Using remote sensing images and random forest algorithm to estimate spatiotemporal patterns of landslides and landslide-derived biomass loss

TCCIP 2023 AR6 新資料說明會 暨 氣候變遷資料應用研討會

Hsueh-Ching Wang (王雪卿)

Department of Earth and Life Science, University of Taipei

hcwang@utaipei.edu.tw

Introduction

Landslide is a wide range of ground movements of a mass rock, debris, soil down a slope, which can cause a catastrophic disaster. Typhoon with extreme precipitation is one of the major factors to trigger large amount of landslide in Taiwan. Climate change may exacerbate the intensity or frequency of extreme precipitation, which might alter the spatiotemporal characteristics and intensity of landslide. Landslide can destroy vegetation and cause the loss of aboveground biomass and the decrease of net primary production (NPP), but few of studies focus on the impacts of landslide on carbon budget. We developed an automated landslide susceptibility mapping by random forest classifier algorithm from open-access 30 m Landsat imagery to assess the long-term (1990–2022) spatiotemporal patterns of landslide activities in the Chilan mountain cloud forest.

Methods

- For landslide classification, Landsat multispectral imagery was used to extract variables, including spectral features such as blue, green, red, near-infrared (NIR), shortwave-infrared 1 (SWIR1), and shortwave-infrared 2 (SWIR2), as well as seasonal vegetation spectral features such as Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), and Enhanced Vegetation Index (EVI).
- Terrain variables such as DEM, aspect, slope, plan, and profile curvature were derived from DEM data. All variables were resampled to 30 m pixel size using the nearest neighbor method.
- We used the Random Forest model combined with the recursive feature elimination (RFE) method.
- The trained model was then used to extract areas of new landslides, which were analyzed in conjunction with TCCIP TReAD daily rainfall data.

Results and discussion

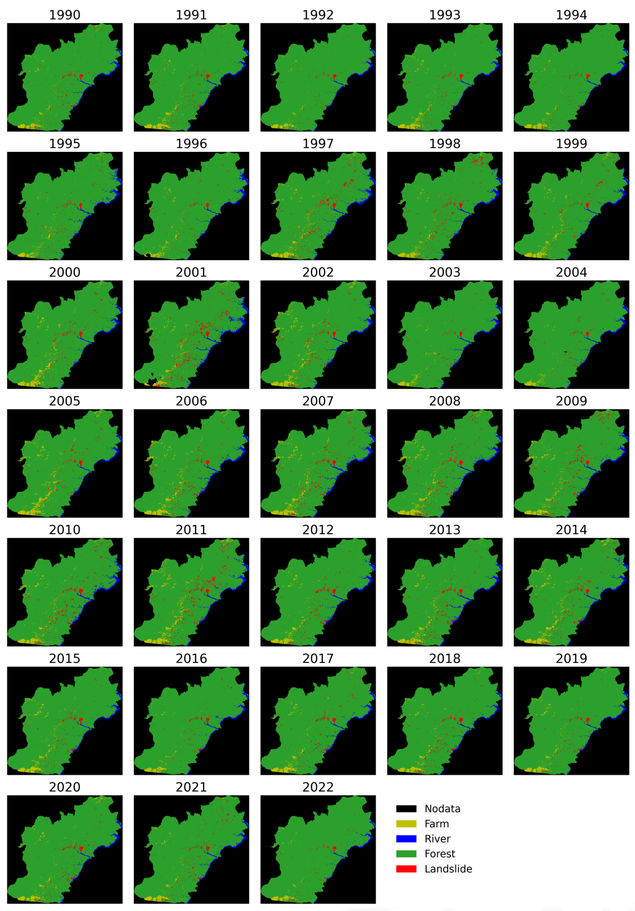

Fig. 1. The land is automatically classified into different categories using the random forest model from 1990 to 2022. Yellow represents farmland, blue represents rivers, green represents forests, and red represents landslides.

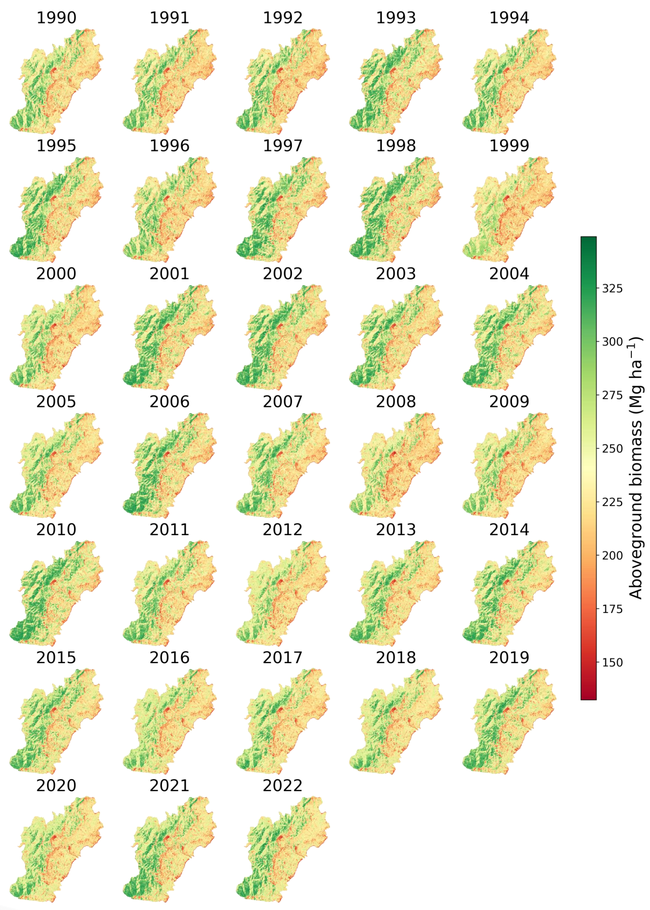

Fig. 2. The spatiotemporal patterns of landslide-derived annual aboveground biomass (AGB) (Mg ha-1) during 1990-2022.

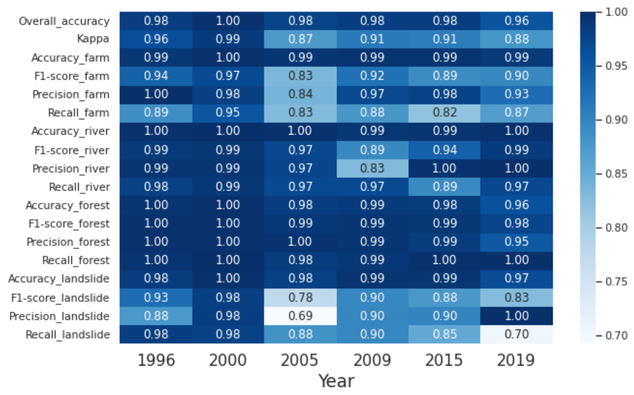

Fig. 3. Validation results of the verification set based on the trained Random Forest model. Performance metrics include overall accuracy, Kappa coefficient, and confusion matrix for each corresponding classification.

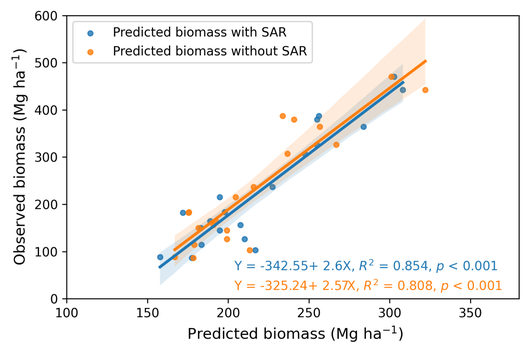

Fig. 4. The relationship between predicted and observed aboveground biomass (Mg ha^{-1}) by random forest algorithm with or without Sentinel-1 SAR dataset.

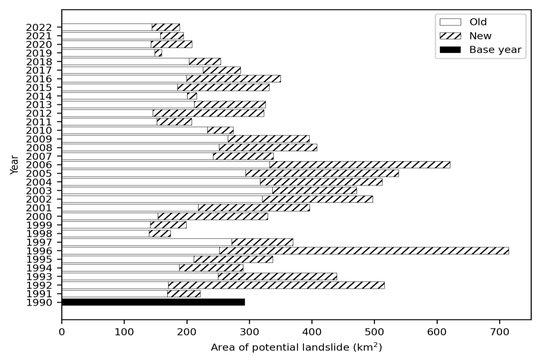

Fig. 5. The estimated landslide area is divided into old landslide area (year t-1: non-landslide and year t: landslide) and new landslide area (year t-1: landslide and year t: non-landslide) from 1990 to 2022.

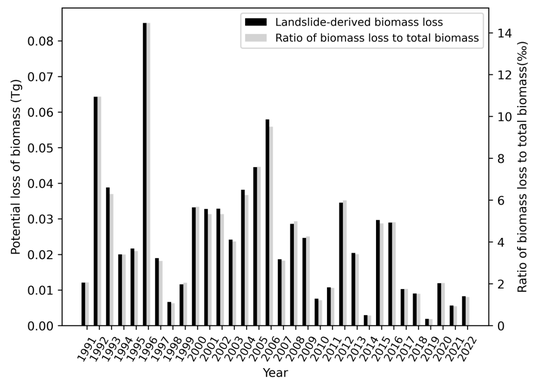

Fig. 6. The ratio of potential biomass loss to total aboveground biomass from 1991 to 2022.

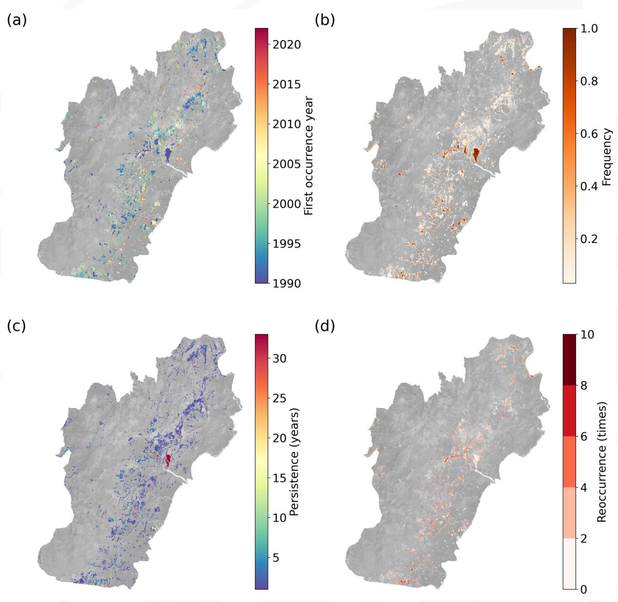

Fig. 7. The spatial patterns of first occurrence, frequency, persistence and reoccurrence of landslide.

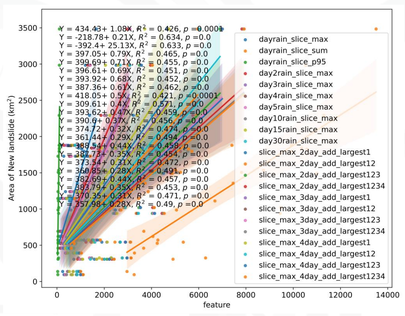

Fig. 8. The relationship between newly increased landslides and daily rainfall based on different rainfall compositions.

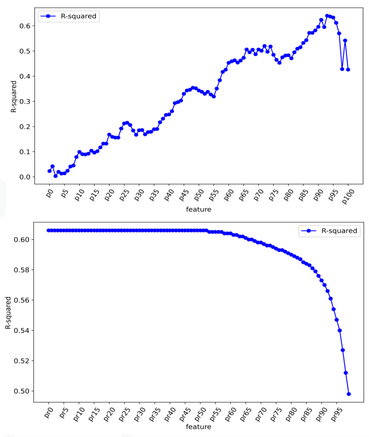

Fig. 9. Trend chart of the R-squared values for the relationship between newly increased landslides and daily rainfall with (a) percentage and (b) cumulative percentile of 0-100.

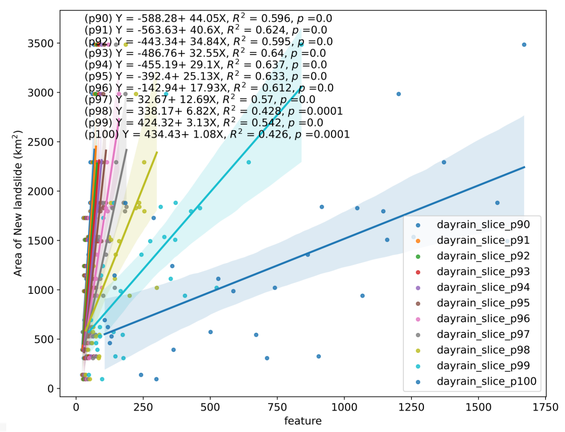

Fig. 10. The relationship between newly increased landslides and a percentile of 90-100.

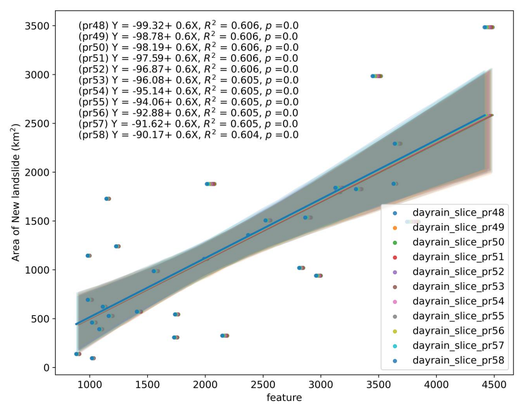

Fig. 11. The relationship between the newly increased landslides and the cumulative values of the percentile of 48-58%.

The Landsat imagery, LiDAR and Sentinel-1 SAR data were used to retrieve the long-term landslide susceptibility mapping from 1990 to 2022 in the Chilan mountain cloud forest. Our findings showed that integrating random forest classifier with feature selection from the recursive feature elimination to remove the redundant features can improve the accuracy of landslide classification with the high overall accuracy (0.96-1) and kappa (0.88-0.99). The predicted biomass can explain 80% variance of observed biomass. For the spatiotemporal patterns of landslide, the extreme precipitation from typhoons was the major disturbance to cause the substantial landslides. The cumulative daily rainfall from 52 percentile (PR52) had a significant explanatory power (R^2 =0.606, p < 0.001) for the occurrence of new increased landslides.

TCCIP 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

指導單位 NSTC 國家科學及技術委員會 National Science and Technology Council 自然科學及永續研究發展處

計畫辦公室 NCDR 行政法人國家災害防救科技中心 National Science and Technology Center for Disaster Reduction