
112年農村水保嘉年華

高精度衛星遙測土壤含水量系統建立

余化龍 教授

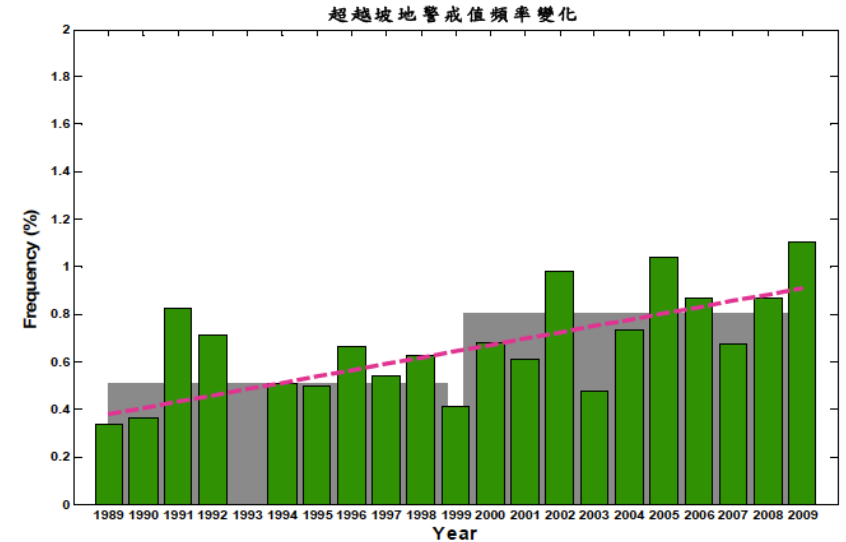
臺大生物環境系統工程學系



前言

研究背景

- IPCC於 2021 年發表的綜合評估報告，揭露地球正面臨嚴峻考驗
 - 區域極端氣象頻率增加
 - 乾濕季降雨量差異增大
- 坡地災害與極端降雨關連性明顯
- 因應未來可能極端氣象發生頻率，研擬創新防災科技，降低坡地災害風險，刻不容緩。



(TCCIP Report, 2011)



前言

擬解決問題

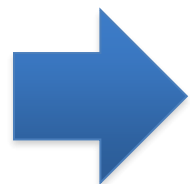
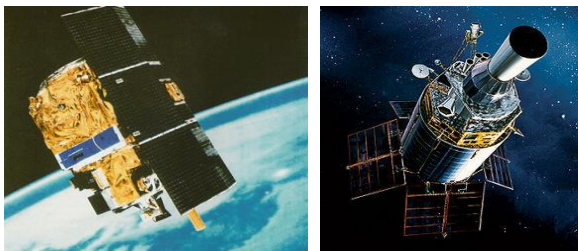
- 為降雨強度與有效累積降雨量雖為土石流主要啟動因子，而土壤臨前含水量亦為決定土石流啟動與規模大小之關鍵
 - 影響土壤剪力強度與入滲容量 (Zhao et al, 2010)
 - 當乾旱過久，含水量過低造成土壤龜裂，突然強降雨，也可能增加土石流規模 (Davis and Milne, 2007; Hu et al., 2015)
- 遙測影像時空解析度對土壤含水量推估有顯著影響
 - Chenyang Xu et al. (2018)依地表溫度與常態化差異植被指數相關性，將 MODIS 與 Landsat 衛星資料融合後產製出更細緻的土壤含水量
 - 常受限於雲朵遮蔽的效應
- 衛星遙測應用於防災之時空間精度不足、土壤含水量現地觀測資料有限

前言

創新研究想法

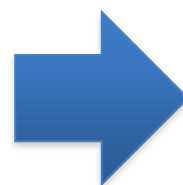
- 近年許多國外研究開始在探討可將不同空間與時間解析度的衛星影像利用統計方法進行融合，產出同時具有高空間解析度以及高觀測頻率的衛星影像(F. Gao, 2006, J. Wang, 2018, Moreno-Martinez, 2020)

不同解析度遙測影像



衛星融合演算技術
時空分析與推估方法

大範圍高精度
土壤含水量推估



考慮高精度土壤含水量之
動態坡地災害預警架構

土石流歷史災害資料庫
氣象局氣象資料庫

前言

計畫目標

- 開發創新之衛星融合演算技術
 - 以影像融合技術進行土壤含水量相關參數之推估，透過大量時間與空間遙測資料分析，建構可考慮不確定性之高時空精度推估架構。
- 融合衛星資料下之土壤含水量時空間分佈推估
 - 配合土壤含水量推估方法，於特定區域透過融合影像進行土壤含水量之高精度推估



工作流程與架構

計畫工作架構

推估模組開發

精準遙測監控土壤含水量系統研發計畫

土壤含水量現地驗證

衛星遙測影像
資料融合模組研發

土壤含水量估算法開發

蒐集及分析土石流
觀測站之觀測值

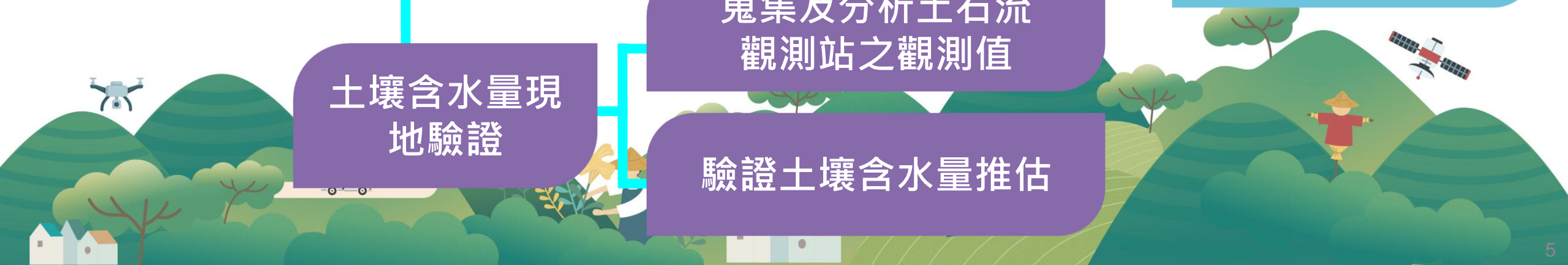
驗證土壤含水量推估

現有衛星影像
蒐集及分析

開發衛星影像
融合演算法

氣象資料
蒐集及前處理

開發土壤含水量
遙測推估方法



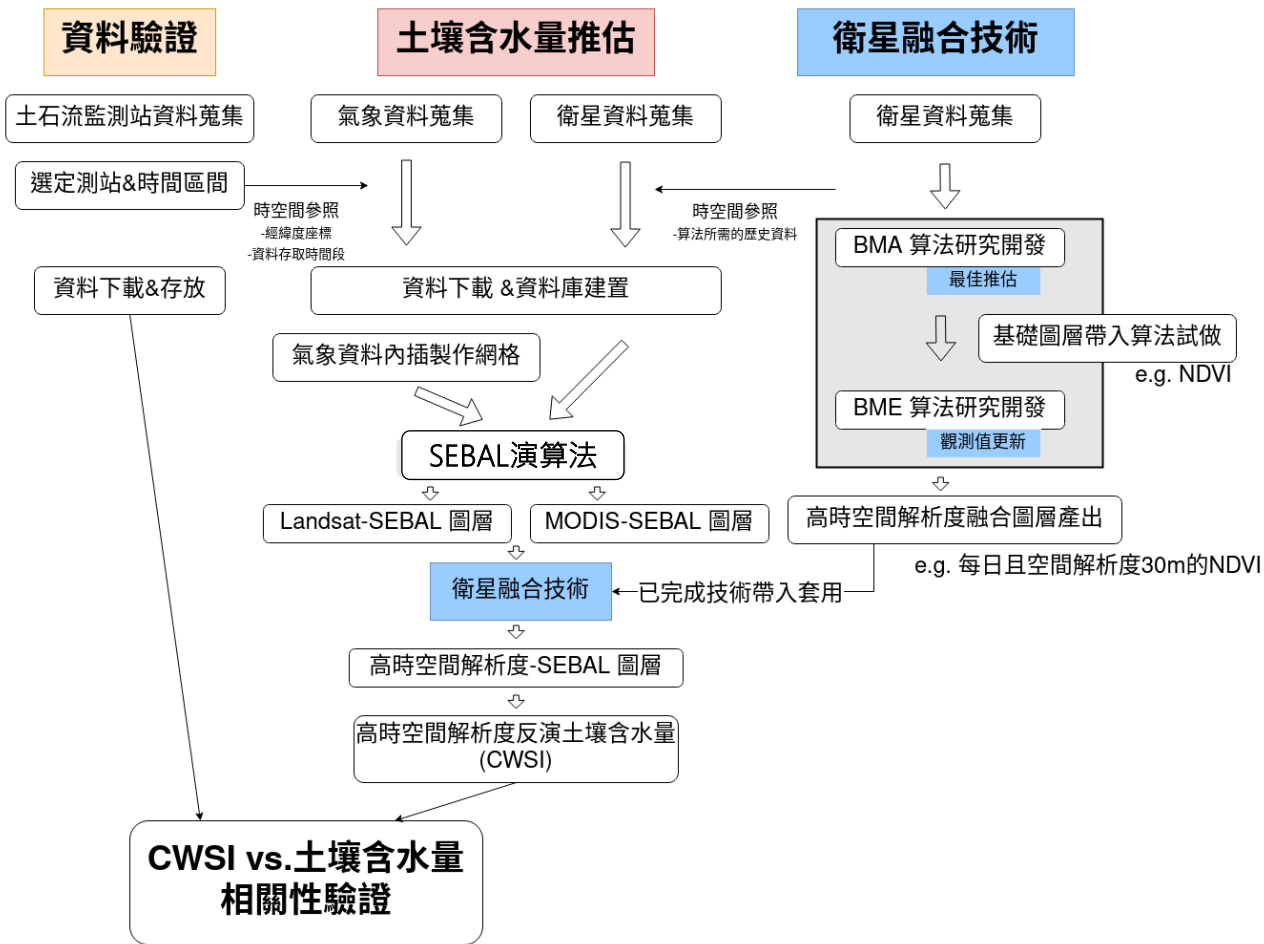
工作流程與架構

整體工作流程

- 資料蒐集與前處理
 - 蒐集土石流觀測站資料、象氣站資料與衛星遙測資料並進行前處理
- 衛星融合技術開發
 - 透過MODIS與Landsat衛星影像並進行藍、紅、近紅、短波紅外線(SWIR)等波段之截取與融合，產製特定區域融合成果。
- 作物缺水指數(CWSI)計算
 - 利用衛星融合技術推估作物缺水指數(CWSI)。
- 高精度土壤含水量推估

透過機器學習演算法(XGBoost)發展土壤含水量推估模型

選定「上安、神木、豐丘、九份二山、坪頂、南豐以及來義」等7站進行CWSI與土壤含水量推估驗證。



資料收集與前處理

資料蒐集類別

• 氣象資料

• 臺灣氣候變遷推估與資訊平臺計畫

- 歷史氣候重建資料(TReAD, Level2/Level4) - 相對濕度、短波輻射、平均溫、2 米溫度、2 米風速、日輻射量
- 中央氣象局自動測站

• 地文資料

- 數位地質高程模型, DEM (高程、坡度、坡向)

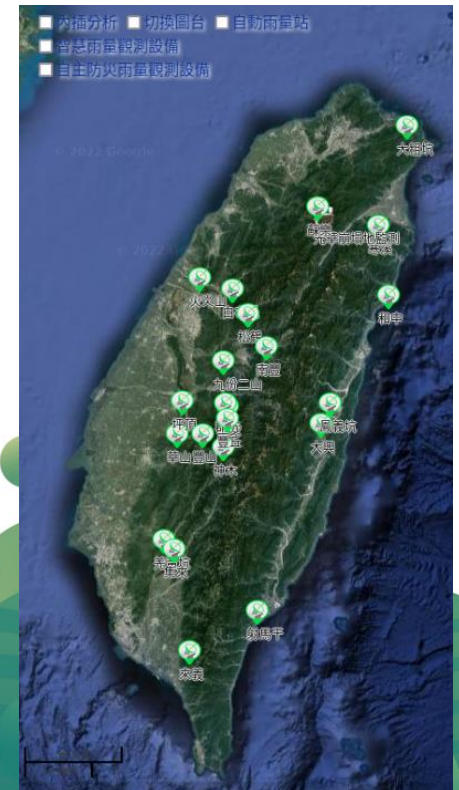
• 衛星資料

• MODIS 以及 Landsat 衛星歷年各波段監測資料下載與前處理

- MODIS-MOD11A1 - 每日地表溫度和發射率，解析度為一公里
- MODIS-MOD09GA - 地表反射，七個波段(0.405 ~ 14.385 um)，解析度為500~1000公尺
- Landsat - 地表反射(surface reflectance)以及地表溫度(surface temperature)

• 土壤含水量資料

- 土石流防災資訊網土石流測站土壤含水量觀測 – 觀測頻率5分鐘

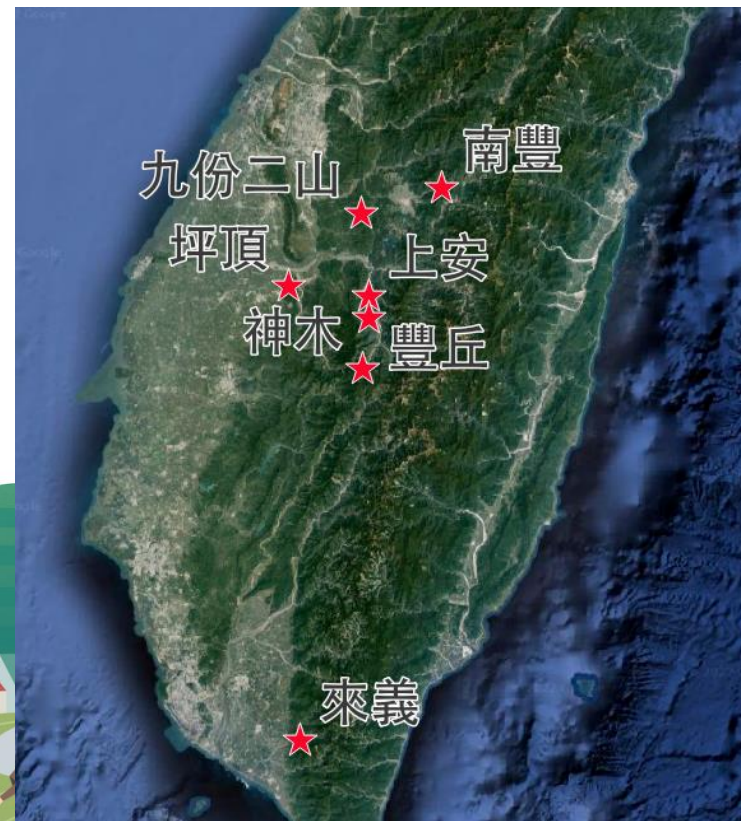
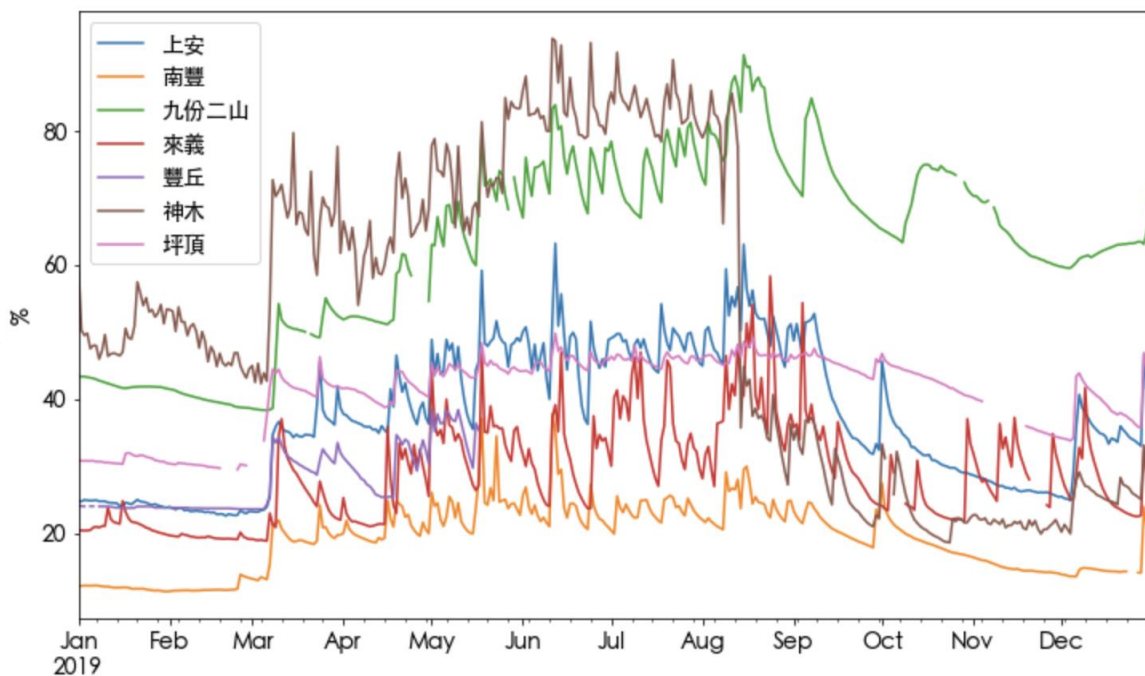


資料收集與前處理

土壤含水量觀測資料

- 土石流觀測站 – 篩選事件紀錄資料，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
 - 篩選衛星資料品質，以 2019 年的資料進行推估驗證
 - 根據遙測資料觀測時間，選定每日上午10~11點間資料進行比對

各測站10-11間
平均土壤含水量

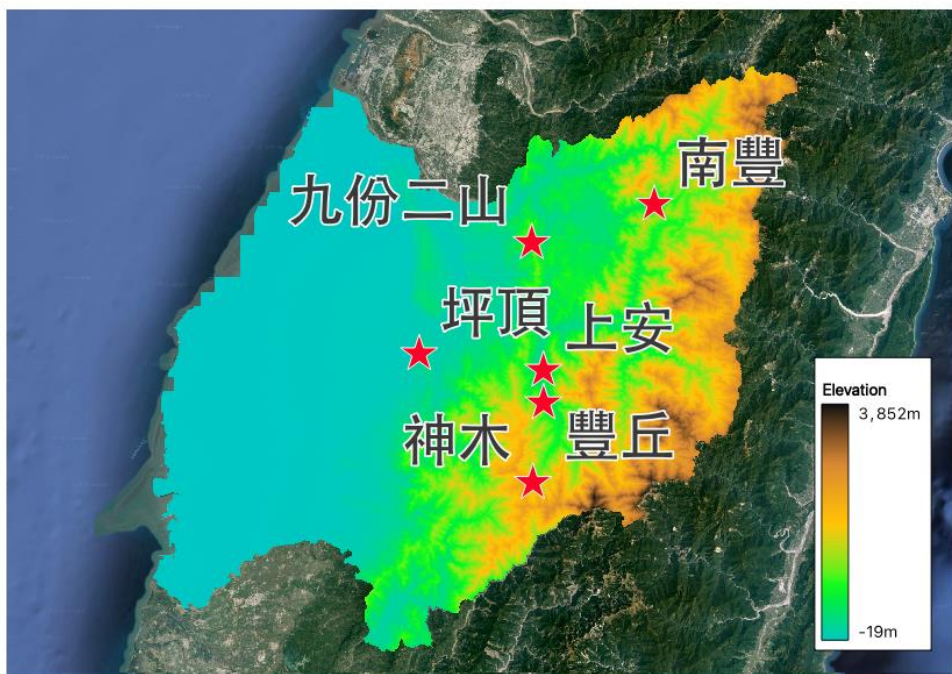


資料收集與前處理

地文資料

- 包括坡度、坡向、高程，利用20m精度之DEM進行計算，並加入空間座標資料。

高程



各測站空間與地形相關參數資訊

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	0.00	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	41.40	340.53	594

資料收集與前處理

氣象資料

- 評估參數包括：24小時臨前累積雨量與12小時臨前平均濕度
- 土石流觀測站鄰近且雨量與濕度缺漏值較少之自動測站

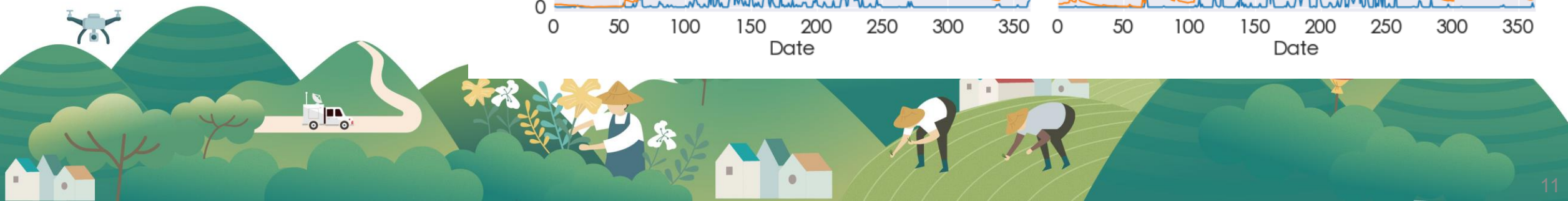
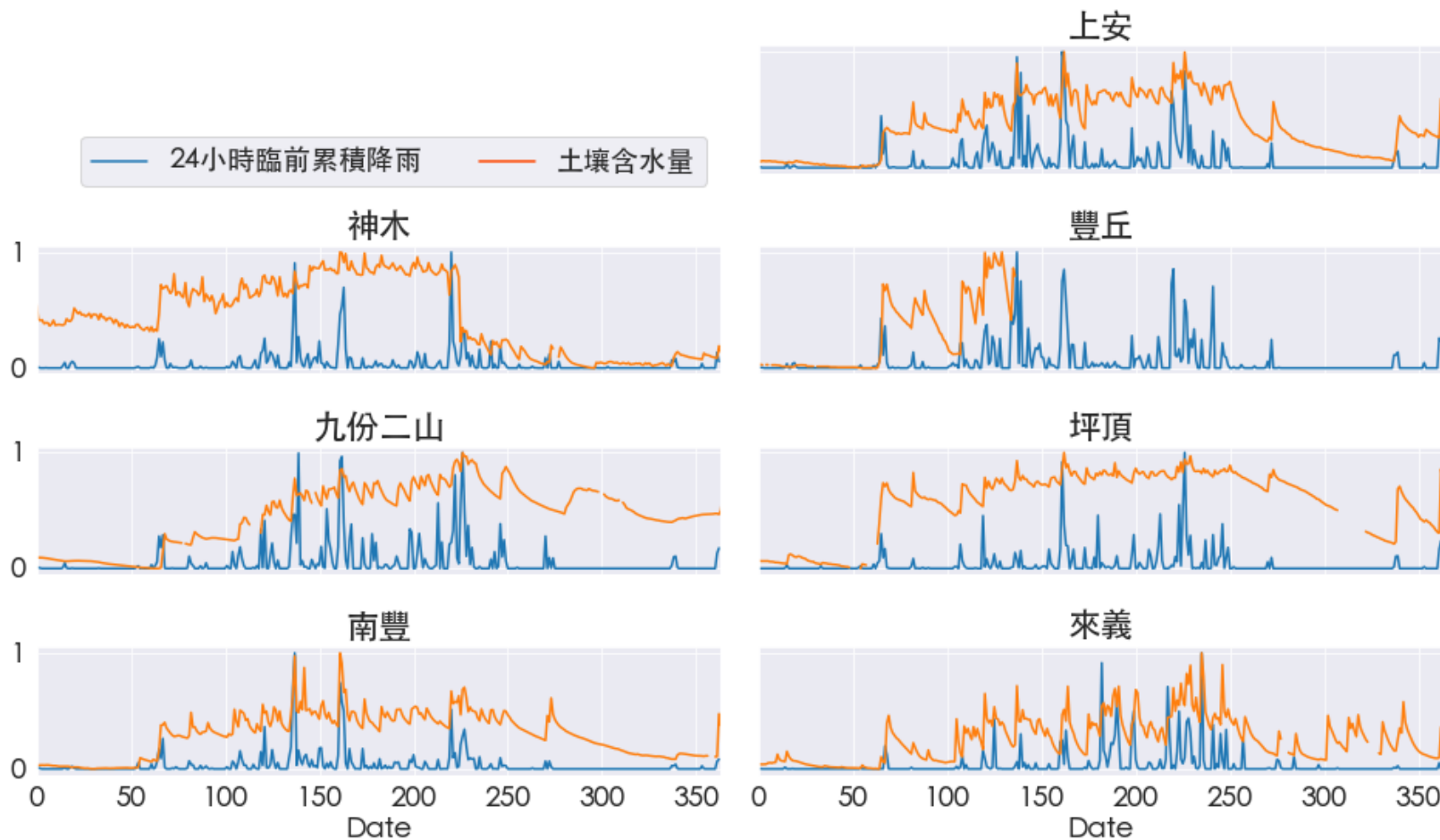
土石流測站	使用之自動測站	距離	缺漏值比例(%)
上安	信義	2677	0.08
神木	神木村	2044	0.03
豐丘	豐丘	2990	0.48
九份二山	蘆竹浦	2790	0.11
坪頂	林內	1769	0.15
南豐	楓樹林	1907	0.46
來義	潮州	12760	2.25

土石流測站	使用之自動測站	距離	缺漏值比例(%)
上安	信義	2677	0.03
神木	神木村	2044	0.06
豐丘	信義	4071	0.03
九份二山	中寮	9378	2.65
坪頂	林內	1769	1.26
南豐	仁愛	6268	0.13
來義	新埤	9487	3.39

資料收集與前處理

氣象資料

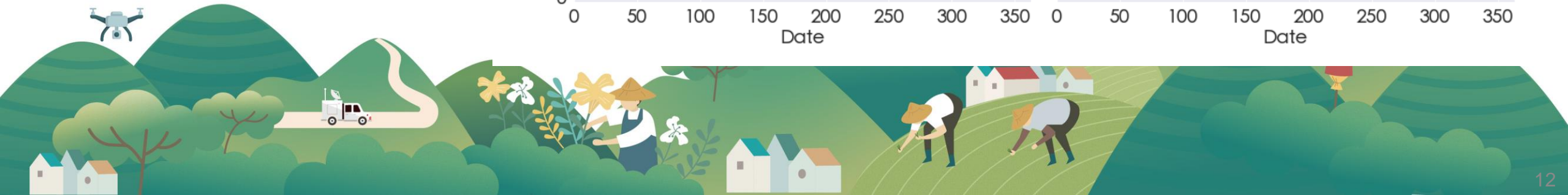
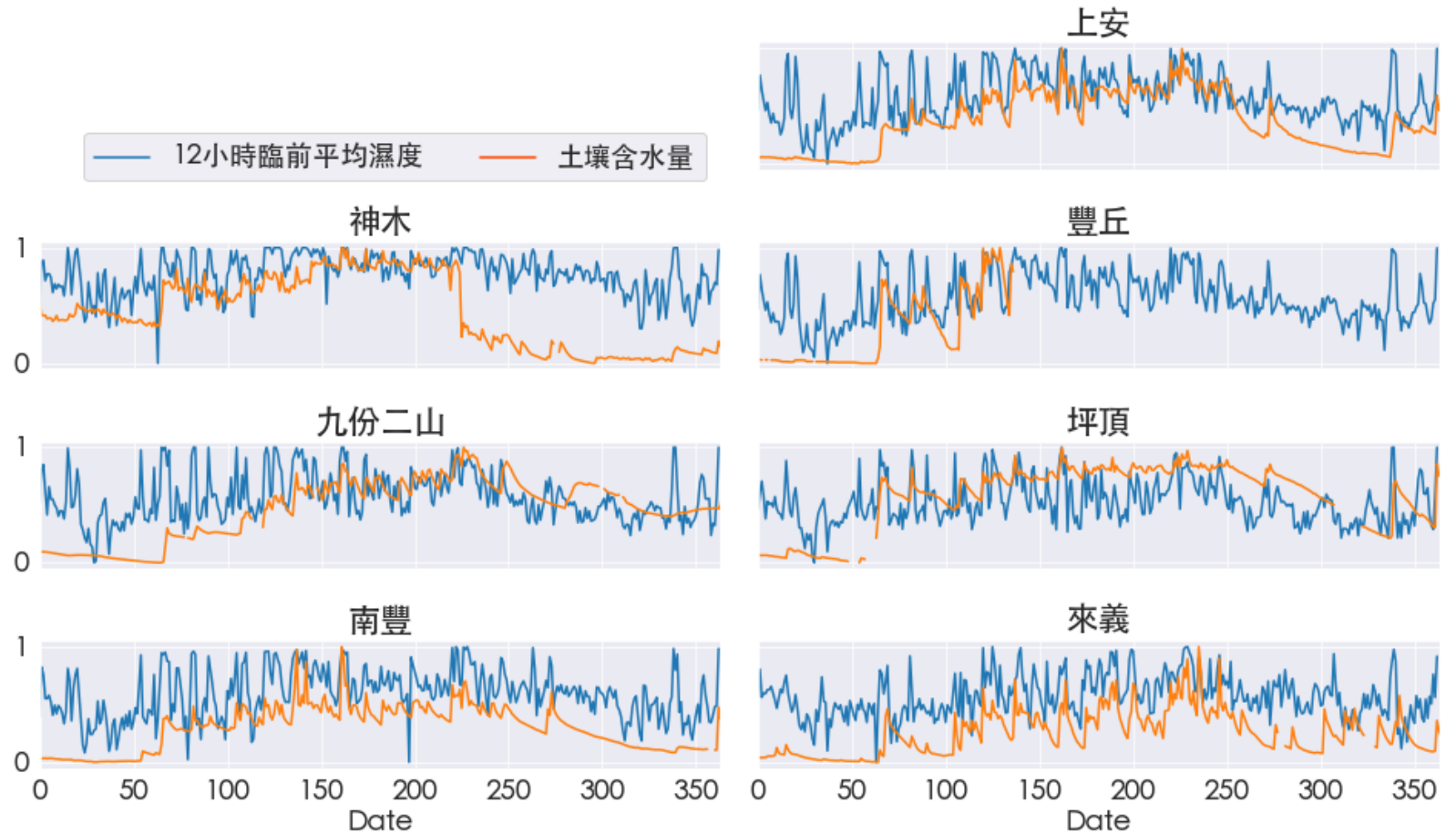
- 計算：**24小時臨前累積雨量**，前一天11 a.m.至今日10 a.m.
- 評估參數包括：**24小時臨前累積雨量與12小時臨前平均濕度**
- 與土壤含水量繪圖比較如右：
- 某些土壤含水量與**24小時臨前累積降雨**有明顯相關
- 如南豐、來義



資料收集與前處理

氣象資料

- 計算：12小時臨前平均濕度，前一天11 p.m.至今日10點 a.m.
- 與土壤含水量繪圖比較如右：
- 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



研究方法

衛星融合技術開發

- 最佳先驗推估 - 貝氏模型平均 (Bayesian model averaging, BMA)

BMA: 利用Landsat(Model A)和MODIS(Model B)在時間k產生一個Priori最佳推估值

Landsat

以2013-2020年的Landsat歷史資料至時間點k之前的資料取得其中位數：最佳推估值
變方：不確定性

MODIS

將2013-2020年之MODIS歷史資料至時間點k之前的資料，和Landsat的資料進行無偏估計，取得其推估值和與Landsat衛星點位之間的共變方

$$\bar{x}_k = z_k \frac{P_{k,MOD}}{P_{k,LS} + P_{k,MOD}} + u_{k,MOD}^{30} \frac{\bar{P}_{k,LS}}{P_{k,LS} + P_{k,MOD}},$$
$$P_k^- = (1 - \gamma) \left(\frac{1}{P_{k,LS}} + \frac{1}{P_{k,MOD}} \right)^{-1},$$

Moreno-Martínez, Á., Izquierdo-Verdiguier, E., Maneta, M. P., Camps-Valls, G., Robinson, N., Muñoz-Mari, J., ... & Running, S. W. (2020). Multispectral high resolution sensor fusion for smoothing and gap-filling in the cloud. *Remote Sensing of Environment*, 247, 111901.

研究方法

衛星融合技術開發

- 即時最佳推估 - 卡曼濾波器 (Kalman filter)
 - 時刻 k 的先驗分布是由外部來源 BMA 的最佳推估取得
 - 加入觀測資料修正為即時最佳推估結果
- 考慮偏差推估，加入 b_k 、 T_k 進行修正

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

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

$$x_k = x_k^- + K_k(z_k - H x_k^-)$$

考慮偏差

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

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

$$T_k = (I - K_k H) T_k^-$$

$$W_k = P_k + T_k$$

$$P_k = (I - K_k H) P_k^-$$

k : 表時刻
 x_k^- : 先驗推估值來自 BMA 的最佳推估值
 z : 觀測資料。
 K : 卡曼增益，代表 x 和 z 相對權重
 P_k^- : 先驗推估值的誤差協方差
 H : 觀測矩陣，用來描述模型和實際觀測數據的相關
 b_k^- : 偏差推估值來自 BMA 的最佳推估值之部分比例
 T_k : 偏差推估的誤差協方差
 W_k : 總推估的誤差協方差

研究方法

遙測資料之應用－CWSI

- 土壤含水量推估-地面能量平衡方程式 & 作物水分逆境指數 (Crop Water Stress Index, **CWSI**)
 - 經地表能量平衡方程式與Penman-Moneteith蒸發散方程式可合併推導得出一**作物葉面溫度所代表之土壤含水量指標**，該指標是利用作物實際蒸發散量與潛勢蒸發散量的比值來做計算。
 - 利用SEBAL方法透過對Landsat與MODIS衛星資料各自計算其蒸發散量及CWSI面資料，再進行資料融合產製在土壤含水量站點之**每日CWSI融合結果**

作物水分逆境指數

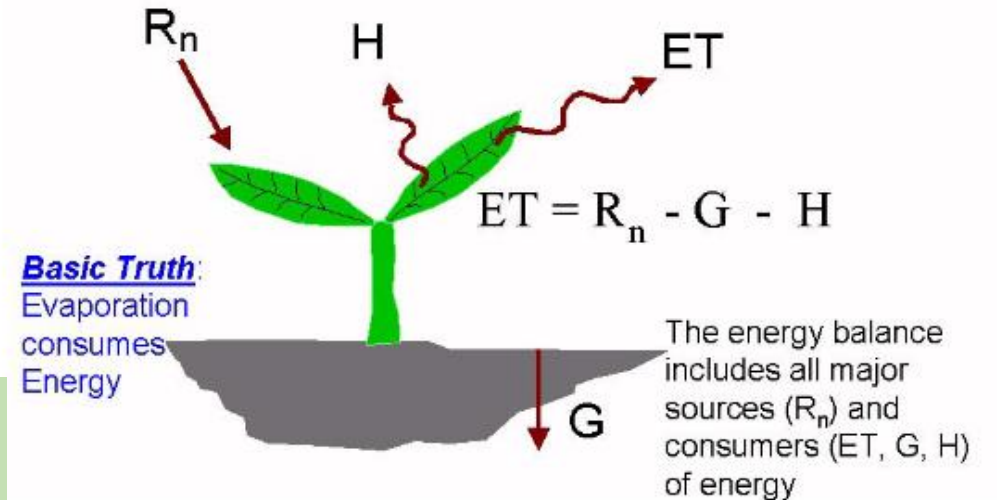
$$CWSI = 1 - \frac{LE}{R_n - G}$$

地面能量平衡方程式

$$R_n = G + H + LE$$

R_n : 淨輻射能量 (Net Radiation, W/m^2) H : 可感熱通量 (Sensible Heat flux, W/m^2)

G : 土壤熱通量 (Soil Heat flux, W/m^2) LE : 潛熱通量 (Latent Heat flux, W/m^2)

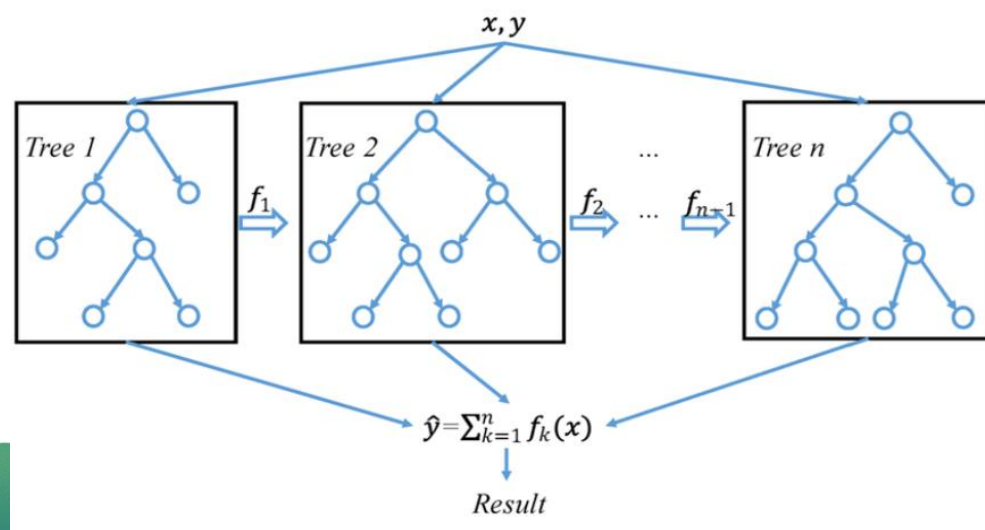


研究方法

機器學習模型建立 – XGboost

- 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

XGBoost eXtreme Gradient Boosting



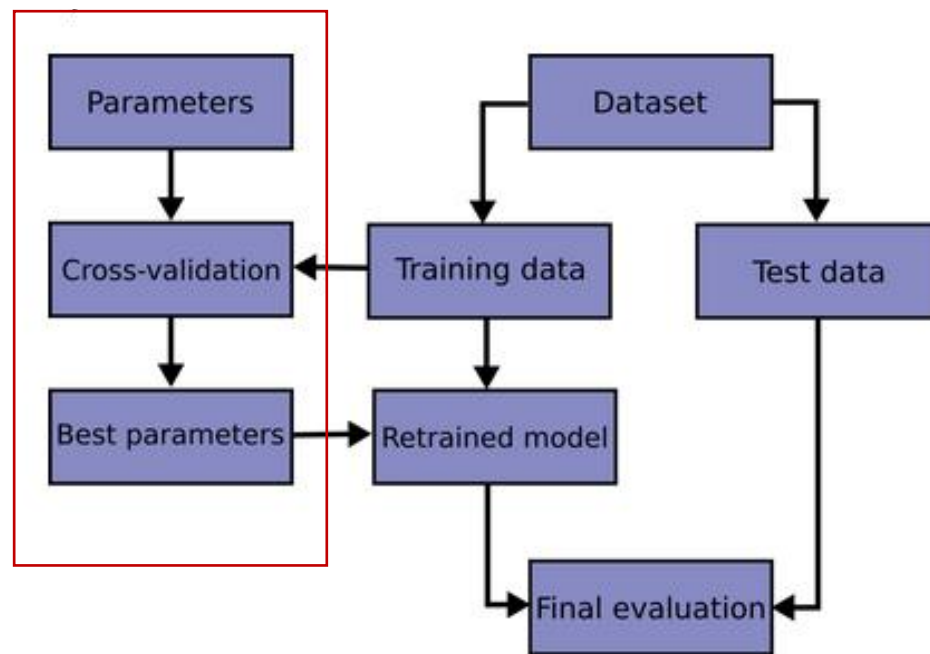
XGBOOST的示意圖

研究方法

機器學習模型建立 – XGboost

- 運用與建立模式同一筆資料評估模式表現唯一方法論上的錯誤，模式只會重複給出他已見過樣本的標籤，但卻無法提供有效的預測給未見過的資料，此現象稱為**過度擬和(overfitting)**
- 為避免以上現象，基本機器學習流程如右：

Feature selection



研究方法

機器學習模型建立 – Feature selection

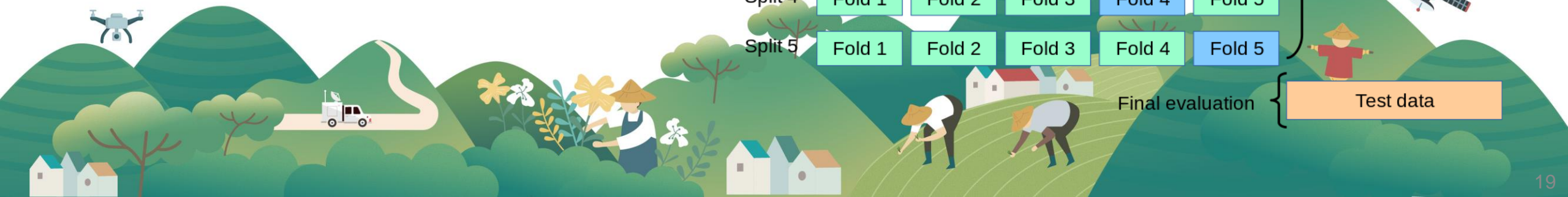
- **Feature selection** 為一常用於機器學習領域之技術，其可從眾多變數中選擇相較重要的變數用以建構模型。
 - The objectives of feature selection include building **simpler** and more **comprehensible models**, **improving data-mining performance**, and **preparing clean, understandable data**. (Jundong Li, et. al. 2017)
- 本計畫用 **Backward Sequential Feature Selection (Backward -SFS)**，依據模式表現，評估各個變數之重要程度
 - **Backward-SFS**：其為一運用迭代方法，依序移除對模式表現相較沒有幫助的變數，直到達到欲保留的變數數量。

研究方法

機器學習模型建立 – Feature selection

• 模式表現的評估依據：**k-fold cross validation & R square**，評估流程如下：

1. 將資料分成訓練集與測試集
2. 從訓練集中分成k等份
3. 利用k-1等分的資料建立模型
4. 並用模型驗證剩下的那一等份，計算R square值
5. 從步驟3開始重複k次，獲得k個R square值並平均

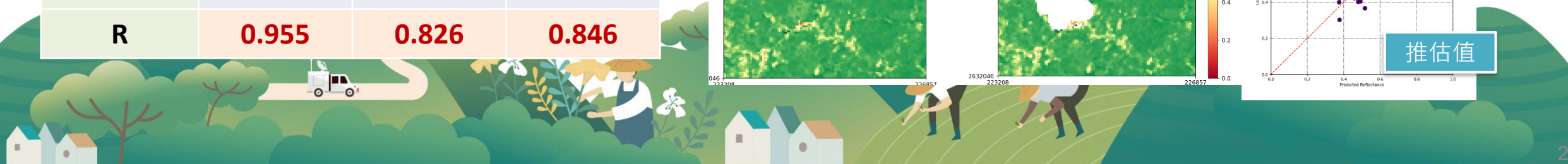
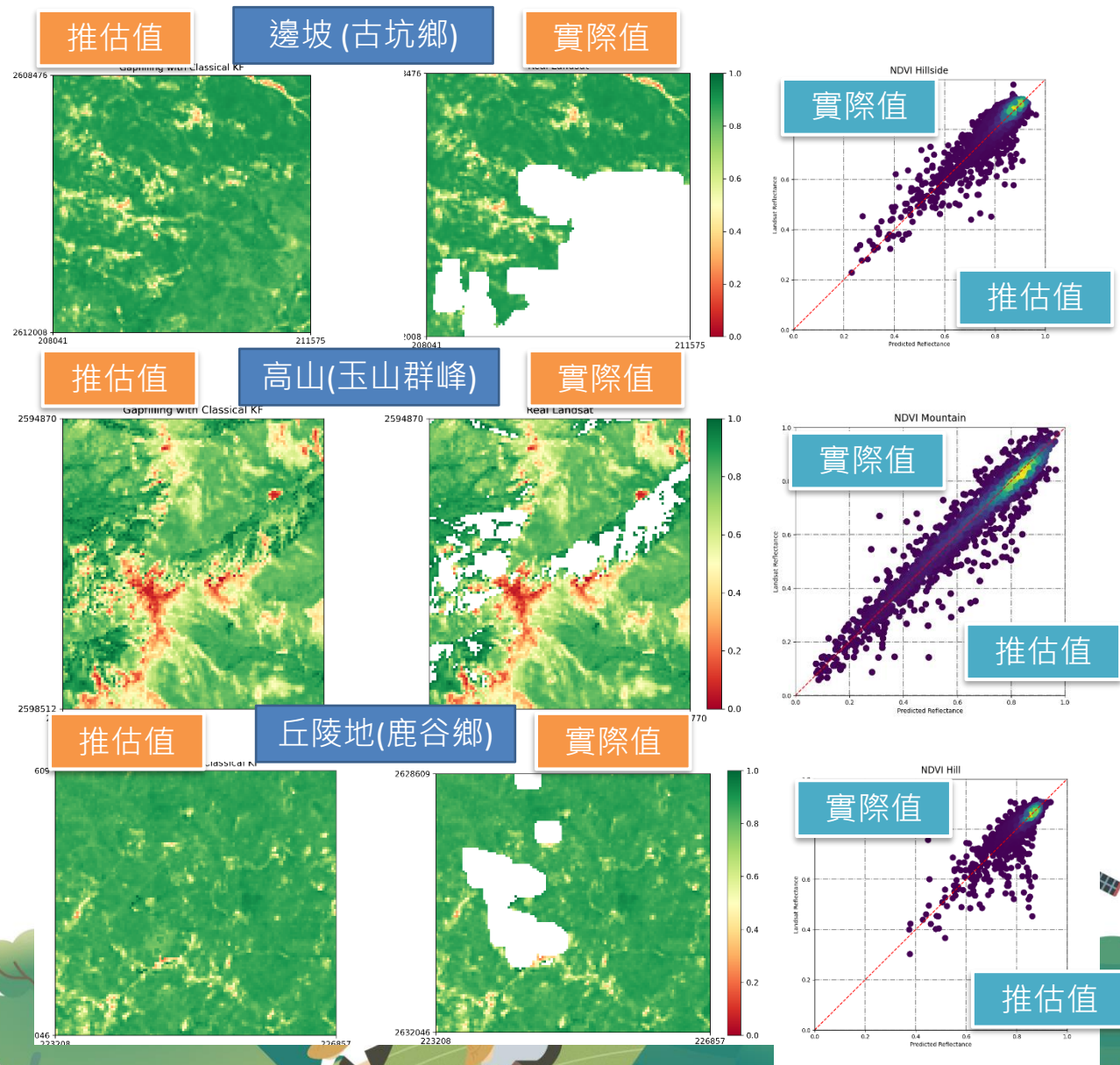


研究成果

衛星融合技術開發

- **NDVI驗證融合技術可靠性**
 - 挑選重要地貌類型進行融合模擬驗證
 1. 2021/11/16, 高山(玉山群峰)、丘陵地(鹿谷鄉)
 2. 2021/09/04, 邊坡(古坑鄉)
- *空白處為雲遮蔽

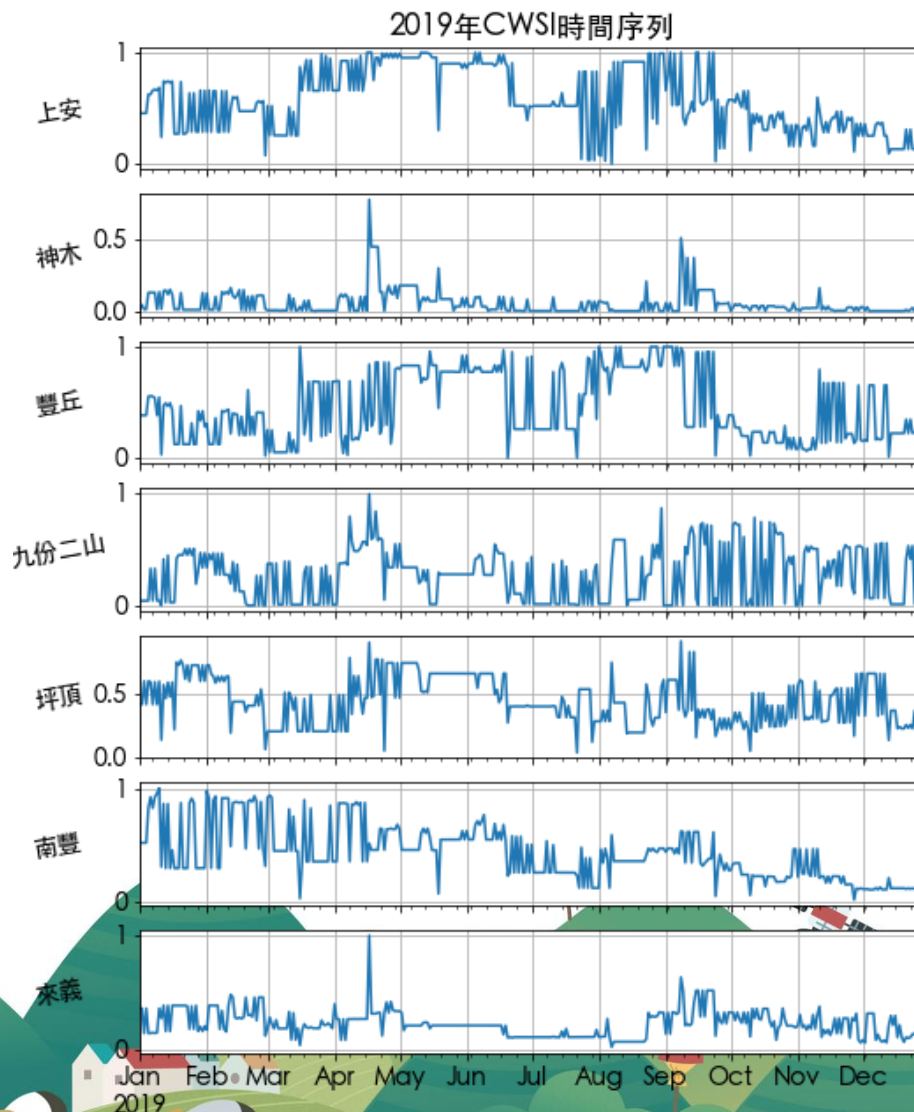
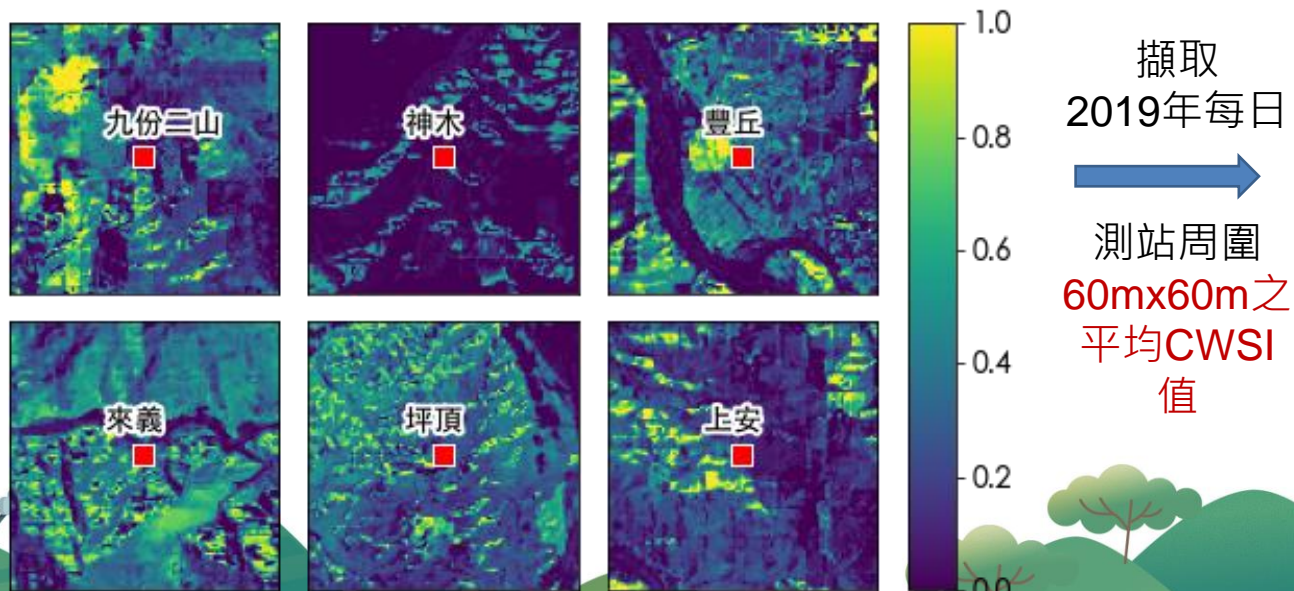
表現衡量指標	高山(玉山群峰)	邊坡(古坑鄉)	丘陵地(鹿谷鄉)
RMSE	0.0659	0.0881	0.0528
MAE	0.0501	0.081	0.0426
R	0.955	0.826	0.846



研究成果

CWSI 計算

- 2019年每日以測站為中心，3km x 3km CWSI 資料融合成果。



研究成果

高精度土壤含水量推估 – 機器學習模型建立

- 機器學習模型：**XGboost**
- 預測變數：**土壤含水量**
- 與土壤含水量相關之因子
 - 根據文獻回顧, 主要影響土壤含水量的因子包括地形、氣象、植被等
- 輸入變數包括：
 - 衛星資料：**CWSI**
 - 氣象資料：**24小時臨前累積雨量與12小時臨前平均濕度**
 - 地文資料：**XY座標、高程、坡度、坡向**
- 建立機器學習之資料之分配
 - 訓練資料：**70%**
 - 驗證資料：**30%**



研究成果

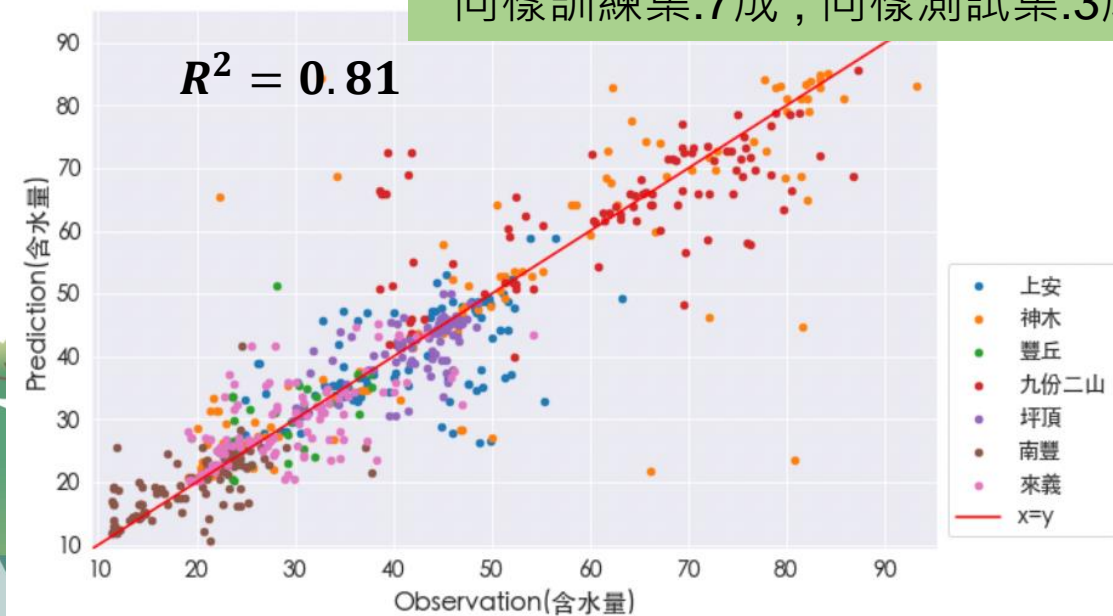
高精度土壤含水量推估 – Feature selection

- 模式使用自變數：**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
- 依據**Backward – SFS**, 同樣**7**成作為訓練集，以**4-fold cross validation** 挑選重要的參數，分別指定挑選前**2**、**3**、**4**重要的參數，結果如下：

只使用**CWSI**與**高程**建立模式
同樣訓練集:**7**成；同樣測試集:**3**成

Backward-SFS 結果

Selected features number	Selected features	4-fold cross validation average
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



結論與建議

結論與建議

- 本研究完成透過**Kalman filter**的架構，發展**MODIS**與**Landsat**衛星資料融合架構，發展高時空解析度衛星資料推估模式。
- 本研究應用高時空解析度衛星資料推估模式，結合**SEBAL**模式推估作物水分逆境指數時空分布。
- 本研究完成以**XGBOOST**模式，結合氣象、地文以及遙測資料建構土壤含水量推估架構。
- 本計畫以水保局土石流觀測站之土壤水分觀測作為研究目標，後續可持續收集其他位置之土壤水分觀測，以降低不同土地利用下對於土壤水分推估之可能偏差。





報告完畢
敬請指教

