國立臺灣大學理學院地理環境資源研究所

碩士論文

Department of Geography

College of Science

National Taiwan University

Master Thesis

運用機器學習整合衛星資料
量化亞熱帶日間山地雲霧森林之雲霧時空分佈

Application of machine learning to integrate satellite data to quantify spatiotemporal distribution of diurnal fog and low stratus occurrence in subtropical montane cloud forests

張桔云

Jie Yun Chong

指導教授：黃倬英 博士

Advisor： Cho-Ying Huang, Ph.D.

中華民國 112 年 2 月

February, 2023

doi:10.6342/NTU202300432

國立臺灣大學碩士學位論文
口試委員會審定書

運用機器學習整合衛星資料
量化亞熱帶日間山地雲霧森林之雲霧時空分佈
Application of Machine Learning to Integrate Satellite Data to Quantify Spatiotemporal Distribution of Diurnal Fog and Low Stratus Occurrence in Subtropical Montane Cloud Forests

本論文係張桔云（R09228025）在國立臺灣大學地理環境資源學系、所完成之碩士學位論文，於民國 111 年 12 月 15 日承下列考試委員審查通過及口試及格，特此證明。

口試委員：黃倬英 （簽名）

（指導教授）

中井太郎

誌謝

兩年多的研究生活讓我覺得充實愉快，感恩大家在研究上的激勵與共同成長！何其有幸接受了黃老師的指導，感謝Choy，讓我在研究中大膽假設，在迷茫中找經驗，在修正中給與耐心與指導。謝謝羅敏輝老師和系中井太郎老師和莊振義老師從棲蘭 meeting 與我探討，教導我從不同角度對待研究，提點細節中的注意事項。以及感謝諸多地理系的老師提供專業的學術討論與建議，讓我收穫許多觀點和激勵自己必須更努力學習。非常感激 605/8 研究室大家的提攜與照顧，在每次開會中的建議與發想，在每個野外中的配合與鼓勵。謝謝欣儒來對我的關懷與教導，謝謝婉瑜在文書上的提醒和協助，當我遇到阻礙時很用心給與我學術和經驗上的意見。謝謝我的家人與朋友的信任與包容，讓我專心投入研究。謝謝棲蘭讓我看到台灣不一樣的山地景觀，在我煩躁和沮喪時獲得療愈能量。感謝這兩年多來許許多多幫助我、關心我的人，感謝你們，謹此獻上我誠摯的謝意！

doi:10.6342/NTU202300432

摘要

山地霧林廣義指被雲霧頻繁籠罩的山地森林，通常為濕暖空氣受到地形抬升作用形成雲霧。獨特的水文氣候特徵為森林提供足夠水源和養分，穩定高濕度環境減少太陽輻射透射量和森林蒸發散進而增加林相的多樣性。近期觀測研究表明氣溫的上升可能會影響雲霧帶的時空分佈，這將會對山地霧林的分佈造成威脅。為瞭解其分佈的變化機制和評估其帶來的影響，光學遙測技術視爲量化雲霧的重要觀測依據。然而，先前的研究受天氣影響限制了多雲區域的評估可行性。本研究針對不同的天氣特徵的山地霧林，提出以利用日本氣象衛星向日葵八號連續時間高密度之可見光、近紅外光、紅外光和太陽幾何角（高度角和天頂角）的觀測資料，搭配雙頻道亮溫差異、植被指數和地表特徵變數，如海拔、坡度、地形位置指數、地表粗糙指數和地形湿度指数（共 31 變數）作特徵萃取，結合監督式機器學習模型以建立日間（7 時至 17 時）霧和低雲(FLS)的偵測算法。研究樣區為植被同質性較高的亞熱帶山地霧林-台灣棲蘭山區，於海拔 1151-1810 公尺之間安裝四個縮時攝像作爲現地雲霧觀測依據，以二元方式分類 FLS 事件。本研究在 2018-2021 期間共蒐集了 53 358 張影像，利用其作爲訓練數據建構三種不同天氣特徵（黃昏/黎明，晴空，多雲）之隨機森林（Ranger）模型,進行獨立測試結果，並以 F_1 分數作爲衡量二分類模型精確度指標。結果顯示晴空條件下 F_1 分數為 0.864，黃昏/黎明和多雲條件下皆爲 0.945，這顯示該模型可在不同天氣條件下穩定地偵測山地霧林之 FLS 事件。這一發現有助於系統性量化山地霧林 FLS 事件之時空變化。

關鍵字：向日葵八號、隨機森林、遙測、台灣、縮時攝影、地形

doi:10.6342/NTU202300432

Abstract

Montane cloud forests (MCFs) are characterised by the presence of persistent, frequent wind-driven, horizontal belt shaped and orographic clouds also known as a cloud band. The hydroclimatic characteristics of MCFs often acts as a water supply to ecosystems due to the prevailing perhumid and dim light, where moisture introduced by depositing cloud can be more efficiently retained. Recent observation has shown that elevated temperatures may lift cloud band, which would cause colossal impacts on MCFs. The first step to assess the potential ramification is to quantify the occurrences of fog and low stratus (FLS) regionally; satellite remote sensing is an ideal tool for the task. However, previous research efforts may only be effective in limited weather condition, confining the feasibility in this cloudy region. In this study, we developed an algorithm to detect diurnal (defined as 07:00-17:00 in this study) FLS occurrence that was insensitive to weather conditions. We used the visible and infrared bands of the Advanced Himawari Imager on board Himawari-8 (H-8), sun geometry (solar zenith [SZA] and azimuth [SAA]) angles of each pixel, dual band differences (DBD), normalised difference vegetation index (NDVI) and local topographic variables (elevation, slope, topographic position index, vector ruggedness measure, topographic wetness index) as input data (31 variables) to model the FLS using “RANdom forest GEneRator” (Ranger), a recently developed machine learning approach derived from random forest). We carried out the study in subtropical MCFs of Chi-Lan Mountain in northeast Taiwan. We installed four ground FLS observation stations across an elevation range of 1151-1811 m a.s.l with 53 358 diurnal time-lapse photographs from 2018-2021 with or without FLS identified by visual assessment. We applied three different model (twilight/dawn, cloudy, clear sky) settings to model FLS occurrence because of the responses of the various bands depending on the time of day and the underlying surface characteristics. We randomly selected 80% of the data for Ranger development and the rest of data for validation. We found that it was possible to detect FLS occurrence in MCFs regardless of the weather conditions using the proposed method with the overall F_1-scores ≥ 0.864, and 0.945 and 0.945 for clear-sky, twilight/dawn and cloudy conditions. The finding may facilitate systematic mapping of FLS occurrence in MCFs.

Keywords

Himawari-8, ranger, remote sensing, Taiwan, time-lapse photography, topography

doi:10.6342/NTU202300432

Table of Contents

doi:10.6342/NTU202300432

doi:10.6342/NTU202300432

List of figures

doi:10.6342/NTU202300432

List of tables

doi:10.6342/NTU202300432

Chapter 1: Introduction

Montane cloud forests (MCFs) are characterised by persistent and frequent wind-driven orographic fog and low stratus (FLS) (Grubb 1977; Hamilton 1995). The distribution of FLS occurrence in MCFs varies depending on winds and rainfall patterns, latitude, distance from the sea and the size of the mountains and local vegetation classification. The frequency and duration of FLS occurrence vary greatly as the irradiance is reduced by 10-50% and the persistent wetting of the leaves leads to reduced photosynthesis (Bruijnzeel and Veneklaas 1998; Frahm and Gradstein 1991) but see Dawson and Goldsmith (2018), which forms unique mountain environments. Moreover, the prolonged availability of perhumid and dim light conditions in this ecosystem, enhance vegetation growth in MCF by shortening the desiccation period and lengthening the period of photosynthetic activity of the plants. Since sea surface temperature and saturated adiabatic lapse rates and their diurnal and seasonality are highly spatially variable across tropics, the altitudinal bands where ground-level clouds occur are also likely to be highly variable (Bruijnzeel and Veneklaas 1998). Accordingly, MCFs is located between 500 m and 3500 m altitude in the humid tropics (Eva et al. 2002; LaBastille and Pool 1978) with major occurrence between 1,200 and 2,500m . The diurnal cycle of FLS occurrence often acts as a water supply to ecosystems (Bruijnzeel et al. 2011; Glasow and Bott 1999) and, by influencing radiative transfer processes in the atmospheric boundary layer, it acts as a modifier in the climate system (Vautard et al. 2009). In particular, the deposition of FLS contributes to nutrient inputs in forests and the deposition carbon flux of these compounds through the canopy-held epiphytic biomass

doi:10.6342/NTU202300432

over forests is a major contributor to ecosystem acidification and eutrophication (Alewell et al. 2000; Chang et al. 2002; Conley et al. 2009). Therefore, MCFs play an important role in regulating the regional hydrological cycle. This unique physical setting harbors high biodiversity and many of them are endemic (Gentry 1992).

A large body of literature has shown that small changes in temperature and rainfall can alter microclimate conditions and threaten existing ecological communities; whereas these ecosystems exist in rugged mountainous regions, narrow microclimate zones can change significantly over short distances (Gotsch et al. 2017; Nadkarni and Solano 2002; Ray et al. 2006; Williams-Linera 2003; Williams et al. 2007). Therefore, the rate of warming is expected to be greater at higher altitudes than at lower altitudes, as it has been reported in the montane around the world (Bradley et al. 2006; Román-Cuesta et al. 2014). Although uncertainties remain about the effects of temperature and humidity conditions on cloud band formation in MCFs (Lawton et al. 2001), the upper displacement of the condensation zone is expected (Halladay et al. 2012; Lawton et al. 2001). Consequently, the altitude of the cloud band may be shifted upward, and the increase in solar radiation cloud also increase the evapotranspiration of the forest and causing reduce the interception of water by the vegetation in MCFs. As a result, changes in the water cycle will lead to droughts throughout the ecosystem and MCFs be recognised as one of the most threaten ecosystems in the world.

One of the basic metrics to quantify FLS in mountainous regions is the occurrence and consequentially the duration; field air temperature, humidity and leaf wetness sensors are commonly used for the measurement (Bruijnzeel et al. 1993; Chu et al. 2014; Gotsch et al. 2014). Some field studies utilised ceilometer lidar and cloud-detection radar to detect FLS occurrence in order to estimate cloud base height (Gaumet et al. 1998; Takano

doi:10.6342/NTU202300432

et al. 2010). However, the high cost of the instruments constrained to identify liquid water clouds due to sensor signals are strongly attenuated in rugged mountain regions. Besides, visibility sensors as ancillary weather monitoring sensors need a specific observation angle and raised flat ground to determine whether they are immersed in clouds and mists based on the output of infrared light and the distance of receiving a specific point (Beiderwieden et al. 2007; Chang et al. 2002). An alternative is time-lapse photography, but cloud-sensitive of capture image characteristics (contrast, the coefficient of variation and the entropy of pixel luminance, and image colourfulness) are required a complex iterative quantisation to quantify spatiotemporal cloud immersion in MCFs (Bassiouni et al. 2017). In addition, the aforementioned instruments of sparse measurement-based can represent the monitoring approach of local areas with low spatial coverage, which is somewhat limited for regional monitoring.

Systematic detection of FLS using satellite imagery permits the quantification over a vast region. There are several approaches available for FLS detection including (1) a physical method using the one-dimensional variational system, (2) a split-window algorithm, and (3) machine learning methods. The physical method is based on modelling the lifting condensation level of cloud to obtain cloud base height from radiosondes data and weather station measurements and compared it with a digital elevation model (DEM) to confirm areas covered by cloud (Stackpole 1967). All parts of the cloud layer during modelling have uniform geometric thickness, but in the case of orographic fog, terrain factors need to be considered. The split-window method is based on the dual-band difference (DBD) of infrared spectral bands to retrieve cloud properties in the atmosphere (Lee et al. 2011). The most common method involves DBD of thermal (10.8 µm) and middle infrared (3.7 µm) spectral bands. The middle infrared has a solar component,

doi:10.6342/NTU202300432

which limits the theoretical use of a combination of brightness temperatures in the middle infrared and thermal infrared to detect spectral information of FLS (Cermak and Bendix 2008). Moreover, a purely thermal infrared -based detection of FLS is not possible, as the brightness temperatures of FLS and land surfaces may be similar (Güls and Bendix 1996). The application of machine learning based has been widely used in the field to quantify FLS using satellite remote sensing (Liu et al. 2021; Tan et al. 2021; Wang et al. 2021). Using ground and satellite meteorological attributes data, Li (2022) compared several machine learning algorithms for monitoring the occurrence of diurnal FLS and finds that random forests outperform others. However, the approach heavily replied on field meteorological data preventing spatially continuous FLS mapping.

Therefore, the objective of this study is to develop a novel satellite-based approach to provide wall-to-wall FLS coverage over subtropical MCFs. We further analysed variables and seasonal variation that were pivotal for FLS occurrence modelling.

doi:10.6342/NTU202300432

Chapter 2: Materials and methods

2.1 STUDY AREAS

Taiwan (21˚85'- 25˚30'N, 120˚00'- 122˚00'E)) is a 36,000 km^2 subtropical mountainous island located in Pacific Asia. According to the National Vegetation Database of Taiwan, MCFs are mainly located 1500-2500 m a.s.l. (Li et al. 2013; Schulz et al. 2017). The study site is the 24 400 ha MCFs of Chi-Lan Mountain located in northeast Taiwan with the elevation range from 310-2845 m a.s.l. (Hu and Huang 2019) (Fig. 1). The regional climate is generally humid with precipitation of 3000-4000 mm yr^{-1} and mean air temperature of 13.9 ℃ (Chu et al. 2014). According to the investigation of vegetation composition, the yellow cypress has the dominant which occupies 82% of the total basal area (Chang et al. 2006). The MCFs of study site are characterised by increased relative air humidity through frequent incidence of FLS, high abundance and species richness of epiphytes, and accumulation of large amounts of epiphytic biomass (Lai et al. 2021; Lai et al. 2020). In addition, nutrient fluxes were estimated more than 50% of the ecosystems input was through FLS deposition (Chang et al. 2002). The topography of study site are relatively homogeneous south-eastern facing slope that extend about 2-km with average slope of 14% (Chu et al. 2014). Due to rugged terrain, study site is often affected by orographic lifting FLS, which occurs approximately 38% of the time of a year (Chang et al. 2006). Furthermore, FLS significantly affects the radiation and horizontal precipitation, causing its interaction with the atmosphere and ecology, thus increasing air humidity and water supply; reducing solar radiation and evapotranspiration (Gu et al.

2021), without distinct wet and dry seasons. Li et al. (2022) depicted that the core cloud band of the study site was narrow mainly distributed at 1514-1670 m a.s.l.

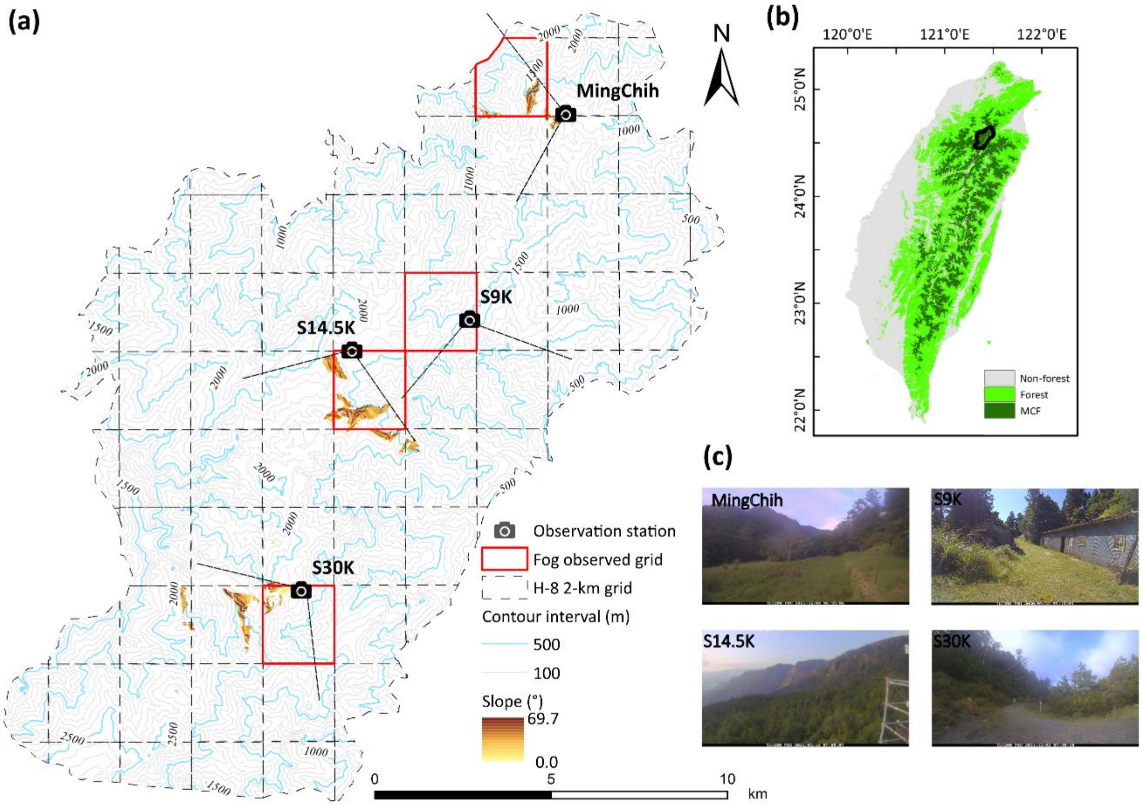

Fig. 1.(a) The study site, MCFs in Chi-Lan Mountain in (b) north-eastern Taiwan (the black polygon). Four fog observation stations installed on the site, and the black-dashed lines and red monochromatic coloured indicated H-8 2-km grid and the areas visible by the stations, respectively. (c) The photographs taken by time-lapse cameras show the field of views of the stations.

2.2 HIMAWARI-8 OBSERVATION DATA

The Advanced Himawari Imager on-board the Himawari-8 (H-8) launched by Japan Meteorological Agency, has 16 spectral bands - four visible (VIS), two near-infrared (NIR), and ten infrared (IR) spectral bands. The satellite enable continuous observation in East Asia and Western Pacific regions with a 10-minute temporal interval and spatial resolutions of VIS (0.5-1 km), NIR (1-2 km) and IR (2 km) spectral bands (Bessho et al. 2016) (Table 1). With these advantages, Himawari-8 AHI provides capability for atmospheric and environmental monitoring using machine learning (Lee et al. 2019; Tan et al. 2021). These applications have shown the geostationary satellites have the potential to draw a spatiotemporally coherent picture of the FLS occurrence (Andersen and Cermak 2018; Egli et al. 2017; Iwabuchi et al. 2018).

To make widespread application of the enhanced monitoring capabilities of Himawari-8, level-1 grid products have been provided by the Japan Aerospace Exploration Agency P-Tree System (https://www.eorc.jaxa.jp/ptree). Corresponding geometric attributes including solar zenith angle (SZA) and solar azimuth angle (SAA) also came with the spectral data and were utilised for the modelling. The brightness temperature of window bands (B_{11}, B_{13}, B_{14} and B_{15}) are related to land and sea surface temperatures, while the water vapor bands (B_8, B_9 and B_{10}) are associated with the distribution of water vapour in three different vertical layers. B_{12} and B_{16} correspond to O_3 and CO_2 absorption. In addition, the differences between the window band and water vapour band ($DBD_{14\text{-}8}$, $DBD_{14\text{-}9}$, $DBD_{14\text{-}10}$) and themselves ($DBD_{14\text{-}11}$, $DBD_{14\text{-}15}$). Studies using ground measurements have shown that the spatial distribution of FLS can be affected by topography and in turn, the distribution of FLS can affect vegetation patterns (Ball and Tzanopoulos 2020; Gultepe et al. 2007). The Normalised Difference Vegetation Index (NDVI) (Rouse et al. 1974) calculated using red (centred at 0.64 μm, Table 1) and NIR (0.86 μm) has been extensively used to monitor vegetation greenness and productivity (Fensholt et al. 2011; Sjöström et al. 2009). Optical satellite images can be used to assess and quantify the impacts of FLS on vegetation at large spatial scales, which

doi:10.6342/NTU202300432

significantly positive relationships in the NDVI were found between FLS and vegetation attributes (Kaseke et al. 2018; Qiao et al. 2020; Spirig et al. 2019).

Yoo (2010) pointed out that the satellite-observed DBD_{7-13} values were varies according to time because B_7 (3.9 µm) measures both reflectivity and emissivity during daytime. Therefore, most of the studies have been separately developed for dawn/twilight or daytime according to SZA (Suh et al. 2017). Purbantoro (2018) compared band pairs of B_{13} (10.4 µm) and B_{15} (12.4 µm) and B_{15} and B_{16} (13.3 µm) to detect different cloud types in summer and winter seasons; the former combination is useful for clear sky and low stratus. The initial threshold values of test elements (e.g., $DBD_{7-13,}$ DBD_{13-15}) through histogram analysis and prepares the background data, such as the ground surface data. Since the response of the various band depends on the time of day and underlying surface characteristics, the algorithm (Fig. 2) was developed to detect diurnal FLS occurrences according to the SZA (dawn/twilight) and a uses the threshold in DBD_{13-15} to distinguish clear sky from a cloudy condition.

2.3 TOPOGRAPHY AND LANDFORM DATA

Field observation showed that the spatial distribution of FLS can be influenced by topography (Gultepe et al. 2007), which may facilitate predicting the spatiotemporal dynamics over a vast region . The topography of Chi-Lan Mountain is highly variable, so we characterised FLS occurrence frequency for a DEM with 2-km spatial resolution. We utilised the Spatial Analyst toolbox in ArcMap v10.7 (Environmental Systems Research Institute, Inc., Redland, CA, USA) to derive an array of topographic metrics from 30-m ASTER-GDEM (the Advanced Spaceborne Thermal Emission and Reflection Radiometer-Global Digital Elevation Model) data. We applied a nearest neighbourhood approach to resample the data to 2-km to match other data. Topographic variables

doi:10.6342/NTU202300432

including elevation, slope, the topographic position index (TPI), the vector ruggedness measure (VRM) and the topographic wetness index (TWI) were used as input variables derived from GDEM V2 (Table 1). The topographic metrics the TPI used different radii derived from elevation to measure the difference between a central cell elevation and the average elevation around it within a predetermined radius (Guisan et al. 1999); the VRM measured terrain ruggedness as the variation in the three-dimensional orientation of grid cells within a neighbourhood (Hobson 2019), and the TWI is a commonly used index in hydrological analysis for describing the tendency of an area to accumulate water and the probability of the area being wet (Beven and Kirkby 1979).

Table 1.Summary of the physical characteristics of H-8 observation data and local topographic variables in the model of detection diurnal FLS occurrence.

Input variable	**Physical characteristics**
H-8 observation data	
B_1 (0.47 μm)	Daytime aerosol over land
B_2 (0.51 μm)	Color composite imagery
B_3 (0.64 μm)	Daytime vegetation/burn scar and aerosols over wate
B_4 (0.86 μm)	Daytime cirrus cloud
B_5 (1.6 μm)	Daytime cloud-top phase and particle size,snow
B_6 (2.3 μm)	Daytime land/loud properties, particle size, vegetation
B_7 (3.9 μm)	Low-level cloud/fog/wildfires
B_8 (6.2 μm)	High-level atmospheric water vapour
B_9 (6.9 μm)	Mid-level atmospheric water vapour
B_{10} (7.3 μm)	Lower-level water vapour
B_{11} (8.6 μm)	Total water for stablility, cloud pahse, dust, SO2
B_{12} (9.6 μm)	Total ozone, turbulence
B_{13} (10.4 μm)	Surface and cloud
B_{14} (11.2 μm)	Imagery, SST, clouds, rainfall
B_{15} (12.4 μm)	Total water, ash SST
B_{16} (13.3 μm)	Air temperature, cloud heights and amounts
DBD_{14-8}	Upper tropospheric moisture
DBD_{14-9}	Mid and upper tropospheric moisture
DBD_{14-10}	Mid tropospheric moisture
DBD_{14-11}	Amount of water vapour
DBD_{14-15}	Split-window bands (amount of water vapour)
SZA	Diurnal observation period
SAA	Temporal and seasonal characteristics
NDVI	Positive relationships between fog and vegetation attributes
Local topographic variables	
Elevation	Vertical distance between the sea level and point of sites
Slope	Difference in elevation and divide it by the horizontal
TPI	Distinguish topographic features by an elevation cell with its surrounding grid cells
VRM	Dispersion of vectors normal (orthogonal) to grid cells within the specified neighbourhood
TWI	Quantify topographic controls on hydrological processes characterising biological processes

2.4 GROUND OBSERVATION DATA

Four FLS observation stations (MingChih, S9K, S14.5K, S30K, Fig. 1) equipped with time-lapse cameras (TLC200 PRO [Brinno Inc., Taipei, Taiwan] and Trophy Cam [Bushnell Corporation, Overland Park, Kansas, USA]) were installed across an elevation gradient (Li et al. 2022). The data collection was initiated on 2020/03/17 (MingChih), 2020/08/17 (S9K), 2018/01/31 (S14.5K), and 2020/05/29 (S30K). FLS was defined as visibility less than 1 km according to the World Meteorological Organization. Data collections were acquired from the field covering the period between 2018/01/31 and 2021/12/31 and are used as the dependent variable and were further classified as binary (FLS vs. non-FLS) categories. Observation angles were able to view at least 1 km away. The time interval of data collection was 10 minutes to correspond with H-8. The observation period was 07:00-17:00 to ensure sufficient solar illumination through seasons for monitoring FLS (Li et al. 2022). The collocation between H-8 data and time-lapse photographs is essential because their different spatial and temporal resolutions may introduce uncertainty. Our matching strategy is to only involve H-8 pixels whose locations are nearest to the camera facing direction. In parallel, the time differences between the match-ups are restricted to 5 min.

2.5 RANGER

Random forest is a machine learning algorithm (Breiman 2001) which has the capability to deal with a complicated relationship between nonlinear input and prediction. Random forest is a non-parametric approach that subdivides data and explanatory variables in decision trees. It is based on a divide and conquer algorithm to separate input feature space into disjoint subsets where improve performance of individually weak decision tree models and can produce good predictions (Chen et al. 2018). The algorithm has been

doi:10.6342/NTU202300432

widely applied in biological research and gradually related to the fields of satellite remote sensing (Bax et al. 2021; Woodall et al. 2015). Li et al. (2022) found that the performance of random forest was superior to other machine learning algorithms (C5.0, CART, Ctree, cubist, neural network and support vector machine) for quantifying spatiotemporal dynamics of FLS occurrence in the study site. In this study, we selected an improved Ranger in r package for the task, which is the fastest and most memory-efficient implementation version of random forest to analyse high-dimensional data (Wright and Ziegler 2015). We conducted satellite (H-8 and ASTER) optical and topographic variables as input variables to model field FLS observation data (Fig. 2). We developed three models to predict FLS occurrence based on different weather conditions: dawn/twilight, cloudy, clear sky.

2.6 MODEL EVALUATION

The train-test split is a technique for evaluating the machine learning algorithm performance. All dataset (dawn/twilight, cloudy, clear sky) is randomly split into a training dataset (80%) and a test dataset (20%) - the averaged model's performance using repeated 10-fold cross-validation and taken as the mean from the number of repeats. The performance of random forests for probability estimation is usually measured using the Brier score (BS), which is the mean squared error between the variable state and the predicted probability:

$$\mathrm{BS} = \frac{1}{\mathrm{N}}\sum_{\mathrm{t}=1}^{\mathrm{N}}(\mathrm{f_t} - \mathrm{o_t})^2 \quad \in [0,1] \qquad \text{(Brier 1950)} \quad (1)$$

The parameter optimisation is conducted based on the BS (eq. 1) for the training dataset, which is a metric used to examine the goodness of a predicted probability score. This is applied for prediction probability scores, whose values range between 0 and 1 and the model closer to 1 perform better. Additionally, to determine the contribution of each input variable to the performance of the three models, relative variable importance indicating how much a given model uses that variable to make accurate predictions is analysed.

We assessed the performance of Ranger using F_1-score (F_1 hereafter) (eqs. 2–4), Accuracy (eq. 5) and the Matthews correlation coefficient (MCC) (eq. 6). We utilised these statistics because we consider different unbalanced situations in our dataset, the following statistical measures were calculated (Delgado and Tibau 2019):

$$F_1 = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \in [0,1] \qquad \text{(Chinchor and Sundheim 1993)} \qquad (2)$$

Where precision and recall are:

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \qquad (3)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \qquad (4)$$

$$\text{Accuracy} = \frac{\text{TN+TP}}{\text{TN+FP+TP+FN}} \in [0,1] \qquad (5)$$

$$\text{MCC} = \frac{\text{TP}\cdot\text{TN-FP}\cdot\text{FN}}{\sqrt{(\text{TP+FP})\cdot(\text{TP+FN})\cdot(\text{TN+FP})\cdot(\text{TN+FN})}} \in [-1, 1] \qquad \text{(Matthews 1975)} \qquad (6)$$

In this study, a binary confusion matrix (Table 2) was used to assess the performance.

Table 2. Confusion matrix for binary classification.

Predicted	FLS (Actual)	Non- FLS (Actual)
FLS	TP	FP
Non- FLS	FN	TN

Where are true positive (TP), true negative (TN), false positive (FP) and false negative (FN), respectively. The F_1 sums up the predictive performance of a model by combining

doi:10.6342/NTU202300432

two otherwise competing metrics, precision and recall for the test dataset (eqs. 2-4). In addition, the accuracy provides an easy-to-understand definition that assumes a binary classification problem used to tell the percentage of accurate predictions (eq. 5). The MCC (eq. 6) is in essence a correlation coefficient between the observed and predicted binary classifications (Chicco and Jurman 2020); it returns a value between −1 and +1. A coefficient of +1 represents a perfect prediction, 0 is no better than random prediction and −1 indicates total disagreement between prediction and observation.

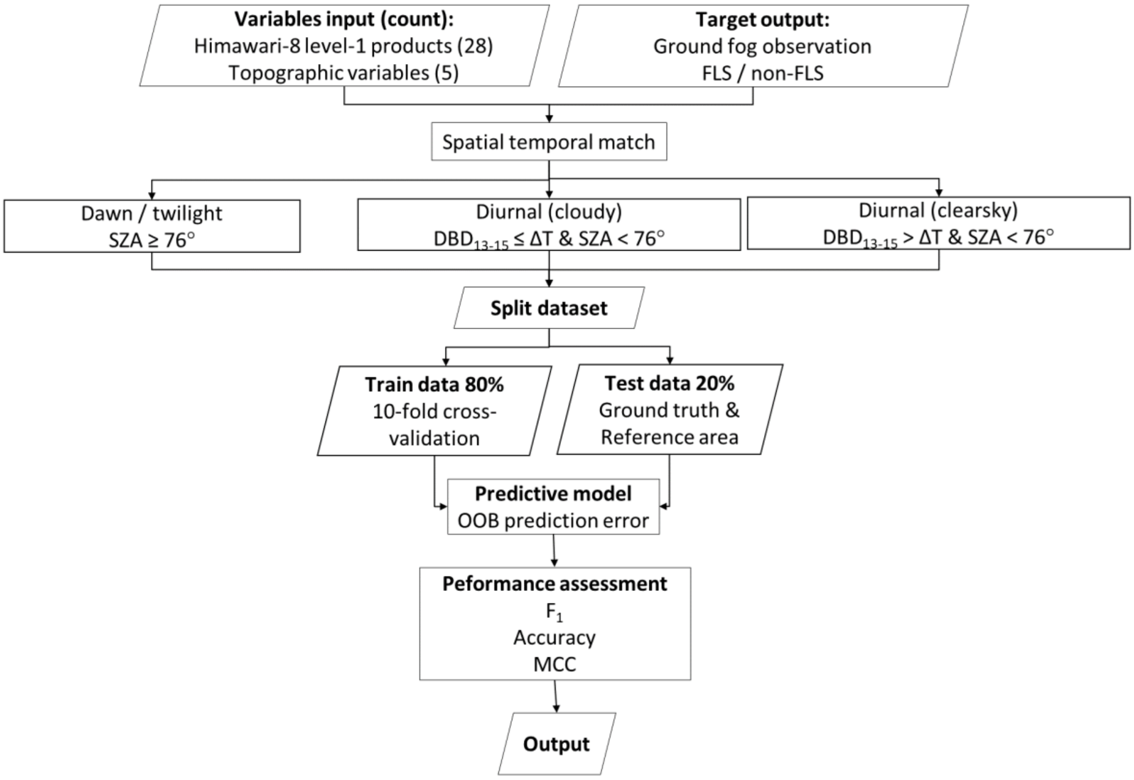

Fig. 2 The workflow for detecting diurnal FLS occurrence. Parallelograms indicate data that are available for input or output as well as representing resources used or generated and the rectangles are a process, action or function. The spatial-temporal matching dataset follows the different flow according to the observation period (SZA greater than 76°) and the threshold value (DBD $_{13\text{-}15}$ greater or less than ΔT).

Chapter 3: Results

3.1 DATA PRE-PROCESSING

Diurnal FLS occurrence probability for the four observation stations was observed in the dataset as 34% (MingChih), 61% (S9K), 57% (S14.5K), and 55% (S30K) ($n = 53363$). First, Ranger applies SZA greater than 76° that is executed for the dawn/twilight model. The frequency distribution of DBD_{7-13} (Fig. 3a) is used to observe the variation of each angle when FLS occurs is affected by cloud emissivity. Second, the algorithm through DBD_{13-15} fixed threshold values (ΔT) as empirical and statistical analysis as 2.38 to distinguish potential clear sky and cloudy pixel (Fig. 3b). Then, these processing steps are performed for each individual satellite image and the threshold values are determined empirically. For these potential clear sky pixels (50% above the median), the calculated ΔT are checked and classified. All remaining pixels of 50% below the median are classified as potential cloudy pixels. The performance of the algorithm may depend on the characteristics of the input variables, mainly relying on the observation data of H-8 for three models of RF algorithms to detect FLS occurring in study site.

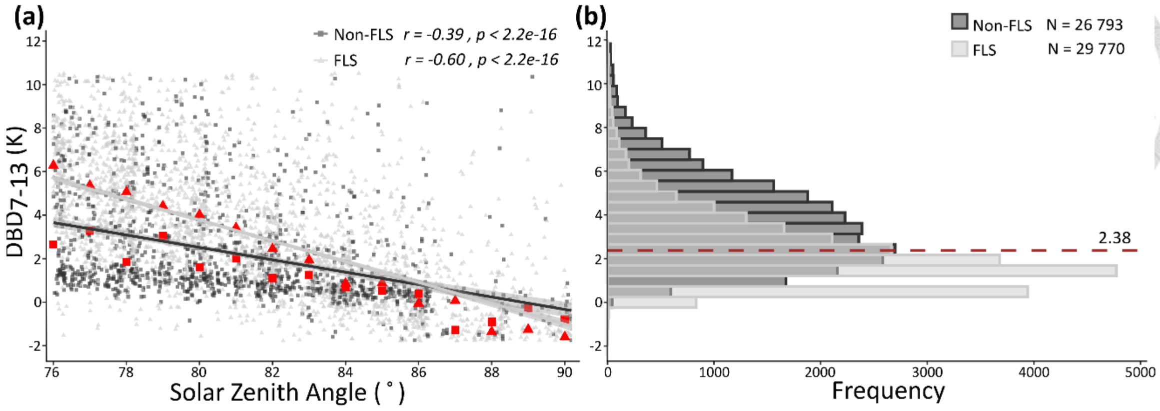

Fig. 3. (a) SZA dependence of DBD$_{7\text{-}13}$ between 76° and 90° with linear regression lines and mean point (red) in non-FLS and FLS condition. Solar radiation becomes weaker through during the FLS condition. (b) Histogram analysis of DBD$_{13\text{–}15}$ performed for FLS and non-FLS conditions and using DBD$_{13\text{–}15}$ threshold values as 2.38 to separate potential clear sky and cloudy pixel.

3.2 MODEL PERFORMANCE ASSESSMENT

The training (80% of the data, n = 3751, 22670, 22582 for dawn/twilight, cloudy, clear sky) and test (20%, 937, 5666, 5645) datasets were randomly split; out-of-bag prediction errors (brier scores) are 0.066 (dawn/twilight), 0.060 (cloudy), and 0.070 (clear-sky). The results reveal that models are of high performance (Table 3). This justifies the use of Ranger for diurnal FLS occurrence modelling, while the performance of model of cloudy was greater than dawn/twilight and clear sky.

Table 3. Model performance assessment of FLS occurrence detection for three models. TP = true positives; TN = true negatives; FP = false positives; FN = false negatives.

Model	**TP**	**TN**	**FP**	**FN**	**F_1**	**Accuracy**	**MCC**
Dawn/twilight	668	191	66	12	0.945	0.917	0.786
Cloudy	3902	1309	343	112	0.945	0.920	0.802
Clear sky	1678	3439	267	261	0.864	0.907	0.793

According to variable importance analysis of Ranger, the NDVI was the most important variable followed by the visible bands (B_1, B_2, B_3) (Fig. 4). Additionally, sun geometry (SAA, SZA) is also playing an important role in diurnal FLS observation.

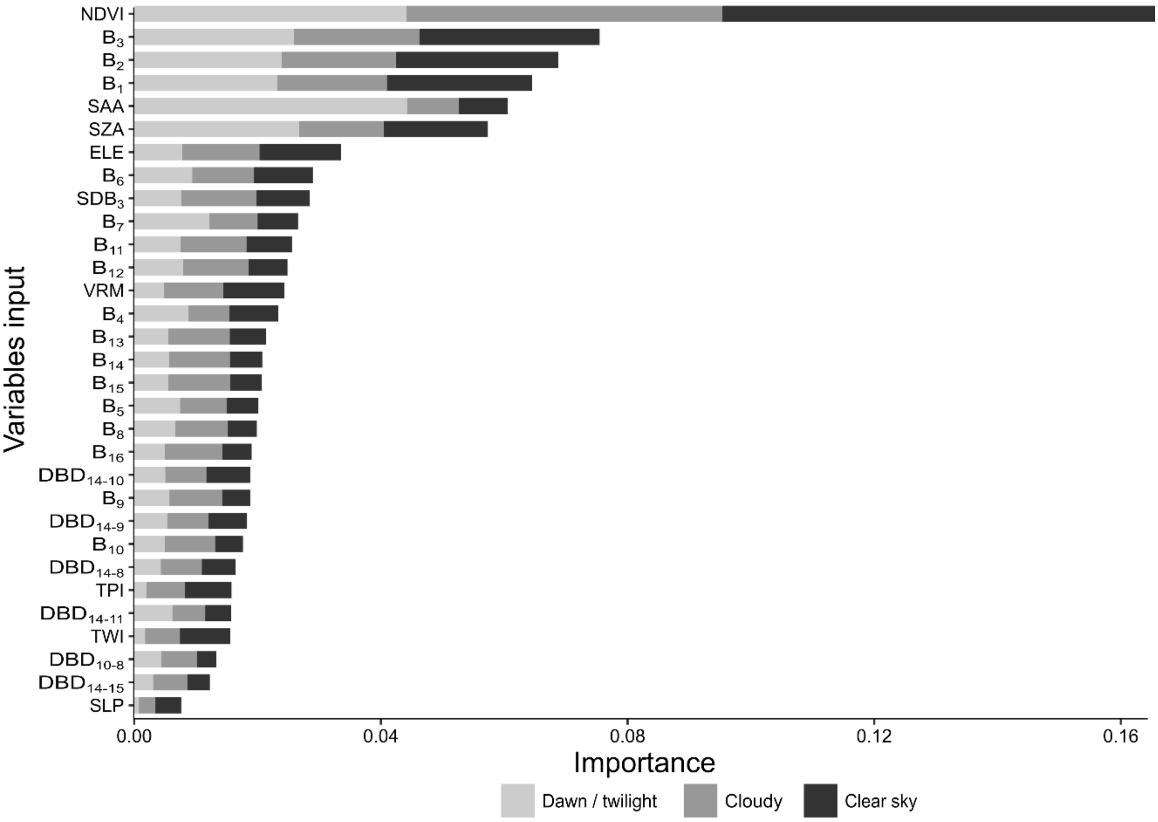

Fig. 4. The variable importance is calculated after the construction of ensemble trees for the rule-based models of the algorithm.

Furthermore, we analysed NDVI that were pivotal for our algorithm through each station and seasons (spring [March-May], summer [June-August], fall [September-November] and winter [December-February]) during the observation period. We conducted the maximum value compositing (MVC) of the NDVI to observe diurnal cycle of each station in non-FLS conditions（Fig. 5) (Holben 1986). As the assumption of MVC method that the NDVI value was reduced by presence of clouds, haze or snow in pixels,

doi:10.6342/NTU202300432

the maximum NDVI composite could be effective to obtain cloud-free information (Gutman and Ignatov 1996). The indices become lower in the afternoon in each station, especially the S30K stations fluctuates vary after 12pm. In addition, some time series NDVI value will not be available because the pixel with the SZA greater than 60° are excluded.

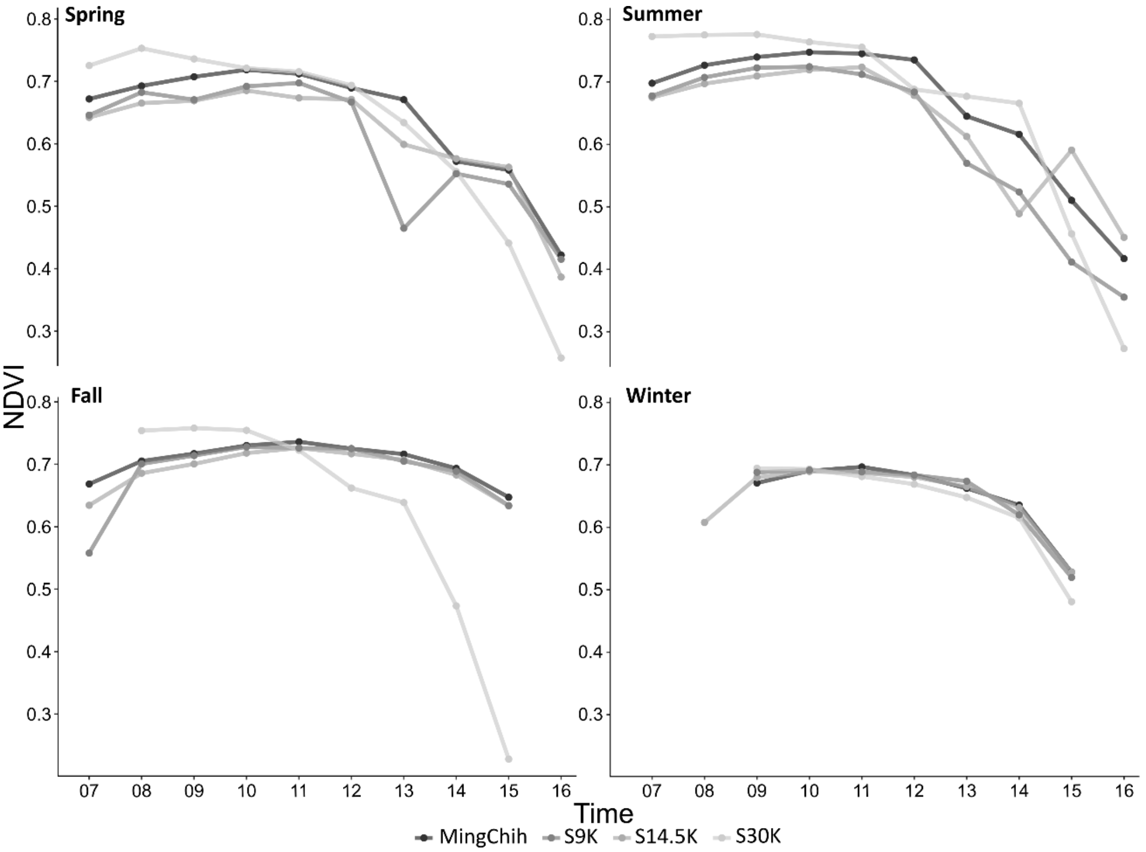

Fig. 5. The intra-diurnal cycle of each station in each season under the non-FLS occurrence conditions.

Non-FLS pixel detects the extensive range of B_{13} and DBD_{13-15} values while the B_{13} was denser in the non-FLS pixel of four seasons (Fig. 6a and b). The difference of DBD_{13-15} is assigned to distinguish between thick and thin clouds that the almost-zero values as thick clouds and positive values as the thin clouds (Akihiro 2020). The variation of the B_{13} between the first quartile was 277 K (269 K) and the third quartile was 290 K (281 K) under the non-FLS (FLS) condition during the observation period, while variation of the DBD_{13-15} show well-represented between the first quartile was 1.97 K (0.91 K) and the third quartile was 4.83 K (2.71 K) under the non-FLS (FLS) condition.

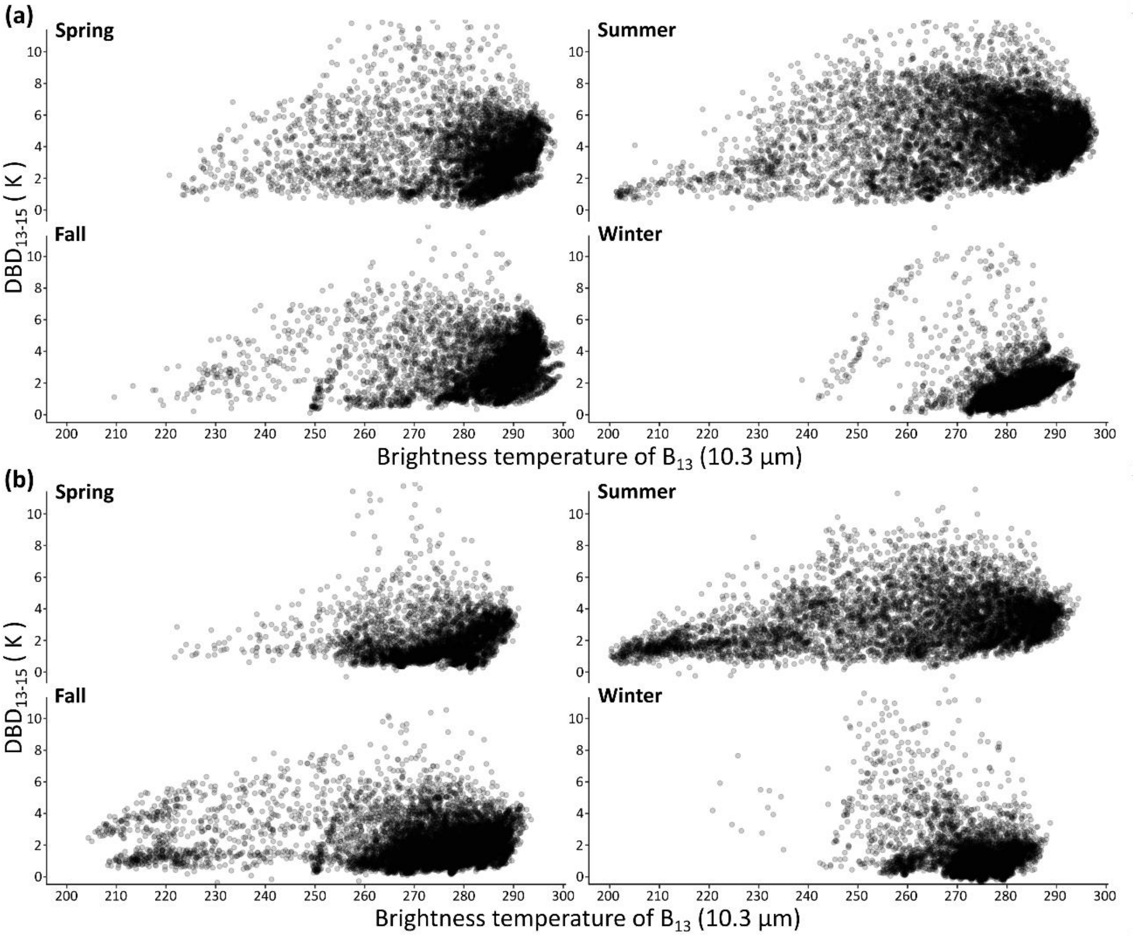

Fig. 6. Scatter plot of the (a) B_{13} and DBD_{13-15} values in non-FLS pixel and (b) FLS pixel for seasonal backgrounds.

3.3 SPATIOTEMPORAL DYNAMICS OF MCFS

The map of FLS probability and spatial variability predicted by the algorithm based on the H-8 observation data and local topographic variables is present in Fig. 7 (a and b). FLS occurrences commonly experience orographic fog to be a widespread phenomenon throughout study site, but the spatial distribution is heterogeneous. Depending on topography and wind conditions, the general increase toward the mountain’s valleys with spatially extensive high values reached and mountain ridges are characterised by considerably fewer FLS occurrences. Specifically, FLS accumulates against steep windward slopes and topography, resulting in hotspots of FLS immersion, while lower FLS probability in leeward side. There was significant seasonality in temperature, precipitation, FLS duration and amounts of radiation at the study sites. Across months, FLS probability is lowest in May, June and July, peaks in October, November and December (Fig. 7c).

We also generated spatial distribution of seasonality FLS probability for mornings (0700-1020 UTC+8), afternoons (1030-1350 UTC+8), and evenings (1400-1700 UTC+8). The magnitude of FLS probability is highest in the fall evening and lowest in the summer morning (Fig. 8). Overall, study sites are largely non-FLS in the mornings. In the evening, FLS can be seen to develop on the valley on the windward sides of the study sites.

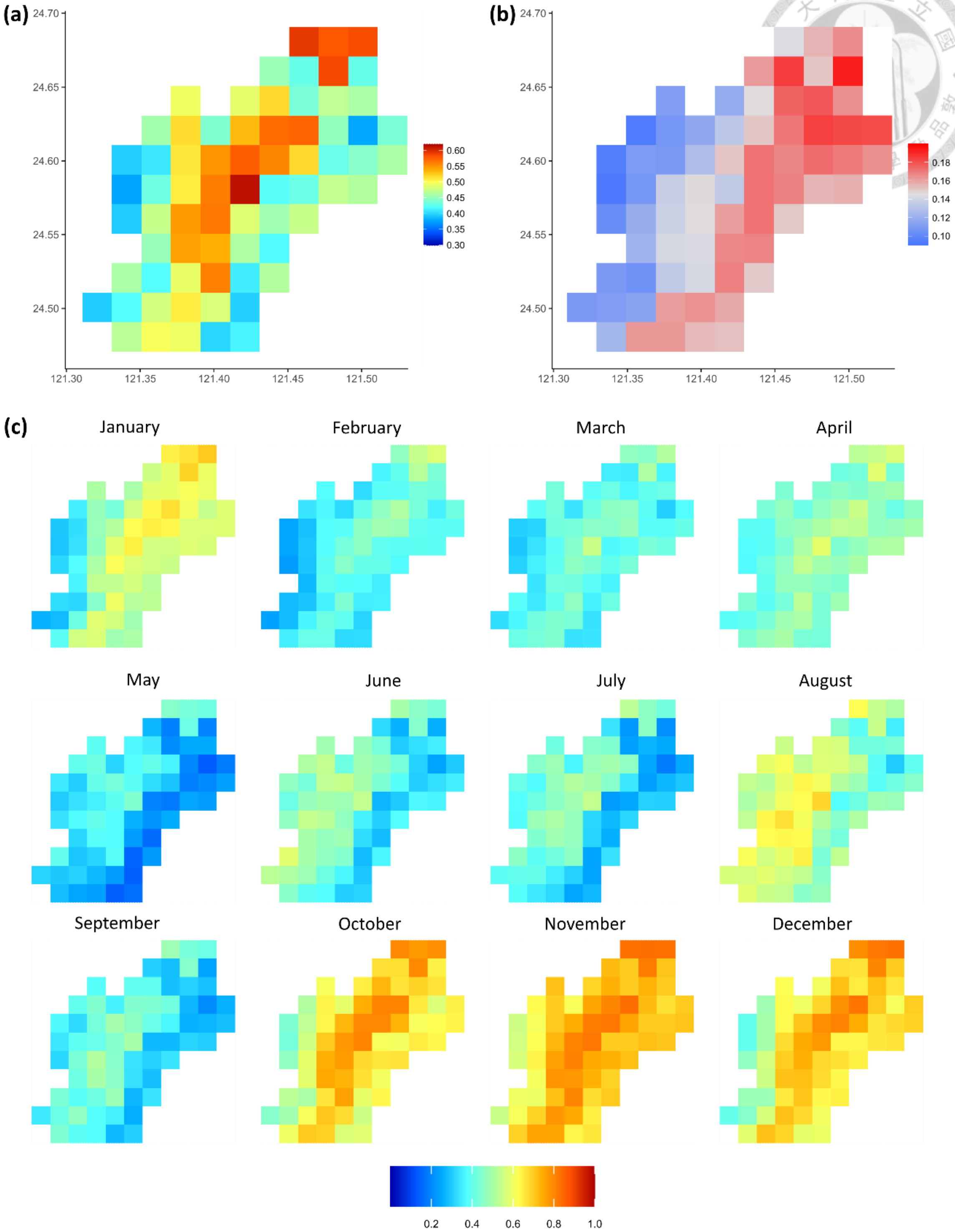

Fig. 7. (a) Mean FLS occurrence probability, (b) FLS occurrence temporal variability of each pixel, and (c) monthly FLS occurrence for Chi-Lan Mountain during the observation period (2018-2021).

doi:10.6342/NTU202300432

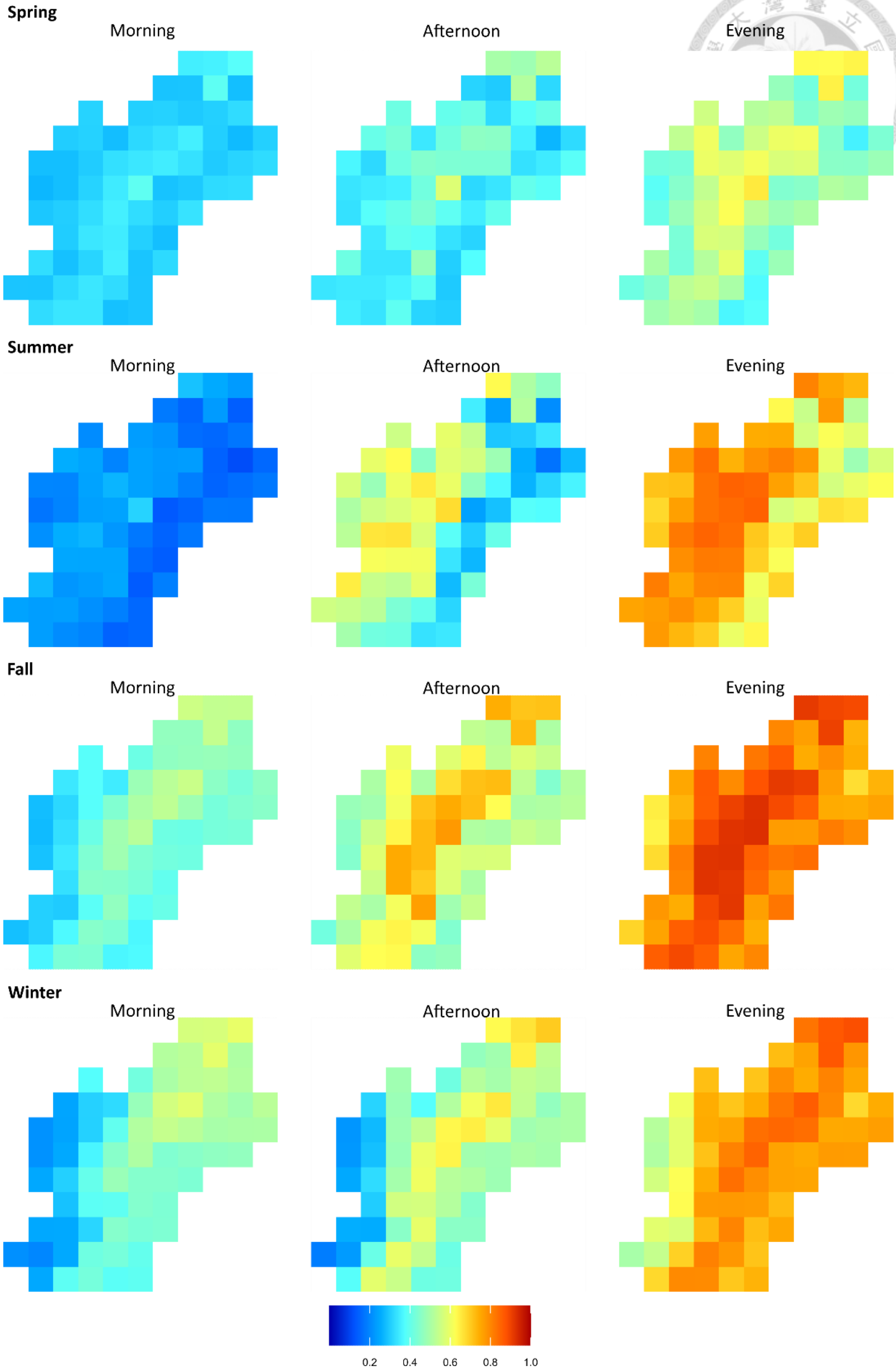

Fig. 8. Diurnal patterns of FLS probability shown for (left) mornings, (middle) afternoons, and (right) evenings of each season.

According to the temporal variability affected by the prevailing winds (Fig. 8), we analysed the FLS probability of windward and leeward sides of intra-diurnal cycle on the pixel located from 1000-2500 m a.s.l. Our results also show that the FLS probability is plotted against FLS temporal variability and then decreasing FLS variability with increasing FLS probability (Fig. 9). Overall, the average of FLS probability of fall is higher than other seasons (winter > summer > spring) during the observation period. The FLS associated with orographic lifting leading to cloud condensation usually occurs in the afternoon. Most of the FLS probability derived on the windward side was higher of a mountain, except summer affected on the leeward side.

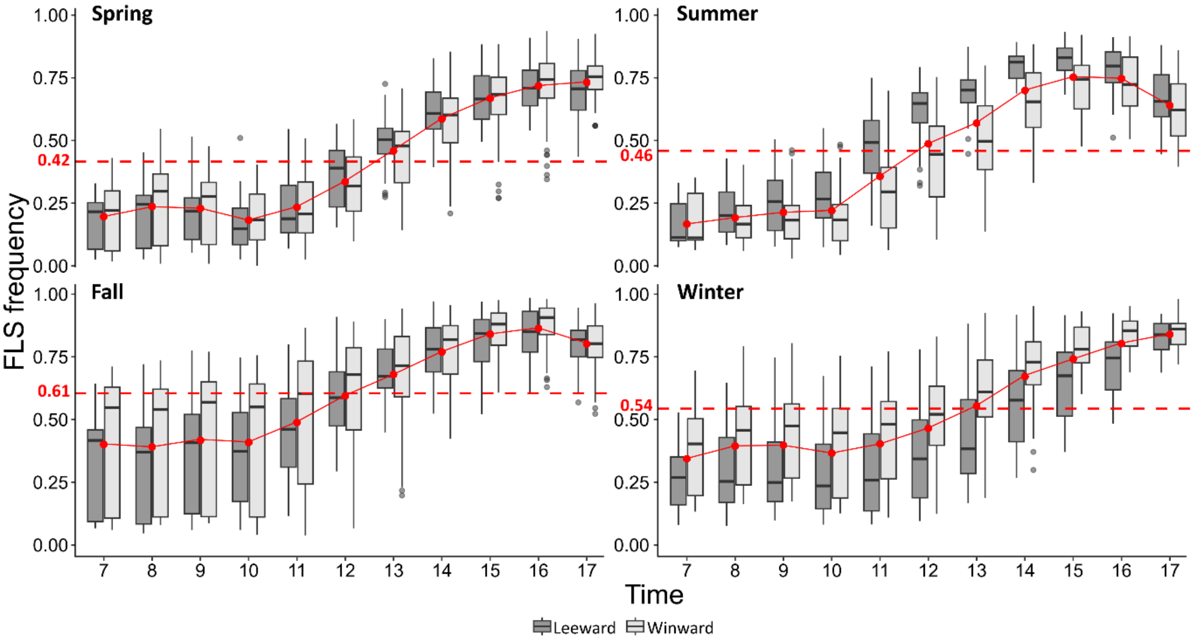

Fig. 9. Predictive results. Hourly diurnal FLS probability of leeward and windward of prevailing winds in four seasons. Red dash lines mean the average of which season and red lines and points mean the hourly average of diurnal FLS probability.

Chapter 4: Discussion

The aim of this study involved assessing the applicability of three satellited-based models to detect diurnal FLS occurrence of MCFs environment in subtropical. We developed a random forest type FLS detection algorithm using H-8 observations and ground fog observation data. In this chapter, we will discuss the potential application for our algorithm, its uncertain about decreasing the accuracy in each model. We also investigate the spatiotemporal pattern of FLS in Chi-Lan Mountains.

4.1 DIFFERENTIATION OF FLS OCCURRENCE DETECTION ALGORITHMS

The algorithm is composed of dawn/twilight, cloudy, clear sky models to detect the diurnal FLS occurrences. This algorithm follows the different flow (Fig. 2) according to the observation period due to the contamination problem by solar radiation during the dawn and twilight on B_7 centred at 3.9 µm (Ebell et al. 2013). There are many other features affecting the value of $DBD_{7\text{-}13}$, such as azimuth angle, surface types, and atmospheric profiles; the SZA is simply considered in this model. A more significant variation indicates the observed $DBD_{7\text{-}13}$ over FLS conditions compared to non- FLS conditions which can be considered very highly correlation of brightness temperature on FLS conditions ($r = -0.60$) is that the 10.4 µm contribution to it is greater than 3.9 µm due to their emissivity difference (Fig. 3a) (Yoo et al. 2006). Furthermore, the performance of machine learning highly depends on the importance variable of the training dataset. SAA is identified as the most contributing variable in this model due to the characteristics position and movement of solar was change over the seasons leading to significant changes in the solar illumination angles for our study site.

In general, many land surface temperature retrieval algorithms focus on the differential absorption in B_{13} and B_{15}. In this study, we identified the clear sky and cloudy pixel using a threshold of DBD_{13-15} because the accuracy and efficiency of land surface temperature retrieval algorithms are well known (Choi and Suh 2018; Yamamoto and Ishikawa 2018). Statistical analysis was used in our process to discern pattern and trend of FLS and non-FLS conditions. Apparently, preliminary application of DBD_{13-15} threshold technology to distinguish clear sky and cloudy pixels has a good classification performance (Fig. 3b).

In addition, for stable FLS detection in the long term, it is necessary to improve the level of FLS detection on complex topographic surface and for weak or local FLS. In particular, the quality of clear sky model and background data should be improved for the accurate separation of mainly FLS from variables of NDVI and visible bands. NDVI is identified as the most contributing variable in all models (Fig. 4). The volatility of time series data is relatively large, which is also suitable for trend analysis and information extraction in the time dimension. Substantial NDVI diurnal variations and negative NDVI hotspot effects were found due to differential red and NIR band sensitivities to diurnal phase angle changes, indicating that FLS tends to be rather bright in the red band and quite dark in the near-infrared band (Tran et al. 2020). Moreover, with the increase of wavelength, the reflectance of clouds decreases slowly while the reflectance of vegetation increases, so the pixels which are clouds or vegetation in the pixel can get preliminarily distinguished by indices (Xiong et al. 2020). Solar radiation is weak during dawn/twilight, resulting in a relatively high probability of scattering factor due to the low surface air temperature and high vapour saturation (Ma et al. 2022). Therefore, we observed the

doi:10.6342/NTU202300432

pattern of intra-diurnal cycle of the NDVI in each season (Fig. 5) except for SZA greater than 60°. However, the vast majority of changes in albedo values are caused by variations in clouds, it still found that in the process of collecting satellite remote sensing data. Of these, it is affected by multiple factors such as water vapour, aerosol, two-way reflection and data transmission, and abnormal vegetation indices will appear, which makes the changes of NDVI time series curves of various types appear irregular (Los et al. 1994).

The NDVI with high importance are drivers of the prediction and their values significantly impact the prediction values. We showed that Ranger variable importance measures are a sensible means for variables selection but are unreliable in situations where potential prediction vary in their measurement scale. Thus, we also analysed the cloud physical under FLS and non-FLS conditions through B_{13} and $DBD_{13\text{-}15}$ (Fig. 6). Obviously, both variable quantifies can vary significantly during the non-FLS condition. By contrast, most of the $DBD_{13\text{-}15}$ are less than 2.71 K under FLS condition.

4.2 SPATIAL AND TEMPORAL VARIATIONS IN FLS OCCURRENCE PROBABILITY

Our study demonstrates the average FLS occurrence probability at Chi-Lan Mountains as retrieved from H-8 satellite data and local topographic variables. It is apparent that the FLS occurs much more frequently than further inland (Fig. 7a). The frequency and fluctuation of FLS formation is significantly higher (reddish areas) on the southern and eastern slopes, especially at low elevations, compared to the western slopes of the study site (Fig. 7b). The probability of FLS mainly high (>0.5) at high elevations between 1500 and 2000 m.a.s.l. and the highest FLS occurrences probability is 0.6, while the very high elevations (>2000m.a.s.l.) are yield lower FLS probability compared to high elevations as less than 0.5. However, it can be observed that FLS frequency stability in

doi:10.6342/NTU202300432

bluish area both (Fig. 7b). As results, we verified that the rising temperatures may reduce the zone of cloud band in MCFs .

A causal relationship with the East Asia monsoon can be assumed since the spatial pattern is particularly pronounced during the month from October and February (Fig. 7c). One explanation for this pattern at eastern slopes is that at lower elevations, the valley is larger and deeper and oriented with their mouths in a windward direction, and therefore can persist FLS occurrences. By contrast, the valley of higher elevations is shallower and oriented in various directions. Thus, FLS passes over them, assisted by the higher wind speeds at elevation. The seasonal pattern of FLS occurrence can be explained by examining the wind direction as easterly or westerly. In summer, the FLS frequency is higher at inland windward slopes because of the southwest monsoon prevailed enhances the local convection. The FLS frequency was derived by eastward winds prevailed in other seasons (Lai et al. 2006).

Furthermore, the study site exhibits affected by the eastward winds prevailed from the morning to the afternoon due to the different thermal effects of the ridges and valleys during the daytime and moist air masses and frequently leading to FLS conditions (Klemm et al. 2006). There is usually more high variability of FLS occurrence at the early morning. In addition, several studies have demonstrated the peak of solar radiation in the cloud forest is in the early morning instead of noon in study site (Gu et al. 2021; Li et al. 2022), and thus the fluctuation as well as extent of FLS is lowest in the morning and highest in the evening(Fig. 8). This can be attributed to convergence of surface wind as it approaches the island and leads to the lifting of surface air to the condensation level, resulting in FLS formation.

doi:10.6342/NTU202300432

Seasonally, the fluctuation of FLS probability is more significant in Fall because the northeast monsoon was pronounced, and the disturbance with increaseed water vapour content. (Fig. 9). On the contrary, there is less water vapour in summer, mainly due to the water vapour brought from the coast by the convection of heterogeneous terrain, so the higher probability is formed in the afternoon. In addition, the variability of temporal distribution in spring is more stable than others, and FLS distribution of both sides is consistent.

As a detailed statistical analysis of the full life cycle of FLS and its interannual behaviour is not within the scope of this paper, this example is intended to illustrate the potential of the novel algorithm for the analysis of spatiotemporal patterns of FLS over land.

4.3 POTENTIAL APPLICATIONS

The novel algorithm is thereby relatively independent from exact DBD thresholds and thus has the potential to be easily applied to other regions and to generate climate datasets. An operational deployment is possible with slight adjustments in algorithm design and holds the potential for the prediction of FLS dissipation. In the future, the value of the derived FLS occurrence may be further enhanced with the retrieval of cloud-based altitudes for the separation of low-level stratus from the ground fog. Furthermore, this algorithm can quickly process large raw data from satellite observations data with a low compute cost, which may be conducive to potential operational applications.

The performance of predictive modelling is dependent on the amount and quality of available data. In this work, we utilise match-ups between H-8 and ground observation data for four years, from January 2018- December 2021, to reduce the effects of seasonal variation. We have some inevitable limitations on the way to observe ground FLS by

doi:10.6342/NTU202300432

time-lapse camera. For example, in the case of our spatial resolution of 2km, the occurrence of FLS events in front of the facing direction of the camera may be inconsistent with the events in front of the lens 1 km. Still, we will be considering high temporal resolution and the FLS occurrence events will be continuous; the misleading events would be well reduced by long-term observation time in order to obtain information about ground fog. In practice, the accuracy of fog observation by the human-judgment is high (90%) and 70.5% using a visibility meter (Egli et al. 2018). To avoid these inevitable factors, we put considerable effort into systematically evaluating the classification of three models. Considering these arguments, the validation results give confidence in the skill of the novel algorithm, which is well suited for the purpose of characterising the spatial and temporal pattern of FLS occurrence probability in MCFs.

Moreover, we explore potential application of the NDVI value in FLS and non-FLS condition in observation region of H-8. The land cover of the main observation region is more suitable for the NDVI value because the most of those areas may not be covered by snow or ice, especially in Southeast Asia where there are important areas of MFC sites. Otherwise, the high temporal resolution of H-8 observation data would be obtained and real-time processing by the algorithm. Then, the Ranger can process data effectively and retrieve the FLS occurrence.

doi:10.6342/NTU202300432

Chapter 5: Conclusions

The central aim of this study was to develop the first satellite-based and local topographic data and thereby diurnally stable satellite retrieval of FLS. The algorithm design uses an observation period (SZA >76°) and DBD threshold condition to model three different atmospheric conditions to retrieve FLS.

The algorithm was applied to detect spatial and temporal patterns of Chi-Lan Mountains FLS and was validated against the confusion matrix. The algorithm shows good overall detection accuracy, F_1 and MCC with imbalance data, which are reliable concordant scores for predictions. The NDVI was more important than other variable and provided an adaptive model for FLS occurrence detection as an indicator of importance for identification.

FLS most frequently occur close to the south-eastern facing ridge, with the cloud belt between 1500 and 2000 m. a.s.l., confirming findings from Chang (2008), Lai (2006) and Li (2022). The diurnal cycle of FLS is described for the location of the four stations. Marked differences in the timing of FLS occurrence and temporal persistence are found. The time lag of FLS occurrence from the low elevation to high elevation region may be attributed to the orographic of FLS from the lowland, a typical feature of the region. FLS persists a longer time in the mid-elevation than in further high altitudes regions but then dissipates more rapidly after evening, with stronger winds due to strong convection.

The study shows the potential of the diurnal FLS algorithm to study FLS patterns and life cycles. Future research efforts should focus on coherently mapping diurnal

doi:10.6342/NTU202300432

characteristics of FLS, further understanding FLS formation processes and potentially detecting changes in FLS occurrence. This may best be achieved by combining the satellite retrievals with numerical modelling and ground-based observation.

doi:10.6342/NTU202300432

References

Akihiro, S. (2020). Introduction to Himawari-8 RGB composite imagery. *Meteorological Satellite Center Technical Note 65*

Alewell, C., Manderscheid, B., Meesenburg, H., & Bittersohl, J. (2000). Is acidification still an ecological threat? *Nature, 407,* 856-857

Andersen, H., & Cermak, J. (2018). First fully-diurnal fog and low cloud satellite detection reveals life cycle in the Namib. *Atmospheric Measurement Techniques, 11,* 5461-5470

Ball, L., & Tzanopoulos, J. (2020). Interplay between topography, fog and vegetation in the central South Arabian mountains revealed using a novel Landsat fog detection technique. *Remote Sensing in Ecology and Conservation, 6*, 498-513

Bassiouni, M., Scholl, M.A., Torres-Sanchez, A.J., & Murphy, S.F. (2017). A method for quantifying cloud immersion in a tropical mountain forest using time-lapse photography. *Agricultural and Forest Meteorology, 243*, 100-112

Bax, V., Castro-Nunez, A., & Francesconi, W. (2021). Assessment of potential climate change impacts on montane forests in the Peruvian Andes: Implications for conservation prioritization. *Forests, 12,* 375

Beiderwieden, E., Schmidt, A., Hsia, Y.J., Chang, S.C., Wrzesinsky, T., & Klemm, O. (2007). Nutrient input through occult and wet deposition into a subtropical montane cloud forest. *Water, Air, and Soil Pollution, 186*, 273-288

Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., Kumagai, Y., Miyakawa, T., Murata, H., Ohno, T., Okuyama, A., Oyama, R., Sasaki, Y.,

doi:10.6342/NTU202300432

Shimazu, Y., Shimoji, K., Sumida, Y., Suzuki, M., Taniguchi, H., Tsuchiyama, H., Uesawa, D., Yokota, H., & Yoshida, R. (2016). An introduction to Himawari-8/9 − Japan’s new-generation geostationary mMeteorological satellites. *Journal of the Meteorological Society of Japan. Ser. II, 94*, 151-183

Beven, K.J., & Kirkby, M.J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin, 24,* 43-69

Bradley, R.S., Vuille, M., Diaz, H.F., & Vergara, W. (2006). Threats to water supplies in the tropical Andes. *Science, 312,* 1755-1756

Breiman, L. (2001). *Machine Learning, 45,* 5-32

Brier, G.W. (1950). Verification of forecasts expressed in terms of probability. *Monthly weather review, 78*, 1-3

Bruijnzeel, L.A., Mulligan, M., & Scatena, F.N. (2011). Hydrometeorology of tropical montane cloud forests: emerging patterns. *Hydrological Processes, 25*, 465-498

Bruijnzeel, L.A., & Veneklaas, E.J. (1998). Climatic conditions and tropical montane forest productivity: The fog has not lifted yet. *Ecology, 79,* 3-9

Bruijnzeel, L.A., Waterloo, M.J., Proctor, J., Kuiters, A.T., & Kotterink, B. (1993). Hydrological observations in montane rain forests on Gunung Silam, Sabah, Malaysia, with special reference to the `Massenerhebung' Effect. *The Journal of Ecology, 81*, 145

Cermak, J., & Bendix, J. (2008). A novel approach to fog/low stratus detection using Meteosat 8 data. *Atmospheric Research, 87*, 279-292

Chang, S.-C., Lai, I.L., & Wu, J.-T. (2002). Estimation of fog deposition on epiphytic bryophytes in a subtropical montane forest ecosystem in northeastern Taiwan. *Atmospheric Research, 64*, 159-167

doi:10.6342/NTU202300432

Chang, S.-C., Tseng, K.-H., Hsia, Y.-J., Wang, C.-P., & Wu, J.-T. (2008). Soil respiration in a subtropical montane cloud forest in Taiwan. *Agricultural and Forest Meteorology, 148*, 788-798

Chang, S.-C., Yeh, C.-F., Wu, M.-J., Hsia, Y.-J., & Wu, J.-T. (2006). Quantifying fog water deposition by in situ exposure experiments in a mountainous coniferous forest in Taiwan. *Forest Ecology and Management, 224*, 11-18

Chen, X., Yang, B., & Lin, Z. (2018). A random forest learning assisted "divide and conquer" approach for peptide conformation search. *Scientific Reports, 8*

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics, 21*, 6

Chinchor, N., & Sundheim, B.M. (1993). MUC-5 evaluation metrics. In, *Fifth Message Understanding Conference (MUC-5). Baltimore, MD, USA.*

Choi, Y.-Y., & Suh, M.-S. (2018). Development of Himawari-8/Advanced Himawari Imager (AHI) land surface temperature retrieval algorithm. *Remote Sensing, 10,* 2013

Chu, H.-S., Chang, S.-C., Klemm, O., Lai, C.-W., Lin, Y.-Z., Wu, C.-C., Lin, J.-Y., Jiang, J.-Y., Chen, J., Gottgens, J.F., & Hsia, Y.-J. (2014). Does canopy wetness matter? Evapotranspiration from a subtropical montane cloud forest in Taiwan. *Hydrological Processes, 28,* 1190-1214

Conley, D.J., Paerl, H.W., Howarth, R.W., Boesch, D.F., Seitzinger, S.P., Havens, K.E., Lancelot, C., & Likens, G.E. (2009). Controlling Eutrophication: Nitrogen and Phosphorus. *Science, 323,* 1014-1015

Dawson, T.E., & Goldsmith, G.R. (2018). The value of wet leaves. *New Phytologist, 219*, 1156-1169

doi:10.6342/NTU202300432

Delgado, R., & Tibau, X.-A. (2019). Why Cohen’s Kappa should be avoided as performance measure in classification. *PLoS One, 14,* e0222916

Ebell, K., Orlandi, E., Hünerbein, A., Löhnert, U., & Crewell, S. (2013). Combining ground-based with satellite-based measurements in the atmospheric state retrieval: Assessment of the information content. *Journal of Geophysical Research: Atmospheres, 118,* 6940-6956

Egli, S., Thies, B., & Bendix, J. (2018). A hybrid approach for fog retrieval based on a combination of satellite and ground truth data. *Remote Sensing, 10,* 628

Egli, S., Thies, B., Drönner, J., Cermak, J., & Bendix, J. (2017). A 10 year fog and low stratus climatology for Europe based on Meteosat Second Generation data. *Quarterly Journal of the Royal Meteorological Society, 143,* 530-541

Fensholt, R., Anyamba, A., Huber, S., Proud, S., Tucker, C., Small, J., Pak, E., Rasmussen, M., Sandholt, I., & Shisanya, C. (2011). Analysing the advantages of high temporal resolution geostationary MSG SEVIRI data compared to Polar Operational Environmental Satellite data for land surface monitoring in Africa. *International Journal of Applied Earth Observation and Geoinformation, 13,* 721-729

Frahm, J.P., & Gradstein, S.R. (1991). An altitudinal zonation of tropical rain forests using byrophytes. *Journal of Biogeography, 18,* 669-678

Gaumet, J.L., Heinrich, J.C., Cluzeau, M., Pierrard, P., & Prieur, J. (1998). Cloud-base height measurements with a single-pulse erbium-glass laser ceilometer. *Journal of Atmospheric and Oceanic Technology, 15,* 37-45

Gentry, A.H. (1992). Tropical forest biodiversity - distributional patterns and their conservational significance. *Oikos, 63,* 19-28

Glasow, R.v., & Bott, A. (1999). Interaction of radiation fog with tall vegetation. *Atmospheric Environment, 33*, 1333-1346

Gotsch, S., Davidson, K., Murray, J., Duarte, V., & Draguljić, D. (2017). Vapor pressure deficit predicts epiphyte abundance across an elevational gradient in a tropical montane region. *American Journal of Botany*, *104*, 1790-1801

Gotsch, S., Asbjornsen, H., Holwerda, F., Goldsmith, G.R., Weintraub, A.E., & Dawson, T.E. (2014). Foggy days and dry nights determine crown-level water balance in a seasonal tropical montane cloud forest. *Plant Cell and Environment, 37*, 261-272

Grubb, P.J. (1977). Control of forest growth and distribution on wet tropical mountains: with special reference to mineral nutrition. *Annual Review of Ecology and Systematics, 8*, 83-107

Gu, R.-Y., Lo, M.-H., Liao, C.-Y., Jang, Y.-S., Juang, J.-Y., Huang, C.-Y., Chang, S.-C., Hsieh, C.-I., Chen, Y.-Y., Chu, H., & Chang, K.-Y. (2021). Early peak of latent heat fluxes regulates diurnal temperature range in montane cloud forests. *Journal of Hydrometeorology*

Guisan, A., Weiss, S., & Weiss, A. (1999). GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology, 143*, 107-122

Güls, I., & Bendix, J. (1996). Fog detection and fog mapping using low cost Meteosat-WEFAX transmission. *Meteorological Applications, 3*, 179-187

Gultepe, I., Tardif, R., Michaelides, S.C., Cermak, J., Bott, A., Bendix, J., Müller, M.D., Pagowski, M., Hansen, B., Ellrod, G., Jacobs, W., Toth, G., & Cober, S.G. (2007). Fog research: A review of past achievements and future perspectives. *Pure and Applied Geophysics, 164*, 1121-1159

doi:10.6342/NTU202300432

Gutman, G., & Ignatov, A. (1996). The relative merit of cloud/clear identification in the NOAA/NASA Pathfinder AVHRR Land 10-day composites. *International Journal of Remote Sensing, 17,* 3295-3304

Halladay, K., Malhi, Y., & New, M. (2012). Cloud frequency climatology at the Andes/Amazon transition: 2. Trends and variability. *Journal of Geophysical Research: Atmospheres, 117* , D23103

Hamilton, L.S. (1995). Mountain cloud forest conservation and research: A yynopsis. *Mountain Research and Development, 15,* 259-266

Hobson, R. (2019). Surface roughness in topography: quantitative approach. *Spatial analysis in geomorphology* (pp. 221-246): Routledge

Holben, B.N. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing, 7*, 1417-1434

Hu, K.-T., & Huang, C.-y. (2019). A metabolic scaling theory-driven remote sensing approach to map spatiotemporal dynamics of litterfall in a tropical montane cloud forest. *International Journal of Applied Earth Observation and Geoinformation, 82*, 101896

Iwabuchi, H., Putri, N.S., Saito, M., Tokoro, Y., Sekiguchi, M., Yang, P., & Baum, B.A. (2018). Cloud property retrieval from multiband infrared measurements by Himawari-8. *Journal of the Meteorological Society of Japan. Ser. II, 96B,* 27-42

Kaseke, K.F., Tian, C., Wang, L., Seely, M., Vogt, R., Wassenaar, T., & Mushi, R. (2018). Fog spatial distributions over the Central Namib Desert - An isotope approach. *Aerosol and Air Quality Research, 18,* 49-61

Klemm, O., Chang, S.-C., & Hsia, Y.-J. (2006). Energy fluxes at a subtropical mountain cloud forest. *Forest Ecology and Management, 224,* 5-10

doi:10.6342/NTU202300432

LaBastille, A., & Pool, D.J. (1978). On the need for a system of cloud-forest parks in Middle America and the Caribbean. *Environmental Conservation, 5*, 183-190

Lai, G.-Y., Liu, H.-C., Chung, C.-H., Wang, C.-K., & Huang, C.-y. (2021). Lidar-derived environmental drivers of epiphytic bryophyte biomass in tropical montane cloud forests. *Remote Sensing of Environment, 253*, 112166

Lai, G.-Y., Liu, H.-C., Kuo, A.J., & Huang, C.-y. (2020). Epiphytic bryophyte biomass estimation on tree trunks and upscaling in tropical montane cloud forests. *PeerJ, 8*, *e9351*

Lai, I., Chang, S., Lin, P., Chou, C., & Wu, J. (2006). Climatic characteristics of the subtropical mountainous cloud forest at the Yuanyang Lake long-term ecological research site, Taiwan. *Taiwania, 51*, 317

Lawton, R.O., Nair, U.S., Pielke, R.A., & Welch, R.M. (2001). Climatic impact of tropical lowland deforestation on nearby montane cloud forests. *Science, 294*, 584-587

Lee, J.-R., Chung, C.-Y., & Ou, M.-L. (2011). Fog detection using geostationary satellite data: Temporally continuous algorithm. *Asia-Pacific Journal of Atmospheric Sciences, 47*, 113-122

Lee, Y., Han, D., Ahn, M.-H., Im, J., & Lee, S.J. (2019). Retrieval of total precipitable water from Himawari-8 AHI data: A comparison of random forest, extreme gradient boosting, and deep neural network. *Remote Sensing, 11*, 1741

Li, C.-F., Chytrý, M., Zelený, D., Chen, M.-Y., Chen, T.-Y., Chiou, C.-R., Hsia, Y.-J., Liu, H.-Y., Yang, S.-Z., Yeh, C.-L., Wang, J.-C., Yu, C.-F., Lai, Y.-J., Chao, W.-C., Hsieh, C.-F., & Bruelheide, H. (2013). Classification of Taiwan forest vegetation. *Applied Vegetation Science, 16*, 698-719

doi:10.6342/NTU202300432

Li, H.-J., Lo, M.-H., Juang, J.-Y., Wang, J., & Huang, C.-y. (2022). Assessment of spatiotemporal dynamics of diurnal fog occurrence in subtropical montane cloud forests. *Agricultural and Forest Meteorology, 317,* 108899

Liu, C., Yang, S., Di, D., Yang, Y., Zhou, C., Hu, X., & Sohn, B.-J. (2021). A machine learning-based cloud detection algorithm for the Himawari-8 spectral image. *Advances in Atmospheric Sciences, 39,* 1994-2007

Los, S.O., Justice, C.O., & Tucker, C.J. (1994). A global 1° by 1° NDVI data set for climate studies derived from the GIMMS continental NDVI data. *International Journal of Remote Sensing, 15,* 3493-3518

Ma, H., Li, Y., Wu, X., Feng, H., Ran, Y., Jiang, B., & Wang, W. (2022). A large-region fog detection algorithm at dawn and dusk for high-frequency Himawari-8 satellite data. *International Journal of Remote Sensing, 43*, 2599-2616

Matthews, B.W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochimica et Biophysica Acta (BBA) - *Protein Structure, 405*, 442-451

Nadkarni, N.M., & Solano, R. (2002). Potential effects of climate change on canopy communities in a tropical cloud forest: an experimental approach. *Oecologia, 131*, 580-586

Purbantoro, B., Aminuddin, J., Manago, N., Toyoshima, K., Lagrosas, N., Sri Sumantyo, J., & Kuze, H. (2018). Comparison of cloud type classification with split window algorithm based on different infrared band combinations of Himawari-8 satellite. *Advances in Remote Sensing, 07,* 218-234

Qiao, N., Zhang, L., Huang, C., Jiao, W., Maggs-Kölling, G., Marais, E., & Wang, L. (2020). Satellite observed positive impacts of fog on vegetation. *Geophysical Research Letters, 47*, e2020GL088428

doi:10.6342/NTU202300432

Ray, D.K., Nair, U.S., Lawton, R.O., Welch, R.M., & Pielke, R.A. (2006). Impact of land use on Costa Rican tropical montane cloud forests: Sensitivity of orographic cloud formation to deforestation in the plains. *Journal of Geophysical Research-Atmospheres*, *111*, D02108

Román-Cuesta, R.M., Carmona-Moreno, C., Lizcano, G., New, M., Silman, M., Knoke, T., Malhi, Y., Oliveras, I., Asbjornsen, H., & Vuille, M. (2014). Synchronous fire activity in the tropical high Andes: An indication of regional climate forcing. *Global Change Biology, 20,* 1929-1942

Schulz, H.M., Li, C.F., Thies, B., Chang, S.C., & Bendix, J. (2017). Mapping the montane cloud forest of Taiwan using 12 year MODIS-derived ground fog frequency data. *PLoS One, 12,* e0172663

Sjöström, M., Ardö, J., Eklundh, L., El-Tahir, B.A., El-Khidir, H.A.M., Hellström, M., Pilesjö, P., & Seaquist, J. (2009). Evaluation of satellite based indices for gross primary production estimates in a sparse savanna in the Sudan. *Biogeosciences, 6*, 129-138.

Spirig, R., Vogt, R., Larsen, J.A., Feigenwinter, C., Wicki, A., Franceschi, J., Parlow, E., Adler, B., Kalthoff, N., Cermak, J., Andersen, H., Fuchs, J., Bott, A., Hacker, M., Wagner, N., Maggs-Kölling, G., Wassenaar, T., & Seely, M. (2019). Probing the fog life cycles in the Namib Desert. *Bulletin of the American Meteorological Society, 100*, 2491-2507

Stackpole, J. D. (1967). Numerical analysis of atmospheric soundings. *Journal of Applied Meteorology,　6*, 464-467

Suh, M.-S., Lee, S.-J., Kim, S.-H., Han, J.-H., & Seo, E.-K. (2017). Development of land fog detection algorithm based on the optical and textural properties of fog using COMS data. *Korean Journal of Remote Sensing, 33*, 359-375

doi:10.6342/NTU202300432

Takano, T., Yamaguchi, J., Abe, H., Futaba, K.-I., Yokote, S.-I., Kawamura, Y., Takamura, T., Kumagai, H., Ohno, Y., Nakanishi, Y., & Nakajima, T. (2010). Development and performance of the millimeter-wave cloud profiling radar at 95 GHz: sensitivity and spatial resolution. *Electronics and Communications in Japan, 93*, 42-49

Tan, Z., Huo, J., Ma, S., Han, D., Wang, X., Hu, S., & Yan, W. (2021). Estimating cloud base height from Himawari-8 based on a random forest algorithm. *International Journal of Remote Sensing, 42*, 2485-2501

Tran, N.N., Huete, A., Nguyen, H., Grant, I., Miura, T., Ma, X., Lyapustin, A., Wang, Y., & Ebert, E. (2020). Seasonal comparisons of Himawari-8 AHI and MODIS vegetation indices over latitudinal Australian grassland sites. *Remote Sensing, 12*, 2494

Vautard, R., Yiou, P., & Van Oldenborgh, G.J. (2009). Decline of fog, mist and haze in Europe over the past 30 years. *Nature Geoscience*, *2*, 115-119

Wang, G., Wang, H., Zhuang, Y., Wu, Q., Chen, S., & Kang, H. (2021). Tropical overshooting cloud-top height retrieval from Himawari-8 imagery based on random forest model. *Atmosphere*, *12*, 173

Williams-Linera, G. (2003). Temporal and spatial phenological variation of understory shrubs in a tropical montane cloud forest. *Biotropica, 35,* 28-36

Williams, J.W., Jackson, S.T., & Kutzbach, J.E. (2007). Projected distributions of novel and disappearing climates by 2100 AD. *Proceedings of the National Academy of Sciences, 104,* 5738-5742

Woodall, C.W., Russell, M.B., Walters, B.F., D’Amato, A.W., Fraver, S., & Domke, G.M. (2015). Net carbon flux of dead wood in forests of the Eastern US. *Oecologia, 177,* 861-874

doi:10.6342/NTU202300432

Wright, M., & Ziegler, A. (2015). ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software, 77*

Xiong, Q., Wang, Y., Liu, D., Ye, S., Du, Z., Liu, W., Huang, J., Su, W., Zhu, D., Yao, X., & Zhang, X. (2020). A cloud detection approach based on hybrid multispectral features with dynamic thresholds for GF-1 remote sensing images. *Remote Sensing, 12,* 450

Yamamoto, Y., & Ishikawa, H. (2018). Thermal land surface emissivity for retrieving land surface temperature from Himawari-8. *Journal of the Meteorological Society of Japan. Ser. II, 96*

Yoo, J.-M., Jeong, M.-J., Hur, Y., & Shin, D.-B. (2010). Improved fog detection from satellite in the presence of clouds. *Asia-Pacific Journal of Atmospheric Sciences, 46,* 29-40

Yoo, J.-M., Jeong, M.-J., & Yun, M.-Y. (2006). Optical properties of fog from satellite observation (MODIS) and numerical stimulation. *Journal of the Korean Meteorological Society, 42(5),* 291-305