PREDICTING EVAPOTRANSPIRATION FROM DRONE-RECORDED LAND SURFACE TEMPERATURES – METHOD TESTING AND DEVELOPMENT

Dissertation
to attain the doctoral degree Dr. rer. nat
of the Faculty of Forest Sciences and Forest Ecology
Georg-August-Universität Göttingen

Submitted by
Florian Johannes Ellsäßer
born on the 18th of May 1988 in Schwäbisch Hall

Göttingen, September 2020

1. Referee: Prof. Dr. Dirk Hölscher
2. Referee: Prof. Dr. Alexander Knohl

Date of oral examination: 20th of August 2020

TABLE OF CONTENTS

Abstract

Evapotranspiration (ET) is a central flux in the hydrological cycle on a regional and on a global scale. Transpiration from plants is the largest water flux from terrestrial surfaces, accounting for the major part of terrestrial ET. This thesis comprises method comparisons of methods for ET and plant-water-use analyses. Established and well-tested methods are compared with recently emerging methods based on drone remote sensing thermography and modelling approaches. The presented studies were developed and realized in the frame work of the Collaborative Research Centre 990 and the subproject A02. Based on the methods comparisons, a standardized workflow for ET prediction by land surface temperatures was implemented in a user friendly open-source software for ET calculation.

The primary goal of this thesis was to test and evaluate the new possibilities that result from the application of drones, radiometric thermal cameras as well as the utilization of causal and machine learning models. Therefore, method comparisons between well-known and tested reference methods and new drone-based methods are implemented. Another objective of this thesis was to streamline measurement efforts in the field and therefore evaluate the most important variables to measure to provide precise predictions. Based on the results of these objectives a standardized workflow for the calculation of ET based on thermal images from a variety of sources is developed and implemented into an open-source software available to a wide range of potential users.

For the first study, thermal images of land surface temperatures were recorded in 61 drone recording flights on five days over a commercially managed oil palm plantation. To predict ET from the thermal images three energy-balance-models (EBMs) were applied: The relatively simple one-source EBM DATTUTDUT (Timmermans et al., 2015) and the more complex two-source EBMs TSEB-PT (Norman et al., 1995) and DTD (Norman et al., 2000). Latent heat flux estimates of the DATTUTDUT model with measured net radiation agreed well with eddy covariance measurements (r^2=0.85; MAE=47; RMSE=60) across variable weather conditions and day times. Confidence intervals for slope and intercept

of a model II Deming regression suggest no difference between drone-based and eddy covariance method, thus indicating interchangeability. TSEB-PT and DTD yielded agreeable results, but all three models are sensitive to the configuration of the net radiation assessment. Overall, we conclude that drone-based thermography with energy-balance modeling is a reliable method complementing available methods for evapotranspiration studies. It offers promising, additional opportunities for fine grain and spatially explicit studies.

The second study focused on plant transpiration as a key element in the hydrological cycle. Widely used methods for its assessment comprise sap flux techniques for whole-plant transpiration and porometry for leaf stomatal conductance. Recently emerging approaches based on surface temperatures and a wide range of machine learning techniques offer new possibilities to quantify transpiration. The focus of this study was to predict sap flux and leaf stomatal conductance based on drone-recorded and meteorological data and compare these predictions with in-situ measured transpiration. Therefore, a comparatively large data set, consisting of 103 drone recording flights, two weeks of sap flux measurements in 10 min intervals and thousands of stomatal conductance measurements was recorded. The data collection was focused on oil palm, as well as on four local tree species that are native to the Sumatra region. Since no causal, but strictly data-driven machine-learning (ML) models were applied, such a large data set would be necessary for accurate prediction results. The models applied were a multiple linear regression that would serve as a simple base line method to compare with the more complex ML algorithms. The ML algorithms used in the study were a support vector machine (SVM), two types of random forest algorithms (RF) and an artificial neural network (ANN). Random forest predictions yielded the highest congruence with measured sap flux (r^2=0.87 for trees and r^2=0.58 for palms) and indicated differences among tree species. Confidence intervals for intercept and slope of a Passing-Bablok regression suggest interchangeability of methods for sap flux prediction using random forest. However, the other algorithms also showed promising results, especially for the prediction of sap flux from oil palm. Predictions for stomatal conductance were less congruent, likely due to spatial and temporal offsets of the measurements. Overall, the applied drone and modelling scheme predicts whole-plant transpiration with high accuracy, especially using random forest algorithms.

Various approaches to compute ET via energy balance models exist, but their handling is often complex and challenging. The prevalent aim of the third study was the development of a user friendly open-source software, that would implement central findings of the first two studies and that would be publicly available as an extension for the geospatial QGIS3 platform. Special emphasis was put

on the option to use thermal maps from a variety of sources including drones, satellites and handheld thermal cameras. As the previously mentioned studies showed, measured radiation and meteorological variables contribute significantly to the prediction accuracy. An option to input measured variables has therefore been added to the software. To test the performance of the software, land surface temperatures from an oil palm plantation were recorded using a drone and a handheld thermal cameras and radiation measurements were further used as optional input. Typical daily ET patterns were found with all model configurations for both recording types. However, the precision of the ET estimates by the software was significantly improved using solar radiation measurements. QWaterModel is compatible with all versions of QGIS3 and is available from the official QGIS Plugin Repository.

In summary, this thesis shows that evapotranspiration and plant-water-use prediction approaches based on drone thermography and subsequent modelling with causal as well as machine learning models are a useful extension or even a potential alternative to well-proven methods such as eddy covariance, sap flux, stomatal conductance. However, using only drone recorded data was often not enough since all predictions benefited from additional information on solar radiation as well as from measurements of relative humidity and air temperature. The open-source software developed in the scope of this thesis is at this time available in four versions and more than 1100 downloads were registered up to this moment.

Zusammenfassung

Evapotranspiration (ET) ist eine zentrale Größe sowohl im regionalen als auch im globalen Wasserkreislauf. Terrestrische ET besteht zum größten Teil aus Transpirationsflüssen von Pflanzen. Diese Dissertation stellt einen Methodenvergleich von Methoden zur Untersuchung von Evapotranspiration und Pflanzenwasserverbrauch an. Verglichen werden etablierte und umfangreich erprobte mit neuen Methoden, die auf dem Einsatz von Thermographie, Drohnen und Vorhersagemodellen basieren. Die wichtigsten Erkenntnisse dieses Methodenvergleichs wurden in einem standardisierten Arbeitsablauf zusammengefasst und in eine nutzer*innenfreundliche Software integriert. Diese Software wurde über die offizielle QGIS3-Erweiterungsplattform (QGIS3 Python Plugin Repository) zur öffentlichen Verfügbarkeit bereitgestellt.

Die primäre Zielsetzung der hier vorgestellten Studien war es, die neuen Möglichkeiten, die durch den Einsatz von Drohnen, radiometrischer Wärmebildkameras, sowie verschiedener kausaler Modelle und durch maschinelles Lernen zur Quantifizierung von Evapotranspiration und Pflanzenwasserverbrauch möglich wurden, zu erproben und gegen etablierte Methoden zu testen. Zudem war es Ziel der Dissertation, die wichtigsten, im Feld zu messenden Variablen festzustellen, um einen minimal notwendigen Datensatz für künftige Studien zu evaluieren. Ein weiteres Ziel war die Erstellung und Veröffentlichung einer quelloffenen Software, die einen standardisierten Arbeitsablauf zur Berechnung von Evapotranspiration auf Basis von Wärmebildern unterschiedlichster Quellen bereitstellt, und so einem möglichst breiten Spektrum von Nutzer*innen zur Verfügung steht. Die aus diesen Zielsetzungen hervorgehenden Fragestellungen wurden in drei Einzelstudien beantwortet:

In der ersten Studie wurden Wärmebildaufnahmen der Oberflächentemperaturen als wichtigste Eingangsvariablen für Energiebalancemodelle (EBMs) über den Zeitraum von fünf Tagen bei 61 Drohnenbefliegungen über einer Palmölplantage aufgezeichnet. Insgesamt wurden dazu drei verschiedene EBMs verwendet: das DATTUTDUT-Modell (Timmermans et al., 2015), bei dem keine Unterscheidung

zwischen Vegetation und Boden vorgenommen wird, und zwei komplexere EBMs, TSEB-PT (Norman et al., 1995) und DTD (Norman et al., 2000), die diese Unterscheidung vornehmen. Vorhersagen zum latenten Wärmefluss stimmten bei unterschiedlichen Wetterlagen in hohem Maße mit den Messungen des Eddy Kovarianz Systems überein (r^2=0.85; MAE=47; RMSE=60). Die Konfidenzintervalle für die Steigung und den Interzept der Regressionsgerade einer Deming-Regression weißen darauf hin, dass zwischen beiden Methoden kein signifikanter Unterschied besteht. Mit dem TSEB-PT- und DTD-Modell konnten auch gute Ergebnisse erzielt werden, allerdings ist die Genauigkeit der Vorhersagen bei allen Modellen von der präzisen Bestimmung der Nettostrahlung abhängig. Zusammenfassend stellen drohnenbasierte Thermographieansätze kombiniert mit EBMs eine verlässliche Methode zur Feststellung von ET dar. Zusätzlich bietet dieser Ansatz die Möglichkeit ET-flüsse in räumlichen Karten zu analysieren und darzustellen.

Die zweite Studie konzentriert sich auf Transpiration von Pflanzen als Schlüsselgröße im Wasserkreislauf. Die typischen Methoden um Transpiration und stomatäre Leitfähigkeit zu messen sind die Saftflussmethode und die Porometie. Kombinationen aus der Nutzung von Oberflächentemperaturen und Machine-Learning-Techniken eröffnen neue Möglichkeiten um Transpiration und stomatäre Leitfähigkeit zu bestimmen. Der Schwerpunkt dieser Studie liegt auf der Vorhersage von Saftfluss und stomatärer Leitfähigkeit mithilfe von mit Drohnen aufgenommener und meteorologischer Daten. Hierfür wurde ein verhältnismäßig großer Datensatz, bestehend aus 103 Drohnenbefliegungen, Saftflussaufzeichnungen in 10 min Intervallen über einen Zeitraum von zwei Wochen und mehreren tausend Messungen der stomatären Leitfähigkeit, angelegt. Der Datensatz wurde sowohl an Ölpalmen als auch an vier auf Sumatra heimischen Baumarten erhoben, die mit Ölpalmen in einem Agroforst-Mischbestand wuchsen. Da hier im Vergleich zur ersten Studie keine Modelle verwendet wurden, die auf kausalen Zusammenhängen basieren, sondern ganz im Gegenteil, rein datengetriebene Algorithmen zum Einsatz kommen sollten, war dieser umfangreiche Datensatz für ein erfolgreiches Trainieren der Modelle notwendig. Die Algorithmen bestanden aus einer multiplen linearen Regression, die als einfache Vergleichsalternative für die deutlich komplexeren Machine-Learning-Algorithmen gedient hat. Als Machine-Learning-Algorithmen wurden jeweils eine Stützvektormaschine (Support Vector Machine, SVM), zwei verschiedene Random-Forest-Modelle (RF) sowie ein künstliches neuronales Netz (ANN) angelegt und mit den Trainingsdatensätzen für Saftfluss und stomatäre Leitfähigkeit trainiert. Das Ergebnis der Evaluation war, dass vor allem die RF-Modelle bei der Quantifizierung von Saftfluss den größten Erfolg versprechen, da hier kongruente Ergebnisse mit den Saftflussmessungen berechnet wurden (r^2=0.87 for trees and r^2=0.58 for palms). Für die meisten

Arten und Artenzusammenstellungen hat sich gezeigt, dass diese Methode vergleichbar gute Ergebnisse wie die Saftflussmethode produziert. Allerdings zeigten auch die anderen Methoden, vor allem bei der Saftflussbestimmung in Ölpalmen, ein großes Potential für weitere Entwicklung. Stomatäre Leitfähigkeit konnte nur mit durchwachsenem Erfolg vorhergesagt werden, was vermutlich an der der relativ großen Zeitdifferenz zwischen den Messungen lag. Allerdings könnte ein größerer Datensatz mit geringerer Zeitdifferenz zu deutlich besseren Ergebnissen führen.

ET lässt sich mit Hilfe von EBMs bestimmen, aber die Nutzung dieser Modelle ist oft relativ komplex. Die letzte Studie hat sich mit der Entwicklung einer quelloffenen Software befasst, die zur Anwender*innen freundlichen Nutzung über eine graphische Nutzer*innenoberfläche (GUI) verfügt und als Erweiterung für die Geoinformationsplattform QGIS3 implementiert worden ist. Wichtigstes Ziel für die Umsetzung der Software war die Möglichkeit, Wärmebilder aus unterschiedlichsten Quellen (z.B. Satelliten- und Drohnenaufnahmen sowie Aufnahmen von Handwärmebildkameras) verwenden zu können. Zudem sollten die Erkenntnisse aus den ersten beiden Studien, vor allem die Option, gemessene Strahlungswerte und Lufttemperatur in den Quantifizierungsprozess der Evapotranspiration zu integrieren, Eingang in die Software finden. Die aus diesen Zielsetzungen entstandene Software „QWaterModel“ steht in der offiziellen QGIS3 Erweiterungsplattform zum Download bereit. Um ihre Anwendungsmöglichkeiten zu testen, wurden Wärmebildaufnahmen einer Palmölplantage mit Drohnen und einer Handwärmebildkamera sowie verschiedene Strahlungsmessungen und -modellierungen als Datengrundlage verwendet. Typische, tägliche Evapotranspirationskurven konnten für beide Aufnahmeoptionen sowie alle Strahlungsoptionen festgestellt werden. Die Präzision der Vorhersage wurde im Vergleich mit der Referenzmethode Eddy-Kovarianz allerdings durch den Einsatz von Strahlungsmessungen erheblich verbessert.

Zusammenfassend zeigt die Arbeit, dass auf Drohnen aufbauende Thermographie- und Modellierungsmethoden zur Quantifizierung von Evapotranspiration und des Pflanzenwasserverbrauchs eine brauchbare Erweiterung oder Alternative zu den weitverbreiteten Methoden (Eddy-Kovarianz, Saftfluss und bedingt auch Porometrie) darstellen. Allerdings sind zusätzliche Messungen, vor allem der Strahlungsintensität sowie der Luftfeuchte und -temperatur, der Quantifizierungspräzision zuträglich. Die im Rahmen dieser Dissertation entwickelte Software „QWaterModel“ ist zu diesem Zeitpunkt bereits in vier Versionen verfügbar und erfreut sich mit mittlerweile weit mehr als 1100 Downloads großer Beliebtheit in der hydrologischen QGIS3 Gemeinschaft.

Chapter 1

Introduction

In 1867 G.J. Symons described the science of evaporation as the most desperate branch of the desperate science of meteorology, in an effort to find 'the almost universal method' to quantify this flux (Symons, 1867). While devices such as the standardized rain gauge had already been installed in many locations in the kingdom, the quantification of evaporation was still very difficult (Symons, 1867). First long term records from evaporation pans were available, but their layout and methodology of use was not standardized and therefore difficult to compare (Symons, 1867). Since the definition of evaporation was much wider at that time and is better described as evapotranspiration (ET) nowadays, this period did not only mark the beginning of ET measurements but it also marked the beginning of the methodological discussion on its assessment and quantification.

Now, 150 years later, many methods to quantify ET have been proposed, tested, standardized and established and the toolbox of available quantification schemes is growing steadily (Verstraeten et al., 2008). In the last decade autonomous flying sensor platforms for near surface remote sensing and miniaturized thermal and multispectral sensors have been added to this toolbox and enabled a large variety of new methods for ET assessment (Berni et al., 2009; Brenner et al., 2018; Sullivan et al., 2007). The use of a combination of unmanned aerial vehicles (UAVs or drones) and miniaturized sensors became very present in many scientific branches and form the basic framework of numerous new methods (Ahongshangbam et al., 2019; Berni et al., 2008; Hoffmann et al., 2016).

This dissertation attempts to compare near-surface drone remote sensing based methods with widely used and established ground-based methods (eddy covariance, sap flux and porometry) to test and add new methods to the available set of tools for evapotranspiration assessment. Furthermore, in the sense of Symons (1867), this study aims to create a standardized tool for ET derivation from

thermal images from best practice workflow evaluations.

ET represents both a mass and an energy flux and describes the process of water transfer from water and land surfaces, sublimation of ice and snow and plant transpiration to the atmosphere (Verstraeten et al., 2008). It is a central flux in the hydrological cycle and consumes almost two-thirds of terrestrial precipitation on a global scale (Oki and Kanae, 2006). Knowledge on ET is essential for a large variety of scientific, environmental and economical applications and fields. In agricultural production ET is a direct indicator of plant water use and affects the choice for land management systems and irrigation schemes in arid regions (Allen et al., 1998). The assessment of thermal patterns and ET can support ecosystem conservation, precision agriculture, and wildfire detection (Baena et al., 2017; Messina and Modica, 2020; Moulianitis et al., 2019). Moreover, ET is closely linked to carbon cycles influencing terrestrial carbon flows from and towards the atmosphere and is therefore important for the assessment of nutrient flows and carbon balance modelling (Verstraeten et al., 2008). Information on ET is further essential for the provision of drinking water, water resource management and food security and therefore directly influencing human life (Kaushal et al., 2017). There is a special urge to better understand ET and its drivers as global climate change is expected to increase atmospheric evaporative demand and droughts are predicted to become more severe and frequent in the future (Prudhomme et al., 2014).

ET is mainly controlled by meteorological and vegetation factors. Therefore, solar radiation, air temperature, wind velocity, relative humidity, leaf area, stomatal conductance and rooting depth play a prevalent role that shape the magnitudes of ET (Verstraeten et al., 2008). Since transpiration from plants represents 80 to 90% of this entire ET magnitude on land, changes in the biosphere and vegetation cover will directly translate to alternations in ET quantities (Jasechko et al., 2013). Therefore, ET is largely affected by land-cover and land-use changes (Good et al., 2015; Hansen et al., 2013). Especially in tropical regions where major conversation efforts foster the transfer from forest to agricultural land, ET is subject to large scale alterations (Hansen et al., 2013; Margono et al., 2014).

The adequate method for ET measurements or its components highly depends on the spatial and temporal scale of interest. For studies that focus on plant metabolisms or water use patterns of a certain species, a small spatial footprint approach on a leaf, individual plant or on a plot level represent the most suitable scale (Dolman, 1988; Waite et al., 2019). Efficient methods for this particular spatial scale are the application of devices such as porometers, sap flux sensors, lysimeters, or deuterium tracing approaches (Mei et al., 2019; Röll et al., 2019;

Teuling, 2018; Waite et al., 2019). If the research object is less focused on individual plants but rather on ecosystems, farms or smaller catchments, a landscape scale is more suitable. The prevalent method for this scale is probably the eddy covariance technique (EC), but also SWAT models and water balancing methods are frequently applied (Meijide et al., 2018; Teuling, 2018; Xia et al., 2016). Large spatial scales are mostly energy or mass balance modelling approaches such as SEBAL or PROMET, which are based on satellite recorded spatial data (Bach et al., 2015; Bastiaanssen et al., 1998).

Furthermore, temporal scales are an important factor to consider. Temporal resolution of data is usually highly dependent on the level of automation of the individual methods. Methods with a high level of automation that (once set up) only require a minimum of direct human interaction are e.g. satellite-based approaches or terrestrial EC, sap flux or lysimeter systems. These methods can provide a constantly updated stream of data, which may last for several decades, as e.g. in case of the Landsat missions (Lechner et al., 2020). Handheld porometry devices or thermal cameras, require high direct involvement of human interaction, since they have to be applied to each measurement point manually. The logging cycle, the time period required to record a single measurement for a certain object or area, represents an important temporal factor. While measurement systems such as EC stations or sap flux devices measure with a very high frequency (every 0.1 to 30 seconds respectively) (Meijide et al., 2017; Röll et al., 2015), the return time of satellites can span up to several weeks (Brenner et al., 2018). Contrary to this, porometry has highly variable measuring periods that depend on external factors such as relative humidity influencing the time required to provide a controlled air humidity in the measurement chamber (Jones, 2014).

Ground based measurements, such as sap flux, stomatal conductance or EC, benefit from the proximity to the measured object. Interference of other variables are usually minor or are well known and can be accounted for. This proximity helps to reduce measurement errors but comes at the cost of a reduced spatial representation. To represent spatial trades more adequately, ground based point measurements require spatial replication and the limited set of samples is often interpolated to estimate the gaps between the measured points (Lechner et al., 2020).

A different option to ground based point measurements is remote sensing, where data is collected without directly contacting the feature of interest (Jones and Vaughan, 2010). Remote sensing approaches to derive ET provide a good spatial representation of data, but often lack a high temporal frequency. In case of satellites, a trade off situation between temporal frequency and resolution exists

(Brenner et al., 2018; Delogu et al., 2012). Using satellite recorded land surface temperature data for ET computations has many advantages, especially when long-term observations are needed (Kalma et al., 2008; Lechner et al., 2020). However, clouds, haze and atmospheric humidity represent a barrier for the thermal spectrum (Still et al., 2019) and hence data availability depends on both, overpass frequency and cloud cover conditions (Delogu et al., 2012). Aerial photography from planes or helicopters can benefit from their relative proximity to the land surface producing a higher resolution compared to satellite recordings. ET has been successfully predicted on plot to landscape scale with thermal images recorded from planes (Xia et al., 2016). Drone-based remote sensing systems, which have recently emerged, could be an option to bridge the gap between plant scale measurement methods and catchment or landscape scale schemes. Drones can operate close to the surface enabling a delimitation of single plant canopies and simultaneously cover considerable areas (Ahongshangbam et al., 2019; Mohan et al., 2017). They can be equipped with a growing variety of light-weight sensors for diverse spectral ranges or structural measurements (Berni et al., 2008; Brenner et al., 2017; Kellner et al., 2019).

This study attempts to bridge the gap between classical methods and recently emerging methods for ET and transpiration assessment. Therefore, well–established ground-based methods for ET and transpiration measurement are compared with drone thermal remote sensing and prediction modelling approaches.

In view of the numerous opportunities originating from drone-based remote sensing with miniaturized thermal sensors and the need for a streamlined and standardized workflow, the following four objectives were set for this dissertation:

(1) For the assessment of bias between well-known and recently emerging approaches, a method comparison of classical methods of energy flux and ET measurements (eddy covariance) with drone remote sensing data based modelling approaches is conducted.

(2) Similar to objective (1) a method comparison of plant scale sap flux and porometry methods with a drone-based transpiration derivation approach, supported by different statistical and machine learning algorithms, is performed to assess the potential bias between the methods.

(3) The most important environmental parameters to predict ET and transpiration from a minimal set of data are determined.

(4) An open-source tool based on the results of objective 1-3 is developed. The aim of this tool is to standardize the workflow for ET computation, and to provide the ability to use thermal input data from different sources (e.g. satellites, drones, handheld thermal cameras). The resulting software is explained and tested against an established reference method.

Working towards objective (1) and (2) it will become clear, whether the drone-based methods are comparable with well-established methods. Objective (3) will show which parameters are relevant to optimize the measurement process in the future. Finally, objective (4) will make some of the results available to a wider public.

Chapter 2

Methods

The choice for the most adequate method highly depends on the feature of interest, scale, availability and cost. In many studies the methods part is used to describe in detail how a method was applied, but often there is no justification why certain methods have been chosen. This part, however, illustrates both reasons and inducements of how the decisions for certain methods were made. The specific methods used in this dissertation are described in detail in each study (Chapter 3-5).

2.1 Canopy temperatures and evapotanspiration

Canopy temperatures are a good indicator for the processes which control the leaf energy balance, since they are linearly related to the rate of water loss (Jones, 2014; Raschke, 1956). Water is lost from the plant surface by evaporation, in the same process energy (or heat) from the leaf is applied to change water from a liquid to a gaseous state (Raschke, 1956). With the evaporation of water, both water and energy are transferred from the plant to the atmosphere, cooling the surface – a process, which is called evaporative cooling. The process of ET therefore affects the mass and energy balance of the leaf (Jones and Vaughan, 2010) and the overall canopy temperatures. The rate of water loss is regulated by stomatal conductance and the coupling of the leaf or canopy to the atmosphere (Raschke, 1956). Transpiration behaves proportional to stomatal conductance, if the plant is well coupled to the atmosphere (Jones, 2014). This is the case for single plants or smaller canopies, where the boundary-layer transfer process is benefited by factors, such as small leaf size, turbulent air movement and a large

canopy roughness (Jones, 2014). For larger continuous canopy areas with homogeneous vegetation these factors are less dominant, since the total ET rate is more dependent on the available energy rather than on stomatal aperture (Jones, 2014).

2.2 Thermal infrared remote sensing and its challenges

To remotely measure canopy temperatures, thermal infrared radiation (TIR) emitted from the canopy is measured using the thermography method. This method measures electromagnetic radiation, which is emitted by all objects with a temperature above 0 K (Planck, 1900) and from most objects on the earth surface, including plants (Kuenzer and Dech, 2013). This emittance peaks at the thermal infrared domain of around 9.7 μm (Sabins and Lulla, 1987).

Remotely sensed TIR data must be corrected for geometric and radiometric distortions (Kuenzer and Dech, 2013). Therefore, the recorded measurement value at the sensor, is converted to a value for the corresponding amount of incoming energy. Then, the temperature on the measured surface is calculated considering the atmospheric path length, atmospheric radiation attenuation, surface and atmospheric emissivity (Kuenzer and Dech, 2013). Thermal radiation is an energy transfer process, that can travel undisturbed through a vacuum but interacts with matter (Sabins and Lulla, 1987). Atmospheric humidity can have a major effect on measurement accuracy, because water vapor does not only attenuate thermal radiation from the surface of interest to the sensor, but also emits its own thermal radiation (Still et al., 2019).

A typical product of remote sensing thermography are land surface temperatures (LST). These LSTs are the key input in a growing number of agricultural, environmental or urban environment studies (Anderson et al., 2011; Brenner et al., 2017; Hoffmann et al., 2016). Typical studies assess plant water stress, yields, evapotranspiration or changes in stomatal conductance (Berni et al., 2009; Hoffmann et al., 2016; Siebert et al., 2014).

2.3 Drones as remote sensing platforms

To record LSTs from remote, a thermal sensor has to be placed on a carrier system to locate it in the right position. The simplest form of sensor location would

be a handheld thermal camera or a camera that is attached to a steady object (Jones et al., 2009; Leinonen et al., 2006; Maes and Steppe, 2012). Camera carriers, such as planes and satellites, are more complex, but have the advantage of covering larger areas (Anderson et al., 2011; Xia et al., 2016). The choice for a sensor carrier is characterized by desired temporal and spatial coverage and resolution (Figure 2.1). Earlier approaches of canopy temperature-based studies either focused on very small scale areas (Leinonen et al., 2006), such as leaves or single plants, or use satellite data on a regional or continental scale (Anderson et al., 2011).

Both approaches have advantages and drawbacks: the small-scale approach offers a high resolution, but does not cover enough area to answer research questions related to the landscape scale; the large-scale approach covers a wider area, but the resolution is too coarse to respond to single plant related research objectives.

Most ecological and agricultural interests require a scale from individual plants over plot scale to a landscape scale. The reasoning behind this scale decision is simple: in this scale range single plant behaviour, as well as plant communities can be observed and analysed in snapshots without an extended time scale. If the area is picked smaller, interactions or patterns in plant communities are difficult to observe. If a larger research area is chosen, time dependent hydrological variables, such as surface runoff, water storage and ground water, have to be considered, requiring longer time series. Figure 2.1 shows the different scales and points out the most relevant scale for ecological research in green. For the studies in this dissertation drones were the tool of choice, because the covered footprint ideally fits most ecohydrological questions.

TIR images and maps usually have a much lower spatial resolution than RGB or near IR cameras. This is caused by the longer wavelength that requires a bigger sensor and larger lens for similar resolutions of images in the TIR range (Kuenzer and Dech, 2013). However, increasing the lens size is costly since they are usually produced from rather expensive Germanium (Luhmann et al., 2013). It is therefore difficult and expensive to obtain high resolution of TIR images with a similar coverage than e.g. an RGB image, if the sensor is placed far away from the surface of interest.

Satellites are good tools for time series data in regional, global or continental scales and their constant recording frequency allows for time series in ET studies (Sobrino and Julien, 2013). The current operational satellite systems offer a thermal resolution of 100m (Brenner et al., 2017) and will not be able to cover single plants or the variability of plot sized areas adequately. For the small-scale

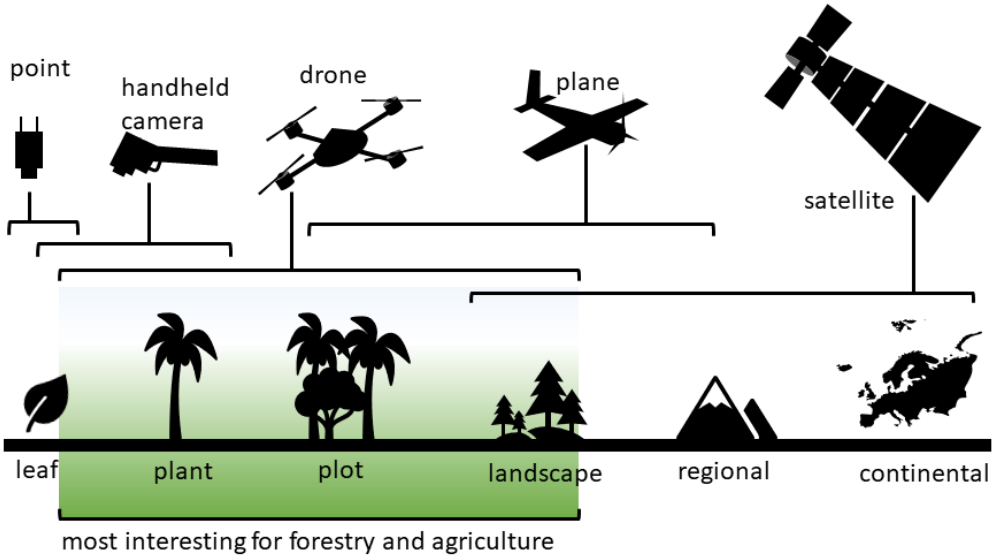


Figure 2.1: *Devices and carriers for land surface temperature recording and their associated scale.*

approach, many small measurements can be up-scaled through different techniques to cover the gaps between the point measurements. But there are still many difficulties to cover e.g. plant community behaviour or thermal patterns in an ecosystem.

In the last decade, an additional option to measure LSTs is the use of unmanned aerial vehicles (UAV) or drones (Berni et al., 2008; Brenner et al., 2017; Hoffmann et al., 2016). The term drone actually covers a much wider variety of autonomous vehicles of all sorts, including autonomous rovers, boats and submarines. However, the term drone is used in this study, since most people would associate this type of carrier platform more with the term 'drone' than the rather technical term 'UAV'. Drones offer a unique opportunity of supporting TIR image acquisition in just the right scale for many ecological, forestry and agricultural questions (Vasterling and Meyer, 2013). Drones operate very close to the surface and can provide images with a higher spatial resolution, compared to satellites (Hoffmann et al., 2016). This proximity reduces the interaction of atmospheric effects with the thermal radiation on its path from the recorded surface to the sensor. A further advantage of close surface remote sensing is the increased resolution, which allows for a delimitation of canopy and non-canopy pixels. Therefore, emissivity can be adjusted individually for e.g. soil and canopy pixels. Depending on the type of drone and its flight and stabilization perfor-

mance, a special camera positioning gear, commonly referred to as gimbal is required. The gimbal can maintain the sensor in a stable position, even if the drone is manoeuvring. This is very important in order to keep a stable angle between the sensor and the recorded surface (usually the sensor is positioned in a nadir view angle, 90° towards the Earth surface). Drones further offer a high degree of flexibility: In a temporal manner, the time and number of overpasses can be planned and adjusted, exact times for overpasses can be planned and operated, which allows for more flexibility compared to satellites (Berni et al., 2008). The area to be covered is not constrained by terrain or flight passes, so decisions on spatial coverage rather depend on drone type and can be chosen flexibly. Pixel size and single image coverage can be adjusted according to the allowed flight height. There is a steadily growing number of sensors that can be attached to the drones and multi-sensor approaches are also an option (Berni et al., 2008; Brenner et al., 2017; Kellner et al., 2019). Radiometric TIR sensors for LST recording have become light-weight and affordable (Brenner et al., 2017), since their production volume and availability increased and miniaturization of parts became cheaper. The uncooled microbolometer, that most of the TIR sensors are based on, can be integrated by the thousands into a tiny chip and the miniaturization of calibration and computation gear benefitted the development of very small, lightweight TIR sensors, which are robust enough to be applied in conjunction with drone carriers (Vasterling and Meyer, 2013).

However, there are disadvantages to this carrier type too: Apart from the challenging technical handling – especially of highly customized drones in science – and the rather small payload of 1-5 kg, a considerable danger of hurting the operation team or bystanders is always involved with the operation of drones. Furthermore, special attention has to be taken for local laws and restrictions. Drones are a great tool for near-surface remote sensing applications, but have to be handled with special care, responsibility and precaution.

2.4 Data and data structure

This section discusses management, alteration and processing of the data and illuminates the basic concepts behind the data structures used. This topic is not discussed very often in scientific publications, but the more data driven and quantitative research becomes, the more work and decisions on data sets are involved, the more 'number crunching' matters. A clear data philosophy helps to increase productivity and to avoid errors. If a database is set up in a well-organized and transparent way, it also provides quick access to all relevant metrics and allows

the generation of data subsets and additional metrics. With the appropriate data structure, the detection of errors is facilitated and corrections are easier to implement. The data for this study can be divided up into different data types: Continuous measurements from EC, meteorological measurements and sap flux, scheduled measurements from drone data acquisition and the handheld thermal camera and less regular semi-timed measurements, such as measurement from stomatal conductance. A further non-time dependent component of data, such as geo-location, LAI, canopy size and flight height, is part of the data set, too.

The data sets arranged for this thesis are usually combinations of all the above-mentioned data types. In the following, the major data sets recorded during the field work for this dissertation are shortly presented. Unfortunately, due to limitations in time, only a small portion of all the gathered data could be used in the studies comprised in this dissertation.

(1) PTPN VI data set I **location:** 103.3914785, -1.6939520

Data used for study in Chapter 3
Recording period: 05.08.2017 – 10.08.2017
Number of flights: 63 (61 used)
Data recorded with drone: TIR, RGB
Other data recorded:
eddy covariance data (in 30 min means)
meteorological data (in 10 min means)
TIR calibration data set

(2) Humusindo data set I **location:** 103.251170, -1.948734

Data used for study in Chapter 4
Recording period: 01.10.2016 — 16.0.2016
Number of flights: 103 (103 used)
Data recorded with drone:
TIR, RGB, NGB
Other data recorded:
sap flux measurements (in 10 min means)
porometry measurements
meteorological data (in 10 min means)

(3) PTPN VI data set II **location:** 103.3914785, -1.6939520

Data used for study in Chapter 5
Recording period: 07.08.2017
Number of flights: 13 (13 used)
Data recorded with drone: TIR, RGB
Other data recorded:
eddy covariance data (in 30 min means)
meteorological data (in 10 min means)
handheld thermal camera data

(4) PT REKI data set **location:** 103.3441548, -2.1898686

Data used for (Ahongshangbam et al., 2020)
Recording periods:
12.11.2016 – 20.11.2016
16.07.2017 – 01.08.2017
05.09.2017 – 12.09.2017
Number of flights: 216 (72,72,72)
Data recorded with drone: TIR, RGB
Other data recorded:
meteorological data (in 10 min means)

(5) Oil palm management experiment **location:** 103.3981297, -1.7065760

Recording period: 08.08.2018 – 15.08.2018
Number of flights: 47
Data recorded with drone:
TIR, RGB, Multispectral
Other data recorded:
meteorological data (in 10 min means)
TIR calibration data set
reference flights over the PTPN VI EC-tower

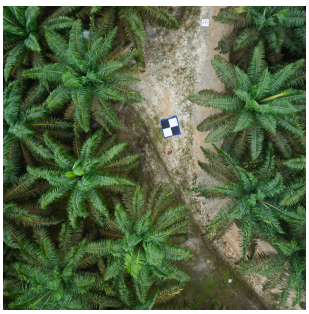
(6) Bungku land use data set **location:** 103.2641877, -1.9387656

Recording period: 13.09.2017 – 19.09.2017
Number of flights: 39
Data recorded with drone: TIR, RGB
Other data recorded:
meteorological data (in 10 min means)
sap flux measurements (in 10 min means)
measurements of tree diameters

(7) PTPN VI area map **location:** 103.3914916, -1.6939556

This data set was used to create a detailed map
Recording period: 01.08.2017
Number of flights: 13 (13 used)
Data recorded with drone: RGB
Other data recorded:
–

(8) PTPN VI oil palm 3D model **location:** 103.3850863, -1.6945917



Recording period: 14.08.2017
Number of flights: 4
Data recorded with drone: TIR, RGB
Other data recorded:
meteorological data (in 10 min means)
3D calibration data

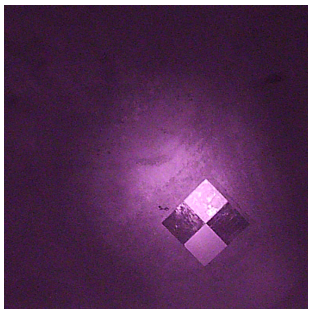
(9) Humusindo data set II **location:** 103.2436324, -1.9539240

Recording period: 23.08.2017 – 29.08.2017
Number of flights: 39
Data recorded with drone: TIR, RGB
Other data recorded:
meteorological data (in 10 min means)
IButton measurements (air temp., rel. hum.)
TIR calibration data set

(10) Humusindo data set III **location:** 103.2482880, -1.9483750

Recording period: 20.09.2017 – 26.09.2017
Number of flights: 67
Data recorded with drone: TIR, RGB
Other data recorded:
meteorological data (in 10 min means)
IButton measurements (air temp., rel. hum.)
TIR calibration data set

(11) Bat observation data set **location:** 103.2530733, -1.9519718



Recording period: 23.09.2020-25.09.2020
Number of flights: 13
Data recorded with drone:
TIR, night vision
Other data recorded:
eddy covariance data (in 30 min means)
meteorological data (in 10 min means)
handheld thermal camera data

The data sets comprised data measured in oil palm plantations with very different management philosophies and practices. The data contain typical examples for commercial management in oil palm plantations of larger companies (1, 3, 4, 7, 8), plantations of small-holder farming (6) and areas that were part of biodiver-

sity enrichment efforts (2, 9, 10). A large portion of the data was recorded in a long-term study in the forest protection area PT REKI (4), where the rainy season as well as the beginning and end of the dry season was covered. Furthermore, the data set (4) covers four flooded and four non-flooded plots. Other data sets covered the 3D modelling of oil palms (8) and the creation of high-resolution spatial maps (7). A very special data set was recorded at night (11) where bats were filmed in a big tree and along a small river using a specially designed night vision camera and a near-IR spotlight.

A calibration data set was recorded with a handheld thermal camera to radiometrically reference the drone TIR recordings. The exact data pre-processing and processing procedure is described in Chapter 3-5. Images were exported to .csv Format using TeAX-Thermoviewer (TeAx Technology GmbH, Wilnsdorf, Germany). All thermal data was checked visually and the images taken before and after the flight were removed. The images were automatically tacked with the geo-location and time (UTC), from the exact site where they had been recorded using the drone GPS output. Blurry pictures were removed visually. The images were then converted from temperature measurements in °C to K and saved as .asc files using a software we especially developed for that purpose. Then, the .asc files were loaded into QGIS3 as a raster layer. Background maps and 3D-Point clouds were created from the RGB images that were recorded with the thermal flights or other mapping missions during the field study. We used Agisoft Photoscan Professional 1.2.6 to create point clouds and subsequently exported orthomosaicked maps.

Following the energy balance modelling procedures (described in detail in Chapter 3 - 5) the results were arranged in an SQL-like database programmed in Python 3.6.9, which uses pandas (McKinney, 2010) and NumPy (Oliphant, 2006; van der Walt et al., 2011) libraries. Graphs and figures were created with Matplotlib (Hunter, 2007) and seaborn (Waskom et al., 2020) libraries.

2.5 Classical statistics and machine learning

The fundamental different concepts of classical statistics and machine learning (ML) are often not discussed. This part was included into this dissertation, because why and when the respective approach is appropriate in science, is rarely explained.

This study comprises two main statistical objectives: inference and prediction.

While inference usually aims at building a mathematical model to formalize a relationship between variables or at testing a hypothesis, prediction aims at forecasting an outcome from previous observations (Bzdok et al., 2018). The goal of inference and classical statistics is to understand the underlying mechanisms of data relations. Prediction in contrary does not necessarily require a deep understanding of the underlying mechanisms and is rather aimed at guessing the unknown based on known parameters (Breiman, 2001).

Back in 2001 Breiman estimated that 98% of statisticians were predominately using classical statistics and inference approaches (Breiman, 2001). While this percentage might have slightly changed in the last two decades, classical statistics are obviously still the dominant procedure for studies in natural sciences. The general idea of classical statistics is routed in the assumption that data are generated by a stochastic data model (probability spaces or probability triples) (Bzdok, 2017; Kolmogorov and Morrison, 2013). In a typical study inference is based on an estimation component - often a linear-regression model – and a test for statistical significance. Typical examples of this approach are the Deming and Passing-Bablok regressions used in Chapter 3 and 4. ML in contrary is an out-of-sample generalization approach, which uses an algorithmic learning model fitted on a bigger training data set and the resulting model is later empirically evaluated using a smaller test data set (Breiman, 2001; Bzdok, 2017).

Unlike classical statistical approaches that typically consider the whole data set, consisting of all available data to reject the null hypothesis, ML procedures derive a model from only a part of the available data (Breiman, 2001; Bzdok, 2017; Wasserstein and Lazar, 2016). A test data set with previously unused data is later used to test predictive accuracy. The validation process in classical statistics is commonly based on a goodness-of-fit test and a residual examination while ML approaches validate the resulting model with metrics, such as predictive accuracy, and based on aforesaid test data set (Breiman, 2001).

Chapter 4 is a typical application of algorithmic modelling with ML, where two target variables sap flux and stomatal conductance are predicted from a large set of different variables. The resulting models are not designed to causally implement or explain all the processes, which take place in the plant, in the leaf to atmosphere boundary-layer transfer process or in the atmosphere. Their purpose is merely the prediction of plant-water-cycle components based on remote sensing and meteorological data, which has been gathered. Classical statistics are then used to evaluate these prediction models against their measured counterparts in a method comparison approach.

2.6 The relevance of method comparisons

The idea of a method comparison is to answer the question whether both methods measure the same values for the same quantity of a feature of interest, or whether there is a systematic difference between the measurements (Passing and Bablok, 1983). This is an important procedure to test a new method against an established method and to verify their comparability. A method comparison is a valuable tool, if a new method is supposed to be established or to replace an older, more labour-intense or more expensive method (Bland and Altman, 1986). For this task, correlation analysis and t-tests are a common procedure in literature, but both methods are not appropriate for the purpose of an actual method comparison (Bland and Altman, 1986). A correlation analysis captures the degree of association of two variables using either the correlation coefficient r or the coefficient of determination r^2. The idea is to provide a metric of how well the data can be explained by a linear relationship. This method is problematic for method comparisons, since the resulting linear model is not necessarily without bias (Simundic, 2016). The linear model could have a bias of e.g. 100% but the degree of association of both variables might still be described by r^2=1. To present the study results the coefficient of determination is still used in this dissertation, mainly as an additional descriptive method in conjunction with the results of the Deming or Passing-Bablok regression. T-tests (paired or for independent samples) are also not the ideal tool for method comparison. The independent t-test compares the means of both methods. If the means of the X and Y method measurements are similar by coincidence, the methods would be declared as comparable in this approach. If the measured values would be exactly opposite, no difference would be detected neither. The paired t-test is already a better choice for a method comparison, but it is highly susceptible to the number of samples (Simundic, 2016). If the number of samples is low, it is biased to neglect a difference between both methods. If the sample size is too big, it will be biased to detect a difference between both methods.

Typical statistical methods to compare two methods are the Deming regression (Deming, 1964) and the Passing Bablok regression (Passing and Bablok, 1983). A simple linear regression assumes no measurement error in the X-variable. The Deming regression extends this concept and considers random measurement errors in both X- and Y-variables (Cornbleet and Gochman, 1979). If the measurement errors of both methods are not known, the default procedure is to assume an error ratio of 1 (Legendre and Legendre, 2003). Deming regression is a parametric method and data is expected to follow a normal distribution. Further, a constant precision over the measurement range and the removal of extreme

outliers are a prerequisite (Simundic, 2016). The sample size should exceed a minimum of n = 60 samples (Legendre and Legendre, 2003). Passing-Bablok regression is a non-parametric method, assumptions about data distribution are not made (Passing and Bablok, 1983). For a statistically unbiased result, it is assumed that both variables are correlated and show a linear relationship (Passing and Bablok, 1984). Both, the Deming and the Passing-Bablok regressions are analysed by verifying if the numbers 1 and 0 are contained in the confidence intervals for intercept and slope respectively. If both numbers are contained in the respective confidence intervals, the methods can be used interchangeably.

This places particular emphasis on the adequate calculation of confidence intervals. For the Deming regression the Jackknife method is suggested to support the calculation of confidence intervals (Linnet, 1993). For the Passing-Bablok regression a nested bootstrap interval is the method of choice (Hinkley and Shi, 1989).

To create adequate data sets for a method comparison, the entire meaningful range of data should be covered (Passing and Bablok, 1983). Measurements should be distributed along the measurement range as evenly as possible. Scatter plots are a good way to check whether there are gaps in the measurements. Figure 5.3 in Chapter 5 shows such a distribution without gaps. Figure 5.4b and 5.4c in Chapter 5 are examples of data sets with a gap. In case of the specific study in Chapter 5, this gap does not present a big problem, since the focus of this particular analysis is not on a method comparison. If a data set for a method comparison is prepared, these gaps should ideally be filled with additional measurements in the missing range. Measurement duplication is advisable but often not a really practical approach in applied science. In case of this dissertation, the data for Chapter 3 cover an entire week of measurements. However, an ideal data set would contain a second data set covering another week – e.g. in another season of the year – using the same procedure and instruments, to be able to calculate measurement errors. Measurements should be taken along the most frequent measurement situations (Simundic, 2016). In case of the methods compared in this dissertation, this implies varying weather conditions, strong and lower solar irradiance and variations of relative humidity. Environmental data is ideally recorded over several days, with a minimum of five days (Simundic, 2016). Except for data duplication in other seasons, all of the requirements mentioned for an ideal method comparison have been fulfilled in the studies presented in Chapter 3 and 4.

There are several options to visualize the agreement between two methods. The two most common difference plots are the Bland-Altman plot and the Krouwer

plot (Giavarina, 2015). The Bland-Altman plot is designed to show the average of the methods measurements on the X-axis and the ratio or percentages between the methods on the Y-axis. In the Krouwer plot the established reference method is plotted on the X-axis and the differences of the methods are plotted on the Y-axis. In this dissertation the results are shown in rather simplified figures that show the limits of the confidence intervals of slope and intercept of a model II regression. These figures simplify the visual observation whether the numbers 0 or 1 are included into the confidence intervals. This is the simplest and most intuitive form of presenting the results of a model comparison.

2.7 Open-source scientific tools

The wide availability of scientific publications allows to build upon research and previous work from other studies. Still, the exact statistical procedures or models are often not published as software or code and hence a lot of effort is required to reconstruct them from publications and other descriptions. One of the guiding principles of the open-source community is to avoid the time-consuming process of reinventing what other people already developed. Many scientific applications and processes have been developed to solve a vast variety of problems or - as Kinkler's second law defines it - 'all the easy problems have already been solved' (Raymond, 1999). Many times, time and energy are spent solving already solved problems and 'reinventing the wheel' over and over again. In the framework of this thesis I'll try to go a slightly different way by not only publishing the results of my work in peer-reviewed publications, but also providing a simple open-source tool, which streamlines the workflow for similar studies. Hereby, I expect the following benefits for my work: (1) An improvement of my code through a collaborative community approach. By publishing code, other developers and researchers can detect, diagnose and resolve software bugs, or as Linus's Law puts it: 'given enough eyeballs, all bugs are shallow' (Raymond, 1999). (2) Increasing the number of potential users and hence encouraging an adaption to a variety of use cases of the software in the long run is a further strategy to improve this software. (3) Foster the possibility for more frequent releases of new versions by accelerating the development using typical modern software development practices such as time and location agnostic, asynchronous development by using online platforms such as github.com.

Chapter 3

Predicting evapotranspiration from drone-based thermography – a method comparison in a tropical oil palm plantation

PREDICTING EVAPOTRANSPIRATION FROM DRONE-BASED THERMOGRAPHY – A METHOD COMPARISON IN A TROPICAL OIL PALM PLANTATION

Florian Ellsäßer [1], Christian Stiegler [2], Alexander Röll [1], Tania June [3], Hendrayanto [4], Alexander Knohl [2,5], Dirk Hölscher [1,5]

[1] *University of Goettingen, Tropical Silviculture and Forest Ecology, Büsgenweg 1, 37077 Göttingen, Germany*
[2] *University of Goettingen, Bioclimatology, Büsgenweg 2, 37077 Göttingen Germany*
[3] *Bogor Agricultural University, Geophysics and Meteorology, Jln. Meranti, 16680 Bogor, Indonesia*
[4] *Bogor Agricultural University, Forest Management, Kampus IPB Darmaga, 16680 Bogor, Indonesia*
[5] *University of Goettingen, Centre of Biodiversity and Sustainable Land Use, Büsgenweg 1, 37073 Göttingen, Germany*

Under review in Biogeosciences.
Preprint available at: https://doi.org/10.5194/bg-2020-159

ABSTRACT

For the assessment of evapotranspiration, near-surface airborne thermography offers new opportunities for studies with high numbers of spatial replicates and in a fine spatial resolution. We tested drone-based thermography and the subsequent application of three energy balance models (DATTUTDUT, TSEB-PT, DTD) using the widely accepted eddy covariance technique as a reference method. The study site was a mature oil palm plantation in lowland Sumatra, Indonesia. For the 61 flight missions, latent heat flux estimates of the DATTUTDUT model with measured net radiation agreed well with eddy covariance measurements (r^2=0.85; MAE=47; RMSE=60) across variable weather conditions and daytimes. Confidence intervals for slope and intercept of a model II Deming regression suggest no difference between drone-based and eddy covariance method, thus indicating interchangeability. TSEB-PT and DTD yielded agreeable results, but all three models are sensitive to the configuration of the net radiation assessment. Overall, we conclude that drone-based thermography with energy-balance modeling is a reliable method complementing available methods for evapotranspiration studies. It offers promising, additional opportunities for fine grain and spatially explicit studies.

3.1 Introduction

Evapotranspiration (ET) is a central flux in the hydrological cycle on a regional and on a global scale. Terrestrial ET consumes almost two-thirds of terrestrial precipitation (Oki and Kanae, 2006). There is an interest in better understanding ET and its drivers as climate change is expected to increase atmospheric evaporative demand and droughts are predicted to become more severe and frequent in the future (Prudhomme et al., 2014). ET is also strongly affected by land-cover and land-use changes, which are currently very pronounced in tropical regions (Hansen et al., 2013).
The eddy covariance technique (EC) is a widely accepted and well-established

method to quantify ET at the stand scale (Baldocchi et al., 2001; Fisher et al., 2017). It results in a single latent heat flux (LE) value integrated over the footprint of the EC tower at a given time that can be converted to an ET estimate. A spatial fine grain attribution of different surface patches to this overall ET value is generally not possible. The EC method is costly and labor intensive, and therefore, a relatively low number of spatial replicates within a given region and among its different ecosystems are typically available. The EC method also has certain constrains regarding topography, atmospheric turbulence and landscape heterogeneity (Göckede et al., 2008).

A complementary approach for assessing LE at larger spatial scales is the use of remotely sensed land surface temperatures (LST) as boundary conditions for energy balance modeling and subsequent conversion to ET (Brenner et al., 2017; Guzinski et al., 2014; Hoffmann et al., 2016; Ortega-Farías et al., 2016; Xia et al., 2016). Compared to the EC method, this approach can potentially increase the number of spatial replicates within and among ecosystems and is also applicable in challenging terrain. Remotely sensed LSTs are regarded as good indicators for plant water use, stress and transpiration (Jones and Vaughan, 2010). One approach to obtain LST data is the use of satellite-based observations (Allen et al., 2007; Bastiaanssen et al., 1998; Ershadi et al., 2013). However, the spatial resolution of satellite data such as Landsat TM, ASTER, MODIS or AVHRR ranges from 90 m to 1 km, limiting the distinction of plant canopies and soil (Berni et al., 2009). A higher temporal resolution of satellite-based thermal infrared (TIR) observations is usually associated with a lower spatial resolution, and TIR data from satellites in both high spatial and high temporal resolution are not yet available (Brenner et al., 2017). Additionally, clouds are barriers for thermal radiation and therefore have a strong effect on the quality and availability of satellite-based TIR observations (Guzinski et al., 2013). This is of particular importance in regions with frequent cloud cover such as in tropical environments.

An alternative, recently emerging approach to measure LSTs is the use of drones. Radiometric TIR sensors for LST recording have become lightweight and affordable, and drones are now capable of carrying adequate payloads for reasonable timespans. Near-surface thermography-based studies allow temporal resolutions in flexible, e.g. hourly time steps and a spatial resolution in the decimeter scale or finer. Drone-based TIR recording and subsequent modeling of LE with energy balance models has previously shown promising results for short grass and crop vegetation in Central Europe (Brenner et al., 2018; Hoffmann et al., 2016). However, remote sensing of LST from drones is challenging and involves careful planning. Recording LST close to the surface results in a high resolution but reduces the area covered in a certain time span compared to surveying from a higher altitude.

Increasing flight altitude reduces spatial resolution of LST images and thus increases the averaging of surface temperatures from individual canopies, soil patches and branches from neighboring canopies into a single pixel (Still et al., 2019). Further, air humidity can have a major effect on measurement accuracy as water vapor does not only attenuate the signals from the surface of interest to the sensor, but also emits its own thermal radiation (Still et al., 2019).

Different energy balance models are available to compute LE from LST and subsequently calculate ET. In the one-source energy balance model DATTUTDUT (Deriving Atmosphere Turbulent Transport Useful To Dummies Using Temperature) (Timmermans et al., 2015) fluxes are estimated by relating single pixel temperatures to local temperature extremes; the hottest and a group of coldest pixels in the image (Timmermans et al., 2015). Two-source energy balance models such as TSEB (Two-Source Energy Balance) (Norman et al., 1995) and DTD (Dual Temperature Difference) (Norman et al., 2000) divide measured LSTs into a vegetation and a soil fraction. Several adaptions of these models were developed; the TSEB-PT model as described in Hoffmann et al. (2016), uses the Priestley-Taylor coefficient (PT) to determine canopy H flux and subsequently calculate the other fractions from the surface energy balance. TSEB-PT is based on the temperature difference between LST and air temperature (Norman et al., 1995). Expanding this concept, DTD uses a dual-temperature difference from an additional early morning set of measurements to account for biases in remotely sensed LSTs (Hoffmann et al., 2016; Norman et al., 2000). Crucial in applying such energy balance models is how the net radiation (Rn) is implemented. In the original formulation of the DATTUTDUT model Rn is fully modeled, assuming a range of prerequisites and environmental conditions (Timmermans et al., 2015). TSEB-PT and DTD models use measured short and long-wave radiation to estimate Rn as a sum of in- and outgoing long- and short-wave radiation (Norman et al., 1995, 2000). Using airplanes or drones to record LSTs, the three models previously showed promising results for grass and crop surfaces in temperate and subtropical regions (Brenner et al., 2017, 2018; Hoffmann et al., 2016; Xia et al., 2016). However, the limited number of temporal replicates for a given study site constrained previous studies to using error terms and correlation coefficients. To our knowledge, a comprehensive method comparison considering potential errors in both reference method (e.g. the EC technique) and novel drone-based approaches is not yet available.

The presented study was conducted in the lowlands of Jambi province (Sumatra, Indonesia) where over the last decades, large areas of rainforest have been converted to rubber and oil palm plantations (Clough et al., 2016; Margono et al., 2012). This resulted in regional-scale changes in transpiration (Röll et al., 2019) and land surface

warming (Sabajo et al., 2017). We assessed energy fluxes in a mature monoculture oil palm plantation and compared the LE estimates of drone-based methods with the established EC method as measured ground-based reference. Three energy-balance models (DATTUTDUT, TSEB-PT, DTD) were tested, each with three different configurations for determination of Rn (fully modeling Rn, Rn estimates based on short-wave irradiance and measuring Rn). The objectives of our study were to compare LE estimates from the drone-based methods to the EC technique, with a special focus on the detection of proportional and continuous errors among the methods and an evaluation of the models prediction performance.

3.2 Methods

3.2.1 Study site

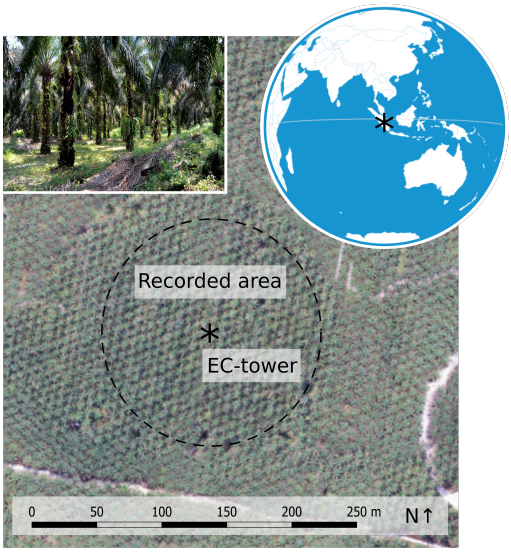


Figure 3.1: *The study site in a mature commercial oil palm plantation in the lowlands of Jambi province, Sumatra, Indonesia.*

The study site is located in the lowlands of Jambi province (Sumatra, Indonesia) near the equator (E 103.3914411, N -1.6929879, 76 m a.s.l.). Average annual air temperature in the region is 26.5°C and average annual precipitation is 2235 mm yr^{-1} (Drescher et al., 2016). At the time of our measurement campaign in August 2017, the studied monoculture oil palm (*Elaeis guineensis*) plantation was 15 years old. Palm stem density was 140 palms ha^{-1}, with an average palm height of 14.3 m and an average canopy radius of 4.5 m. Leaf area index (LAI) was estimated at 3.64 m^2 m^{-2} (Fan et al., 2015) and canopy cover was estimated to be 90%. Plantation management included the

removal of older and non-vital leaves from the oil palms, herbicide application to remove most understory plants and fertilization (196 kg N ha^{-1} yr^{-1}) (Meijide et al., 2017). The average annual oil palm yield is 27.7 Mg ha^{-1}. An EC tower (22 m height) is situated in the center of the site with a fetch of up to 500 m in each direction (Meijide et al., 2017) (Figure 3.1).

3.2.2 Drone-based image acquisition

We used an octocopter drone (MK EASY Okto V3; HiSystems, Germany) equipped with a thermal and an RGB camera mounted in a stereo setup on a gimbal to ensure nadir perspective. The radiometric thermal camera was a FLIR Tau 2 640 (FLIR Systems, USA) attached to a TeAx Thermo-capture module (TeAx Technology, Germany). The sensor covers spectral bands ranging from 7.5 to 13.5 μm with a thermal accuracy of 0.04 K. The RGB camera was based on an Omnivision OV12890 CMOS Sensor (Omnivision, USA) with a 170° FOV fish–eye lens. Instead of the mosaicking approaches applied in most of the mentioned previous studies, we used a single image recording concept as faster image acquisition allows for a denser temporal resolution of LSTs. To capture an area of 100 m radius around the EC tower in a single shot of the thermal camera, images were taken from 260 m altitude. Image corners were removed due to vignetting effects. During a consecutive five-day flight campaign in August 2017, 61 LST data sets and matching EC measurements were recorded. Flights were conducted between 9 am and 4 pm local time, in accordance with the 30 min intervals of the EC averaging cycles, resulting in 10 to 14 flights per day (Table A1). All LSTs were measured using a fixed emissivity of one as the energy balance models would introduce specific soil and vegetation emissivities in the process.

3.2.3 Energy balance models

LSTs are recorded as 'snapshots' representing an instantaneous state of surface temperatures. Soil-Vegetation-Atmosphere Transfer models use these instantaneous observations of LST to solve the energy balance equation and estimate instantaneous fluxes. In our study the one-source energy balance model DATTUTDUT (Timmermans et al., 2015) and two two-source energy balance models, TSEB-PT (Norman et al., 1995) and DTD (Norman et al., 2000), were applied.

The key input for the DATTUTDUT model is a LST map from which the hottest and coldest pixels are extracted, assuming that hot pixels are a result of very little to no ET and cold pixels origin from high ET (Timmermans et al., 2015). Apart from the LST map, all further inputs for the model such as albedo, atmospheric transmissivity and surface emissivity are either

calculated from temperature extremes in the map, location and time or are fixed values as suggested in Timmermans et al. (2015). A version of the DATTUTDUT model has recently been implemented in form of a QGIS3 plugin (QWaterModel) with focus on easy usability (Ellsäßer et al., 2020). TSEB-PT calculates surface-energy budgets from the recorded LSTs splitting observations into a canopy and a soil fraction (Norman et al., 1995; Song et al., 2016; Xia et al., 2016). The Priestley-Taylor (PT) approximation is used to calculate the sensible heat flux for the canopy fraction from net radiation divergence estimates (Hoffmann et al., 2016). With the sensible heat flux known, canopy and soil temperature are calculated and with known resistances fluxes are computed (Hoffmann et al., 2016). Calculation of aerodynamic temperature by using an excess resistance term is not needed, since TSEB-PT uses directional radiometric temperature as input (Hoffmann et al., 2016). The TSEB-PT model requires additional in situ meteorological measurements of long- and short-wave radiation, wind speed, barometric pressure and relative humidity, which in our case were recorded at the EC tower. Further, data on LAI as well as surface and canopy albedo are required. In the DTD model, the absolute temperatures of land surface and air (as used in the TSEB-PT) are supplemented with a second set of early morning reference measurements of LST and air temperature, thus creating a dual-temperature difference (Norman et al., 2000). This relates measurements at any time during the day to measurements recorded in the morning, when fluxes are assumed to be minimal, and thereby accounts for measurement biases of LST (Anderson et al., 1997; Hoffmann et al., 2016). H flux is then calculated using the time-differential temperature and a series resistance network as it is recommended for densely vegetated regions to consider interaction of soil and canopy fluxes (Guzinski et al., 2014; Li et al., 2005). We used two thermal cameras attached to the EC tower (see EC methodology for details and Sect. 2.7 for the limitations) for the necessary early morning reference readings of absolute temperature and used the averaged LSTs from the thermal images to create a uniform map as input for the DTD model (similar as e.g. in Hoffmann et al., 2016). Ground heat flux (G) was computed in the same way for all three models, i.e. as a linear function of Rn (Liebethal and Foken, 2007). More details on the applied models are provided as supplementary information (Appendix II).

Modeled LE estimates are highly sensitive to the type of Rn estimates used. Consequently, in our study we compared three different configurations to include Rn into each of the three mentioned energy balance models: a) fully modeled Rn from sun-earth-geometrics (Rn_mod) as in the original procedure of the DATTUTDUT model with no option to consider clouds, b) measuring only incoming short-wave radiation and calculating net short-wave using

the surface albedo, while net long-wave is calculated from measured air temperature, LST, the Stefan Bolzmann equation and atmospheric emissivity to estimate Rn (Rn_sw) as in Guzinski et al. (2013), and c) measuring the four components of the radiation budget independently and calculating Rn (Rn_mes), as is the case for the original procedure for the TSEB-PT and DTD model and the reference EC method. From the presented results (LE flux densities normalized by area, in W m^{-2}), we further calculated hourly ET rates (amount of water, mm h^{-1}); in this, we assumed a stable relationship of the fluxes during the estimation period (Cammalleri et al., 2014), and followed Timmermans et al. (2015) for calculating the latent heat of vaporization.

3.2.4 Eddy covariance measurements

The micrometeorological tower is located in the center of the study site (Fig. 1). The EC technique was used to measure LE and H fluxes from high frequency (10 Hz) measurements of above-canopy water vapor concentration, sonic temperature, and 3-D wind components. The flux system consisted of a sonic anemometer (Metek uSonic-3 Scientific, Elmshorn, Germany) and a fast response open-path CO2/H2O infrared gas analyzer (Li-Cor7500A, LI-COR Inc. Lincoln, USA) installed at 22 m height. Meteorological variables were measured every 10 sec, averaged to 10 min means and stored on a DL16 Pro data logger (Thies Clima, Göttingen, Germany). Rn and its components were measured with a net radiometer (CNR4, Kipp & Zonen, Delft, The Netherlands) at 22 m height. Air temperature and relative humidity were measured with thermohygrometers (type 1.1025.55.000, Thies Clima, Göttingen, Germany) at 22 m height. Further, a wind direction sensor (Thies Clima, Göttingen, Germany) (22 m height) and 3-cup anemometers (Thies Clima, Göttingen, Germany) (18.5, 15.4, 13, and 2.3 m height) for wind speed measurements were installed on the tower. Two fixed thermal cameras (IR100 Radiometer, Campbell Scientific Inc., Logan, USA) on top of the tower (22 m height) were used for early morning measurements of LST. Ground heat flux was measured using heat flux plates (HFP01, Huxeflux, Delft, The Netherlands) at 10 cm depth. Additional soil moisture and temperature measurements (Trime-Pico 32, Imko, Ettlingen, Germany) above the heat flux plate at 5 cm depth were used to calculate heat flux at the soil surface. EC data recording, filtering and processing were carried out identical to the methodology described in Meijide et al. (2017) for the same study site. As the applied drone-based models all assume full energy balance closure, we used the Bowen ratio closure method (Pan et al., 2017; Twine et al., 2000) to compute full closure for the EC measurements.

3.2.5 Statistical analyses

We applied the model II Deming regression method to consider uncertainties in both x and y variables (Cornbleet and Gochman, 1979; Glaister, 2001) with the assumption that the error ratio ($\sigma\varepsilon^2/\sigma\delta^2$) of the variances ($\sigma$) of errors on y ($\varepsilon_i$) and on x ($\delta_i$) would not differ from 1 which is the standard procedure if both uncertainties are unknown (Legendre and Legendre, 2003). We used the interquartile range method with a factor k=1.5 to remove outliers from the regression. A Durbin-Watson test was applied to test for correlation in error terms. We checked for heteroscedasticity visually and using a White test. Normal distribution of error terms was tested visually plotting standardized residuals vs. theoretical quantities and performing a Shapiro-Wilk test. Standard errors and confidence intervals for slope and intercept of the Deming regression were calculated using analytical methods (parametric) and the jackknife method (Armitage et al., 2001; Linnet, 1993). As further indicators of model performance, we calculated the coefficients of determination (r^2), the Mean Absolute Error (MAE), the Root Mean Square Error (RMSE) and slope and intercept from the Deming regression. Statistics such as r^2 have their limitations in method comparison since they are designed to indicate how well the resulting model of the regression describes the outcome and are not necessarily a good measure for systematical bias between methods. However, they are used as a statistic in this study since they represent an additional indicator for interpretation. Linearity was checked visually plotting residuals vs. fitted values. To examine the spatial heterogeneity of the LE from drone-data based estimates, we calculated standard deviation, coefficient of variation, Kurtosis and median-based Fisher-Pearson Coefficient of Skewness (FPCS) for each model output to characterize the respective distributions in relation to a normal distribution (Doane and Seward, 2011; Legendre and Legendre, 2003).

All modeling procedures and parts of the statistical analyses were computed using Python version 3.7.1 (Python Software Foundation), involving the libraries NumPy 1.14.2, SciPy 1.1.0, pandas 0.23.1, scikit-learn 0.19.1, gdal 2.3.2, Astropy 3.2.2 and tkinter 8.6. The Deming regression was computed using the MethComp and mcr v2.2.1 package (Manuilova et al., 2014) in R version 3.6.1 (R Development Core Team, 2019). Graphic representation was processed in Python using the Matplotlib 3.0.2 and Seaborn 0.9.0 libraries.

3.2.6 Data set characteristics

The dataset offers a comparatively high number of replicates from 61 drone recording flights and the corresponding eddy covariance measurements enabling a method comparison which requires at least n = 60 observations (Legendre and Legendre, 2003). The

data was recorded in a 30 min frequency, to facilitate the analysis of daily courses of evapotranspiration behavior creating a trade-off situation of more flights per day with shorter flight times per flight. Because flight times were so short, only a smaller footprint with a radius of 100 m around the eddy covariance station was covered, while the footprint recorded with the eddy covariance system ranged up to a 500 m radius around the tower. Therefore, the reduced area of the drone recorded LST maps is often smaller than the extent of the eddy covariance footprint. We have several reasons to assume that this doesn't cause major problems for the comparison though: the study area is very homogenous with an elevation difference of 5 m in the eddy covariance footprint and the biosphere is strongly dominated by only one species (oil palm). The plantation is very well managed, so that all oil palm canopies are vital, no oil palms have died and only dry leaves are removed. A further limitation of the data set is the lack of morning or night LST measurements that could not be recorded with the drone due to security concerns and limited access to the plantation at night. This doesn't affect the procedure of the DATTUTDUT and TSEB-PT model, but morning measurements are an important factor for the DTD model. We were able to record night and morning measurements with two stationary thermal cameras that were attached to the tower. As for the DTD model, morning and later recordings should ideally be recorded with the same camera. To check whether both cameras would measure similar temperatures, we compared a total of 122 LST maps from the drone and the stationary cameras and plotted the measured and adjusted temperatures of both recording systems (Fig. A1). There is a small deviation of the measured temperatures resulting in a mean absolute error (MAE) and root mean squared error (RMSE) of 1.59 and 2.15 K respectively. Since LST measurements are subject to a certain degree of uncertainty and thermal cameras usually have a measurement error of up to $\pm 1^{\circ}$C we decided to use the morning measurements from the tower cameras as input for the morning temperature reference (Aubrecht et al., 2016). The implementation of the DTD model is therefore strictly experimental and has to be interpreted with the uncertainties of the morning measurements in mind.

3.3 Results

3.3.1 Meteorology

During our 61 flight missions, cloudiness was variable from clear sky to full cloud cover; short-wave irradiance ranged from 204 to 1110 W m^{-2}. The prevailing wind direction was from north-east, at an average wind speed of 1.7 m s^{-1}. Canopy air temperature ranged from 22.5 to 32.3°C and relative humidity varied between 62 and 99%. The energy balance closure of the reference EC measurements was 0.77 (r^2 = 0.87).

3.3.2 Drone-based modeling methods vs. eddy covariance method

At the time of the drone flights, LE from the EC method ranged between 87 and 596 W m^{-2} (mean: 337 W m^{-2}). Congruence of the LE estimates of the three applied models differed in their congruence with EC measurements, depending on the configuration of the Rn assessment (Appendix figure 3). Generally, error metrics were reduced and congruence was increased the more measurement-controlled the Rn determination process was.

For the Rn_mod configuration the daily patterns of DATTUTDUT LE estimates closely agreed with EC measurements around noon but resulted in higher LE fluxes in the morning and afternoon where TSEB-PT LE estimates are much higher than EC measurements (Figure 3.2a). While some of the LE predictions from the DTD model in Rn_mod configuration are similar to the EC measurements, many of its LE predictions are over- or underestimations. Models with Rn_sw configuration produced LE estimates that matched LE from EC more closely. DATTUTDUT computed higher estimates of LE compared to the EC measurements during noon, while TSEB-PT produced more congruent LE estimates for the midday hours but rather underestimated LE fluxes especially in the morning.

The DTD model underestimated LE fluxes for all daytimes (Figure 3.2b). All models with Rn_sw configuration yielded comparably low estimates during the morning and afternoon hours. With Rn_mes configuration, DATTUTDUT computed closely matching LE estimates at all times of day across the five-day measurement period, while TSEB-PT consistently produced much higher estimates than EC around noon and the DTD model underestimated LE fluxes especially in the morning hours while otherwise producing accurate results (Figure 3.2c).

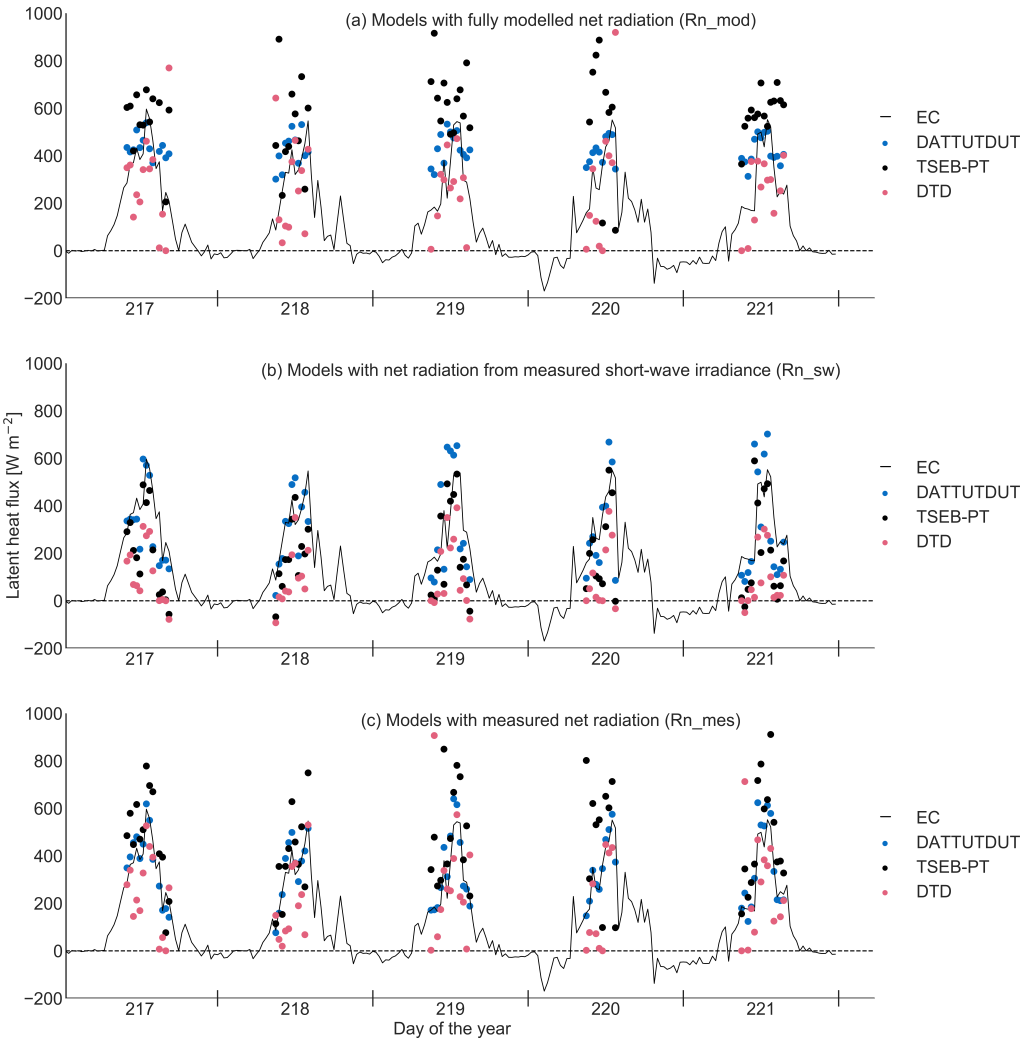


Figure 3.2: *Latent heat flux (LE) from energy balance models (DATTUTDUT, TSEB-PT, DTD) and three different configurations of net radiation (Rn) determination (Rn_mod, Rn_sw, Rn_mes) and eddy covariance measurements (EC) over five consecutive days (n = 61 flight missions).*

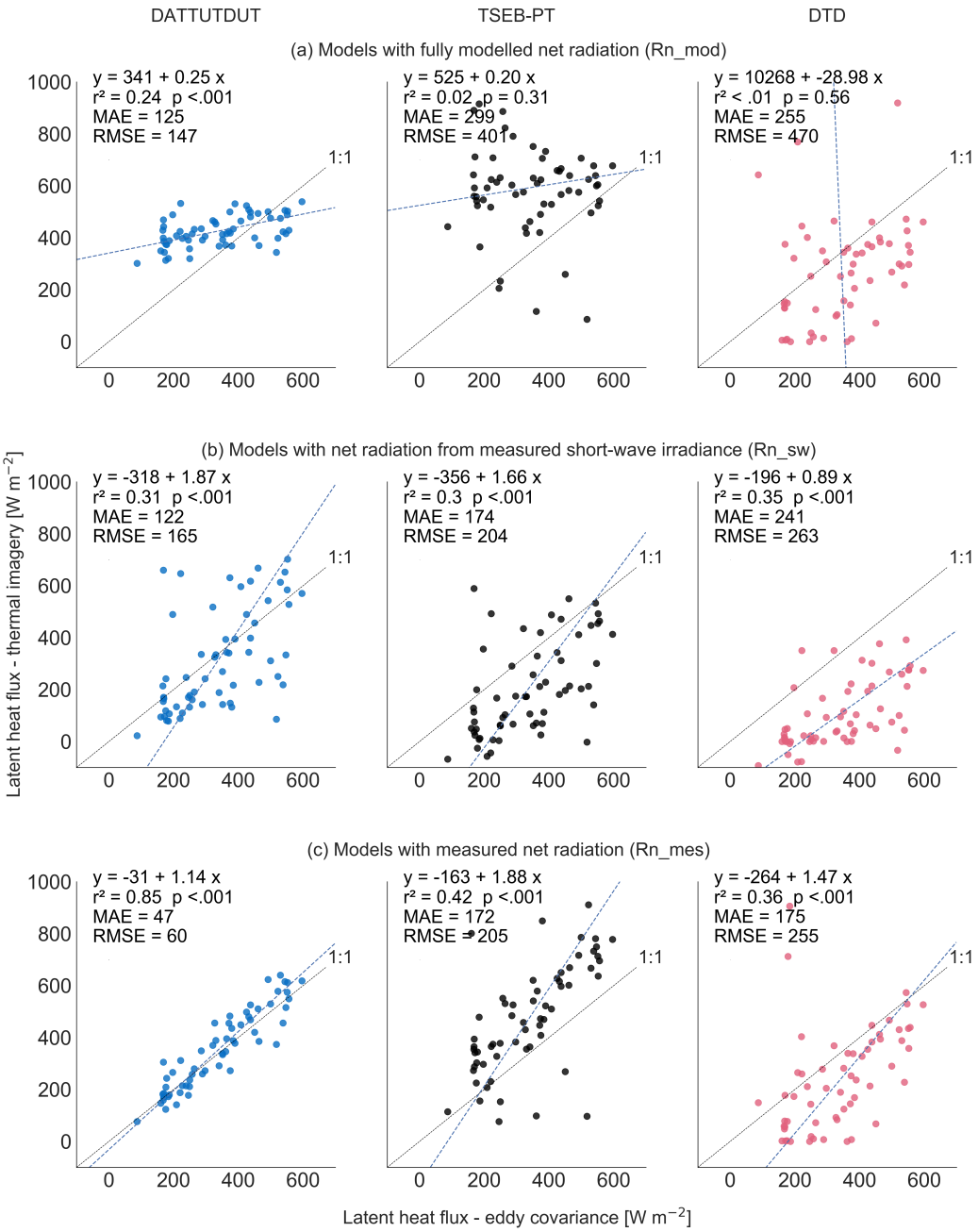


Figure 3.3: *Model II Deming regression of latent heat flux estimates from drone-based energy balance models (DATTUTDUT, TSEB-PT, DTD) and different configurations of net radiation (Rn_mod, Rn_sw, Rn_mes) with the eddy covariance method (n = 61 flight missions).*

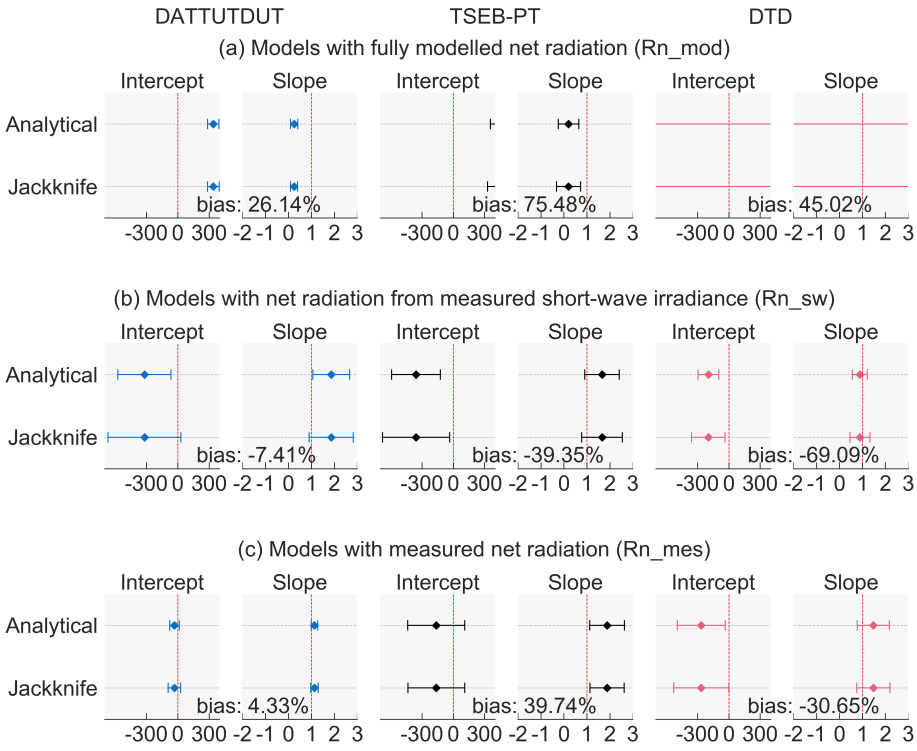


Figure 3.4: *Confidence intervals for intercept and slope of Deming regression for the different LE estimation approaches compared with EC measurements. X-level for the bias is the mean of the established EC reference method. The intercept is displayed in W* m^{-2}.

Across all daytimes and weather conditions (n=61 flight missions), congruence among drone-based LE estimates and reference EC measurements was highest for the DATTUTDUT model with Rn_mes configuration (r^2=0.85); MAE and RMSE were 47 and 60 W m^{-2}, respectively (Figure 3.3). To compare the model predictions and the eddy covariance measurements, we computed a Deming regression between both LE predictions and LE estimates. The methods are considered to be statistically interchangeable if the confidence intervals of the slope and intercept include one and zero respectively. Deming regression of the LE estimates of the DATTUTDUT model with Rn_mes configuration showed no significant proportional or constant error compared to EC measurements as the values one and zero lay within the respective 99% confidence interval ranges of slope and intercept (Figure 3.4). It is thus indicated that there is no significant difference between DATTUTDUT with Rn_mes configuration and the EC technique. The TSEB-PT model in Rn_mes configuration also showed no significant continuous errors but was subject to proportional bias (Figure 3.4c) predictions tend to overestimate LE around noon when fluxes are very high (Figure 3.2c and 3.3c). The DTD model showed no proportional bias but indicated a continuous error in the analytical method and the Jackknife method (Figure 3.4c). In the Rn_sw configuration, all three models showed no significant proportional error of LE estimates compared to EC measurements (Figure 3.4b). However, all TSEB-PT and DTD model estimates deviated from the EC measurements by a significant constant amount (Figure 3.2a and 3.3b). All models in the Rn_mod configuration showed significant proportional and constant errors or large biases compared to EC measurements, as well as very large confidence intervals (Figure 3.3a and 3.4a).

3.3.3 Spatial distribution of LE

For 9th of August 2017, 12.30 h, the DATTUDUT in Rn_mes configuration suggested a mean of 526 W m^{-2} (minimum of 0 on the corrugated iron roof of the EC tower system, maximum of 637 W m^{-2}, coefficient of variation 7.53%, for the analyzed 18,383 pixels) (Figure 3.5), which translates to a mean ET of 0.778 mm m^{-2} h^{-1}. Locally, i.e. in the center of oil palm crowns, high LE of $>$ 400 W m^{-2} was observed, while LE from soil and ground vegetation areas between oil palm canopies was lower. The LE fluxes of all pixels were almost normally distributed for the one-source energy balance model DATTUDUT (Figure 3.6), whereas the distributions of the two-source energy balance model TSEB-PT (for the same LST dataset) was more skewed, with more LE observations at the upper end of the range. The DTD model resulted in an almost normal distribution but with values far more evenly distributed over the entire range from minimal to maximal value and with a

less pronounced peak. While the averaged FPCS ranged around ±1 for all three models and model configurations over the 61 flights indicating only a minor skewness, Kurtosis was always within the platykurtic distribution for the DATTUTDUT and DTD models revealing a very low number of outliers whereas the distribution was highly variable for TSEB-PT (Appendix figure A2).

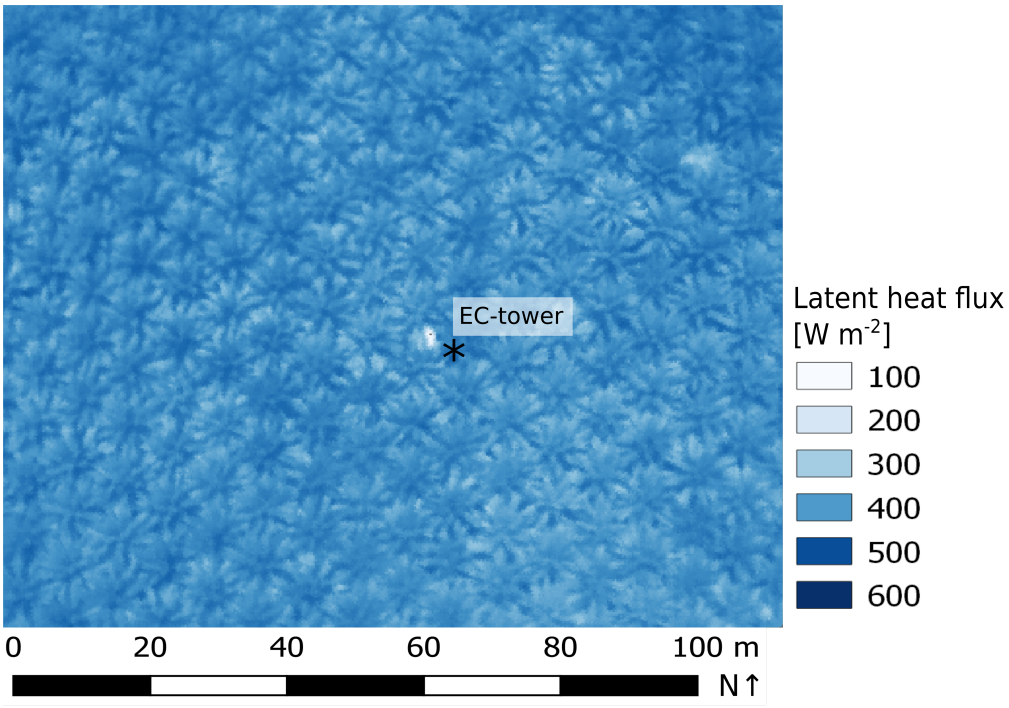


Figure 3.5: *Spatial distribution of latent heat flux from drone-based thermography and subsequent energy balance modeling (DATTUTDUT with Rn_mes configuration, 9 August 2017, 12.30 h).*

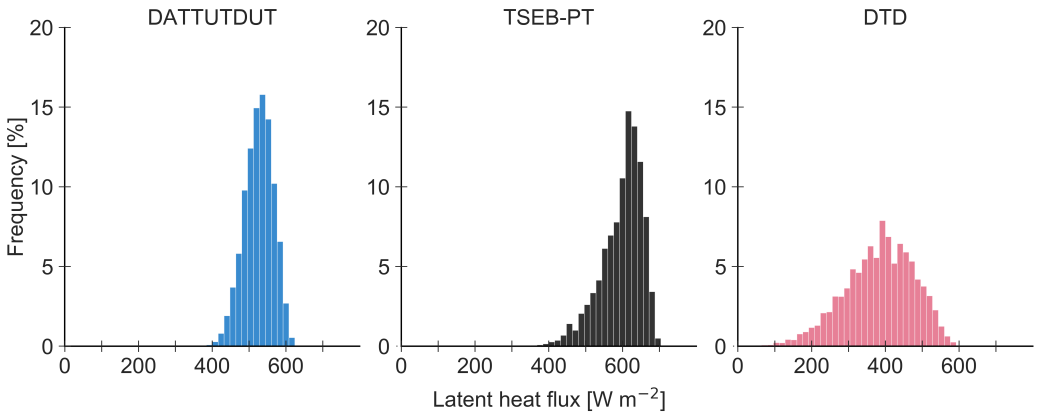


Figure 3.6: *Frequency distribution of latent heat flux for the model output images from the same thermal image as shown in Fig. 5 (9 August 2017, 12.30 h). Absolute histogram bin size was set to 16 W m^{-2}, we used 50 bins from 0 to 800 W m^{-2}.*

3.4 Discussion

Our study indicates a high agreement between latent heat fluxes assessed by drone-based thermography and the eddy covariance technique. However, the performance of the three applied energy balance models differed among each other and among different configurations of net radiation assessments in the models (Figure 3.2 and Appendix figure A3). Model II Deming regression analyses and associated quality assessments suggest interchangeability between the DATTUTDUT model in Rn_mes configuration and the EC technique (Figure 3.3 and 3.4). Applying this configuration, a fine grain spatial analysis of latent heat fluxes suggests relatively low heterogeneity of LE in the studied tropical oil palm plantation (Figure 3.5).

3.4.1 Drone-based LE modeling vs. eddy covariance measurements

The confidence intervals of slope and intercept of the Deming regression indicate that the one-source energy balance model DATTUTDUT with Rn_mes configuration is statistically interchangeable with the established EC method for estimating LE fluxes. There are advantages and limitations to both methods. For example, the DATTUTDUT model provides insights on the spatial distribution of LE fluxes within the full extent of the available LST maps, whereas the EC technique averages the LE fluxes within its footprint to a single value. On the other hand, the DATTUTDUT model is temporally limited to the availability of LST maps, whereas the EC method can measure fluxes continuously over

several years once the equipment is in place. The DATTUTDUT model with Rn_mes configuration further requires additional measurements of short- and long-wave radiation budgets. In our study, these measurements were taken with the EC equipment, but future stand-alone drone approaches are possible by using on-board miniaturized radiation sensors (Castro Aguilar et al., 2015; Suomalainen et al., 2018). The two-source energy balance models TSEB-PT and DTD in the Rn_mes configuration showed different behaviors. TSEB-PT was found to have no significant continuous errors, but proportional errors compared to the reference EC method. This is largely rooted in the overestimation of relatively high fluxes around noon, while lower fluxes during the morning and afternoon hours were predicted more accurately.

An opposite situation was found for the DTD model that showed no proportional errors but a continuous error for the analytical method. Intercept confidence intervals from the Jackknife method included zero and suggest no continuous error for the DTD model in the Rn_mes configuration. Due to the mismatch between the results of the Jackknife and the analytical method, a considerable bias of -30.65% over the mean and remaining accuracy concerns of the morning LST measurements, this method configuration cannot yet be considered as statistically interchangeable.

All models with the Rn_sw configuration showed a significant constant error compared to EC measurements, i.e. all modeled LE estimates derived from this configuration underestimate measured fluxes by a certain fixed amount (about 200 W m^{-2} on average compared to EC measurements). These underestimations of Rn translated directly to an underestimation of turbulent fluxes (Appendix figure A3). Previous studies have pointed out that Rn derivation based on short-wave irradiance measurements is challenging as long-wave radiation budgets are often completely independent from their short-wave counterparts (Hoffmann et al. 2016). Estimation errors in long-wave radiation budgets have e.g. been reported to be related to high relative air humidity, when some of the original model assumptions are no longer met (Hoffmann et al., 2016). We observed a negative correlation ($r^2 = 0.46$) between incoming long-wave irradiance and relative humidity and assume that the high relative humidity in our tropical study area may have affected the determination of Rn when using the Rn_sw configuration through inaccuracies in estimating long-wave radiation budgets, therefore causing the observed significant continuous errors. Such constant errors in Rn estimation can be reduced or even eliminated by enhanced calibration of the models. We thus also consider the Rn_sw configuration valuable for future research, particularly because measurements of incoming short-wave radiation are much easier to implement than assessing complete short- and long-wave radiation budgets as necessary for the Rn_mes configuration. The application of the Rn_sw

configuration for a one-source energy balance model such as DATTUTDUT was also tested in two previous studies, with similar results to our study, i.e. a reduction of errors compared to its original formulation with fully modeled Rn_mod (Brenner et al., 2018; Xia et al., 2016).

Lastly, the model configuration Rn_mod did not produce accurate LE estimates for all three models, as many of the basic assumptions for fully modelled Rn determination are not met in tropical environments such as our equatorial study area. As such, the sky is often cloudy, while haze frequently occurs during periods without rainfall. Even if no clouds are visible, relative humidity is often high, which interferes with the clear-sky assumptions of the Rn_mod configuration (Still et al., 2019).

LE estimates from TSEB-PT and DTD are sometimes close to zero or even negative (for all three Rn configurations) and thus deviate substantially from the EC measurements (Figure 3.2). This occurs when the difference between surface aerodynamic temperature and air temperature becomes very small, which causes the evaporative fraction to approach zero, resulting in an overestimation of H and an underestimation of LE in the fragmentation process of turbulent fluxes (see Eq. 1 in Norman et al. 2000). This effect is especially pronounced for the DTD model, when the morning reference air and land surface temperature measurements are very similar to each other. We assume that the problem is rather typical for tropical environments as daily temperature changes are often not as pronounced as e.g. in semiarid areas where, conversely to our findings substantial overestimations of LE fluxes were observed (Kustas et al., 2016; Morillas et al., 2013).

Among the three models applied in our study, the relatively simple DATTUTDUT model produced the most precise LE estimates compared to eddy covariance reference measurements. Similar conclusions were reached by Brenner et al. (2018), where DATTUTDUT marginally outperformed the more complex TSEB-PT model. On the other hand, contrasting observations were made by Xia et al. (2016) in vineyards with more extreme temperature divergences between soil and vegetation, where the TSEB-PT model produced more precise estimates of LE than the DATTUTDUT model. This was explained by the better physical representation of energy and radiative exchange in the TSEB-PT model. The authors further point out that Rn determination is not the only source of error in the DATTUTDUT model (Xia et al., 2016). In our study, the TSEB-PT model slightly outperformed the more complex DTD model in the Rn_mes configuration regarding error terms, congruence and continuous errors. The DTD model on the other hand showed less bias and no proportional error and is therefore the more promising approach as its continuous error can be reduced through calibration.

We used the Bowen-ratio method to close the energy balance for the refer-

ence EC measurements. As reported by Xia et al. (2016), agreement between measured EC and modeled LE estimates could potentially be increased by using the residual method from Twine et al. (2000) for energy balance closure. Further potential improvements include the aerial sampling alignment with the EC measurement logging cycles. We compared snapshot measurements of LST to 30 min averages of EC for the corresponding times in an environment where key variables such as solar irradiance can change very quickly. Better matching the measurement cycle duration may further improve agreement between the methods and was already suggested in a previous study (Brenner et al., 2018). Further, in our study the aerial-derived LST images represented only the center area of the (at times quite variable and large) EC footprint. Covering the whole potential area of the footprint in all directions could also increase agreement between the measurements, but would require even higher flight altitude or longer flight times to cover the whole area; both options would reduce the number of temporal replicates and increase errors from measurements and processing, but could nonetheless be viable approaches for other research questions.

Only few previous studies have demonstrated applicability and limitations of estimating LE with the three energy balance models from non-satellite data. In these studies, LSTs were e.g. recorded from drones for European grasslands and croplands (Brenner et al., 2018; Hoffmann et al., 2016) and from drones or airplanes for taller vegetation including olive orchards and vineyards (Ortega-Farías et al., 2016; Xia et al., 2016). Our study adds to this an application of these models in a tropical environment, for higher vegetation (i.e. oil palm) and across variable daytimes and weather conditions. We further analyzed for the first time whether drone-data based models and EC measurements can be used interchangeably, as our large sample size of n=61 flights allowed for a method comparison based on a model II Deming regression (Legendre and Legendre, 2003). We conclude that this is the case for some models and configurations, above all for the DATTUTDUT with Rn_mes configuration.

3.4.2 Spatial distribution of latent heat fluxes

A particular strength of drone-based thermal imagery is the high spatial resolution which allows for spatially explicit assessments of evapotranspiration, within and potentially also beyond the footprints of EC towers. The outlines of the single oil palm canopies are clearly visible in the LE flux map (Figure 3.5), with the highest LE fluxes occurring in the center of the oil palm canopies. We assume that this spatial pattern is caused by an increased local LAI in the centers of the oil palm canopies, while leaf area density decreases towards the outer canopies. Further, the central areas of oil palm canopies are more ex-

posed to sunlight and wind throughout most of the day, increasing their potential for (evapo)transpiration compared to canopy edges. Mixed pixel effects (among soil and canopy) likely also contribute to the observed lower LE fluxes towards the borders of oil palm canopies. Further contributing factors to higher LE fluxes in the centers of oil palm canopies could be leaf age (with younger leaves in the center) and additional ET from pockets in the axils of pruned leaves along the stem, which contain small water reservoirs and epiphytes (Meijide et al., 2017; Tarigan et al., 2018).

In the histograms of LE fluxes from all pixels within the single studied footprints (Figure 3.6), the DATTUTDUT and DTD models results in a bell-shaped normal distribution but very different value ranges. While the DATTUTDUT histogram shows only few pixel values of zero and most pixels closely distributed around the mean, the DTD histogram is much wider distributed and the peak is much more moderate. Mean and median are very similar indicating close to zero skewness for both, the DATTUTDUT and DTD model. Such a distribution tending towards unimodality is also considered typical for landscapes where ET is highly dominated by one species (Xia et al., 2016). The TSEB-PT model shows a different, more skewed distribution of LE fluxes (for the same dataset of LST), with the median of the LE estimates located between the mean and the upper end of the LE flux range. Kurtosis is much more pronounced for the TSEB-PT model in all net radiation configurations than in the other two models (Appendix figure A2).

The increased Kurtosis is a very strong sign towards more extreme outliers that emerge in the TSEB-PT model. We assume that the TSEB-PT model is more sensitive to dry surfaces and hence produces more extreme outliers. Since the DTD model references the LST measurements with a second set of land surface and air temperatures it is less affected by extreme outliers in the LST data.

We see great potential in the drone-based remote sensing applications presented in this study; especially when recent developments in drone-environment interaction, mobile edge computing (potentially on-board of the drone) and communication technologies such as LoRaWan (Long Range Wide Area Network) or 5G are combined (Becerra, 2019; Marchese et al., 2019). Autonomous acquisition of LSTs over EC stations and the surrounding areas can be supplemented by on-board and ground sensors and energy-balance models can be computed on the edge enabling a dense temporal resolution of LST, flux and ET maps in almost real-time. This concept can e.g. be used for the attribution of fluxes in mixed species plant communities, the study of edge effects in landscapes, and further be adapted e.g. to detect water stress in agriculture and forests.

3.5 Conclusions

Drone-based thermography and subsequent energy balance modeling under certain configurations can be considered a highly reliable method for estimating latent heat flux and evapotranspiration; for some configurations statistical interchangeability is suggested with the established eddy covariance technique. They thus complement the asset of available methods for evapotranspiration studies by fine grain and spatially explicit assessments.

Data availability:

The final data used for the statistical tests were uploaded in Göttingen Research Online Data with a doi: https://doi.org/10.25625/JY9AFT
Raw thermal images, orthomosaics and terrain data, georeferenced rasters and model configurations are available upon request to the corresponding author.

Author Contribution:

The study was conceptualized by DH in cooperation with H (drone measurements) and AK in cooperation with TJ (eddy covariance measurements). FE led the writing of the paper with help from AR and DH supervised the work. FE collected and processed the drone data and CS the eddy covariance data. FE conducted data processing, model application, statistical analysis and production of plots in cooperation mainly with DH and AR. FE, DH and AR created a first version of the manuscript, which was further improved in a cooperation of all authors.

Acknowledgements:

This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 192626868 – SFB 990 (subprojects A02 and A03) and the Ministry of Research, Technology and Higher Education (Ristekdikti). We thank Ristekdikti for providing the research permit for field work (No. 322/SIP/FRP/E5/Dit.KI/IX/2016, No. 329/SIP/FRP/E5/Dit.KI/IX/2016 and No. 28/EXT/SIP/FRP/E5/Dit.KI/VII/2017). We thank our field assistants Zulfi Kamal, Basri, Bayu and Darwis for great support during the field campaigns and Edgar Tunsch, Malte Puhan, Frank Tiedemann and Dietmar Fellert for their technical support. We thank Hector Nieto for publishing the code for TSEB-PT and DTD (pyTSEB) on www.github.com. We also thank Perseroan Terbatas Perkebunan Nusantara VI, Batang Hari Unit (PTPN6) for giving us permission to conduct our research at the oil palm plantation. Thanks to all 'EFForTS' colleagues and friends in Indonesia, Germany, and around the world.

REFERENCES

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56. FAO, Rome.

Allen, R.G., Tasumi, M., Trezza, R., 2007. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)—Model. J. Irrig. Drain. Eng. 133, 380–394. https://doi.org/10.1061/(ASCE)0733-9437(2007)133:4(380)

Anderson, M.C., Norman, J.M., Diak, G.R., Kustas, W.P., Mecikalski, J.R., 1997. A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. Remote Sens. Environ. 60, 195–216. https://doi.org/10.1016/S0034-4257(96)00215-5

Armitage, P., Berry, G., Matthews, J.N.S., 2001. Statistical methods in medical research, 4th ed. ed. Blackwell Science, Malden, MA.

Aubrecht, D.M., Helliker, B.R., Goulden, M.L., Roberts, D.A., Still, C.J., Richardson, A.D., 2016. Continuous, long-term, high-frequency thermal imaging of vegetation: Uncertainties and recommended best practices. Agric. For. Meteorol. 228–229, 315–326. https://doi.org/10.1016/j.agrformet.2016.07.017

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw U, K.T., Pilegaard, K., Schmid, H.P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. Bull. Am. Meteorol. Soc. 82, 2415–2434. https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2

Bastiaanssen, W.G.M., Menenti, M., Feddes, R.A., Holtslag, A.A.M., 1998. A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. J. Hydrol. 212–213, 198–212. https://doi.org/10.1016/S0022-1694(98)00253-4

Becerra, V.M., 2019. Autonomous control of unmanned aerial vehicles. Electronics 8, 452. https://doi.org/10.3390/electronics8040452

Berni, J.A.J., Zarco-Tejada, P.J., Sepulcre-Cantó, G., Fereres, E., Villalobos, F., 2009. Mapping canopy conductance and CWSI in olive orchards using high resolution thermal remote sensing imagery. Remote Sens. Environ. 113, 2380–2388. https://doi.org/10.1016/j.rse.2009.06.018

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. Int. J. Remote Sens. 38, 3003–3026. https://doi.org/10.1080/01431161.2017.1280202

Brenner, C., Zeeman, M., Bernhardt, M., Schulz, K., 2018. Estimation of evapotranspiration of temperate grassland based on high-resolution thermal and visible range imagery from unmanned aerial systems. Int. J. Remote Sens. 39, 5141–5174. https://doi.org/10.1080/01431161.2018.1471550

Brutsaert, W., 1982. Evaporation into the Atmosphere. Theory, History, and Applications. Reidel Publishing Co.

Burchard-Levine, V., Nieto, H., Riaño, D., Migliavacca, M., El-Madany, T.S., Perez-Priego, O., Carrara, A., Martín, M.P., 2019. Adapting the thermal-based two-source energy balance model toestimate energy fluxes in a complex tree-grass ecosystem. Hydrol. Earth Syst. Sci. Discuss. 1–37. https://doi.org/10.5194/hess-2019-354

Burridge, D.M., Gadd, A.J., 1977. The Meteorological Office operational 10-level numerical weather prediction model (December 1975), Scientific paper - Meteorological Office. British Meteorological Office, Bracknell, England.

Cammalleri, C., Anderson, M.C., Kustas, W.P., 2014. Upscaling of evapotranspiration fluxes from instantaneous to daytime scales for thermal remote sensing applications. Hydrol. Earth Syst. Sci. 18, 1885–1894. https://doi.org/10.5194/hess-18-1885-2014

Campbell, G.S., Norman, J.M., 1998. An Introduction to Environmental Biophysics. Springer, New York, New York.

Castro Aguilar, J.L., Gentle, A.R., Smith, G.B., Chen, D., 2015. A method to measure total atmospheric long-wave down-welling radiation using a low cost infrared thermometer tilted to the vertical. Energy 81, 233–244. https://doi.org/10.1016/j.energy.2014.12.035

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., Holtkamp, A.M., Hörandl, E., Irawan, B., Jaya, I.N.S., Jochum, M., Klarner, B., Knohl, A., Kotowska, M.M., Krashevska, V., Kreft, H., Kurniawan, S., Leuschner, C., Maraun, M., Melati, D.N., Opfermann, N., Pérez-Cruzado, C., Prabowo, W.E., Rembold, K., Rizali, A., Rubiana, R., Schneider, D., Tjitrosoedirdjo, S.S., Tjoa, A., Tscharntke, T., Scheu, S., 2016. Land-use choices follow profitability at the expense of ecological functions in Indonesian smallholder landscapes. Nat. Commun. 7, 1–12. https://doi.org/10.1038/ncomms13137

Cornbleet, P.J., Gochman, N., 1979. Incorrect least-squares regression coefficients in method-comparison analysis. Clin. Chem. 432–438.

Doane, D.P., Seward, L.E., 2011. Measuring skewness: a forgotten statistic? J. Stat. Educ. 19, 3. https://doi.org/10.1080/10691898.2011.11889611

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., Krashevska, V., Krishna, V., Leuschner,

C., Lorenz, W., Meijide, A., Melati, D., Nomura, M., Pérez-Cruzado, C., Qaim, M., Siregar, I.Z., Steinebach, S., Tjoa, A., Tscharntke, T., Wick, B., Wiegand, K., Kreft, H., Scheu, S., 2016. Ecological and socio-economic functions across tropical land use systems after rainforest conversion. Philos. Trans. R. Soc. B Biol. Sci. 371, 20150275. https://doi.org/10.1098/rstb.2015.0275

Ellsäßer, F., Röll, A., Stiegler, C., Hendrayanto, Hölscher, D., 2020. Introducing QWater-Model, a QGIS plugin for predicting evapotranspiration from land surface temperatures. Environmental Modelling & Software. 130, https://doi.org/10.1016/j.envsoft.2020.104739

Ershadi, A., McCabe, M.F., Evans, J.P., Walker, J.P., 2013. Effects of spatial aggregation on the multi-scale estimation of evapotranspiration. Remote Sens. Environ. 131, 51–62. https://doi.org/10.1016/j.rse.2012.12.007

Fan, Y., Roupsard, O., Bernoux, M., Le Maire, G., Panferov, O., Kotowska, M.M., Knohl, A., 2015. A sub-canopy structure for simulating oil palm in the Community Land Model (CLM-Palm): phenology, allocation and yield. Geosci. Model Dev. 8, 3785–3800. https://doi.org/10.5194/gmd-8-3785-2015

Fisher, J.B., Melton, F., Middleton, E., Hain, C., Anderson, M., Allen, R., McCabe, M.F., Hook, S., Baldocchi, D., Townsend, P.A., Kilic, A., Tu, K., Miralles, D.D., Perret, J., Lagouarde, J.-P., Waliser, D., Purdy, A.J., French, A., Schimel, D., Famiglietti, J.S., Stephens, G., Wood, E.F., 2017. The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources: The future of evapotranspiration. Water Resour. Res. 53, 2618–2626. https://doi.org/10.1002/2016WR020175

Garratt, J.R., 1992. The Atmospheric Boundary Layer. Cambridge University Press, Cambridge.

Glaister, P., 2001. . Least Sq. Revisit. Math. Gaz. 85. https://doi.org/10.2307/3620485

Göckede, M., Foken, T., Aubinet, M., Aurela, M., Banza, J., Bernhofer, C., Bonnefond, J.M., Brunet, Y., Carrara, A., Clement, R., Dellwik, E., Elbers, J., Eugster, W., Fuhrer, J., Granier, A., Grünwald, T., Heinesch, B., Janssens, I.A., Knohl, A., Koeble, R., Laurila, T., Longdoz, B., Manca, G., Marek, M., Markkanen, T., Mateus, J., Matteucci, G., Mauder, M., Migliavacca, M., Minerbi, S., Moncrieff, J., Montagnani, L., Moors, E., Ourcival, J.-M., Papale, D., Pereira, J., Pilegaard, K., Pita, G., Rambal, S., Rebmann, C., Rodrigues, A., Rotenberg, E., Sanz, M.J., Sedlak, P., Seufert, G., Siebicke, L., Soussana, J.F., Valentini, R., Vesala, T., Verbeeck, H., Yakir, D., 2008. Quality control of CarboEurope flux data - Part 1: Coupling footprint analyses with flux data quality assessment to evaluate sites in forest ecosystems. Biogeosciences 5, 433–450. https://doi.org/10.5194/bg-5-433-2008

Guzinski, R., Anderson, M.C., Kustas, W.P., Nieto, H., Sandholt, I., 2013. Using a thermal-based two source energy balance model with time-differencing to estimate surface energy fluxes with day–night MODIS observations. Hydrol. Earth Syst. Sci. 17, 2809–2825. https://doi.org/10.5194/hess-17-2809-2013

Guzinski, R., Nieto, H., Jensen, R., Mendiguren, G., 2014. Remotely sensed land-surface energy fluxes at sub-field scale in heterogeneous agricultural landscape and coniferous plantation. Biogeosciences 11, 5021–5046. https://doi.org/10.5194/bg-11-5021-2014

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. Science 342, 850–853. https://doi.org/10.1126/science.1244693

Hoffmann, H., Nieto, H., Jensen, R., Guzinski, R., Zarco-Tejada, P., Friborg, T., 2016. Estimating evaporation with thermal UAV data and two-source energy balance models. Hydrol. Earth Syst. Sci. 20, 697–713. https://doi.org/10.5194/hess-20-697-2016

Jones, H.G., Vaughan, R.A., 2010. Remote sensing of vegetation: principles, techniques, and applications. Oxford University Press, Oxford; New York.

Kustas, W.P., Nieto, H., Morillas, L., Anderson, M.C., Alfieri, J.G., Hipps, L.E., Villagarcía, L., Domingo, F., Garcia, M., 2016. Revisiting the paper "Using radiometric surface temperature for surface energy flux estimation in Mediterranean drylands from a two-source perspective." Remote Sens. Environ. 184, 645–653. https://doi.org/10.1016/j.rse.2016.07.024

Kustas, W.P., Norman, J.M., 1999. Evaluation of soil and vegetation heat flux predictions using a simple two-source model with radiometric temperatures for partial canopy cover. Agric. For. Meteorol. 17.

Legendre, P., Legendre, L., 2003. Numerical Ecology, 2/20. ed. Elsevier.

Lhomme, J.-P., 1997. A theoretical basis for the Priestley-Taylor coefficient. Bound.-Layer Meteorol. 82, 179–191. https://doi.org/10.1023/A:1000281114105

Li, F., Kustas, W.P., Prueger, J.H., Neale, C.M.U., Jackson, T.J., 2005. Utility of remote sensing–based two-source energy balance model under low- and high-vegetation cover conditions. J. Hydrometeorol. 6, 878–891. https://doi.org/10.1175/JHM464.1

Liebethal, C., Foken, T., 2007. Evaluation of six parameterization approaches for the ground heat flux. Theor. Appl. Climatol. 88, 43–56. https://doi.org/10.1007/s00704-005-0234-0

Linnet, K., 1993. Evaluation of regression procedures for method comparison studies. Clin Chem 424–432.

Manuilova, E., Schuetzenmeister, A., Model, F., 2014. mcr: Method Comparison Regression. Retrieved from:

https://www.rdocumentation.org/packages/mcr/versions/1.2.1 (accessed on 02.02.2020)

Marchese, M., Moheddine, A., Patrone, F., 2019. IoT and UAV integration in 5G hybrid terrestrial-satellite networks. Sensors 19, 3704. https://doi.org/10.3390/s19173704

Margono, B.A., Turubanova, S., Zhuravleva, I., Potapov, P., Tyukavina, A., Baccini, A., Goetz, S., Hansen, M.C., 2012. Mapping and monitoring deforestation and forest degradation in Sumatra (Indonesia) using Landsat time series data sets from 1990 to 2010. Environ. Res. Lett. 7, 034010. https://doi.org/10.1088/1748-9326/7/3/034010

Meijide, A., Röll, A., Fan, Y., Herbst, M., Niu, F., Tiedemann, F., June, T., Rauf, A., Hölscher, D., Knohl, A., 2017. Controls of water and energy fluxes in oil palm plantations: Environmental variables and oil palm age. Agric. For. Meteorol. 239, 71–85. https://doi.org/10.1016/j.agrformet.2017.02.034

Morillas, L., García, M., Nieto, H., Villagarcia, L., Sandholt, I., Gonzalez-Dugo, M.P., Zarco-Tejada, P.J., Domingo, F., 2013. Using radiometric surface temperature for surface energy flux estimation in Mediterranean drylands from a two-source perspective. Remote Sens. Environ. 136, 234–246. https://doi.org/10.1016/j.rse.2013.05.010

Norman, J.M., Kustas, W.P., Humes, K.S., 1995. Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. Agric. For. Meteorol. 77, 263–293. https://doi.org/10.1016/0168-1923(95)02265-Y

Norman, J.M., Kustas, W.P., Prueger, J.H., Diak, G.R., 2000. Surface flux estimation using radiometric temperature: A dual-temperature-difference method to minimize measurement errors. Water Resour. Res. 36, 2263–2274. https://doi.org/10.1029/2000WR900033

Ogée, J., Lamaud, E., Brunet, Y., Berbigier, P., Bonnefond, J.M., 2001. A long-term study of soil heat flux under a forest canopy. Agric. For. Meteorol. 106, 173–186. https://doi.org/10.1016/S0168-1923(00)00214-8

Oki, T., Kanae, S., 2006. Global hydrological cycles and world water resources. Am. Assoc. Adv. Sci. 313, 1068–1072. https://doi.org/10.1126/science.1128845

Ortega-Farías, S., Ortega-Salazar, S., Poblete, T., Kilic, A., Allen, R., Poblete-Echeverría, C., Ahumada-Orellana, L., Zuñiga, M., Sepúlveda, D., 2016. Estimation of energy balance components over a drip-irrigated olive orchard using thermal and multispectral cameras placed on a helicopter-based Unmanned Aerial Vehicle (UAV). Remote Sens. 8, 638. https://doi.org/10.3390/rs8080638

Pan, X., Liu, Y., Fan, X., Gan, G., 2017. Two energy balance closure approaches: applications and comparisons over an oasis-desert ecotone. J. Arid Land 9, 51–64. https://doi.org/10.1007/s40333-016-0063-2

Prudhomme, C., Giuntoli, I., Robinson, E.L., Clark, D.B., Arnell, N.W., Dankers, R., Fekete, B.M., Franssen, W., Gerten, D., Gosling, S.N., Hagemann, S., Hannah, D.M., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y., Wisser, D., 2014. Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. Proc. Natl. Acad. Sci. 111, 3262–3267. https://doi.org/10.1073/pnas.1222473110

Röll, A., Niu, F., Meijide, A., Ahongshangbam, J., Ehbrecht, M., Guillaume, T., Gunawan, D., Hardanto, A., Hendrayanto, Hertel, D., Kotowska, M.M., Kreft, H., Kuzyakov, Y., Leuschner, C., Nomura, M., Polle, A., Rembold, K., Sahner, J., Seidel, D., Zemp, D.C., Knohl, A., Hölscher, D., 2019. Transpiration on the rebound in lowland Sumatra. Agric. For. Meteorol. 274, 160–171. https://doi.org/10.1016/j.agrformet.2019.04.017

Sabajo, C.R., le Maire, G., June, T., Meijide, A., Roupsard, O., Knohl, A., 2017. Expansion of oil palm and other cash crops causes an increase of the land surface temperature in the Jambi province in Indonesia. Biogeosciences 14, 4619–4635. https://doi.org/10.5194/bg-14-4619-2017

Song, L., Liu, S., Kustas, W.P., Zhou, J., Xu, Z., Xia, T., Li, M., 2016. Application of remote sensing-based two-source energy balance model for mapping field surface fluxes with composite and component surface temperatures. Agric. For. Meteorol. 230–231, 8–19. https://doi.org/10.1016/j.agrformet.2016.01.005

Still, C., Powell, R., Aubrecht, D., Kim, Y., Helliker, B., Roberts, D., Richardson, A.D., Goulden, M., 2019. Thermal imaging in plant and ecosystem ecology: applications and challenges. Ecosphere 10, e02768. https://doi.org/10.1002/ecs2.2768

Suomalainen, J., Hakala, T., Alves de Oliveira, R., Markelin, L., Viljanen, N., Näsi, R., Honkavaara, E., 2018. A novel tilt correction technique for irradiance sensors and spectrometers on-board unmanned aerial vehicles. Remote Sens. 10, 2068. https://doi.org/10.3390/rs10122068

Tarigan, S., Wiegand, K., Sunarti, Slamet, B., 2018. Minimum forest cover required for sustainable water flow regulation of a watershed: a case study in Jambi Province, Indonesia. Hydrol. Earth Syst. Sci. 22, 581–594. https://doi.org/10.5194/hess-22-581-2018

Timmermans, W.J., Kustas, W.P., Andreu, A., 2015. Utility of an automated thermal-based approach for monitoring evapotranspiration. Acta Geophys. 63, 1571–1608. https://doi.org/10.1515/acgeo-2015-0016

Twine, T.E., Kustas, W.P., Norman, J.M., Cook, D.R., Houser, P.R., Meyers, T.P., Prueger, J.H., Starks, P.J., Wesely, M.L., 2000. Correcting eddy-covariance flux underestimates over a grassland. Agric. For. Meteorol. 103, 279–300. https://doi.org/10.1016/S0168-1923(00)00123-4

Xia, T., Kustas, W.P., Anderson, M.C., Alfieri, J.G., Gao, F., McKee, L., Prueger, J.H., Geli,

H.M.E., Neale, C.M.U., Sanchez, L., Alsina, M.M., Wang, Z., 2016. Mapping evapotranspiration with high-resolution aircraft imagery over vineyards using one- and two-source modeling schemes. Hydrol. Earth Syst. Sci. 20, 1523–1545. https://doi.org/10.5194/hess-20-1523-2016

Zhang, L., Hu, Z., Fan, J., Zhou, D., Tang, F., 2014. A meta-analysis of the canopy light extinction coefficient in terrestrial ecosystems. Front. Earth Sci. 8, 599–609. https://doi.org/10.1007/s11707-014-0446-7

Chapter 4

Predicting tree sap flux and stomatal conductance in a mixed agroforestry system from drone recorded surface temperatures – a machine learning approach

PREDICTING TREE SAP FLUX AND STOMATAL CONDUCTANCE IN A MIXED AGROFORESTRY SYSTEM FROM DRONE RECORDED SURFACE TEMPERATURES – A MACHINE LEARNING APPROACH

Florian Ellsäßer [1], Alexander Röll[1], Joyson Ahongshangbam[1], Pierre-André Waite[2], Hendrayanto[3], Bernhard Schuldt[4], Dirk Hölscher[1,5]

[1] *University of Goettingen, Tropical Silviculture and Forest Ecology, Büsgenweg 1, 37077 Göttingen, Germany*
[2] *University of Goettingen, Plant Ecology and Ecosystems Research, Untere Karspüle 2, 37073 Göttingen, Germany*
[3] *Bogor Agricultural University, Forest Management, Kampus IPB Darmaga, 16680 Bogor, Indonesia*
[4] *University of Würzburg, Julius-von-Sachs-Institute of Biological Sciences, Julius-von-Sachs-Platz 3, 97082 Würzburg, Germany*
[5] *University of Goettingen, Centre of Biodiversity and Sustainable Land Use, Büsgenweg 1, 37073 Göttingen, Germany*

In preparation.

ABSTRACT

Plant transpiration is a key element in the hydrological cycle. Widely used methods for its assessment comprise sap flux techniques for whole-plant transpiration and porometry for leaf stomatal conductance. Recently emerging approaches based on surface temperatures and a wide range of machine learning techniques offer new possibilities to quantify transpiration. The focus of this study was to predict sap flux and leaf stomatal conductance based on drone-recorded and meteorological data and compare these predictions with in-situ measured transpiration. To build the prediction models, we applied statistical approaches and machine learning algorithms. The field work was conducted in an oil palm agroforest in lowland Sumatra. Random forest predictions yielded the highest congruence with measured sap flux (r^2=0.87 for trees and r^2=0.58 for palms) and indicated differences among tree species. Confidence intervals for intercept and slope of a Passing-Bablok regression suggest interchangeability of methods for sap flux prediction using random forest. Predictions for stomatal conductance were less congruent, likely due to spatial and temporal offsets of the measurements. Overall, the applied drone and modelling scheme predicts whole-plant transpiration with high accuracy, especially using random forest algorithms.

4.1 Introduction

Transpiration is the largest water flux from terrestrial surfaces, accounting for 80 to 90% of terrestrial evapotranspiration (Jasechko et al., 2013). Transpiration is strongly affected by changes in land cover and land use (Good et al., 2015). In many tropical regions conversions of forests to agricultural land are ongoing (Hansen et al., 2013; Margono et al., 2014), resulting in large scale alterations of the water cycle and transpiration as a key flux. To measure, model and understand the effects of altered transpiration on the hydrological cycle, measurements at the plant and leaf scale with sap flux probes

and porometers are frequently applied (Ahongshangbam et al., 2019; Röll et al., 2019; Waite et al., 2019). While these hydrometric methods are commonly implemented at the leaf, plant or plot level, measuring transpiration at larger scales remains a challenging task (Ford et al., 2007; Good et al., 2015; Kume et al., 2010). Remote sensing techniques are considered to be more cost- and labour effective than their ground-based counterparts, particularly for applications in agricultural and forest landscapes (Virnodkar et al., 2020). Remote sensing data e.g. from satellites offers opportunities for spatial extrapolation of point measurements but is associated with large uncertainties, especially for diverse mixed stands and agricultural areas (Wei et al., 2017).

Recently emerging drone-based remote sensing systems could be an option to bridge the gap between leaf and plant scale measurement methods and catchment or landscape scale schemes (Suab and Avtar, 2020). Drones can operate close to the surface enabling a delimitation of single plant canopies and simultaneously cover considerable areas (Khokthong et al., 2019; Mohan et al., 2017). They can be equipped with a growing variety of light-weight sensors for diverse spectral ranges or structural measurements (Berni et al., 2008; Brenner et al., 2017; Kellner et al., 2019). For evapotranspiration, modelled results based on thermal sensor data often follow a linear relationship with ground based measurements from eddy covariance systems (Ellsäßer et al., 2020b). For complex non-linear structures and relationships e.g. between evapotranspiration and its controlling factors statistical regression models can be supplemented using machine learning (ML) algorithms (Dou and Yang, 2018; Fernandes et al., 2016). ML models are not explicitly programmed to represent biological processes but are rather trained using a training data set and can later be applied to previously unknown data (Virnodkar et al., 2020). ML algorithms such as support vector machines (SVM), random forests (RF) and artificial neural networks (ANN) have frequently been used e.g. to estimate evapotranspiration or to predict vegetation structural characteristics in a wide range of ecosystems (Antonopoulos et al., 2016; dos Reis et al., 2018; Dou and Yang, 2018; Feng et al., 2017; Pan et al., 2019; Wu et al., 2016). Despite many successful applications of ML algorithms to predict hydrological fluxes in the biosphere, the reliable quantification of the non-linear processes that govern water fluxes remains a challenging task (Dou and Yang, 2018). A successful quantification of water fluxes would highly depend on the choice of algorithm, a set of representative prediction variables and the availability of an appropriately sized data set (Granata, 2019).

The objectives of this study were:

(1) to compare linear statistical and ML approaches to calibrate models to predict sap flux and stomatal conductance measurements from drone remote sensing data and meteorological measurements
(2) to delineate the most important prediction variables for these models
(3) to compare the direct measurement methods (sap flux, porometry) with the modelling and drone-based methods for bias and interchangeability.

4.2 Methods

4.2.1 Study site

The study was conducted in the lowlands of Sumatra, in Jambi province, Indonesia. Average elevation of the area is 47 m asl., mean annual precipitation is 2,235 mm year^{-1} and average annual temperature 26.7°C (Drescher et al., 2016).

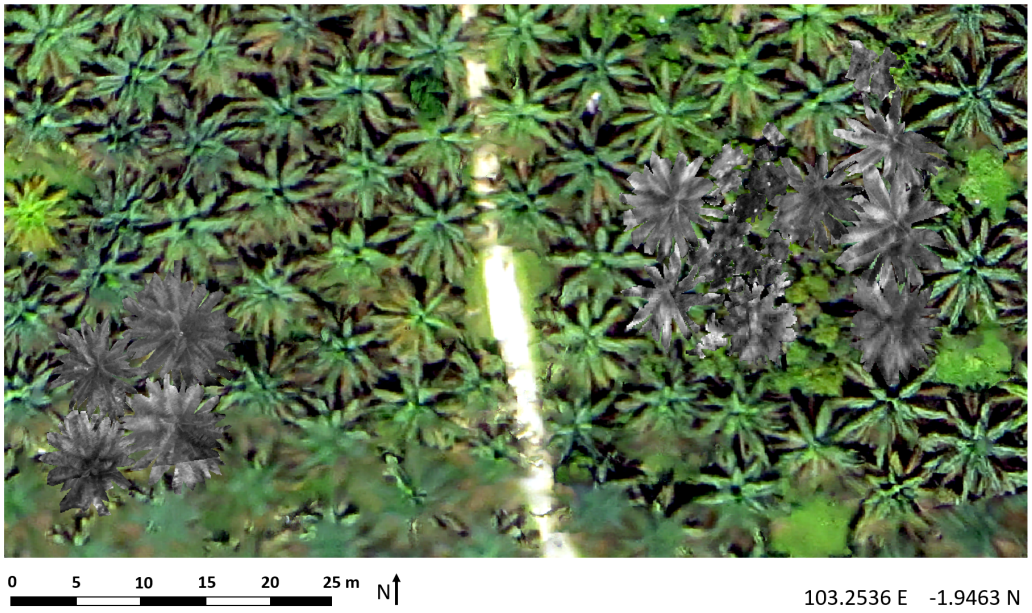


Figure 4.1: *The studied monoculture oil palm plantation on the left-hand side, and an oil palm agroforest site on the right side. Canopies of some of the measured trees and palms are shown with greyscale thermal images. Canopies were marked with 8-bit barcodes (small white dots in the image) to be recognizable from the air. The map shows an exemplary selection of all sampled canopies.*

The sites were situated in an oil palm plantation of the company PT Humusindo in a conventional monocultural oil palm plantation and oil palm agroforests, that were established in the context of a biodiversity enrichment experiment (EFForts-BEE, 103.2536 E, -1.9463 N) (Teuscher et al., 2016). At the experimental agroforest sites, three years prior to this study, 40% of the oil palms were cut and native tree species were planted. In the conventional oil palm monoculture palms were between 9 and 15 years old and the stem density was approximately 140 palms per hectare. At the time of our study in October 2016, average tree height was 4.7 m and average oil palm meristem height was 6.8 m.

4.2.2 Data acquisition

Two groups of data were acquired: the target variables, i.e. point measurements of sap flux and stomatal conductance with in situ sensor applications, and a set of prediction variables recorded with a drone and a meteorological station.

Sap flux measurements

Eight oil palms (*Elaeis guineensis*) and 16 trees of the four native species *Archidendron pauciflorum*, *Parkia speciosa*, *Peromena canescens* and *Shorea leprosula* were equipped with sap flux sensors. Four palms were located in a conventional monoculture area and four palms and all 16 trees in the agroforest sites. Oil palm sap flux was assessed with thermal dissipation probes (Granier, 1985) installed in the leaf petioles as described in Niu et al. (2015); sap flux density was calculated using calibrated, oil palm specific parameters (Niu et al., 2015). For the dicot trees, we used the heat ratio method (Burgess et al., 2001); sensors were installed radially into the xylem at breast height. Sap flux was then calculated using Sap Flow Tool version 1.4.1 (ICT International, Australia). Further details on the applied sap flux methods are provided in Ahongshangbam et al. (2019). Plant water use is expressed as mm h^{-1} to provide a fine temporal resolution format that allows the prediction of variations during the course of the day. Sap flux measurements were recorded every 10 min over the course of two weeks. We used barcode markers on sample trees and palms to facilitate their identification in the aerial images.

Stomatal conductance measurements

Stomatal conductance measurements were conducted using three porometers (AP4-UM-3, Delta-T Devices Ltd, Burwell, Cambridge, UK). The sunlit areas of the canopies were reached using scaffoldings. Three palms were assessed using nine fronts per palm as described in Waite et al. (2019). For trees we successfully measured three *Archidendron pauciflorum*, three *Peromena canescens* and four *Shorea leprosula*. The species *Parkia speciosa* is not included in the stomatal conductance data set because

leaflets were too small to fully cover the porometer measurement chamber. Seven to ten sunlit leaves each were measured on three branches per tree and mean values were calculated. Each sample palm or tree was measured between three and five times per day, over the course of two weeks. This variation was mainly caused by differences in relative humidity which affected the time period to complete the measurement cycle of the porometer devices. Measured leaves were marked with barcodes to facilitate identification in the drone recorded images.

Drone-based image acquisition

We used a multicopter drone (MK EASY Okto V3; HiSystems, Moormerland, Germany) equipped with a thermal and an RGB camera mounted in a stereo setup on a gimbal set to nadir perspective. The radiometric thermal camera was a FLIR Tau 2 640 (FLIR Systems Inc., Wilsonville, Oregon, USA) attached to a TeAx Thermo-capture module (TeAx Technology GmbH, Wilnsdorf, Germany). The sensor covers spectral bands ranging from 7.5 to 13.5 μm with a thermal accuracy of 0.04 K. The RGB camera was a Sony A5000 (Sony Corporation, Tokio, Japan) with an EPZ 16 – 55mm F3.5-5.6 OSS lens that was set to 16 mm focal length. During an 11–day campaign in October 2016 a total of 103 flights were conducted to record canopy surface temperatures simultaneously to the sap flux and stomatal conductance measurements. RGB images were merged using Photoscan 1.3.0 (Agisoft LLC, St. Petersburg, Russia) to create a geo-referenced orthomosaic map. All TIR images were then referenced to the RGB map, the canopies were delineated and single palm and tree canopies were extracted using QGIS3 version 3.6 Noosa (QGIS Development Team, 2020).

Meteorological measurements

Meteorological measurements were conducted at a station located in the studied oil palm plantation. The station was equipped with a global radiation sensor (CMP3 Pyranometer, Kipp & Zonen, Delft, The Netherlands), two thermohygrometers (type 1.1025.55.000, Thies Clima, Göttingen, Germany), a net radiometer (NR Lite2, Kipp & Zonen) and a 3-cup anemometer and a wind direction sensor (both Thies Clima). Data were stored on a logger every 10 min (see Meijide et al. (2018) for details).

Data pre-processing

Single canopies were marked with 8-bit barcodes that were visible in the RGB as well as the thermal spectrum and canopy surface temperatures were cut from the thermal images for each canopy and flight. We used the QGIS3-plugin QWaterModel (Ellsäßer et al., 2020a) to analyse the canopy surface temperatures, extract key metrics and apply the energy balance model DATTUTDUT (Timmermans et al., 2015) based on the canopy

temperatures. We applied both a fully modelled net radiation approach and a short-wave irradiance-based estimation approach (Ellsäßer et al., 2020b). We used the pandas library (McKinney, 2010) in Python 3.6.9 to merge the datasets according to the individual plant and recording time.
We set the maximum time delta to an hour and matched 1710 datasets for sap flux and 877 for stomatal conductance. The average time offsets for sap flux and stomatal conductance with the remote sensing data were 70 and 1481 seconds, respectively.
The final data set contained 94 variables including the two target variables (sap flux and stomatal conductance), meteorological variables (short-wave irradiance, barometric pressure, air temperature, wind speed and direction, relative humidity) and a multitude of variables from drone recorded data and modelled products such as land surface temperatures, evaporative fraction and different evapotranspiration estimates (incl. various dispersion metrics such as mean, median, standard deviation, coefficients of variance, kurtosis and the Fisher-Pearson Coefficient of Skewness (FPCS)). Furthermore, the data set contained local time encoded as cyclical features (with sinus and cosine as a variable), the canopy area, the number of pixels and atmospheric emissivity and transmissivity.

Prediction models

The data sets for sap flux prediction (n = 1710) and stomatal conductance prediction (n = 877) were each split into a training (or fitting) set and a test (or validation) set (70% and 30% of the data, respectively). We performed a multicollinearity test based on a variance inflation factor with Pearson's r (cut-off value > 0.9) and backward elimination of variables based on the Akaike information criterion (significance level 0.05) and removed variables accordingly reducing the input variables to 42. We decided to use a multiple linear regression (MLR) as our baseline method as it represents a well-known standard approach in statistics. We further applied support vector machine regressors (SVM), random forest regressors (RF) and artificial neural network regressors (ANN) to build prediction models.

Multiple linear regression

Multiple linear regression (MLR) is an approach to model the relationship between the explanatory input variables and a response variable by fitting a linear equation expressed as a regression plane (Nguyen and Zeigermann, 2018). In our study we used a least-squares model that minimizes the sum of squared vertical deviations from each point to this regression plane (Legendre and Legendre, 2003). MLR was computed using the *LinearRegression* regression class from the Scikit-learn package in python (Pedregosa et al., 2011). MLR were previously successfully used to predict daily transpiration, evapotranspiration and basal area of trees from spatial data (dos Reis et al.,

2018; Fernandes et al., 2016; Pan et al., 2019).

Support Vector Machine

SVMs model the relationship between explanatory variables and the response variable by fitting a regression plane with parallel margins to the data in order to include as many instances as possible between the two margins (Guo et al., 2012; Vapnik, 2000). This regression plane is also referred to as a hyperplane if several explanatory variables are used (Raschka, 2017). SVMs are mostly known as classifier algorithms, but can also be applied for regression problems (Shrestha and Shukla, 2015). SVM regressors were previously successfully applied to regression problems with spatial data (dos Reis et al., 2018; Garcia-Gutierrez et al., 2014; Wu et al., 2016). We used the SVR method of Scikit-learn package and a linear kernel to build SVMs (Pedregosa et al., 2011).

Random Forest

Decision trees are predictive models that use recursive partitioning for classification or regression tasks (Rokach and Maimon, 2015). Random forests combine the results of a set of individually trained decision trees (Breiman, 2001). This principle is called ensemble learning and has shown to improve the predictive performance (Opitz and Maclin, 1999; Raschka, 2017). Two widely used methods of ensemble learning are bootstrap aggregation referred to as Bagging (Breiman, 1996a) and Boosting (Freund and Schapire, 1996; Schapire, 1990).

In the Bagging method multiple replicates of the original learning data set are created by bootstrapping with replacement (Breiman, 1996b). The decision trees are then trained with these different variations of the original data set and the results of the individual decision trees are averaged (Breiman, 1996b). Bagging reduces the variance of simple models and helps to avoid overfitting of more complex models (Ghojogh and Crowley, 2019). We trained 2000 trees and used the *RandomForestRegressor* method of the Scikit-learn package (Pedregosa et al., 2011) to build our model. The idea behind Boosting is to combine a set of weak and moderately inaccurate decision trees and average their predictions to create very accurate predictions (Freund and Schapire, 1999). We trained 4000 trees and used the adaptive boosting algorithm (AdaBoost) introduced by (Freund and Schapire, 1997) from the Scikit-learn package (Pedregosa et al., 2011). Different designs of RF algorithms have previously been used to successfully predict e.g. crop water stress, evapotranspiration, above ground-biomass and basal area (dos Reis et al., 2018; Feng et al., 2017; Pan et al., 2019; Virnodkar et al., 2020; Wu et al., 2016).

Artificial Neural Network

ANNs are inspired by biological brains and consist of multiple neurons that are

organized in a set of layers (Nguyen and Zeigermann, 2018). ANNs are ideal to identify complex non-linear relationships between in- and output data sets and particularly useful in regression problems with processes that are difficult to capture entirely (Antonopoulos et al., 2016; Fernandes et al., 2016). We used the *Sequential*, *Dense* and *KerasRegressor* methods of the keras library (Chollet, 2015) to build an ANN with the typical multiple perceptron type (MLP) architecture which was recently used to predict tree metrics from spatial data (dos Reis et al., 2018; Tavares Júnior et al., 2019). Similar ANN designs have been used to estimate evapotranspiration and transpiration in a wide range of ecosystems (Antonopoulos et al., 2016; Feng et al., 2017; Fernandes et al., 2016). We used rectified linear units (ReLU) for the activation function in the input and the three hidden layers and a linear activation function for the output layer.

Variable importance

To estimate the importance of each predictor variable for the regression model, we decided against a removal-based approach as described in e.g. dos Reis et al. (2018) and opted for a randomization-based permutation test (also called Mean Decrease Accuracy) using the *PermutationImportance* method from the eli5 package (Breiman, 2001; Strobl et al., 2008). Hereby, the values of a single input variable of the prediction dataset are randomized (and not left out) and its effect on the prediction accuracy of the regression model is measured (Strobl et al., 2008).

Statistical analyses of predicted and measured values

To evaluate the prediction models, we compared measured and predicted values from the test data set and calculated model accuracy in %, mean absolute percentage error (MAPE), root mean squared error (RMSE) and R^2. Variable importance of each predictor variable was assessed using a permutation test (Strobl et al., 2008). Single variable importance was averaged over all data sets for each sap flux and stomatal conductance and might vary for single species. To compare the prediction methods we used a non-parametric Passing-Bablok regression (Bilic-Zulle, 2011; Passing and Bablok, 1984, 1983). The python *MethComp package* (van Doorn, 2020) was used for the computation of the Passing-Bablok regressions. The Passing-Bablok regression outputs a regression line of which the confidence intervals of slope and intercept are especially interesting for a model comparison analysis. If the confidence intervals of the slope and the intercept include 1 and 0 respectively, there is no statistically significant bias between the methods (Legendre and Legendre, 2003; Passing and Bablok, 1983). Linearity of the data is a crucial assumption for Passing-Bablok regressions (Passing and Bablok, 1983); and was checked visually. All statistical analyses were computed with Python

3.6.9 using pandas (McKinney, 2010), NumPy (Oliphant, 2006; van der Walt et al., 2011), SciPy (SciPy 1.0 Contributors et al., 2020), statsmodels, scikit learn (Pedregosa et al., 2011), keras and eli5 packages and libraries. Graphs and figures were created with Matplotlib (Hunter, 2007) and seaborn (Waskom et al., 2020) libraries.

4.3 Results

4.3.1 Prediction accuracy

Highest model accuracy was achieved by Random Forest models (both bagging and boosting) predicting sap flux for individual tree species, especially *Archidendron pauciflorum* where a model accuracy of well over 90% was reached (Figure 4.2). Across all tree species, accuracy was 60% and 52% for oil palm (RF bagging).

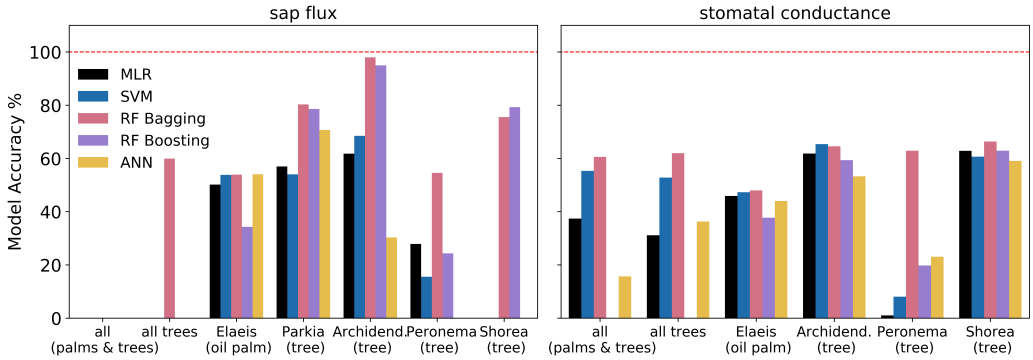


Figure 4.2: *Model accuracy % comparing measured and predicted values of sap flux and stomatal conductance.*

Prediction of sap flux across all canopies, including dicot trees and palms, was unsuccessful with all applied algorithms. For predictions of stomatal conductance, RF bagging was again the method with the highest model accuracy, consistently yielding prediction accuracies of 60% for the tree species and 48% for oil palm (Figure 4.2). For stomatal conductance, predictions across all canopies were successful, with 60% accuracy for RF bagging. A comparison of error metrics (MAE and RMSE) showed that errors for the random forest algorithms and particularly RF bagging were comparatively low for both sap flux and stomatal conductance prediction (Figures 4.3 and 4.4). The simpler algorithms such as MLR and SVM but also the more complex ANN frequently produced higher errors. Generally, the errors for stomatal conductance predictions are more evenly distributed across algorithms and target groups than for sap flux prediction. For instance, MAE and RMSE for Shorea leprosula were relatively small and homogeneous across algorithms for stom-

atal conductance, but ranged from under 30% (RF bagging) to over 100% (ANN) for sap flux.

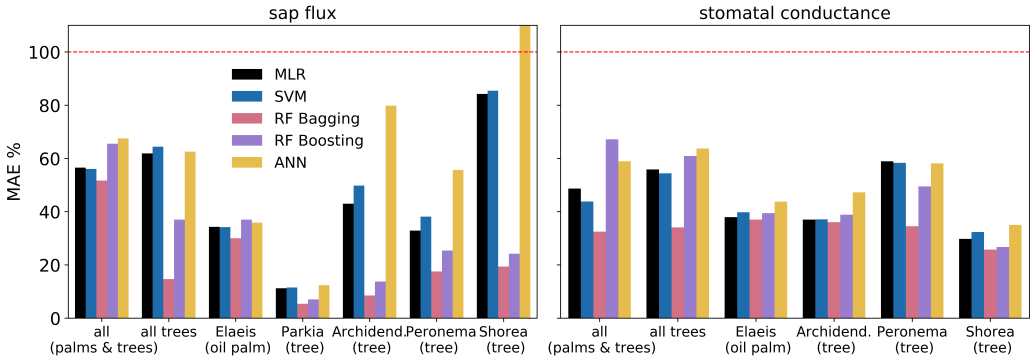


Figure 4.3: *Mean Absolute Error (MAE) of predicted and measured values for sap flux and stomatal conductance.*

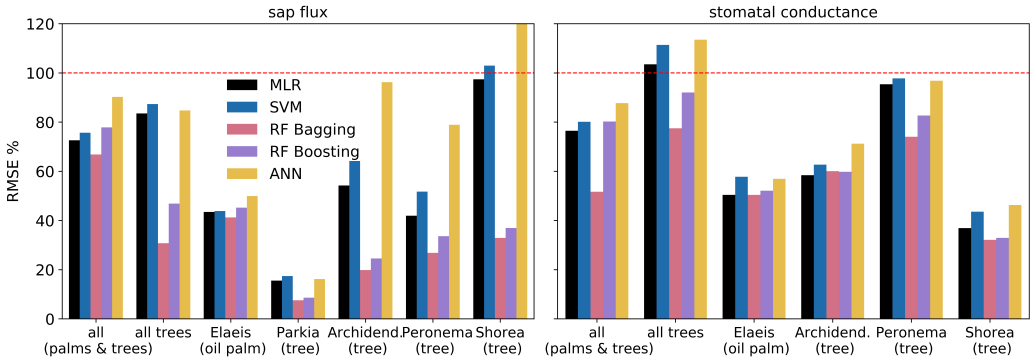


Figure 4.4: *Root Mean Square Error (RMSE) of predicted and measured values for sap flux and stomatal conductance.*

Both RF algorithms resulted in predictions highly congruent with the sap flux measurements in terms of R^2 (Figure 4.5); across all tree species as well as for each individual tree species, R^2s were close to or higher than 0.8 (RF bagging), while they were close to 0.6 for oil palm. The predictions for stomatal conductance are generally less congruent with the measurements than for sap flux. However, the RF bagging algorithm does achieve R^2s around 0.5 or higher across all canopies and across all tree species, as well as individually for two of the four species.

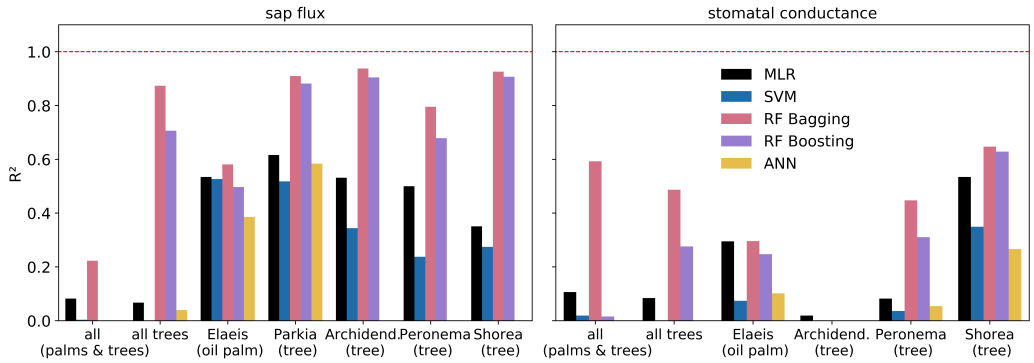


Figure 4.5: *Coefficient of determination (R^2) for predicted and measured values for sap flux and stomatal conductance.*

4.3.2 Variable Importance

A total of 94 input variables (or features) were used to train the algorithms. After applying a multicollinearity test and a backward elimination the remaining 42 variables were used to train the prediction algorithms. To improve the distribution of measurement efforts and the relevance of input features we performed a permutation importance analysis. Therein, the numbers of most important variables that explain 95% of the model outcome are highly variable (Figure 4.6).

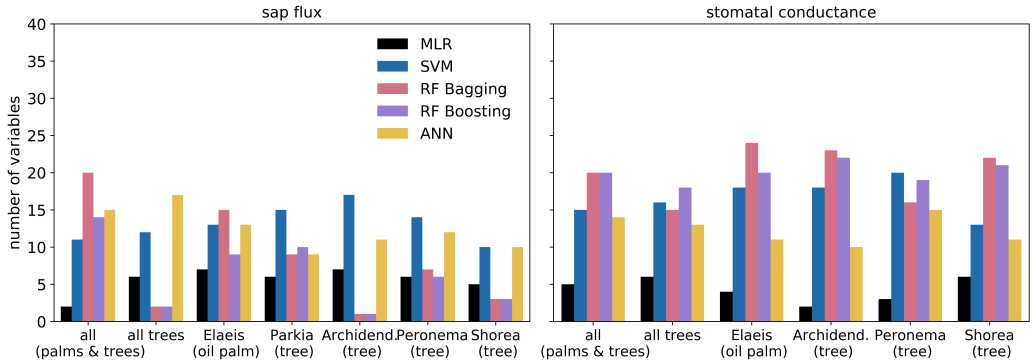


Figure 4.6: *Number of most important input variables that explain 95% of the corresponding models.*

For the MLR a very low number of only up to 7 input variables are required to explain most of the model result for both sap flux and stomatal conductance. For the SVM algorithm between 10 and 20 variables explain 95% of the model outcome; a smaller number of variables explain the sap flux results, while the number of variables was higher for stomatal conductance predictions. The RF algorithms showed the highest variations ranging

from only one variable up to 20 for sap flux and generally using 15 to 20 main variables for stomatal conductance to explain most of the model's outcome. The ANN algorithm uses an intermediate number of around 9 to 17 important variables for both sap flux and stomatal conductance prediction.
Variable importance (expressed in %) for predicting sap flux averaged over all sap flux and stomatal conductance data sets shows big differences for the MLR algorithm, where variables for sap flux are very homogeneous and the prediction of stomatal conductance is dominated by the latent heat flux as derived from the DATTUTDUT model (Figure 4.7 and 4.8). A similar but less pronounced pattern can be observed for the SVM regressor where the main prediction variables for sap flux are very homogeneous and for the prediction of stomatal conductance barometric pressure is the prevalent variable.

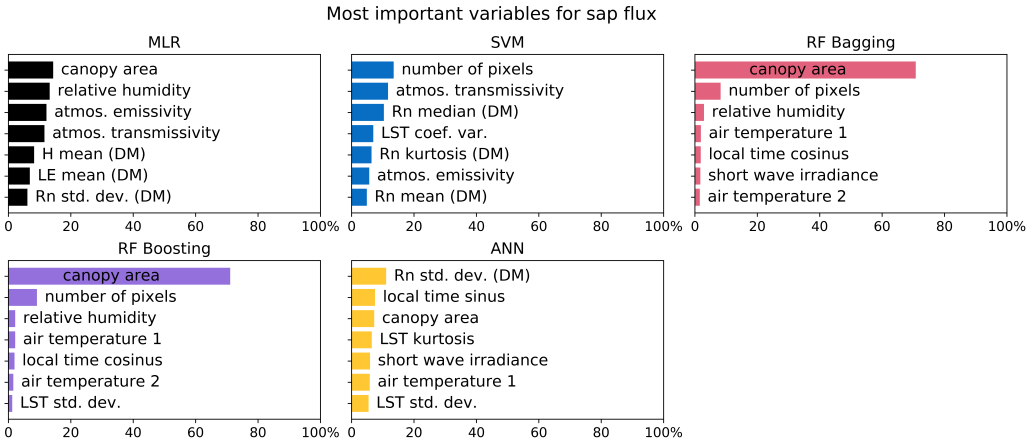


Figure 4.7: *The seven most important input variables averaged over all data sets (different canopies and combinations) for the corresponding models when predicting sap flux*

Both RF algorithms mainly rely on the canopy area as a main prediction variable for both sap flux and stomatal conductance. Another observation is the second variable that is of a much less pronounced importance for the model results: both algorithms use the number of pixels for sap flux predictions and the barometric pressure for stomatal conductance predictions (Figure 4.7 and 4.8). The ANN algorithm shows a very homogeneous distribution for both sap flux and stomatal conductance and no prevalent variable that is especially important for the results (Figure 4.7 and 4.8).

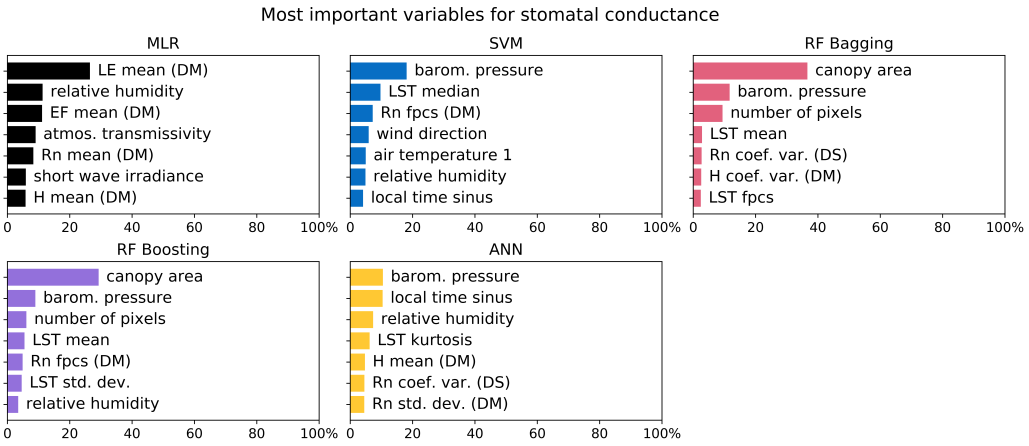


Figure 4.8: *The seven most important input variables averaged over all data sets (different canopies and combinations) for the corresponding models when predicting stomatal conductance.*

4.3.3 Method comparison

Predicted values for sap flux from the RF algorithms almost always showed a linear relationship with their measured counterparts, whereas this can only be observed for the MLR and SVM algorithms for oil palm datasets and was never detected for the ANN algorithm (Figure 4.9). There was no linear relationship of measured and predicted values for stomatal conductance for all prediction algorithms (Figure 4.10). Sap flux predictions for all trees and oil palm (but not all canopies) as well as for *Archidendron pauciflorum* showed no significant continuous or systematic bias from the measurements when the RF Bagging algorithm was applied (Figures 4.9 and 4.11), which indicates interchangeability of the methods.

Except for some outliers, the measured and RF bagging predicted values for sap flux in trees closely follow the 1:1 line, while showing more variance for oil palm (Figure 4.11). The same was observed for oil palm canopies where the MLR and RF Boosting algorithms predicted sap flux without significant bias or errors (Figure 4.9). For species specific predictions, linearity was found for both RF algorithms but except for *Archidendron pauciflorum* a significant systematic and continuous bias was detected. None of the algorithms showed potential in predicting sap flux of *Parkia speciosa*.

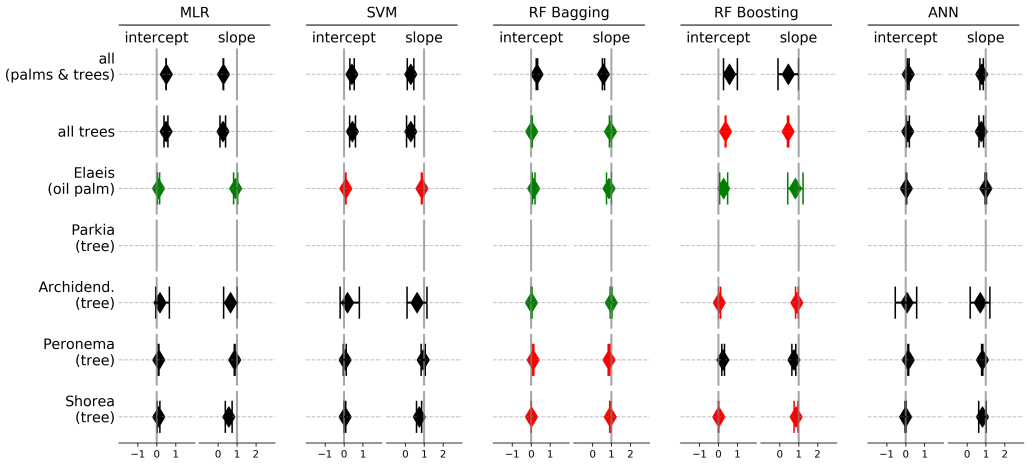


Figure 4.9: *The 99% confidence intervals of slope and intercept from model II Passing-Bablok regression for sap flux. The results in green show no significant bias between measured and predicted values. The results indicated in red fulfilled the assumptions of linearity but a significant difference between measured and predicted values was found. The results displayed in black did not meet the assumption of linearity and are therefore not to be considered.*

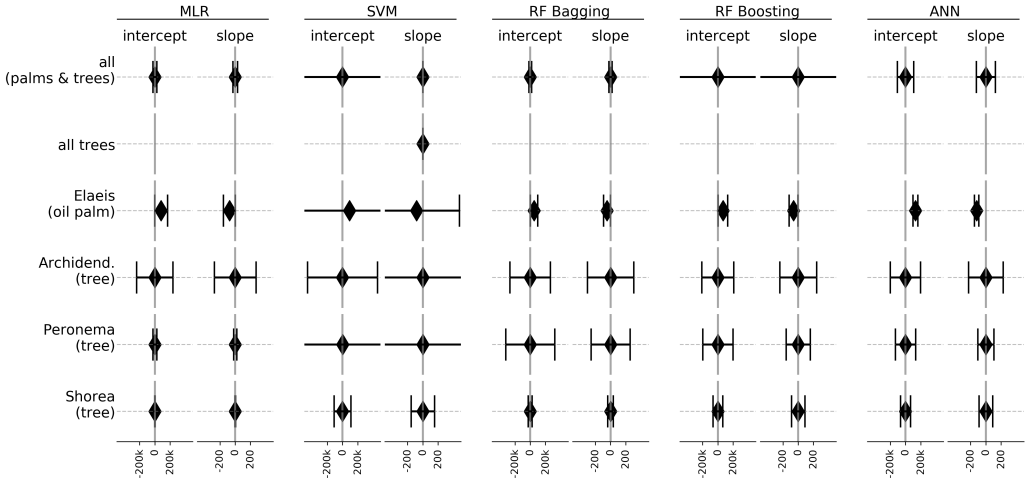


Figure 4.10: *The 99% confidence intervals of slope and intercept from model II Passing-Bablok regression for stomatal conductance. None of the results for stomatal conductance fulfilled the assumption of linearity necessary to conduct a valid Passing-Bablok Regression.*

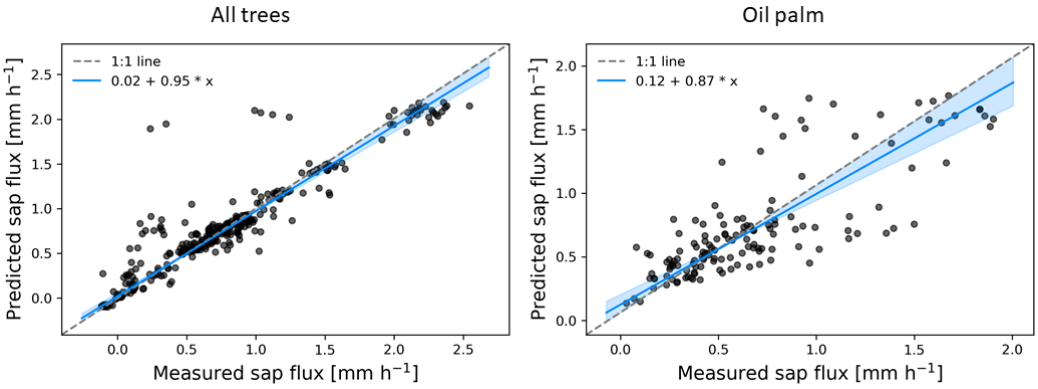


Figure 4.11: *Measured and predicted sap flux for all tree and palm canopies from RF Bagging algorithm.*

4.4 Discussion

Our study demonstrates the applicability of thermal remote sensing data from drones and meteorological data for predicting sap flux and stomatal conductance, two key processes related to the water balance of individual plants and ecosystems. Model accuracy was generally higher and errors were lower for sap flux than for the stomatal conductance predictions. Therein, different prediction algorithms showed substantially different behaviour in terms of input requirements and variable importance. The prediction results especially from the RF Bagging method were found to be interchangeable with ground-based measurements for some canopy types and species.

4.4.1 Prediction accuracy

Using the initially reduced input data set of 42 variables, the five applied prediction algorithms estimated the target variables sap flux and stomatal conductance with varying results. In our study, MLR, representing the simplest algorithm, produced results of intermediate quality, often with lower errors than the more complex SVM algorithm. This stands in contrast to previous studies where the SVM algorithm produced more congruent estimates, e.g. sap flux from a range of directly measured variables or biomass from remote sensing data (Fernandes et al., 2016; Wu et al., 2016). Since the creation of a representative model with MLR requires a prevalent linear relationship of at least some input variables with the target variables, we assume that given a supplementary set of ground-based measurements MLR can still be a simple alternative to more complex machine learning algorithms. The relatively poor results of MLR predictions might be caused due to the fact that especially the remotely sensed data often does not directly correlate with either sap flux or stomatal conductance measurements.

Compared to previous studies that

e.g. modelled evapotranspiration from hydro-climatic variables (Shrestha and Shukla, 2015), the SVM algorithm did not produce satisfying results. On average, model accuracy was higher for the SVM when predicting stomatal conductance than when predicting sap flux, coefficients of determination indicate that neither model predicts the target variable accurately. Since we used a linear kernel, the performance of the SVM algorithm highly depends on the linear relationship of at least some input variables with the according target variable.

Overall, both RF algorithms (Bagging and Boosting) resulted in similar, high quality predictions, with the Bagging algorithm commonly performing slightly better. Previous studies predicting water stress from remote sensing data showed similar results, i.e. the bootstrapping based Bagging option slightly outperforming the boosting algorithm, but small overall differences between the methods were observed (Ismail and Mutanga, 2010). In our study both RF algorithms show a comparatively high model accuracy and low errors with a high congruence of measured and predicted values, particularly for sap flux. Similar outcomes were found in studies where above-ground biomass was estimated from remote sensing data and where RF algorithms had the lowest errors and highest congruence and generally outperformed all other algorithms (dos Reis et al., 2018; Wu et al., 2016).

The ANN algorithm did not produce sufficiently accurate predictions in our study, except for the prediction of sap flux for one tree species, Parkia speciosa. A previous study applying similar feed-forward networks but using ground-based input data approaches showed satisfactory results for sap flux prediction (Liu et al., 2009). While we expected better results from the ANN algorithm due to its complexity, the parameterization of the method is known to be difficult and highly sensitive to variations in input parameters (dos Reis et al., 2018; Rodriguez-Galiano et al., 2015). Further research in parameter setting options might therefore produce more accurate results.

4.4.2 Variable importance evaluation

The output quality of prediction algorithms highly depends on how well the input variables represent the ecosystem and the target variables (Pan et al., 2019). For sap flux and stomatal conductance, one would generally expect a strong influence of environmental and physiological variables. The overall number of variables that explained 95% of our MLR model was low, never surpassing 7 variables (Figure 4.6). This is a strong indicator for either no or non-linear relationships between most input variables and the output variables. The seven most important variables averaged over all input data set configurations (mixed and single species) of the MLR model in our study show that canopy area as well as

relative humidity have the greatest impact when predicting sap flux (Figure 4.6). These results only partly resemble variables from other studies where air temperature and VPD related variables such relative humidity played a key role in MLR derived models to predict sap flux (Liu et al., 2009). The most important variable for stomatal conductance predictions is the mean latent heat flux derived from the DATTUTDUT model (Figure 4.7). Surprisingly wind speed was not amongst the most important input variables for the MLR algorithm whereas an influence of wind speed on both target variables was shown in several previous studies (e.g. Chu et al., 2009). Asides from meteorological variables, soil moisture is considered to be a classic driver of sap flux and thus transpiration (e.g. Giambelluca et al., 2016; Kobayashi et al., 2014). While our data set unfortunately did not comprise soil moisture data, in a previous study soil moisture played a key role for predicting transpiration with a MLR in a semiarid region (Fernandes et al., 2016). However, in our equatorial study region in lowland Sumatra previous assessments showed no significant influence of the typically rather small soil moisture fluctuations on (evapo)transpiration of oil palms and trees (Röll et al., 2019; Stiegler et al., 2019), with exception of a strong El Niño event (Stiegler et al., 2019).

We could not identify dominating input variables that would explain the model built by the SVM algorithm for sap flux. The prediction from the SVM algorithm for stomatal conductance was slightly dominated from the barometric pressure variable (Figure 4.8). None of the variables explained more than 20% of the final model. In contrast, a previous study successfully applied a SVM to predict potential evaporation using solar radiation, relative humidity, air temperature and wind speed as input variables (Kişi and Çimen, 2009). As both sap flux and stomatal conductance are related to evapotranspiration we expected the SVM derived model to be based on these variables. However, compared to the MLR algorithm, none of these variables was among the most important variable for sap flux prediction and only relative humidity played a minor role for stomatal conductance prediction.

The number of variables that explains 95% of the model results was very low for both RF algorithms when predicting sap flux. In contrast, the RF algorithms used among the highest numbers of variables when predicting stomatal conductance. Among them, canopy area was the (clearly) dominating variable for the prediction of both sap flux and stomatal conductance. With simple linear regression analysis, no relationship between canopy size and sap flux could be found, mainly because the canopy size is a constant variable while sap flux quantity varies during the day. We assume a non-linear relationship between canopy size and sap flux which requires auxiliary variables to be predicted. Remotely sensed canopy size is an important factor for plant transpiration and was already found to be a suitable predictor in previous studies

(Ahongshangbam et al., 2019, 2020). A general advantage of RF algorithms is their ability to produce accurate predictions even from small samples and with a large set of independent input variables (dos Reis et al., 2018); in our study, they performed better than all other applied algorithms.
The ANN algorithm mainly uses an intermediate number of input variables compared to the small set of MLR based models and the large sets of the SVM models. Compared to the other algorithms, the ANN does not show any clearly dominant input variable. An interesting observation is that the ANN algorithm uses cyclic local time variables among its most important input variables. This seems logical approach as sap flux is known to follow a daily course with near-noon maxima and night time near-zero minima in the study region (e.g. Niu et al., 2017; Röll et al., 2015). One previous study found that the best results for stand transpiration prediction were achieved by using climate data, soil water content and canopy properties as key input variables (Fernandes et al., 2016), which are partly also represented in our ANN derived prediction models.

4.4.3 Method comparison

Compared to the ground measurements, most algorithms were able to predict sap flux in oil palm without major errors, three algorithms even showed no bias in its prediction. Simple MLR was sufficient to predict sap flux from oil palm canopies without bias. Compared to the other species oil palm had by far the biggest sample size. This illustrates how important adequate sample sizes are for model training. While the SVM algorithm produced rather poor predictions in our study, it was found to predict better results for e.g. potential ET when compared with conventional ANN approaches (Kişi and Çimen, 2009). Using RF algorithms, sap flux can be predicted from remote sensing data in close congruence to measurements without producing many outliers (Figure 4.11). While analysing tree and palm canopies separately, sap flux can well be predicted by the RF bagging algorithm, while for the whole data set including both, trees and palms, sap flux was not adequately predicted. The reasons are likely underlying differences in physiology between dicot trees and monocot palms; this e.g. includes size and distribution of water-conductive vessels in the stem and crown and leaf architecture. As such, previous ecohydrological assessments in the study region pointed to vast differences in water use between oil palms and trees including rubber trees of similar age, with substantially higher per-tree and stand transpiration rates of oil palms (Röll et al., 2015; Niu et al., 2017; Röll et al., 2019). Drone-derived crown metrics of oil palms and adjacent trees further suggested that oil palms transpire two-times more water per unit of crown volume as agroforest trees (Ahongshangbam et al., 2019) and about five-times more per unit of crown surface area than rain-

forest trees (Ahongshangbam et al., 2020), which may well be the reason why the joined analysis of trees and palms was unsuccessful. The good prediction performance of RF algorithms compared to e.g. ANN algorithms was previously also described for predicting potential evapotranspiration (Feng et al., 2017). RF was further found to be the best of several algorithms for above ground biomass estimation from remote sensing data (Wu et al., 2016). The rather complex ANN algorithm showed no convincing results in our study. Potential reasons include that variables might not have been captured due to a lack of a sufficing number of free weights (Fernandes et al., 2016), or that a surplus of free weights might have caused an over-fitting of the model which lacked the generalization to predict reasonable results from our test data set (Kumar et al., 2011). More than the other algorithms optimal layout of the ANN highly depends on the input variables, the problem that one tries to solve and is often subject to a trial and error process (Maier and Dandy, 2000).

None of the applied models could predict sap flux in *Parkia speciosa*, despite comparatively low errors and a promising model accuracy as well as high congruence with measurements. However, confidence intervals tended towards infinity and intercept and slope were far off their regular scope. Reviewing the thermal canopy pictures of *Parkia speciosa* canopies it is quite obvious that the canopies are not dense and partly 'invisible' in the thermal images and most pixels are either mixed or just represent the soil below.

Since a core advantage of ML algorithms is that they can continuously be improved when new data is available (dos Reis et al., 2018), better predictions for single tree species are possible opening up new options to predict sap flux for a vast variety of species. In our study, major challenges remain with the prediction of stomatal conductance. The data set for stomatal conductance differed strongly from the data set for sap flux. Average time offsets of directly and remotely measuring stomatal conductance were much bigger than for sap flux and measured and modelled input variables were not ideally suited for stomatal conductance prediction. The use of an additional set of multispectral images and resulting indices such as the NDVI (normalized difference vegetation index) or the EVI (enhanced vegetation index) could potentially enhance the prediction of stomatal conductance (Panda et al., 2014). Further, the sample size for sap flux was much larger than for stomatal conductance due to a limited amount of porometry devices and the non-automated nature of the measurements. Furthermore, the stomatal conductance measurements represent only a small portion (centimetre scale) of a given canopy. The identification and subsequent upscaling from this small leaf area to a whole canopy thus likely introduces substantial errors into the predictions.

A general drawback of machine learning models is that they usually lack causal relations that would enable a bi-

ological interpretation (dos Reis et al., 2018; Özçelik et al., 2013). Our study is based on approaches that are often described as the traditional statistical (MLR) and traditional machine learning (SVM, RF, ANN) methods (Dou and Yang, 2018), while more complex approaches such as the stochastic gradient boosting, which combines the advantages of boosting and bagging (Wu et al., 2016), were not within the scope of this study. In summary, RF approaches worked best for predicting sap flux, while a model that can predict stomatal conductance without bias was not found.

4.5 Conclusions

Drone remote sensing and in situ meteorological observations in conjunction with machine learning algorithms can be considered reliable methods for the prediction of sap flux. For tree and palm canopies random forest regressors predicted interchangeable results without significant bias compared to direct sap flux measurements. The prediction of stomatal conductance from remotely sensed data was less successful and requires further research. Our study complements the asset of available sap flux approaches by a reliable method for drone-based sap flux prediction.

REFERENCES

Ahongshangbam, J., Khokthong, W., Ellsäßer, F., Hendrayanto, H., Hölscher, D., Röll, A., 2019. Drone-based photogrammetry-derived crown metrics for predicting tree and oil palm water use. Ecohydrology 12. https://doi.org/10.1002/eco.2115

Ahongshangbam, J., Röll, A., Ellsäßer, F., Hendrayanto, Hölscher, D., 2020. Airborne tree crown detection for predicting spatial heterogeneity of canopy transpiration in a tropical rainforest. Remote Sens. 12, 651. https://doi.org/10.3390/rs12040651

Antonopoulos, V.Z., Gianniou, S.K., Antonopoulos, A.V., 2016. Artificial neural networks and empirical equations to estimate daily evaporation: application to Lake Vegoritis, Greece. Hydrol. Sci. J. 61, 2590–2599. https://doi.org/10.1080/02626667.2016.1142667

Berni, J.A.J., Zarco-Tejada, P.J., Suárez, L., González-Dugo, V., Fereres, E., 2008. Remote sensing of vegetation from UAV platforms using lightweight multispectral and thermal imaging sensors. Inter-Comm. WG IV 38, 6.

Bilic-Zulle, L., 2011. Comparison of methods: Passing and Bablok regression. Biochem. Medica 49–52. https://doi.org/10.11613/BM.2011.010

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., 1996a. Stacked regressions. Mach. Learn. 24, 49–64. https://doi.org/10.1007/BF00117832

Breiman, L., 1996b. Bagging predictors. Mach. Learn. 24, 123–140. https://doi.org/10.1007/BF00058655

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. Int. J. Remote Sens. 38, 3003–3026. https://doi.org/10.1080/01431161.2017.1280202

Burgess, S.S.O., Adams, M.A., Turner, N.C., Beverly, C.R., Ong, C.K., Khan, A.A.H., Bleby, T.M., 2001. An improved heat pulse method to measure low and reverse rates of sap flow in woody plants. Tree Physiol. 21, 589–598. https://doi.org/10.1093/treephys/21.9.589

Chollet, F. and others, 2015. Keras. Retrieved from: https://github.com/fchollet/keras (accessed: 12.05.2020)

Chu, C.R., Hsieh, C.-I., Wu, S.-Y., Phillips, N.G., 2009. Transient response of sap flow to wind speed. J. Exp. Bot. 60, 249–255. https://doi.org/10.1093/jxb/ern282

dos Reis, A.A., Carvalho, M.C., de Mello, J.M., Gomide, L.R., Ferraz Filho, A.C., Acerbi Junior, F.W., 2018. Spatial prediction of basal area and volume in Eucalyptus stands using Landsat TM data: an assessment of prediction methods. N. Z. J. For. Sci. 48, 1. https://doi.org/10.1186/s40490-017-0108-0

Dou, X., Yang, Y., 2018. Evapotranspiration estimation using four different machine learning approaches in different terrestrial ecosystems. Comput. Electron. Agric. 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., Krashevska, V., Krishna, V., Leuschner, C., Lorenz, W., Meijide, A., Melati, D., Nomura, M., Pérez-Cruzado, C., Qaim, M., Siregar, I.Z., Steinebach, S., Tjoa, A., Tscharntke, T., Wick, B., Wiegand, K., Kreft, H., Scheu, S., 2016. Ecological and socio-economic functions across tropical land use systems after rainforest conversion. Philos. Trans. R. Soc. B Biol. Sci. 371, 20150275. https://doi.org/10.1098/rstb.2015.0275

Ellsäßer, F., Röll, A., Stiegler, C., Hendrayanto, Hölscher, D., 2020a. Introducing QWaterModel, a QGIS plugin for predicting evapotranspiration from land surface temperatures. Environ. Model. Softw. 104739. https://doi.org/10.1016/j.envsoft.2020.104739

Ellsäßer, F., Stiegler, C., Röll, A., June, T., Hendrayanto, Knohl, A., Hölscher, D., 2020b. Predicting evapotranspiration from drone-based thermography - a method comparison in a tropical oil palm plantation (preprint). Biogeosciences. https://doi.org/10.5194/bg-2020-159

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. Agric. Water Manag. 193, 163–173. https://doi.org/10.1016/j.agwat.2017.08.003

Fernandes, T., Campo, A., García-Bartual, R., González-Sanchis, M., 2016. Coupling daily transpiration modelling with forest management in a semiarid pine plantation. IForest - Biogeosciences For. 9, 38–48. https://doi.org/10.3832/ifor1290-008

Ford, C.R., Hubbard, R.M., Kloeppel, B.D., Vose, J.M., 2007. A comparison of sap flux-based evapotranspiration estimates with catchment-scale water balance. Agric. For. Meteorol. 145, 176–185. https://doi.org/10.1016/j.agrformet.2007.04.010

Freund, Y., Schapire, R.E., 1999. A short introduction to boosting. Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence, 1401–1406, Morgan Kaufmann

Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. 55, 119–139. https://doi.org/10.1006/jcss.1997.1504

Freund, Y., Schapire, R.E., 1996. Experiments with a New Boosting Algorithm. ICML'96: Proceedings of the Thirteenth International Conference on International Conference on Machine Learning, Pages 148–156

Garcia-Gutierrez, J., Martínez-Álvarez, F., Troncoso, A., Riquelme, J.C., 2014. A comparative study of machine learning regression methods on LiDAR data: A case study, in: Herrero, Á., Baruque, B., Klett, F., Abraham, A., Snášel, V., de Carvalho, A.C.P.L.F., Bringas, P.G., Zelinka, I., Quintián, H., Corchado, E. (Eds.), International Joint Conference SOCO'13-CISIS'13-ICEUTE'13. Springer International Publishing, Cham, pp. 249–258. https://doi.org/10.1007/978-3-319-01854-6_26

Ghojogh, B., Crowley, M., 2019. The theory behind overfitting, cross validation, regularization, bagging, and boosting: tutorial. ArXiv190512787 Cs Stat.

Giambelluca, T.W., Mudd, R.G., Liu, W., Ziegler, A.D., Kobayashi, N., Kumagai, T., Miyazawa, Y., Lim, T.K., Huang, M., Fox, J., Yin, S., Mak, S.V., Kasemsap, P., 2016. Evapotranspiration of rubber (Hevea brasiliensis) cultivated at two plantation sites in Southeast Asia: Rubber evapotranspiration in SE Asia. Water Resour. Res. 52, 660–679. https://doi.org/10.1002/2015WR017755

Good, S.P., Noone, D., Bowen, G., 2015. Hydrologic connectivity constrains partitioning of global terrestrial water fluxes. Am. Assoc. Adv. Sci. 349, 175–177. https://doi.org/10.1126/science.aaa5931

Granata, F., 2019. Evapotranspiration evaluation models based on machine learning algorithms—A comparative study. Agric. Water Manag. 217, 303–315. https://doi.org/10.1016/j.agwat.2019.03.015

Granier, A., 1985. Une nouvelle méthode pour la mesure du flux de sève brute dans le tronc des arbres. Ann. Sci. For. 42, 193–200. https://doi.org/10.1051/forest:19850204

Guo, Y., Li, Z., Zhang, X., Chen, E., Bai, L., Tian, X., He, Q., Feng, Q., Li, W., 2012. Optimal Support Vector Machines for forest aboveground biomass estimation from multisource remote sensing data, in: 2012 IEEE International Geoscience and Remote Sensing Symposium. Presented at the IGARSS 2012 - 2012 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Munich, Germany, pp. 6388–6391. https://doi.org/10.1109/IGARSS.2012.6352721

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. Science 342, 850–853. https://doi.org/10.1126/science.1244693

Hunter, J.D., 2007. Matplotlib: A 2D Graphics Environment. Comput. Sci. Eng. 9, 90–95. https://doi.org/10.1109/MCSE.2007.55

Ismail, R., Mutanga, O., 2010. A comparison of regression tree ensembles: Predicting Sirex noctilio induced water stress in Pinus patula forests of KwaZulu-Natal, South Africa. Int. J. Appl. Earth Obs. Geoinformation 12, S45–S51. https://doi.org/10.1016/j.jag.2009.09.004

Jasechko, S., Sharp, Z.D., Gibson, J.J., Birks, S.J., Yi, Y., Fawcett, P.J., 2013. Terrestrial water fluxes dominated by transpiration. Nature 496, 347–350. https://doi.org/10.1038/nature11983

Kellner, J.R., Armston, J., Birrer, M., Cushman, K.C., Duncanson, L., Eck, C., Falleger, C., Imbach, B., Král, K., Krůček, M., Trochta, J., Vrška, T., Zgraggen, C., 2019. New opportunities for forest remote sensing through ultra-high-density drone Lidar. Surv. Geophys. 40, 959–977. https://doi.org/10.1007/s10712-019-09529-9

Khokthong, W., Zemp, D.C., Irawan, B., Sundawati, L., Kreft, H., Hölscher, D., 2019. Drone-based assessment of canopy cover for analyzing tree mortality in an oil palm agroforest. Front. For. Glob. Change 2, 12. https://doi.org/10.3389/ffgc.2019.00012

Kişi, O., Çimen, M., 2009. Evapotranspiration modelling using support vector machines / Modélisation de l'évapotranspiration à l'aide de 'support vector machines.' Hydrol. Sci. J. 54, 918–928. https://doi.org/10.1623/hysj.54.5.918

Kobayashi, N., Kumagai, T., Miyazawa, Y., Matsumoto, K., Tateishi, M., Lim, T.K., Mudd, R.G., Ziegler, A.D., Giambelluca, T.W., Yin, S., 2014. Transpiration characteristics of a rubber plantation in central Cambodia. Tree Physiol. 34, 285–301. https://doi.org/10.1093/treephys/tpu009

Kumar, M., Raghuwanshi, N.S., Singh, R., 2011. Artificial neural networks approach in evapotranspiration modeling: a review. Irrig. Sci. 29, 11–25. https://doi.org/10.1007/s00271-010-0230-8

Kume, T., Tsuruta, K., Komatsu, H., Kumagai, T., Higashi, N., Shinohara, Y., Otsuki, K., 2010. Effects of sample size on sap flux-based stand-scale transpiration estimates. Tree Physiol. 30, 129–138. https://doi.org/10.1093/treephys/tpp074

Legendre, P., Legendre, L., 2003. Numerical Ecology, 2/20. ed. Elsevier.

Liu, X., Kang, S., Li, F., 2009. Simulation of artificial neural network model for trunk sap flow of Pyrus pyrifolia and its comparison with multiple-linear regression. Agric. Water Manag. 96, 939–945. https://doi.org/10.1016/j.agwat.2009.01.003

Maier, H.R., Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environ. Model. Softw. 15, 101–124. https://doi.org/10.1016/S1364-8152(99)00007-9

Margono, B.A., Potapov, P.V., Turubanova, S., Stolle, F., Hansen, M.C., 2014. Primary forest cover loss in Indonesia over 2000–2012. Nat. Clim. Change 4, 730–735. https://doi.org/10.1038/nclimate2277

McKinney, W., Wes and others 2010. Data structures for statistical computing in Python. Proceedings of the 9th Python in Science Conference, 445, 51–56.

Meijide, A., Badu, C.S., Moyano, F., Tiralla, N., Gunawan, D., Knohl, A., 2018. Impact of forest conversion to oil palm and rubber plantations on microclimate and the role of the 2015 ENSO event. Agric. For. Meteorol. 252, 208–219. https://doi.org/10.1016/j.agrformet.2018.01.013

Mohan, M., Silva, C., Klauberg, C., Jat, P., Catts, G., Cardil, A., Hudak, A., Dia, M., 2017. Individual tree detection from Unmanned Aerial Vehicle (UAV) derived canopy height model in an open canopy mixed conifer forest. Forests 8, 340. https://doi.org/10.3390/f8090340

Nguyen, C.N., Zeigermann, O., 2018. Machine Learning: kurz & gut, 1. Auflage. ed, O'Reillys Taschenbibliothek. O'Reilly, Heidelberg.

Niu, F., Roll, A., Hardanto, A., Meijide, A., Kohler, M., Hendrayanto, Holscher, D., 2015. Oil palm water use: calibration of a sap flux method and a field measurement scheme. Tree Physiol. 35, 563–573. https://doi.org/10.1093/treephys/tpv013

Niu, F., Röll, A., Meijide, A., Hendrayanto, Hölscher, D., 2017. Rubber tree transpiration in the lowlands of Sumatra. Ecohydrology 10, e1882. https://doi.org/10.1002/eco.1882

Oliphant, T.E., 2006. A guide to NumPy. Trelgol Publ.

Opitz, D., Maclin, R., 1999. Popular ensemble methods: an empirical study. J. Artif. Intell. Res. 11, 169–198. https://doi.org/10.1613/jair.614

Özçelik, R., Diamantopoulou, M.J., Crecente-Campo, F., Eler, U., 2013. Estimating Crimean juniper tree height using nonlinear regression and artificial neural network models. For. Ecol. Manag. 306, 52–60. https://doi.org/10.1016/j.foreco.2013.06.009

Pan, S., Pan, N., Tian, H., Friedlingstein, P., Sitch, S., Shi, H., Arora, V.K., Haverd, V., Jain, A.K., Kato, E., Lienert, S., Lombardozzi, D., Ottle, C., Poulter, B., Zaehle, S., 2019. Evaluation of global terrestrial evapotranspiration by state-of-the-art approaches in remote sensing, machine learning, and land surface models (preprint). Global hydrology/Modelling approaches. https://doi.org/10.5194/hess-2019-409

Panda, S., Amatya, D.M., Hoogenboom, G., 2014. Stomatal conductance, canopy temperature, and leaf area index estimation using remote sensing and OBIA techniques. Journal of Spatial Hydrology 12, 25.

Passing, H., Bablok, W., 1984. Comparison of several regression procedures for method comparison studies and determination of sample sizes application of linear regression procedures for method comparison studies in clinical chemistry, Part II. Clin. Chem. Lab. Med. 22. https://doi.org/10.1515/cclm.1984.22.6.431

Passing, H., Bablok, W., 1983. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. Clin. Chem. Lab. Med. 21. https://doi.org/10.1515/cclm.1983.21.11.709

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., 2011. Scikit-learn: Machine Learning in Python. Scikit-Learn Mach. Learn. Python 12, 6.

Raschka, S., 2017. Machine Learning mit Python: das Praxis-Handbuch für Data Science, Predictive Analytics und Deep Learning, 1. Auflage. ed. mitp, Frechen.

Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., Chica-Rivas, M., 2015. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geol. Rev. 71, 804–818. https://doi.org/10.1016/j.oregeorev.2015.01.001

Rokach, L., Maimon, O., 2015. Data mining with decision trees: theory and applications, Second edition. ed. World Scientific, Hackensack, New Jersey.

Röll, A., Niu, F., Meijide, A., Ahongshangbam, J., Ehbrecht, M., Guillaume, T., Gunawan, D., Hardanto, A., Hendrayanto, Hertel, D., Kotowska, M.M., Kreft, H., Kuzyakov, Y., Leuschner, C., Nomura, M., Polle, A., Rembold, K., Sahner, J., Seidel, D., Zemp, D.C., Knohl, A., Hölscher, D., 2019. Transpiration on the rebound in lowland Sumatra. Agric. For. Meteorol. 274, 160–171. https://doi.org/10.1016/j.agrformet.2019.04.017

Röll, A., Niu, F., Meijide, A., Hardanto, A., Hendrayanto, Knohl, A., Hölscher, D., 2015. Transpiration in an oil palm landscape: effects of palm age. Biogeosciences 12, 5619–5633. https://doi.org/10.5194/bg-12-5619-2015

Schapire, R.E., 1990. The strength of weak learnability. Mach Learn 5, 197–227. https://doi.org/10.1007/BF00116037

SciPy 1.0 Contributors, Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods. https://doi.org/10.1038/s41592-019-0686-2

Shrestha, N.K., Shukla, S., 2015. Support vector machine based modeling of evapotranspiration using hydro-climatic variables in a sub-tropical environment. Agric. For. Meteorol. 200, 172–184. https://doi.org/10.1016/j.agrformet.2014.09.025

Stiegler, C., Meijide, A., Fan, Y., Ashween Ali, A., June, T., Knohl, A., 2019. El Niño–Southern Oscillation (ENSO) event reduces CO 2 uptake of an Indonesian oil palm plantation. Biogeosciences 16, 2873–2890. https://doi.org/10.5194/bg-16-2873-2019

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC Bioinformatics 9, 307. https://doi.org/10.1186/1471-2105-9-307

Suab, S.A., Avtar, R., 2020. Unmanned Aerial Vehicle System (UAVS) applications in forestry and plantation operations: experiences in sabah and sarawak, Malaysian Borneo, in: Avtar, R., Watanabe, T. (Eds.), Unmanned Aerial Vehicle: Applications in Agriculture and Environment. Springer International Publishing, Cham, pp. 101–118. https://doi.org/10.1007/978-3-030-27157-2_8

Tavares Júnior, I., Rocha, J., Ebling, Â., Chaves, A., Zanuncio, J., Farias, A., Leite, H., 2019. Artificial neural networks and linear regression reduce sample intensity to predict the commercial volume of eucalyptus clones. Forests 10, 268. https://doi.org/10.3390/f10030268

Teuscher, M., Gérard, A., Brose, U., Buchori, D., Clough, Y., Ehbrecht, M., Hölscher, D., Irawan, B., Sundawati, L., Wollni, M., Kreft, H., 2016. Experimental biodiversity enrichment in oil-palm-dominated landscapes in Indonesia. Front. Plant Sci. 07. https://doi.org/10.3389/fpls.2016.01538

Timmermans, W.J., Kustas, W.P., Andreu, A., 2015. Utility of an automated thermal-based approach for monitoring evapotranspiration. Acta Geophys. 63, 1571–1608. https://doi.org/10.1515/acgeo-2015-0016

van der Walt, S., Colbert, S.C., Varoquaux, G., 2011. The NumPy array: a structure for efficient numerical computation. Comput. Sci. Eng. 13, 22–30. https://doi.org/10.1109/MCSE.2011.37

Vapnik, V.N., 2000. Introduction: four periods in the research of the learning problem, in: The Nature of Statistical Learning Theory. Springer New York, New York, NY, pp. 1–15. https://doi.org/10.1007/978-1-4757-3264-1_1

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. Precis. Agric. https://doi.org/10.1007/s11119-020-09711-9

Waite, P.-A., Schuldt, B., Mathias Link, R., Breidenbach, N., Triadiati, T., Hennings, N., Saad, A., Leuschner, C., 2019. Soil moisture regime and palm height influence embolism resistance in oil palm. Tree Physiol. 39, 1696–1712. https://doi.org/10.1093/treephys/tpz061

Waskom, M., Botvinnik, O., Ostblom, J., 2020. mwaskom/seaborn: seaborn v0.8.1, Zenodo, https://doi.org/10.5281/zenodo.883859

Wei, Z., Yoshimura, K., Wang, L., Miralles, D.G., Jasechko, S., Lee, X., 2017. Revisiting the contribution of transpiration to global terrestrial evapotranspiration: Revisiting Global ET Partitioning. Geophys. Res. Lett. 44, 2792–2801. https://doi.org/10.1002/2016GL072235

Wu, C., Shen, H., Shen, A., Deng, J., Gan, M., Zhu, J., Xu, H., Wang, K., 2016. Comparison of machine-learning methods for above-ground biomass estimation based on Landsat imagery. J. Appl. Remote Sens. 10, 035010. https://doi.org/10.1117/1.JRS.10.035010

Chapter 5

Introducing QWaterModel, a QGIS plugin for predicting evapotranspiration from land surface temperatures

INTRODUCING QWATERMODEL, A QGIS PLUGIN FOR PREDICTING EVAPOTRANSPIRATION FROM LAND SURFACE TEMPERATURES

Florian Ellsäßer [1],Alexander Röll [1], Christian Stiegler [2], Hendrayanto [3], Dirk Hölscher [1,4]

[1] *University of Goettingen, Tropical Silviculture and Forest Ecology, Büsgenweg 1, 37077 Göttingen, Germany*
[2] *University of Goettingen, Bioclimatology, Büsgenweg 2, 37077 Göttingen Germany*
[3] *Bogor Agricultural University, Forest Management, Kampus IPB Darmaga, 16680 Bogor, Indonesia*
[4] *University of Goettingen, Centre of Biodiversity and Sustainable Land Use, Büsgenweg 1, 37073 Göttingen, Germany*

Published in Environmental Modelling & Software
Available at: https://doi.org/10.1016/j.envsoft.2020.104739

ABSTRACT

Evapotranspiration (ET) is a central flux in the hydrological cycle. Various approaches to compute ET via energy balance models exist, but their handling is often complex and challenging. We developed QWaterModel as an easy-to-use tool to make ET predictions available to broader audiences. QWaterModel is based on the DATTUTDUT energy balance model and uses land surface temperature maps as an input. Such maps can e.g. be obtained from satellite, drone or handheld camera imagery. In the present study, we successfully tested QWaterModel for predicting ET in a tropical oil palm plantation against the well-established eddy covariance method. QWaterModel is compatible with all versions of QGIS3 and is available from the official QGIS Plugin Repository.

Software Availability:
Version: 1.0 / Software licence: GNU - General Public License
Availability: plugins.qgis.org/plugins/qwatermodel / github.com/FloEll/QWaterModel

5.1 Introduction

Evapotranspiration (ET) is the combined water flux of evaporation from soil, plant and water surfaces as well as transpiration from plants (Allen et al., 1998). Terrestrial ET is a major flux in the hydrological cycle consuming about 60% of terrestrial precipitation (Oki and Kanae, 2006). Associated with climate and land-use change, major transformations in the global hydrological cycle are projected; therefore, a broad understanding and knowledge of ET and its patterns are of paramount importance (Kaushal et al., 2017; Ziegler et al., 2003). The current understanding of how ecosystems respond to such changes is limited by insufficient monitoring capabilities (Fisher et al., 2017). The development of effective adaption strategies for agriculture, ecosystems and water management will depend on the availability of ET assessment schemes that can be readily ap-

plied from local to global scales (Fisher et al., 2017). ET can be measured locally using e.g. the eddy covariance method or estimated at larger spatial scales by applying energy balance models. For many energy balance models remotely-sensed land surface temperatures (LST) are used as principal input, therein assuming that hot pixels are the result of low ET and cold pixels indicate high ET (Timmermans et al., 2015). Current developments such as increasing computation power, extensive availability of free satellite imagery (e.g. from Landsat 7 and 8) and a large variety of portable thermal cameras that can be attached to drones or used as handheld devices foster the use of energy balance modeling for ET estimation (Hoffmann et al., 2016; Maes and Steppe, 2012; Timmermans et al., 2015; Xia et al., 2016). However, most energy balance models are complex and their implementation requires advanced programming skills and technical expertise, which constitutes a barrier for the application of energy balance models for ET estimation. To overcome this barrier, we developed the QGIS plugin 'QWaterModel', which addresses the following objectives: to facilitate the use of energy balance and evapotranspiration modeling, to take LST images and maps from a wide range of sources including satellites, planes, drones and handheld thermal cameras as input, to reduce the use of complementary data to a feasible level and to use a documented open source structure to encourage further development. In the present study, we explain and test key-features of QWaterModel in a scientific context, using exemplary LST data from different sources recorded in an oil palm plantation in Indonesia to compute ET estimates. We then compare the ET predictions from the plugin with simultaneous ET reference measurements with the well-established eddy covariance method.

5.2 Methods

5.2.1 Software design

To facilitate the use of energy balance modeling for ET estimation for a broad audience, we chose QGIS3 (QGIS Development Team, 2020) as a platform. QGIS is a free and open source geographic information system that provides a versatile environment for work flows with spatial data such as LST maps. It offers subsystems for data input and management, analysis, easy visualization of spatial data, has a large community of developers and is supported by most operating systems (Bhatt et al., 2014; Criollo et al., 2019). The use of QGIS is frequently taught in scientific institutions and its functionality can be extended with a large variety of available plugins. For example, the WET (Water Ecosystems Tool) plugin (Nielsen et al., 2017) provides easy access to complex watershed modeling. To our knowledge, no such easy-to-use plugins exist in the QGIS environment for instantaneous flux modeling or ET prediction. The presented plugin, QWaterModel, was

developed to fill this gap. It is based on the energy balance model DATTUTDUT (Deriving Atmosphere Turbulent Transport Useful To Dummies Using Temperature) (Timmermans et al., 2015), which uses LST maps as main input and, except for time and location of the LST recording, requires no further input of ancillary data. In the original DATTUTDUT model the radiation budget is simply modeled from a set of parameters and sun-earth geometrics (Timmermans et al., 2015). This approach works well for cloud-free conditions. For a broader usability under conditions of cloudy skies and high relative air humidity we extended this original model concept, giving the user the possibility to configure model parameters according to in-situ field measurements. The radiation budget can be complemented by measurements of short-wave irradiance or net radiation at the study sites via manual input. The DATTUTDUT model has successfully been tested for a broad range of LST input types, from satellite recorded images (Timmermans et al., 2015) to plane and drone recorded images (Brenner et al., 2018; Xia et al., 2016). The plugin was built using Plugin Builder 3.2.1 (GeoApt LLC, 2019) and Plugin Reloader 0.7.9 (Jurgiel, 2020). We used Qt Creator 4.11.0 (The Qt Company, 2020) to develop the graphical user interface (GUI) and Python 3.8 with the QGIS3/Python standard libraries (gdal, math, numpy, datetime and os) for the associated functionality. Detailed installation instructions and the source code can be found on: github.com/FloEll/QWaterModel/blob/master/README.md

5.2.2 Application and testing

To test the plugin, we used data recorded in a mature monoculture oil palm (Elaeis guineensis Jacq.) plantation located in the lowlands of Sumatra (Jambi province, Indonesia, 103.3914411 E, -1.6929879 N, 76 m a.s.l., see Meijide et al. (2017) for further details). We exemplarily use LST recordings from two independent sources (drone and handheld camera) to test the QWaterModel plugin against eddy covariance reference measurements:

1. Images recorded with an octocopter drone (MK EASY Okto V3; HiSystems, Germany) equipped with a radiometric thermal camera FLIR Tau 2 640 (FLIR Systems, USA) attached to a TeAx Thermo-capture module (TeAx Technology, Germany). The data have a spatial resolution of 0.2 m and cover a footprint of 0.8 ha and are thus suitable for tree-scale to stand-scale assessments of ET. 13 images from 09:00 to 15:00 h local time on 7th of August 2017 were recorded in 30 min intervals.

2. Images taken from a tower at 10 m above the canopy with a handheld Fluke ti100 thermal camera (Fluke Systems, USA). The data have an approx. spatial resolution of 0.06 m and a footprint of 0.006 ha and are thus suitable for leaf to canopy assessments of ET. 5 images from 11:00 to 15:00 h on 7th of August 2017 were taken in 60 min intervals.

Data are provided for download at: github.com/FloEll/QWaterModel/tree/master/Data_Examples

We did not include satellite images into this study since this was already included in the original DATTUT-DUT model study (Timmermans et al., 2015). In addition, satellite (Landsat 7 and 8) recorded LST maps for our study region contain a very high cloud cover fraction (for 2017: min. 23%, mean 72%, max. 100%). Our study thus focuses on drone and handheld thermal camera acquired images.
A widely accepted ground-based reference method for ET assessments is the eddy covariance method, which provides measurements of ET at the stand-scale. Eddy covariance data at the oil palm site was recorded, filtered and processed according to the methodology described in Meijide et al. (2017). Because the energy-balance model used in the plugin assumes full energy balance closure, eddy covariance data was processed using the Bowen ratio closure method (Pan et al., 2017; Twine et al., 2000). Horizontal energy flows or incomplete energy balance closure might introduce certain errors to this reference method (Loescher et al., 2006). To derive ET from latent heat fluxes, latent heat of vaporization was calculated using in-situ air temperature measurements for eddy covariance measurements and the lowest pixel temperatures from the LST maps for the plugin estimates following the methodology described in Timmermans et al. (2015).

5.3 Results

5.3.1 Software implementation results

The QWaterModel plugin can be installed from the official QGIS3 python plugin repository. The plugin GUI consists of a main window with four different input sections (A, B, C, D, Figure 5.1) and an information window, which are both organized using tabs. In the main window, section A contains the inputs that are essential for the plugin to work with in a minimal data approach: a thermal image with temperatures in Kelvin (e.g. from satellite, drone or handheld camera), the definition of an output raster and an output file where key statistics will be summarized, and the coordinated universal time (UTC) when the picture was taken. Providing information in boxes B, C and D is optional, but may improve the quality of the ET estimates. The default values that are visible in the input fields of box B are taken from Timmermans et

al. (2015) and can be adjusted if more local data are available from field measurements or previous studies. The info tab offers detailed information on input, output and usage of the plugin. The plugin outputs a raster with 6 bands (net radiation Rn [W m^{-2}], latent heat flux LE [W m^{-2}], sensible heat flux H [W m^{-2}], ground heat flux G [W m^{-2}], evaporative fraction EF in [%], evapotranspiration ET [mm $time^{-1}$]). Input raster details and output statistics are stored in a .csv file.
Figure 5.2 shows a workflow example with LST data from different sources, i.e. recorded from drone and handheld thermal camera and the corresponding output.

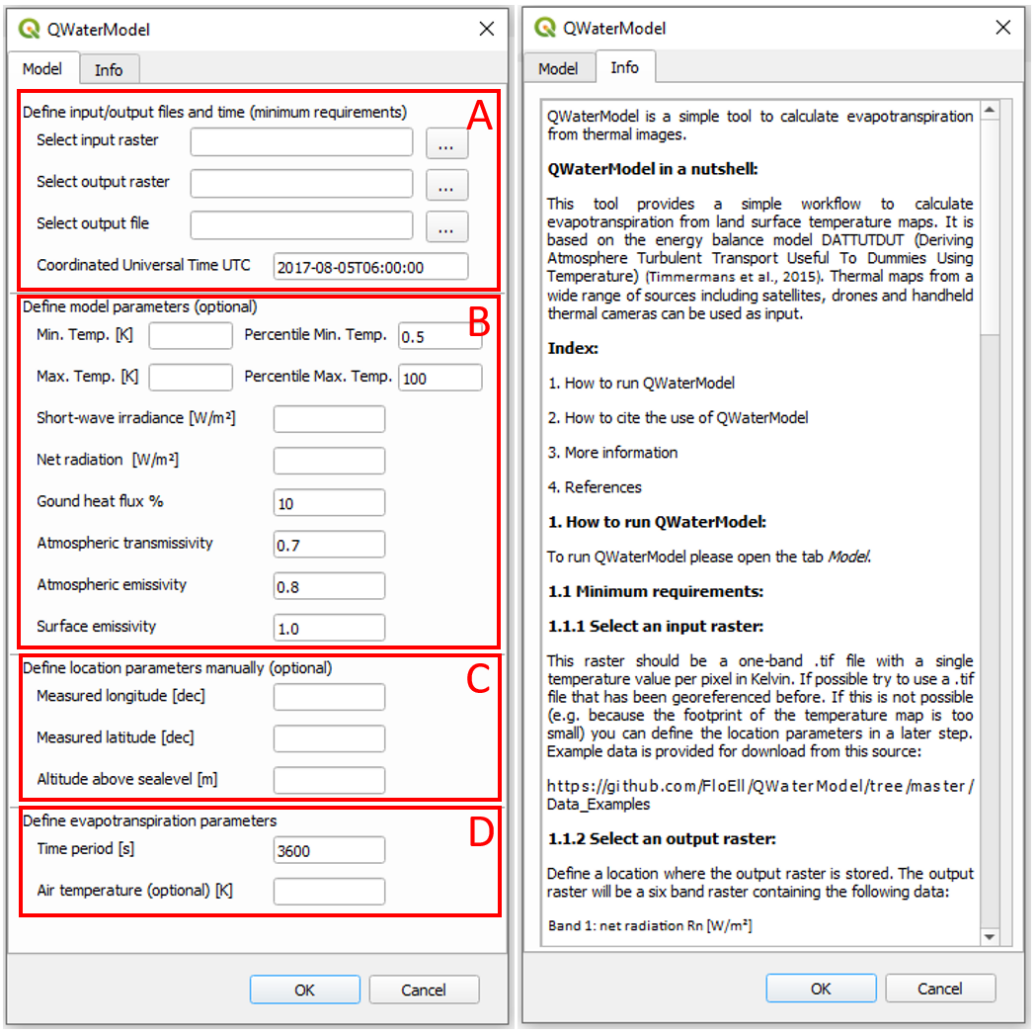


Figure 5.1: *Both tabs of the graphical user interface (GUI) of the plugin version 1.0 with the 'Model'-tab input boxes highlighted in red.*

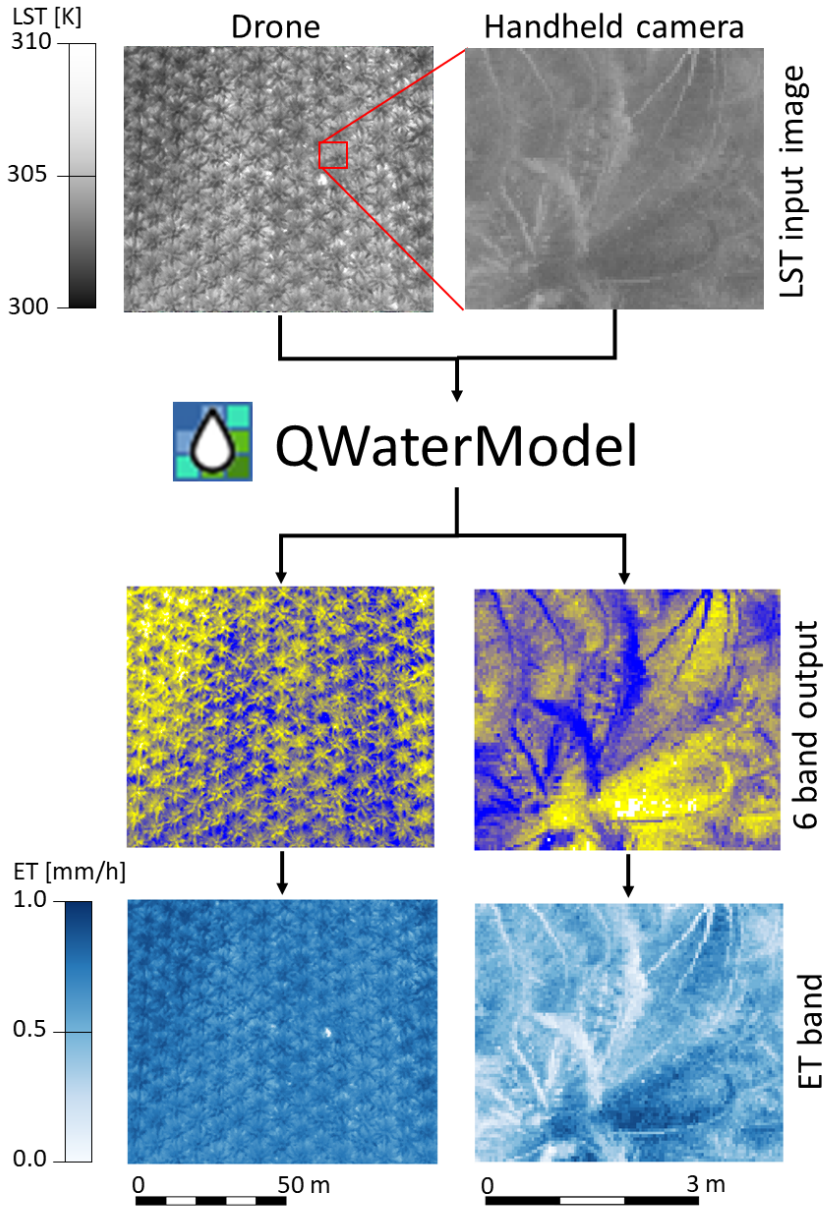


Figure 5.2: *The ET modelling workflow with input and output examples. The input files are one-layer .tif files where each pixel contains a temperature value in Kelvin. The output files are raster files with 6 bands of which the ET estimates are contained in the 6th band, which is presented separately in the lower panels. ET was extracted from the 6-band output raster using the Rearrange Bands tool from the Processing Toolbox in QGIS 3.10.*

5.3.2 Software experimental testing results

Land surface temperatures differed with the time of recording. Thermal variability was high in both drone and handheld camera obtained maps (Table 1). To compare the ET estimates from drone and handheld thermal camera maps with a measured reference, we plotted them against the daily course of eddy covariance derived ET values (Figure 5.3) and into a scatterplot comparing measurements and predictions of ET (Figure 5.4). Applying the energy balance model in its original version with minimal data input resulted in overestimations of ET during the morning and afternoon hours, but in acceptable estimates around noon (Figure 5.3a). Errors were smaller for the estimates based on handheld camera pictures than on drone recordings (Figure 5.4 and Table 2). More precise results were achieved by adding measured short-wave irradiance (Figure 5.3b) or net radiation measurements (Figure 5.3c) to the model inputs, i.e. ET estimates closely follow the reference method if additional radiation data is supplied. The handheld thermal camera maps generally produce slightly more congruent results with the eddy covariance measurements than drone maps; therein, around noon ET estimates from drone maps are slightly higher than eddy covariance ET measurements. In the minimal data approach, ET predictions are most congruent with eddy covariance measurements when ET is high (Figure 5.4a), while in contrast predictions are most congruent at low ET when measurements of short-wave irradiance or net radiation are used in the model (Figure 5.4b and 5.4c). Congruence generally increases with increasing measurement-based input; errors are thus smallest for models computed with directly measured net radiation (Table 2).

Table 1: Input key-values of land surface temperatures from both sources.

Temperatures	Drone (13 maps)	Handheld camera (5 maps)
Mean Temp. [K]	298.3 – 305.5	300.8 – 308.3
Minimum Temp. [K]	297.2 – 303.0	299.6 – 305.5
Maximum Temp. [K]	301.2 – 325.1	302.4 – 317.0

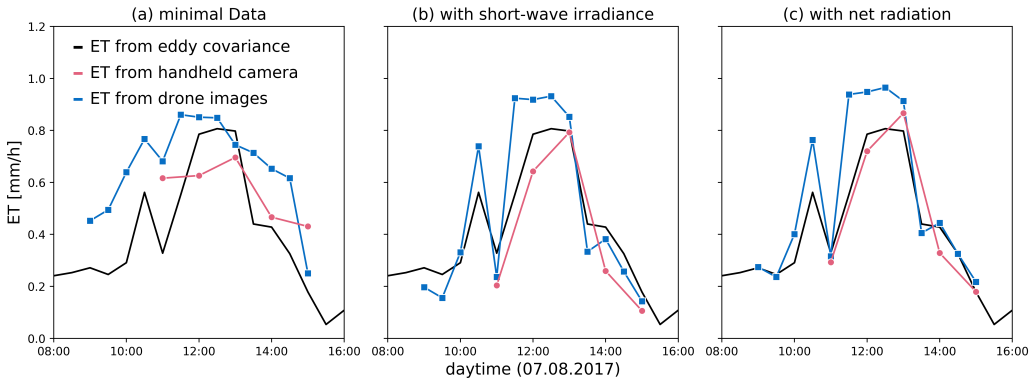


Figure 5.3: *Temporal comparison of measured reference ET from eddy covariance and estimated ET from different LST sources during the course of a day.*

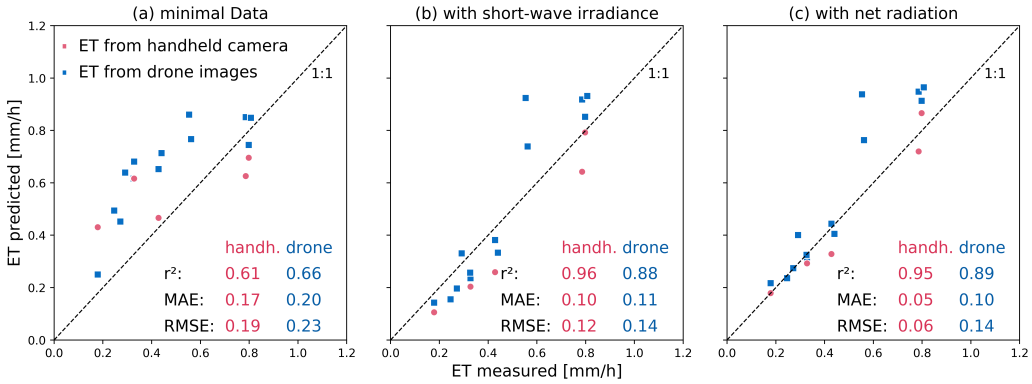


Figure 5.4: *Comparison of measured reference ET from eddy covariance and estimated ET from different LST sources; metrics on the respective relationships are shown in the figure and provided in Table 2.*

Table 2: Key metrics of linear regression between in-situ measurements and model output for different data input options (a-c) and recording methods (handheld thermal camera and drone).

Metrics	(a) minimal data		(b) with short wave irradiance		(c) with net radiation	
	handheld	drone	handheld	drone	handheld	drone
r^2	0.61	0.66	0.96	0.88	0.95	0.89
MAE	0.17	0.2	0.1	0.11	0.05	0.1
RMSE	0.19	0.23	0.12	0.14	0.06	0.14

5.4 Discussion and conclusions

The QWaterModel plugin is an easy-to-use open-source tool for predicting evapotranspiration from land surface temperatures. It is aimed at applications in ecology, bioclimatology and land-use science, but is also useful for a broader utilization including hydrological and ecosystem management. It offers a link between science and practical application of energy balance modelling for ET prediction. Meeting our objectives, a variety of LST data from very different sources and spatial extends can be used as input and complementary data input requirements were kept at a minimum level. The case study showed that images taken at the same time but with different camera setups (drone and handheld) result in comparable results for ET estimates and show high congruence with reference eddy covariance measurements. Because in the minimal data approach the radiation budget is calculated from sun-earth geometry, clouds and relative humidity or haze are not considered. This explains the overestimation of ET in the morning and afternoon hours (Figure 5.3a). In-situ measurements of short-wave irradiance or net radiation significantly improved the ET estimates (Figure 5.3b and 5.3c). Similar observations were made in other studies e.g. on European grasslands where additional short-wave irradiance measurements improved the overall accuracy of the DATTUTDUT model (Brenner et al., 2018). A method comparison of the DATTUTDUT model with the eddy covariance method over varying weather conditions and day times showed the general applicability of this approach in the tropics and suggested no difference between drone-based and eddy covariance method for certain configurations (Ellsäßer et al., under review). However, even when using measured short-wave or net radiation, ET estimates based on drone recorded LST data showed overestimations around noon, which did not occur in the estimates based on handheld images. A possible reason is the presence of artefacts (e.g. roofs, cars, rocks) in the drone images, which have a much larger footprint than the handheld images. Temperatures on such non-canopy components can differ greatly from surface temperatures of vegetation, and these temperature outliers have a strong effect on the quality of predictions. This was also observed in a study with big temperature differences conducted in vineyards (Xia et al., 2016). A potential solution would be to manually exclude known artefacts from the images before analysis or to manually define minimal and maximal temperatures in the GUI of the plugin. Unfortunately, no matching satellite images were available for the time and date when drone and handheld camera maps were recorded therefore further analyses are not performed in this study. This should be followed up upon in future studies, even though a major restriction in our study region is that satellite images are generally hard to ac-

quire since a considerable cloud cover is present on most days of the year. A further challenge is the fixed overpass time of satellites, which is often not around noon and therefore not ideal for ET derivation (Delogu et al., 2012). To overcome these limitations, complementary ET methods from planes, drones and handheld or fixed thermal cameras can potentially be calibrated against simultaneously acquired satellite images (e.g. using QWaterModel) and then be applied to increase temporal and spatial coverage.

The scatterplots reveal a gap between lower and higher ET estimates (Figure 5.4b and 5.4c). This gap and its absence in the ET estimates based on modelled Rn in Figure 5.4a demonstrates the effects of clouds on prediction accuracy. If clouds are present, and measurements of short-wave irradiance or net radiation can represent this adequately in the model, ET estimates are clearly lower than with the Rn modelling approach. Future improvements of the software could include a referencing tool that links ground measurements of hot and cold surfaces with LST data. This would allow for a simple radiometric correction of LST images and maps. To make the plugin applicable to the needs of a broader range of users, upcoming versions could further include other LST map-based indices such as the CWSI (crop water stress index) (Bian et al., 2019; Idso et al., 1981; Jones, 2014). We appreciate all user input and ideas to further develop the plugin and encourage to modify the plugin for other ET related applications or to report bugs via our bug tracker. In the present study, the software was only tested at one site, a commercial monoculture oil palm plantation in lowland Sumatra. We strongly encourage the testing and use of the software in other climatic regions and on other crop and ecosystem types for further validation. We provide the source code for the plugin and plan to consistently add more functionality to the QWaterModel.

REFERENCES

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56. FAO, Rome.

Bhatt, G., Kumar, M., Duffy, C.J., 2014. A tightly coupled GIS and distributed hydrologic modeling framework. Environ. Model. Softw. 62, 70–84. https://doi.org/10.1016/j.envsoft.2014.08.003

Bian, J., Zhang, Z., Chen, J., Chen, H., Cui, C., Li, X., Chen, S., Fu, Q., 2019. Simplified evaluation of cotton water stress using high resolution unmanned aerial vehicle thermal imagery. Remote Sens. 11, 267. https://doi.org/10.3390/rs11030267

Brenner, C., Zeeman, M., Bernhardt, M., Schulz, K., 2018. Estimation of evapotranspiration of temperate grassland based on high-resolution thermal and visible range imagery from unmanned aerial systems. Int. J. Remote Sens. 39, 5141–5174. https://doi.org/10.1080/01431161.2018.1471550

Criollo, R., Velasco, V., Nardi, A., Manuel de Vries, L., Riera, C., Scheiber, L., Jurado, A., Brouyère, S., Pujades, E., Rossetto, R., Vázquez-Suñé, E., 2019. AkvaGIS: An open source tool for water quantity and quality management. Comput. Geosci. 127, 123–132. https://doi.org/10.1016/j.cageo.2018.10.012

Delogu, E., Boulet, G., Olioso, A., Coudert, B., Chirouze, J., Ceschia, E., Le Dantec, V., Marloie, O., Chehbouni, G., Lagouarde, J.-P., 2012. Reconstruction of temporal variations of evapotranspiration using instantaneous estimates at the time of satellite overpass. Hydrol. Earth Syst. Sci. 16, 2995–3010. https://doi.org/10.5194/hess-16-2995-2012

Ellsäßer, F., Stiegler, C., Röll, A., June, T., Hendrayanto, Knohl, A., Hölscher, D., submitted. Predicting evapotranspiration from drone-based thermography – a method comparison in a tropical oil palm plantation. Biogeosciences.

Fisher, J.B., Melton, F., Middleton, E., Hain, C.,

Anderson, M., Allen, R., McCabe, M.F., Hook, S., Baldocchi, D., Townsend, P.A., Kilic, A., Tu, K., Miralles, D.D., Perret, J., Lagouarde, J.-P., Waliser, D., Purdy, A.J., French, A., Schimel, D., Famiglietti, J.S., Stephens, G., Wood, E.F., 2017. The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources: The future of evapotranspiration. Water Resour. Res. 53, 2618–2626. https://doi.org/10.1002/2016WR020175

Hoffmann, H., Nieto, H., Jensen, R., Guzinski, R., Zarco-Tejada, P., Friborg, T., 2016. Estimating evaporation with thermal UAV data and two-source energy balance models. Hydrol. Earth Syst. Sci. 20, 697–713. https://doi.org/10.5194/hess-20-697-2016

Idso, S.B., Jackson, R.D., Pinter, P.J., Reginato, R.J., Hatfield, J.L., 1981. Normalizing the stress-degree-day parameter for environmental variability. Agric. Meteorol. 24, 45–55. https://doi.org/10.1016/0002-1571(81)90032-7

Jones, H.G., 2014. Plants and microclimate: a quantitative approach to environmental plant physiology, Third edition. ed. Cambridge University Press, Cambridge; New York.

Kaushal, S., Gold, A., Mayer, P., 2017. Land use, climate, and water resources—global stages of interaction. Water 9, 815. https://doi.org/10.3390/w9100815

Loescher, H.W., Law, B.E., Mahrt, L., Hollinger, D.Y., Campbell, J., Wofsy, S.C., 2006. Uncertainties in, and interpretation of, carbon flux estimates using the eddy covariance technique. J. Geophys. Res. 111, D21S90. https://doi.org/10.1029/2005JD006932

Maes, W.H., Steppe, K., 2012. Estimating evapotranspiration and drought stress with ground-based thermal remote sensing in agriculture: a review. J. Exp. Bot. 63, 4671–4712. https://doi.org/10.1093/jxb/ers165

Meijide, A., Röll, A., Fan, Y., Herbst, M., Niu, F., Tiedemann, F., June, T., Rauf, A., Hölscher, D., Knohl, A., 2017. Controls of water and energy fluxes in oil palm plantations: Environmental variables and oil palm age. Agric. For. Meteorol. 239, 71–85. https://doi.org/10.1016/j.agrformet.2017.02.034

Nielsen, A., Bolding, K., Hu, F., Trolle, D., 2017. An open source QGIS-based workflow for model application and experimentation with aquatic ecosystems. Environ. Model. Softw. 95, 358–364. https://doi.org/10.1016/j.envsoft.2017.06.032

Oki, T., Kanae, S., 2006. Global hydrological cycles and world water resources. Am. Assoc. Adv. Sci. 313, 1068–1072. https://doi.org/10.1126/science.1128845

Pan, X., Liu, Y., Fan, X., Gan, G., 2017. Two energy balance closure approaches: applications and comparisons over an oasis-desert ecotone. J. Arid Land 9, 51–64. https://doi.org/10.1007/s40333-016-0063-2

Timmermans, W.J., Kustas, W.P., Andreu, A., 2015. Utility of an automated thermal-based approach for monitoring evapotranspiration. Acta Geophys. 63, 1571–1608. https://doi.org/10.1515/acgeo-2015-0016

Twine, T.E., Kustas, W.P., Norman, J.M., Cook, D.R., Houser, P.R., Meyers, T.P., Prueger, J.H., Starks, P.J., Wesely, M.L., 2000. Correcting eddy-covariance flux underestimates over a grassland. Agric. For. Meteorol. 103, 279–300. https://doi.org/10.1016/S0168-1923(00)00123-4

Xia, T., Kustas, W.P., Anderson, M.C., Alfieri, J.G., Gao, F., McKee, L., Prueger, J.H., Geli, H.M.E., Neale, C.M.U., Sanchez, L., Alsina, M.M., Wang, Z., 2016. Mapping evapotranspiration with high-resolution aircraft imagery over vineyards using one- and two-source modeling schemes. Hydrol. Earth Syst. Sci. 20, 1523–1545. https://doi.org/10.5194/hess-20-1523-2016

Ziegler, A.D., Sheffield, J., Maurer, E.P., Nijssen, B., Wood, E.F., Lettenmaier, D.P., 2003. Detection of intensification in global- and continental-scale hydrological cycles: temporal scale of evaluation. J. Clim. 16, 13.

Chapter 6

Conclusions

This dissertation comprises a combination of three individual studies on evapotranspiration (ET) and plant-water-use prediction with remote sensing thermography using drones. The first two studies (Chapter 3 and 4) are focused on method comparisons where the well-known and established methods – eddy covariance, sap flux and porometry – are compared with modelling and prediction approaches, based on drone recorded and auxiliary data. The last study (Chapter 5) summarizes key ideas and findings, which result from the first two studies, and crafts them into an easy to use software. From the findings in all three papers, overall conclusions on the potential of drone recorded data and predictions in hydrological modelling can be drawn. Furthermore, possible future use case scenarios and applications of this methodology and its development options are illuminated.

This dissertation presents two very common prediction approaches for hydrological features of interest: (1) The energy balance model (EBM) prediction methods for ET are based on causal relations between the plant surface, the atmosphere and the boundary layer between both. (2) The statistical and machine learning (ML) prediction approaches for sap flux and stomatal conductance, which use emerging patterns among multiple variables in the data set to compute predictions. All prediction methods have overall shown very promising results. Given enough data on solar radiation, ET could well be predicted by all EBMs applied in Chapter 3 and 5. Daily evapotranspiration patterns were accurately represented by the EBMs model results, especially when radiation budget measurements were available. Frequency distributions of latent heat fluxes showed, that increasing complexity of the EBM also resulted in a more heterogenous representation of flux variability, while not necessarily contributing to an overall more accurate result, compared to the eddy covariance reference. This more complex representation of frequency distribution can be interpreted as a result of a better physical

representation of energy and radiative exchange by the model (Xia et al., 2016).

Since the statistical and ML based prediction models in Chapter 4 lack the representation of causal relations of the biosphere and are purely data driven, they represent an entirely different approach. Sap flux was overall well predicted, and random forest prediction methods have proofed to be by far the most promising approach. While a prediction of both dicot trees and monocot palms in the same data set remains a challenging matter, the prediction of sap flux from trees and oil palm in separate data sets resulted in very precise predictions. Vast differences in water use between trees and oil palm have been observed for this region (Niu et al., 2017; Röll et al., 2019, 2015) and might be the main reason for the challenges to predict sap flux from both dicot and monocot canopies in the same data set. Compared to the EBMs, where the simplistic DATTUTDUT approach computed the most congruent predictions, and the more complex EBMs better represented diverse canopy patterns, an increase in complexity of the prediction models from the simple multiple linear regression to the random forest regressors improved prediction quality significantly. However, when the much more complex artificial neural network has been applied to the same data, prediction accuracy was largely reduced. The prediction success of data driven models is highly dependent on the quantity of input data. Random forest models are known to be able to compute decent predictions already from a very limited number of samples (dos Reis et al., 2018), whereas artificial neural networks benefit from a large data base. The data set used for the study in Chapter 4 was already very large in terms of a conventional analysis. However, it might still not have a sufficient size for the application of artificial neural networks yet.

Major challenges to predict stomatal conductance from the recorded data set remain. Compared to the sap flux data base, the data set for stomatal conductance was only half the size, and temporal offsets were by average 20 times larger. Another uncertainty was introduced by up-scaling 7-10 measurements of stomatal conductance covering a centimetre scale of leaf area to an entire canopy. Since we applied strictly data driven algorithms, the amount of data and its accuracy are the most crucial components of the input data set and might have directly led to the high variance in prediction results for stomatal conductance.

The applied model II regressions (Deming and Passing-Bablok) found interchangeability of the well-known methods with the drone data-based approaches for both causal and data driven prediction strategies. However, such method comparisons allowing measurement bias in both X and Y variable methods are rarely applied in forest hydrology studies. Our comparatively large data sets have opened up the possibilities for full method comparisons. However, a profound

comparison of the presented results of the model II method comparisons with similar studies was not possible, since no such studies – using equally large or larger data sets or similar model II method comparisons – have yet been published.

Performing up to 16 survey flights per day, it was possible to record the daily course of land surface temperatures and model evapotranspiration in a very high frequency. This temporal resolution is denser than in any other published drone-based study on plant water traits, and almost resembles studies with constantly installed thermal cameras (Aubrecht et al., 2016), while it additionally covers a much larger area. This dense observation of ET and plant water use allows for a detailed insight into water use patterns and sets the presented studies apart from others with a much lower temporal or spatial resolution. The daily course of latent heat fluxes in oil palm from the EBM approach peaked between 600 and 700 W m^{-1}, with some outliers exceeding these values. These findings are highly congruent with eddy covariance measurements, where some EBM configurations are more likely to match these measurements taken in the morning and afternoon hours and others during the peak ET at noon.

The original version of the DATTUTDUT model (Timmermans et al., 2015) works only under clear sky conditions, therefore assuming the absence of clouds. However, since the sky has usually been clouded and a high air humidity has been present at our study site, modelling ET from only a thermal image, time and location proved very difficult. The original DATTUTDUT model was therefore modified to account for cloud cover and humidity. These adapted versions of the DATTUTDUT model, which have been modified to include radiation measurements and therefore consider cloud cover and relative humidity, predicted latent heat fluxes with remarkable accuracy. The results from predicting ET emphasized that EBMs benefit significantly from radiation measurements, as well as from knowledge on reflective and absorptive properties from surfaces. For one-source EBMs, such as DATTUTDUT, the minimal set of collected variables would therefore be the land surface temperatures (as collected with the thermal camera mounted on the drone), as well as local time, location, short-wave irradiance or, for the most accurate results, the net radiation as a budget of short- and long wave radiation. The more complex two-source EBMs (TSEB-PT and DTD) further require wind speed, relative humidity and air temperature and their respective measurement altitudes, as well as information on the canopy properties. Not all of these variables can simply be measured from a drone platform. This fact has also influenced the decision of which EBM to implement into the software described in Chapter 5.

The prediction models for sap flux were mainly based on canopy area, relative hu-

midity, the number of pixels and the air temperature. The importance of canopy area for transpiration estimates has already been shown (Ahongshangbam et al., 2020, 2019). However, canopy size cannot fully explain the good prediction results of the presented prediction models. ET is subject to strong variations during the course of the day, while canopy size remains steady, even during extreme conditions e.g. dryness or flooding. Therefore, canopy area might be an important indicator, but it is surely not the only variable that led to the observed accuracy in the predictions. This fact emphasizes the importance of auxiliary variables, such as relative humidity, air temperature and solar radiation, for precise estimates. Stomatal conductance has mainly been predicted by barometric pressure, canopy area, relative humidity, latent heat flux estimates and land surface temperature derivates. The influence of soil moisture has not been investigated in the scope of this study, since no significant influence on (evapo)transpiration of oil palm and trees in lowland Sumatra had been found in non-ENSO periods in previous studies (Röll et al., 2019; Stiegler et al., 2019). In other study regions, soil moisture is known to be an important driver of transpiration (Giambelluca et al., 2016; Kobayashi et al., 2014) and therefore a key variable in prediction models (Fernandes et al., 2016).

In conclusion, drones as recording tools for land surface temperatures, have shown many advantages over other methods. For studies on forestry hydrology, where the determination of water fluxes in the forest ecosystem and water use on stand level are of major importance, drones can cover just the right scale, considering resolution and coverage area. Compared to remote sensing from satellites, drones are able to record a high spatial resolution which allows to easily distinguish between canopies, while still covering considerably large areas. Drones are extremely mobile and versatile and the recording process is usually not constrained by terrain or remoteness of the studied area. For this dissertation, a multicopter drone was used to record land surface temperatures and RGB images in plantations and forest environments. The payload options, consisted mainly of thermal infrared sensors and RGB cameras in a stereo-camera setup. Moreover, a multispectral camera and an on-board sensor to record meteorological parameters were applied, but the data was not used in the scope of this dissertation. Chapter 3 and 5 showed, that EBMs can benefit significantly from additional measurements, especially incoming short-wave radiation and the net radiation budget of short- and long wave radiation. Sensors to measure these variables can easily be mounted on the drone, as several studies measuring similar variables have shown (Suomalainen et al., 2018, 2017). In the scope of this dissertation, a small set of on-board drone sensors has been developed and tested, but the recorded solar radiation data has not been used for the studies after all. Extending this concept, auxiliary modelling data can further be gathered by en-

vironmental sensor networks on the ground. These networks record and transmit the data, using e.g. the LoRaWAN protocol to directly communicate with the drone and to supply further variables, such as soil moisture, in-canopy air temperature and relative humidity to the modelling process. A big disadvantage of drones is, that their complex handling requires training and experience. But with progressing development of more autonomous systems, less human interaction will be required in the future. Fully automated drone-based survey flights were not used for the data acquisition in this dissertation yet, but they will surely be part of the frame work of hydrological studies in the coming years.

A fully automated data acquisition approach will not only increase the frequency of recording, but further enable a new branch of scientific studies and publications. A constant stream of near-real-time recorded data, similar to the reliability of satellite maps will be available. However, since drones are much more flexible carrier vehicles than satellites, the choice for payloads and locations can easily be adapted and temporal and spatial resolution can be increased dramatically. Since ML models usually boost their prediction quality when more data is available, they can continuously be improved using the available continuous stream of automated drone recorded data. These constantly evolving prediction models might also change the concept of scientific publications. Publications might therefore not be immutable if published, but rather frequently adapted to the latest findings based on the ever-evolving data base.

The urge for standardized and simple-to-use tools for ET prediction dates back even further, then the example from Symons (1867), mentioned in the introduction of this work. With the open-source software QWaterModel from Chapter 5, we introduced our own tool to enable a simple and easy-to-use workflow for EB-modelling of mass and energy fluxes from thermal images. The development of the software followed Occam's razor principle, for which the minimum of complexity needed to get decent ET prediction results, has been implemented. The combination of an easy access via download from the official QGIS3 Plugin repository, a simple and self-explanatory graphical user interface (GUI), and the code published, reviewable and adaptable on the biggest public code data base, resulted in a wide user base (up to the submission of this dissertation four versions of QWaterModel exist and more than 1000 downloads have been registered), ensuring a broad and standardized use and a future interactive development of the software.

The drone-based remote sensing applications presented in this dissertation and the modelling and prediction approaches entail a large potential for agriculture and forestry studies. As shown in Chapter 3, adding a spatially explicit map to the

eddy covariance footprints is a very useful supplement to understand flux distributions. Autonomous acquisition of LSTs over EC stations and the surrounding areas can be supplemented by on-board and ground sensors, and energy-balance models can be computed on the edge, enabling a dense temporal resolution of LST, flux and ET maps in almost real-time. This concept can e.g. be used for the attribution of fluxes in mixed species plant communities, the study of edge effects in landscapes, and can further be adapted e.g. to detect water stress in agriculture and forests. The very promising results from Chapter 4 support point measurements of sap flux and stomatal conductance and enable an extrapolation of the point measurements to large landscape sized areas with a solid spatially explicit representation. The results from both Chapter 3 and 4 can help to optimize irrigation, save water, prevent water stress, reduce crop fatality and increase yields in precision agriculture applications. Moreover, an urban application is the option to adjust irrigation and to detect plant diseases in urban green lands and parks. It is difficult to estimate the impact of an open-source software such as 'QWaterModel' on the spatial ecohydrology community. Since this software is the first of its kind in terms of free and open-source availability and user-friendliness in the QGIS3 environment, there is a small chance that it might actually set a standard for ET modelling using thermal images. However, it might also be replaced by a more accurate or specialized approach in the long run, if user-friendliness and availability of the new software can match or improve the simple GUI-based concept of 'QWaterModel'.

References

Ahongshangbam, J., Khokthong, W., Ellsäßer, F., Hendrayanto, H., Hölscher, D., Röll, A., 2019. Drone-based photogrammetry-derived crown metrics for predicting tree and oil palm water use. Ecohydrology 12. https://doi.org/10.1002/eco.2115

Ahongshangbam, J., Röll, A., Ellsäßer, F., Hendrayanto, Hölscher, D., 2020. Airborne tree crown detection for predicting spatial heterogeneity of canopy transpiration in a tropical rainforest. Remote Sens. 12, 651. https://doi.org/10.3390/rs12040651

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56. FAO, Rome.

Anderson, M.C., Kustas, W.P., Norman, J.M., Hain, C.R., Mecikalski, J.R., Schultz, L., González-Dugo, M.P., Cammalleri, C., d'Urso, G., Pimstein, A., Gao, F., 2011. Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. Hydrol. Earth Syst. Sci. 15, 223–239. https://doi.org/10.5194/hess-15-223-2011

Aubrecht, D.M., Helliker, B.R., Goulden, M.L., Roberts, D.A., Still, C.J., Richardson, A.D., 2016. Continuous, long-term, high-frequency thermal imaging of vegetation: Uncertainties and recommended best practices. Agric. For. Meteorol. 228–229, 315–326. https://doi.org/10.1016/j.agrformet.2016.07.017

Bach, H., Klug, P., Ruf, T., Migdall, S., Schlenz, F., Hank, T., Mauser, W., 2015. Satellite image simulations for model-supervised, dynamic retrieval of crop type and land use intensity. ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XL-7/W3, 1–7. https://doi.org/10.5194/isprsarchives-XL-7-W3-1-2015

Baena, S., Moat, J., Whaley, O., Boyd, D.S., 2017. Identifying species from the air: UAVs and the very high resolution challenge for plant conservation. PLOS ONE 12, e0188714. https://doi.org/10.1371/journal.pone.0188714

Bastiaanssen, W.G.M., Menenti, M., Feddes, R.A., Holtslag, A.A.M., 1998. A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. J. Hydrol. 212–213, 198–212. https://doi.org/10.1016/S0022-1694(98)00253-4

Berni, J.A.J., Zarco-Tejada, P.J., Sepulcre-Cantó, G., Fereres, E., Villalobos, F., 2009. Mapping canopy conductance and CWSI in olive orchards using high resolution thermal remote sensing imagery. Remote Sens. Environ. 113, 2380–2388. https://doi.org/10.1016/j.rse.2009.06.018

Berni, J.A.J., Zarco-Tejada, P.J., Suárez, L., González-Dugo, V., Fereres, E., 2008. Remote sensing of vegetation from UAV platforms using lightweight multispectral and thermal imaging sensors. Inter-Comm. WG IV 38, 6.

Bland, J.M., Altman, D.G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement International journal of nursing studies. 47. https://doi.org/10.1016/j.ijnurstu.2009.10.001

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324

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. Int. J. Remote Sens. 38, 3003–3026. https://doi.org/10.1080/01431161.2017.1280202

Brenner, C., Zeeman, M., Bernhardt, M., Schulz, K., 2018. Estimation of evapotranspiration of temperate grassland based on high-resolution thermal and visible range imagery from unmanned aerial systems. Int. J. Remote Sens. 39, 5141–5174. https://doi.org/10.1080/01431161.2018.1471550

Bzdok, D., 2017. Classical statistics and statistical learning in imaging neuroscience. Front. Neurosci. 11, 543. https://doi.org/10.3389/fnins.2017.00543

Bzdok, D., Altman, N., Krzywinski, M., 2018. Statistics versus machine learning. Nat. Methods 15, 233–234. https://doi.org/10.1038/nmeth.4642

Cornbleet, P.J., Gochman, N., 1979. Incorrect least-squares regression coefficients in method-comparison analysis. Clin. Chem. 432–438.

Delogu, E., Boulet, G., Olioso, A., Coudert, B., Chirouze, J., Ceschia, E., Le Dantec, V., Marloie, O., Chehbouni, G., Lagouarde, J.-P., 2012. Reconstruction of temporal variations of evapotranspiration us-

ing instantaneous estimates at the time of satellite overpass. Hydrol. Earth Syst. Sci. 16, 2995–3010. https://doi.org/10.5194/hess-16-2995-2012

Deming, W.E., 1964. Statistical adjustment of data, Dover Books on Mathematics Series. Dover Publications.

Dolman, A.J., 1988. Transpiration of an oak forest as predicted from porometer and weather data. J. Hydrol. 97, 225–234. https://doi.org/10.1016/0022-1694(88)90117-5

dos Reis, A.A., Carvalho, M.C., de Mello, J.M., Gomide, L.R., Ferraz Filho, A.C., Acerbi Junior, F.W., 2018. Spatial prediction of basal area and volume in Eucalyptus stands using Landsat TM data: an assessment of prediction methods. N. Z. J. For. Sci. 48, 1. https://doi.org/10.1186/s40490-017-0108-0

Fernandes, T., Campo, A., García-Bartual, R., González-Sanchis, M., 2016. Coupling daily transpiration modelling with forest management in a semiarid pine plantation. IForest - Biogeosciences For. 9, 38–48. https://doi.org/10.3832/ifor1290-008

Giambelluca, T.W., Mudd, R.G., Liu, W., Ziegler, A.D., Kobayashi, N., Kumagai, T., Miyazawa, Y., Lim, T.K., Huang, M., Fox, J., Yin, S., Mak, S.V., Kasemsap, P., 2016. Evapotranspiration of rubber (Hevea brasiliensis) cultivated at two plantation sites in Southeast Asia: Rubber evapotranspiration in SE Asia. Water Resour. Res. 52, 660–679. https://doi.org/10.1002/2015WR017755

Giavarina, D., 2015. Understanding Bland Altman analysis. Biochem. Medica 25, 141–151. https://doi.org/10.11613/BM.2015.015

Good, S.P., Noone, D., Bowen, G., 2015. Hydrologic connectivity constrains partitioning of global terrestrial water fluxes. Am. Assoc. Adv. Sci. 349, 175–177. https://doi.org/10.1126/science.aaa5931

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. Science 342, 850–853. https://doi.org/10.1126/science.1244693

Hinkley, D.V., Shi, S., 1989. Importance sampling and the nested bootstrap. Biometrika, 76, 3, 435–446. https://doi.org/10.1093/biomet/76.3.435

Hoffmann, H., Nieto, H., Jensen, R., Guzinski, R., Zarco-Tejada, P., Friborg, T., 2016. Estimating evaporation with thermal UAV data and two-source energy balance models. Hydrol. Earth Syst. Sci. 20, 697–713. https://doi.org/10.5194/hess-20-697-2016

Hunter, J.D., 2007. Matplotlib: A 2D graphics environment. Comput. Sci. Eng. 9, 90–95. https://doi.org/10.1109/MCSE.2007.55

Jasechko, S., Sharp, Z.D., Gibson, J.J., Birks, S.J., Yi, Y., Fawcett, P.J., 2013. Terrestrial water fluxes dominated by transpiration. Nature 496, 347–350. https://doi.org/10.1038/nature11983

Jones, H.G., 2014. Plants and microclimate: a quantitative approach to environmental plant physiology, Third edition. ed. Cambridge University Press, Cambridge; New York.

Jones, H.G., Serraj, R., Loveys, B.R., Xiong, L., Wheaton, A., Price, A.H., 2009. Thermal infrared imaging of crop canopies for the remote diagnosis and quantification of plant responses to water stress in the field. Funct. Plant Biol. 36, 978. https://doi.org/10.1071/FP09123

Jones, H.G., Vaughan, R.A., 2010. Remote sensing of vegetation: principles, techniques, and applications. Oxford University Press, Oxford; New York.

Kalma, J.D., McVicar, T.R., McCabe, M.F., 2008. Estimating land surface evaporation: a review of methods using remotely sensed surface temperature data. Surv. Geophys. 29, 421–469. https://doi.org/10.1007/s10712-008-9037-z

Kaushal, S., Gold, A., Mayer, P., 2017. Land use, climate, and water resources—global stages of interaction. Water 9, 815. https://doi.org/10.3390/w9100815

Kellner, J.R., Armston, J., Birrer, M., Cushman, K.C., Duncanson, L., Eck, C., Falleger, C., Imbach, B., Král, K., Krůček, M., Trochta, J., Vrška, T., Zgraggen, C., 2019. New opportunities for forest remote sensing through ultra-high-density drone lidar. Surv. Geophys. 40, 959–977. https://doi.org/10.1007/s10712-019-09529-9

Kobayashi, N., Kumagai, T., Miyazawa, Y., Matsumoto, K., Tateishi, M., Lim, T.K., Mudd, R.G., Ziegler, A.D., Giambelluca, T.W., Yin, S., 2014. Transpiration characteristics of a rubber plantation in central Cambodia. Tree Physiol. 34, 285–301. https://doi.org/10.1093/treephys/tpu009

Kolmogorov, A.N., Morrison, N., 2013. Foundations of the theory of probability. Martino Fine Books.

Kuenzer, C., Dech, S. (Eds.), 2013. Thermal infrared remote sensing: sensors, methods, applications, Remote sensing and digital image processing. Springer, Dordrecht.

Lechner, A.M., Foody, G.M., Boyd, D.S., 2020. Applications in remote sensing to forest ecology and management. One Earth 2, 405–412. https://doi.org/10.1016/j.oneear.2020.05.001

Legendre, P., Legendre, L., 2003. Numerical Ecology, 2/20. ed. Elsevier.

Leinonen, I., Grant, O.M., Tagliavia, C.P.P., Chaves, M.M., Jones, H.G., 2006. Estimating stomatal conductance with thermal imagery. Plant Cell Environ. 29, 1508–1518. https://doi.org/10.1111/j.1365-3040.2006.01528.x

Linnet, K., 1993. Evaluation of Regression Procedures for Method Comparison Studies. Clin Chem 424–432.

Luhmann, T., Piechel, J., Roelfs, T., 2013. Geometric Calibration of Thermographic Cameras, in: Kuenzer, C., Dech, S. (Eds.), Thermal infrared remote sensing. Springer Netherlands, Dordrecht, pp. 27–42. https://doi.org/10.1007/978-94-007-6639-6_2

Maes, W.H., Steppe, K., 2012. Estimating evapotranspiration and drought stress with ground-based thermal remote sensing in agriculture: a review. J. Exp. Bot. 63, 4671–4712. https://doi.org/10.1093/jxb/ers165

Margono, B.A., Potapov, P.V., Turubanova, S., Stolle, F., Hansen, M.C., 2014. Primary forest cover loss in Indonesia over 2000–2012. Nat. Clim. Change 4, 730–735. https://doi.org/10.1038/nclimate2277

McKinney, W., Wes and others 2010. Data structures for statistical computing in Python. Proceedings of the 9th Python in Science Conference, 445, 51–56.

Mei, T., Fang, D., Röll, A., Hölscher, D., 2019. Bamboo water transport assessed with deuterium tracing. Forests 10, 623. https://doi.org/10.3390/f10080623

Meijide, A., Badu, C.S., Moyano, F., Tiralla, N., Gunawan, D., Knohl, A., 2018. Impact of forest conversion to oil palm and rubber plantations on microclimate and the role of the 2015 ENSO event. Agric. For. Meteorol. 252, 208–219. https://doi.org/10.1016/j.agrformet.2018.01.013

Meijide, A., Röll, A., Fan, Y., Herbst, M., Niu, F., Tiedemann, F., June, T., Rauf, A., Hölscher, D., Knohl, A., 2017. Controls of water and energy fluxes in oil palm plantations: Environmental variables and oil palm age. Agric. For. Meteorol. 239, 71–85. https://doi.org/10.1016/j.agrformet.2017.02.034

Messina, G., Modica, G., 2020. Applications of UAV thermal imagery in precision agriculture: State of the art and future research outlook. Remote Sens. 12, 1491. https://doi.org/10.3390/rs12091491

Mohan, M., Silva, C., Klauberg, C., Jat, P., Catts, G., Cardil, A., Hudak, A., Dia, M., 2017. Individual tree detection from Unmanned Aerial Vehicle (UAV) derived canopy height model in an open canopy mixed conifer forest. Forests 8, 340. https://doi.org/10.3390/f8090340

Moulianitis, V. C., Thanellas, G., Xanthopoulos, N., Aspragathos, N.A., 2019. Evaluation of UAV based schemes for forest fire monitoring. Aspragathos, N.A., Koustoumpardis, P.N., Moulianitis, Vassilis C. (Eds.), Advances in Service and Industrial Robotics. Springer International Publishing, Cham, pp. 143–150. https://doi.org/10.1007/978-3-030-00232-9_15

Niu, F., Röll, A., Meijide, A., Hendrayanto, Hölscher, D., 2017. Rubber tree transpiration in the lowlands of Sumatra. Ecohydrology 10, e1882. https://doi.org/10.1002/eco.1882

Norman, J.M., Kustas, W.P., Humes, K.S., 1995. Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. Agric. For. Meteorol. 77, 263–293. https://doi.org/10.1016/0168-1923(95)02265-Y

Norman, J.M., Kustas, W.P., Prueger, J.H., Diak, G.R., 2000. Surface flux estimation using radiometric temperature: A dual-temperature-difference method to minimize measurement errors. Water Resour. Res. 36, 2263–2274. https://doi.org/10.1029/2000WR900033

Oki, T., Kanae, S., 2006. Global hydrological cycles and world water resources. Am. Assoc. Adv. Sci. 313, 1068–1072. https://doi.org/10.1126/science.1128845

Oliphant, T.E., 2006. A guide to NumPy. Trelgol Publ.

Passing, H., Bablok, W., 1984. Comparison of several regression procedures for method comparison studies and determination of sample sizes application of linear regression procedures for method comparison studies in clinical chemistry, Part II. Clin. Chem. Lab. Med. 22. https://doi.org/10.1515/cclm.1984.22.6.431

Passing, H., Bablok, W., 1983. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. Clin. Chem. Lab. Med. 21. https://doi.org/10.1515/cclm.1983.21.11.709

Planck, M., 1900. Entropie und Temperatur strahlender Wärme. Ann. Phys. 306, 719–737. https://doi.org/10.1002/andp.19003060410

Prudhomme, C., Giuntoli, I., Robinson, E.L., Clark, D.B., Arnell, N.W., Dankers, R., Fekete, B.M., Franssen, W., Gerten, D., Gosling, S.N., Hagemann, S., Hannah, D.M., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y., Wisser, D., 2014. Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. Proc. Natl. Acad. Sci. 111, 3262–3267. https://doi.org/10.1073/pnas.1222473110

Raschke, K., 1956. Über die physikalischen Beziehungen zwischen Wärmeübergangszahl, Strahlungsaustausch, Temperatur und Transpiration eines Blattes. Planta 48, 200–238.

Raymond, E., 1999. The cathedral and the bazaar. Knowl. Technol. Policy, 12, 23–49. https://doi.org/10.1007/s12130-999-1026-0

Röll, A., Niu, F., Meijide, A., Ahongshangbam, J., Ehbrecht, M., Guillaume, T., Gunawan, D., Hardanto, A., Hendrayanto, Hertel, D., Kotowska, M.M., Kreft, H., Kuzyakov, Y., Leuschner, C., Nomura, M., Polle, A., Rembold, K., Sahner, J., Seidel, D., Zemp, D.C., Knohl, A., Hölscher, D., 2019. Transpiration on the rebound in lowland Sumatra. Agric. For. Meteorol. 274, 160–171. https://doi.org/10.1016/j.agrformet.2019.04.017

Röll, A., Niu, F., Meijide, A., Hardanto, A., Hendrayanto, Knohl, A., Hölscher, D., 2015. Transpiration in an oil palm landscape: effects of palm age. Biogeosciences 12, 5619–5633. https://doi.org/10.5194/bg-12-5619-2015

Sabins, Lulla, K., 1987. Remote sensing: Principles and interpretation. Geocarto Int. 2, 66–66. https://doi.org/10.1080/10106048709354087

Siebert, S., Ewert, F., Eyshi Rezaei, E., Kage, H., Graß, R., 2014. Impact of heat stress on crop yield—on the importance of considering canopy temperature. Environ. Res. Lett. 9, 044012. https://doi.org/10.1088/1748-9326/9/4/044012

Simundic, A.M., 2016. Statistical analysis in method comparison studies – Part one. Retrieved from URL: https://acutecaretesting.org/en/articles/statistical-analysis-in-method-comparison-studies-part-one , (accessed 02.06.2020)

Sobrino, J.A., Julien, Y., 2013. Time series corrections and analyses in thermal remote sensing, in: Kuenzer, C., Dech, S. (Eds.), Thermal Infrared Remote Sensing. Springer Netherlands, Dordrecht, pp. 267–285. https://doi.org/10.1007/978-94-007-6639-6_14

Stiegler, C., Meijide, A., Fan, Y., Ashween Ali, A., June, T., Knohl, A., 2019. El Niño–Southern Oscillation (ENSO) event reduces CO 2 uptake of an Indonesian oil palm plantation. Biogeosciences 16, 2873–2890. https://doi.org/10.5194/bg-16-2873-2019

Still, C., Powell, R., Aubrecht, D., Kim, Y., Helliker, B., Roberts, D., Richardson, A.D., Goulden, M., 2019. Thermal imaging in plant and ecosystem ecology: applications and challenges. Ecosphere 10, e02768. https://doi.org/10.1002/ecs2.2768

Sullivan, D.G., Fulton, J.P., Shaw, J.N., Bland, G., 2007. Evaluating the sensitivity of an unmanned thermal infrared aerial system to detect water stress in a cotton canopy. Trans. ASABE 50, 1963–1969. https://doi.org/10.13031/2013.24091

Suomalainen, J., Hakala, T., Alves de Oliveira, R., Markelin, L., Viljanen, N., Näsi, R., Honkavaara, E., 2018. A novel tilt correction technique for irradiance sensors and spectrometers on-board unmanned aerial vehicles. Remote Sens. 10, 2068. https://doi.org/10.3390/rs10122068

Suomalainen, J., Hakala, T., Honkavaara, E., 2017. Measuring incident irradiance on-board an unstable UAV platform - first results on virtual horizontalization of multiangle measurement. Online Proc. Int. Conf. Unmanned Aer. Veh. Geomat. 2.

Symons, G.J., 1867. Evaporators and evaporation. British Rainfall 7, 9–10.

Teuling, A.J., 2018. A forest evapotranspiration paradox investigated using lysimeter data. Vadose Zone J. 17, 170031. https://doi.org/10.2136/vzj2017.01.0031

Timmermans, W.J., Kustas, W.P., Andreu, A., 2015. Utility of an automated thermal-based approach for monitoring evapotranspiration. Acta Geophys. 63, 1571–1608. https://doi.org/10.1515/acgeo-2015-0016

van der Walt, S., Colbert, S.C., Varoquaux, G., 2011. The NumPy array: a structure for efficient numerical computation. Comput. Sci. Eng. 13, 22–30. https://doi.org/10.1109/MCSE.2011.37

Vasterling, M., Meyer, U., 2013. Challenges and opportunities for UAV-borne thermal imaging, in: Kuenzer, C., Dech, S. (Eds.), Thermal Infrared Remote Sensing. Springer Netherlands, Dordrecht, pp. 69–92. https://doi.org/10.1007/978-94-007-6639-6_4

Verstraeten, W., Veroustraete, F., Feyen, J., 2008. Assessment of evapotranspiration and soil moisture content across different scales of observation. Sensors 8, 70–117. https://doi.org/10.3390/s8010070

Waite, P.-A., Schuldt, B., Mathias Link, R., Breidenbach, N., Triadiati, T., Hennings, N., Saad, A., Leuschner, C., 2019. Soil moisture regime and palm height influence embolism resistance in oil palm. Tree Physiol. 39, 1696–1712. https://doi.org/10.1093/treephys/tpz061

Waskom, M., Botvinnik, O., Ostblom, J., 2020. mwaskom/seaborn: seaborn v0.8.1, Zenodo, https://doi.org/10.5281/zenodo.883859

Wasserstein, R.L., Lazar, N.A., 2016. The ASA Statement on p -Values: Context, Process, and Purpose. Am. Stat. 70, 129–133. https://doi.org/10.1080/00031305.2016.1154108

Xia, T., Kustas, W.P., Anderson, M.C., Alfieri, J.G., Gao, F., McKee, L., Prueger, J.H., Geli, H.M.E., Neale, C.M.U., Sanchez, L., Alsina, M.M., Wang, Z., 2016. Mapping evapotranspiration with high-resolution aircraft imagery over vineyards using one- and two-source modeling schemes. Hydrol. Earth Syst. Sci. 20, 1523–1545. https://doi.org/10.5194/hess-20-1523-2016

Acknowledgements

Completing this work would have never been possible without the help and support of so many great people. In the following, I want to express my gratitude towards everybody who supported, encouraged, assisted and helped me on the long road of completing this dissertation. Some people are able to express their gratitude in a short and concise way. However, I'm not one of them:

First of all, I want to thank **Dirk Hölscher** who continuously guided me through all ups and downs of this four year long journey until the completion of this work. Thank you so much for your endless trust and optimism, your encouragement and your support in all the minor and major things that now resulted in this thesis and a really creative, nice and productive time. Further thanks goes to **Alexander Knohl** and **Holger Kreft** for agreeing to be on the committee and for the fruitful collaboration and many helpful discussions and critical comments.

A big thank you also to **Eva Siegelkow** for all her efforts to help me with the administrative tasks and continuously supporting me with establishing a circular cake themed π-day in the institute.

Huge thanks goes to **Krissi** and **Joyson** for being my office buddies and really, really good friends. Thank you Krissi for so many great discussions and conversations and for never getting tired to explain me the difference between spruces and firs. Thank you Joyson for the motorbike lessons and an exceptional time in Jambi and for always providing me with awesome egg-curry and empanadas. Further thanks goes to **Alex R.** for barbecuing during field work, for many helpful suggestions and for never getting tired to read and revise my manuscripts.

I would further like to heartily thank all my dear friends and colleagues from the CRC 990s second phase for all the good times in Jambi and in Göttingen: **Pierre** and **Greta** for countless evenings with (too many) good drinks and nice conversations, **Dirk B.** for never getting tired to explain me whether 'melodic speed pirate metal' exists, **Johannes B.** for building the fanciest 80ies roof-top

bar in Jambi with me, **Fenna** for tea on her Ostfriesensofa, **Carina** for painting flowers on things, **Clara** for adopting every cat she could find, **Christian** for all the Eddy knowledge and being a great co-author, **Tati** and **Rhami** for the best karaoke singing, **Isa** for countless invitations to nice parties, **Joost** for keeping these parties alive, **Katrin** for Rhabarberkuchen, **Watit** for the fancy Jambi lifestyle and a further thank you goes to **Britta**, **Jenny** and **Kevin** and all the others :).

I am grateful to the **DFG** for funding the EFForTS project and making all this possible. Special thanks also goes to my assistants **Ozy** and **Zulfi** for the huge support during the field work, and all the great times in Jambi and on the road.

Surely no thanksgivings would be complete without explicitly thanking my non-project friends too. You make the universe a shiny place:

A huge thank you goes to **Hanni** and **Schlodde** for all the good times, for visits to wherever anyone of us was roaming, for surviving unexpected snowstorms and for finishing the biggest breakfasts, for the boundless love for Pepsi and goats, for countless hours on the phone and for being the greatest friends I could ever have wished for. A very special thanks is further reserved for that other **Flo** for self-accommodating on my couch 11 years ago and being a great and very supportive friend since then. I further owe buckets of gratitude to **Roli** for snowball fights in Thessaloniki and for constant technical and emotional support.
Special thanks goes to **Mirka** and **Johannes** for always inviting me to their garden after work. If you ever star in *Zu Tisch in Göttingen* I'll provide you with a free fish.
I am especially grateful to my flat mates who really made Göttingen a home! Thank you **Jan**, **Lisa**, **Leonie**, **Maria**, **Lotta** and **Detlef** for countless nice moments, memories and more than four great years in Göttingen.

Most importantly, my greatest thanks goes to my family for supporting me during all those years. To my sister **Sophie** for her everlasting patience explaining things, for countless holidays and for encouragement and for cheering me up when I needed it. To my parents **Syl** and **Volle** who never doubted that I would finish this thesis and who supported me during my whole life, to my grandparents **Mädi** and **Pa** and to my aunt and uncle **Andrea** and **Ande**.

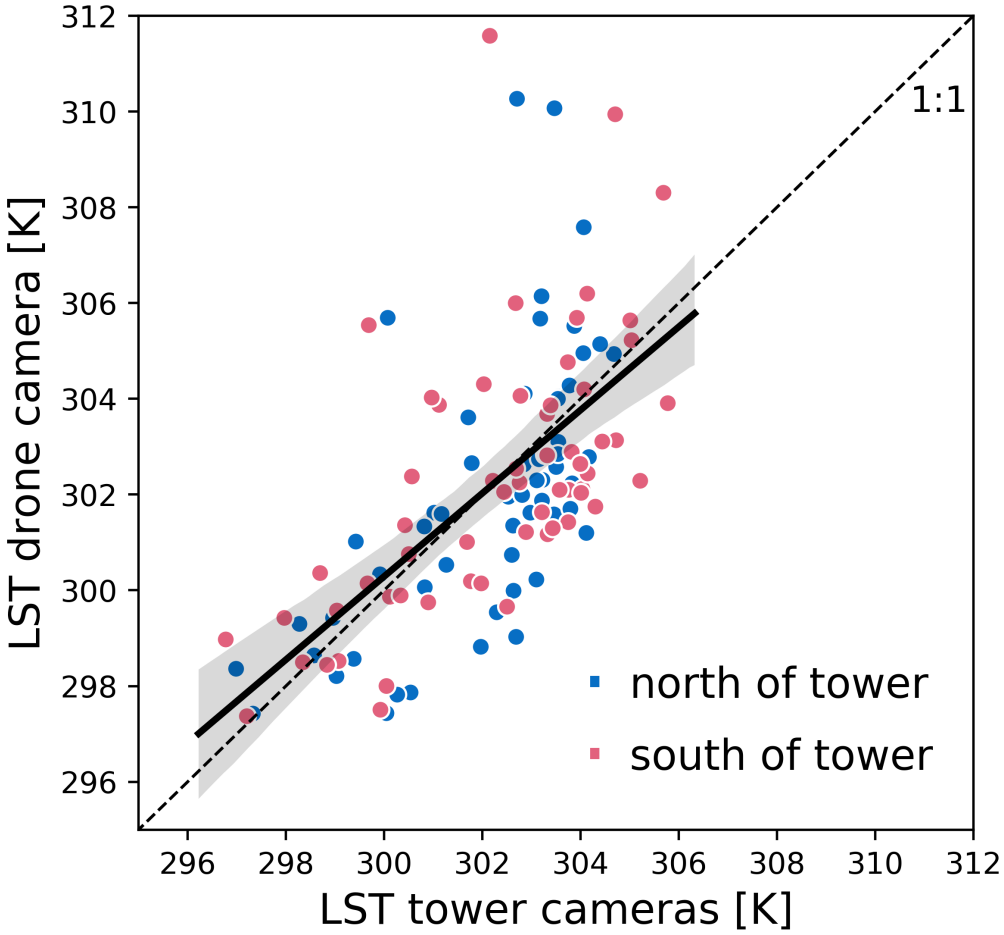
Appendix

Appendix I

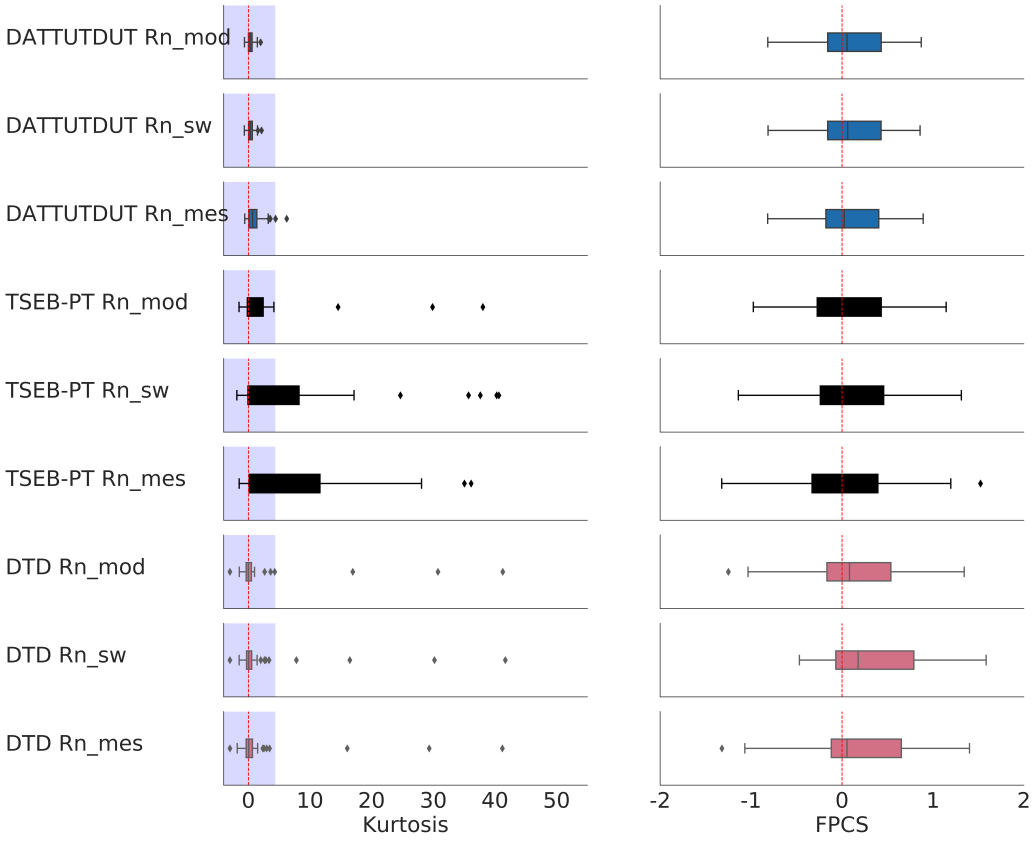
Table A1: All flight missions (n = 61) indicated with their flight registration numbers (176–236) conducted from DOY 217 to DOY 221 in 2017. (c) indicates cloudy or partly cloudy skies as observed on site.

Timetable PTPN VI Flights (176–236)

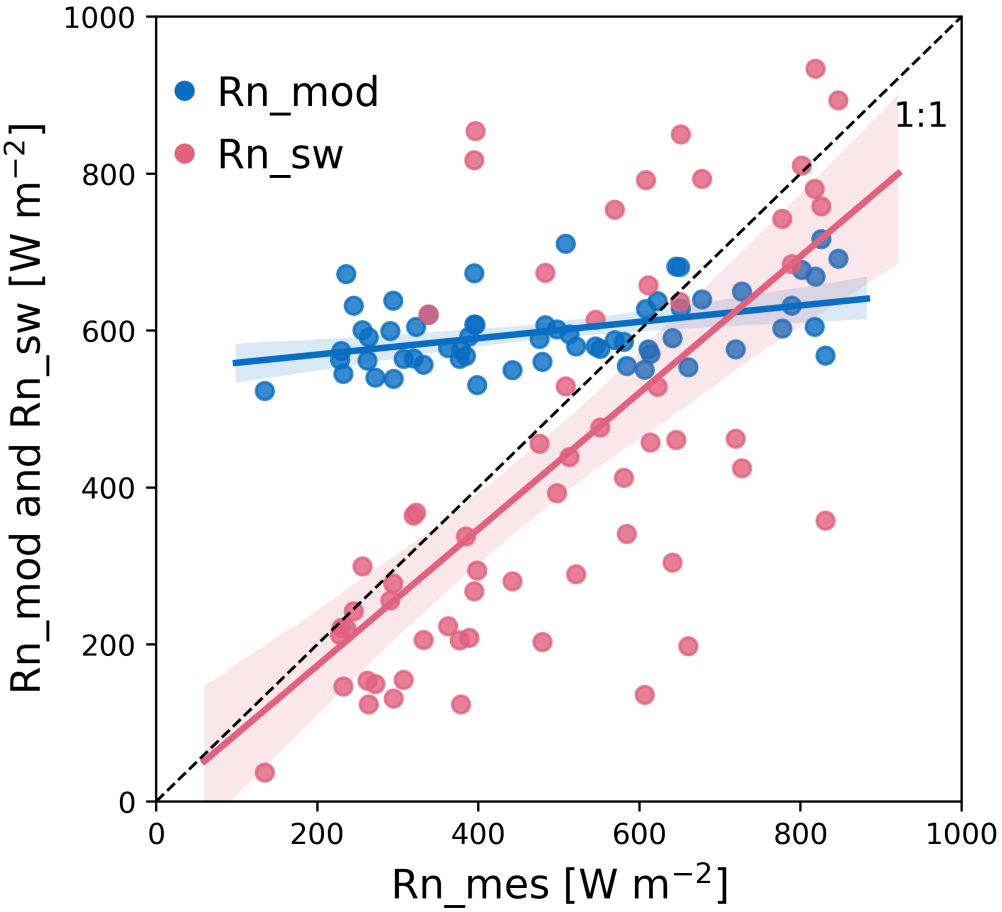
local time	217	218	219	220	221
9:00	-	(c) 189	(c) 200	(c) 213	(c) 223
9:30	-	(c) 190	(c) 201	(c) 214	(c) 224
10:00	176	(c) 191	(c) 202	(c) 215	(c) 225
10:30	177	192	203	(c) 216	(c) 226
11:00	178	193	(c) 204	(c) 217	227
11:30	(c) 179	194	205	218	228
12:00	(c) 180	195	206	219	(c) 229
12:30	181	(c) 196	207	220	230
13:00	182	197	208	221	231
13:30	183	198	(c) 209	(c) 222	(c) 232
14:00	184	199	(c) 210	-	(c) 233
14:30	-	-	(c) 211	-	(c) 234
15:00	(c) 185	-	(c) 212	-	(c) 235
15:30	(c) 186	-	-	-	(c) 236
16:00	(c) 187	-	-	-	-
16:30	(c) 188	-	-	-	-



Appendix Figure A 1: *Scatterplot and linear regression of LSTs obtained with the drone and tower cameras. Mean absolute error (MAE) is 1.59 K and root mean squared error (RMSE) is 2.15 K, relationship of both measurements is* $r^2 = 0.4$.



Appendix Figure A 2: *Boxplot distribution of histogram Kurtosis and Fisher-Pearson Coefficient of Skewness (FPCS) for the different model settings, each with n=61. Zero for normal distribution is indicated with a red line, the blue area indicates platykurtic distribution.*



Appendix Figure A 3: *Measured net radiation (Rn_mes) plotted against fully modeled net radiation (Rn_mod) and net radiation estimates based on short-short wave irradiance (Rn_sw).*

Appendix II

Method appendix: model descriptions

All models in this study use instantaneous land surface temperatures (LST) to solve the energy balance equation:

$$Rn = G + H + LE \tag{1}$$

Where Rn is the net radiation, G is the ground heat flux and the turbulent fluxes H and LE represent sensible and latent heat flux, respectively. Rn is estimated by calculating the budget of incoming (↓) and outgoing (↑) long- (l) and short-wave (s) radiation:

$$Rn = R_s^{\downarrow} + R_s^{\uparrow} + R_l^{\uparrow} + R_l^{\downarrow} = (1-\alpha) * R_s^{\downarrow} + \varepsilon_{surf} * \varepsilon_{atm} * \sigma * T_{air}^4 - \varepsilon_{surf} * \sigma * T(\theta)_{surf}^4 \tag{2}$$

where the short-wave component is calculated by multiplying incoming shortwave radiation Rs↓ [W m^{-2}] with its absorption ratio deducted from the combined soil and vegetation albedo α. The long-wave radiation budget is calculated from surface (soil and vegetation) emissivity ε_{surf} and atmospheric emissivity ε_{atm}, the Stefan-Boltzmann constant σ (5.6704*10^{-8} W m^{-2}*K^{-4}), air temperature T_{air} and radiometric land surface temperature $T(\theta)_{surf}$ (both in K). The available energy consisting of the turbulent fluxes H and LE is calculated by subtracting G from Rn. G is computed as a linear function of Rn similar as described in Liebethal and Foken, (2007):

$$G = a * R_n - b \tag{3}$$

We set a = 0.1 for ground heat flux under canopies and b = 0 (Ogée et al., 2001). A time offset ΔRn as in the original formulation of Liebethal and Foken, (2007) is not included for simplicity reasons. With the fraction of turbulent fluxes (Rn – G / Rn) known, radiometric LSTs are used to calculate H and obtain LE as a residual.

DATTUTDUT

Key input for the DATTUTDUT model is a LST map from where the hottest and the 0.5% of coldest pixels are extracted, assuming that hot pixels are a result of

very little to no evapotranspiration and cold pixels origin in a high evapotranspiration rate (Timmermans et al., 2015). Fully modeled Rn is calculated based on down-welling short-wave radiation estimates calculated using sun-earth geometry to solve eq. 2. Surface albedo P0 is calculated as in Timmermans et al. (2015) based on the assumption that dense vegetation appears colder than rocks or soil in the thermal imagery (Brutsaert, 1982; Garratt, 1992):

$$P_0 = 0.05 + ((T_0 - T_{min})/(T_{max} - T_{min})) * 0.2 \quad (4)$$

Downwelling shortwave radiation $R_s \downarrow$ is calculated from the dimensionless atmospheric transmissivity τ and the exo-atmospheric shortwave radiation SWexo = 1360 W m^{-2} (Timmermans et al., 2015). Transmissivity τ is calculated as described in Burridge and Gadd (1977) using the solar elevation angle α that was determined from the geographic position of our site and the coordinated universal time (UTC) of the measurements:

$$\tau = 0.6 + 0.2 * sin(\alpha) \quad (5)$$

$$R_s \downarrow = \tau * SW_{exo} \quad (6)$$

Timmermans et al. (2015) suggest using a constant value of 0.7 for τ and 0.8 atmospheric emissivity (ε_{atm}), but as our flight times range from 09:00 to 16:30 h local time we decided to include the solar elevation angle as in eq. 5. Further, we used a constant surface emissivity (ε_{surf}) of 0.98 and not 1.0 as in the original formulation of the DATTUTDUT model. Air temperature T_{air} was calculated as mean temperature from 0.5% of the coldest pixels in the image.
Calculation of Rn using short wave irradiance or measured Rn for DATTUTDUT:

As the original DATTUTDUT formulation doesn't account for cloud cover, eq. 6 is replaced by measured short-wave irradiance as in Brenner et al. (2018) for model runs with Rn_sw. For model runs with Rn_mes eq. 2 was replaced by Rn measurements recorded at the EC-tower.

Calculation of turbulent fluxes in DATTUTDUT:

The sum of the turbulent fluxes is calculated by subtracting G from Rn. The result is fractioned into its components H and LE, using the evaporative fraction (EF) (Timmermans et al., 2015):

$$EF = LE/(LE+H) = LE/(Rn-G) = (T_{max}-T(\theta)_{surf})/(T_{max}-T_{min}) \quad (7)$$

TSEB-PT

The TSEB-PT model divides eq. 1 into a canopy and a soil fraction (Kustas and Norman, 1999; Song et al., 2016; Xia et al., 2016). The model consists of two parts: First an initialization part where all parameters that do not depend on soil and canopy temperature partition and knowledge of atmospheric stability are computed. Afterwards an iterative part where the Monin-Obukhov length is stabilized and the fluxes are finally derived. To begin this process vegetation cover $f_c(\theta)$ is computed as in Campbell and Norman, (1998):

$$f_c(\theta) = 1 - exp((-0.5\omega(\theta) * LAI)/(cos(\theta))) \quad (8)$$

where LAI is leaf area index, θ is the sun zenith angle and ω is a nadir view clumping factor to represent the cross-row structure in which the oil palm is planted (Kustas and Norman, 1999). Guzinski et al., (2014) suggest a maximum limit of 0.95 for $f_c(\theta)$, so that a small fraction of the soil is still visible and extreme magnitudes for soil temperature are avoided. Roughness parameters are calculated from vegetation height. Eq. 2 is used for the original calculation of Rn from short-wave irradiance for TSEB-PT. Tair was measured at the EC-tower, T(θ)surf was recorded with the drone both similar to descriptions in Hoffmann et al. (2016). The canopy emissivity E_{leaf} was set to 0.98 and soil emissivity E_{soil} to 0.95. The three resistances in the soil-canopy-atmosphere heat flux network, the aerodynamic resistance to heat transport (R_A), the resistance to heat transport from the soil surface (R_S) and the total boundary layer resistance of the leaf canopy (R_X) are calculated as in Norman et al. (2000, 1995). Net radiation and the three resistances remain constant during the model runs. After finishing the computation of all constant parameters, the iterative part of the model starts assuming Monin-Obukhov length tends to infinity. In the first iteration Rn is partitioned into a soil and canopy fraction by calculating net radiation divergence ΔRn (Hoffmann et al., 2016; Norman et al., 2000):

$$\Delta Rn = Rn * (1 - exp((-K * LAI * \omega_0)/\sqrt{((2cos(\theta_s))))}) \quad (9)$$

where K is an extinction coefficient that varies according to LAI (Hoffmann et al., 2016). We are aware of the fact, that the determination of K using LAI is disputed as other studies found no significant correlation of K and LAI (Zhang

et al., 2014). With ΔRn known, sensible heat flux is then estimated using the Priestley-Taylor approximation:

$$H_c = \Delta Rn * (1 - \alpha_{PT} * f_G * (D/(D + \gamma)) \quad (10)$$

Where αPT is the Priestley-Taylor coefficient that was calculated as described in Lhomme (1997). The psychrometric constant γ and the slope of the saturation pressure curve D were both calculated as in Allen et al. (1998). Canopy temperature T_C was computed by summing up the results of the linear approximation in equation (A7) for $T_{C,lin}$ and ΔT_C from equation (A11) both from Norman et al. (1995). Knowing canopy temperature T_C and fraction of view covered by vegetation fθ as in Hoffmann et al. (2016), soil temperature T_S can be calculated:

$$T_s = (T(\theta)R^4 - f_\theta * T_C^4)/(1 - f_\theta)^{(1/4)} \quad (11)$$

Now with soil and canopy temperatures and the resistances of the soil-canopy-atmosphere heat flux network known, fluxes can be calculated with equations (9), (10), (11) and (13) from Hoffmann et al. (2016). Soil heat flux G is computed as in equation (3). Total latent and sensible heat fluxes are calculated as the sums of canopy and soil fluxes. In the following iterations, a recalculation of Monin-Obukhov length takes place until a stable value is reached and the resulting fluxes are derived.

Turbulent fluxes are calculated and the original EF is also used to estimate LE from fully modeled and measured Rn. For model runs with fully modelled Rn, Rn is calculated as in DATTUTDUT using equation (2), for model runs with measured Rn, Rn is measured at the EC-tower.

DTD

The Dual-Temperature-Difference (DTD) model works very similar to TSEB-PT and differs mainly in the way how sensible heat flux is calculated (Hoffmann et al., 2016). A detailed description of the model can be found in Guzinski et al. (2014) and Norman et al. (2000). It uses two observations of LST and air temperature in the same day. The first observation is recorded in the early morning hours, when fluxes are known to be minimal and the second observation is recorded later on the same day at any given time. To account for inaccuracies and bias in the measurements a double-difference of LSTs and air temperatures is calculated to avoid the direct use of absolute LST data. Equation (A37) from Guzinski et al. (2014) describes how sensible heat flux is calculated using a series

resistance network instead of the parallel from the original model formulation in Norman et al. (1995).

Calculation of evapotranspirated amount of water:

The actual amount of evapotranspirated water (ET_w) in mm h^{-1} was calculated as in Timmermans et al. (2015):

$$ET_w = ((LE * t)/1000000)/(2.501 - 0.002361 * (T_{air} - 273.15)) \quad (12)$$

Where LE is the latent heat flux in W m^{-2}, t is the respective timespan in seconds and T_{air} is the air temperature in Kelvin.