# Observing the spatiotemporal dynamics of tea plantations in a tropical mountainous region using machine learning

Cho-Ying Huang1,2,\*, Zih-Yu Shen1, Tzu-Hsin Karen Chen3

¹ Department of Geography, National Taiwan University, Taipei 10617, Taiwan

² Research Center for Future Earth, National Taiwan University, Taipei 10617, Taiwan

³ Department of Urban Design and Planning, University of Washington, Seattle, WA 98195, USA

\* choyhuang@gmail.com

**KEY WORDS:** tea farm, time series, satellite images, continuous change detection and classification, deep learning

**ABSTRACT:**

Tea is a popular drink worldwide and a major cash crop in mountainous agricultural regions in Taiwan. However, due to the rugged terrain, these areas are difficult to manage, and frequent fog, cloud cover, and spectral complexity hinder remote sensing applications. With only 92 observations in 221 satellite images (2019-2021) on average for each pixel in the Greater Ali-Mountain tea plantation region of Southern Taiwan (the study site), a systematic method may be necessary to effectively map tea farms. This study aims to evaluate the feasibility of classifying tea plantations in the study region using satellite imagery with machine learning. We utilized Sentinel-2 surface reflectance images to generate annual, seasonal, and continuous change detection and classification land use and land cover (LULC) maps. Two machine learning methods, Random Forests (RF), and U-net with Resnet-18 backbone, were employed to classify five types of LULC. After validating the results, we found that U-net had higher accuracy than RF with significantly higher efficiency in identifying tea farms compared to RF. Using U-net with the seasonal approach resulted in the highest overall accuracy of 0.949, with the tea farm producer's accuracy and user's accuracy being 0.916 and 0.939, respectively. Our findings suggest that U-net is suitable for identifying tea farms due to its ability to augment training data, to use an encoder-decoder structure, and to incorporate skip-connections, which capture image features more effectively and prevent the loss of critical information. This approach offers significant advantages in image interpretation. Moreover, the method shows promising potential for mapping other mountain evergreen crops, such as fruit and coffee trees.

# INTRODUCTION

Tea is the second most popular drink next to water (Hicks, 2009) and a major cash crop in the mountainous region of Taiwan. However, these areas confront challenges due to rugged terrain, hindering accessibility and obstructing remote sensing techniques often impeded by fog and clouds. Habitat fragmentation further complicates managing this delicate human-nature system. To effectively address these complexities and manage tea farms, monitoring the long-term spatiotemporal dynamics of these agricultural areas is essential.

In this study, we propose utilizing satellite imagery paired with machine learning models for precise classification of diverse land types, encompassing tea farms, forests, soil, manmade and crops. By harnessing the high spatial and temporal resolution inherent in satellite imagery, we can actively monitor and analyze the ever-evolving landscapes of tea farming regions. This amalgamation of technology and data-driven approaches holds remarkable potential for advancing our comprehension and stewardship of tea cultivation areas.

# METHODS

* 1. **Research area**

The study focuses on the Greater Ali-Mountain tea region in the central Taiwan (Fig. 1). Publicly available Sentinel-2 imagery with six spectral bands and ALOS global digital surface model (DSM) 30 m data were obtained.

The research focuses on four districts of Chiayi, Taiwan, commonly referred to as the "Greater Alishan Tea Area." This region is renowned as the primary tea production area in Taiwan, with tea farms predominantly situated between 800 to 1500 meters in elevation. The area experiences foggy and cold conditions throughout the year.

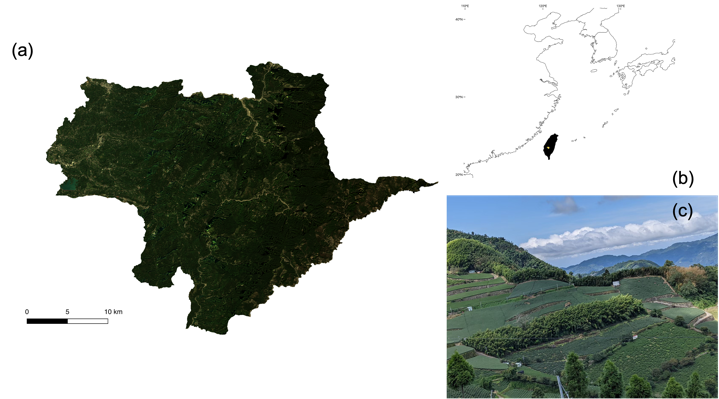


Fig. 1 (a) The study site, the Greater Ali-Mountain tea region in (b) the central Taiwan. The background true colours Sentinel-2 imagery was acquired on 2020/01/16. (c) A typical view of the tea country consisting of tea plantations, secondary evergreen forests and built environments.

* 1. **Continuous change detection and classification (CCDC) algorithm**

The CCDC method is employed to create a seasonal model for each pixel within the image. Depending on the number of valid values, it can produce three types of models: a simple model (12), an advanced model (15), and a full model (18). Each model is equipped to calculate the root-mean-square error and identify significant deviations from the model values. When such deviations occur, the method classifies it as a change in land cover and subsequently generates a new model for the subsequent timeline (Zhu & Woodcock, 2014).

What distinguishes this method is its ability to analyze the entire time series rather than a single image. As a result, it can concurrently detect land cover changes and the seasonal variations in each pixel. By selecting a specific time point, we can obtain the seasonal parameters for that day, facilitating the evaluation of machine learning model training. Furthermore, the CCDC method is adept at handling NA (not available) values within the formula simulation by leveraging existing data. This adaptability proves advantageous, particularly in high-altitude mountainous regions where satellite data often go missing due to fog and cloud cover. In such cases, the CCDC method not only resolves data gaps but also maximizes the utility of available data (Zhu Zhe, 2020).

(1)

where x = day of year

i = the ith band

T = number of days in per year (T = 365)

c1,i = coefficient for overall value for the ith band

a1,i, b1,i, a2,i, b2,i, a3,i, b3i = coefficients for intra-annual change for the ith band

= predicted value for the ith band at Julian date x

* 1. **Satellite images processing**

We combined three years of satellite imagery and utilized annual, seasonal, and continuous change detection and classification (CCDC) methods for classification training. The study utilizes Sentinel-2 imagery, specifically six spectral bands: blue, green, red, near-infrared, and two shortwave infrared bands. These images have a spatial resolution of either 10 or 20 meters and a temporal resolution of 5 days. Additionally, ALOS global DSM 30m data is employed to obtain terrain information. All imagery data is sourced from Google Earth Engine (Gorelick et al., 2017).

Due to the influence of fog and cloud cover in the mountainous climate, a strategy is employed to enhance the clarity of pixels. To achieve this, imagery from three consecutive years (2019-2021) is combined. Three different approaches are utilized for land use training: annual image, seasonal image, and continuous change detection and classification (CCDC) methods. For annual and seasonal images, the first step involves cloud masking, followed by selecting the 40th percentile value of each pixel in each spectral band. Subsequently, normalization is performed by scaling the values of each band using the 0.5th and 99.5th percentiles. In the case of seasonal images, the following assignments are made: March to May as spring, June to August as summer, September to November as fall, and December to February as winter. The bands information of each model is shown in Table 1.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Data | Resolution | Bands | Index |
| RF | Annual | 10, 20, 30 | 7 | Blue, Green, Red,  NIR, SWIR1, SWIR2, DSM |
| RF | Seasonal | 10, 20, 30 | 25 | Blue, Green, Red,  NIR, SWIR1, SWIR2 for 4 seasons, DSM |
| RF | CCDC | 10, 20, 30 | 49 | CCDC 0~7(0 and 1 combined), RMSE for 6 bands, DSM |
| U-net | Annual | 10, 20, 30 | 7 | Blue, Green, Red,  NIR, SWIR1, SWIR2, DSM |
| U-net | Seasonal | 10, 20, 30 | 25 | Blue, Green, Red,  NIR, SWIR1, SWIR2 for 4 seasons + DSM |
| U-net | CCDC | 10, 20, 30 | 17 | Blue: 0&1, 2, RMSE, Green: 0&1, RMSE, Red: 0&1, 4, NIR: 0&1, 3, 4, RMSE, SWIR1: 0&1, 5, RMSE, SWIR2: 0&1, 3, DSM |

Table 1. Satellite imagery bands of each model after pre-processing.

* 1. **Training points and validation points**

We employed ground truth data to train and validate our models. The ground truth data for tea plantations were derived from Aerial Survey and Remote Sensing Branch’s aerial photographs. After a process of artificial selection and digitization, we obtained precise tea plantation area measurements. For other land cover classes, such as forests, soil, impermeable surface, and crops, we utilized National Land Surveying and Mapping Center’s predefined land use maps and digitized the corresponding polygons.

To establish our training and validation datasets, we initially divided the polygons into distinct areas. Subsequently, we randomly generated training and validation points within these ground truth polygons, ensuring that each point was separated by a minimum distance of 90 meters. To enforce this spatial separation, we employed GIS software, specifically QGIS (Version 3.16; QGIS Development Team, 2021). The minimum distance between points was configured using QGIS. The distribution of validation points for each land cover class is summarized in Table 2.

|  |  |  |
| --- | --- | --- |
|  | Training points | Validation points |
| Forest | 2967 | 442 |
| Soil | 330 | 50 |
| Tea plantation | 670 | 50 |
| Crops | 458 | 234 |
| Impermeable surface | 695 | 50 |
| Total | 5120 | 826 |

Table 2 Distribution of Training and Validation Points for Each Class. It calculates the total number of validation points for the research area and assigns each class based on area ratios, with a minimum of 50 validation points for each class (Olofsson et al., 2014).

* 1. **Machine learning models**

Two machine learning models, Random Forests (RF) and U-net + Resnet-18 (U-net), were utilized in our study. RF is a highly effective method based on decision trees (Breiman, 2001). One of its advantages is its ability to detect data correlations, obviating the need for extensive image pre-processing. We employed the randomForest function (v 4.7-1.1) in the R language for computation, validation, and result output.

U-net is specialized in feature and texture extraction through image segmentation, allowing it to consider surrounding pixels during identification (Ronneberger et al., 2015). This feature enhances its effectiveness in image processing. Prior to model training, we divided the entire image dataset into patches and further separated these patches into training and validation sets in a near 6:4 ratio. This training data is distinct from the final validation dataset. Additionally, we augmented the training dataset by rotating the original data, enabling the model to recognize objects from various angles. We chose Resnet-18 as the backbone architecture (He et al., 2015), and after assessing the model's stability, we determined that model accuracy stabilized after 40 epochs, prompting us to run 50 epochs during each training session.

* 1. **Data evaluation**

To assess the effectiveness of our models, we utilized the same validation dataset, ensuring a fair comparison between Random Forests (RF) and U-net. Each test was repeated 10 times to gauge model stability. To maintain fairness, we employed an identical validation dataset for both RF and U-net. Specifically, we randomly selected 80 percent of the validation points and repeated this process 20 times to calculate the average and standard deviation.

Notably, we encountered the highest confusion between forest and tea plantation classes. Therefore, we employed several evaluation metrics to gauge model effectiveness, including Overall Accuracy (OA), User's Accuracy (UA) for tea plantations and forests, and Producer's Accuracy (PA). These metrics offer valuable insights into model accuracy and performance. Additionally, we constructed each model 10 times, assessing both its accuracy and standard deviation. By considering OA, UA, and PA for tea plantations, along with the standard deviation of these metrics, we determined the stability of each model.

# RESULT

The results demonstrate the superiority of U-net over RF in land classification (Fig. 2 and Fig. 3). RF tends to mix tea plantations and forests in its classifications, while U-net consistently outperforms RF across various evaluation metrics. RF shows high OA but struggles to accurately identify tea plantations (Fig. 4). In contrast, U-net exhibits high effectiveness, particularly in seasonal image classification, accurately identifying tea plantations and other land classes. U-net is recommended for land classification in mountainous tea farming regions. Both models exhibit high stability (Table 1), with standard deviations (STD) < 2%. For tea plantations, U-nets show even better stability, with accuracy STD values of 5-6%.

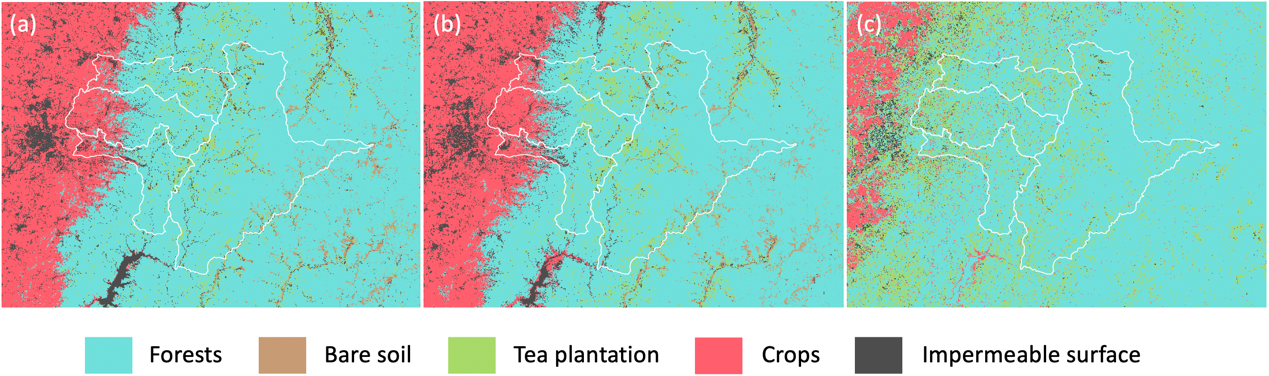


Fig. 2 Training results of RF models using (a) annual, (b) seasonal and (c) CCDC images.

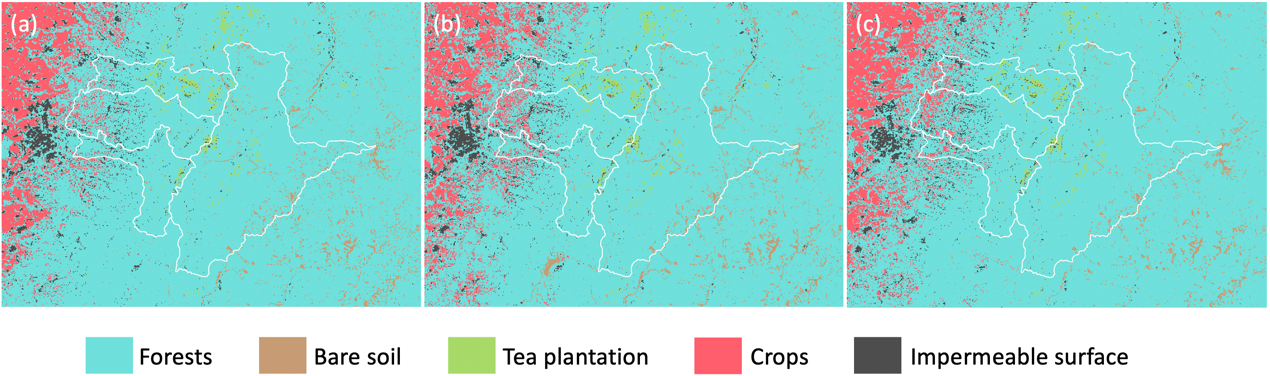


Fig. 3 Training results of U-net models using (a) annual, (b) seasonal and (c) CCDC images.

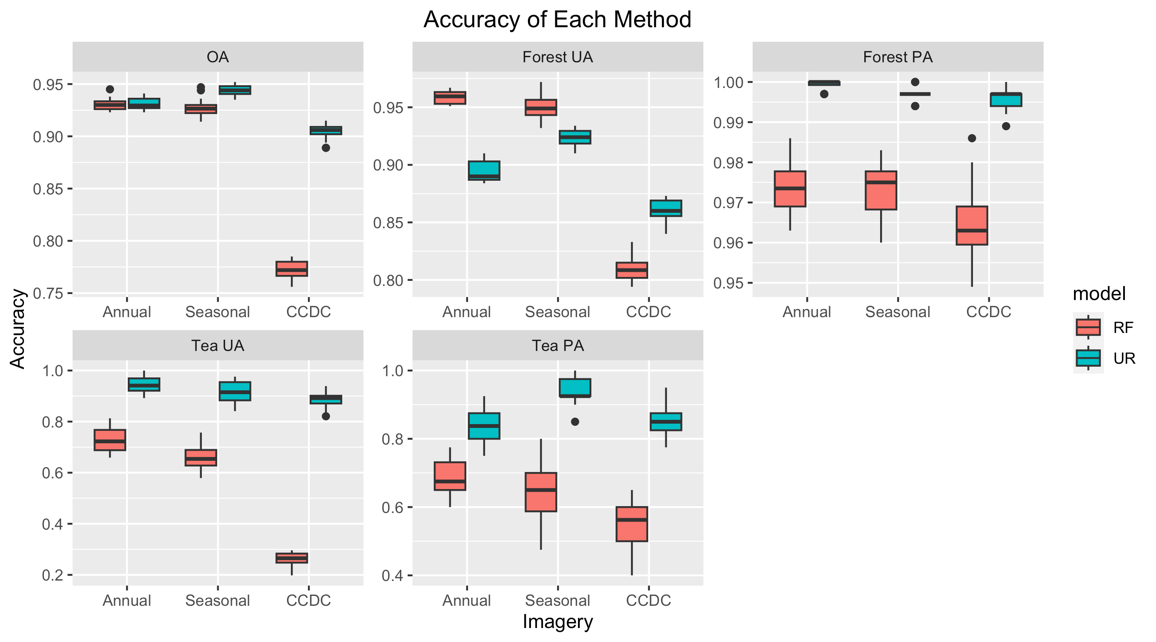


Fig. 4 The OA, UA and PA of tea plantations and forests of each model and imagery.

Table 3 The stability of each model and imagery

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Imagery | OA | OA std | Tea UA | Tea UA STD | Tea PA | Tea PA STD |
| RF | Annual | 0.937 | 0.006 | 0.753 | 0.056 | 0.698 | 0.045 |
| RF | Seasonal | 0.927 | 0.006 | 0.693 | 0.047 | 0.605 | 0.078 |
| RF | CCDC | 0.775 | 0.01 | 0.271 | 0.029 | 0.543 | 0.079 |
| U-net | Annual | 0.946 | 0.012 | 0.927 | 0.043 | 0.865 | 0.049 |
| U-net | Seasonal | 0.92 | 0.021 | 0.9 | 0.041 | 0.933 | 0.037 |
| U-net | CCDC | 0.923 | 0.011 | 0.912 | 0.028 | 0.85 | 0.058 |

# DISCUSSION

U-net excels in accurately identifying tea plantations within mountainous regions, effectively capturing image features, and preserving vital information. We attribute this superiority to U-net's utilization of data augmentation, which extends available training data, and its incorporation of a skip-connection structure between the encoder and decoder. This structure prevents feature loss when reducing and evaluating resolution (Ronneberger et al., 2015). This discovery suggests potential applications for mapping other mountain cash crops, such as coffee plantations. However, it's essential to exercise caution when relying exclusively on CCDC methods in foggy and cloud-prone areas. These conditions can lead to decreased accuracy due to limited valid pixels, especially if invalid pixels are not uniformly distributed throughout the months, causing the CCDC algorithm to lack sufficient seasonal information during formula fitting. In summary, U-net shows promise for quantifying tea plantations in tropical mountainous regions.

# CONCLUSION

U-net is superior to RF for accurately identifying tea plantations in mountainous regions. It effectively captures image features and maintains critical information. The finding to this study suggests potential applications for mapping other cash crops in mountains such as coffee plantations. However, caution is advised when relying solely on CCDC methods in foggy and cloud-prone areas due to lower accuracy caused by limited valid pixels. In summary, U-net holds promises for tea plantation quantification in tropical mountainous regions.

**REFERENCES**

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5-32. <https://doi.org/https://doi.org/10.1023/A:1010933404324>

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. . *Remote Sensing of Environment.* <https://doi.org/10.1016/j.rse.2017.06.031>

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. <https://doi.org/https://doi.org/10.48550/arXiv.1512.03385>

Hicks, A. (2009). Current Status and Future Development of Global Tea Production and Tea Products. *AU Journal of Technology*, *12*, 251-264.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, *148*, 42-57. <https://doi.org/https://doi.org/10.1016/j.rse.2014.02.015>

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. <https://doi.org/https://doi.org/10.48550/arXiv.1505.04597>

Zhu, Z., & Woodcock, C. E. (2014). Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, *144*, 152-171. <https://doi.org/https://doi.org/10.1016/j.rse.2014.01.011>

Zhu Zhe, D. C. (2020). Continuous subpixel monitoring of urban impervious surface using Landsat time series. *Remote Sensing of Environment*, *238*. <https://doi.org/https://doi.org/10.1016/j.rse.2018.10.011>