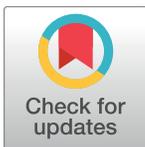RESEARCH ARTICLE

Mapping the montane cloud forest of Taiwan using 12 year MODIS-derived ground fog frequency data

Hans Martin Schulz[1]*, Ching-Feng Li[2], Boris Thies[1], Shih-Chieh Chang[3], Jörg Bendix[1]

1 Laboratory for Climatology and Remote Sensing, Philipps-Universität Marburg, Marburg, Germany,
2 School of Forestry and Resource Conservation, National Taiwan University, Taipei, Taiwan, 3 Department of Natural Resources and Environmental Studies, National Dong Hwa University, Hualien, Taiwan

* martin.schulz@staff.uni-marburg.de

Abstract

Up until now montane cloud forest (MCF) in Taiwan has only been mapped for selected areas of vegetation plots. This paper presents the first comprehensive map of MCF distribution for the entire island. For its creation, a Random Forest model was trained with vegetation plots from the National Vegetation Database of Taiwan that were classified as "MCF" or "non-MCF". This model predicted the distribution of MCF from a raster data set of parameters derived from a digital elevation model (DEM), Landsat channels and texture measures derived from them as well as ground fog frequency data derived from the Moderate Resolution Imaging Spectroradiometer. While the DEM parameters and Landsat data predicted much of the cloud forest's location, local deviations in the altitudinal distribution of MCF linked to the monsoonal influence as well as the Massenerhebung effect (causing MCF in atypically low altitudes) were only captured once fog frequency data was included. Therefore, our study suggests that ground fog data are most useful for accurately mapping MCF.

Introduction

Montane cloud forest (MCF) is characterized by significant precipitation input from the canopy interception of frequently or persistently occurring foggy conditions at ground (tree) level [1]. On a regional scale, most endemic species can be found in MCF areas due to unique hydrological processes and how they interact with the biological communities [2, 3]. At the same time, MCF has been recognized as one of the world's most endangered ecosystems [4, 5]. Climate observations from the past decade show that MCF areas suffer from a decreasing trend in ground fog occurrence that is likely related to global warming [6–11]. Further, the habitats of this ecosystem are heavily impacted by anthropogenic influences such as deforestation for timber harvesting and agriculture [5, 12–15]. Hence, worldwide efforts to reliably map the actual distribution of MCF are urgently required for the purpose of natural resources management.

Subtropical MCF in Taiwan has never been comprehensively mapped. It is, for example, absent from a map of global MCF [5] distributed by the United Nations Environment

Programme (UNEP). Today, the National Vegetation Database of Taiwan provides the most reliable data about its extent. However, the data are spatially restricted to vegetation plots distributed across the island (cf. Sect. *Training data for the MCF conditions map*). Therefore, our study aims to map the entire distribution of Taiwanese MCF.

Simple MCF mapping approaches (e.g., [16] or the UNEP approach) use altitude as the predicting variable for MCF occurrence as it is a proxy for climatic (temperature, rainfall, and particularly ground fog frequency) and edaphic (soil water status and acidity) factors that are associated with the occurrence of MCF [17]. Altitude is not a perfect proxy, however. One important reason for this in Taiwan is the influence of the East Asian Monsoon which creates local deviations in the altitudinal distribution of ground fog occurrence. The Massenerhebung effect also seems to have similar impacts. Both have been considered as an explanation for the occurrence of MCF at atypically low altitudes (cf. Sect. Study area). An approach used by Mulligan and Burke [18] to map MCF in the global tropics could not capture these effects in the tropical south of Taiwan. The authors mapped MCF based on ground fog frequency modelled from reanalysis data [19–21], the resolution of which is too low to reproduce the influence either of the monsoon or the Massenerhebung effect. To investigate the suitability of remote sensing methods for MCF mapping, Nair et al. [22] compared a ground fog frequency map derived from 1-km Moderate Resolution Imaging Spectroradiometer (MODIS) data to 13 points from the UNEP-WCMC cloud forest locations database [23] in Costa Rica, southern Nicaragua, and northern Panama. They were able to capture local leeward effects and found a connection between high frequency of ground fog and MCF occurrence. This suggests that local altitudinal deviations of ground fog occurrence can be mapped using MODIS data. Thies et al. [24] and Wilson and Jetz [25] examined the relationship between MODIS-based 1-km cloud climatologies and MCF occurrence for Taiwan and for the global tropics (including the tropical south of Taiwan), respectively. While Thies et al. could not show a relationship between cloud frequency and MCF occurrence for most types of MCF, Wilson and Jetz created a global map of tropical MCF incorporating cloud frequency, inter-annual cloud variability, intra-annual cloud variability and elevation. Both studies mention that their approaches fail to separate clouds with and without ground contact, as MODIS cloud products do not include the height of the cloud base. Therefore, we present the first map of Taiwanese MCF distribution that incorporates remote sensing-based ground fog frequency data.

Study area

Taiwan (21°85′—25°30′N, 120°00′—122°00′E, $\sim$ 36,000 km^2) is a subtropical island in East Asia with a maximum elevation of 3952 m a.s.l. and more than 200 peaks above 3000 m a.s.l. Climatic variation in Taiwan is mostly correlated to altitude and monsoonal influence.

Monthly mean temperatures range from -1.1˚C in the Central Mountain Range (CMR) in January (measured at an altitude of around 3850 m a.s.l. by Yushan Weather Station; YSW in Fig 1) to about 29˚C in the lowlands in July. Annual mean temperatures as low as 4.2˚C can be observed in the mountains. In the lowlands, the annual mean temperatures are about 17˚C in the north and 20˚C in the south [26].

The summer monsoon between May and August approaches the island from the southwest, bringing, together with 3—4 typhoons per year, heavy rains to the entire island. The winter monsoon brings mild rainfall from the northeast direction to the windward areas of the island between September and April. Annual precipitation in the eastern lowlands exceeds 2000 mm, while the western part, which is less affected by the winter monsoon, receives less than 2000 mm. Most Taiwanese weather stations at altitudes higher than 1500 m a.s.l. measure annual precipitations greater than 2500 mm [28].
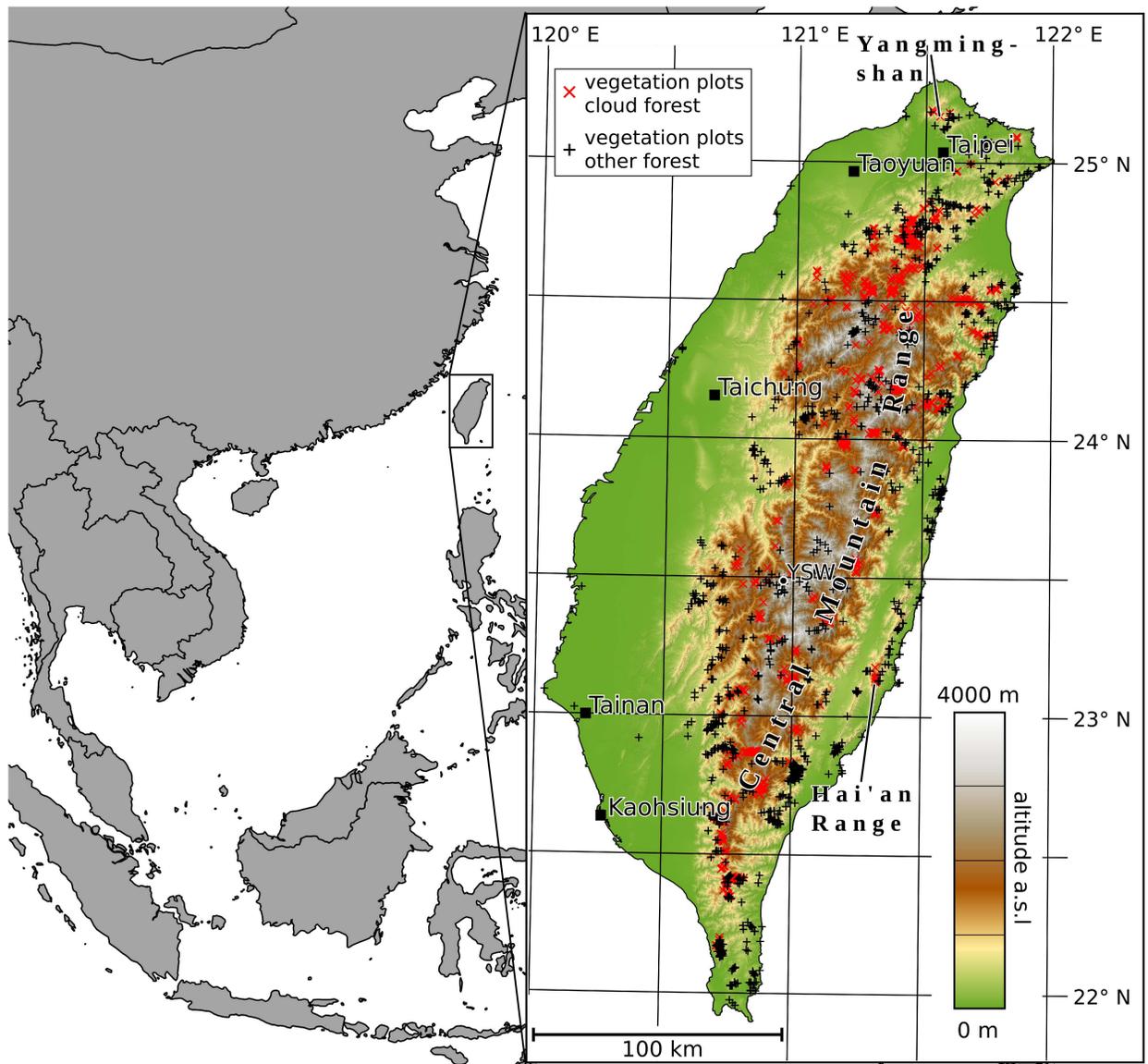
Fig 1. Topography and geographical location of Taiwan. The topography was derived from the ASTER GDEM 2 digital elevation model (cf. Sect. *Digital elevation model and related inputs*). The depicted vegetation plots are described in Sect. *Training data for the MCF conditions map*. Country borders were taken from OpenStreetMap [27].

doi:10.1371/journal.pone.0172663.g001

Forests in Taiwan can be classified into five vegetative zones based on local climate, which is primarily driven by altitude. At altitudes of ∼ 0—500 m a.s.l. and ∼ 500—1500 m a.s.l., respectively, the foothill zone and sub-montane zone are dominated by evergreen broad-leaved trees. At altitudes of ∼ 1500—2500 m a.s.l. the montane zone is characterized by frequent ground fog occurrence. It is dominated by either a mix of deciduous broad-leaved, coniferous, and evergreen broad-leaved forests or purely evergreen broad-leaved forests, both with characteristic features of MCF. Coniferous forests grow in the high-montane zone and subalpine zone at altitudes of ∼ 2500—3300 m a.s.l. and ∼ 3300—3700 m a.s.l., respectively. These zonal forests can be further classified into subtropical and tropical types. Northern Taiwanese flora belongs to the Holarctic Kingdom, while flora of the Paleotropical Kingdom can

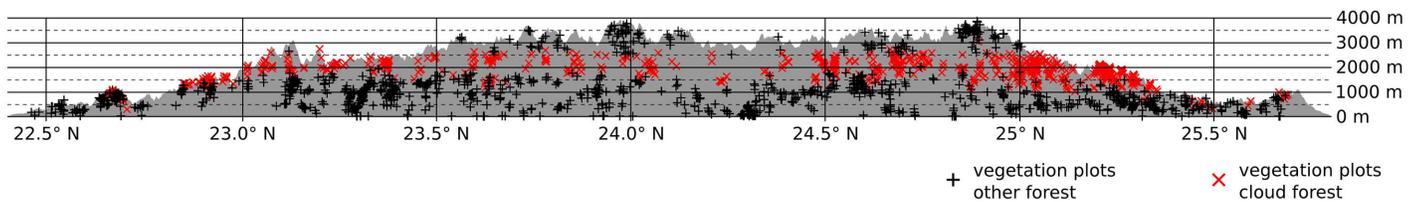**Fig 2. Altitudinal occurrence of MCF and non-MCF vegetation plots.** The depicted vegetation plots are described in Sect. *Training data for the MCF conditions map*.

be found in the south. Azonal forests correspond less to the climatic influence (mostly related to altitude) than the zonal forests mentioned above. The distribution of azonal forests is related to non-climatic factors such as fire regime, landslides, human disturbance, and rocky outcrops [29].

As the plots from the National Vegetation Database of Taiwan show, large areas of MCF can be found outside of the 1500 to 2500 m a.s.l. range. The distribution reaches altitudes as low as 1000 m a.s.l. in the northern and southern parts of the CMR. Isolated occurrences (particularly in the north) can even be found below 500 m a.s.l. (cf. Fig 2). Two factors can be assumed to be responsible for this distribution:

- The winter monsoon brings heavy fog to the windward parts of the island [3, 30]. This was linked to MCF occurrence at atypically low altitudes in northeastern Taiwan (cf. Figs 1 and 2)[3, 29].

- In isolated mountain terrain—especially if it is located near the ocean—the Massenerhebung effect causes a contraction/lowering of vegetation zones. This is particularly the case for MCF as condensation levels (and thus the altitude of frequent cloud immersion of the terrain) are shifted downwards. This is explained by a steep temperature lapse rate of isolated mountain terrain and a high water vapor concentration in the atmosphere around isolated coastal mountains [6, 31–33]. As the mountainous terrain is more isolated to the north and south of the CMR than in the center of Taiwan, the Massenerhebung effect is considered an explanation for lowered altitudes of the vegetation zones here [34].

Data and methods

Overall methodology

The map of Taiwanese MCF presented in this study was created by the intersection of two binary maps:

- A map of MCF conditions.

- A forest map (Sect. *Creation of the forest map*).

The usage of a machine learning approach allows additional inputs (besides ground fog frequency) to be incorporated. A Random Forest classifier [35] (implemented in the R [36] package "randomForest" [37]) was used to create the map of MCF conditions. Several studies have already shown the effectiveness of this ensemble machine learning method in vegetation mapping using remotely sensed data [38–40]. The Random Forest model was trained to distinguish the classes "MCF" and "non-MCF" using training samples extracted from different raster inputs with a resolution of 250 m per pixel. These inputs were monthly ground fog frequency maps (cf. Sect. *Ground fog frequency inputs*), different DEM parameters (Sect. *Digital elevation*

model and related inputs) as well as mosaics of the Landsat 7 solar bands (Sect. *Landsat channels*) and texture measures calculated from them (Sect. *Geostatistical texture features*). The positions from which training data were extracted were given by a point data set, which contains information about the presence or absence of MCF (cf. Sect. *Training data for the MCF conditions map*). After training, the number of input variables was reduced in order to avoid the "curse of dimensionality" [41] and increase the quality classification. This was done using the recursive feature elimination method implemented in the R package "caret" [42]. The reduced inputs were used to predict the occurrence of MCF conditions in the 250 m resolution of the input raster data.

Pixels determined to have MCF conditions are not necessarily covered by forest, since the data points used to create the map of MCF conditions do not distinguish between forested and non-forested areas. Since cloud forest can only exist where there is indeed forest, the map of MCF conditions was combined with the forest map.

Ground fog frequency inputs

Data from the Moderate-resolution Imaging Spectroradiometer (MODIS) onboard the near-polar orbiting satellites Terra (daytime Taiwan overflight $\sim$ 9:30—11:30 UTC + 8) and Aqua (daytime Taiwan overflight $\sim$ 12:30—14:30 UTC + 8) were used to create the ground fog frequency maps. Bands 1 and 2 of the instrument each have a resolution of 250 m per pixel, which produce a detailed image of ground fog and the shape of its extent as restricted by complex terrain. The distribution of MCF is highly dependent on altitude and, in mountainous areas, a coarse resolution implies a high altitudinal span covered by each pixel. Therefore, MODIS was chosen instead of the geostationary satellite Himawari 7, which covers Taiwan with a spatial resolution of up to 1000 m and a temporal sampling rate of 30 minutes. Although the newer Himawari 8 has a spatial resolution of up to 500 m, this data was disregarded as the satellite has only been in service since July 2015. This active service period is too short for the creation of meaningful ground fog frequency maps.

Common remote sensing-based ground fog detection schemes rely on assumptions regarding the microphysical properties of fog clouds, their vertical distribution within the cloud, and a certain relationship between the cloud top temperature and the cloud top height. As they are based on observations of radiation fog, which is common in temperate latitudes, those assumptions are not suited for fog in Taiwan, which is mostly orographic. Schulz et al. [43] therefore developed the algorithm *Detection of Ground Fog in Mountainous Areas* (DOGMA), which is custom tailored and validated for ground fog detection in the mountains of Taiwan. DOGMA detects fog in MODIS daytime scenes based on the DEM and the MODIS MOD 06 cloud optical thickness product (nighttime scenes can not be processed for the lack of a nighttime optical thickness input) using a statistical approach. The method sharpens the MODIS input data using the high-resolution MODIS bands 1 and 2 resulting in cloud masks with a resolution of 250 m.

DOGMA was used to create ground fog masks for every MODIS daytime scene covering Taiwan between 1 January 2003 and 31 December 2014. These fog masks were used to create ground fog frequency maps for the whole year as well as for each individual month, in order to capture the influence of the seasonality of ground fog occurrence.

Digital elevation model and related inputs

Other factors that are linked to the distribution of MCF (temperature, rainfall, soil water status and acidity) are not available as high-resolution raster inputs. Therefore, altitude was included in the machine learning approach as a proxy for those variables (cf. Sect. Introduction). It was

taken from the ASTER GDEM 2 (property of METI and NASA) distributed via the USGS global data explorer [44]. ASTER GDEM 2 originally has a resolution of 1 arcsecond ($\sim$ 30 m) and was resampled to the 250 m resolution of DOGMA for our machine learning approach.

Li et al. [3] successfully used an ordinal classification of the topography of Taiwan (1 = ridge, 2 = upper slope, 3 = middle slope, 4 = lower slope, 5 = valley, 6 = plain) as a predictor variable for the floristic composition of cloud forests in Taiwan. This classification system was used as a proxy for soil water availability and light input. Several vegetation plots were manually classified by Li et al., but this is not expedient for the creation of an input file for all of Taiwan. Therefore, several quantitative inputs that essentially contain the same information as the ordinal classification by Li et al. were calculated from the original 30 m version of the ASTER GDEM 2 and transferred to the 250 m resolution afterwards:

- The **sky view factor** is related to topographic shading [45]. Its values are high on ridges and low in valleys.

- The **slope** corresponds to the steepness of the terrain. It is low on valley bottoms and plains as well as on mountain tops. It is high on valley flanks.

- The **distance to the ridge** gives information about the segment of a slope in which a pixel is located. In order to calculate the distance to the ridge, terrain was divided into 4 main classes of its aspect using k-means clustering. For each entity of spatially connected pixels of the same class, the vertical distance to the highest pixel of that entity was calculated.

- The **distance to the valley bottom** is calculated in the same way as the distance to the ridge. Here, the vertical distance to the lowest pixel of an entity is calculated instead of the distance to the highest pixel.

Landsat channels

Landsat 7 Enhanced Thematic Mapper Plus (ETM+) visible and shortwave infrared bands 1, 2, 3, 4, 5 and 7 were used in the machine learning approach to account for spectral characteristics of MCF. Mosaics covering the whole area of Taiwan were compiled from 25 relatively cloudless ETM+ scenes captured between 1999 and 2003 (newer imagery has data gaps due to the failure of the ETM+ scan line corrector in May 2003).

Effects of atmospheric absorption, reflection, and scattering were removed from the Landsat 7 bands using the Second Simulation of a Satellite Signal in the Solar Spectrum code (6S code [46]) modified as described in [47]. The modified version is able to cope with a wide range of terrain altitudes in a scene. The Landsat bands were topographically corrected using a modified version of an approach suggested by Teillet et al. [48]. We masked out clouds as described in [49]. Terrain shadows were masked out by calculating their position based on the ASTER GDEM 2 and the sun azimuth and zenith.

After the atmospheric and topographic correction, bands 1, 2, 3, 4, 5 and 7 of all Landsat scenes were each stitched together in a cloud and shadow free mosaic. The mosaics are publicly available under doi:10.5678/LCRS/DAT.283. For the usage in the machine learning approach, the mosaics were resampled to the 250 m resolution of the other machine learning inputs. Texture measures were calculated from the mosaics of each band in the original 30 m Landsat resolution.

Geostatistical texture features

Geostatistical texture features characterize textures based on the relationship between the similarity and distance between pixels of remotely sensed images [50]. Various studies [39, 51]

have shown that their usage can enhance the quality of forest classifications. They could also be useful for MCF mapping as trees often have a reduced stature under MCF conditions [1], which may affect the texture of Landsat bands.

We created texture images of the Landsat composites using the texture features variogram, madogram, rodogram, cross variogram, and pseudo cross-variogram [50, 51]. While the variogram, madogram and rodogram compare pixels of the same raster input to each other, the cross variogram and the pseudo-cross variogram characterize the spatial variability based on pixels of different raster inputs. The variogram, madogram and rodogram texture parameters were calculated from all Landsat bands. The cross variogram and pseudo cross-variogram parameters were calculated from each possible combination of two different Landsat bands. Each texture input was calculated twice, using lag distances of 1 and 2 pixels and a window size of 7 pixels.

Training data for the MCF conditions map

The data points used to train the Random Forest model for MCF conditions mapping were taken from the National Vegetation Database of Taiwan (Global Index of Vegetation-Plot Databases ID: AS-TW-001). Azonal forest types identified by Li et al. [29] were removed from the original data set because their presence or absence is not related to climatic factors such as the frequency of fog formation. The remaining 2367 points were classified into the categories "MCF" and "non-MCF" based on the composition of their woody species (trees and shrubs).

According to [29] there are four types of montane cloud forests in Taiwan: C2A03 *Chamaecyparis* montane mixed cloud forest, C2A04 *Fagus* montane deciduous broad-leaved cloud forest, C2A05 *Quercus* montane evergreen broad-leaved forest and C3A09 *Pasania-Elaeocarpus* montane evergreen broad-leaved forest. The first three types have a subtropical flora and the last one has a tropical flora. Each type is defined in [29] using a cocktail formula that requires the presence of certain species groups (and the absences of others) for a specific vegetation type [52–54].

In the current study, the MCF types C2A04 and C3A09 from [29] and species groups defined in [3] for the other montane cloud forest types were used. The latter definition was used because the study by [3] led to a better understanding of *Chamaecyparis* and *Quercus* cloud forest resulting in a better definition of these two types of MCF. Altogether, 834 plots were selected as MCF points in the training data set (cf. Fig 1). All other zonal forest plots were defined as non-MCF. From those, plots containing montane cloud forest species groups defined in [3] were further excluded to make these points purely represent forests uninfluenced by fog. In the end, there were 1533 points selected as non-MCF in the training data set (cf. Fig 1).

The classified data points are publicly available under doi:10.5678/LCRS/DAT.147.

Creation of the forest map

The training samples described in the previous section only include vegetation plots. The locations of unforested areas needed to be known as well to create the forest map. Therefore, 408 training locations for two classes of point data (forest and non-forest) distributed over the whole island were manually set in a GIS using the Landsat layers presented in Sect. *Landsat channels* and (in cases where the class could not be unambiguously identified based on the Landsat data) Google Earth as well as Google Street View data as reference. A Random Forest model was trained based on those point locations to detect forest in the Landsat data. After a recursive feature elimination was performed, the model was used to delineate the forested areas of Taiwan.

Methodology of validation

One advantage of Random Forest classifiers is that the same data can be used for training and for validation using an out-of-bag approach [55]. This was done for the map of MCF conditions as well as for the forest map. Using confusion matrices comparing the prediction results ("MCF", "non-MCF" / "forest", "non-forest") with the training data set, the following statistical measures were calculated [56]:

- Proportion correct (PC): Fraction of the point data that was correctly predicted by the binary classification;

- Bias: Over- and underestimation (values above and below 1, respectively) of MCF / forest occurrence;

- Probability of detection (POD): Fraction of MCF / forest points that were correctly predicted as MCF / forest;

- Probability of false detection (POFD): Fraction of non-MCF / non-forest points that were falsely predicted as MCF / forest;

- False alarm rate (FAR): Fraction of points for which MCF / forest was predicted that were actually non-MCF / non-forest;

- Matthews correlation coefficient (MCC) [57]: A measure for the overall quality of a binary classification. 1 = perfect agreement between classification and point data; -1 = total disagreement between classification and point data.

Results and discussion

Ground fog frequency maps

Frequency maps created from the DOGMA ground fog masks (publicly available under doi: 10.5678/LCRS/DAT.145) are presented in Fig 3. The frequency of ground fog is generally high at altitudes between 1500 and 2500 m a.s.l. (cross-hatched area). This corresponds to the montane cloud zone with the highest MCF occurrence. Distinct local deviations in the altitudinal distribution of ground fog occurrence do exist, however. On the western slopes of the CMR, fog clearly forms less frequently than on the northern and eastern slopes. More fog occurs also at low altitudes on the northern and eastern slopes. This is particularly pronounced during the northeasterly winter monsoon between September and May. Due to the spatial pattern, it can be assumed that a causal link to the winter monsoon exists. In the summer months, fog generally appears less frequently and is more evenly distributed across the entire island. Areas with frequent occurrence of ground fog in isolated mountain terrain at altitudes below 1500 m a.s.l., e.g. in the Hai'an Range on the east coast or the Yangmingshan in northern Taiwan (cf. Fig 1), could be caused by the Massenerhebung effect.

As cloud frequency is independent from frequent cloud immersion of the terrain, it is—in contrast to the frequency of ground fog—hardly correlated to the altitude of the terrain. Therefore, the map of full-year ground fog frequency presented in the this study differs strongly from the cloud frequency maps presented by Thies et al. [24] and Wilson and Jetz [25] (cf. Sect. Introduction). In the latter, the cloud frequency between 1500 and 2500 m a.s.l. is, for example, similar to that in wide areas of the lowlands. However, the inter-annual cloud variability calculated by Wilson and Jetz shows low values on the northern and eastern slopes of the CMR in particular, as well as on the Hai'an Range, where our map reveals high ground fog frequency values. This must be caused by the spatial distribution of ground fog in Taiwan
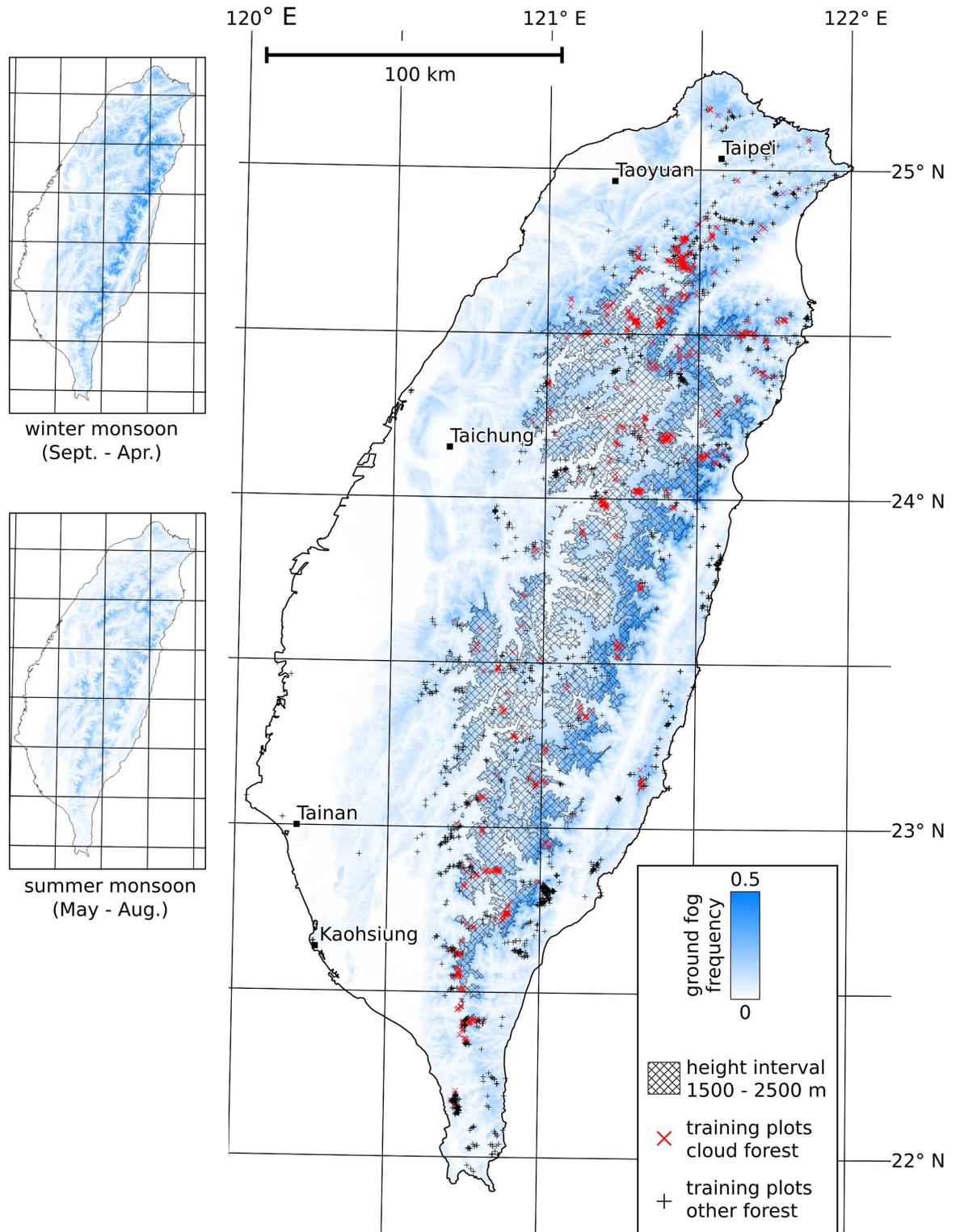
Fig 3. Ground fog frequency map calculated from MODIS data captured between 1 January 2003 and 31 December 2014.

doi:10.1371/journal.pone.0172663.g003

following yearly patterns more predictable than other clouds, most likely due to the influence of the winter monsoon.

Maps of MCF conditions

Slope and sky view factor were the only DEM-based parameters excluded from the input data set by the recursive feature elimination. Bands 4 and 7 as well as all texture measures except six variants of the pseudo cross-variogram with different lag sizes and combinations of bands were removed from the Landsat-based inputs. The frequency input data set remained completely.

The map of MCF conditions predicted by the Random Forest model based on the remaining inputs is presented in Fig 4A. It is publicly available under doi:10.5678/LCRS/DAT.278. A second map (Fig 4B) was created for comparison. The ground fog frequency inputs were excluded from the model inputs before the recursive feature elimination was conducted for this map.

Out-of-bag validation for both maps can be found in Table 1. The results reveal that both maps are of high quality. Map A has a slightly higher overall quality than map B in a direct comparison of the Matthews correlation coefficients and the proportion correct. The proportion of MCF plots correctly reproduced in the map (probability of detection) is also slightly higher in map A than for map B. The fraction of the plots for which MCF conditions were predicted although they are actually non-MCF plots (false alarm rate) is lower for map A. The fraction of non-MCF plots that were mistaken for MCF (probability of false detection) as well as the degree of underestimation of the area of MCF conditions (Bias) are the same for both maps.

In order to quantitatively assess the altitudinal dependency of the prediction quality, validation was performed a second time for vegetation plots at altitudes below 1500 m a.s.l. and above 1500 m a.s.l., respectively. The results are shown in Table 1. The validation based on plots above 1500 m a.s.l. shows a pattern that is similar to the validation result based on all plots. Here, however, the probability of detection is better in map B than in map A. Overall, the benefit of using frequency inputs seems insignificant at altitudes above 1500 m a.s.l. This is due to a relatively small transition zone from MCF to non-MCF conditions at an altitude of $\sim$ 2500 m a.s.l. (cf. Fig 2). Above this transition zone, MCF conditions hardly occur. Below this transition zone, almost all plots within the > 1500 m a.s.l. class are MCF. Therefore, DEM-related inputs alone are well suited to distinguish MCF plots from non-MCF plots here. As most of the MCF plots used in validation are located above 1500 m a.s.l., this is strongly reflected in the validation including all plots and explains similar validation results for both maps.

In the < 1500 m a.s.l. class, the overall prediction quality as measured by the Matthews correlation coefficient and the proportion correct is higher for map A. The difference in the Matthews correlation coefficient between map A and map B is much more distinct than in the validation including all plots. This is mostly due to a clearly higher probability of detection resulting in a higher (= better) Bias. In summary, frequency raster inputs clearly contribute to the correct prediction of MCF plots at low heights, where altitude is not a useful predictor. Significant visual differences between both maps reflect this finding. The area of MCF conditions in Fig 4B corresponds very well with the interval between 1500 and 2500 m a.s.l. There is no visible influence of local deviations in the altitudinal distribution of ground fog occurrence on the distribution of MCF. It is, however, captured in Fig 4A. More areas with MCF conditions can be found at altitudes below 1500 m a.s.l on the northern and eastern slopes of the CMR and in low areas in northern Taiwan (domain 1 in Fig 4 including Yangmingshan) as well as
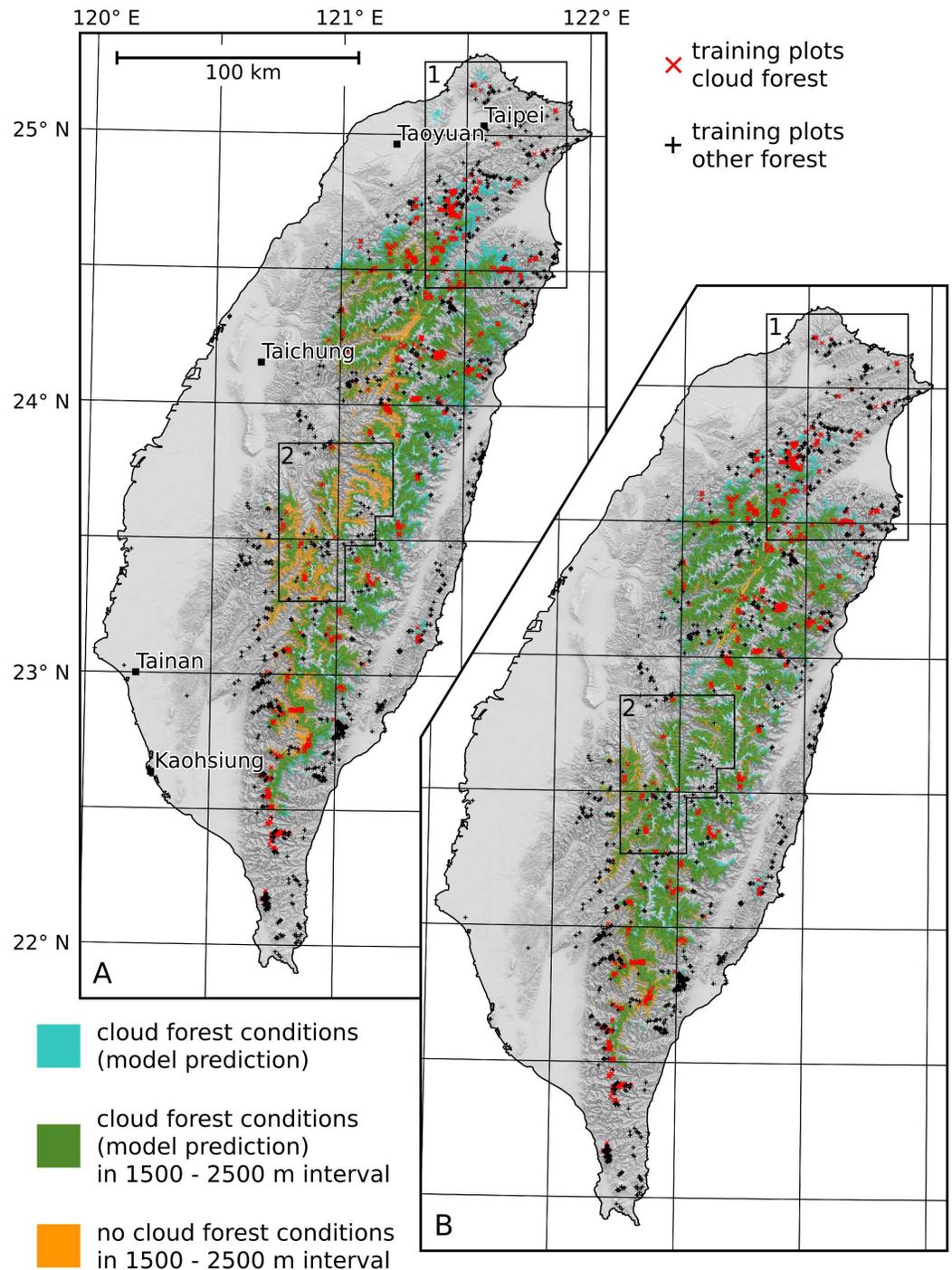
Fig 4. Maps of MCF conditions with (A) and without (B) the ground fog frequency being included.

in the Hai'an Range. All of these areas are exposed to the winter monsoon and some of them have isolated mountain terrain. In contrast, MCF conditions can mainly be found at altitudes above 2000 m a.s.l. in the western part of the CMR in central Taiwan (domain 2 in Fig 4) that is less influenced by the winter monsoon.

Table 1. Validation results for MCF conditions maps A and B. The best value of each statistical parameter is written in bold type. TP = true positives; TN = true negatives; FP = false positives; FN = false negatives; MCC = Matthews correlation coefficient; PC = Proportion correct (PC); POD = Probability of detection; POFD = Probability of false detection; FAR = False alarm rate (FAR).

TP	TN	FP	FN	MCC	PC	Bias	POD	POFD	FAR	Validated map and height interval
765	1477	56	69	**0.88**	**0.95**	**0.99**	**0.92**	**0.04**	**0.07**	A
130	1259	28	50	**0.74**	**0.95**	0.88	**0.72**	**0.02**	**0.18**	A (<1500 m a.s.l.)
635	218	28	19	**0.87**	**0.95**	**1.01**	0.97	**0.11**	**0.04**	A (≥1500 m a.s.l.)
754	1466	67	80	0.86	0.94	**0.99**	0.90	**0.04**	0.08	B
112	1258	29	68	0.67	0.93	0.79	0.62	**0.02**	0.21	B (<1500 m a.s.l.)
642	208	38	12	0.86	0.94	1.04	**0.98**	0.15	0.06	B (≥1500 m a.s.l.)

doi:10.1371/journal.pone.0172663.t001

In order to assess the explanatory power of the different input raster sets, our study independently validated the monthly ground fog frequencies, the DEM-based inputs as well as the Landsat bands and textures. The input raster sets were reduced in their dimensionality through recursive feature elimination and were each used to train a separate Random Forest model, which were then validated using the out-of-bag approach. The results are shown in Table 2. Most statistical parameters calculated from the frequency input raster set have the best values. As the monthly ground fog frequencies are better suited to create a map of MCF conditions than the DEM-based inputs, the relatively small advantage of the inclusion of ground fog frequency maps for the mappings of MCF conditions above 1500 m a.s.l. must be a result of the strong height dependence of ground fog frequency in these altitudes.

Clearly, the monthly ground fog frequency maps are not a perfect predictor for MCF conditions. Otherwise the Random Forest model based solely on frequency data would be as good as the models including more than one input raster set. This could be the result of MCF conditions depending on environmental factors other than ground fog frequency. Altitude, for example, is not only a proxy for the fog frequency but also for other climatic and edaphic parameters with impact on the floristic composition. Furthermore, the low temporal sampling rate (particularly the lack of nighttime scenes) of the satellite data used to create the ground fog frequency maps as well as imperfections in the fog detection algorithm (e.g., problems with clouds with an optical thickness greater than 40 [43]) result in flaws in the frequency data that reduce their explanatory power.

The forest map and final MCF map

The forest map (publicly available under doi:10.5678/LCRS/DAT.278) is presented as the light grey and green area in Fig 5. The out-of-bag validation (223 true positives, 180 true negatives, 3 false positives, 3 false negatives) verifies that it is overall of very high quality (Matthews correlation coefficient = 0.97, proportion correct = 0.99). Almost all forest training samples were predicted correctly (probability of detection = 0.99) and almost none of the non-forest samples were mistaken for forest (probability of false detection = 0.02). In addition, nearly no sample

Table 2. Validation results for Random Forest models trained with different raster input sets. The best value of each statistical parameter is written in bold type. TP = true positives; TN = true negatives; FP = false positives; FN = false negatives; MCC = Matthews correlation coefficient; PC = Proportion correct (PC); POD = Probability of detection; POFD = Probability of false detection; FAR = False alarm rate (FAR).

TP	TN	FP	FN	MCC	PC	Bias	POD	POFD	FAR	used raster input set
723	1446	87	111	**0.82**	**0.92**	0.97	**0.87**	**0.06**	**0.11**	monthly ground fog frequency maps
719	1438	95	115	0.80	0.91	**0.98**	0.86	**0.06**	0.12	DEM-based inputs
686	1415	118	148	0.75	0.89	0.96	0.82	0.08	0.15	Landsat-based inputs

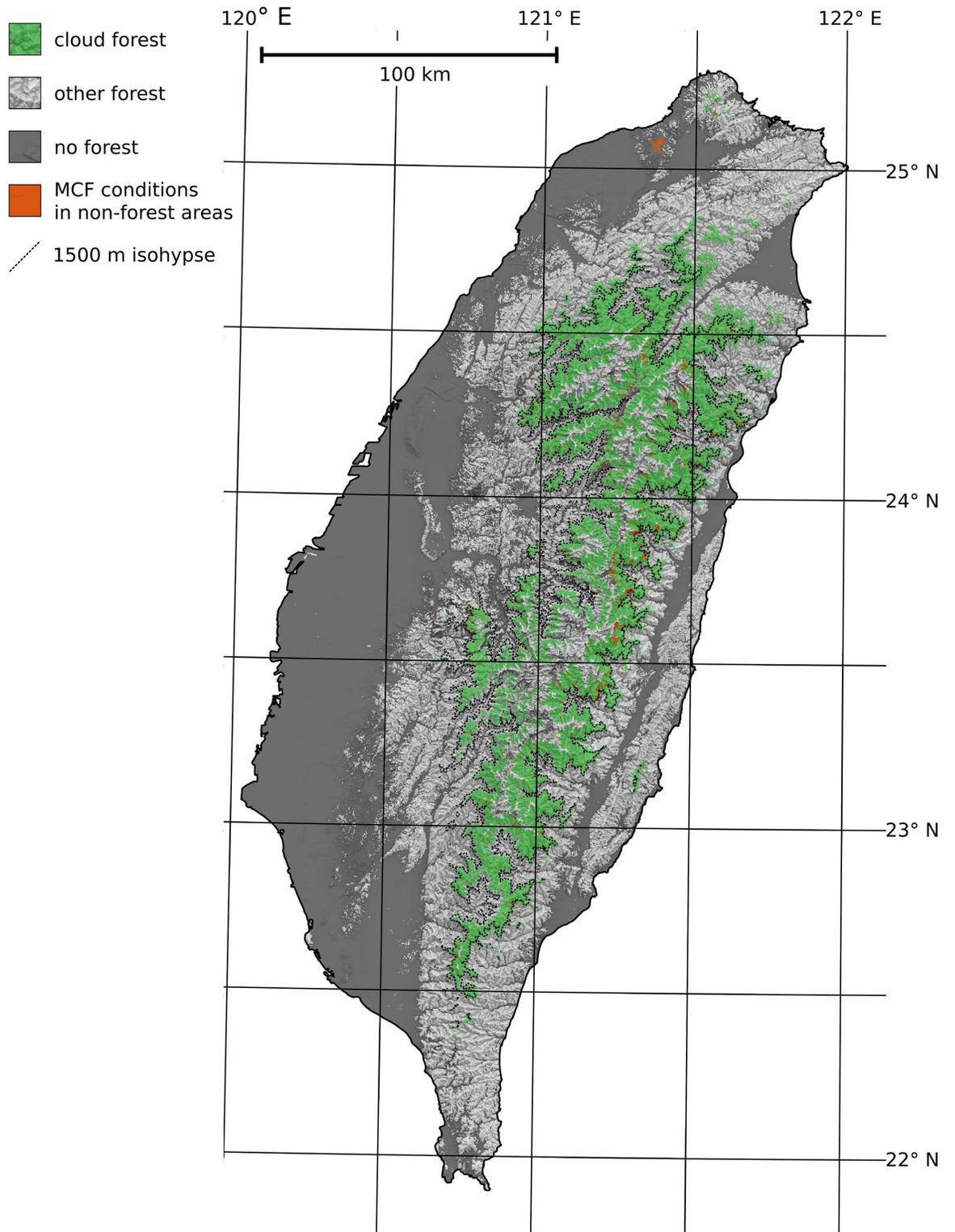doi:10.1371/journal.pone.0172663.t002

Fig 5. Final MCF map (green area), created from the combination of the MCF conditions map B (green and orange area) and the forest map (white and green area).

for which forest was predicted was actually forest free (false alarm rate = 0.01). No systematic over- or underestimation occurred (Bias = 1.00).

The final MCF map—the intersection of the map of MCF conditions (Fig 4A) and the forest map—is presented as the green area in Fig 5. It is publicly available under doi:10.5678/LCRS/ DAT.278. This map displays only pixels where both MCF conditions and actual forest occur. As MCF conditions mostly occur in the wooded center of Taiwan, the difference (orange area = pixels with MCF conditions but no forest occurrence) between the final MCF map and the map of MCF conditions is small. In general, relatively small features such as river beds or bedrock on steep slopes were removed from the map of MCF conditions. Only in the urban north of the island was a larger cluster of pixels removed.

Despite substantial differences in the frequency inputs and different spatial resolutions, the southern part of the MCF map presented in Fig 5 and the Taiwanese areas of tropical MCF mapped by Wilson and Jetz [25] mostly overlap. As the cloud frequency input used by Wilson and Jetz does distinguish the lower boundary of MCF conditions, it is reasonable to assume that the southern Taiwanese MCF was correctly mapped mostly due to the elevation input and the inter-annual cloud variability.

Conclusion and outlook

This study produced the first comprehensive map of montane cloud forests in Taiwan. To the best of our knowledge, it is also the first study to successfully use a high-resolution ground fog frequency climatology in vegetation mapping. An MCF map based on DEM-derived parameters and Landsat data alone correctly locates much of the Taiwanese cloud forest. Including frequency data for ground fog in the machine learning inputs captures what we consider to be the influence of the Massenerhebung effect as well as the monsoonal impact on the island. This significantly enhances the detection of montane cloud forest at low altitudes.

As stated in Sect. *Maps of MCF conditions*, the low temporal sampling rate of MODIS data is an issue of the mapping approach presented in this study. MODIS was chosen over Himawari 7 due to its higher spatial resolution. Although the geostationary satellite Himawari 8 has a relatively high spatial resolution of 500 m per pixel and high temporal sampling rate of 10 minutes, it has only been in operational service above the western Pacific Ocean since July 2015. Once it has delivered sufficient data for the creation of meaningful ground fog frequency maps, it will help to further refine the mapping of MCF in Taiwan, e.g., by capturing intradiurnal variations of the ground fog frequency as described in [58]. Additional products could also be included in the machine learning approach, e.g. daily fog duration thanks to the high temporal sampling rate of Himawari 8. Also, the inclusion of further environmental parameters that can be derived from remotely sensed data might further enhance the correct delineation of Taiwanese MCF areas in future studies (Mulligan and Burke [18], for example, stressed the importance of the cloud cover frequency as a proxy for rainfall and the reduction of incoming radiation).

Author Contributions

Conceptualization: JB SCC BT.

Data curation: CFL.

Formal analysis: HMS.

Funding acquisition: BT JB SCC.

Methodology: HMS CFL BT.

Project administration: JB SCC.

Software: HMS.

Supervision: JB SCC BT.

Validation: HMS.

Visualization: HMS.

Writing – original draft: HMS CFL.

Writing – review & editing: HMS.

References

1. Hamilton LS. Mountain cloud forest conservation and research: a synopsis. Mt Res Dev. 1995; 10: 259–266. doi: 10.2307/3673933

2. Bruijnzeel LA, Kappelle M, Mulligan M, Scatena FN. Tropical montane cloud forests: state of knowledge and sustainability perspectives in a changing world. In: Bruijnzeel LA, Mulligan M, Scatena FN, editors. Hydrometeorology of tropical montane cloud forests: emerging patterns. Cambridge: Cambridge University Press; 2010. pp. 691–740.

3. Li CF, Zelený D, Chytrý M, Chen MY, Chen TY, Chiou CR, et al. Chamaecyparis montane cloud forest in Taiwan: ecology and vegetation classification. Ecol Res. 2015; 30: 771–791. doi: 10.1007/s11284-015-1284-0

4. Aldrich M, Billington C, Edwards M. Tropical montane cloud forests: an urgent priority for conservation. WCMC Biodiversity Bulletin No. 2. Cambridge: UNEP World Conservation Monitoring Centre; 1997.

5. Bubb P, May I, Miles L, Sayer J. Cloud forest agenda. Cambridge: UNEP World Conservation Monitoring Centre; 2004.

6. Foster PN. The potential negative impacts of global climate change on tropical montane cloud forests. Earth-Sci Rev. 2001; 55: 73–106. doi: 10.1016/S0012-8252(01)00056-3

7. Hu J, Riveros-Iregui DA. Life in the clouds: are tropical montane cloud forests responding to changes in climate? Oecologia 2016; doi: 10.1007/s00442-015-3533-x PMID: 26739003

8. Loope LL, Giambelluca TW. Vulnerability of island tropical montane cloud forest to climate change, with special reference to East Maui, Hawaii. Clim Chang. 1998; 39: 503–517. doi: 10.1023/A:1005372118420

9. Ponco-Reyes R, Reynoso-Rosales VH, Watson JEM, VanDerWal J, Fuller RA, Pressey RL, Possingham HP. Vulnerability of cloud forest reserves in Mexico to climate change. Nat Clim Chang 2012; 2: 448–452. doi: 10.1038/nclimate1453

10. Still C, Foster PN, Schneider SH. Simulating the effects of climate change on tropical montane cloud forests. Nature. 1999; 398: 608–610. doi: 10.1038/19293

11. Wey TH, Lai YJ, Chang CS, Shen CW, Hong CY, Wand YN, Chen MC. [Preliminary studies on fog characteristics at Xitou region of central Taiwan]. [Journal of the Experimental Forest of National Taiwan University]. 2011; 25:149–160. Chinese.

12. Chen YF. [Vegetation of Taiwan—4th Volume: Chamaecyparis Cloud Forest Zone]. Taipei: Avanguardbook; 2001. Chinese.

13. Lawton RO, Nair US, Pielke RA Sr., Welch RM. Climatic impact of tropical lowland deforestation on nearby montane cloud forest. Science 2001; 294: 584–587. PMID: 11641496

14. Peh KSH, Soh MCK, Sodhi NS, Laurance WF, Ong DF, Clements R. Up in the clouds: is sustainable use of tropical montane cloud forests possible in Malaysia? Bioscience. 2011; 61: 27–38. doi: 10.1525/bio.2011.61.1.8

15. Pope I, Bowen D, Harbor J, Shao G, Zanotti L, Burniske G. Deforestation of montane cloud forest in the Central Highlands of Guatemala: contributing factors and implications for sustainability in Q'eqchi' communities. Int J Sust Dev World; 22: 201–212. doi: 10.1080/13504509.2014.998738

16. Campanella R. The role of GIS in evaluating contour-based limits of cloud forest reserves in Honduras. In: Hamilton LS, Juvik JO, Scatena FN, editors. Tropical Montane Cloud Forests, Ecological Studies, vol. 110. New York: Springer-Verlag; 1995. pp 116–124.

17. Scatena FN, Bruijnzeel LA, Bubb P, Das S. Setting the stage. In: Bruijnzel LA, Scatena FN, Hamilton LS, editors. Tropical montane cloud forests: science for conservation and management. Cambridge: Cambridge University Press; 2010. pp 3–13.

18. Mulligan M, Burke SM. DFID FRP Project ZF0216 Global Cloud Forests and Environmental Change in a Hydrological Context. Final report. London: DFID; 2005 Dec.

19. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol. 2005; 25: 1965–1978. doi: 10.1002/joc.1276

20. New MG, Hulme M, Jones PD. Global monthly climatology for the twentieth century. Database: ORNL DAAC [Internet]. Accessed: http://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=551.

21. Allan R, Ansell T. A new globally-complete monthly historical gridded mean sea level pressure data set (HadSLP2): 1850-2004. J Clim. 2006; 19: 5816–5842. doi: 10.1175/JCLI3937.1

22. Nair US, Asefi S, Welch RM, Ray DK, Lawton RO, Manoharan VS, et al. Biogeography of tropical montane cloud forests. Part II: mapping of orographic cloud immersion. J Appl Meteorol Climatol. 2008; 47: 2183–2197. doi: 10.1175/2007JAMC1819.1

23. Aldrich M, Billington C, Edwards M, Laidlaw R. A global directory of tropical montane cloud forest, Cambridge: UNEP World Conservation Monitoring Centre; 2004.

24. Thies B, Groos A, Schulz M, Li CF, Chang SC, Bendix B. Frequency of low clouds in Taiwan retrieved from MODIS data and its Relation to cloud forest occurrence. Remote Sens., 2015; 7: 12986–13004. doi: 10.3390/rs71012986

25. Wilson AM, Jetz W. Remotely sensed high-resolution global cloud dynamics for predicting ecosystem and biodiversity distributions. PLOS Biol. 2016; doi: 10.1371/journal.pbio.1002415

26. Central Weather Bureau. Monthly mean temperatures, Available: http://www.cwb.gov.tw/V7e/climate/monthlyMean/tx.htm

27. OpenStreetMap Contributers. OpenStreetMap. Available: http://www.openstreetmap.org/

28. Chen TC, Yen MC, Hsieh JC, Arrit RW. Diurnal and seasonal variations of the rainfall measured by the automatic rainfall and meteorological telemetry system in Taiwan. B Am Meteorol Soc. 1999; 80: 2299–2312. doi: 10.1175/1520-0477(1999)080%3C2299:DASVOT%3E2.0.CO;2

29. Li CF, Chytrý M, Zelený D, Chen MY, Chen TY, Chiou CR, et al. Classification of Taiwan forest vegetation. Appl Veg Sci. 2013; 16: 698–719. doi: 10.1111/avsc.12025

30. Lai IL, Chang SC, Lin PH, Chou CH, Wu JT. Climatic characteristics of the subtropical mountainous cloud forest at Yuanyang Lake long-term ecological research site, Taiwan. Taiwania 2006; 51: 317–329.

31. Bruijnzeel LA, Waterloo MJ, Proctor J, Kuiters AT, Kotterink B. Hydrological observations in montane rain forest on Gunung Silam, Sabah, Malaysia, with special reference to the 'Massenerhebung' effect. J Ecol. 1993; 81: 145–167. doi: 10.2307/2261231

32. Flenley JR. Cloud forest, the Massenerhebung effect and ultraviolet insolation. In: Hamilton LS, Juvik JO, Scatena FN, editors. Tropical Montane Cloud Forests, Ecological Studies, vol. 110. New York: Springer-Verlag; 1995. pp 150–155.

33. Ohsawa M. The montane cloud forest and its gradational changes in southeast Asia. In: Hamilton LS, Juvik JO, Scatena FN, editors. Tropical Montane Cloud Forests, Ecological Studies, vol. 110. New York: Springer-Verlag; 1995. pp 254–265.

34. Su HJ. Studies on the climate and vegetation types of the natural forests in Taiwan (II): Altitudinal vegetation zones in relation to temperature gradient. Q J Chin Ecol. 1984; 17: 57–73.

35. Breiman L. Random forests. Mach Learn. 2001; 45: 5–32. doi: 10.1023/A:1010933404324

36. R Core Team. R version 3.2.2: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2015. Accessed: https://www.R-project.org/

37. Liaw A, Wiener M. RandomForest version 4.6.12; 2015. Accessed: https://cran.r-project.org/web/packages/randomForest/index.html

38. Feng Q, Liu J, Gong J. UAV remote sensing for urban vegetation mapping usng random forest and texture analysis. Remote Sens. 2015; 7: 1074–1094. doi: 10.3390/rs70101074

39. Galiano-Rodriguez VF, Chica-Olmo M, Abarca-Hernández F, Atkinson PM, Jeganathan C. Random forest classification of mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. Remote Sens Environ. 2012; 121: 93–107. doi: 10.1016/j.rse.2011.12.003

40. Lawrence RL, Wood SD, Sheley RL. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (random forest). Remote Sens Environ. 2006; 100: 356–362. doi: 10.1016/j.rse.2005.10.014

41. Alonso MC, Malpica JA, de Agirre AM. Consequences of the Hughes phenomenon on some classification techniques. Proceedings of the ASPRS 2001 annual conference; 2011 May 1-5, Milwaukee, Wi.

42. Kuhn M, Wing J, Weston S, Williams A, Keefer C, Engelhardt A, et al. Caret: Classification and Regression Training. R package version 6.0.62; 2015. Accessed: http://cran.r-project.org/package=caret

43. Schulz HM, Chang SC, Thies B, Bendix J. Detection of ground fog in mountainous areas from MODIS (Collection 051) day-time data using a statistical approach. Atmos Meas Tech. 2016; 9: 1135–1152. doi: 10.5194/amt-9-1135-2016

44. United States Geological Survey. ASTER GDEM 2; 2011. Database: Global Data Explorer [Internet]. Accessed: http://gdex.cr.usgs.gov/

45. Dubayah R, Paul MR. Topographic solar radiation models for GIS. Int J Geogr Inf Syst. 2007; 9: 405–419. doi: 10.1080/02693799508902046

46. Vermote EF, Tanré D, Deuzé JL, Herman M, Morcette JJ. Second simulation of the satellite signal in the solar spectrum, 6S: an overview. IEEE T Geosci Remote. 1997; 35: 675–686. doi: 10.1109/36.581987

47. Curatola Fernández GF, Obermeier WA, Gerique A, López Sandoval MF, Lehnert LW, Thies B., Bendix J. Land cover change in the Andes of Southern Ecuador—patterns and drivers. Remote Sens. 2015; 7: 2509–2542. doi: 10.3390/rs70302509

48. Teillet P, Guindon B, Goodenough DG. On the slope-aspect correction of multispectral scanner data. Can. J. Remote Sens. 1981; 8: 84–106.

49. Martinuzzi S, Gould WA, Ramos González OM. Creating cloud-free Landsat ETM+ data sets in tropical landscapes: cloud and cloud-shadow removal. General technical report. Washington, DC: USDA; 2007 Feb. Report No.: IITF-32.

50. Chica-Olmo M., Abarca-Hernández F. Computing geostatistical image texture for remotely sensed data classification. Comput Geosci. 2001; 26: 373–383. doi: 10.1016/S0098-3004(99)00118-1

51. Wijaya A, Marpu PR, Gloaguen R. Geostatistical texture classification of tropical rainforest in Indonesia. In: Stein A, Bijker W, editors. Quality aspects in spatial data mining. Boca Raton, Florida: CRC Press; 2008. pp. 199–210.

52. Bruelheide H. [The grasslands of the Harz and their site conditions. Including a report about the classification system based on statistically derived groups of species]. Diss Bot. 1995; 244: 1–338. German.

53. Chytrý M. Vegetation of the Czech Republic. 1st ed. Prague: Academia; 2007.

54. Kočí M, Chytrý M, Tichý L. Formalized reproduction of an expert-based phytosociological classification: a case study of subalpine tall-forb vegetation. J Veg Sci. 2003; 14: 601–610. doi: 10.1111/j.1654-1103.2003.tb02187.x

55. Breiman L. Out-of-bag estimation. Technical report. Berkley, Ca: University of California, Berkley; 1996.

56. Mason IB. Binary events. In: Joliffe IT, Stephensons DB, editors. Forecast verification—a practioner's guide in atmospheric science. Chicester, UK: John Wiley & Sons; 2003. pp. 37–76.

57. Matthews BW. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochim Biophys Acta. 1975; 405: 442–451. doi: 10.1016/0005-2795(75)90109-9 PMID: 1180967

58. Chang SC, Lai IL, Wu JT. Estimation of fog deposition on epiphytic bryophytes in a subtropical montane forest ecosystem in northeastern Taiwan. Atmos Res. 2002; 64: 159–167.