣候變遷推估與資訊平台計畫 (Taiwan Climate Change Projection and Information Platform, TCCIP) 的作為歷史氣象資料來源。TCCIP 中提供 Level 2 以及 Level 4 的臺灣歷史氣候重建資料 (TReAD)。Level 2 為日資料，包含 SEBAL 模型中需要的相對濕度、短波輻射、平均溫、2 米溫度、2 米風速、日輻射量。Level 4 為小時資料，與 Level 2 資料型態相似，但僅擷取上午 10 點至 11 點間的資料區段，該區段的資料對應於 MODIS 以及 Landsat 衛星於觀測區域的拍攝時間，因此作為 SEBAL 模型中的瞬時 (Instantaneous) 資料使用。

2、Level 2 資料下載及前處理

Level 2 資料由 TCCIP 網站之氣候變遷資料商店中點選下載 (https://tccip.ncdr.nat.gov.tw/ds_03.aspx)，選擇指定參數以及空間尺度 (2km) 後手動下載包含所有年份的歷史資料。資料為.zip 壓縮檔，內部各檔案為 csv 檔，透過 Python 自動化解壓縮，再利用 OSGeo4W 程式轉換成網格資料、使用近鄰演

算法 (Nearest-Neighbour Algorithm) 降空間尺度至 30mx30m，存成.tiff 檔案輸出至資料庫。

3、Level 4 資料前處理

由於 Level 4 有其資料產製之假設與使用限制，需與 TCCIP 方開會討論資料的使用、限制以及不確定性是否適合本研究計畫，目前已配合 TCCIP 召開線上討論會議，瞭解使用之假設與限制。將可陸續取得 2013 至 2019 年全臺之溫濕度、風場、日射量以及蒸發散、土壤含水量等氣候重建資料。預計將取得之 netcdf 格式資料進行前處理，將溫濕度、風場、日射量等變數轉為適合的格式，透過 SEBAL 進行蒸發散量之計算，而由 TCCIP 提供之蒸發散、土壤含水量等變數則可提供研究推估結果比對，探討利用不同方法產製之資料間差異。

4、衛星資料庫蒐集及建置

(a)MODIS 衛星資料下載

經由 MODIS 網站上可連結至其衛星遙測演算產品資料庫 (<http://modis.gsfc.nasa.gov>)，資料庫內存有全世界所有地區(圖 2-3)觀察資料之 HDF 檔。若由人工手動方式抓取資料將耗費大量時程選取並下載，本研究利用開源程式語言-Python，運用自行研發之 MODIS 資料下載指定時間及地區的資料。程式中可指定所需資料之日期、時間範圍、產品編號，即可由 MODIS 資料庫中下載所需資料。MODIS 所提供之資料格式為 HDF 檔，HDF 全名為 Hierarchical Data Format，用於儲存與組織較大量且種類繁多的資料。由於 MODIS 產品之資料與諸多輻射頻段及觀測條件有關，資料的儲存格式上便選擇使用 HDF 檔。此種檔案格式無法直接利用，需經過特別的地理資訊資料-GDAL(Geospatial Data Abstraction Library) 來進行處理。本研究利用 OSGeo4W 程式來進行 HDF 檔案的處理。

MODIS 衛星有許多不同的產品，本計畫將選用其中 MOD11A1 和 MOD09GA 作為本研究的資料來源。MOD11A1 全名為 Terra Land Surface Temperature and Emissivity Daily Global 1km，該產品提供每日地表溫度和發射率，解析度為一公里；



圖 2-3: MODIS 衛星遙測演算產品資料庫服務畫面

MOD09GA 全名 MOD09GA.006 Terra Surface Reflectance Daily Global 1km and 500m，該產品為提供地表反射，總共有七個頻帶，頻帶介於 $0.459 \sim 2.155 \mu m$ ，涵蓋可見光、近紅外光與熱紅外光。

(b)Landsat 衛星資料下載

經由 NASA 以及 USGS 提供的開源資料平臺 earthexplorer(<https://earthexplorer.usgs.gov/>) 可以連接上至 Landsat 衛星之資料。但同樣因為手動耗時的原因，本研究使用開源的 Landsat 衛星下載的 Python 套件來下載需要的衛星資料(<https://pypi.org/project/landsatxplore/>)。Landsat 資料所提供的格式為 TAR 檔，TAR 檔的全名為 Tape Archive，為一種封包檔案，用於儲存資料，而不會壓縮它，是相當常見的歸檔格式。由 earthexplorer 下載的 Landsat 檔案中，各光譜帶以及空間相關資訊的頻帶 (bands) 被儲存成 Tiff 檔於其中，Tiff 的全名為 Tagged Image File Format，為一種壓縮圖像格式，至於衛星資料的儲存檔案由於擁有 GPS 元資料 (metadata)，故儲存的檔案又被稱為 GeoTiff。

Landsat 衛星有兩個主要的產品，Level 1 以及 Level 2。Level 1 是原始資料，以數位 (Digital Number, DN) 紀錄感測器測量到的數值，其本身不具有實際數值意義，所以第一步必需根據不同世代的 Landsat 和每張照片的原始資料，將 DN 轉

換為輻射值。第二步則是將輻射數據轉換為大氣頂層 (Top of atmosphere, ToA) 反射率，該資料取得為即時; Level 2 產品是大氣校正後的標準數值產品，產品中資料集為各地表反射率以及地表溫度，該資料因為需要進行大氣校正，且需要透過 MODIS 衛星提供大氣校正的修正因子，故取得影像的時間為拍攝時間的八天之後。

(c)MODIS 次級圖層處理計算

- NDVI 運算

本研究使用 MOD09GA 提供的每日地表反射數值中的紅光 B4 和遠紅光 B5 計算每日常態化差值植生指標 (Normalized difference vegetation index)，由於 MODIS 產品已經經過大氣校正故其透過反射數值計算出的 NDVI 為地表 NDVI。計算公式如下

$$NDVI = (B5_{nir} - B4_{red}) / (B5_{nir} + B4_{red}) \quad (2.30)$$

- SAVI 運算

SAVI 為土壤植生指數 (Soil Adjusted Vegetation Index, SAVI)，該指數最小化土壤亮度對於植被指數的影響，本研究使用 MOD09GA 提供的每日地表反射數紅光和遠紅光計算之，由於 MODIS 產品已經經過大氣校正故其透過反射數值計算出的 SAVI 即為地表 SAVI，計算公式如下：

$$SAVI = (B5_{nir} - B4_{red}) / (B5_{nir} + B4_{red} + L) \times (1 + L) \quad (2.31)$$

其中 L 為土壤亮度校正指數，設定為 0.5 基於多種不同地表覆蓋的結果 (Huete, 1988)。

- 反照率 (Surface Albedo) 運算

本研究對於地表反照率使用 Tasumi, M., Allen, R. G., &

Trezza, R. (2008) 對於 MODIS 產品基於每日地表反照率的計算公式求得，該公式被證明與 MODIS 自身地表反照率產 MCD43A3.006 的誤差為-0.035 至 +0.033(信賴區間 95

$$\begin{aligned} Surf\ albedo = & 0.254 \times B1_L + 0.149 \times B2_L + 0.147 \times B3_L \\ & + 0.311 \times B4_L + 0.103 \times B5_L + 0.036 \times B7_L \end{aligned} \quad (2.32)$$

(d)Landsat 次級圖層處理計算

本研究使用 Landsat 的 Level 2 資料產品計算 NDVI，公式與 MODIS 用於計算 NDVI 相同。Landsat 產品具有品質鑑定圖層 (quality assessment, QA)，產品中的名稱為 QA_PIXEL。計畫考量下，所有不被分類為清晰地表的像素皆被剔除。雲遮蔽處理後的圖層會進行重新縮放，以 Landsat 產品提供的縮放尺度進行調整，公式如下：

$$P_{L,rescaled} = P_L \times 2.75 \times 10^{-5} - 0.2 \quad (2.33)$$

P_L 為 Level 2 產品中地表反射率原始值， $P_{L,rescaled}$ 為尺度縮放的數值。

• 反照率 (Surface albedo) 運算

根據 Tasumi, M., Allen, R. G., & Trezza, R. (2008) 的文獻，Landsat 衛星反照率的計算公式由 Liang (2000) 衍生，在考慮綠譜帶的參數下，該方法在密集植被、綠色覆蓋、以至於農業用地上有較佳的推估，該方法的推導公式如 (Liang, 2001; Tasumi et al., 2008)：

$$\begin{aligned} Surfalbedo = & 0.254 \times B1_L + 0.149 \times B2_L + 0.147 \times B3_L \\ & + 0.311 \times B4_L + 0.103 \times B5_L + 0.036 \times B7_L \end{aligned} \quad (2.34)$$

(e) 衛星地表溫度銳化 (Kustas et al., 2003)

- NDVI 和表面溫度的二次多項式擬合

衛星資料的地表溫度與地表植被指數存在關係當運用於空間分布的地表能量通量 (Surface energy fluxes) 和可用水分。Kustas, et al. (2003) 提出用於計算蒸發散輻及地表能量通量的地表溫度銳化模型，利用地表溫度和地表植被指數的二階多項式擬合，擬合公式如：

$$\hat{T}(NDVI) = a_1 + a_2 NDVI + a_3 NDVI^2 \quad (2.35)$$

$\hat{T}(NDVI)$ 為以 NDVI 為解釋變數的預測值，透過於實際觀測溫度 T 的普通最小乘估計取得地表溫度與地表植被指數的二階多項式相關係數 a_1 、 a_2 以及 a_3 。

- 線性映射 (Linear mapping)

取得地表溫度和植被指數之相關係數後，地表溫度銳化將以線性映射達成，透過高解析度的地表植被指數圖層透過式推估出高解析度之地表溫度。

- 異常值去除 (Outlier removal)

最後，為了消除異常值，銳化後的地表溫度會剔除異常值，任何像素數值小於 250k 以及大於 400k 被剔除。

(f) 衛星地表溫度校正 (UNESCO-IHE, 2018)

- 基於高程校正地表溫度

地表溫度受到海拔高度的影響，由衛星遙測熱感應量測出來的溫度需要透過數值高程模型 (Digital elevation model,

DEM) 進行校正，公式為下：

$$T_{s,dem} = T_s + \gamma_o \times DEM \quad (2.36)$$

γ_o 為氣溫垂直遞減率， DEM 為像素點的數值高程值， T_s 為地表溫度， $T_{s,dem}$ 為經過數值高程校正後的地表溫度。

除了數值高程模型校正之外，本推估考量地球外太陽輻射 (Extraterrestrial solar radiation) 在傾斜和平坦地形上的差異，以解釋以坡度引起的溫度變化，計算公式如下：

$$T_{s,corrected} = T_{s,datum} + \frac{G_{sc} \times d_r \times \tau_{sw} \times (\cos\theta_{slope} - \cos\theta_{flat})}{\rho_{air} \times 1004 \times 0.05} \quad (2.37)$$

$T_{s,datum}$ 為地表溫度的基準值， G_{sc} 為太陽常數， d_r 為地球-太陽的距離反比， ρ_{air} 為空氣密度， θ_{slope} 為傾斜地形的太陽入射角，坡度 (Slope) 由 DEM 推算得知， θ_{flat} 為平坦地形的太陽入射角。

溫度校正後會將異常值去除，將地表溫度範圍限縮在 273k 和 350k 之間。

5、數值高程模型及前處理

數值高程模型 (Digital Elevation Model, DEM) 資料取自內政部 20 公尺網格數值地形模型 (<http://dtm.moi.gov.tw/> 不分幅_全台及澎湖.zip)，該資料於 2016 年 11 月上架。該壓縮檔案包含全臺的 DEM TIFF 資料，元資料 (Metadata) 以及澎湖地區的 DEM TIFF 資料。

由於 DEM 的解析度與衛星產品不同，近鄰演算法 (Nearest-Neighbour Algorithm) 被套用進行像素點位的重新選取，以利於與衛星資料對齊像素點位。由於本計畫產出之衛星產品空間解析度為 30 公尺，DEM 圖層被降尺度至 30 公尺解析度。

(七) 衛星融合驗證

衛星融合驗證旨在評估融合演算法對於空缺值的估計精準度，又依據不同地貌在地表反射數值變異上相異，本研究將著重於臺灣各式主要地貌以模擬重複取樣的方法驗證在融合演算法推估數值精準度。此驗證分成四個步驟：

1、同質區域選擇

利用肉眼檢視臺灣土地利用狀態，於 QGIS 框選紀錄四角座標，透過座標繪製邊界框 (Bounding box)。總共選取五個地貌種類：高山、丘陵地、農地、草原以及坡地。挑選後，每個地貌資料挑選中央長寬為 100 個像素的區域，由於融合結果之空間尺度為 3000m，故面積為 9 km^2 。

2、資料隨機挖空

選取一張 Landsat Level 2 的觀測資料，對於每個區域個別進行資料隨機挖空，除了保留原始觀測值中的缺失值，另外在隨機挖取長寬皆為 3 個像素的區域。

3、融合資料推估

使用該資料的 BMA 最佳推估，分別使用卡曼濾波器以及 BME 進行資料推估，步驟 2 & 3 重複五百次。推估結果計算 $RMSE$ 、 MAE 以及 R^2 。

4、替換觀測值

將 Landsat Level 2 的觀測資料替換為同一天以 MODIS 進行線性映射後的觀測資料，重複上述步驟 1-3 計算 $RMSE$ 、 MAE 以及 R^2 。

(八) 土壤含水量推估模型建立方法

機器模型常被應用在遙測領域用來辨識地貌或者農作物，此方式不只能節省實際調查所消耗的人力物力，更能將過去以點狀的調查行提升到面狀。本計畫將以所推估之 CSWI 配合地文資料（如：高度、坡度等）以及水文資料（如降雨、濕度等）以機器學習方式建立土壤含水量之推估模型。機器模型雖然近幾年被廣泛應用在許多領域去完

成傳統演算法無法達成的任務,但其基本概念並不難理解,以下將簡要的對機器學習作介紹

1、機器學習術語和定義

通常,ML 方法涉及學習過程,其目標是從“經驗”(訓練數據 training data)中學習以執行任務。ML 中所需使用的數據通常有不同類別,這不同類別也被稱為不同特徵 (Features)、變數 (Variables) 或者屬性 (Attributes)。特徵可以是二進制的 (即 0 或 1)、類別型的 (如貓、狗、鳥)、數字的 (整數、實數等) 或是序列的 (例如 A+ 或 B- 等),ML 模型在特定任務中的性能是通過性能指標來衡量的,一般該指標會隨著學習時間或數據的增加而提高。ML 模型和算法的性能,可使用統計量和數學模型加以量化,如 AIC、Root mean square error, P value 等。學習過程結束後,訓練好的模型可以利用訓練過程中獲得的經驗,對新的數據或者說測試數據 (testing data) 進行分類 (classification)、預測 (prediction) 或聚類 (clustering)。一張典型且簡易的 ML 概念顯示如圖 2-4。

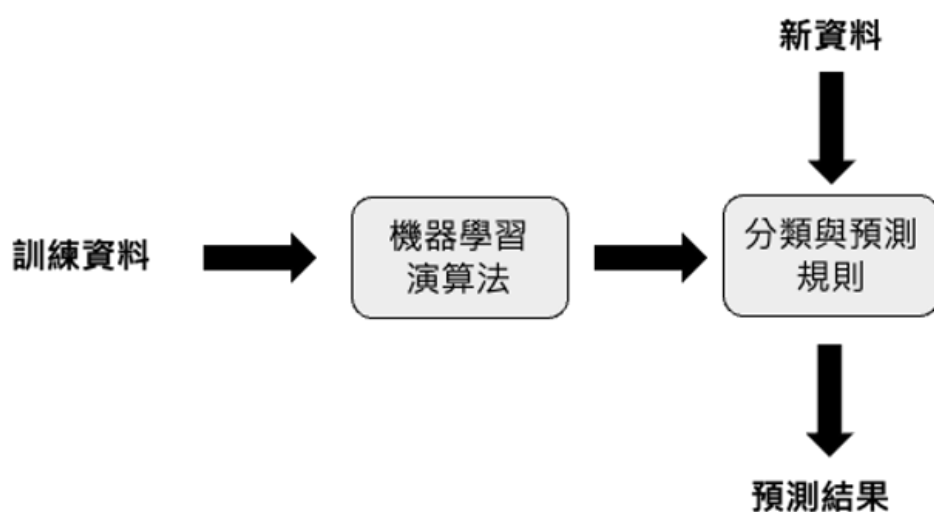


圖 2-4: 機器學習模型概念圖

2、監督式學習與非監督式學習

根據學習的方式,ML 的學習可分為兩大類,即監督學習 (supervised learning) 和非監督學習 (unsupervised learning)。在監督學習中,必須事先給定輸入特徵和其對應的輸出值,目標是構建輸入與輸出資料間的通用規則。在監督學習中,由訓練資料訓練好的模型降用於預測測試資料的輸出值。然而,在非監督學習中,訓練資料

和測試資料之間沒有區別，不需事先給定標記過的訓練範例，其會自動對輸入的資料進行分類或分群。

3、維度災難 (curse of dimensionality)

維度災難最早由理查德·貝爾曼 (Richard E. Bellman) 在考慮優化問題時首次提出，在如採樣、機器學習 (ML) 和資料探勘 (Data mining) 領域也都有提及到這個名字的現象。其概念為隨維度增加，空間體積將會急速膨脹，若資料量沒有相對應的增加，資料在空間中就會相對變得稀疏，而稀疏性在任何要求有統計意義的方法而言都是一個問題，因此，在機器學習模型中用來預測的輸入變數並非越多種越好。為精簡變數量，資料的降維 (Dimensionality reduction, DR) 方法或者特徵選擇 (Feature selection) 方法常被用於機器學習領域中。

.

降維 (Dimensionality reduction, DR) DR 為一不論在監督或非監督式學習中都可執行的分析，旨在降低維度的同時盡可能多地保留原始數據中的信息，讓數據有更緊湊的表示，通常在應用分類或回歸模型之前執行，以避免維度的影響。一些最常見的 DR 算法如下：主成分分析 (Principal component analysis, PCA)，偏最小平方迴歸 (partial least squares regression, PLS)，以及線性判別分析 (Linear discriminant analysis, LDA)。

.

特徵選擇 (Feature selection) 在機器學習和統計學中，特徵選擇也稱為變量選擇 (variable selection) 或屬性選擇 (attribute selection)，是挑選與輸出變數較相關或者對預測較有幫助的特徵用以模構建型。使用特徵選擇技術有以下幾個原因：(1) 簡化模型：使它們更容易被研究人員/用戶解釋；(2) 縮短訓練時間；(3) 避免維數災難；(4) 提高數據與學習模型類的兼容性等等。特徵選擇的方法主要分三大類：過濾法 (Filter methods)、包裝法 (Wrapper methods) 與嵌入法 (Embedded methods)。

- (a) 過濾法 (Filter methods)

其篩選透過評估輸入變數與輸出變數之間的相關性，排除相關性較低的輸入變數。衡量相關性可用一些計算相關性統計量如皮爾森相關係數 (Pearson correlation coefficient)、斯皮爾曼等級相關係數 (Spearman's rank correlation coefficient)；或者一些統計上用於檢測相關性的檢定如常見的卡方檢定 (Chi-squared Test) 與 ANOVA (ANalysis Of VAriance) 單變數檢定。

然而過濾法並不計算輸入變數間的相關性，因此容易選取的彼此相關性高的輸入變數，且其變數篩選亦不考慮模型表現之好壞。

- (b) 包裝法 (Wrapper methods)

包裝法透過評估模式表現的方式篩選特徵，即選擇能使模式表現較好之變數組合。依據篩選方式又可分為前進選取法 (Forward Selection)、反向淘汰法 (Backward Elimination) 與遞迴淘汰法 (Recursive Feature Elimination)。

- 前進選取法 (Forward Selection)：

首先必須決定一個評斷模式表現的指數如：R square 或者均方根誤差 (Root mean square error, RMSE) 等。而後從零個開始，依序加入可使模式表現最佳好的特徵，直到達到欲保留特徵之數量。

- 反向淘汰法 (Backward Elimination)：

同樣首先必須決定一個評斷模式表現的指數如：R square 或者均方根誤差 (Root mean square error, RMSE) 等。而後從所有特徵開始，依序刪減掉對模式表現最沒有幫助得特徵，直到達到欲保留特徵之數量。

- 遞迴淘汰法 (Recursive Feature Elimination)：

此方法與前述稍不一樣，是依據給定推估模型中不同特徵的權重 (例如線性模型之係數) 迭代

篩去特徵。首先以所有特徵訓練一模型，而後將 n 個權重較低的特徵剔除後，再用剩餘的特徵訓練一模型，而後再剔除 n 個權重較低的特徵，重複迭代直到達到欲保留特徵之數量。每次迭代剔除之特徵數量 n 可由使用者自行決定。

- (c) 嵌入法 (Embedded methods)：

過濾法在篩選特徵後才進行模式建構；包裝法則是在模式建構完成與特徵選擇中來回游移，即選擇特徵組合與模型建構是分開的；而嵌入法故名思義是將特徵選擇嵌入在演算法中，即特徵選擇與模式建構同時完成。常見的演算法為 LASSO (Least absolute shrinkage and selection operator)。其即為一種結合正則化 (regularization) 與特徵選擇的迴歸分析法，一般迴歸式之參數組合會以使損失函數 (Loss function) 最小的為目標，而所謂正則化就是將原本最小化損失函數的目標改為最小化損失函數 + 懲罰參數 (penalty parameter, P)，即 $\min(\text{loss function} + P)$ 。其包含了過濾法與包裝法的優點，透過懲罰參數將不重要的特徵係數降為零，並同時考量模式表現。另外基於樹狀 (Tree-based) 的機器學習方法亦為屬於嵌入法之演算法，因為其在建構樹同時也會提供各特徵的重要性，常見的為決策樹 (decision tree), 隨機森林 (random forest) 和一些梯度提升的演算法 (gradient boosting algorithms) 如 XGBoost, Catboost 等。

4、過度擬和 (overfitting)

過度擬和為一種機器學習中不甚理想的行為，一般指模型在訓練階段能有極佳的表現，卻在加入新資料進行預測或推估時不慎理想。造成過度擬和的原因有許多，如訓練資料包含大量不相關的資訊 (稱為雜訊資料)，或者模型複雜度太高，模型學習到訓練資料中的雜訊。為避免以上現象，一般機器學型流程如圖 2-5：

5、K 折交叉驗證 (k-fold cross-validation)

K 折交叉驗證為一種評估模型表現的方式，若作為特徵選擇之評估標準，一般作用於訓練集，如圖 2-6。圖 2-6 中亦展示 K 折交叉驗證之計算概念，首先將資料平均分為 k 等份，選定一模式表現評

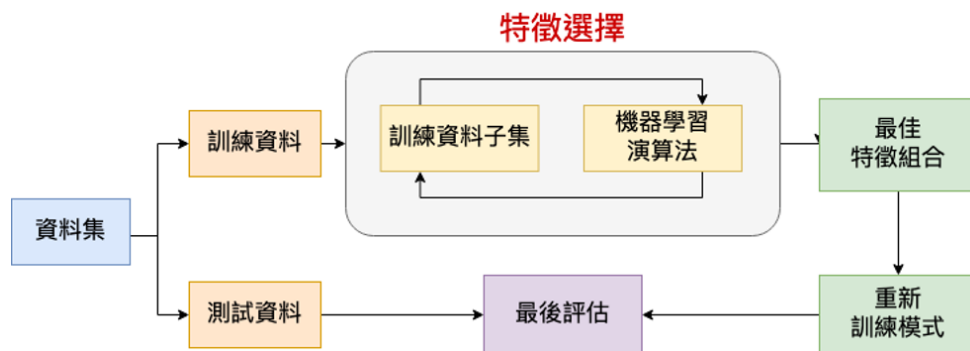


圖 2-5: 機器學習模式建構示意圖

估指標如 R square，並依照下列步驟計算：

- 利用 $k-1$ 等分的資料建立模型
- 用剩下的一等份，計算模式表現評估指標
- 重複上述兩步驟 k 次，獲得 k 個模式表現評估指標並平均

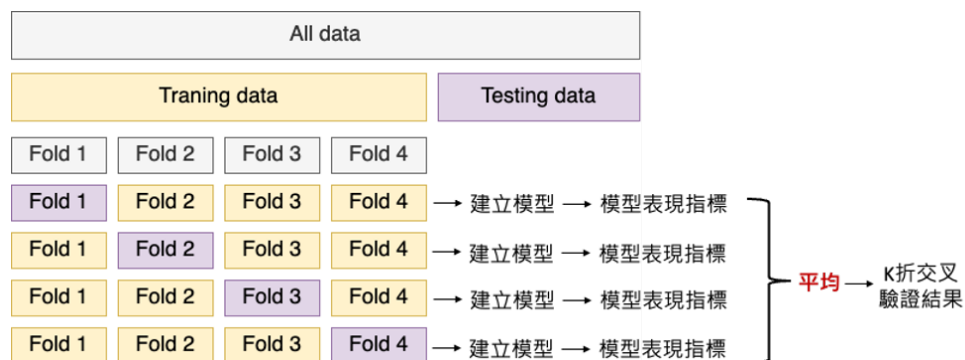


圖 2-6: k 折交叉驗證示意圖 ($k=4$)

6、XGBoost(eXtreme Gradient Boosting)

XGboost 全名 eXtreme Gradient Boosting，是一個優化的分佈式梯度提升庫 (distributed gradient boosting)，旨在高效、靈活和便攜。它在 Gradient Boosting 框架下實現了機器學習算法。

- 屬於 boosting 集成學習的其中一種，是將許多樹模型 (弱分類器) 集成在一起，形成一個很強的分類器。而所用到的樹模型則是 CART 回歸樹模型。
- 該算法思想就是不斷地添加樹，不斷地進行特徵分裂來生長一棵樹，每次添加一個樹，其實是學習一個新函數，去擬合上次預測的殘差

(九) 相關資料處理

本計畫建立土壤含水量推估模型時, 不只利用衛星影像所計算之 CWSI, 亦加入可能與土壤含水量相關之水文與地文資料, 介紹如下：

1、土壤含水量觀測資料

本研究擷取土石流觀測站之土壤含水量，為配合衛星影像之時間，主要使用 2019 年每八日上午 10-11 點之資料，選定 7 座土壤含水量測站：上安、神木、豐丘、九份二山、坪頂、南豐以及來義進行推估驗證，其測站分布圖如圖 2-7，各測站含水量時間序列如圖 2-8



圖 2-7: 使用土石流測站分布圖

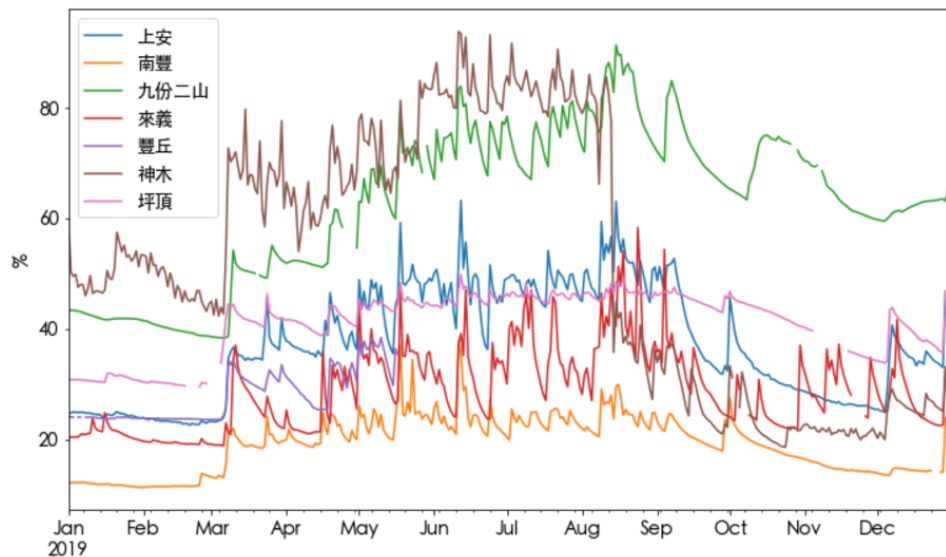


圖 2-8: 使用土石流測站之含水量觀測時間序列

2、地文資料

主要包括坡度、坡向與高程，皆以 20m 精度之 DEM 進行計算，下圖 2-9、圖 2-10、圖 2-11 分別展示坡度、坡向與高程之計算結果，由於來義測站較其餘測站遠，考量若將測站一併顯示比例尺會過小，故僅擷取使用測站較為密集之中部山區做展示。各測站之地文資訊則列於表 2-1

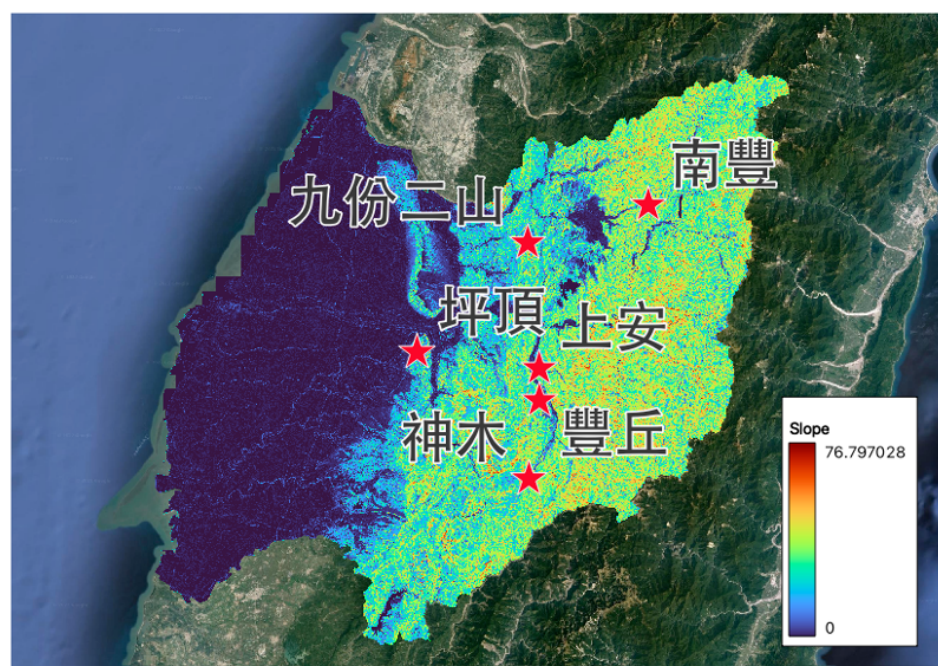


圖 2-9: 中台灣坡度空間分布

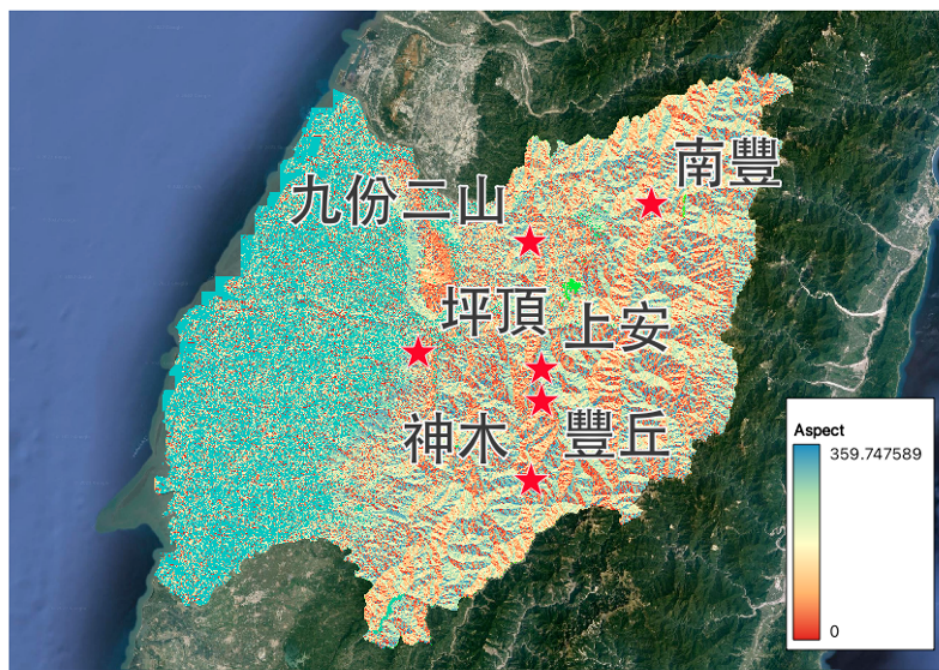


圖 2-10: 中台灣坡向空間分布

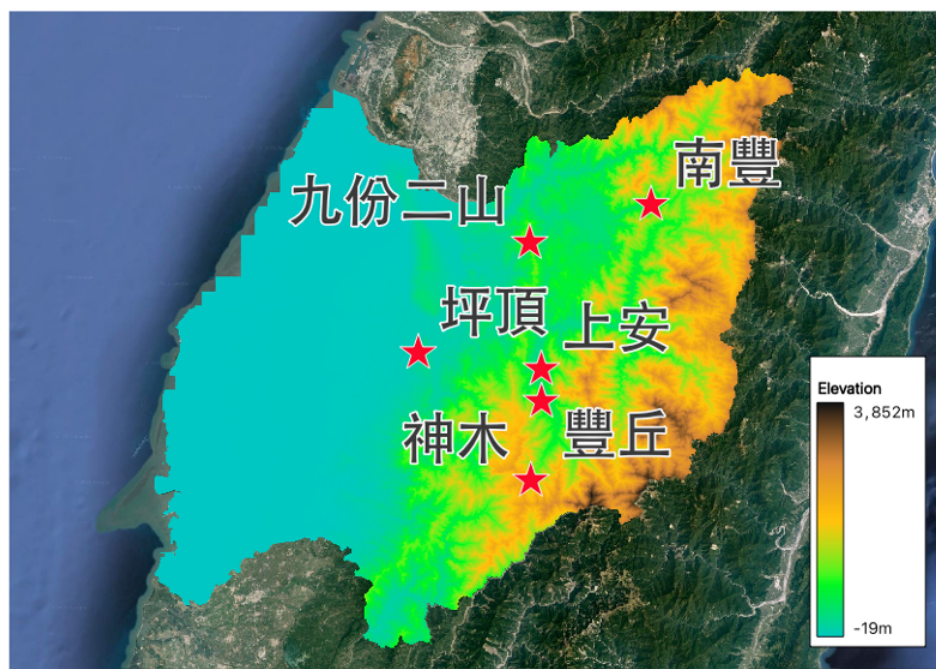


圖 2-11: 中台灣高程空間分布

表 2-1: 使用測站之地文資訊

土石流測站	TWD97X	TWD97Y	Slope (m)	Aspect	Elevation(m)
上安	235912	2623105	28.12	349.21	617
神木	233867	2601249	31.48	0	1629
豐丘	235943	2616757	36.26	171.17	1073
九份二山	233642	2647969	20.18	324.68	949
坪頂	211715	2626317	17.09	296.56	143
南豐	257516	2655457	37.85	242.19	969
來義	215204	2490323	41.4	340.53	594

3、水文資料

本研究使用之水文資料主為中央氣象局之雨量與濕度資料，將用以作為輸入資料與土壤含水量資料建構機器學習模式，因此根據兩項標準 (1) 資料品質與 (2) 距離土石流觀測站之距離進行篩選，所使用之雨量資料與濕度資料必須滿足缺漏值少以及鄰近土石流觀測站兩個條件，據此篩選出之測站展示如表 2-2

表 2-2: 各土石流測站所使用之 (左) 雨量與 (右) 濕度的自動測站

土石流測站	降雨使用之測站	距離	缺漏值比例 (%)	土石流測站.1	濕度使用之測站	距離.1	缺漏值比例 (%) .1
上安	信義	2677	0.08	上安	信義	2677	0.03
神木	神木村	2044	0.03	神木	神木村	2044	0.06
豐丘	豐丘	2990	0.48	豐丘	信義	4071	0.03
九份二山	蘆竹湳	2790	0.11	九份二山	中寮	9378	2.65
坪頂	林內	1769	0.15	坪頂	林內	1769	1.26
南豐	楓樹林	1907	0.46	南豐	仁愛	6268	0.13
來義	潮州	12760	2.25	來義	新埤	9487	3.39

表 2-2 中可發現所挑選出來之自動測站的大部分都距離土石流測站 5000m 且缺漏值不到 1%。一般來說，土壤含水量的多寡與瞬時降雨無關，與累積降雨較有關，且作為學習推估模型的輸入，與輸出相關性越高越佳，因此本計畫嘗試計算觀測時間點前 24 小時累積降

雨以及 12 小時前平均濕度作為輸入變數，並展示兩變數與對應測站之土壤含水量觀測值比較如圖 2-12與圖 2-13。

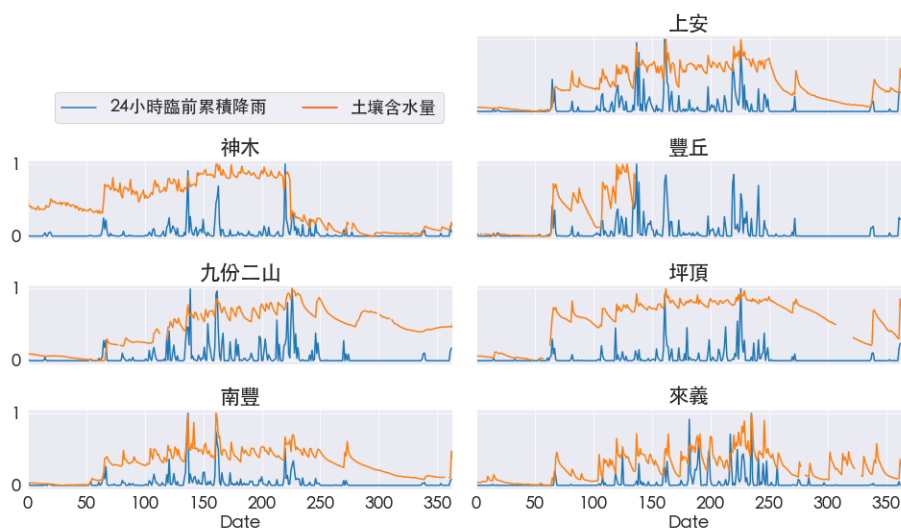


圖 2-12: 小時累積降雨與對應測站之土壤含水量比較圖

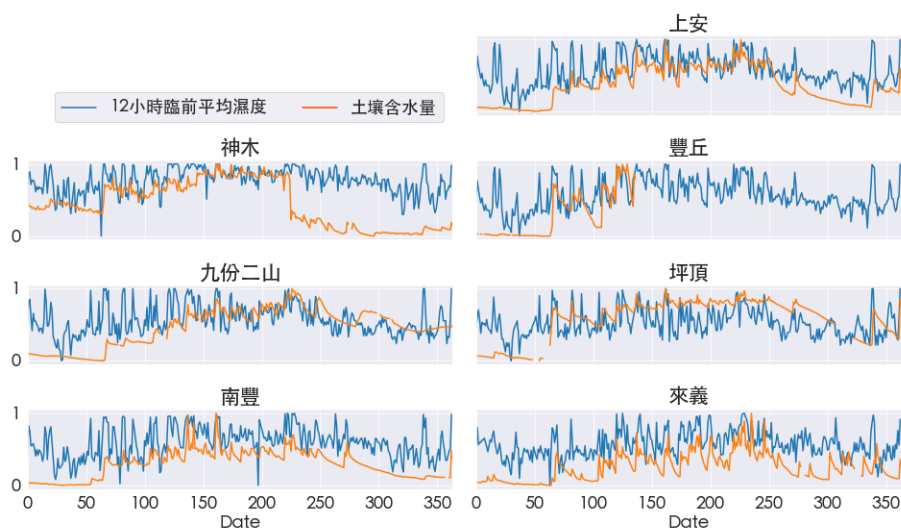


圖 2-13: 12 小時前平均濕度與對應測站之土壤含水量比較圖

從圖 2-12與圖 2-13中可發現某些土壤含水量與 24 小時臨前累積降雨有明顯相關，如南豐、來義；亦有某些土壤含水量與 12 小時臨前平均濕度有明顯相關，如來義、上安。

(十) 土壤含水量驗證

本計畫採用行政院農業委員會水土保持局土石流防災資訊網所提供的土壤含水量資料來進行推估結果的率定。本計畫將選擇至少四個測站，蒐集該處的土壤含水量歷史觀測數據，利用該歷史數據的時間和空間資訊，來擷取至本計畫推估取得之該點位的CWSI，繪製推估之CWSI以及土壤含水量的散布圖並且計算皮爾遜相關係數（Pearson correlation coefficient），以此驗證推估結果趨勢是否吻合與儀器監測結果。

1、土石流監測站資料蒐集

水土保持局共架設了 26 座土石流觀測站，分別位於大粗坑、蘇樂、司馬庫斯、玉峰、下田埔、白布帆、松鶴、上安、郡坑、神木、豐丘、九份二山、南豐、華山、坪頂、豐山、羌黃坑、集來、來義、和中、大興、鳳義坑、射馬干、大烏、寒溪及火炎山，測站資訊如表 2-4，空間分布如圖 2-4 所示。其中位於寒溪及火炎山的測站資料並未開放申請，其他 24 座觀測站之觀測資料皆可申請。由觀測資料供應平臺中，可查詢到最早的土壤含水量資料為 2008 年的觀測資料。土壤含水量計設定為每 5 分鐘會紀錄一筆觀測資料，不過偶有漏測之情形。

本計畫考量土壤含水量時空間資料的完整度，以求驗證的全面性，故蒐集所有測站的土壤含水量儀器數值作為驗證分析之用。

2、土壤含水量資料時空間參照

土壤含水量的數據為每五分鐘一筆，而高精度的土壤含水量推估為時間尺度為每日，表 2-4 整理了兩者的時空間差異。

土壤含水量儀偵測空間上的一個點位，透過二度分帶座標，不同測站的土壤含水量儀可以標記在高精度土壤含水量推估值的特定點位上，由此完成空間參照。至於時間參照，雖然說由衛星遙測所提供的土壤含水量結果是日資料，但其

表 2-3: 土石流監測站基本資訊

觀測站	高程 (m)	二度分帶 X 座標	二度分帶 Y 座標
寒溪	160	319453	2722634
大粗坑	196	334625	2776134
蘇樂	679	286109	2732677
火炎山	180	222486	2694463
白布帆	503	240423	2688002
松鶴	770	248525	2674657
上安	828	237412	2624605
郡坑	647	236766	2626219
神木	1187	235367	2602749
豐丘	633	237443	2618257
九份二山	538	235142	2649469
南豐	790	259016	2656957
華山	542	209977	2610061
坪頂	306	213215	2627817
豐山	933	223941	2609683
羌黃坑	274	201720	2552550
集來	400	207449	2547710
來義	400	216704	2491823
和中	19	325315	2685218
大興	259	288581	2615400
鳳義坑	236	293458	2626390
射馬干	229	253895	2513847
大烏	190	216704	2491823

表 2-4: 土壤含水量測量的時空間差異

	土壤含水量測量儀	高精度土壤含水量推估
時間尺度	每五分鐘	每日
空間尺度	點資料	面資料 (30 公尺解析度)

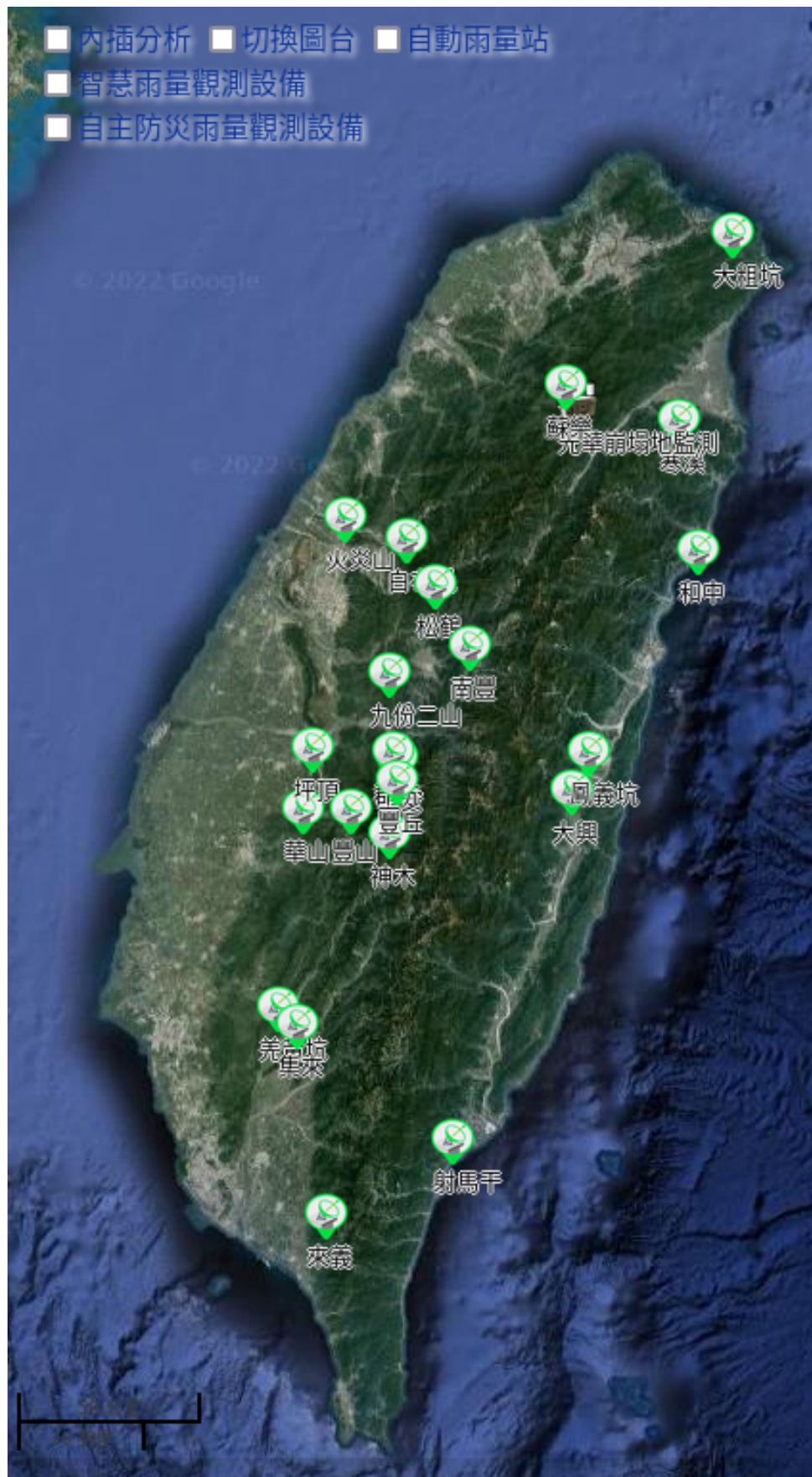


圖 2-14: 土石流觀測站空間分布圖。資料來源：土石流防災資訊網

實更為精準地說，該資料是於衛星拍攝時間點位作為推估依據，也就是每日當地時間上午 10 點至 11 點，由此依據，由土壤含水量儀器得到的土壤含水量數值，使用該地區每日上午 10 點至 11 點的所有觀測資料進行平均，以該數值當作當日土壤含水量儀器數值作為驗證資料使用。

第三章 研究進度及成果

一、衛星融合技術驗證

衛星驗證方法資料由 NDVI 作為融合資料的候選數值，由於 NDVI 可以代表植被生長的健康狀況，並且該數值在之後研究欲融合的地表實際蒸發散量中扮演重要的角色。本計畫透過人工隨機挖空區域以及真實觀測值的相關性與誤差，來檢視衛星融合推估模型的可靠性。

(一) 同質區域選取

本計畫挑選五種地貌進行融合算法的模擬，分別是高山(玉山群峰)、丘陵地(鹿谷鄉)、邊坡(古坑鄉)、草原(墾丁龍磐草原)、農地(嘉南平原)，該些區域的二度分帶座標於表 3-1 所示。

表 3-1: 同質區域的二度分帶座標邊界框

	邊界框 TWD97
高山	244510, 2594870, 247770, 2598512
丘陵地	223208, 2628609, 226857, 2632046
農地	178156, 2620069, 182022, 2623319
草原	232238, 2434186, 235783, 2437666
邊坡	208041, 2608476, 211575, 2612008

(二) 資料隨機挖空

選取單張 Landsat 的照片，高山、丘陵地、草原分布於 Landsat Path 117 的拍攝範圍，而農地、邊坡在 Landsat Path 118 的拍攝範圍。故依照 Landsat 不同 Path 拍攝之照片品質選擇一適合的日期作為觀測日期，Path 117 選擇 2021 年 11 月 16 日，而 Path 118 選擇 2021 年 9 月 4 日。

各區域之觀測值皆有些許雲遮蔽的情況發生，去除雲遮蔽的區域後，本研究使用重複選取的方式進行挖空模擬，每次模擬中選取 3x3 的區域。

(三) 融合資料推估

鑑於衛星融合空間解析度的需求，使用 Landsat 衛星資料作為真實觀測值進行驗證。再考慮到空間解析度，本研究以 MODIS 衛星資料作為衛星融合模型的輸入值，並且以 Landsat 衛星資料當日的前後兩天，共三天的 MODIS 衛星資料進行平均當作當日的觀測值。

MODIS 觀測值中隨機挖空 3x3 區域後利用 HISTARFM 方法進行補值，以下稱作預測地表反射值 (Predicted reflectance)。預測地表反射值推估出來後，會與原本點位上被挖空的數值進行比較。本研究使用 $RMSE$ 、 MAE 、 MSE 、 MBE 以及 r 判斷推估準確度，同時在補遺的區域，計算區域內的不確定性，以此得知補遺演算法在各個地貌區域上的穩定度差異。

透過 HISTARFM 的卡曼濾波器，產出最佳即時推估的結果，再透過具有時序性的偏差修正取得最終推估成果。推估結果依據各個地貌有些許的差異，結果列於表。高山區域使用卡曼濾波器作為最佳即時推估的結果最佳， $RMSE=0.0659$ 、 $MAE=0.0501$ 、 $MBE=0.0423$ 、 $MSE=0.00434$ 以及 $r=0.955$ 。相反地，農田在目前模型中的推估效果最差， $RMSE=0.209$ 、 $MAE=0.165$ 、 $MBE=0.0722$ 、 $MSE=0.0437$ 以及 $r=0.0949$ 。表顯示各個地貌區域的推估精準度結果，由佳到差排序依序為高山、丘陵地、邊坡、草原，最後是農田。此外，針對各個區域選定時間進行驗證，圖 3-1 顯示農田 (a) 和邊坡 (b) 在 2022 年 9 月 4 日的推估預測值和實際觀測值的散布圖，而圖 3-2 顯示高山 (a)、草原 (b) 和丘陵地 (c) 的在 2022 年 11 月 16 日的推估預測值和實際觀測值的散布圖。所有區域的預測結果皆有正偏差的結果產生，由 MBE 的結果從 0.00278 至 0.0437 的結果可得知，表示推估結果相較於實際觀測值較低。

邊坡、高山和丘陵的數值較為無偏估並且有較佳的準確度 (accuracy) 及精確度 (precision)，草原則是在實際 NDVI 較高的區域 (0.8-1.0) 預測值較難準確預測，預測結果的數值有較為劇烈的波動。農田的預測最為不精準，預測結果雖由正相關的趨勢但卻發散，呈現低度正相關的情況，由該區域的分布情勢可知，農田

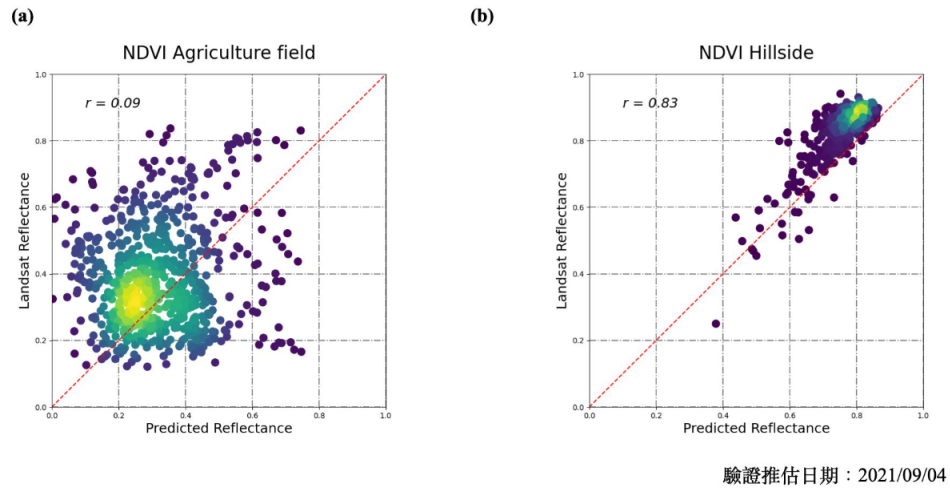


圖 3-1: 農田 (a) 以及邊坡 (b) 的預測地表反射值與實際 Landsat 觀測值之散布圖

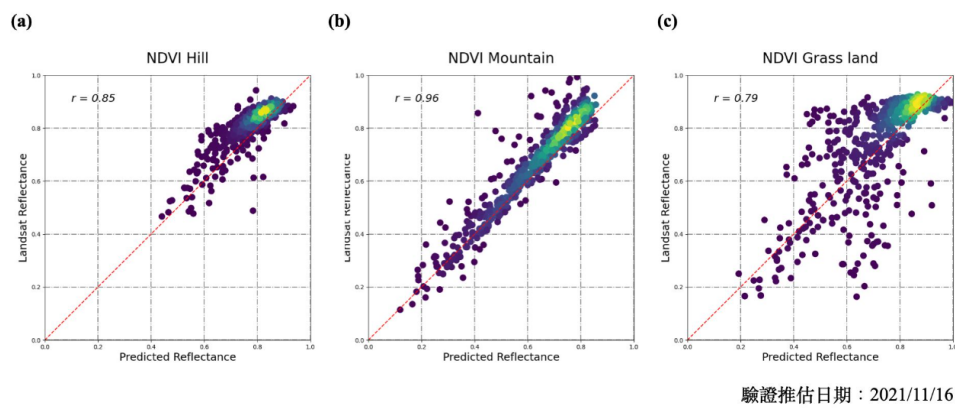


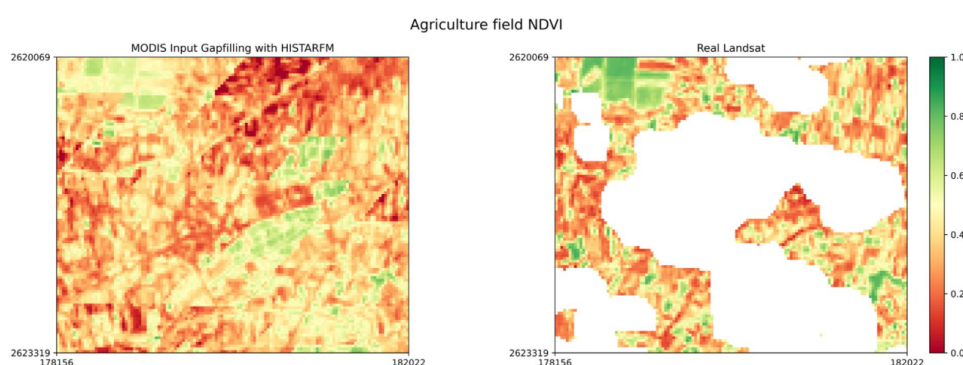
圖 3-2: 丘陵地 (a)、山區 (b) 和草原 (c) 的預測地表反射值與實際 Landsat 觀測值之散布圖

在以 MODIS 為底的 HISTARFM 的卡曼濾波器中，以 MAE 的結果來看，平均的絕對誤差高達 0.165。

表 3-2: 同質區域的模型表現性衡量指標

	農田	邊坡	高山	草原	丘陵
RMSE	0.209	0.0881	0.0659	0.11	0.0528
MAE	0.165	0.081	0.0501	0.074	0.0426
MBE	0.0722	0.0787	0.0423	0.024	0.0359
MSE	0.0437	0.00776	0.00434	0.0121	0.00278
r	0.0949	0.826	0.955	0.789	0.846

針對不同的區域，本研究繪製 HISTARFM 的推估結果圖。圖 3-3 為 2021 年 9 月 4 號於農田的推估結果，左邊為以 MODIS 為底的 HISTARFM 結果，圖 3-3 右為 Landsat 於當日農田的實際觀測值，空白區域為雲遮蔽的區域。推估結果雖然有達到 Landsat 衛星 30 公尺的解析度，不過在坵塊的推估，MODIS 為底的 HISTARFM 結果 (圖 3-3 左) 和 Landsat 實際觀測值 (圖 3-3 右) 有顯著的不同，並且 MODIS 為底的 HISTARFM 結果 (圖 3-3 左) 有明顯的斜狀橫帶。

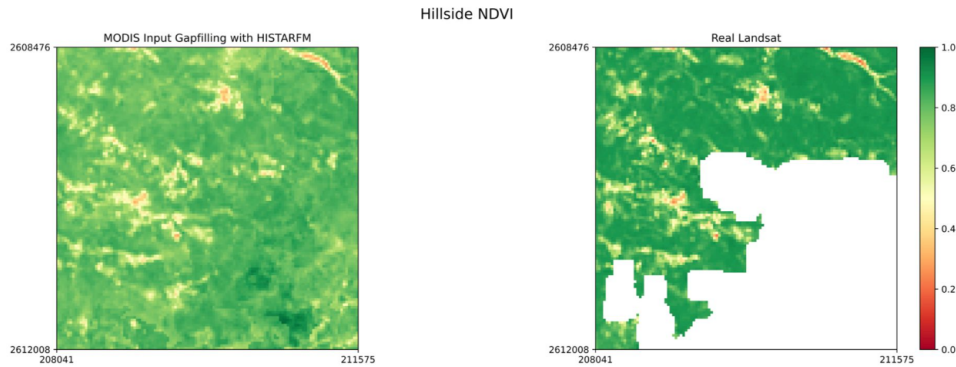


驗證推估日期：2021/09/04

圖 3-3: 農田的 HISTARFM 推估結果。(左) 以 MODIS 空間降級的 HISTARFM 推估結果，(右)Landsat 的觀測值

圖 3-4 左為 2021 年 9 月 4 號邊坡推估的結果，圖 3-4 右為 Landsat 於當日邊坡的實際觀測值，空白區域為雲遮蔽的區域。

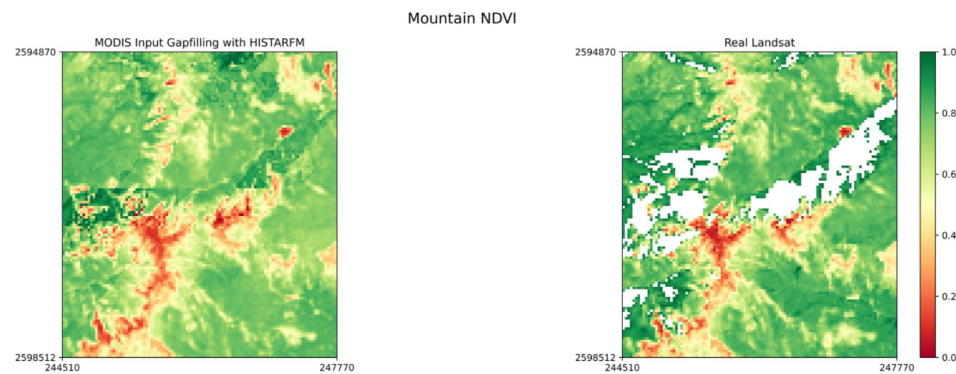
在邊坡的推估當中，整體推估的 NDVI 數值低於低於觀測值，在雲遮蔽缺值補遺的區域，HISTARFM 的推估結果 (圖 3-4 左) 無明顯的補遺痕跡。



驗證推估日期：2021/11/16

圖 3-4: 邊坡的 HISTARFM 推估結果。(左) 以 MODIS 空間降解的 HISTARFM 推估結果，(右)Landsat 的觀測值

圖 3-5 左為 2021 年 11 月 16 號於高山的推估結果，圖 3-5 右為 Landsat 於當日高山的實際觀測值。高山區域有部分的區域帶有平行四邊形的橫帶，在 NDVI 數值較低到區域 (圖中紅色區域)，以 MODIS 為底的 HISTARFM 結果具有不錯的推估補遺成果。

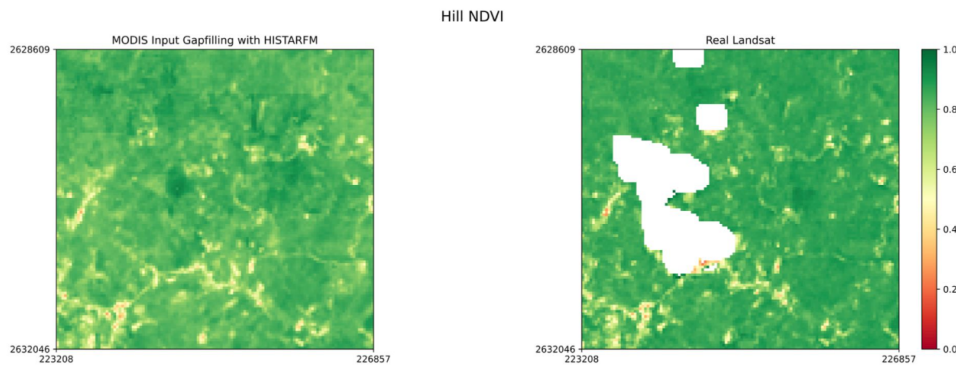


驗證推估日期：2021/11/16

圖 3-5: 高山的 HISTARFM 推估結果。(左) 以 MODIS 空間降解的 HISTARFM 推估結果，(右)Landsat 的觀測值

圖 3-6 左上為 2011 年 11 月 16 日丘陵地的推估結果，圖 3-6 右上為 Landsat 於當日丘陵地的實際觀測值。丘陵地上以 MODIS

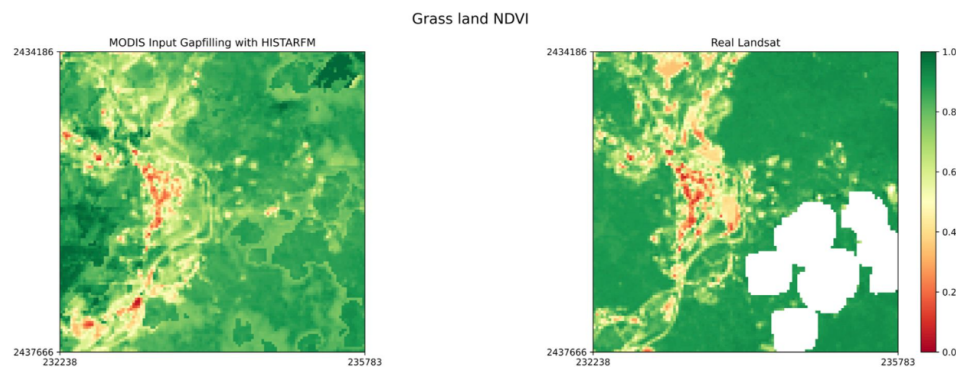
為底的 HISTARFM 結果依舊在高 NDVI 值域區域有低估的狀態。此外，丘陵地的推估結果並無明顯的橫帶出現。



驗證推估日期：2021/09/04

圖 3-6: 丘陵地的 HISTARFM 推估結果。(左) 以 MODIS 空間降級的 HISTARFM 推估結果，(右)Landsat 的觀測值

圖 3-7 左上為 2011 年 11 月 16 日草原的推估結果，圖 3-7 右上為 Landsat 於當日草原地的實際觀測值。Landsat 實際觀測值中，草地呈現均一化的植被健康狀態，但在以 MODIS 為底的 HISTARFM 結果卻有些許區域，沿著綠色區塊邊緣的區域呈現推估出較低的 NDVI 數值。



驗證推估日期：2021/11/16

圖 3-7: 草原的 HISTARFM 推估結果。(左) 以 MODIS 空間降級的 HISTARFM 推估結果，(右)Landsat 的觀測值

此外，針對補遺區域的數值推估，本研究擷取最後更新的誤差協方差 W_k ，作為判定不確定性的依據，圖 3-8 為所有觀測區

域的誤差協方差 W_k 。當中農田區域中有較多不確定性較高的區域，並且從圖 3-8 (a) 中可以看出是因為耕犁所產生的劇烈地表植被狀態變動所造成的，在其他地區則有少數區域有高不確定性。

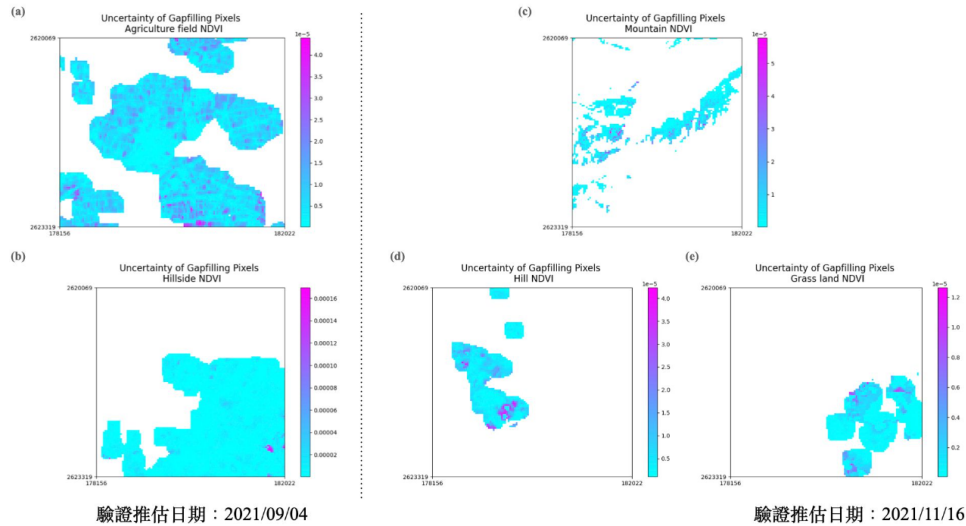


圖 3-8: 所有觀測區域的誤差協方差

二、土壤含水量推估驗證

PySEBAL 具有自動化挑選冷熱像素的演算功能，藉此可以實現自動化產製 SEBAL 圖層的目標，不過 PySEBAL 演算系統需要讀取衛星產品的元資料 (metadata)。由此特性，預先產製各個衛星產品衍生出的 SEBAL 圖層後再進行融合較為恰當。

本研究分別採用 MODIS 以及 Landsat 的衛星資料，再輔以 TC-CIP 的氣象資料，推估出地表的各蒸散量的面資料。藉由蒸發散量資訊，本研究先推估出個別衛星的作物水分逆境指數 CWSI，而後使用 HISTARFM 進行衛星融合產出每日高精度的 CWSI 圖層。

(一) 土壤含水量資料時空間參照

為了驗證該 CWSI 能夠代表土含水量，本研究對比水保局土石流測站以及衛星資料推估結果，最終選定以 2019 年的全年資料進行推估驗證。

首先，選定 2019 年運作中的土石流測站，擷取土壤含水量的資料和該些測站的地理座標。選定地理坐標後，本研究以該些

點作為中心點，選定 100x100 的空間資料進行全年度的 CWSI 圖層之衛星融合。由於某些測站周邊的空間資料點位尚有資料缺失的情況，最終選定 7 座土石流測站進行最後推估的驗證，分別為上安、神木、豐丘、九份二山、坪頂、南豐以及來義，該些測站延伸出來的關注區 (area of interest) 如表 3-3

表 3-3: 目標土石流觀測站的關注區

土石流測站	土壤含水量關注區 TWD97
上安	235912, 2623105, 238912, 2626105
神木	233867, 2601249, 236867, 2604249
豐丘	235943, 2616757, 238943, 2619757
九份二山	233642, 2647969, 236642, 2650969
坪頂	211715, 2626317, 214715, 2629317
南豐	257516, 2655457, 260516, 2658457
來義	215204, 2490323, 218204, 2493323

(二) 土壤含水量資料推估

由於驗證需求，本研究挑選小區域，以欲預測的數值點 (目標土石流觀測站) 為中心點，向外延伸 100x100 的小區域，藉由 HISTARFM 時空間關聯的特性，最終推估目標點位的 CWSI 指標。針對目標的七個測站，本研究擷取每日 CWSI 衛星融合圖層，對於土石流測站的點資料作為每日的衛星推估 CWSI 數值。另一方面，針對土石流測站的土壤含水量，本研究擷取每日的資料中 10 a.m. 至 11a.m. 的資料，將五分鐘一筆的資料進行平均，作為現地觀測值。

本計畫以 CWSI、高程、坡度、坡向、x 座標、y 座標、臨前 24 累積降雨、臨前 12 小時平均濕度共八種輸入變數，與土壤含水量建立 XGBoost 機器學習模型，流程圖參照一般機器學習流程如圖 3-9，並以 Backward elimination 搭配特 k-fold cross validation 的方式進行特徵選擇，評估模式優劣的指標則選擇 R square，方式如下：

- 1. 首先隨機將資料的七成作為訓練集, 三成作為測試集
- 2. 於訓練集中利用 4-fold cross validation 挑選重要特徵
- 3. 分別指定挑選前 2、3、4 重要的參數，結果如表三

- 4. 利用最重要的兩個特徵以測試集驗證模式表現，結果如圖 3-10

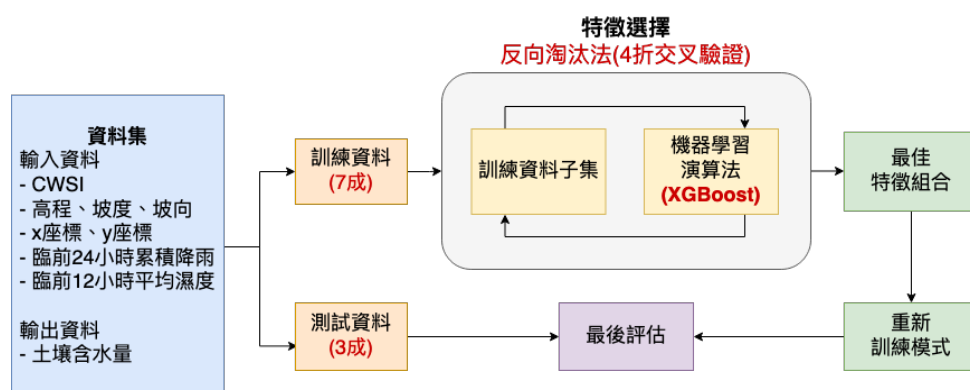


圖 3-9: 本計畫建構機器學習模型流程圖

表 3-4: 特徵選擇結果

特徵數量	所選擇之特徵	4-fold 交叉驗證的平均 R^2
2	CWSI、高程	0.76
3	CWSI、高程、24 小時臨前累積雨量	0.805
4	CWSI、高程、24 小時臨前累積雨量、坡度	0.818

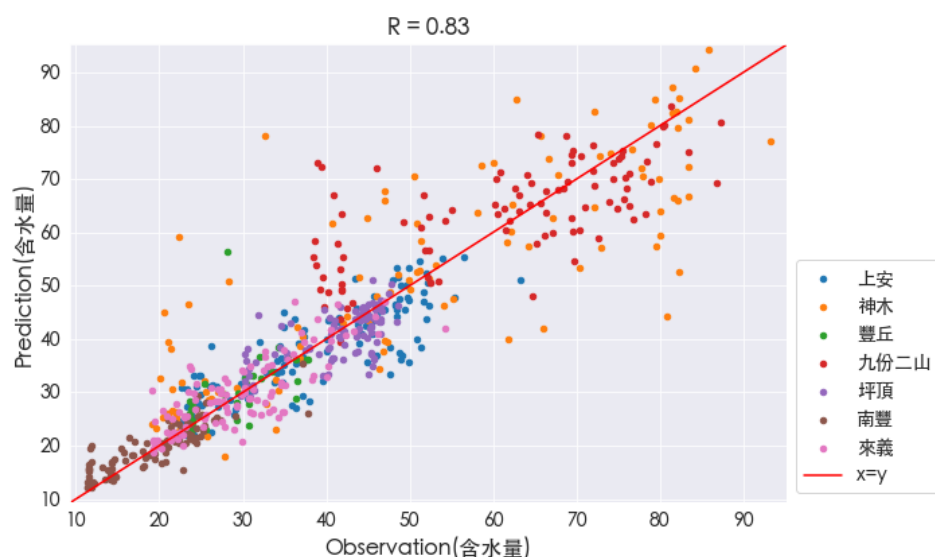


圖 3-10: 以 CWSI 與高程為輸入特徵之模型驗證結果

由表三中之特徵選擇結果來看，僅以 CWSI 與高程兩特徵以 k-

fold cross validation 評估模式表現對於土壤含水量預測來說已經足夠好，R square 的結果為 0.76，若選擇保留三個特徵，結果是 CWSI、高程與 24 小時累積降雨會被選入，恰巧是衛星資料、地文與水文資料都各被納入一種，R square 可再多增加 0.045，達到 0.8 以上。後續若選擇保留四個特徵，會再額外保留坡度，然而 R square 的增加就比較不顯著，話句話說其所提供的資訊有限。為使參數精簡化，本研究僅以 CWSI 與高程兩特徵作為輸入變數建立 XGBoost 模型，以剩下三成之測試資料進行驗證，結果如圖 3-10，其 R square 值為 0.81，換句話說，僅以 CWSI 與高程兩變數即可達到不錯的預測精度，綜上分析，本計畫基於衛星影像所計算之 CWSI 能有效的提供預測土壤含水量所需的資訊。

第四章 研究貢獻與侷限性

一、討論與建議

(一) 衛星融合技術

本研究開發之衛星融合演算法，經由融合算法的驗證結果得知，在農地的補遺精準度和準確性皆欠佳，該原因為衛星融合技術中尚無法偵測過於頻繁變動的區域，為了若要將此衛星融合技術應用於高變動的區域，以下三點需要再深入研究。

1、貝氏模型平均的優化

目前貝氏模型平均，在偵測地表變動的模型為 MODIS 衛星所衍生的空間分解 (spatial disaggregation) 後的資料作為輸入。該資料有兩個缺點，一是當該年度特定點位的歷史資料因為雲遮蔽或是其他異常值出現而導致資料量低，與 Landsat 相對應點位進行線性映射時，該擬和會有較差的穩健性。第二點，和第一點相關，當線性映射的結果較不穩健，在生成斜方差矩陣時，自身的變異數就大，此現象導致在模型平均時，來自於 MODIS 衛星的高時間解析度資料被採納的權重較低，故較難偵測到變化劇烈的區域。

本研究認為若能將原本利用線性映射方法產生的高時間解析度模型加以優化，將有助於提升農田至於其他高度平凡變動區域的最佳推估結果。

2、 γ 的動態調整

本研究在卡曼濾波器的算法開發中，尚未考慮 γ 值所帶來的影響，將卡曼濾波器的 γ 值統一設定為 0.6。 γ 在 HISTARFM 當中扮演著調控最佳即時推估和偏差修正的佔比多寡。理論上，針對區域的變動程度，可以透過調整 γ 來增進模型的預測準確度。

3、偏差修正的優化

本研究在五個區域（高山、丘陵地、農地、草原以及坡地）的推估結果，*MBE* 的數值皆為正，表示在每個區域，此推估方法造成了預測低估的情況發生。

回顧期中報告的結果。當時研究進度是先透過 HIS-TARFM 的卡曼濾波器，產出最佳即時推估的結果，尚未加入偏差修正來更新最終結果。圖 4-1 顯示期中報告時，尚未使用偏差修正即產出之卡曼濾波器最佳即時推估結果，圖中各個區域尚未有預測數值出現偏差的情況。由此可知，偏差修正的運算之中，尚有未釐清的地方，導致加入後推估造成每個區域皆有一致性偏差產生，這點需要後續研究持續校正與優化。值得一提的是，當時期中報告所帶入，作為衛星當日觀測的資料為 Landsat，而不是 MODIS 透過空間分解的資料，故該操作也有可能是造成結論出現一致性偏差的原因。

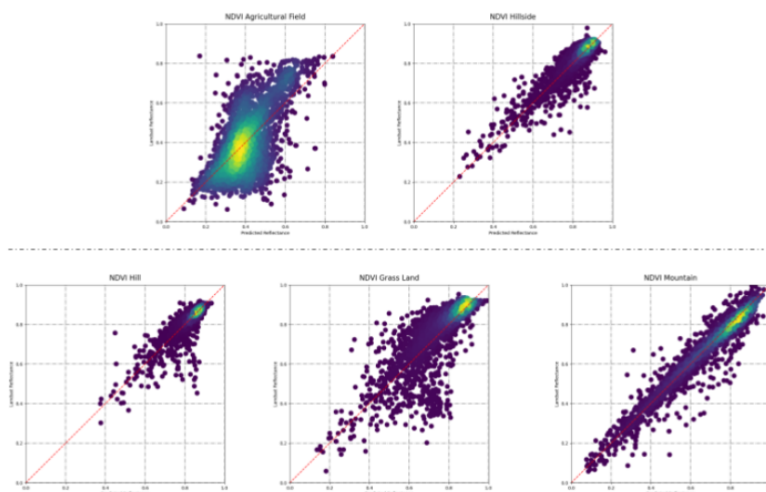


圖 4-1: 實際與卡曼濾波器最佳即時推估產製之預測地表反射值 (NDVI) 之散布圖

二、成果貢獻

本研究率先將兩個不同類型的演算推估方法合併進行臺灣山坡地區的土壤含水量推估，以達到高時空間解析度的土壤含水量推估圖層。在參考文獻之際，針對於臺灣的環境，兩個別演算方法皆有創新精進，以達到的高時空解析度土壤含水量監測系統為目的。

(一) 初步衛星融合技術的功能建置

本研究使用兩個開源衛星，Landsat 以及 MODIS，該衛星不僅僅提供多光譜的地表反射數值 (surface reflectance)，亦有提供地表反照率 (surface albedo) 以及地表溫度 (surface temperature) 等資訊。雖然兩個衛星在時間或是空間解析度上並非現有產品中最佳，但是卻是最易取得且公開資訊完整的開源衛星。

HISTARFM 衛星融合技術，有辦法產出每日 30 公尺的衛星觀測資料。整體來看，目前的推估結果在人類較少擾動的區域，例如山區邊坡、高山以及丘陵地等有不錯的推估結果，且圖層整體沒有出現明顯的橫帶，表示在補遺區域，HISTARFM 技術有辦法良好的捕捉空間中的相關性，使得整體推估結果較為平滑。

未來該技術將不受限於 Landsat 以及 MODIS 兩個衛星，也不受限於推估本次所需求的 SEBAL 圖層作物水分逆境指數 CWSI 或者是常態化植被指數 NDVI。可以針對需求，選定目標圖層，選取目標衛星，適應不同情況融合衛星資料取得高解析度的遙測資料。

1、土壤含水量預測參數研發

作物水分逆境指數 CWSI 是一個用來判斷植物缺水逆境的一個指標，該指標透過植物生長狀態、植被覆蓋地區的地表蒸發散量以及各項氣候因子求得。該方法雖尚無法明確的指出土壤的實際含水量，但可以透過和現地觀測站中土壤含水量觀測儀的數值比對，得到其相關性。再者，透過此指標亦可以大面積的取得地區面資料的土壤缺少或豐水程度。

本研究透過 PySEBAL，一個可以自動選取冷熱像素點位後推估植物蒸發散量的演算法。相較於過往的 SEBAL 演算法，需要用肉眼去挑選定位冷熱的錨點位才可計算，PySEBAL 可以更有效的自動產生圖層，對於防災預警等需自動化的系統頗有助益。

CWSI 擷取氣象資料 (風速、日照、氣溫以及濕度) 以及遙測資料 (地表溫度、地表反射率以及反照率) 等資訊，該圖層

今後可以與前人研究，例如合成孔徑雷達 (Synthetic aperture radar, SAR)、觀測站附近的地貌以及土壤質地等因素共同來預測現地的土壤含水量。

三、未來展望

本年度的創新專案開發了衛星融合的演算法以及自動化的植被蒸發散推估。因為開發時間時程的關係，兩者皆尚有需要調整優化的地方。

在融合演算法中，本研究計畫完善了個別衛星資料的資料庫建置，和各項地理空間資料的存取及運算功能。這些基本的功能顯著地對日後加速推估演算法的修正和優化等各項的開發有所助益。具體的開發，本研究後續將針對貝氏模型平均中，代表即時資訊的模型進行優化（即為 MODIS 空間降解為資料之模型）。本研究希望取代原本簡單線性映射的方式，以下三個方向發展。(1) 從空間特性下手，由與空間資料相關的空間降解方法，例如卷積神經網路 (Convolutional neural network, CNN)。(2) 從時間特性下手，考慮地表變化的季節性，將同個季節特性的圖層進行擬和。(3) 增加最佳先驗推估的資料量，嘗試融入不只兩顆衛星的其他衛星資料，來增進最佳先驗推估的結果。

本計畫後續將深入研究 CWSI 在土壤含水量偵測的預期深度，預計將建置新的土壤含水量偵測儀至研究的關注區，並且將土壤深度、土壤性質以及植被狀態等等因素考慮進來，以求得 CWSI 確切所代表的土壤深度含水量為目標，這將能夠有助於土石流災害預警上，在不同深度以及不同土壤質地上的土壤含水量預測，延伸至災害預警上起到很大的幫助。

参 考 文 献

- Baranowski, P., Mazurek, W., & Walczak, R. T. (2001). The relation of actual to potential evapotranspiration as an indicator of plant water stress for meadow plant cover. *Acta Agrophysica*, 53.
- Dhungel, S., & Barber, M. E. (2018). Estimating calibration variability in evapotranspiration derived from a satellite-based energy balance model. *Remote sensing*, 10(11), 1695.
- Feng, L. (2015). *Sensitivity analysis of hot/cold pixel selection in sebal model for et estimation* (Doctoral dissertation). Virginia Tech.
- Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006). On the blending of the landsat and modis surface reflectance: Predicting daily landsat surface reflectance. *IEEE Transactions on Geoscience and Remote sensing*, 44(8), 2207–2218.
- Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: A tutorial (with comments by m. clyde, david draper and ei george, and a rejoinder by the authors. *Statistical science*, 14(4), 382–417.
- Huete, A. (1988). A soil-adjusted vegetation index (savi). *remote ssensing of environment*, 25, 295-309.
- Johnson, F. S. (1954). The solar constant. *Journal of Atmospheric Sciences*, 11(6), 431–439.
- Kustas, W. P., Norman, J. M., Anderson, M. C., & French, A. N. (2003). Estimating subpixel surface temperatures and energy fluxes from the vegetation index–radiometric temperature relationship. *Remote sensing of environment*, 85(4), 429–440.
- Liang, S. (2001). Narrowband to broadband conversions of land surface albedo i: Algorithms. *Remote sensing of environment*, 76(2), 213–238.
- Mangus, D. L., Sharda, A., & Zhang, N. (2016). Development and evaluation of thermal infrared imaging system for high spatial and temporal resolution crop water stress monitoring of corn within a greenhouse. *Computers and Electronics in Agriculture*, 121, 149–159.

- Moreno-Martínez, Á., Izquierdo-Verdiguier, E., Maneta, M. P., Camps-Valls, G., Robinson, N., Muñoz-Marí, J., Sedano, F., Clinton, N., & Running, S. W. (2020). Multispectral high resolution sensor fusion for smoothing and gap-filling in the cloud. *Remote Sensing of Environment*, 247, 111901.
- Organization, I. C. A. (1993). *Manual of the icao standard atmosphere: Extended to 80 kilometres (262 500 feet)* (Vol. 7488).
- Ru, C., Hu, X., Wang, W., Ran, H., Song, T., & Guo, Y. (2020). Evaluation of the crop water stress index as an indicator for the diagnosis of grapevine water deficiency in greenhouses. *Horticulturae*, 6(4), 86.
- Tasumi, M., Allen, R. G., & Trezza, R. (2008). At-surface reflectance and albedo from satellite for operational calculation of land surface energy balance. *Journal of hydrologic engineering*, 13(2), 51–63.
- UNESCO-IHE. (2018). Pysebal.
- Wang, J., & Huang, B. (2017). A rigorously-weighted spatiotemporal fusion model with uncertainty analysis. *Remote Sensing*, 9(10), 990.
- Wukelic, G., Gibbons, D., Martucci, L., & Foote, H. (1989). Radiometric calibration of landsat thematic mapper thermal band. *Remote sensing of environment*, 28, 339–347.
- 交通部中央氣象局. (2009). 衛星遙測台灣地區土壤含水量與降雨量之研究.
- 余化龍. (2016). 時空巨量資料推估技術於坡地災害風險預警之發展應用.
- 吳怡瑩，劉哲欣，張志新. (2013). 降雨量與表層土壤含水量關係之研究.
- 行政院國家科學委員會. (2012). 遙測衛星影像應用於地表乾燥指標及地表土壤含水量之研究.
- 行政院農業委員會水土保持局. (2010). 以 smap 衛星遙測配合 sar 雷達回波評估土壤含水量分布與崩塌關聯分析.

附錄一 審查意見回覆

期初審查

1. 開會事由：111 年度創新研究計畫 (新興科技應用) 期初審查會議 (場次一)
2. 開會時間：111 年 4 月 13 日 (星期三) 下午 2 時
3. 開會地點：本局府前辦公室第二會議室 (南投市省府路 2 號)
4. 主持人：陳副總工程司振宇
5. 紀錄：高副工程司百毅
6. 參加單位及人員：(詳會議簽名冊)
7. 主持人致詞：(略)
8. 業務單位報告：(略)
9. 國立臺灣大學簡報：(略)
10. 審查意見及討論：

一、張竝瑜委員

(一) 可見光及紅外光影像的測深如何？反映的土壤含水量是深度多少範圍內的資料？如何排除地表逕流影響？

研究方法為透過山坡地上覆蓋植生的狀態搭配環境的氣象因子去推導，測量的數值在地表階層而並未像雷達穿透土表。透過該方法可以取得植物根部部分的土壤含水量情況，文獻指出在 0-40 公分的土壤含水量較為精確 (余化龍, 2016)，已經在第一章第二節頁數 1-4 之其他環境參數對土壤含水量預測之影響中說明。此外，衛星觀測的地面數據主要為植被狀態，且時間為每天，若發生土壤逕流產生異常值考慮將會剔除並用前後時間進行內插。

(二) 為何不納入雷達影像反射率的資料，因為波長較長，測深可達 1 公尺左右。

關於使用合成孔徑雷達 (Synthetic aperture radar, SAR)，本研究經過文獻回顧和資料蒐集後，尚未找到與 Sentinel-1 衛星產品相似的衛星產品，故在這點上本研究方法的融合方式較為困難。該方法前人研究正在進行，將 Sentinel-1 SAR 觀測數據加入進行土壤含水量的推估 (Ru et al., 2020)。未來期許在不同層測的土壤含水量監測起到互補的作用。

(三) 建議可再設立長期土壤含水率測試站 2 到 3 處，利用 TDR 進行長期資料監測作為校正衛星遙測所判釋的含水量之用。

謝謝建議，本研究目前由於經費的問題，暫無計畫架設土壤含水率儀，但已經擴蒐集至全臺灣各處土石流監測站之土壤含水量，內容修改見第二章第一節頁數 2-2 之土壤含水量驗證。

(四) 山區受植被影響，如何排除植被因素？(例如不同植生，生長狀況)

預計後續加入調查觀測地區植被的研究，將觀測站附近的植被分化成森林、草原、裸露土地等等 (已實際視察狀況而定)，加入參數檢視植被對於此方法觀測土壤含水量是否有影響。同時也會把 NDVI 的數值嘗試納入建模當中，其對於植被生長期況具有充分的代表性。

(五) 如何排除大氣狀況或人為擾動 (如污染影響)？

我們選用 Landsat 和 Modis 大氣校正後的產品進行研究，在建模以及後續資料的導入流程，皆會撰寫異常值 (受到雲、空氣汙染等) 偵測及資料清洗的步驟，確保模型能保有最佳的推估。

二、楊明德委員

(一) 如何蒐集到較廣值域之土壤含水量現地資料？以利模式適用範圍。

目前對於土壤含水量的資訊，目的是要反映山坡地的土壤含水量情況，於此蒐集土石流觀測站的對應數據是較能被應用在實際的

情況下。於此對應的蒐集對策，我們將從僅蒐集一站的資料擴增到所有土石流觀測站的土壤含水量數據，藉此可以取得更廣值域的數值，內容修改見第二章第一節頁數 2-2 之土壤含水量驗證。

(二) 如何有系統排除環境及植生影響效應。

預計後續加入調查觀測地區植被的研究，將觀測站附近的植被分化成森林、草原、裸露土地等等(已實際視察狀況而定)，加入參數檢視植被對於此方法觀測土壤含水量是否有影響。同時也會把 NDVI 的數值嘗試納入建模當中，其對於植被生長期況具有充分的代表性。另外會再加入其他環境因子例如坡地、坡向、土壤質地等因子。

(三) 高精度需定義，在空間、時間或土壤含水量值域上之精度？

高精度定義為將時空間提升至單一衛星無法達成的成果。舉 Landsat 和 Modis 為例，該兩個產品能夠透過模式取得之最高精度的觀測值(每日 30m 空間解析度)。若未來取得許可加入其他衛星如 SPOT 或是 Sentinel-2 進行融合，將可在各個衛星中取其所長加以提升精度，不過前提是在觀測資料上，衛星之間所觀測的數據必須是相同或是近乎相同的。舉例來說，在計算 CWSI 的方法中需要用到地表溫度的圖層，該圖層目前僅有 Landsat 和 Modis 的產品中有。

三、葉昭憲委員

(一) 影像資料融合模組之推估器修改，如何比較方法間之準確性？

準確性 (Accuracy) 的評估，將會透過兩個模型推算出的結果與實際數值進行比較，該驗證方法將參照 Moreno-Martínez (2020) 的研究方法，但是把該模型的參照從原本 STARFM-HISTARFM 替換成 HISTARFM-HISTARFM modified by BME 這兩個模型，對於同一個點位同一個時段進行數據比對，透過 RMSE、MAE 以及 R^2 等指標驗證推估準確度的精進程度 (Ru et al., 2020)，詳細驗證方法見第二章第三節頁數 2-22 執行步驟第 (七) 項之衛星融合驗證。

(二) 氣象資料影響模式建立，建議評估土石流監測站既有氣候觀測資料進行之可能性。

謝謝委員的建議。目前的研究對於其他影響因子納入考量，並且整理至第一章第二節頁數 1-4 之其他環境參數對土壤含水量預測之影響。

(三) 建議期末能呈現模式參數檢定及模式驗證之成果。

謝謝委員的建議。模式驗證的方法將採用與 Moreno-Martínez (2020) 論文中的準確性評估，並且提供視覺化圖表和推估結果展現，詳細驗證方法見第二章第三節頁數 2-22 執行步驟第(七)項之衛星融合驗證，展現內容形式將以期中報告成果中第三章第一節頁數 3-2 執行步驟項目(七)衛星融合驗證之融合資料推估類似的方法展現。

四、劉正千委員

(一) 降雨和地震是觸發土砂災害的關鍵因子，但都受到土壤含水率的影響，如果能夠透過遙測的手段來推估土壤含水率，會是很重要的資訊。

謝謝委員的建議。這點我們將列入未來研究項目的考量當中，對於山坡地淺層土壤含水量與坡地災害的關係，針對此項目進行文獻回顧，於第一章第二節頁數 1-4 之其他環境參數對土壤含水量預測之影響之中。

(二) 考慮遙測影像來源的空間分辨率與時間間隔，以及目前推估方法可以達到的精度範圍，請總結是否真正能夠達到這項研究的目標。

謝謝委員的提問，我們認為每天的資料間隔去除掉可能的干擾因子，對於山坡地的觀測可以達到每周不只一筆的資料，在未來配合氣象預報，即可獲得近日的土壤含水量納入預警的參照中。

五、陳振宇委員

(一) 使用衛星資料進行面的土壤含水量之變化，如以現地土壤含水量計之單點變化資料作為驗證，是否需要特別處理？

考量到現地土壤含水量計每五分鐘一筆的時間尺度，以及在地圖上僅有測量一個像素大小的空間尺度，本研究匯透過時間空參照來對於現地土壤含水量資料進行處理，詳細處理方式於第二章第三節頁數 2-33 之土壤含水量資料時空間參照中有詳細說明。

(二) 是否考慮將坡度、植生等地形因子納入推估，或作為修正因子？

這點將會納入考量最為修正因子。本研究之方法對於土壤含水量推估為其中一個變因，旨在取得該變因的資料。參考前人對於土壤含水量的推估研究，後續有考慮坡度、坡向、地質地貌等地形因子以及植被狀態 (NDVI 及地表覆蓋資訊) 納入成為修正因子，相關的文獻回顧整理於第一章第二節頁數 1-4 之其他環境參數對土壤含水量預測之影響。

六、技研小組 (莊承穎)

(一) 文中雖列出各式衛星應用及方法，建議增加說明遙測含水量的系統之建立方法，以利瞭解此計畫內容。

謝謝委員的建議。在往後的報告當中將會區分衛星的資料處理與融合以及遙測土壤含水量推估之算法。本計畫實質上是將兩個獨立的演算法銜接在一起以達到高精度推估土壤含水量的目的，簡要的流程圖可以參照第二章第二節頁數 2-2 之研究方法。

(二) 建議說明此系統可應用於遙測土壤含水量的土壤深度。

謝謝委員的建議。透過本計畫之方法可以取得植物根部部分的土壤含水量情況，文獻指出在 0-40 公分的土壤含水量較為精確 (余化龍, 2016)，於第一章第二節頁數 1-4 之其他環境參數對土壤含水量預測之影響有做文獻探討及整理。

期中審查

七、張竝瑜委員

(一) 報告敘述中請清楚列出此項技術，運用於台灣山區可以達到之空間解析度及時間解析能力，並應比較與先前文獻所採用之方法，在時間與空間解析度上，本研究所增進之部分。

比較先前研究使用 MODIS 單顆衛星進行實際蒸發散量的推估，本研究融合 Landsat 衛星將空間解析度由 500m 降至 30m 而時間解析度維持不變。

附表 1-1: 解析度

研究名稱	時間解析度	空間解析度
時空巨量資料推估技術於坡地災害風險預警之發展應用 (余化龍, 2016)	1 day	500 m
高精度衛星遙測土壤含水量系統建立	1 day	30 m

(二) 對於衛星資料，建議團隊除了由資料庫中下載分析完成之反射率數據資料外，可以適度了解 MODIS Terra 及 Landsat 衛星不同感測器與頻段所提供之數據物理意義，在處理資料時，可以再選擇其中較能反映土壤含水量之光譜頻段資料加以分類分析，或能較為貼近實際之物理分析需求。

謝謝委員的建議，本研究首先目標首為建置演算流程，後續會針對衛星探究其物理意義以及衛星資料的校正方法，進行更詳盡的分析，或是開發針對土壤含水量之光譜頻段以及其衍生數據指標。

(三) 研究團隊預計剔除地表逕流之異常值，然而發生土石流破壞之前後期間，可能就是產生持續降雨，發生大量地表逕流之時段，若加以剔除，反而可能對於土壤飽和度之極限值附近之觀測產生偏差，建議研究團隊應考量如何針對此一可能偏差，進行估算與校正。

謝謝委員的提問。的確若地表逕流時段的資料被剔除，那就無法反饋當時已經超出臨界值的土壤含水量，使其模型推論的結果產生誤差。本計畫將於推估結果出來後，針對該狀況進行分析，了解它對模型的影響。

關於地表逕流對於土石流發生自然是極為重要，而關於這點本研究開發之工具可能不適用。不只如此，本研究的結果時間解析度落在一天，此尺度在即將山洪爆發時也依然是不足即時偵測預警。但反過來說，本研究雖不能反饋一天以內即時的訊息，卻能以每天的資料尺度來觀測山坡地的狀況。透過這個資訊，配合中央氣象局的大雨警報，可以在大雨發生之前事先定位高度風險區域，使在大雨發生之前預警土石流發生。

(四) 台灣目前運用合成孔徑雷達資料進行土壤含水率推估之研究，上有許多其他團隊進行中，建議研究團隊可尋求相關團隊合作，利用雷達遙測資料推估之結果，比較驗證本研究之產出成果。

謝謝委員的建議。本計畫旨在開發一套土壤含水量的推估方法以及衛星融合技術，的確可以在往後與相關團隊合作，綜合多方的技術建立更全面的山坡地土石流預警。但目前此技術尚未考慮納入雷達遙測資料，不過本計畫目前臺灣氣候變遷推估資訊與調適知識平台計畫 (Taiwan Climate Change Projection and Information Platform, TCCIP) 與取得土壤含水量，預計使用該資料進行面資料的驗證。

八、黃效禹委員

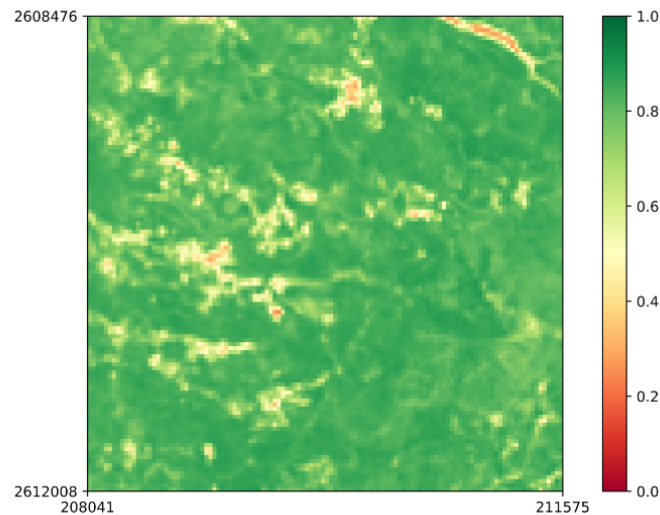
(一) 本計畫所需資料處理、模型建立說明完整，惟欠缺整體執行項目之工作進度說明，不易瞭解完成度，請補充。

謝謝委員的建議。在期末報告的結果部份，本計畫將加入整體執行項目之工作進度說明，於此先說明期中完成的進度。

一、融合衛星最佳先驗推估：貝氏模型平均

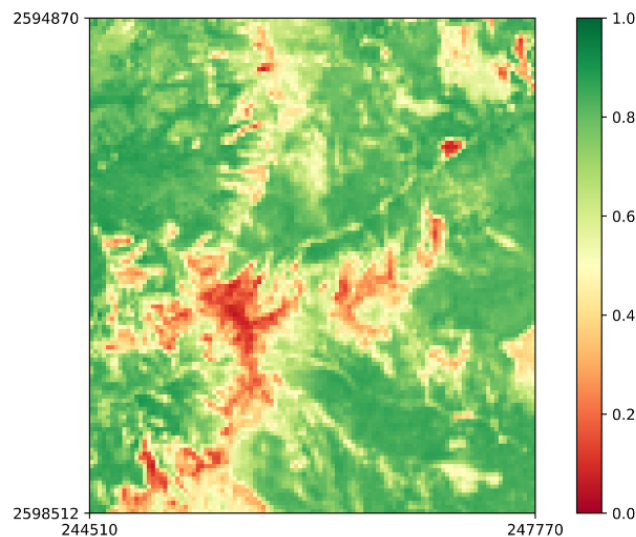
開發完成。本計畫建立 MODIS 和 Landsat 的衛星資料庫，透過使用並且修改部份 HISTARFM 文章中演算法的參數，求得當前最佳之先驗推估成果。圖 1-5 為五個目標區域(邊坡、高山、丘陵、草原以及農田)之 BMA 估算成果。

Bayesian Model Averaging (BMA) : Hillside



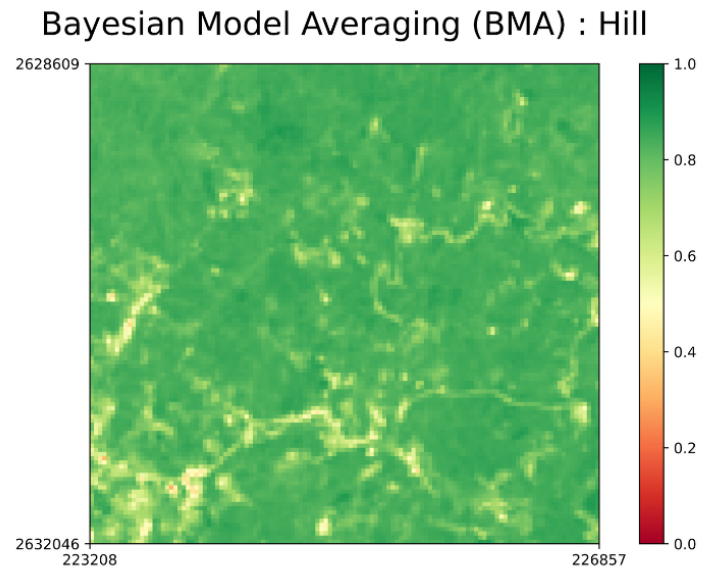
附圖 1-1: 邊坡古坑鄉的最佳先驗推估結果 (20220904)

Bayesian Model Averaging (BMA) : Mountain

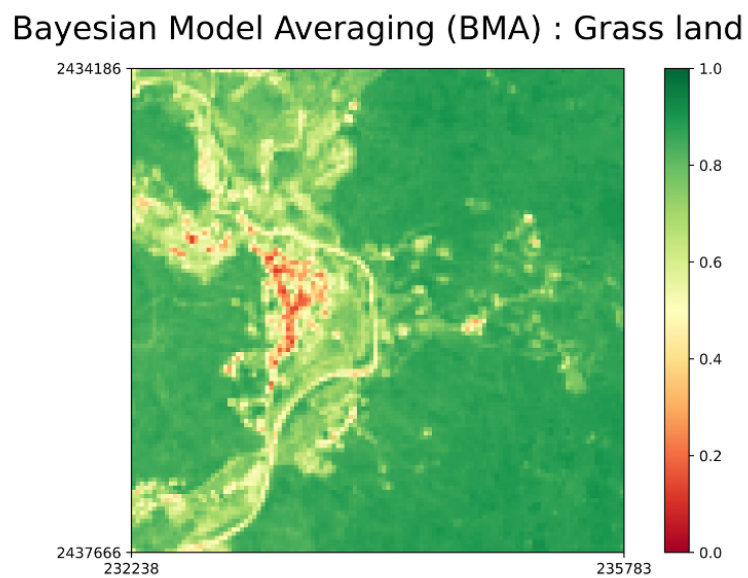


附圖 1-2: 高山-玉山群峰的最佳先驗推估結果 (20221116)

二、即時最佳推估：卡曼濾波器

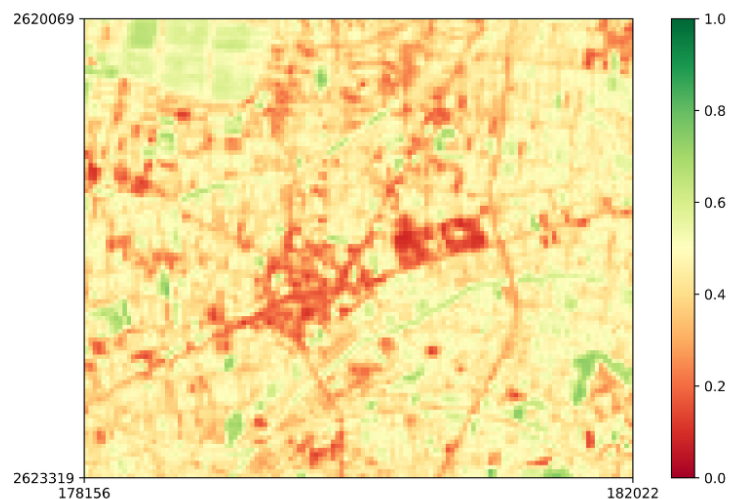


附圖 1-3: 丘陵地-鹿谷鄉的最佳先驗推估結果 (20221116)



附圖 1-4: 草原-墾丁龍磐草原的最佳先驗推估結果 (20221116)

Bayesian Model Averaging (BMA) : Agriculture field



附圖 1-5: 農地-嘉南平原的最佳先驗推估結果 (20220904)

開發完成。使用最佳的先驗推估結果，再融合即時 (該日) 的觀測值，得到即時的最佳推估。本步驟同樣參考 HISTARFM 演算法中的運算邏輯，並針對參數進行修改取得準確度較高的推估結果。結果圖在章節-當前研究成果中卡曼濾波器的推估結果圖。

六、資料參數蒐集

目前已經完成 2013-2021 年之 Landsat 和 MODIS 的衛星資料下載與前處理，並且針對土壤含水量推估方法所需的圖層，產製常態植被指數、土壤校正植被指數、地表溫度以及地表反照率等全臺灣資料。

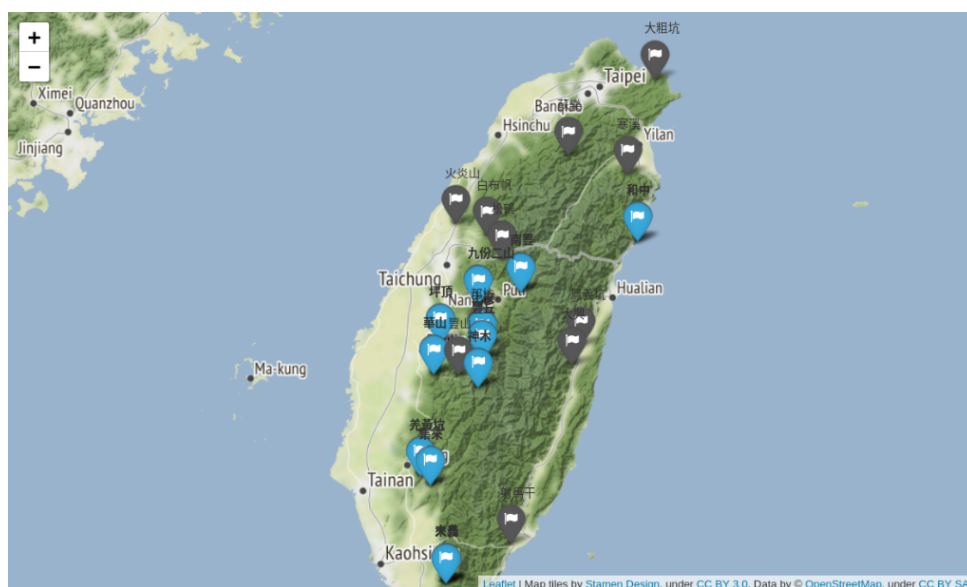
氣象資料方面，以從氣象觀測站取得 2013-2021 年的地面氣象站日及時資料。此外，也從 TCCIP 取得臺灣歷史氣候重建資料 (TReAD)。兩筆資料來接有代表土壤含水量推估模型的參數 (濕度、氣溫、風速以及輻射量)，經由單位轉換後儲存至資料庫之中。

七、衛星融合推估模型驗證

推估模型驗證的方法開發完畢，透過此方法分析卡曼濾波器的推估結果，已撰寫於當前研究成果-融合資料推估。

(二) 本研究需用到土石流監測站有關土壤含水量計觀測資料，目前僅有約 10 站有此項觀測，其空間分布及佈設方式，是否符合本研究需求？

謝謝委員的提問。透過該方法可以取得植物根部部分的土壤含水量情況，文獻指出在 0-40 公分的土壤含水量較為精確 (Ru et al., 2020)。目前以透過申請取得目標時間的所有測站的紀錄，但如委員所講，僅有 10 站左右有紀錄並且集中在中部，如圖 6 所示。因此，本計畫除了針對現有土石流觀測站進行驗證之外，另外也會加入 TCCIP 的土壤含水量資料進行比對驗證。



附圖 1-6: 土石流觀測站，目標測站 (藍色標籤) 以及非目標測站 (灰色標籤) 的分布圖

(三) 本研究成果在系統建立後可提供坡地哪些相關業務使用？建議提出其效益。

謝謝委員的提問。透過植被的實際蒸發散量推估土壤含水量的方法建立，可以進行坡地的土石流監測預警。土壤含水量的建立方法相當多元，本計畫提出的方法是針對較為淺層 0-40 公分土壤含水量進行監測，並且需要為植被覆蓋的區域。它具有預警的效益且針對尚未坍方區域有較好的監測，可以防止高風險區域在無預警情況下造成的悲劇發生。

除此之外，衛星影像融合技術也可再延伸研究，例如對於坡地

的每日高精度資料，此方面的資料可以定位山區可能發生濫墾濫伐的區域，並且以日單位產生，使主管機關可以更快速的反應。還有，透過雙衛星的時空間資料特性，本系統也有助於分析山坡地之中的資料不確定性，近一步了解在哪些坡地上是容易發生變化，尋找高敏感的山坡地區。

(四) 本案研究動機指出掌握各地區土壤含水量情形，可提高對於土石流預警的準確度？此部分研究並未納入土石流災害案例分析，如何找到相關性。

謝謝委員提問。本研究的基礎建立在過去土壤含水量與土石流發生的相關性，基於該假設才推論出透過分析土壤含水量的多寡來建立災害可能發生的危險程度。未來若此計畫的結果良好，會深入進行該資料與土石流災害的案例方析，透過衛星歷史資料完整的特性，去尋找土石流發生時該地區的土壤含水量以及相關空間參數與土石流發生的關係。

九、陳振宇委員

(一)P1-5，水保局 2016 年的土壤含水量推估方法，請列入參考文獻。

謝謝委員的提點，已將文獻附上。

(二)P3-1 P3-4，內文有關圖表之編號及說明文字請再檢視；又如圖 3-1 無法分辨這 5 張圖分別對應之類型(高山？草原？)，且無顏色之圖例等說明，同時，圖 3-1 圖 3-4 太小，請放大。

謝謝委員的提點，針對文字說明進行修改，也已經調整圖片的設定，將於期末報告中呈現。

(三) 向日葵衛星每 10 分鐘有一筆資料，儘管其空間解析度為 1 km，但若能與 QPE 雷達降雨網格資料 (1.3km × 1.3 km) 整合，是否能得到 1 km 網格之 10 分鐘土壤含水量變化？如可，亦能有助於防災預警。

謝謝委員的建議。委員的研發構想確實是有機會增進防災預警的能力，向日葵衛星每 10 分鐘一筆的特性，過去適合進行颱風路徑等氣象方面的監測，再加上 QPE 雷達所提供的降雨量網格預測，可以取得近一公里的雨量網格資料，該數值總的來說，可以解讀成本

次降雨的嚴重程度，尚無法直接當成土石流的發生機率。

土壤含水量與降雨的關係尚隔了一層紗，取得降雨資訊的確有助於增進土石流預警，但何處會是高敏感區域？何處的風險在本次降雨發生時最大？受到地表特性的影響（土壤質地、坡度、植被狀態等等），使土壤含水量和降雨存在著不確定性。這些問題則需要透過更高空間解析度的衛星遙測資料取得，在取得降雨預警前先分析風險較高的區域，這樣才能使最即時的向日葵和 QPE 雷達降雨發揮最大程度的效用。

十、技研小組(李哲宇)

(一)P2-3，三、研究執行步驟-(一) 融合衛星最佳先驗推估: 貝式模型平均之敘述是否有句尾尚未說明完成?

謝謝委員的提點，已經修改該敘述完成句子。

(二)P3-3，圖 3-2 農田和邊坡推估與實際結果比較上較其他區域白雲明顯遮蔽較多，是否會影響模式驗證正確性? 是否考慮更換樣點位置?

謝謝委員的提問。該狀況目前已有納入考量，雲遮蔽情況是本計畫模式需要處理的項目，目前的驗證方法透過隨機挖空小區域的方式進行，只要選取區域中觀測值尚有足夠的觀測值，在驗證預測的正確性上就不會有太大的影響，故目前不考慮更換樣點位置。但日後進行土壤含水量推估時，會更換樣點位置至目標土石流監測站的位置進行此演算法。

期末審查

1. 開會事由：111 年度創新研究計畫 (新興科技應用) 期末聯合審查暨成果發表會
2. 開會時間：111 年 12 月 7 日 (星期三) 上午 9 時
3. 開會地點：本局府西辦公室 (南投市府西路 71 號)
4. 主持人：林副局長長立、王副局長晉倫、陳主任秘書重光、陳總工程司振宇
5. 紀錄：高副工程司百毅
6. 參加單位及人員：(詳會議簽名冊)
7. 主持人致詞：(略)
8. 業務單位報告：(略)
9. 國立臺灣大學簡報：(略)
10. 審查意見及討論：

(一) 值得繼續探討、精進，俾導入相關防災業務使用。

謝謝委員肯定。

(二)1-4、2-3、2-12、2-16、2-17 頁的參考文獻為問號，請補充說明。

謝謝委員說明。本文章透過 latex 語法撰寫，文獻參照之功能於期末成果初稿使用有誤，現已進行修正。

(三)2-11 頁 NDVI 為何選擇跟隨前人選用 0.1 請補充說明。

NDVI 跟隨前人選用 0.1 主要為經驗參考，此門檻可過濾掉低植生指數，也就是非植被地區，該些區域不適合進行後續的蒸發散量運算，後續可嘗試不同的門檻來比較運算結果的差異。

(四)2-14 頁文中 (圖??)、網址未括號，請修正。

謝謝委員，已將「經由 MODIS 網站上可連結至其衛星遙測演算產品資料庫 (<http://modis.gsfc.nasa.gov>)，資料庫內存有全世界所有地區 (圖??) 觀察資料之 HDF 檔。」修正為「經由 MODIS 網站上可連結至其衛星遙測演算產品資料庫 (<http://modis.gsfc.nasa.gov>)，資料庫內存有全世界所有地區 (圖 2-3) 觀察資料之 HDF 檔。」網址未括號部分也由「本研究使用開源的 Landsat 衛星下載的 Python 套件來下載需要的衛星資料 <https://pypi.org/project/landsatxplore/>。」修正為「本研究使用開源的 Landsat 衛星下載的 Python 套件來下載需要的衛星資料 (<https://pypi.org/project/landsatxplore/>)。」。

(五)2-15 2-22 頁排版與文字紊亂不易讀懂意思，請重新整理排列。

謝謝委員說明。經檢視原文章中分節較不明確。已重新整理排列，增加子標號以方便閱讀。

(六)3-7 頁為何選定 2019 年的全年資料為驗證推估？

謝謝委員提問。選定 2019 年係同時配合各方面蒐集資料的最終結果。遙測資料本身量較充足，選定 2019 年除配合其它資料限制(後述)外，其本身在我們所選定的測試場域也有較少的雲遮蔽；PySEBAL 計算所使用的參數由臺灣氣候變遷推估資訊與調適知識平台所取得，資料時間範圍最新至 2020 年；以及由土石流監測站申請之事件土壤含水量時間範圍中，也以 2019 年資料較為充足，故以 2019 年做為推估及驗證之年份。

(七)3-8 頁開始的土壤含水量變化圖，請試說明個別圖說成果。

謝謝委員提問。經調整後之詳細方法說明已補充於 2-23 頁至第三章結束；圖說成果已整合說明於 3-7 頁至第三章結束。以下節錄部分說明予委員參考：以 CWSI 與高程兩特徵以 k-fold cross validation 評估模式表現對於土壤含水量預測來說已經足夠好，R

square 的結果為 0.76，若選擇保留三個特徵，結果是 CWSI、高程與 24 小時累積降雨會被選入，恰巧是衛星資料、地文與水文資料都各被納入一種，R square 可再多增加 0.045，達到 0.8 以上。而本研究以 CWSI 與高程兩特徵作為輸入變數建立 XGBoost 模型，以剩下三成之測試資料進行驗證，結果如圖 3-10，R square 值為 0.81。