


RESEARCH ARTICLE

Combining UAV thermography, point cloud analysis and machine learning for assessing small-scale evapotranspiration patterns in a tropical rainforest

Álvaro Cortés-Molino^{1,2} | Alejandra Valdés-Uribe² | Florian Ellsäßer^{3,2} |
 Medha Bulusu² | Joyson Ahongshangbam^{4,2} | Hendrayanto⁵ | Dirk Hölscher^{2,6} |
 Alexander Röhl² 

¹Departamento de Botánica y Fisiología Vegetal, Universidad de Málaga, Málaga, Spain

²Tropical Silviculture and Forest Ecology, University of Göttingen, Göttingen, Germany

³Department for Natural Resources, ITC, University of Twente, Enschede, Netherlands

⁴Institute for Atmospheric and Earth System Research, University of Helsinki, Helsinki, Finland

⁵Forest Management Department, Bogor Agricultural University, Bogor, Indonesia

⁶Centre of Biodiversity and Sustainable Land Use, University of Göttingen, Göttingen, Germany

Correspondence

Alexander Röhl, Tropical Silviculture and Forest Ecology, University of Göttingen, Büsgenweg 1, 37077 Göttingen, Germany.
 Email: aroell@gwdg.de

Funding information

Deutsche Forschungsgemeinschaft, Grant/Award Number: CRC990; University of Málaga; University of Göttingen

Abstract

Microclimate and vegetation structure control evapotranspiration (ET) from land surfaces at stand and landscape scales. Tropical rainforests are among the most diverse and complex terrestrial ecosystems, harbouring vast plant and animal species throughout their dense multistory canopy. They contribute substantially to global precipitation through their high ET. However, there is little information about ET influences at very small spatial scales under given climatic conditions. In a tropical rainforest on Sumatra, we studied the relationship between pixel-level ET as derived from high-resolution (~10 cm), near-surface thermography from an unmanned aerial vehicle (UAV) and canopy structure as derived from red–green–blue (RGB) image and three-dimensional (3D) point cloud analyses. The 16 derived potential predictors encompassed vegetation height, height variability, vegetation density and reflectance variables. Using regression models, several of the studied variables had a significant linear relationship with ET, but the explained variance was only marginal. However, applying a random forest algorithm including forward feature selection and target oriented cross validation explained substantial parts of the pixel-level variance in ET ($R^2 = 0.56\text{--}0.65$), thus indicating multiple non-linear relationships with interactions among predictor variables. Therein, green leaf index, leaf area density and vegetation height were often the most important variables for the prediction outcome, but their sequence varied among the four study plots. Overall, combining canopy structure variables derived from RGB photogrammetry explained relatively large parts of spatial ET variations. Our study thus indicates the large potential of combining UAV-based thermography and photogrammetry techniques with machine learning approaches to better understand ET but also suggests that more work remains to be done in explaining ET patterns at very small spatial scales.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Ecohydrology* published by John Wiley & Sons Ltd.

KEYWORDS

canopy structure, drone, leaf area density, photogrammetry, Sumatra, vegetation height

1 | INTRODUCTION

Climate change is projected to lead to an intensification of the hydrological cycle (Allan et al., 2020; Yeh & Wu, 2018) and to increases in frequency and severity of extreme meteorological events such as droughts or floods (Barichivich et al., 2018; Donat et al., 2019; Kundzewicz, 2008; Trenberth, 1999). Terrestrial evapotranspiration (ET) as a key component in the hydrological cycle accounts for approximately 40% of global precipitation (Schlesinger & Jasechko, 2014). ET is the combination of water evaporation from surfaces and transpiration from plants. The contribution of transpiration to ET varies across biomes, around 50% in drylands and up to 70% in tropical rainforests (Schlesinger & Jasechko, 2014). Assessments of ET in tropical rainforests are generally very challenging due to the remoteness and high topographical and structural complexity of the ecosystem.

ET dynamics are driven by micrometeorological variables including solar radiation, vapour pressure deficit and wind speed. The combination of these variables can be expressed in the maximum possible ET rate under ideal conditions, the potential evapotranspiration (PET; Allen et al., 1994). PET assumes an unlimited supply of water, while actual ET is strongly governed by water availability, that is, precipitation and soil moisture, and is further strongly influenced by topographical and physiological aspects (Dimitriadou & Nikolakopoulos, 2021). The transpiration component of ET largely depends on leaf surface and canopy properties. Crown leaf density and vertical vegetation structure (e.g. tree or crown height) have been proposed as key regional controls of ET (Coronel et al., 2011), while studies at the local level also report strong influences of canopy conductance and plant health (e.g. Marques et al., 2020; Nazarbaksh et al., 2019). While ET controls at larger spatial scales thus are relatively well studied (e.g. Nelson et al., 2020; Williams et al., 2012), small-scale controls of ET have received less attention, particularly in tropical rainforests. Studies at the tree level indicate a dominant influence of tree and crown size on per-tree transpiration (Ahongshangbam et al., 2020; Kotowska et al., 2021; Meinzer et al., 2005). Local differences in tree density and sizes and the associated differences in local ‘canopy packing’ were reported to lead to considerable small-scale variability in transpiration patterns at eight lowland rainforest sites (Ahongshangbam et al., 2020). Assessments of rainforest ET at even higher spatial resolution (i.e. in cm or dm scales) are to our knowledge not yet available. From the studies at larger spatial scales, one may speculate that key variables such as vegetation height and its variability and vegetation and leaf density may also influence ET at smaller spatial scales under given micrometeorological conditions.

A widely used method to measure ET fluxes is the eddy covariance technique, where ET is measured at very high temporal resolution and over long periods of time. However, the technique averages

ET in a single value over a footprint of typically a few hectares around the measurement tower (Baldocchi, 2020). At larger spatial scales, ET is often assessed via thermography from satellites, but even modern products such as ECOSTRESS have only moderate spatial and temporal resolution, that is, 70 m and one overflight a day (Fisher et al., 2020). A further major constraint of passive satellite methods is cloud cover, which can result in only few usable images per year for a given (tropical) region. A promising method for the assessment of ET, at landscape scale and with high temporal (e.g. hourly) and very high spatial resolution (i.e. cm scale), is thermography from unmanned aerial vehicles (UAV), also referred to as drones (Rauneker & Lischeid, 2012). While establishing high-resolution temporal ET time series from multiple UAV flights is logistically more challenging, spatial ET patterns at a given sites can often be assessed from a single flight, for example, carried out close to noon under conditions of relatively high solar irradiance. To derive ET from land surface temperatures, energy balance models with varying levels of complexity can be applied. Timmermans et al. (2015) developed the ‘Deriving Atmosphere Turbulent Transport Useful To Dummies Using Temperature’ (DATUTDUT) model, a simplified one-source energy balance model that requires only surface temperature data and few auxiliary variables such as time and location as input. The DATUTDUT model was implemented in the QGIS plugin QWaterModel (Ellsäßer, Röhl, Stiegler, et al., 2020) to provide a simple-to-use graphical interface for ET assessments. The approach showed high agreement with reference ground methods across spatial scales from leaf to ecosystem when tested in a tropical oil palm plantation (Ellsäßer et al., 2021; Ellsäßer, Röhl, Stiegler, et al., 2020) and in a tropical agroforest (Ellsäßer, Röhl, Ahongshangbam, et al., 2020).

Using the structure from motion (SfM) technique, the images recorded with UAVs can be used to calculate high-density three-dimensional (3D) point clouds and orthomosaic maps (Westoby et al., 2012). Commonly, this is carried out with UAV-recorded red-green-blue (RGB) images. In analogy to point clouds derived from the typically much more expensive LiDAR technology, forest structure variables and crown and canopy metrics can be extracted from the photogrammetrically derived point clouds (Iglhaut et al., 2019). Previous studies at the tree level showed that crown metrics derived from RGB 3D point cloud analysis were better predictors of per-tree and per-palm transpiration than conventionally applied variables like stem diameter (Ahongshangbam et al., 2019, 2020). We are not aware of any previous studies connecting ecohydrological forest functions such as transpiration or ET to forest structure at even smaller spatial scales. Potential variables of interest derived from RGB images and point clouds in the context of small-scale ET patterns include simple indices such as the green leaf index (GLI; Louhaichi et al., 2001) and visual atmospherically resistance index (VARI; Gitelson et al., 2002), point cloud variables related to, for example, voxel density or height distribution

(McGaughey & Carson, 2003) as well as classic ecological variables such as the leaf area index or leaf area density (LAD) (Kricher, 2011).

High-resolution UAV imagery produces large datasets with potentially millions of georeferenced pixels and multiple variables derived from different sensors, with complex, often non-linear relationships and variable interactions. For analysing such large datasets, machine learning techniques are widely employed across the ecological sciences because of their typically much better model performance and prediction accuracy compared to conventional methods (e.g. Camps-Valls, 2009; Maxwell et al., 2018; Meyer & Pebesma, 2022). Regardless of the applied algorithm, spatial predictions via machine learning are time consuming and computationally intensive, limiting the application of multiple algorithms at the same time and highlighting the importance of the choice of an appropriate algorithm for the given study context. For spatial predictions in ecological studies, random forest stands out among the available algorithms as particularly well performing (Ahmad et al., 2017; Fernandez-Delgado et al., 2014), for example, when applied to predict reference ET (Dias et al., 2021; Feng et al., 2017), ET of tropical mountain forests (Valdés-Urbe et al., 2023), water stress (Virnodkar et al., 2020), sap flux and leaf stomatal conductance (Ellsäßer, Röhl, Ahongshangbam, et al., 2020), net ecosystem exchange (Reitz et al., 2021) or land-cover change (Aide et al., 2013). Recent studies have proposed solutions to previous autocorrelation and overfitting issues in spatial predictions via forward feature selection (FFS) and target oriented cross validation, thus minimizing the risk of spatial overfitting with random forests and showing realistic overall model performances (Gasch et al., 2015; Meyer et al., 2018, 2019).

In our study, we used an UAV to capture high-resolution, close-to-surface thermal and RGB images of the canopy at four study plots in a lowland rainforest in Sumatra, Indonesia. We applied the DAT-TUTDUT model to assess pixel-level ET and derived spatially

matching structural canopy variables from RGB image and point cloud analysis. The objectives were to further explore small-scale patterns in ET, with a focus on (i) assessing the relationship between ET and single structural variables and (ii) evaluating whether ET can be predicted reliably from multiple canopy variables with a random forest machine learning approach.

2 | MATERIAL AND METHODS

2.1 | Study region and plots

The study area was in the Harapan rainforest in Jambi province (Sumatra, Indonesia) (Figure 1a). The climate is tropical humid with mean annual precipitation of 2235 mm year⁻¹ and average annual temperature of 26.7°C, with two peak rainy seasons in March and December, and a drier period from June to September. The soils in this area are loamy Acrisols (Drescher et al., 2016). The Harapan rainforest is a tropical lowland rainforest that has a legacy of selective logging and wood extraction (Harrison & Swinfield, 2015). Today, the region almost has no natural forest left and has largely been converted to agricultural land including large proportions of oil palm and rubber plantations (Clough et al., 2016). The Harapan rainforest is a protected area managed by the governmental PT REKI company.

Our study plots (plot codes: HF1, HF2, HF3, HF4; Drescher et al., 2016; Figure 1) were upland rainforest sites located inside the Harapan rainforest. The plots were 50 × 50 m² in size, with a mean elevation of 65 m a.m.s.l. (for more details, see Ahongshangbam et al., 2020; Drescher et al., 2016). A previous assessment found over 200 tree species with a diameter at breast height (DBH) of more than 10 cm across the four plots (Rembold et al., 2017). Stand densities

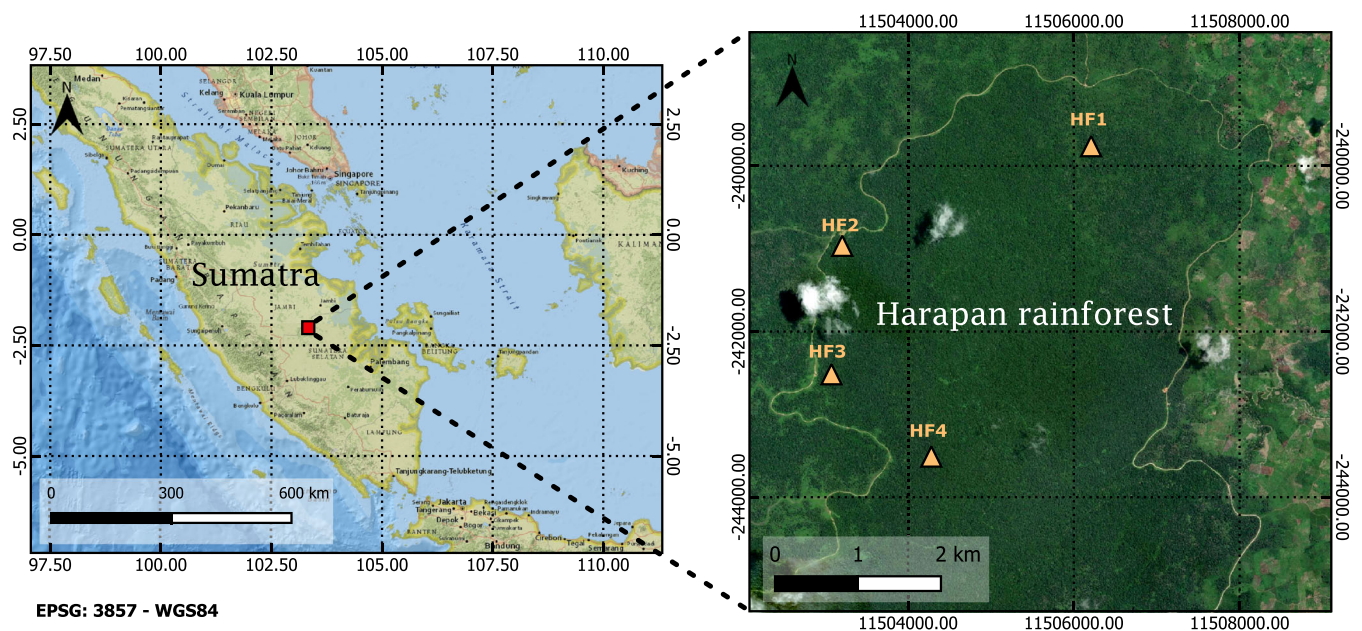


FIGURE 1 Study region in Jambi province, Sumatra, Indonesia (a) and location of the four lowland rainforest study plots (HF1, HF2, HF3, HF4) in the Harapan rainforest (b).

and average DBH were similar among the plots, ranging from 616 to 728 trees ha⁻¹ and 19.5 to 22.0 cm, respectively (Kotowska et al., 2015).

2.2 | Thermal and RGB image acquisition and processing

At each of the four study plots, an octocopter UAV (MK EASY Okto V3, HiSystems, Germany) was used for data acquisition. The UAV was equipped with a radiometric thermal sensor FLIR Tau 2 640 (FLIR Systems, USA) and a RGB camera with an Omnivision OV12890 CMOS sensor (Omnivision, USA). Both cameras were simultaneously mounted in a stereo configuration and on a gimbal to ensure nadir view. GPS location was recorded with an onboard GPS device (MKBNSS V3 GPS/Glonass/Beidou, HiSystems, Germany). Flight planning was conducted using the MikroKopter-Tool V2.14b software. Flight paths were designed as a series of overlapping circular and grid patterns to ensure high overlap (>80%) between images. Flight altitude was set to 80 m above ground, which corresponds to approximately 30–40 m above the canopy. One flight was carried out over each plot in August 2017, always close to noon and preferably during clear-sky conditions. At the respective times of the data capture, air temperature ranged from 31 to 38°C, and short-wave radiation from 700 to 935 W·m⁻² (Stiegler, 2021; Table S1).

The ThermoViewer 3.0.7 software (TeAX Technology, 2021) was used to delete blurry or non-relevant thermal images, to export the images to TIFF format and to convert the data to degrees Kelvin. To process the individual georeferenced RGB images (220–300 per site) and thermal images (330–370 images per site) to orthomosaic maps, AgiSoft MetaShape 1.8.2 (AgiSoft LLC, 2021) was used. It applies the SfM technique (Westoby et al., 2012) and further outputs digital elevation models (DEM) and RGB point clouds. We then used QGIS 3.16 (QGIS Development Team, 2020) to resample all RGB-derived raster layers to match the native 10 cm resolution of the thermal orthomosaics and to clip all layers to the respective plot boundaries (coordinates from Ahongshangbam et al., 2020).

2.3 | ET assessment via thermography

From the thermal orthomosaic maps of each of the four plots, ET was modelled with the QGIS plugin QWaterModel (Ellsäßer, Röhl, Stiegler, et al., 2020), which bases on the one-source energy balance model DATTUTDUT (Timmermans et al., 2015). We used the default settings of QWaterModel (version 1.3) in deriving ET. We used a previously established and tested modelling workflow to incorporate measured short-wave irradiance data into the modelling process (Ellsäßer et al., 2021; Ellsäßer, Röhl, Stiegler, et al., 2020) and also included measured air temperature data. For all four plots, we used in situ measured data from a nearby microclimatic station operated by the EFForTS project (Stiegler, 2021) where short-wave irradiance was measured in an open area (at 3 m height) with a CMP3 pyranometer

(Kipp & Zonen, The Netherlands) in 10 min intervals and air temperature was measured (at 2 m height) with a thermohygrometer (type 1.1025.55.000, Thies Clima, Göttingen, Germany); therein, we used the climate data that matched the time stamp of the respective UAV flight at each plot. From the thermal orthomosaic maps of the study plots and the short-wave radiation and air temperature data, QWater-Model then processes ET maps.

2.4 | Structural variables derived from photogrammetry

From the RGB orthomosaics and point clouds of each plot, several canopy-related variables were derived at the pixel level, in alignment with the 10 cm resolution of the thermal orthomosaics. From the RGB orthomosaics, we obtained the VARI (García-Martínez et al., 2020; Gitelson et al., 2002); (Equation 1) and the GLI (Louhaichi et al., 2001; Raymond Hunt et al., 2011) (Equation 2), which are measures of vegetation fraction based on normalized RGB bands.

$$\text{VARI} = (\text{Green} - \text{Red}) / (\text{Green} + \text{Red} - \text{Blue}) \quad (1)$$

$$\text{GLI} = (2 \times (\text{Green} - \text{Red} - \text{Blue})) / (2 \times (\text{Green} + \text{Red} + \text{Blue})) \quad (2)$$

From the RGB 3D point clouds, we obtained several height-related metrics (Table 1) at the pixel level with the *Gridmetrics* algorithm from the Fusion software (McGaughey, 2018), that is, the ‘relative height’ [Height_{rel}, the maximum z value among all voxels belonging to a given pixel in the two-dimensional (2D) plane, divided by the mean z value observed in the respective plot], the absolute minimum (Height_{abs_min}) and maximum z values (Height_{abs_max}) within a given pixel, the percentiles 25 (P25), 50 (P50), 75 (P75) and 90 (P90) of the z-value distribution within a given pixel as well as their coefficient of variation (CV_Height) and the percentage of points above mean height (PM). Further outputs of the algorithm include the canopy cover of a given pixel above 10 m (Cover10), the canopy relief ratio (CRR), which characterizes canopy shape (McGaughey, 2018) as well as point cloud density (PCD) per pixel. Pixel-level LAD was estimated with the *leafR* package (Almeida et al., 2019).

2.5 | Data analysis

Due to the non-simultaneous nature of the data acquisition and the therewith associated varying climatic conditions that can potentially affect ET, the analysis was carried out separately for the four study plots. The dataset for each plots consisted of up to 226,000 pixels, with information on the target variable ET and the 16 potential explanatory variables (Table 1) for each 10-cm pixel. We first examined the ET distributions at the four plots in violin plots.

For subsequent analysis, all pixels were removed from the according datasets that had at least one no data (NA) entry among the

TABLE 1 Variables from thermography, RGB orthomosaics, 3D point clouds and digital elevation models (DEM) at each of the four study plots.

Variable category	Abbreviation	Units	Variable	Source	Reference
Target variable	ET	mm h ⁻¹	Evapotranspiration	Thermal orthomosaic	Timmermans et al., 2015
RGB related	GLI	Dimensionless	Green leaf index	RGB orthomosaic	Louhaichi et al., 2001
RGB related	VARI	Dimensionless	Visual atmospherically resistance index	RGB orthomosaic	Gitelson et al., 2002
Vegetation height	Height_abs_max	m	Maximum canopy height	Point cloud	McGaughey, 2018
	Height_abs_min	m	Minimum canopy height	Point cloud	McGaughey, 2018
	Height_abs_sm	m	Absolute canopy height obtained from the smoothed digital elevation model (DEM)	DEM	McGaughey, 2018
	Height_rel	Dimensionless	Relative canopy height	Point cloud	McGaughey, 2018
	P25	m	25th percentile	Point cloud	McGaughey, 2018
	P50	m	50th percentile	Point cloud	McGaughey, 2018
	P75	m	75th percentile	Point cloud	McGaughey, 2018
Vegetation height variability	P90	m	90th percentile	Point cloud	McGaughey, 2018
	CRR	Dimensionless	Canopy relief ratio calculated as: (mean (h) – min (h))/(max (h) – min (h)), where h is canopy height	Point cloud	McGaughey, 2018
	CV_height	%	Coefficient of variance of height distribution	Point cloud	McGaughey, 2018
Vegetation density	Cover10	%	Canopy cover of a given pixel above 10 m	Point cloud	McGaughey, 2018
	LAD	m ² .m ⁻³	Leaf area density	Point cloud	Almeida et al., 2019
	PCD	Points	Point cloud density: number of points in each pixel	Point cloud	McGaughey, 2018
	PM	%	Percentage of points above mean height	Point cloud	McGaughey, 2018

17 variables. As an initial step, we computed correlation matrices among all variables with the R package *corrplot* (Wei & Simko, 2021). To examine the influence of single variables on pixel-level ET, we applied linear regressions between ET and each of the 16 potential predictor variables, separately for each study plot. For visualization in scatterplots, we selected the variables LAD and maximum vegetation height (Height_abs_max) due to their importance for modelling forest structure (Andersen et al., 2005; McGaughey, 2018; Mutlu, 2006) and their reported strong influence on ET in studies at larger spatial scales (Coronel et al., 2011).

For predicting pixel-level ET from multiple variables, we applied a random forest approach, therein closely following best-practice methodology as outlined in previous studies (e.g. Valdés-Urbe et al., 2023). To consider spatial autocorrelation and avoid overfitting, we implemented the following approach: (i) The assignation of spatial ID to 1 × 1 m blocks (Roberts et al., 2017), (ii) The random selection of spatial predictors through FFS and *k-fold* leave-location-out cross validation (LLO-CV; Meyer et al., 2018, 2019) with the *CAST* R package (Meyer et al., 2021) and (iii) random forest modelling and *k-fold* LLO-CV with the *caret* package (Kuhn, 2021). First, using the R package ‘blockCV’ (Valavi et al., 2018), we partitioned each plot into spatial blocks of 1 m² (the minimum area accepted by the function). For every pixel inside that block, a distinct ID number between 1 and

5 was allocated, which corresponds to the desired number of cross-validation folds (set to 5 by default) (Roberts et al., 2017). Then, each dataset was split randomly into 60% training and 40% model validation data before feature selection. To reduce computation times for the FFS, we subsampled 50,000 pixels from the training datasets through stratified random sampling (Ludwig et al., 2019). The subsamples were divided by *k-fold* = 5 for spatial cross validation (Roberts et al., 2017). At each split, the maximum combination of predictor variables (*mtry*) was set to two. The FFS algorithm is based on pairing feature combinations, wherein it saves the best initial model and adds additional features through LLO-CV improvement detection. Thus, the model improves incrementally until no further decreases in root mean square errors (RMSE) are detected (Meyer et al., 2018). For each plot, the combination of features selected by the FFS was used as input for the final random forest with *k-fold* LLO-CV and 1000 trees (following Valdés-Urbe et al., 2023). We performed model validation by predicting over the 40% testing dataset. To evaluate model performance, we calculated RMSE and *R*² as an indicator of the variance explained by each model. We examined variable importance (in % of contribution to final model outcome) to assess the contribution of each feature at each plot using the *varImp* wrapper function from *caret* R package (Kuhn, 2021).

For all statistical analyses, R 4.3.0 (R Core Team, 2023) was used.

3 | RESULTS

3.1 | Spatial ET patterns

ET was highly variable among the pixels of a given plot and among the four study plots. At similar times of day, pixel-level ET varied from 0.0 to 1.2 mm h⁻¹ across the plots. Pixel-level minima of ET were below 0.1 mm h⁻¹ at all four plots, while the respective maxima were larger than 1.0 mm h⁻¹, resulting in at least 15-fold within-plot variations in ET (Figures 2 and S1). The within-plot variation of ET expressed as the coefficient of variation (CV) was similar among the plots, ranging from 16.2% to 18.9%. Mean ET values among the four plots varied 1.7-fold, ranging from 0.5 to 0.8 mm h⁻¹ (Figure S2). The higher ET values at the plots HF2 and HF4 correspond to higher air temperature and short-wave radiation at the time of overflight than at HF1 and HF3 (Table S1).

3.2 | Prediction of ET

Among the 16 canopy variables derived from RGB orthomosaics and point clouds, correlations with the target variable ET were very weak

across all plots ($R = -0.21$ – 0.21 ; Figure S3). Initial speculations that pixel-level variations of ET may be largely driven by vegetation height, its variability or vegetation density (incl. leaf density) were not confirmed at the level of single variables when applying linear regressions. Therein, although often significant ($P < 0.05$), the models explained only marginal amounts of the observed variance in pixel-level ET ($R^2 < 0.05$; Table S2), that is, there was no suitable single predictor of pixel-level ET among the studied variables.

The FFS approach for the random forest modelling resulted in the selection of 9–10 out of the 16 available variables per plot. The selection differed across plots (see Figure 3), but five variables related to vegetation height (P25, P50, P75, P90, Height_abs_min) were consistently present as predictors in all plots. Three further variables related to vegetation height (Height_abs_max), height variability (CV_Height) and reflectance (GLI) were present in three out of the four models. Using the final sets of variables for each plot for pixel-level ET prediction with a random forest approach resulted in fair to good model performance, with RMSE between 0.04 and 0.09 mm h⁻¹ and 56% (HF1), 64% (HF3 and HF4) and 65% (HF2) of the observed variance in ET explained by the prediction models. Model performance was similar for training and prediction datasets across all plots (Table 2), indicating that

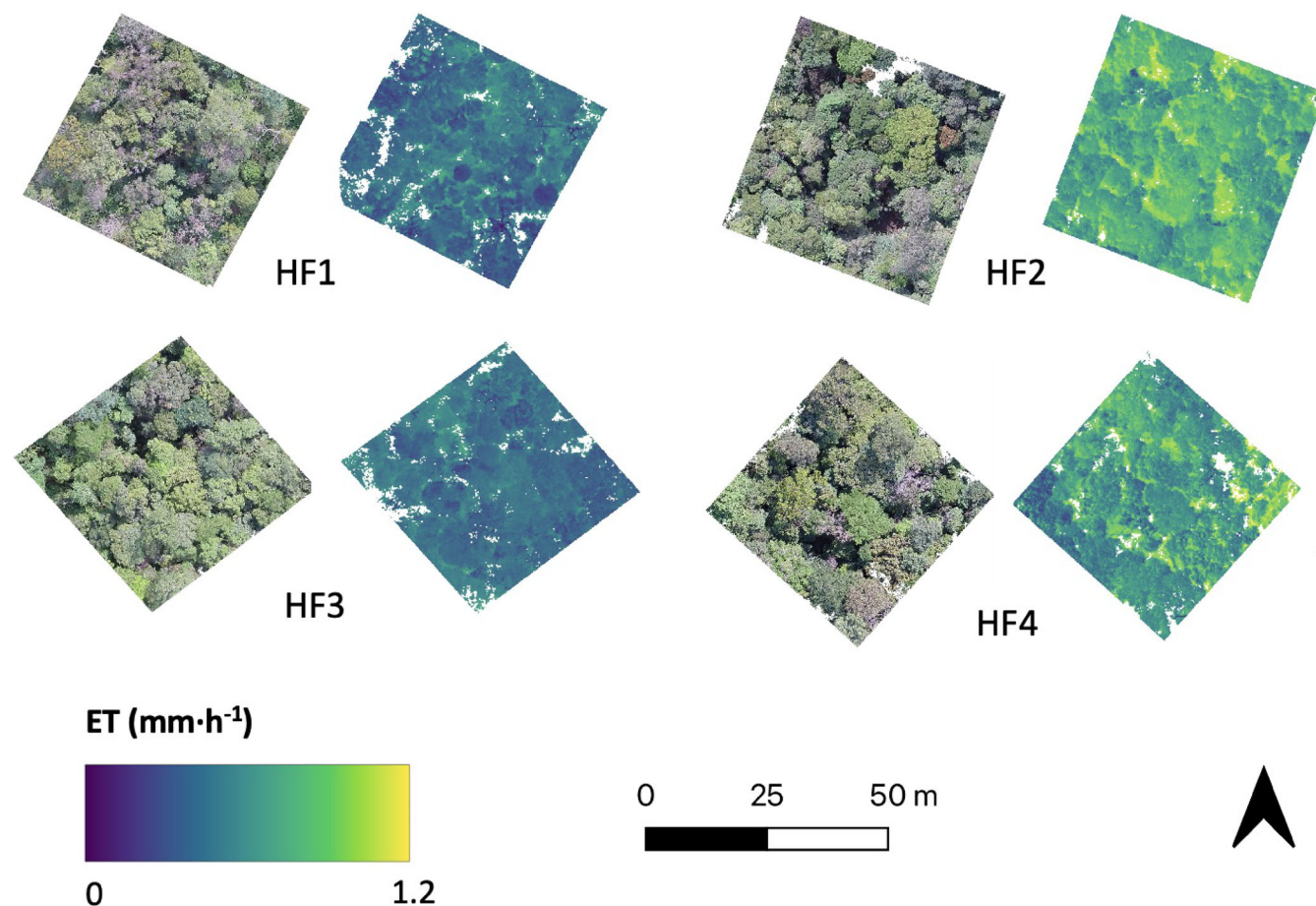


FIGURE 2 RGB (left) and evapotranspiration (ET) orthomosaics (right) for the four study plots. Hourly ET (mm h⁻¹) as indicated in the legend was derived from UAV-based thermal images with the plugin QWaterModel (Ellsäßer, Röhl, Stiegler, et al., 2020). The flights were carried out close to noon on successive days in August 2017. Air temperatures at the respective times of flight were 31.5°C (HF1), 36.5°C (HF2), 32.2°C (HF3) and 38.4°C (HF4), and short-wave radiation was at 704 W m⁻² (HF1), 936 W m⁻² (HF2), 858 W m⁻² (HF3) and 893 W m⁻² (HF4).

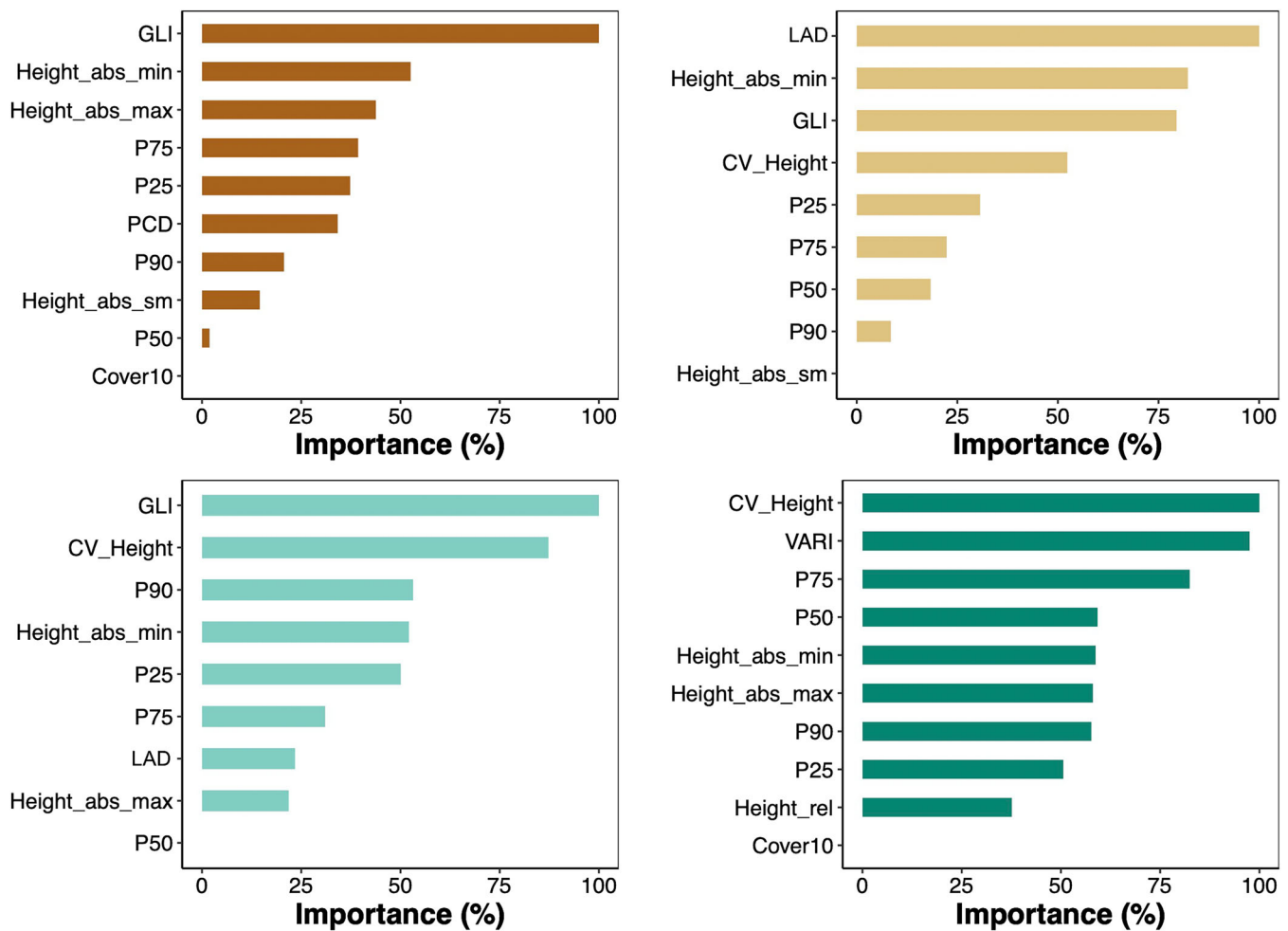


FIGURE 3 Variable importance (%) for the random forest model outcome at the four study plots. Refer to Table 1 for abbreviations and further variable details across the variable categories vegetation height (Height_abs_min, Height_abs_max, Height_abs_sm, P25, P50, P75, P90), height variability (CV_Height, Cover10), vegetation density (PCD, LAD) and reflectance (GLI, VARI).

TABLE 2 Performance of random forest models at the four study plots. The variables as selected with the forward feature selection approach were applied. The total amount of pixels at a given site was randomly split into training (60%) and prediction (40%) datasets.

Plot	Total number of pixels	Training models		Prediction models	
		RMSE (mm h ⁻¹)	R ²	RMSE (mm h ⁻¹)	R ²
HF1	169,248	0.03	0.56	0.06	0.56
HF2	167,857	0.04	0.65	0.07	0.65
HF3	211,345	0.02	0.64	0.04	0.64
HF4	169,887	0.05	0.65	0.09	0.64

Note: The number of pixels varied across plots and differs from the presented ET pixel numbers in Figure S2 because of no data value removal (7%–23% of pixels removed).
Abbreviation: RMSE, root mean square error.

spatial overfitting did not occur and that the models can predict ET for pixels that were not part of the respective training data.

Variable importance for the random forest model outcome differed substantially across the four plots (Figure 3), with no clear dominance of any variable or variable category but some observed common patterns. One to three variables per plot were of very high importance (>75%), and another one to four variables of high

importance (50%–75%). Variables of high or very high importance came from all categories, that is, vegetation height, height variability, vegetation density and reflectance. Specifically, the variables GLI (in three plots), CV_Height (two plots) and Height_abs_min, LAD, VARI and P75 (one plot each) were of very high importance for model outcome, and further variables of high importance in at least one of the plots included P50, P90 and Height_abs_max.

4 | DISCUSSION

4.1 | Spatial ET patterns

We assessed pixel-level ET patterns across four study plots in a tropical lowland rainforest with a UAV-based thermography approach and subsequent energy balance modelling. The therein applied energy balance model DATTUTDUT (Timmermans et al., 2015) and the associated workflow from image acquisition to high-resolution ET orthomosaic maps had previously been established and validated against ground-based reference measurements such as eddy covariance and sap flux measurements across different land-use systems in the same study region (Ellsäßer, Röhl, Ahongshangbam, et al., 2020; Ellsäßer, Röhl, Stiegler, et al., 2020; Ellsäßer et al., 2021) and including study sites in lowland rainforest (Bulusu et al., 2023). The mean mid-day hourly ET rates for our lowland rainforest study plots under mostly cloud-free conditions ($0.5\text{--}0.8\text{ mm h}^{-1}$) are similar to values from a nearby commercial, mature oil palm plantation, with mean ET of 0.4 mm h^{-1} and peaks up to 0.9 mm h^{-1} (Ellsäßer et al., 2021). This is in line with previous findings that stand-scale (evapo)transpiration of oil palm plantations can match or even surpass rates observed in (previously logged) lowland rainforests (Röhl et al., 2019; Tarigan et al., 2020). Somewhat lower diurnal ET peak values (around 0.5 mm h^{-1}) than in our study were, for example, reported for temperate ecosystems such as grasslands and coniferous forests (Kelliher et al., 1993). Overall, the mean ET values in our study are within expectation, giving credibility to the applied methods and to the observed spatial ET patterns.

At the coarsest spatial scale in our study, that is, when looking at plot-to-plot differences in ET, mean ET was approx. 50% higher in the study plots HF2 and HF4 than in HF1 and HF3. The higher ET values for HF2 and HF4 are partially driven by the climatic conditions at the time of data recording, that is, air temperature was on average 18% higher and short-wave radiation was 17% higher than for HF1 and HF3 (Table S1). However, the site-to-site differences in ET are higher than the differences in temperature and radiation, thus indicating spatial variability in mean ET after accounting for the non-simultaneous nature of the measurements and the associated climatic differences. Therein, while our results are consistent with previous sap flux-based transpiration assessments at the same four study sites in terms of substantial observed plot-to-plot variability (Ahongshangbam et al. 2020; Röhl et al., 2019), they do not match the previous reports of lower transpiration rates at HF4. However, HF4 has higher stand basal area, stand structural complexity, canopy cover and total biomass compared to the other three study plots (Drescher et al., 2016; Kotowska et al., 2015; Röhl et al., 2019), which is in line with the observed high ET at the plot in our study. Correspondingly, an enhanced UAV-based sap-flux scaling method suggests the highest stand transpiration rates at HF4, with 7%–13% lower transpiration at the other three plots (Ahongshangbam et al., 2020). Magnitude and patterns of plot-to-plot variability in ET in our study thus are mostly in line with previous studies, with some divergences. Those may be due to inaccuracies and uncertainties associated with all applied methods, which were

reported to be 30%–57% for estimating transpiration in tropical rainforests with ground-based sap-flux scaling schemes (Ahongshangbam et al., 2020; Granier et al., 1996; Röhl et al., 2019) and to be around 20% when scaling sap flux via UAV data-derived crown assessments (Ahongshangbam et al., 2020). While the uncertainties of the applied UAV-based ET method have not yet been quantified specifically for lowland rainforests, in a nearby oil palm plantation, a comparison to eddy covariance reference measurements indicated high agreement between the two methods (Ellsäßer et al., 2021).

At very fine spatial resolution ($\sim 10\text{ cm}$), we also found substantial differences in ET. At all four plots, pixel-level minima of ET were below 0.1 mm h^{-1} , while the respective maxima were larger than 1.0 mm h^{-1} . Taking the CV as a measure of within-plot variability in ET, we obtained values of 18.9% (HF1), 16.2% (HF2), 16.4% (HF3) and 16.9% (HF4). These values are lower than the reported within-plot variability of canopy transpiration (31% CV) at the same study plots, notably at a much coarser spatial resolution of 3 m (Ahongshangbam et al., 2020). Interestingly, when exemplarily resampling our ET maps to 3 m resolution, within-plot variability of ET remains lower (10.9%–22.0% CV) than reported for transpiration by Ahongshangbam et al. (2020). We are not aware of any further studies assessing rainforest (evapo)transpiration at such fine spatial resolution, that is, at the sub-canopy scale. At the level of individual trees, several studies also point to substantial spatial variability in tree-to-tree transpiration patterns due to differences in tree size (e.g. Kotowska et al., 2021; Meinzer et al., 2005), which introduces spatial variability in stand-scale (evapo)transpiration patterns in dependence of stand density, tree sizes and associated local canopy packing (Ahongshangbam et al., 2020). Overall, we consider the applied UAV-thermography-based scheme a powerful tool for the spatially explicit, high-resolution analysis and prediction of ET within and across sites. New generations of UAVs and imaging technology allow for further enhanced spatial resolutions over increasingly large areas, synchronized multi-site campaigns and fully integrated cameras that reduce operation effort and thereby contribute to better temporal coverage.

4.2 | Prediction of ET

We extracted 16 potential predictor variables via RGB image and point cloud analysis to further examine the drivers of small-scale ET patterns under given climatic conditions. The UAV-based image acquisition and the processing to georeferenced orthomosaics and 3D point clouds followed previously established and tested workflows (Ahongshangbam et al., 2020). In correlation and regression analyses, we found no clear influence of the studied predictors on pixel-level ET across the variable categories vegetation height, height variability, vegetation density and reflectance in any of the four study plots. While many of the studied variables had a significant linear relationship with ET ($P < 0.05$), the models explained less than 5% of the observed variance and thus are not suitable as single predictors of pixel-level ET. To our knowledge, there are no studies available for

comparison that examine the influence of forest structural variables on (rain)forest ET patterns at very small, sub-canopy spatial scales. A related UAV-based study in a tropical oil palm agroforest showed an important influence of canopy characteristics on pixel-level predictions of sap flux and stomatal conductance, but the role of structural variables for small-scale patterns was not further assessed (Ellsäßer, Röhl, Stiegler, et al., 2020).

From results of previous studies at larger spatial scales, that is, examining tree-to-tree and site-to-site patterns in rainforest (evapo)transpiration, one could well speculate that forest structural variables may also influence ET under given climatic conditions at very small spatial scales. As such, given previous results that tree-level transpiration mainly scales with tree size (e.g. Kotowska et al., 2021; Meinzer et al., 2005) and that taller vegetation generally has access to windier and more turbulent atmospheric exchange layers (Kricher, 2011), a pattern of increasing ET with increasing vegetation height would seem reasonable. However, the relationships between several available vegetation height variables and pixel-level ET instead suggested no clear influence of vegetation height variables on ET. Likewise, height variability variables such as the canopy relief ratio or the CV_Height had no clear influence on pixel-level ET, despite evidence from studies at larger spatial scales that increasing canopy heterogeneity (as, e.g. characterized by the variable canopy roughness length) leads to enhanced (evapo)transpiration (Bauerle et al., 2004; Lawrence et al., 2022; Tan et al., 2019). Similarly, the lack of a clear influence of vegetation density variables such as PCD and specifically the derived variable LAD on pixel-level ET stands out, with further studies at larger spatial scales reporting enhanced (evapo)transpiration at higher biomass and higher crown leaf density (Ruhoff et al., 2013; Tian et al., 2015).

Reasons for a lack of relationships between small-scale (pixel level) ET and structural canopy variables could be of ecological nature, due to methodological limitations, or a combination of both. From an ecohydrological point of view, it is well conceivable that a complex process such as rainforest ET, which is influenced by and interacting with a multitude of micrometeorological, pedological and forest structural processes, among others, is not explained by single structural variables, particularly at such small spatial scales as in our study. Possibly, the heterogeneity and dynamics within the rainforest canopy regarding micrometeorological conditions, tree species and age and associated biomass and leaf area (density), among other factors, are too large, and influences of the direct and indirect pixel-level 3D environment (e.g. understory, neighbouring pixels) are too prominent to see clear relationships between pixel-level ET and single structural variables. From a methods point of view, the applied UAV-based thermography method is limited to recording the uppermost canopy surface temperature of a given pixel, from which ET is subsequently modelled. It remains unclear how or to what degree ET of the multiple lower canopy strata microclimatically affects land surface temperatures of the recorded upper canopy and what levels of uncertainty this introduces into remotely sensed ET maps. It also remains disputed how high the contribution of lower canopy strata to (evapo)transpiration in tropical rainforests is in general, with estimates ranging from as

low as 20% to as high as 35% (Iida et al., 2020). While the problem is common to all above-canopy thermography methods, studies across several different ecosystems reported high congruence of UAV-derived ET estimates with ground-based reference measurements such as the eddy covariance technique (Aboutalebi et al., 2019; Brenner et al., 2017; Ellsäßer, Röhl, Ahongshangbam, et al., 2020), thus indicating the reliability of the method. Despite such methodological limitations, our study demonstrates the large potential of bringing together UAV-based thermography and photogrammetry methods for examining the relationship between ecosystem function (ET) and forest structure at very small spatial scales.

The lack of clear linear relationships between pixel-level ET and any of the studied canopy structural variables across all four study plots point to complex controls potentially involving non-linear relationships and interactions among multiple variables. Indeed, some previous studies reported nonlinearity in ET relationships (e.g. Boers et al., 2015; Zhou, 2011) and one recent study examining remotely sensed ET across 100,000 s of pixels also found no clear influences of single topographic, climatic or forest structure variables on ET across a large tropical forest region, concluding that accurate ET prediction is too complex for conventional statistical methods such as linear regressions (Valdés-Urbe et al., 2023). Machine learning approaches are a powerful tool to analyse such large datasets with complex relationships. Machine learning algorithms can be trained to predict ET (or other ecosystem processes) from available ancillary variables, therein often yielding much better prediction performance than conventional prediction methods (e.g. Camps-Valls, 2009; Maxwell et al., 2018; Meyer & Pebesma, 2022). Unlike conventional methods, machine learning models do not give information on predictor significance, effect direction or size; however, an analysis of model feature importance shows the importance of each variable for the overall prediction outcome. There are several machine learning algorithms that are widely applied throughout the environmental sciences (Dou & Yang, 2018; Kanevski, 2009; Lary et al., 2016). Random forest is one of the most popular algorithms and often performs well in ecological contexts (Ahmad et al., 2017; Fernandez-Delgado et al., 2014), for example, when predicting reference ET, water stress, sap flux, leaf stomatal conductance, net ecosystem exchange or land-cover change (Aide et al., 2013; Dias et al., 2021; Ellsäßer, Röhl, Ahongshangbam, et al., 2020; Feng et al., 2017; Reitz et al., 2021; Virnodkar et al., 2020). The random forest algorithm is particularly suitable for spatial predictions and was thus the machine learning algorithm of choice in our study. We applied state-of-the-art techniques such as FFS and target-oriented cross validation to reduce the risk of spatial overfitting (Gasch et al., 2015; Ludwig et al., 2019; Meyer et al., 2018, 2019; Roberts et al., 2017), therein achieving an overall realistic model performance.

In our study, the FFS reduced the dataset for pixel-level ET prediction from the original 16 variables to a final set of 9–10 variables per plot, with a relatively high consistency of selection, that is, eight variables (mostly related to vegetation height) occurred in at least three out of the four models. The similar model performance between training and testing (prediction) outcomes in the subsequent random

forest modelling suggests that spatial overfitting did not occur in the models and that they thus can reliably predict ET for locations (i.e. pixels) that were not part of the model training. Overall, the performance of the random forest models was fair for HF1 and good for the other three study plots (HF2, HF3, HF4), with RMSE between 0.03 (HF1) and 0.05 (HF4) and model prediction accuracies (R^2) between 0.56 (HF1) and 0.65 (HF4). These accuracy metrics are comparable to other studies applying random forest modelling for predictions of ecological target variables such as reference ET ($R^2 = 0.91$; Dias et al., 2021), daily ET ($R^2 = 0.63$; Yang et al., 2021), sap flux ($R^2 = 0.80$) and stomatal conductance ($R^2 = 0.50$; Ellsäßer, Röhl, Ahongshangbam, et al., 2020) as well as actual ET across a large tropical forest region ($R^2 = 0.64$; Valdés-Urbe et al., 2023).

Variable importance for the random forest model outcome differed across the four study plots, with no clear dominance of any variable or variable category. One to three variables per plot were of very high importance (>75% feature importance for final model outcome), and another one to four variables of high importance (50%–75%). Therein, variables of high or very high importance came from all variable categories (vegetation height, height variability, vegetation density, reflectance). The most prominently occurring variables of very high importance were GLI (reflectance) followed by CV_height (height variability), followed by Height_abs_min and P75 (vegetation height), LAD (vegetation density) and VARI (reflectance). To our knowledge, there are no studies available for comparison involving such a set of structural canopy variables for predicting ET under given climatic conditions at such small spatial scales. A random forest study predicting ET at much larger spatial scales (70 m pixel size) found a dominant influence of topographic variables, and to a lesser extent of forest structure variables such as the leaf area index, on spatial predictions of ET (Valdés-Urbe et al., 2023); while climatic variables were found to be of relatively minor importance in that study, some further previous studies, also at much larger spatial scales, successfully applied machine learning techniques to model ET (dynamics) from climatic datasets (Granata, 2019; Tikhamarine et al., 2019). Both the random forest ET study by Valdés-Urbe et al. (2023) and our study show varying variable importance across different studied days in tropical forests, which may (partially) be related to differences in the given climatic conditions. While in our study the reflectance variable GLI was for example of much higher variable importance on days receiving less solar irradiance (flights of plots HF1 and HF3) than on sunnier days (HF2 and HF4), the underlying mechanisms of vastly varying variable importance are not yet understood. We consider our study only a first step in understanding and modelling pixel-level ET patterns, with many challenges remaining for future studies on the way to holistic spatial ET models that perform well across all spatial and temporal scales.

We conclude that combining multiple canopy structure variables that can relatively easily be derived from RGB photogrammetry explain substantial parts of the observed variance in pixel-level ET across four lowland rainforest plots in our study but that much work remains to be done. With 36%–44% of ET variance remaining unexplained, future studies could further improve our understanding of

small-scale (rain)forest ET controls by, for example, including more sophisticated indices characterizing vegetation greenness and plant health (e.g. NDVI) derived from simultaneous high-resolution multi- or hyperspectral UAV imaging. Overall, our study adds to the available knowledge on ET and its drivers that small-scale ET patterns under given climatic conditions can well be predicted from structural variables alone when accounting for complex multiple non-linear relationships.

AUTHOR CONTRIBUTIONS

Álvaro Cortés-Molino, Dirk Hölscher, Hendrayanto and Alexander Röhl conceived the ideas and designed methodology; Florian Ellsäßer collected the data; Álvaro Cortés-Molino, Alejandra Valdés-Urbe, Florian Ellsäßer, Joyson Ahongshangbam, Medha Bulusu and Alexander Röhl analysed the data; Álvaro Cortés-Molino and Alexander Röhl led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

We would like to thank the Ministry of Research, Technology and Higher Education, Indonesia, for issuing the research permit for fieldwork (No. 285/SIP/FRP/E5/Dit.KI/VIII/2016 and No. 322/SIP/FRP/E5/Dit.KI/IX/2016).

We would also like to thank our field assistant Erwin Pranata for supporting us during the field campaigns and the EFForTS staff in general for their support over the years.

The study was financially supported by the Deutsche Forschungsgemeinschaft (DFG) (CRC990, sub-project A02).

Alvaro Cortés-Molino received a grant for the achievement of international mention in his doctoral thesis by the University of Málaga (2022).

We acknowledge support by the Open Access Publication Funds of the University of Göttingen.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon request. They are permanently stored in the data repository of the University of Göttingen 'GRO.data' and become available at [10.25625/IXCGWZ](https://doi.org/10.25625/IXCGWZ) once the study is published. This includes the R code applied for the analysis.

ORCID

Alexander Röhl  <https://orcid.org/0000-0001-9457-4459>

REFERENCES

- Aboutaleb, M., Torres-Rua, A. F., McKee, M., Kustas, W. P., Nieto, H., Alsina, M. M., White, A., Prueger, J. H., McKee, L., Alfieri, J., Hipps, L., Coopmans, C., & Dokoozlian, N. (2019). Incorporation of unmanned aerial vehicle (UAV) point cloud products into remote sensing evapotranspiration models. *Remote Sensing*, 12(1), 50. <https://doi.org/10.3390/rs12010050>
- Agisoft LLC. (2021). Agisoft Metashape.
- Ahmad, M. W., Mourshed, M., & Rezgui, Y. (2017). Trees vs neurons: Comparison between random forest and ANN for high-resolution

- prediction of building energy consumption. *Energy and Buildings*, 147, 77–89. <https://doi.org/10.1016/j.enbuild.2017.04.038>
- Ahongshangbam, J., Khokthong, W., Ellsäßer, F., Hendrayanto, H., Hölscher, D., & Röhl, A. (2019). Drone-based photogrammetry-derived crown metrics for predicting tree and oil palm water use. *Ecohydrology*, 12(6), Portico. <https://doi.org/10.1002/eco.2115>
- Ahongshangbam, J., Röhl, A., Ellsäßer, F., Hendrayanto, & Hölscher, D. (2020). Airborne tree crown detection for predicting spatial heterogeneity of canopy transpiration in a tropical rainforest. *Remote Sensing*, 12(4), 651. <https://doi.org/10.3390/rs12040651>
- Aide, T. M., Clark, M. L., Grau, H. R., López-Carr, D., Levy, M. A., Redo, D., Bonilla-Moheno, M., Riner, G., Andrade-Núñez, M. J., & Muñiz, M. (2013). Deforestation and reforestation of Latin America and the Caribbean (2001–2010). *Biotropica*, 45(2), 262–271. <https://doi.org/10.1111/j.1744-7429.2012.00908.x>
- Allan, R. P., Barlow, M., Byrne, M. P., Cherchi, A., Douville, H., Fowler, H. J., Gan, T. Y., Pendergrass, A. G., Rosenfeld, D., Swann, A. L. S., Wilcox, L. J., & Zolina, O. (2020). Advances in understanding large-scale responses of the water cycle to climate change. *Annals of the New York Academy of Sciences*, 1472(1), 49–75. <https://doi.org/10.1111/nyas.14337>
- Allen, R. G., Smith, M., Perrier, A., & Pereira, L. S. (1994). An update for the definition of reference evapotranspiration. *ICID Bulletin*, 43(2), 1–34.
- Almeida, D. R. A., Stark, S. C., Shao, G., Schietti, J., Nelson, B. W., Silva, C. A., Gorgens, E. B., Valbuena, R., Papa, D. A., & Brancalion, P. H. S. (2019). Optimizing the remote detection of tropical rainforest structure with airborne lidar: Leaf area profile sensitivity to pulse density and spatial sampling. *Remote Sensing*, 11(1), 92. <https://doi.org/10.3390/rs11010092>
- Andersen, H. E., McGaughey, R. J., & Reutebuch, S. E. (2005). Estimating forest canopy fuel parameters using LIDAR data. *Remote Sensing of Environment*, 94(4), 441–449. <https://doi.org/10.1016/j.rse.2004.10.013>
- Baldocchi, D. D. (2020). How eddy covariance flux measurements have contributed to our understanding of global change biology. *Global Change Biology*, 26(1), 242–260. <https://doi.org/10.1111/gcb.14807>
- Barichivich, J., Gloor, E., Peylin, P., Brienen, R. J. W., Schöngart, J., Espinoza, J. C., & Pattayak, K. C. (2018). Recent intensification of Amazon flooding extremes driven by strengthened Walker circulation. *Science Advances*, 4(9), eaat8785. <https://doi.org/10.1126/sciadv.aat8785>
- Bauerle, W. L., Bowden, J. D., McLeod, M. F., & Toler, J. E. (2004). Modeling intra-crown and intra-canopy interactions in red maple: Assessment of light transfer on carbon dioxide and water vapor exchange. *Tree Physiology*, 24(5), 589–597. <https://doi.org/10.1093/treephys/24.5.589>
- Boers, N., Marwan, N., Barbosa, H., & Kurths, J. (2015). How Amazonian deforestation can alter the South American circulation regime: Insights from a non-linear moisture transport model. 10922. <https://ui.adsabs.harvard.edu/abs/2015EGUGA..1710922B>
- Brenner, C., Thiem, C. E., Wizemann, H.-D., Bernhardt, M., & Schulz, K. (2017). Estimating spatially distributed turbulent heat fluxes from high-resolution thermal imagery acquired with a UAV system. *International Journal of Remote Sensing*, 38(8–10), 3003–3026. <https://doi.org/10.1080/01431161.2017.1280202>
- Bulusu, M., Ellsäßer, F., Stiegler, C., Ahongshangbam, J., Marques, I., Hendrayanto, H., Röhl, A., & Hölscher, D. (2023). UAV-based thermography reveals spatial and temporal variability of evapotranspiration from a tropical rainforest. *Frontiers in Forests and Global Change*, 6, 1232410. <https://doi.org/10.3389/ffgc.2023.1232410>
- Camps-Valls, G. (2009). Machine learning in remote sensing data processing. 2009 IEEE International Workshop on Machine Learning for Signal Processing, 1–6. <https://doi.org/10.1109/MLSP.2009.5306233>
- Clough, Y., Krishna, V. V., Corre, M. D., Darras, K., Denmead, L. H., Meijide, A., Moser, S., Musshoff, O., Steinebach, S., Veldkamp, E., Allen, K., Barnes, A. D., Breidenbach, N., Brose, U., Buchori, D., Daniel, R., Finkeldey, R., Harahap, I., Hertel, D., ... Scheu, S. (2016). Land-use choices follow profitability at the expense of ecological functions in Indonesian smallholder landscapes. *Nature Communications*, 7(1), 13137. <https://doi.org/10.1038/ncomms13137>
- Coronel, C., Tapia Silva, O., Hernández, G., Madrigal, J. M., Rosales, E., Toledo, A., Galeana, M., López Caloca, A., & Silvan, J. L. (2011). Conceptual elements and heuristics from complexity paradigm suitable to the study of evapotranspiration at the landscape level. In L. Łabędzki (Ed.), *Evapotranspiration*. InTech. <https://doi.org/10.5772/14056>
- Dias, S. H. B., Filgueiras, R., Fernandes Filho, E. I., Arcanjo, G. S., Silva, G. H. D., Mantovani, E. C., & Cunha, F. F. D. (2021). Reference evapotranspiration of Brazil modeled with machine learning techniques and remote sensing. *PLoS ONE*, 16(2), e0245834. <https://doi.org/10.1371/journal.pone.0245834>
- Dimitriadou, S., & Nikolakopoulos, K. G. (2021). Evapotranspiration trends and interactions in light of the anthropogenic footprint and the climate crisis: A review. *Hydrology*, 8(4), 163. <https://doi.org/10.3390/hydrology8040163>
- Donat, M. G., Angéilil, O., & Ukkola, A. M. (2019). Intensification of precipitation extremes in the world's humid and water-limited regions. *Environmental Research Letters*, 14(6), 065003. <https://doi.org/10.1088/1748-9326/ab1c8e>
- Dou, X., & Yang, Y. (2018). Evapotranspiration estimation using four different machine learning approaches in different terrestrial ecosystems. *Computers and Electronics in Agriculture*, 148, 95–106. <https://doi.org/10.1016/j.compag.2018.03.010>
- Drescher, J., Rembold, K., Allen, K., Beckschäfer, P., Buchori, D., Clough, Y., Faust, H., Fauzi, A. M., Gunawan, D., Hertel, D., Irawan, B., Jaya, I. N. S., Klarner, B., Kleinn, C., Knohl, A., Kotowska, M. M., Krashevskaya, V., Krishna, V., Leuschner, C., ... Scheu, S. (2016). Ecological and socio-economic functions across tropical land use systems after rainforest conversion. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 371(1694), 20150275. <https://doi.org/10.1098/rstb.2015.0275>
- Ellsäßer, F., Röhl, A., Ahongshangbam, J., Waite, P. A., Hendrayanto, S., Schuldt, B., & Hölscher, D. (2020). Predicting tree sap flux and stomatal conductance from drone-recorded surface temperatures in a mixed agroforestry system—a machine learning approach. *Remote Sensing*, 12(24), 4070. <https://doi.org/10.3390/rs12244070>
- Ellsäßer, F., Röhl, A., Stiegler, C., Hendrayanto, & Hölscher, D. (2020). Introducing QWaterModel, a QGIS plugin for predicting evapotranspiration from land surface temperatures. *Environmental Modelling & Software*, 130, 104739. <https://doi.org/10.1016/j.envsoft.2020.104739>
- Ellsäßer, F., Stiegler, C., Röhl, A., June, T., Hendrayanto, K., Knohl, A., & Hölscher, D. (2021). Predicting evapotranspiration from drone-based thermography—a method comparison in a tropical oil palm plantation. *Biogeosciences*, 18(3), 861–872. <https://doi.org/10.5194/bg-18-861-2021>
- Feng, Y., Cui, N., Gong, D., Zhang, Q., & Zhao, L. (2017). Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling. *Agricultural Water Management*, 193, 163–173. <https://doi.org/10.1016/j.agwat.2017.08.003>
- Fernandez-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research*, 15, 3133–3181.
- Fisher, J. B., Lee, B., Purdy, A. J., Halverson, G. H., Dohlen, M. B., Cawse-Nicholson, K., Wang, A., Anderson, R. G., Aragon, B., Arain, M. A., Baldocchi, D. D., Baker, J. M., Barral, H., Bernacchi, C. J., Bernhofer, C., Biraud, S. C., Bohrer, G., Brunsell, N., Cappelaere, B., ... Hook, S. (2020). ECOSTRESS: NASA's next generation mission to measure evapotranspiration from the international space station. *Water Resources Research*, 56(4), 1–20. <https://doi.org/10.1029/2019WR026058>
- García-Martínez, H., Flores-Magdaleno, H., Ascencio-Hernández, R., Khalil-Gardezi, A., Tijerina-Chávez, L., Mancilla-Villa, O. R., & Vázquez-

- Peña, M. A. (2020). Corn grain yield estimation from vegetation indices, canopy cover, plant density, and a neural network using multi-spectral and rgb images acquired with unmanned aerial vehicles. *Agriculture (Switzerland)*, 10(7), 277. <https://doi.org/10.3390/agriculture10070277>
- Gasch, C. K., Hengl, T., Gräler, B., Meyer, H., Magney, T. S., & Brown, D. J. (2015). Spatio-temporal interpolation of soil water, temperature, and electrical conductivity in 3D + T: The Cook Agronomy farm data set. *Spatial Statistics*, 14, 70–90. <https://doi.org/10.1016/j.spasta.2015.04.001>
- Gitelson, A. A., Kaufman, Y. J., Stark, R., & Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. <https://digitalcommons.unl.edu/natrespapers/149>
- Granata, F. (2019). Evapotranspiration evaluation models based on machine learning algorithms—A comparative study. *Agricultural Water Management*, 217, 303–315. <https://doi.org/10.1016/j.agwat.2019.03.015>
- Granier, A., Biron, P., Breda, N., Pontallier, J.-Y., & Saugier, B. (1996). Transpiration of trees and forest stands: Short and long-term monitoring using sapflow methods. *Global Change Biology*, 2(3), 265–274. <https://doi.org/10.1111/j.1365-2486.1996.tb00078.x>
- Harrison, R. D., & Swinfield, T. (2015). Restoration of logged humid tropical forests: An experimental programme at Harapan rainforest, Indonesia. *Tropical Conservation Science*, 8(1), 4–16. <https://doi.org/10.1177/194008291500800103>
- Iglhaut, J., Cabo, C., Puliti, S., Piermattei, L., O'Connor, J., & Rosette, J. (2019). Structure from motion photogrammetry in forestry: A review. *Current Forestry Reports*, 5(3), 155–168. <https://doi.org/10.1007/s40725-019-00094-3>
- Iida, S., Shimizu, T., Tamai, K., Kabeya, N., Shimizu, A., Ito, E., Ohnuki, Y., Chann, S., & Levia, D. F. (2020). Evapotranspiration from the understory of a tropical dry deciduous forest in Cambodia. *Agricultural and Forest Meteorology*, 295, 108170. <https://doi.org/10.1016/j.agrformet.2020.108170>
- Kanevski, M. (2009). *Machine learning for spatial environmental data: Theory, applications, and software*. EPFL Press. [10.1201/9781439808085](https://doi.org/10.1201/9781439808085)
- Kelliher, F. M., Leuning, R., & Schulze, E. D. (1993). Evaporation and canopy characteristics of coniferous forests and grasslands. *Oecologia*, 95(2), 153–163. <https://doi.org/10.1007/BF00323485>
- Kotowska, M. M., Leuschner, C., Triadiati, T., Meriem, S., & Hertel, D. (2015). Quantifying above- and belowground biomass carbon loss with forest conversion in tropical lowlands of Sumatra (Indonesia). *Global Change Biology*, 21(10), 3620–3634. <https://doi.org/10.1111/gcb.12979>
- Kotowska, M. M., Link, R. M., Röhl, A., Hertel, D., Hölscher, D., Waite, P.-A., Moser, G., Tjoa, A., Leuschner, C., & Schuldt, B. (2021). Effects of wood hydraulic properties on water use and productivity of tropical rainforest trees. *Frontiers in Forests and Global Change*, 3, 598759. <https://doi.org/10.3389/ffgc.2020.598759>
- Kricher, J. (2011). *Tropical ecology*. Princeton University Press.
- Kuhn, M. (2021). caret: Classification and regression training.
- Kundzewicz, Z. W. (2008). Climate change impacts on the hydrological cycle. *Ecohydrology & Hydrobiology*, 8(2–4), 195–203. <https://doi.org/10.2478/v10104-009-0015-y>
- Lary, D. J., Alavi, A. H., Gandomi, A. H., & Walker, A. L. (2016). Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, 7(1), 3–10. <https://doi.org/10.1016/j.gsf.2015.07.003>
- Lawrence, D., Coe, M., Walker, W., Verchot, L., & Vandecar, K. (2022). The unseen effects of deforestation: Biophysical effects on climate. *Frontiers in Forests and Global Change*, 5, 756115. <https://doi.org/10.3389/ffgc.2022.756115>
- Louhaichi, M., Borman, M. M., & Johnson, D. E. (2001). Spatially located platform and aerial photography for documentation of grazing impacts on wheat. *Geocarto International*, 16(1), 65–70. <https://doi.org/10.1080/10106040108542184>
- Ludwig, M., Morgenthal, T., Detsch, F., Higginbottom, T. P., Lezama Valdes, M., Nauß, T., & Meyer, H. (2019). Machine learning and multi-sensor based modelling of woody vegetation in the Molopo area, South Africa. *Remote Sensing of Environment*, 222, 195–203. <https://doi.org/10.1016/j.rse.2018.12.019>
- Marques, T. V., Mendes, K., Mutti, P., Medeiros, S., Silva, L., Perez-Marin, A. M., Campos, S., Lúcio, P. S., Lima, K., dos Reis, J., Ramos, T. M., da Silva, D. F., Oliveira, C. P., Costa, G. B., Antonino, A. C. D., Menezes, R. S. C., Santos e Silva, C. M., & Bezerra, B. (2020). Environmental and biophysical controls of evapotranspiration from Seasonally Dry Tropical Forests (Caatinga) in the Brazilian Semi-arid. *Agricultural and Forest Meteorology*, 287, 107957. <https://doi.org/10.1016/j.agrformet.2020.107957>
- Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, 39(9), 2784–2817. <https://doi.org/10.1080/01431161.2018.1433343>
- McGaughey, R. (2018). FUSION/LDV: Software for LIDAR data analysis and visualization.
- McGaughey, R. J., & Carson, W. W. (2003). Fusing LIDAR data, photographs and other data using 2D and 4D visualization techniques. *Proceedings of Terrain Data: Applications and Visualization-Making the Connection*, 16–24.
- Meinzer, F. C., Bond, B. J., Warren, J. M., & Woodruff, D. R. (2005). Does water transport scale universally with tree size? *Functional Ecology*, 19(4), 558–565. <https://doi.org/10.1111/j.1365-2435.2005.01017.x>
- Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., & Nauss, T. (2018). Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environmental Modelling & Software*, 101, 1–9. <https://doi.org/10.1016/j.envsoft.2017.12.001>
- Meyer, H., Reudenbach, C., Wöllauer, S., & Nauss, T. (2019). Importance of spatial predictor variable selection in machine learning applications – Moving from data reproduction to spatial prediction. *Ecological Modelling*, 411, 108815. <https://doi.org/10.1016/j.ecolmodel.2019.108815>
- Meyer, H., & Pebesma, E. (2022). Machine learning-based global maps of ecological variables and the challenge of assessing them. *Nature Communications*, 13(1). <https://doi.org/10.1038/s41467-022-29838-9>
- Meyer, H., Reudenbach, C., Ludwig, M., Nauss, T., Pebesma, E., & Meyer, M. H. (2021). Package ‘CAST’. R package version 0.5. 2021 Apr 7;1. <https://hannameyer.github.io/CAST/>
- Mutlu, M. (2006). Mapping surface fuels using lidar and multispectral data fusion for fire behavior modeling (Número December). Texas A&M University.
- Nazarbakhsh, M., Ireson, A. M., & Barr, A. G. (2019). Controls on evapotranspiration from jack pine forests in the Boreal Plains Ecozone. *Hydrological Processes*, 34(4), 927–940. Portico. <https://doi.org/10.1002/hyp.13674>
- Nelson, J. A., Pérez-Priego, O., Zhou, S., Poyatos, R., Zhang, Y., Blanken, P. D., Gimeno, T. E., Wohlfahrt, G., Desai, A. R., Gioli, B., Limousin, J., Bonal, D., Paul-Limoges, E., Scott, R. L., Varlagin, A., Fuchs, K., Montagnani, L., Wolf, S., Delpierre, N., ... Jung, M. (2020). Ecosystem transpiration and evaporation: Insights from three water flux partitioning methods across FLUXNET sites. *Global Change Biology*, 26(12), 6916–6930. <https://doi.org/10.1111/gcb.15314>
- QGIS Development Team. (2020). QGIS geographical information system. Open Source Geospatial Foundation Project.
- R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Rauneker, P., & Lischeid, G. (2012). Spatial distribution of water stress and evapotranspiration estimates using an unmanned aerial vehicle (UAV). 10477. <https://ui.adsabs.harvard.edu/abs/2012EGUGA..1410477R>

- Raymond Hunt, E., Daughtry, C. S. T., Eitel, J. U. H., & Long, D. S. (2011). Remote sensing leaf chlorophyll content using a visible band index. *Agronomy Journal*, 103(4), 1090–1099. <https://doi.org/10.2134/ agronj2010.0395>
- Reitz, O., Graf, A., Schmidt, M., Ketzler, G., & Leuchner, M. (2021). Upscaling net ecosystem exchange over heterogeneous landscapes with machine learning. *Journal of Geophysical Research - Biogeosciences*, 126(2), e2020JG005814. <https://doi.org/10.1029/2020JG005814>
- Rembold, K., Mangopo, H., Tjitrosoedirdjo, S. S., & Kreft, H. (2017). Plant diversity, forest dependency, and alien plant invasions in tropical agricultural landscapes. *Biological Conservation*, 213, 234–242. <https://doi.org/10.1016/j.biocon.2017.07.020>
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillaera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929. <https://doi.org/10.1111/ecog.02881>
- Röll, A., Niu, F., Meijide, A., Ahongshangbam, J., Ehbrecht, M., Guillaume, T., Gunawan, D., Hardanto, A., Hendrayanto, H., Hertel, D., Kotowska, M. M., Kreft, H., Kuzyakov, Y., Leuschner, C., Nomura, M., Polle, A., Rembold, K., Sahnner, J., Seidel, D., ... Hölscher, D. (2019). Transpiration on the rebound in lowland Sumatra. *Agricultural and Forest Meteorology*, 274, 160–171. <https://doi.org/10.1016/j.agrformet.2019.04.017>
- Ruhoff, A. L., Paz, A. R., Aragao, L. E. O. C., Mu, Q., Malhi, Y., Collischonn, W., Rocha, H. R., & Running, S. W. (2013). Assessment of the MODIS global evapotranspiration algorithm using eddy covariance measurements and hydrological modelling in the Rio Grande basin. *Hydrological Sciences Journal*, 58(8), 1658–1676. <https://doi.org/10.1080/02626667.2013.837578>
- Schlesinger, W. H., & Jasechko, S. (2014). Transpiration in the global water cycle. *Agricultural and Forest Meteorology*, 189–190, 115–117. <https://doi.org/10.1016/j.agrformet.2014.01.011>
- Stiegler, C. (2021). A03_Reki Meteo Data 2017 (2) [data set]. Göttingen Research Online/Data. <https://doi.org/10.25625/CXMO30>
- Tan, Z.-H., Zhao, J.-F., Wang, G.-Z., Chen, M.-P., Yang, L.-Y., He, C.-S., Restrepo-Coupe, N., Peng, S.-S., Liu, X.-Y., Da Rocha, H. R., Kosugi, Y., Hirano, T., Saleska, S. R., Goulden, M. L., Zeng, J., Ding, F.-J., Gao, F., & Song, L. (2019). Surface conductance for evapotranspiration of tropical forests: Calculations, variations, and controls. *Agricultural and Forest Meteorology*, 275, 317–328. <https://doi.org/10.1016/j.agrformet.2019.06.006>
- Tarigan, S., Stiegler, C., Wiegand, K., Knohl, A., & Murtlaksono, K. (2020). Relative contribution of evapotranspiration and soil compaction to the fluctuation of catchment discharge: Case study from a plantation landscape. *Hydrological Sciences Journal*, 65(7), 1239–1248. <https://doi.org/10.1080/02626667.2020.1739287>
- TeAX Technology. (2021). ThermalCapture.
- Tian, X., Van Der Tol, C., Su, Z., Li, Z., Chen, E., Li, X., Yan, M., Chen, X., Wang, X., Pan, X., Ling, F., Li, C., Fan, W., & Li, L. (2015). Simulation of forest evapotranspiration using time-series parameterization of the surface energy balance system (SEBS) over the Qilian mountains. *Remote Sensing*, 7(12), 15822–15843. <https://doi.org/10.3390/rs71215806>
- Tikhamarine, Y., Malik, A., Kumar, A., Souag-Gamane, D., & Kisi, O. (2019). Estimation of monthly reference evapotranspiration using novel hybrid machine learning approaches. *Hydrological Sciences Journal*, 64(15), 1824–1842. <https://doi.org/10.1080/02626667.2019.1678750>
- Timmermans, W. J., Kustas, W. P., & Andreu, A. (2015). Utility of an automated thermal-based approach for monitoring evapotranspiration. *Acta Geophysica*, 63(6), 1571–1608. <https://doi.org/10.1515/acgeo-2015-0016>
- Trenberth, K. E. (1999). Conceptual framework for changes of extremes of the hydrological cycle with climate change. In *Weather and climate extremes* (pp. 327–339). Springer. https://doi.org/10.1007/978-94-015-9265-9_18
- Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillaera-Arroita, G. (2018). blockCV: An R package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Methods in Ecology and Evolution*, 10, 225–232.
- Valdés-Urbe, A., Hölscher, D., & Röll, A. (2023). ECOSTRESS reveals the importance of topography and forest structure for evapotranspiration from a tropical region of the Andes. *Remote Sensing*, 15, 2985. <https://doi.org/10.3390/rs15122985>
- Virnodkar, S. S., Pachghare, V. K., Patil, V. C., & Jha, S. K. (2020). Remote sensing and machine learning for crop water stress determination in various crops: A critical review. *Precision Agriculture*, 21(5), 1121–1155. <https://doi.org/10.1007/s11119-020-09711-9>
- Wei, T., & Simko, V. (2021). R package «corrplot»: Visualization of a correlation matrix.
- Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). ‘Structure-from-motion’ photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, 179, 300–314. <https://doi.org/10.1016/j.geomorph.2012.08.021>
- Williams, C. A., Reichstein, M., Buchmann, N., Baldocchi, D., Beer, C., Schwalm, C., Wohlfahrt, G., Hasler, N., Bernhofer, C., Foken, T., Papale, D., Schymanski, S., & Schaefer, K. (2012). Climate and vegetation controls on the surface water balance: Synthesis of evapotranspiration measured across a global network of flux towers. *Water Resources Research*, 48(6), W06523. <https://doi.org/10.1029/2011WR011586>
- Yang, Y., Sun, H., Xue, J., Liu, Y., Liu, L., Yan, D., & Gui, D. (2021). Estimating evapotranspiration by coupling Bayesian model averaging methods with machine learning algorithms. *Environmental Monitoring and Assessment*, 193(3), 156. <https://doi.org/10.1007/s10661-021-08934-1>
- Yeh, P. J. -F., & Wu, C. (2018). Recent acceleration of the terrestrial hydrologic cycle in the U.S. Midwest. *Journal of Geophysical Research-Atmospheres*, 123(6), 2993–3008. <https://doi.org/10.1002/2017JD027706>
- Zhou, M. (2011). Estimates of evapotranspiration and their implication in the Mekong and Yellow River basins. In L. Łabędzki (Ed.), *Evapotranspiration*. InTech. <https://doi.org/10.5772/14791>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Cortés-Molino, Á., Valdés-Urbe, A., Ellsäßer, F., Bulusu, M., Ahongshangbam, J., Hendrayanto, Hölscher, D., & Röll, A. (2024). Combining UAV thermography, point cloud analysis and machine learning for assessing small-scale evapotranspiration patterns in a tropical rainforest. *Ecohydrology*, 17(1), e2604. <https://doi.org/10.1002/eco.2604>