UAV Applications for Intensive Forest Monitoring

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Stuart Krause

aus Nanaimo, Kanada

Bonn, Februar 2023

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Klaus Greve

2. Gutachter: Prof. Dr. Zbynek Malenovsky

Tag der Promotion: 22.11.2023

Erscheinungsjahr: 2024

Acknowledgements

This work would not have been possible without the assistance from various individuals and institutions. First of all, I would like to thank my supervisor Prof. Dr. Klaus Greve for his valuable guidance and positive encouragement throughout my PhD journey. I also would like to express my sincerest gratitude to Dr. Tanja Sanders for her unending support and counsel through thick and thin, and providing me the opportunity to follow my research goals.

Furthermore, I am grateful to all of my colleagues at the Thünen-Institut of Forest Ecosystems for the acquired friendships and collaborations especially within the remote sensing work group. I would like to thank Prof. Dr. Andreas Bolte for his motivational leadership with regard to current forest research topics as well as the team at the Britz Research Station including Dr. Jürgen Müller. I am also very appreciative to Dr. Inken Krüger for reviewing publications and offering valuable advice with respect to intensive forest monitoring. Also, I am indebted to Dietmar Fenske, Kinga Jánosi, and Jens Wiesehahn for their assistance in the field and Karsten Dünger for his support with providing access to the high-performance computing cluster.

Additionally, I would like to thank Prof. Dr. Jan-Peter Mund and the Eberswalde University of Sustainable Development (HNEE) for the invaluable opportunity to explore forest-related geospatial technologies as well as for the support during the initial stages of this work. Thanks, is also due to Dr. Stephan Raspe from the LWF for sharing his knowledge on phenology and ICP Forests.

Last but not least, I would like to thank my parents Kathryn and Helmar Krause for their support and understanding throughout my career change as well as my brother Erick for his unlimited positive feedback. Finally, I am thankful to my dog Fred for his companionship during field work and his constant reminders to leave the computer and go outdoors.

Abstract

Intensive forest monitoring at Level II plots facilitate the acquisition of Europe-wide information on forest health, diversity and productivity. The implementation of such a comprehensive monitoring program requires a considerable amount of uniformity and standardization in terms of how forest parameters are observed and measured at monitoring sites. The harmonization of such data at the European level allows for potential forest assessments that cover large areas not only limited to regional and national scales. Additionally, the use of intensive monitoring plots as ground-truthing for the upscaling to open access remote sensing (RS) Earth Observation (EO) platforms (i.e. Sentinels, Landsat) is of significant interest. The translation of data acquired at monitoring sites is however challenging, and site coverage is typically insufficient in the representation of the ever-increasing diverse forests in terms of structure, species configuration and management practices. Furthermore, intensive monitoring programs are time-consuming and costly, especially in terms of the systematic field campaigns required to acquire uniform data.

Thanks to recent developments in RS technology, in particular Unmanned/Unoccupied Aerial Vehicle-RS (UAV-RS), we have the opportunity to acquire primary close-range RS data at a very high temporal and spatial resolution with a wide-range of active and passive sensors. With UAV-RS emerges the prospect of acquiring typical forest measurements such as tree height and crown diameter, as well as the enhancement of phenological observations and tree condition assessments by providing an aerial perspective complimenting the view from the ground. Additionally, individual tree mapping with a high absolute geolocation accuracy can potentially solve the problems of mapping sites under dense canopies which is typically a challenge with geodetic instruments from the ground-level. In some cases, UAV-RS can not only enhance ground observations but in come case replace them, notably with reference to tree height and crown measurements. Furthermore, specific measurements and observations at intensive monitoring plots can be extended with targeted external plots for the purpose of acquiring additional forest types representative of a particular region.

This dissertation explores the possibility to enhance, or in some cases replace traditional ground observations and measurements at intensive forest monitoring plots using UAV-RS techniques. In particular, three main research questions will be explored. The first being with regard to UAV-

based photogrammetric tree height extraction and whether it is comparable to traditional methods. The second being whether UAV thermal imaging can accurately detect drought stress in Beech. And the third research question being whether UAV-derived phenological machine learning (ML) models are accurate enough to replace ground observations.

In the first investigation, photogrammetric tree heights (*Pinus sylvestris*) were compared to field measurements with a hypsometer. Both methods were validated using the unique opportunity of destructive techniques where selected trees were felled for thinning purposes, and the tree length measured horizontally along ground. Here it was shown that traditional tree height measurements tend to over-estimate with a Root-Mean-Squared-Error (RMSE) of 0.304 (1.82 %) and photogrammetric tree heights are inclined to underestimate tree heights with a RMSE of 0.34 m (2.07 %) (n = 34). These findings are significant, as the over-estimation of field-based tree height measurement could cause uncertainty when used as a validation method of photogrammetric or Laser Scanning tree height extraction methods.

The second study involves an exploration into the acquisition of UAV-based thermal imaging with the aim of tree drought stress detection quantification and Tree Water Deficit (TWD) modelling. The initial stages of the study involved indoor experiments with two plants involving drought-stress treatments validated with leaf temperature sensors. It was discovered that thermal pixel extraction using less than three pixels can result in error of up to 1 K whereas increasing the "spot size" can reduce the mean difference in error to 0.02 K. Additionally, it was discovered that drought-induced leaves, due to a lower emissivity can also propagate thermal imaging error. During field trials, validated with tree crown-mounted leaf temperature sensors, it was found that imagery acquired from gridded flights produced higher standard deviations due to an increase in distance and angle to validation trees. Thermal imagery acquired from close-range hovering resulted in lower standard deviations. With respect to the machine learning modeling of TWD, training and validation data was acquired using point dendrometers alongside input features derived from meteorological data and close-range individual tree crown thermal imaging. The best results were achieved with a Generalised Additive Model (GAM) and relative humidity, air temperature and tree canopy temperature as input features with a RMSE of 4.92 (µm) and R^2 of 0.87. This study presents a promising methodology in quantifying drought stress in European beech (*Fagus sylvatica*) with thermal imaging and meteorological data with an emphasis on reducing UAV-based thermography error propagation.

The third study, investigates the prediction of high-resolution spring phenological phase data for European beech with the use of UAV-based multispectral indices and machine learning. Using field observation data acquired from 2019 - 2021 as training and validation, various machine learning algorithms were tested with selected vegetation indices derived from multispectral and RGB sensors. Training datasets were sorted into various partitions during the feature selection process in order to detect year specific observation error which could affect modeling accuracy. The final GAM boosting model using the Green Chromatic Coordinate (GCC) vegetation index derived from the RGB sensors, was capable in predicting spring phenological phases on unseen datasets within a RMSE threshold of 0.5. Considerable limitations were however discovered with indices implementing the near-infrared (NIR) band due to oversaturation and an improved field calibration is necessary.

It can be concluded that the application of UAV-RS for intensive forest monitoring purposes will most probably become an integral part of monitoring programs in the near future, however the challenges presented in this dissertation in terms of uniform radiometric calibration, thermal imaging error propagation and the translation of ground observations to RS data will however require extensive exploration and experimentation. A coordinated effort amongst research institutions is recommended, alongside the establishment of a governing body to develop and maintain standardization goals for UAV-based intensive forest monitoring.

Zusammenfassung

Das intensive forstliche Monitoring (Level II) ermöglicht den Zugang zu europaweiten Informationen über Baumvitalität, Baumartendiversität und -produktivität. Die Umsetzung eines solch umfassenden Monitoringprogrammes erfordert ein beträchtliches Maß an Einheitlichkeit und Standardisierung in Bezug darauf, wie Waldparameter an Überwachungsstellen beobachtet und gemessen werden. Die Harmonisierung solcher Daten auf europäischer Ebene ermöglicht es potenziell großflächige Waldflächen zu beurteilen und hebt die Beschränkung von regionalen und nationalen Maßstäben auf. Darüber hinaus ist die Nutzung intensiver Monitoringflächen als verlässliche Bodeninformation für die Hochskalierung auf Open-Access-Fernerkundung-Erdbeobachtungsplattformen (z. B. Sentinels, Landsat) von großem Interesse. Die Übersetzung der an Überwachungsstellen erfassten Daten auf die Fernerkundungsebene, ist jedoch eine Herausforderung, und die Erfassung vor Ort ist in der Regel unzureichend, um die immer größer werdende Vielfalt der Wälder in Bezug auf Struktur, Artenkonfiguration und Bewirtschaftungspraktiken darzustellen. Darüber hinaus sind intensive Überwachungsprogramme zeitaufwändig und kostspielig, insbesondere im Hinblick auf die systematischen Feldkampagnen, die zur Erhebung einheitlicher Daten erforderlich sind.

Dank der jüngsten Entwicklungen im Bereich der Fernerkundung-Technologie, insbesondere der unbemannten/unbesetzten Luftfahrzeug-Fernerkundung (Drohnen, UAV-RS), haben wir die Möglichkeit, primäre Nahbereichs-Fernerkundungsdaten mit einer sehr hohen zeitlichen und räumlichen Auflösung mit einem breiten Spektrum an aktiven und passive Sensoren zu verschneiden. Durch UAV-RS ergibt sich die Möglichkeit, typische Waldmessungen wie Baumhöhe und Kronendurchmesser, sowie die Verbesserung phänologischer Beobachtungen und Baumzustandsbewertungen durch Bereitstellung einer Luftperspektive zu erhalten, die eine Ansicht vom Boden ergänzt. Darüber hinaus kann die Kartierung einzelner Bäume mit einer hohen absoluten Geolokalisierungsgenauigkeit möglicherweise die Probleme der Kartierung von Standorten unter dichten Baumkronen lösen, was typischerweise eine Herausforderung bei dem Einsatz geodätischer Messinstrumente vom Boden aus darstellt. In einigen Fällen kann UAV-RS die Bodenbeobachtung nicht nur verbessern, sondern diese insbesondere in Bezug auf Baumhöhen- und Kronenmessungen ersetzen. Darüber hinaus können spezifische Messungen und Beobachtungen auf Level -II-Flächen mit gezielten externen Parzellen erweitert werden, um weitere Waldtypen zu erfassen, die für eine bestimmte Region repräsentativ sind.

Diese Dissertation untersucht die Möglichkeit, traditionelle Bodenbeobachtungen und -messungen auf Flächen intensiver Waldüberwachung mit UAV-RS-Techniken zu verbessern oder in einigen Fällen zu ersetzen. Insbesondere werden drei Hauptforschungsfragen untersucht. Das erste betrifft die UAV-basierte photogrammetrische Baumhöhenextraktion und ob sie mit traditionellen Methoden vergleichbar ist. Die zweite Frage ist, ob die UAV-Wärmebildgebung Trockenstress in Buchen genau erkennen kann. Und die dritte Forschungsfrage lautet, ob UAV-abgeleitete phänologische ML-Modelle genau genug sind, um Bodenbeobachtungen zu ersetzen.

In der ersten Untersuchung wurden photogrammetrische Baumhöhen (*Pinus sylvestris*) mit Feldmessungen mit einem Hypsometer verglichen. Beide Methoden wurden anhand der einzigartigen Gelegenheit des Einsatzes destruktiver Techniken validiert, bei denen ausgewählte Bäume zu Durchforstungszwecken gefällt und die Baumlänge horizontal entlang des Bodens gemessen wurde. Hier zeigte sich, dass herkömmliche Baumhöhenmessungen mit einem Root-Mean-Squared-Error (RMSE) von 0,304 (1,82 %) tendenziell überschätzen und photogrammetrische Baumhöhen dazu neigen, Baumhöhen mit einem RMSE von 0,34 m (2,07 %) zu unterschätzen (n = 34). Diese Ergebnisse sind bedeutsam, da die Überschätzung der feldbasierten Baumhöhenmessung zu Unsicherheiten führen könnte, wenn sie als Validierungsmethode für photogrammetrische oder Laserscanning-Baumhöhenextraktionsmethoden verwendet wird.

Die zweite Studie umfasst eine Untersuchung der Erfassung von UAV-basierter Wärmebildgebung mit dem Ziel der Quantifizierung von Trockenstress in Bäumen und der Modellierung des Baumwasserdefizits oder „Tree Water Deficit“ (TWD). Die Anfangsphase der Studie umfasste Indoor-Experimente mit zwei Pflanzen, bei denen Trockenstressbehandlungen durchgeführt wurden, die mit Blatttemperatursensoren validiert wurden. Es wurde entdeckt, dass die thermische Pixelextraktion mit weniger als drei Pixeln zu einem Fehler von bis zu 1 K führen kann, während eine Erhöhung der „Spot Size“ (Punktgröße) den mittleren Fehlerunterschied auf 0,02 K reduzieren kann. Weiterhin zeigten die Versuche, dass Trockene Blätter durch Trockenheit verursacht wird, aufgrund eines geringeren Emissionsgrades können sich auch Wärmebildfehler ausbreiten. Bei Feldversuchen, die mit an Baumkronen montierten Blatttemperatursensoren validiert wurden, wurde festgestellt, dass Bilder, die von Rasterflügen erfasst wurden, aufgrund einer Zunahme des Abstands und des Winkels zu Validierungsbäumen höhere Standardabweichungen erzeugten. Wärmebilder, die im Schwebeflug aus nächster Nähe

aufgenommen wurden, führten zu geringeren Standardabweichungen. In Bezug auf die maschinelle Lernmodellierung von TWD wurden Trainings- und Validierungsdaten unter Verwendung von Punktdendrometern neben Eingabemerkmalen erfasst, die aus meteorologischen Daten und individuellen Baumkronen-Wärmebildaufnahmen aus nächster Nähe abgeleitet wurden. Die besten Ergebnisse wurden mit dem Generalised Additive Model (GAM) und relativer Luftfeuchtigkeit, Lufttemperatur und Baumkronentemperatur als Eingabemerkmale mit einem RMSE von 4,92 (µm) und R^2 von 0,87 erzielt. Diese Studie stellt eine vielversprechende Methode zur Quantifizierung von Trockenstress bei Rotbuche (*Fagus sylvatica*) mit Wärmebild- und meteorologischen Daten vor, wobei der Schwerpunkt auf der Reduzierung der UAV-basierten Thermografie-Fehlerausbreitung liegt.

Die dritte Studie untersucht die Vorhersage hochauflösender phänologischer Frühjahrsphasendaten für Rotbuche unter Verwendung von UAV-basierten multispektralen Indizes und maschinellem Lernen. Feldbeobachtungsdaten, welche von 2019 bis 2021 zu Trainings- und Validierungszwecken erfasst wurden, wurden verschiedene maschinelle Lernalgorithmen mit ausgewählten Vegetationsindizes getestet, die von multispektralen und RGB-Sensoren abgeleitet wurden. Trainingsdatensätze wurden während des Merkmalsauswahlprozesses in verschiedene Partitionen sortiert, um jahresspezifische Beobachtungsfehler zu erkennen, die die Modellierungsgenauigkeit beeinträchtigen könnten. Das endgültige GAM-Boosting-Modell unter Verwendung des Green Chromatic Coordinate (GCC)-Vegetationsindex, der von den RGB-Sensoren abgeleitet wurde, war in der Lage, phänologische Frühlingsphasen auf unsichtbaren Datensätzen innerhalb eines RMSE-Schwellenwerts von 0,5 vorherzusagen. Bei Indizes, die das Nahinfrarotband (NIR) implementieren, wurden jedoch aufgrund von Übersättigung erhebliche Einschränkungen entdeckt, und eine verbesserte Feldkalibrierung ist erforderlich.

Es kann geschlussfolgert werden, dass die Anwendung von UAV-RS für ein intensives forstliches Monitoring höchstwahrscheinlich in naher Zukunft ein fester Bestandteil von werden wird, jedoch die Herausforderungen, die in dieser Dissertation in Bezug auf eine einheitliche radiometrische Kalibrierung, Wärmebildfehlerausbreitung und Die Übersetzung von Bodenbeobachtungen in RS-Daten erfordert jedoch umfangreiche Erkundungen und Experimente. Neben der Einrichtung eines Leitungsgremiums zur Entwicklung und Aufrechterhaltung von Standardisierungszielen für die UAV-basierte intensive

Waldüberwachung wird eine koordinierte Anstrengung zwischen Forschungseinrichtungen empfohlen.

Table of Contents

List of Figures

List of Tables

List of Acronyms

µm	Micrometer
AI	Artificial Intelligence
AIC	Akaike Information Criterion
ALS	Aerial/Airborne Laser Scanning
AOI	Area-of-Interest
AT	Air Temperature
AVP	Active Vapor Pressure
CHM	Canopy Height Model
CIR	Colour-Infrared
DAP	Digital Aerial Photogrammetry
DLS	Downwelling Light Sensor
DL	Deep Learning
DOY	Day of Year
DSM	Digital Surface Model
DTM	Digital Terrain Model
EO	Earth Observation
EVI	Enhanced Vegetation Index
GAM	Generalised Additive Model
GBM	Gradient Boosting Machine
GCC	Green Chromatic Coordinate
GCP	Ground Control Point
GNDVI	Green Normalized Difference Vegetation Index
GSD	Ground Sampling Distance
HPC	High Performance Computing
IPCC	Intergovernmental Panel on Climate Change
K	Kelvin
LAI	Leaf-Area-Index
LiDAR	Light Detection and Ranging
LVPD	Leaf-to-Air Vapour Pressure Deficit
LWIF	Longwave Infrared

MAE	Mean Squared Error
ME	Mean Error
ML	Machine Learning
NDRE	Normalized Difference Red-edge Red Index
NDREI	Normalized Difference Red-edge Index
nDSM	Normalized Digital Surface Model
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NGRDI	Normalized Green Red Difference Index
NIR	Near-Infrared
NN	Neural Network
NUC	Non-uniform Correction
PPC	Photogrammetric Point Cloud
PPK	Post-processing Kinematic
RGB	Red-Green-Blue
RH	Relative Humidity
RMSE	Root-Mean-Squared-Error
RS	Remote Sensing
RTK	Real-time Kinematic
RTK-GNSS	Real-time Kinematic Global Navigation Satellite System
SAR	Synthetic Aperture Radar
SVP	Saturation Vapor Pressure
SD	Standard Deviation
SDE	Standard Deviation Error
SfM	Structure-from-Motion
SR	Solar Radiation
TIR	Thermal Infrared
TS	Total Station
TWD	Tree Water Deficit
UAS	Unoccupied/Unmanned Aerial System
UAV	Unoccupied/Unmanned Aerial Vehicle
UAVP	Unoccupied/Unmanned Aerial Vehicle Photogrammetry

UAV-HI	Unoccupied/Unmanned Aerial Vehicle Hyperspectral Imaging
UAV-RS	Unoccupied/Unmanned Aerial Vehicle Remote Sensing
VPD	Vapour Pressure Deficit
WS	Wind Speed

1 Introduction

In these times of exponential technological growth, advancements in high performance computing and remote sensing (RS) platforms are paralleled with an accumulating threat to European forests due to rising temperatures and extreme weather patterns. In order to insure forest resilience in a changing climate, a holistic form of forest assessment and management is required considering not only typical tree parameters for determining growth and health but also climatic drivers in terms of environmental conditions and climate change modelling scenarios (Achim et al., 2021). Additionally, in-situ terrestrial monitoring is essential for the calibration of remote sensing data (Lausch et al., 2016) for the large-scale wall-to-wall mapping of forest condition. Unmanned/Unoccupied Aerial Vehicles (UAV) mounted with passive and active sensors have recently opened up new avenues for forest mapping and research for the purpose of primary data collection particularly in the area of intensive forest monitoring (i.e. Level II) which was founded on traditional field-based observation methods in the 1990's, and has been gradually adding technological advances (i.e. electronic dendrometers, hemispherical photography, Airborne Laser Scanning). Low-altitude UAV-based sensors can aid in intensive monitoring programs in terms of simply providing a "view from above" for the enhancement of phenological and tree crown condition observations in terms of annotations linking expert qualitative observations to high resolution imagery, as well as deliver quantitative data modelled from high resolution below canopy sensors (i.e. dendrometers and sap flow sensors), meteorological data, and litterfall. Additionally, UAVs can enhance, and in some cases, replace traditional ground-based data collection methods (i.e. tree height) as well as acquire individual tree geopositioning for plot-level mapping and time-sensitive archival photo documentation. Such enhancements to forest monitoring systems through the incorporation of standardised low-altitude UAV-based data could prove an important step towards realising European forest monitoring goals.

The use of UAVs for the purpose of enhancing intensive forest monitoring programs has, until recently, been deemed as experimental and implementation as a standardized data acquisition methodology is not yet, but at the brink of becoming a viable forest monitoring tool. Reasons for the lack of acceptance among monitoring program managers in terms of UAVs being implemented as a recognized data acquisition tool, lies not only in the scepticism of a new and unproven technology but also in the lack of established universally documented methods to carry out specific intensive monitoring tasks such as RS-based phenological observations, tree crown

condition assessments and tree height measurement. Additionally, the shortage of specialized personnel to carry out technical hardware and software operations is a considering factor as well as the costly hardware required to accomplish the complex workflow of acquiring UAV-based data beginning with flight campaigns to the high processing loads and increased storage capacity. This is not only to mention the analysis and modelling capabilities required to handle such datasets in particular RS data which does not necessarily simply translate directly to ground observation data typically acquired at intensive forest monitoring sites. In this dissertation, the possibility to develop standardised RS-based forest monitoring practices is explored with the objective of incorporating UAV technology into intensive forest monitoring programs at the operational level.

1.1 Intensive Forest Monitoring: A New Perspective

With renewed interest in the establishment of a harmonised European forest monitoring system, it is an imperative to incorporate well established systems while strengthening the remote sensing component (Ferretti, 2021). Level II intensive monitoring plots, alongside highly instrumented "core plots" or "supersites" (Fischer et al., 2011), though effective for monitoring forest status and change, often lack the link to small-scale Earth Observation (EO) platforms due to inadequate sampling coverage, plot positional inaccuracies, absence of individual tree mapping and the linkage of in-situ measurements and observations to the aerial perspective. According to governmental agencies (European Commission, 2021), the future of understanding and assessing forests lies in a holistic approach in monitoring which entails understanding full tree and stand dynamics. Current methods such as those carried out at intensive monitoring plots, undertake in-depth analysis of forest ecosystem dynamics, however acquired data and models are spatially limited especially when we are interested in undertaking small-scale forest assessments. Furthermore, we require to not only understand the current status of forests but also what and where action is necessary in order to create and maintain resilient forests for the times ahead. RS, especially open access satellite-based RS (i.e. Sentinels, Landsat) has been of great interest as we are able to examine forests at an immensely large scale at a high temporal resolution (Holzwarth et al., 2020). Here, the usage of big data applications implementing machine learning (ML) and deep learning (DL) algorithms trained and validated with ground truthing posed to solve real world problems with generalizable models (X. X. Zhu et al., 2017). The challenge is however in that existing forest monitoring programs possess valuable long-term ground truthing data from observations and measurements which do not necessarily translate automatically to the aerial perspective. The problem lies for the most part in a combination of the "mixed pixel" effect (Choodarathnakara et al., 2012; S. Liang & Wang, 2020) and the capabilities of aerial

sensors to capture information relating to terrestrial in-situ observations and measurements. The mixed pixel effect arises typically from coarser resolutions where for example tree species in a mixed stand cannot be separated for phenological mapping purposes. With regard to the difficulties in the translation of ground-truthing data to the aerial view, an example would be in the modelling of daily drought-stress related stem diameter fluctuations with canopy thermal measurements, as the sensor-visible sun leaves of the upper canopy will not necessarily be representative of processes occurring in the entire tree crown (see Section 3). Another example would be that of crown damage observations from the ground perspective which are unable to detect upper canopy discoloration normally visible from the aerial perspective. Here lies the potential for the misinterpretation of ground events in terms of aerial-based qualitative and quantitative assessments when carried out in absence of expert knowledge possessed by ground crews.

The possibility to link in-situ terrestrial observations with concurrent high-resolution UAV data could potentially bridge the gap to EO platforms. The difficulty with upscaling intensive forest monitoring plots (i.e. Level II) is that due to operating expenses, there are a limited number of plots available for RS-related ground-truthing purposes. In Germany, there are 68 Level II plots nation-wide (Sanders et al., 2020) and approximately 800 plots in the European-wide ICP Forests network (ICP-Forests, 2022). In order to account for data-hungry machine learning algorithms as well as increasingly species rich and structurally diverse forests, a large amount of training and validation data is required. Figure 1 shows the conceptual framework for the extension of Level II plots for the purpose of expanding available sampling plots. It should be stated however that these “external” plots would be in effect a reduced version of the original intensive monitoring plots designed to acquire only specific parameters which are conducive to UAV-RS methods.

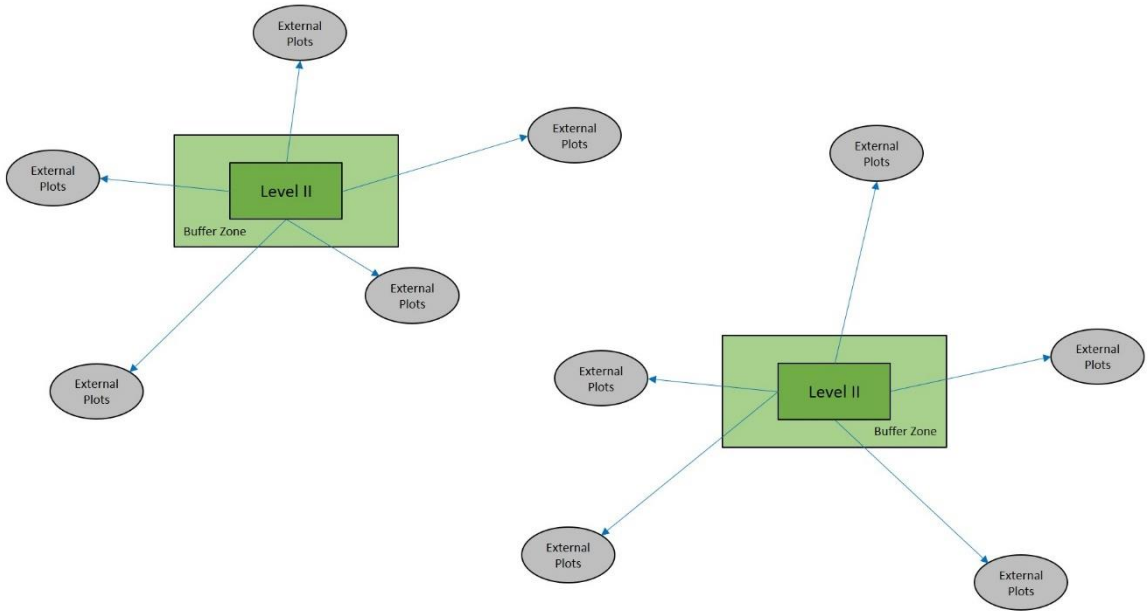


Figure 1: Extending Level II plots with less intensive UAV-based external plots could solve the problem of limited amounts of training and validation ground-truthing data for the upscaling to EO platforms.

Implementing UAV-RS to actually replace measurements and observations at intensive monitoring plots is not the aim here, yet enhancing specific observations could be beneficial (i.e. phenological and tree condition observations). On the other hand, plot mapping and georeferencing on an individual tree basis is of great interest, as here a digital twin is created which can be updated on a yearly basis. This in turn facilitates the extraction of individual tree heights as well as crown diameter measurements at similar and even better accuracies than traditional methods (see Section 2). In terms of external plots however, the acquisition of specific parameters for the purpose of upscaling to satellite platforms could facilitate large-scale mapping projects of for example tree condition or phenology. Table 1 gives an overview of the potential survey parameters (adapted from Lorenz, 1995) which could be suitable for upscaling at external plots. Methods and models would require to be developed and trained at the main plots and additional validation data may be necessary for external plots which vary from main plots in terms of structure, age and species configuration. It is assumed that parameters which would require hyperspectral data are deemed at this stage impractical due to difficulties in the availability of specialized sensors as well as challenges in developing a standardized acquisition methodology. Additionally, access to Light Detection and Ranging (LiDAR) sensors may not be feasible for all member states however is increasingly becoming common usage (Maltamo et al., 2014; J. White et al., 2016). The usage of multispectral sensors is for the most part in significant use and the utilization of low-cost RGB sensors is wide spread. The acquisition of meteorological

data with onboard UAV sensors is also an interesting development, although will only be limited to data acquired during the flight mission.

Table 1: Overview of parameters surveyed at Level II sites where UAV-RS could be a possibility (adapted from Lorenz, 1995).

Survey	***Provide data on***	***UAV-RS***	***Description***	***Sensors***
Plot description	Location, size and status of the plot	Yes	Individual Tree Mapping (sub-decimeter accuracy)	RGB, Multispectral, LiDAR
Stand description	Basic characteristics of the stand	Partially	Digital Twin - Top View only with photogrammetry	RGB, Multispectral, LiDAR
Management operations and natural disturbances	Forest management operations and natural disturbances	Partially	Wind throw, Missing Trees, Gaps	RGB, Multispectral, LiDAR
Tree condition	Indicators of crown, branches and stem status of the trees	Partially	Defoliation, Discoloration, Drought stress, qualitative and quantitative	RGB, Multispectral, LiDAR
Tree growth and yield	Actual periodic growth of the stand and of individual trees	Partially	Tree Height, Crown Diameter, Allometric equations for DBH estimation	RGB, Multispectral, LiDAR
Tree phenology	Timing of the annual development stages of forest trees (plot level)	Yes	Flushing/Colouring, qualitative and quantitative	RGB, Multispectral, LiDAR
Tree phenology (intensive)	Timing of the annual development stages of forest trees (individual tree level)	Yes	Flushing/Colouring, qualitative and quantitative	RGB, Multispectral, LiDAR
Ground vegetation	Species richness and abundance	No	Issues of occlusion, possibility with LiDAR	LiDAR
Ozone injury on plants	Presence on visible injury attributable to tropospheric ozone	No	Issues of occlusion, hyperspectral bands required	Hyperspectral
Meteorological measurements	Basic (T, P, wind speed) meteorological variables, soil moisture and temperature	Partially	On board measurements possible but only during flight mission	On board sensor (i.e. Rasberry Pi)
Foliar sampling and analysis	Chemical concentration of elements in foliage of trees	No	Hyperspectral Sensors not wide-spread available	Hyperspectral
Ambient air quality	Concentration of SO_2, NO_x, O_3 in the air	Partially	Possible with onboard measurements but only during flight mission	On board sensor (i.e. Rasberry Pi)
Leaf area index	Total canopy leaf area	Yes	Adaptation of Hemispherical methods, ML training from litterfall	RGB, Multispectral, LiDAR

1.2 Unoccupied Aerial Vehicles and Sensors

A prominent feature of a multicopter is the possibility to fly at lower altitudes while hovering. This in turn enables the acquisition of high-resolution single-shot imagery. These unrectified images, although subjected to radial distortion, can deliver highly detailed, and in some cases, radiometrically calibrated tree crown images facilitating qualitative assessments as well as automated measurements of various parameters in terms of tree condition, drought stress (see Section 3) and phenological phase prediction (see Section 4). On the other hand, the usage of Structure-from-Motion (SfM) to create point clouds and derived Digital Surface Models (DSM) and Canopy Models (CHM) as well as Orthomosaics enables the mapping of individual trees and tree height and tree crown diameter extraction. Here, the resolution will typically not be as high as single image acquisition for the reason that in order accomplish image mosaicking with homogeneous textures, as is the case in forested scenes, a substantial flying height is

required. Figure 2 gives an overview of the products available from typical UAV-based RGB and multispectral sensors.

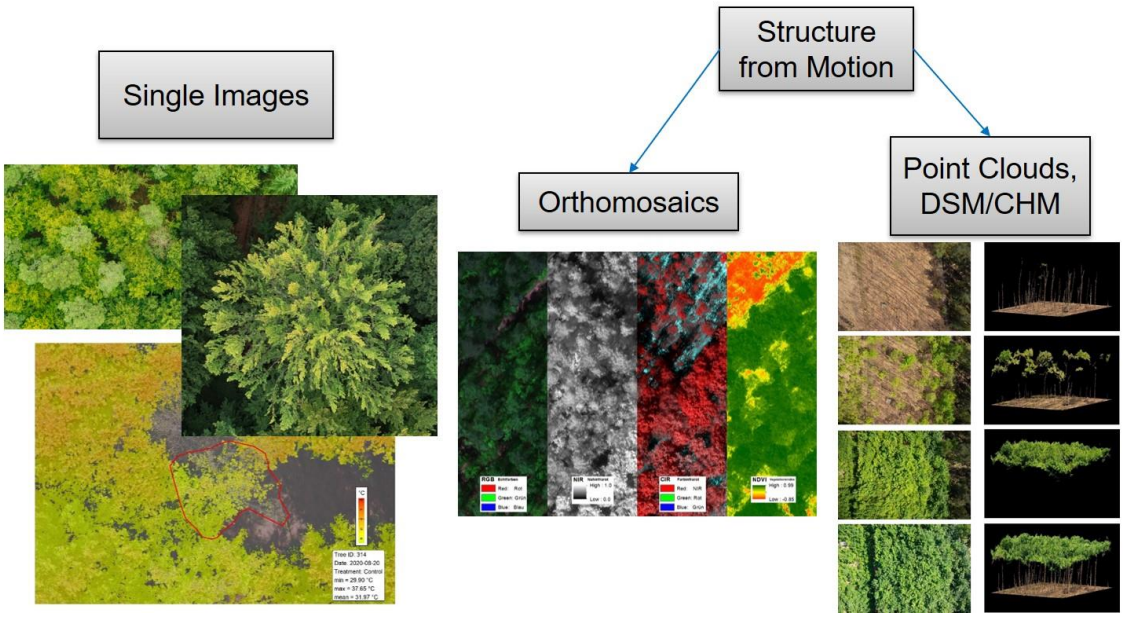


Figure 2: UAV-derived single images as opposed to typical orthorectified images and other photogrammetric products.

The inclination to deploy close-range aerial sensors has long been seen as an invaluable venture with developments ranging from air balloons to helicopters over decades of experimentation (McGlone & Lee, 2013). Recent development in UAVs has provided the scientific community with a cost-effective method to acquire primary RS data, however is often still seen as experimental (J. White et al., 2016) and not in use at the operational level at least for intensive forest monitoring purposes. With various UAV platforms available, the possibilities to get an array of sensors airborne is ever increasing however it is important to identify which UAV platforms are conducive to Level II intensive monitoring plots. The typical size of a Level II plot is between 0.25 and 1 ha, with an allowed minimum size of 0.25 ha (Ferretti & Fischer, 2013) which is conducive to the use of Multicopters (i.e. Quadrocopters). Multicopters possess shorter flight durations than fixed-wing platforms (Eisenbeiß, 2009), and have the capabilities to mount a wide range of sensors at nadir and oblique angles as well as accommodate multiple sensor configurations. Typical sensors at the current state-of-the-art for close-range aerial forest remote sensing are as follows (X. Liang et al., 2022):

- Multispectral Imaging
- Laser Scanning (LiDAR)
- Hyperspectral Imaging
- Synthetic Aperture Radar (SAR)

In terms of multispectral imaging, there can be a division between RGB and multispectral sensors where for multispectral, near-infrared (NIR), red-edge and even Longwave Infrared (LWIR) bands are made available in addition to RGB bands within one imaging system (i.e. *Micasense Altum*). RGB sensors would represent sensors which only possess RGB bands of the visible spectrum and can often be found on low-cost platforms with readily available sensors such as the *DJI Phantom Pro 4* (dji.com) however also high-resolution sensors such as the *Zenmuse P1* (dji.com) or *Phase One* (phaseone.com).

The use of Laser Scanning for UAV-based forest monitoring has shown an increase recently with various commercially available light-weight laser scanners adapted for UAV usage (Kellner et al., 2019). Here, the possibility to acquire highly accurate point cloud data with limited occlusion issues is a reality, however costs can be very high in comparison with optical sensors. Furthermore, UAV Laser Scanning (UAV-LS) does however necessitate more testing areas of scan angle and flight planning (Brede et al., 2022).

With regards to UAV-based hyperspectral remote sensing, hundreds of bands are made available where high-resolution spectroscopy can be implemented for the characterization of chemical and physical properties (Adão et al., 2017). UAV Hyperspectral imaging (UAV-HI) does require extensive calibration methods (H. Li et al., 2015) which can be complex involving custom workflows (Proctor & He, 2015). Alongside challenges in radiometric calibration, emerge further difficulties due to systematic noise, variable speed of acquisition, and poor orthorectification among other factors (Iwanoczko, 2017). Current research into the uses of UAV-HI is a worthwhile pursuit for various institutions, however with regards to the development of a standardized acquisition methodology at intensive forest monitoring plots could prove challenging.

With respect to recent advancements in Microwave sensing technology, in particular the miniaturization of sensors for the acquisition of high-resolution UAV-based Synthetic Aperture Radar (UAV-SAR) is of interest for the assessment of biophysical parameters such as canopy height, Leaf Area Index (LAI) (Pandey & Arellano, 2023) and other aspects of forest structure as well as soil and vegetation moisture (Clewley et al., 2013; Pulliainen et al., 2004). These applications are still however in the experimental stages (Xu & Zhu, 2018) and more research and testing as well as sensor availability is required before wide-spread usage is established.

1.3 Digitalization of Intensive Monitoring Plots

The incorporation of UAV technology for the purpose of plot digitalisation is of interest for intensive monitoring plots in terms of absolute geolocation as well as tree height and crown diameter extraction among other possibilities. In order to realise these recent technological developments in individual tree mapping and measurement, a digital framework is required for the UAV to operate. 3D mapping enables the automated flight planning where the UAV can with repeated fight missions for time-series applications cover the same area for grid mapping as well as repeated individual tree image capture using waypoints. Here lies the critical aspect of UAV Real-time Kinematic (UAV-RTK) and/or Post-processing Kinematic (PPK) capabilities which enable sub-decimate absolute horizontal accuracies without the usage of GCPs (Tomaštík et al., 2019). Mapping intensive monitoring plots (i.e. Level II) from above not only increases the positional accuracy in comparison to typical ground-based methods, but also facilitates an Orthorectified map where the analysis and measurements of individual trees can be carried out at a digital workstation.

1.3.1 Stand and Individual Tree Mapping

An important step in the incorporation of UAVs into an intensive forest monitoring plot network is the digital mapping of the stands as well as individual tree positions. In order to plan and execute various UAV-related tasks, a spatially accurate digital framework is required for the UAV to operate effectively for the purpose of enabling repeatable (time-series) data acquisition. A digital framework in the case of a monitoring plot would be in the creation of a "digital twin" of the stand which is essentially a virtual version of the plot. Here the Users can analyse and manage the plot as well as plan UAV missions using a 2D or even 3D virtual representation via a computer desktop.

To follow, a brief description is provided of the overall workflow in the digitalisation of a monitoring plot and creating a digital twin framework for UAV flight planning. To begin, the corner points of the plot require to be established in order to determine the boundaries for creation of the initial Orthomosaic which is derived from a grid-based flight plan (see Figure 3). These corner points can be established using a ground-based Real-time Kinematic Global Navigation Satellite System (RTK-GNSS) receiver or also from the aerial perspective through low-cost single shot RGB imagery. Here a visible target such as a contrasted ground target or even a field crew member with an orange hat can be positioned at the corner boundaries and an image taken using the live view of the target positioned in the center of the display of the UAV ground control

system. Cross-hairs can typically be activated via the systems settings of the Unmanned Aerial System (UAS) ground station. The cross-hair positioning should be tested beforehand to determine if the crosshairs actually line up with the center of the digital image. In some cases, an offset could be necessary to account for RTK-GNSS antennae positioning in relation to the sensors mounting position on the gimbal. With at least three corner positions the coordinates can then be extracted from the Image EXIF Metadata and used to create a rectangle boundary in the form of a comma-separated values *(.csv*), Shape *(.shp*) or Keyhole Markup Language *(.kml*) file. Once established, the plot boundary can be used to create a grid-based flight plan which when carried out with a RTK-GNSS-equipped UAS, can deliver sub-decimeter absolute geolocation accuracies with the resulting Orthomosaic. In the case of the absence of a RTK-enabled UAV, a ground control point (GCP) network could be established however this can be less cost-effective due to increased labour costs, and accuracy is dependent on ground-based RTK-GNSS capabilities as well as GCP visibility from the aerial perspective.

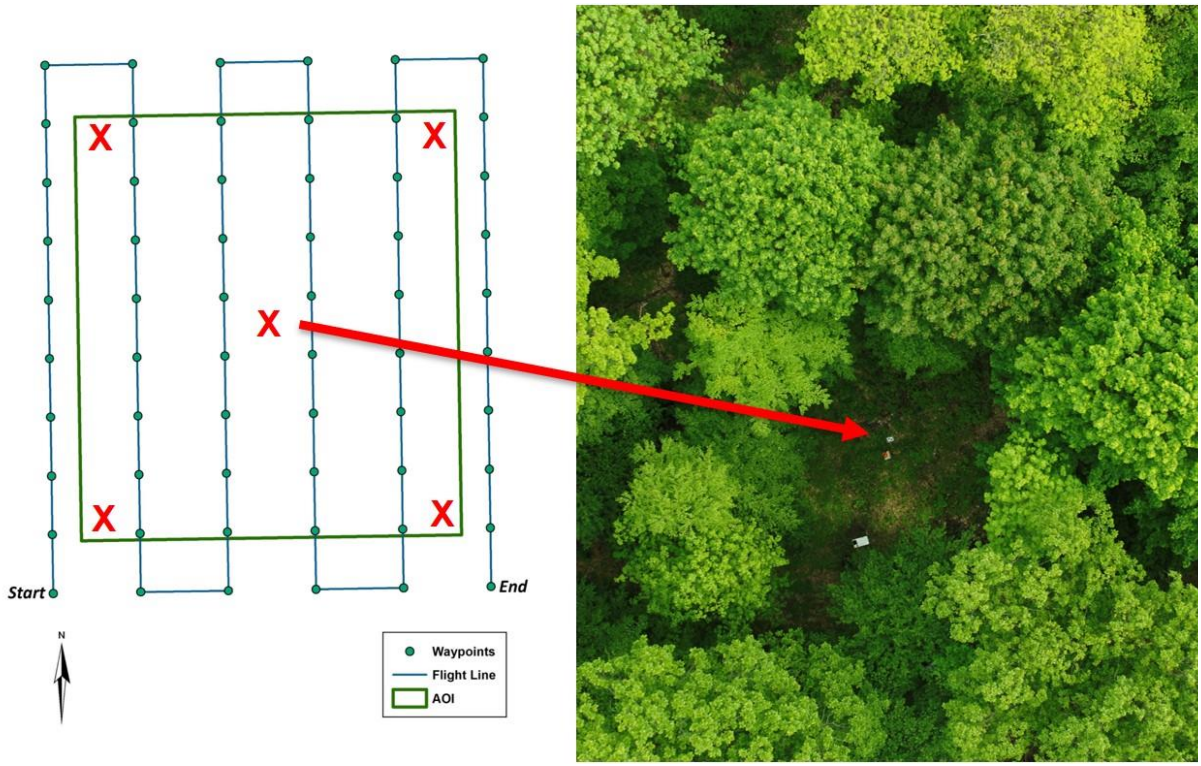


Figure 3: Typical Ground Control Point (GCP) configuration (red X's) of a one ha intensive monitoring plot. The use of UAV-RTK/PPK can eliminate the necessity for GCP usage and can facilitate the measurement of the ground position of the plot center (see red arrow). The extent of the plot (outer X's) must however be determined for flight planning purposes. The position of the plot center or near center is dependent on ground visibility.

Locating individual trees for the purpose of ID assignment can be challenging in homogeneous forest stands (i.e. monocultures). At intensive monitoring plots it is important to be able to relocate individual "focus" trees for time-series applications which is particularly relevant for trees where various sensors are installed (e.g. dendrometers or sap-flow sensors) or phenology

and crown conditions observations take place. In the case that trees with existing IDs are not easily located in the Orthomosaics, a similar method as described previously can be implemented where the center of the tree crown of a particular focus tree is determined via single shot imagery.

1.3.2 Flight Planning

Gone are the days of the necessity of manual flight planning calculations. Today, numerous flight planning applications are available which can automatically determine an appropriate flight plan grid depending on User input, typically comprising of an Area-of-Interest (AOI), sensor, desired overlap and Ground Sampling Distance (GSD). An imperative for forest photogrammetry is a constant GSD during the flight mission as this can affect image overlap in especially mountainous terrain. It is recommended to maintain frontal and side image overlap of approximately 80 - 85 % in forested scenes and in some cases a higher-flying height and/or shorter focal length may be required due to limited pixel matching possibilities in homogeneous forest landscapes. Inadequate image overlap can result in artefacts or "holes" in the resulting Orthomosaics as well as inconsistencies in the resulting point cloud.

With regard to area coverage outside of the AOI, it is recommended to maintain at least 50 - 75 % of the image ground footprint along the outer boundary providing a buffer (i.e. 5 trees) has already been maintained in reference to focus trees. The GSD is estimated by means of the flying height (*H*), which is the distance from the ground nadir point to the perspective center, the focal length (*f*), and the sensor's pixel size (*p*).

$$\boldsymbol{GSD = \frac{H * p}{f}}$$

1.3.3 Tree Height Extraction and Crown Measurements

The estimation of photogrammetric tree heights is dependent on the normalisation of the DSM. Forest photogrammetry tends to have limited capabilities due to occlusion in creating a complete Digital Terrain Model (DTM) where the forest floor is fully represented. As opposed to Aerial Laser Scanning (ALS), it is also challenging to classify the ground points which is typically determined from last returns originating from LiDAR-derived point clouds. Here, it is essential to implement an existing LiDAR-derived DTM for calculating the normalised DSM (nDSM),

also known as the CHM where elevation (Z) values are represented in values beginning from 0-*n* meters. In deciduous stands, it can be possible to create a photogrammetric DTM during the leaf-off season providing year-round ground vegetation is not a hinderance. LiDAR-derived DTMs are however typically superior as occlusion is less of a factor. The CHM is calculated as follows:

$$CHM = DSM - DTM$$

Once the CHM is created, the highest pixel within the individual tree crown is assumed to be the apical meristem and highest point of the tree. When locating the highest pixel within a tree crown particular attention should be taken for focus trees where the crown is forked as can be the case with some deciduous trees.

Tree height accuracy will be dependent on the quality of the 3D information of the point cloud where individual tree crowns are adequately represented. Additionally, the quality of the DTM is critical in establishing the base of tree stems. It is recommended that during field campaigns, the tree height from the ground is measured for focus trees ($n > 10$) using a hypsometer and compared with photogrammetric tree heights. Here a RMSE within 0.5 meters is acceptable, however it should be considered that for example a hypsometers height estimates could over-estimate heights and Photogrammetric methods could under-estimate tree heights resulting in error uncertainty when a high accuracy is required (e.g. yearly tree heights) (see Section 2). For this reason, it is recommended to test photogrammetric methods at core plots where known tree heights are available and the repeatability of photogrammetric methods can be tested. Yearly tree height extraction could be possible when an accuracy of under at least a RMSE of 30 cm is attainable. For faster growing species such as *Salix* or *Populus* a higher RMSE could be possible.

Tree crown measurements can typically be carried out using the Orthomosaic or CHM providing standards for horizontal accuracy are met and the tree crown pixels are clearly represented in the Orthomosaic. Here, an accurate segmentation of the tree crown is required which is followed by a calculation of the average diameter of the segmented polygon.

Once the digitisation of intensive forest monitoring plots is acquired, as described above, a framework is established which enables the further exploration of various UAV-RS methods for the purpose of extracting and predicting forest monitoring parameters as well as upscaling to EO platforms.

1.4 UAV-RS Linkage to Forest Parameters

Many of the previously mentioned UAV-RS technologies involved in the digitisation of individual trees and forest stands are well established as RS-based mapping technologies, however the translation into data which can interpret forest parameters reliably at least the same level of accuracy as a human observer is not yet well established. At the pixel level, especially for high resolution multispectral imaging, a close-range (i.e. < 10 m) image of a tree crown could be comprised of hundreds of pixels. A Sentinel 2 image could possess typically multiple tree crowns within one pixel. At such a high-resolution, UAV-RS has the capabilities to carry out highly complex analysis of a single tree crown where various specific leaf morphological aspects as well as leaf coloration traits can be accounted for. Additionally, the absence of leaves, or "woodiness" as well as branch formation can deliver insights in past damage effects including drought events and xylem cavitation or perhaps be only representing Autumn leaf senescence. Such high-resolution pixel information could be interpreted by an expert at a computer workstation however not typically feasible for large datasets. Herein lies the challenge of automating such interpretations as traditional statistical analysis methods may not be limited in capturing such complex relationships evident with the assortment of available high-resolution geospatial features. For this reason, the implementation of ML models should enable more complex predictive modelling which can incorporate an array of features which would not only be limited to geospatial datasets but also other meteorological and environmental variables. Furthermore, ML modelling, thanks to also current high-performance computing capabilities, can handle large datasets much more efficiently and economically than traditional methods. In this regard, the evaluation of current UAV-RS technologies is not the main focus of this dissertation, but rather the exploration of acquisition methods alongside feature engineering and ML modelling to interpret forest monitoring parameters is placed at the forefront.

1.5 Research Aims and Objectives

With rapid advancements in RS technology, as well as societal aims to better understand the effects of climate change on forests, there is an ever-increasing requirement to digitalize forest information in order to improve analysis possibilities not just in terms of increasing accuracy and efficiency, but also better understand forests on a large scale. A direct translation of forest information into useful information is however a challenge, and is particularly difficult to adapt UAV-RS methods to already well-established ground observation systems such as long-term intensive forest monitoring programs (Level II). RS methods are becoming more accurate and

reliable than previous experimental studies which is causing a level of uncertainty to develop as RS techniques become just as accurate as ground observations (Persson et al., 2022) and in some cases are more reliable. An obvious example here would be the case of potentially subjective ground observations such as phenology and tree condition assessments. Additionally, the traditional measurements of geometric features of an individual tree in terms of tree height and crown diameter have limitations in terms of accuracy and are also considerably time- and labor-intensive. The use of sensors which are mounted directly to trees, as with point dendrometers for example, can deliver highly accurate information but typically restricted to a limited number of trees per research plot. This dissertation explores the possibility to enhance or in some cases replace traditional ground observations and measurements at intensive forest monitoring plots using UAV-RS techniques. In particular, the following research questions will be explored:

(i) Is UAV-based photogrammetric tree height extraction comparable to traditional methods?
(ii) Can UAV thermal imaging accurately detect drought stress in Beech?
(iii) Are UAV-derived phenological ML models accurate enough to rival ground observations?

To follow, the answers to these research questions will be synthesized and a framework for standardization is presented.

1.6 Thesis Structure and Outline

As shown in Figure 4, this dissertation is divided into five main Chapters. Chapter 1 gives an overview of some new concepts with regard to the use of RS technology in conjunction with intensive forest monitoring plots in particular the use of UAVs and available sensors. Additionally, the digitalization of intensive forest monitoring plots as well as technical considerations are covered followed by the main aims and objectives of this dissertation.

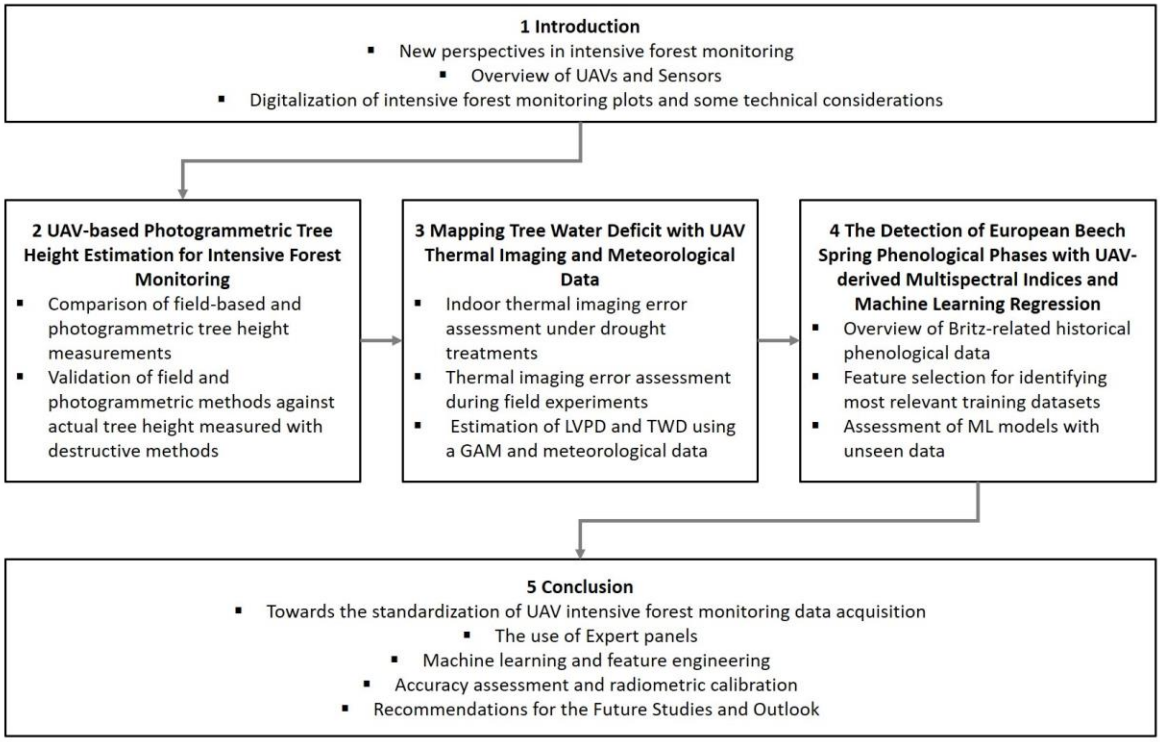


Figure 4: Structure of the thesis and content overview.

Chapter two is published as a peer review article in the journal *Remote Sensing* in the special issue *Remote Sensing Techniques for Precision Forestry* (Krause et al., 2019). As of the February 2023, according to *Google Scholar*, the article has been cited 125 times. The popularity of the article is most probably due to the reliable method of field validation presented in the study. Traditional field methods for tree height measurements alongside UAV-based photogrammetric measurements were compared and validated against actual tree lengths where the trees were felled and measured for length on the ground. Results showed a similar RMSE however photogrammetric tree heights tended to underestimate while traditional field measurements with a hypsometer tended to overestimate tree heights. This study is of significance, as it shows that UAV photogrammetry can acquire tree heights at a similar accuracy to that of field methods, with an increased standard of repeatability.

Chapter three delves into the acquisition of high-resolution thermal imagery for the purpose of mapping drought stress. The article (Krause & Sanders, 2022a) has been submitted to the journal *Remote Sensing in Earth Systems Sciences* and is currently under review. In this study, the uncertainties of thermal image acquisition accuracy of a UAV-mounted multispectral sensor were evaluated under controlled indoor conditions as well as field trials. Leaf temperature derived from the thermal sensor was validated using leaf temperature sensors. Furthermore, the possibility to predict Tree Water Deficit (TWD) was tested by means of thermal imaging and

meteorological data as input features and detrended and normalised point dendrometer data as the dependent variable. This study is significant for the scientific community as it explores the origins of potential thermal imaging errors while pioneering the prediction of the TWD with UAV-derived tree crown temperature and meteorological data.

Chapter four involves a study of the prediction of spring phenological phases with UAV-based multispectral indices and has been submitted to the journal *Annals of Forest Science* which is currently under review (Krause & Sanders, 2022b). Here, multispectral indices and ground-based spring phenological observation data from three acquisition years was applied to create predictive models. It was shown that low-cost RGB sensors and the Green Chromatic Coordinate (GCC) with ML can be utilized to predict spring phenological phases on unseen data.

Chapter five provides a general conclusion to the dissertation while providing insights into the importance of standardization of UAV-based intensive monitoring data acquisition. Here a framework for feature engineering/selection and usage of machine learning is presented with an emphasis on training and validation data approved through expert panel evaluations. Furthermore, the importance of accuracy assessment and reliable radiometric calibration is highlighted followed by recommended future research.

2 UAV-based Photogrammetric Tree Height Measurement for Intensive Forest Monitoring

Abstract

The measurement of tree height has long been an important tree attribute for the purpose of calculating tree growth, volume, and biomass, which in turn deliver important ecological and economical information to decision makers. Tree height has traditionally been measured by indirect field-based techniques, however these methods are rarely contested. With recent advances in UAV remote sensing technologies, the possibility to acquire accurate tree heights semi-automatically has become a reality. In this study, photogrammetric tree height and field-based tree height measurements of a Scots Pine stand were validated using destructive methods. Field-based measurements resulted with a similar level of error to that of the photogrammetric measurements with a RMSE of 0.304 m (1.82 %) and 0.34 m (2.07 %) respectively (n = 34). The photogrammetric tree height measurements of the remaining trees (n = 285) were validated against the field-based measurements and resulted with a RMSE of 0.479 m (2.78 %). Additionally, two separate photogrammetric tree height datasets were compared (n = 251), and a very low amount of error was observed with a RMSE of 0.138 m (0.79 %) suggesting a high potential for repeatability.

2.1 Introduction

Tree height is an important parameter required to quantify timber resources and is essential in evaluating the economic and ecological value of a forest stand. In particular, height plays an important role in the calculation of individual and total stand volumes (Kershaw et al., 2016) as well as assessing the social status of individual trees (Thenkabail, 2015) and productive capacity of a site (West, 2015). Furthermore, tree height is critical in the estimation of forest biomass and carbon stocks (Andersen et al., 2006), which are factors of increasing relevance for the mitigation of climate change through forest activities (Bolte et al., 2009; Spathelf et al., 2018; Picard et al., 2012).

As a typical measurement parameter for forest inventory and monitoring programs, tree height (h) is measured in the field by means of the direct or indirect measurement of the distance between the base (ground-level) and the tip (apical meristem) of a tree (Bragg, 2014; Laar &

Akça, 2007; Watts & Tolland, 2005). The direct measurement of tree heights can be accomplished with telescopic measuring poles, however the measurement accuracy and precision is limited to heights of up to 23 meters and the method is slow and cumbersome (Kershaw et al., 2016). Alternatively, direct measurements can be carried out with destructive methods where trees are required to be harvested and the length determined along the ground with a measuring tape. Indirect measurement on the other hand, would be that of non-destructive geometric or trigonometric methods (West, 2015) carried out via field measurements with a hypsometer (Apostol et al., 2016), laser device (Falkowski et al., 2006), or Total Station (TS) (Andersen et al., 2006). Indirect tree measurement is also possible using remote sensing techniques as in traditional photogrammetric measurements derived from analog aerial imagery (Rogers, 1947; Spurr, 1960), digital aerial photogrammetry (DAP) (Guerra-Hernandez et al., 2016; St-Onge et al., 2015), active sensor remote sensing techniques such as Light Detection and Ranging (LiDAR) (Andersen et al., 2006; Kaartinen et al., 2012; Maltamo et al., 2014; Næsset et al., 2004) or Interferometric Synthetic Aperture Radar (SAR) (Huanmin et al., 2010; Pohl & Loong, 2015).

Recent developments in remote sensing technologies such as LiDAR and Digital Photogrammetry opened up new possibilities in not only the estimation of individual tree heights, but also in the estimation of tree crown diameter (Falkowski et al., 2006; Hyyppä et al., 2005a). Tree height and crown diameter can be further utilized to estimate individual tree characteristics such as stem diameter and volume (Hyyppä et al., 2005b; Jucker et al., 2017). Such technological advances in forestry-related remote sensing are partly attributable to the arrival of relatively new terms such as “Precision Forestry” (Corona et al., 2017; Dash et al., 2016; Holopainen et al., 2014; Smaltschinski & Becker, 2009; Šumarstvo, 2010; Taylor et al., 2002). Based on reliable and accurate data, precision forestry delivers detailed information on the structural (Moskal et al., 2018) and ecological aspects of forests at a high spatial and temporal resolution even at the individual tree level (Holopainen et al., 2014). Remote sensing platforms are being implemented to obtain such data from various satellite, aerial and terrestrial platforms which are making use of active and passive sensors (Dash et al., 2016; J. White et al., 2016). LiDAR-based active sensor systems such as Airborne Laser Scanning (ALS) are of great interest due to the possibility to acquire highly detailed 3D information on the structural aspects of forests (Maltamo et al., 2014).

Typically, the interest in more detailed 3D forest information has mostly been driven by the prospect of profitability (Holopainen et al., 2014) within the forest industry, however for research purposes, the prospect of obtaining highly accurate and timely forest structural data is also of

interest (Ferretti & Fischer, 2013). Forest inventories and monitoring on the national and international level are nonetheless often limited to public funding which can restrict the accessibility to costlier methods of remote sensing such as for example ALS. DAP, in particular Unmanned Aerial Vehicle Photogrammetry (UAVP) however, offers a cost-effective alternative to ALS, where passive consumer grade sensors can be mounted on a civilian UAV (Eisenbeiß, 2009). A typical UAV mounted with passive sensors can be acquired and operated at a fraction of the cost of ALS with similar results and accuracy (Wallace et al., 2016).

Numerous studies involving the extraction of individual tree heights through ALS and DAP have been carried out with promising results (Apostol et al., 2016; Dempewolf et al., 2017; Falkowski et al., 2006; Gatziolis et al., 2010; Imai et al., 2004; G. Liu et al., 2018; Mielcarek et al., 2018; Panagiotidis et al., 2016; Popescu et al., 2002). One aspect common among all of these studies is that they all implemented indirect field-based tree measurements for validation. However, field-based measurements can themselves also potentially propagate error (Boi, 2005; Bragg, 2014; Butt et al., 2013; Larjavaara & Muller-Landau, 2013). To the authors' knowledge, there are few studies that implement direct measurements for the verification of ALS and DAP tree height estimations as this would require destructive methods which can prove costly or impractical. Sibon et al. (Sibona et al., 2016) however, assessed tree height estimation using ALS and field-based survey methods by implementing 100 harvested trees for validation. Results showed that the ALS derived tree heights were closer to the direct tree measurements than traditional field-based measurements. In a recent study conducted by Wang et al. (Y. Wang et al., 2019), the authors implement a non-destructive statistical method to show that field measurements tend to overestimate tree heights for codominant trees when correlated with ALS and Terrestrial Laser Scanning (TLS).

In this study, we explore the possibility of enhancing or even replacing field-based tree height measurements for long-term forest monitoring plots where a permanent Ground Control Point (GCP) configuration is feasible. Using the Britz Research Station as a case study, our aim was as follows: (i) determine the accuracy of traditional field-based tree height measurements against direct measurements (ii) determine the accuracy of photogrammetrically derived tree heights against direct measurements (iii) compare photogrammetric tree heights against field-based measurements (iv) assess the repeatability of temporal photogrammetric tree heights.

2.2 Materials and Methods

2.2.1 Study Site

The study site is located within the Schorfheide Biosphere reservation near the village of Britz, Brandenburg (52.87 °N 13.83 °S, 42 m above sea level) approx. 50 km north-east of Berlin, Germany. The Britz research station was originally established in 1972 by the *Institute for Forest Sciences Eberswalde* for the purpose of conducting forest hydrological research with the use of large-scale lysimeters (Müller et al., 2009). Currently, the Britz Research Station is under management of the *Thünen Institute for Forest Ecosystems* and consists of nine plots of varying species and structure (see Figure 5).

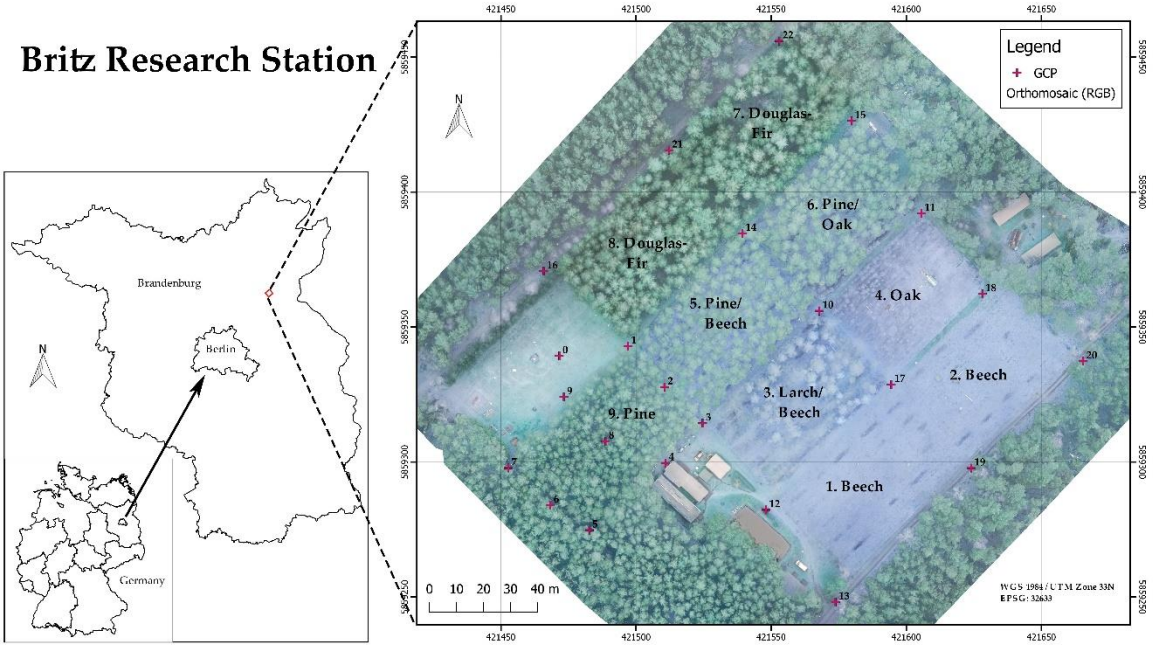


Figure 5: Overview of the Britz Research Station. Plots 1 & 2: European Beech (Fagus sylvatica), Plot 3: European Larch (Larix decidua) with a Beech understory, Plot 4: Sessile Oak (Quercus petraea), Plot 5: Scots Pine (Pinus sylverstris) with a Beech understory.

The dominating tree species for plots 1 to 3 and 5 to 9 were planted as 2-year-old seedlings in 1974. For this study, the 0.24 ha monoculture Scots Pine stand (plot 9) was chosen and consisted of 289 trees before selective harvesting took place in late March, 2018. Ten of the trees are equipped with analog and digital dendrometers and one tree is capable of sending its daily dendrometric measurements to a twitter feed (https://treewatch.net/). Five trees are marked for long-term phenological analysis and three trees are artificially watered to field capacity twice a week during the growth season (see Figure 6).

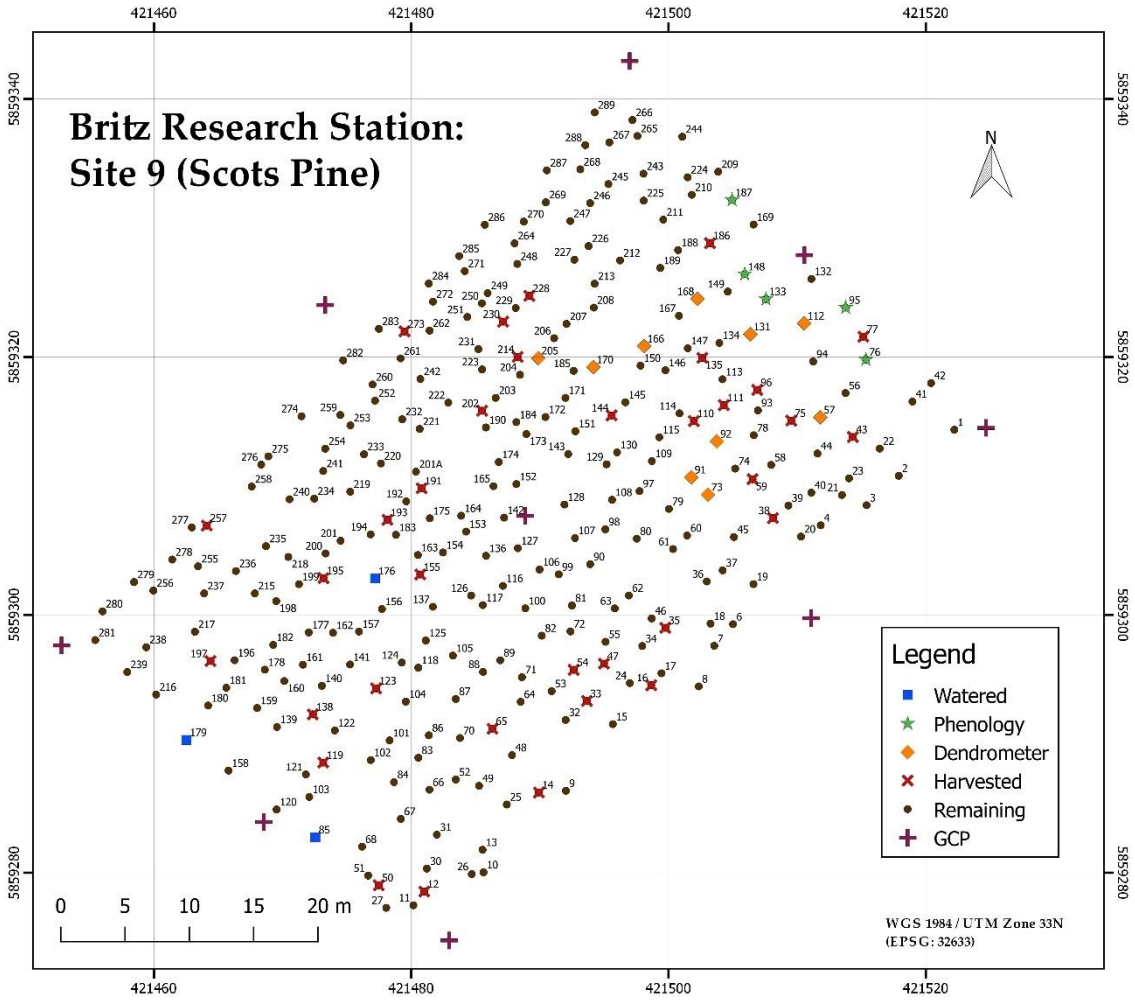


Figure 6: Overview of the 289 trees of the Scots Pine monoculture.

2.2.2 Field Measurements

Each individual tree height of the entire stand was measured indirectly prior to felling with a Vertex IV hypsometer (Haglöf, www.haglof.se) by an experienced forest inventory specialist from the *German Forest Inventory* (*Bundeswaldinventur*) in February, 2018. Field measurements were carried out prior to felling as to account for the entire stand including the trees to be felled. Direct measurements were carried out in late March, 2018 after selective felling where the length of each tree was measured as the length on the ground from the base of the cut to the apical bud. The stump height was added to the length of the harvested tree while taking the undercut in to consideration.

2.2.3 Ground Control Points and Tree Stem Positions

Due to challenges in acquiring a fixed RTK-GNSS (*Topcon GR-3*) position under a dense forest canopy, the GCPs were measured with a *Sokkia Set 5*$_{10}$ TS. Three GCPs were recorded on an open field adjacent to the study plot (see Figure 5) where a fixed RTK-GNSS signal was possible. The TS was then positioned on a GCP and the azimuth was calibrated using a second GCP. The third GCP was used as a control. Following the TS setup process, each of the 9 GCPs for the study plot where measured with the TS. Special care was taken in insuring the visibility between the TS and the target prism in relation with the aerial nadir visibility of GCPs through small gaps in the canopy. This procedure was best carried out during the winter months as ground vegetation and lower story deciduous trees can hinder visibility. Concurrent with the GCP measurements, the individual tree stem positions were also recorded with the TS. Four different GCP positions were required to measure tree stems due to limitations in visibility.

2.2.4 Remotely Sensed Data Acquisition

Aerial UAV imagery was acquired using a 36-megapixel full format *Sony* A7r RGB camera with a fixed lens (35 mm *Zeiss Sonnar* 2.8). The camera was focused manually at the hyperfocal distance and fixed permanently to inhibit variations in focal length. Camera settings were as follows: Shutter speed: 1/650, F-stop: f8, ISO: Auto (100 – 800), Format: RAW, Intervalometer: 2 seconds. The sensor was mounted onboard an *OctoXL 6S12* Octocopter (see Figure 7) from *Mikrokopter* (*HiSystems GmbH*) and flown autonomously with a predetermined flight plan. The flight plan was comprised of seven flight lines (at nadir) calculated in accordance to a focal length of 35.3 mm, 7360 x 4912 sensor size (pixel), 80 % frontal- and side-overlap, and a flying height of 75 meters. The theoretical Ground Surface Distance (GSD) was estimated at 1 centimeter. Images were triggered with an in-camera intervalometer calculated to shoot every two seconds and was deemed essential in order to eliminate any miss firing or latency in waypoint activated triggering (Krause et al., 2016). The same flight plan was flown for the pre-harvest mission on the 13th of February, 2018 and the post-harvest mission on the 29th of April, 2018. Each mission was flown within ± 60 minutes of solar noon and at an average maximum wind speed of 2.25 m/s for the pre-harvest mission and 3.83 m/s after harvesting. The maximum average wind speed was determined from onsite wind speed measurements at a height of 18 meters.

Figure 7: OctoXL 6S12 Octocopter mounted with a fixed lens Sony A7r RGB camera.

2.2.5 Processing and Tree Height Extraction

The original raw images (*Sony ARW*) for both datasets were indirectly georeferenced (Siebert & Teizer, 2014) by synchronizing the image time stamps with the Octocopters log file. Images were preprocessed to match histograms and corrected for Vignetting before being converted to 16-bit Tagged Image File Format (TIFF). Point clouds and orthomosaics were generated using the Structure from Motion (SfM) software *Pix4D* version 4.1.25 (*Pix4D*, Lausanne, Switzerland). At least five calibrated images were manually marked for each GCP and initial processing settings were set to *Full* image scale with *Geometrically Verified Matching* activated. The point clouds were processed at an *Optimal* point density with the image scale set at half image size.

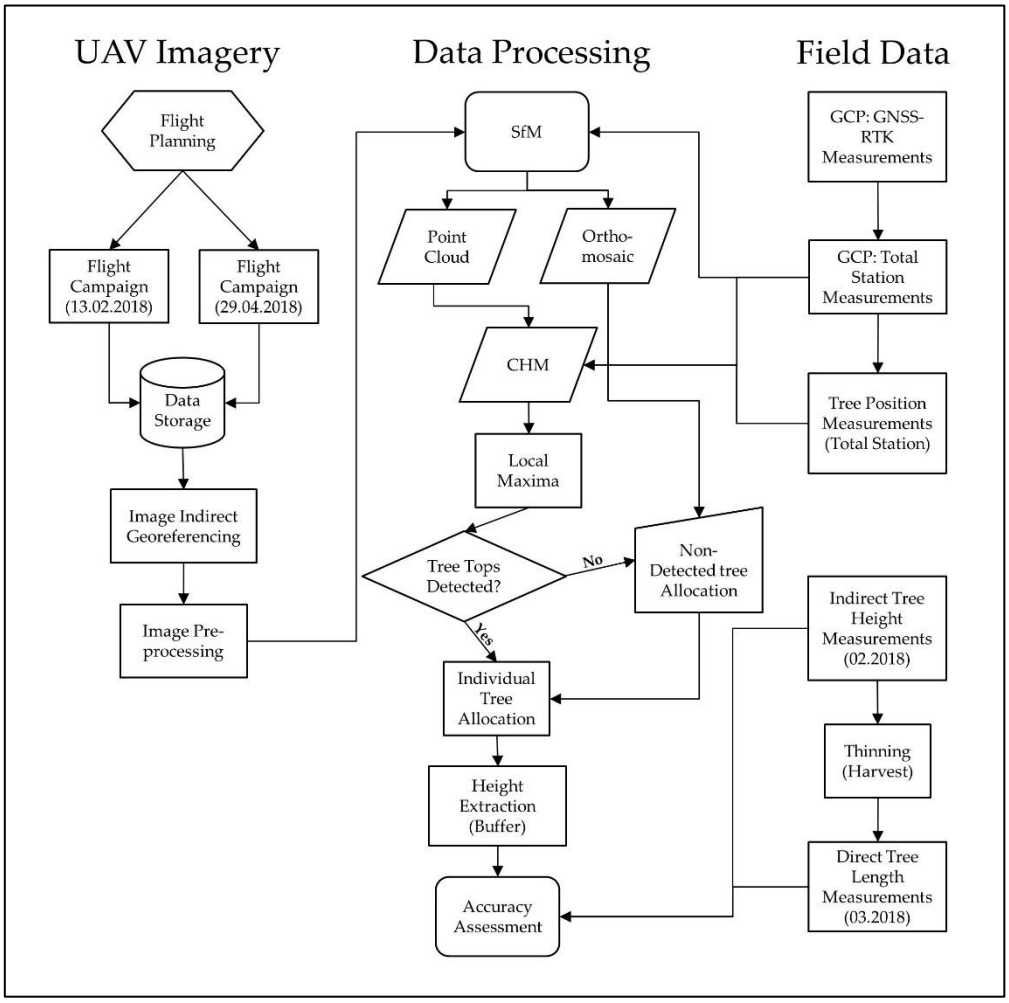


Figure 8. UAV image processing workflow in relation to the image acquisition and field data collection campaigns.

Point clouds and derivatives were processed with *FUSION/LDV* (McGaughey, 2018), *rLiDAR* and *lidR* packages in the *R* software package (R Core Team, 2022; Roussel & Auty, 2018; Silva et al., 2015). A Digital Terrain Model (DTM) raster was created by interpolating (nearest neighbor) a combination of the tree positions and GCPs and implemented to calculate the Canopy Height Models (CHMs). The DTM and CHMs were created at a resolution of 0.5 m and 0.3 m respectively. Treetops were detected using the local maximum algorithm and allocated an ID in reference to the TS tree positions. Missed or falsely detected trees were found by manually adjusting TS tree positions with aid of the Orthomosaic and CHM (Figure 9) and creating a 0.1 m buffer to capture the maximum pixel values. Typically, we have found that at least 80% of trees will be automatically detected which is in accordance with Mohan et al. (Mohan et al., 2017). Simply creating buffers around the TS positions to extract tree heights was not possible due an offset of the apical meristem to the stem position especially evident with codominant trees. Figure 8 shows the full processing workflow in relation to the image acquisition and field data collection campaigns.

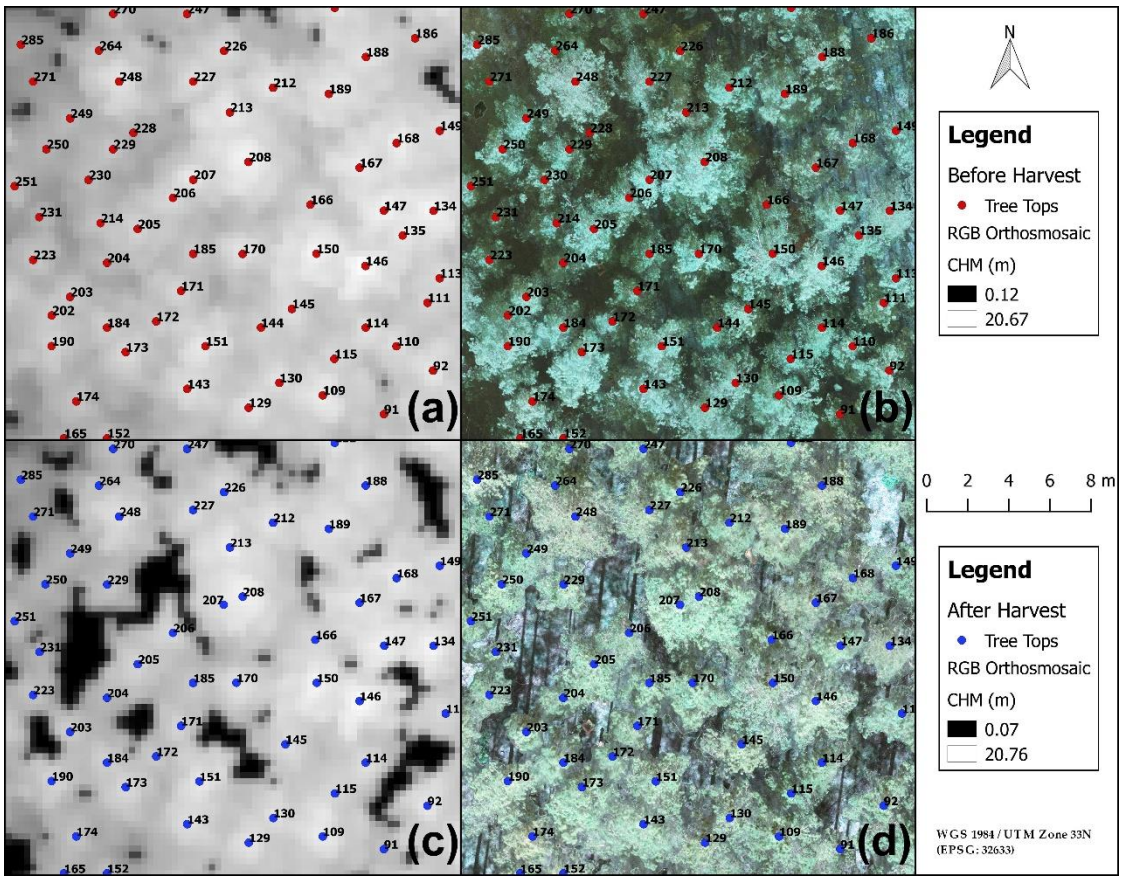


Figure 9. The majority of the treetops (> 80%) were detected with the local maxima algorithm. Non-detected trees required positional adjustments dependent on the highest pixel values in proximity to the tree stem positions and visual assessment of the tree crowns in the Orthomosaic. (a) The CHM before harvesting with automatically detected and adjusted tree tops. (b) Treetop positions shown over the Orthomosaic before harvesting (c) CHM after harvesting with noticeable gaps in the canopy (d) Orthomosaic after harvesting.

2.2.6 Statistical Analysis

The statistical analysis in this study was carried out with the *R* software package (R Core Team, 2022) and based upon the validation of the tree heights derived from the two photogrammetric datasets acquired before and after harvesting as well as a validation of the field-based indirect measurements. The direct measurements of the felled trees were used as validation for the indirect field measurements as well as the photogrammetric dataset before harvesting. The indirect field measurements were implemented as validation for all tree heights of the photogrammetric dataset prior to harvesting as well as the photogrammetric dataset containing the remaining trees after harvesting.

The datasets to be validated were assessed by quantifying the total error, systematic error, and random error (Larjavaara & Muller-Landau, 2013) as well as calculating the coefficient of determination R-squared (Matasci et al., 2018). We adopted the method from Larjavaara et al. (Larjavaara & Muller-Landau, 2013) and use the terms systematic error (bias) and random error

in order to clarify the difference between precision (lower random error) and accuracy (lower systematic error). Total error is determined by calculating the root mean squared error (RMSE):

$$\boldsymbol{RMSE} = \sqrt{\frac{\mathbf{1}}{\boldsymbol{n}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}(\boldsymbol{y_i} - \boldsymbol{\hat{y}_i})^{\mathbf{2}}}, \quad (1)$$

$$\boldsymbol{RMSE}\ [\%] = \frac{\boldsymbol{RMSE}}{\boldsymbol{\bar{y}}} * \mathbf{100}, \quad (2)$$

where $\boldsymbol{y_i}$ = predicted values, $\boldsymbol{\hat{y}_i}$ = observed values, $\boldsymbol{\bar{y}}$ = mean of observed *n* values, $\boldsymbol{n}$ = total number of samples (Larjavaara & Muller-Landau, 2013; Straub et al., 2013; Vastaranta et al., 2014; J. White et al., 2015). The mean error (ME) is synonymous with the systematic error and calculated as following:

$$\boldsymbol{ME} = \frac{\mathbf{1}}{\boldsymbol{n}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}(\boldsymbol{y_i} - \boldsymbol{\hat{y}_i}), \quad (3)$$

$$\boldsymbol{ME}\ [\%] = \frac{\boldsymbol{ME}}{\boldsymbol{\bar{y}}} * \mathbf{100}. \quad (4)$$

The standard deviation of error (SDE) depicts the random error and is calculated as following:

$$\boldsymbol{SDE} = \boldsymbol{SD}(\boldsymbol{y_i} - \boldsymbol{\hat{y}_i}), \quad (5)$$

$$\boldsymbol{SDE}\ [\%] = \frac{\boldsymbol{SDE}}{\boldsymbol{\bar{y}}} * \mathbf{100}. \quad (6)$$

R-squared was determined by calculating a linear model and provides a measure in how well the data is fitted to the regression line.

2.3 Results

2.3.1 Processing Results and Statistical Analysis

The two resulting CHMs were named PPC_{pre} and PPC_{post} and represent the photogrammetrically calculated tree heights before and after harvesting. Table 2 displays some of the statistical parameters of the resulting point clouds and orthomosaics. Important to note is that the RMSE values represent only the relative accuracy with regard to the GCPs and calibrated images rather than absolute geolocation accuracy. Furthermore, a decrease in point density for the PPC_{post} dataset is most probably due to absent trees in the point cloud after harvesting.

Table 2: Overview of the resulting statistical parameters for the photogrammetric products before and after harvesting.

Mission	Calibrated Images	RMSE	GSD (cm/pixel)	Points	Point Density (m^3)
February 13th, 2018 (PPC_{pre})	143	0.009	0.90	27,576,865	2,731.39
April 29th, 2018 (PPC_{post})	161	0.01	1.02	27,171,260	1,727.18

The four main datasets for the study were labeled (see Table 3) and adjusted for sample size in terms of the harvested trees (n = 34), all trees of the stand prior to harvesting (n = 285) and the remaining trees after harvesting (n = 251). Table 4 displays the tree height datasets adjusted for sample size.

Table 3: The four main datasets for analysis.

	Description	*n*	Acq. Date
Direct	Direct measurements of harvested trees	34	March, 2018
$Field_{all}$	Field-based indirect measurements	285	February, 2018
PPC_{pre}	UAVP mission before harvesting	285	February, 2018
PPC_{post}	UAVP mission after harvesting	251	April, 2018

Table 4: Overview of the statistical results of the tree height datasets (in meters) adjusted in sample size for analysis. Fieldharv and PPCpre_harv were adjusted corresponding to direct measurements, Fieldpre_rem and PPCpre_rem were selected corresponding to the sample size after harvesting (PPCpost).

	Min (*m*)	**1st Qu.** (*m*)	**Median** (*m*)	**Mean** (*m*)	**3rd Qu.** (*m*)	**Max** (*m*)	***n***	**Orig. Dataset**
Direct	12.12	15.74	16.60	16.58	17.53	19.60	34	-
$Field_{harv}$	12.50	15.85	16.85	16.72	17.95	19.50	34	**$Field_{all}$**
PPC_{pre_harv}	13.00	15.54	16.37	16.44	17.43	19.20	34	**PPC_{pre}**
$Field_{all}$	11.10	16.40	17.60	17.58	18.80	21.90	285	-
PPC_{pre}	10.37	16.07	17.25	17.22	18.43	21.12	285	-
$Field_{all\_rem}$	11.10	16.45	17.80	17.70	19.00	21.90	251	**$Field_{all}$**
PPC_{pre_rem}	10.37	16.15	17.34	17.32	18.54	21.12	251	**PPC_{pre}**
PPC_{post}	10.45	16.10	17.36	17.33	18.62	21.09	251	-

The validation of indirect field measurements against the direct measurements was labeled $Field_{harv}$_Direct and the validation of the photogrammetric point cloud (PPC) derived tree heights against the direct measurements was labelled PPC_{pre_harv}_Direct. Both comparisons implemented the 34 felled trees as samples. The photogrammetric tree height dataset prior to harvesting (PPC_{pre}_$Field_{all}$) was compared to the indirect field measurements for the entire stand (n = 285) as well as with an adjusted sample size (n = 251) excluding the harvested trees (PPC_{pre_rem}_$Field_{all_rem}$). This was done as to enable a comparison of the photogrammetric datasets before and after harvesting (PPC_{pre_rem}_$Field_{all_rem}$ and PPC_{post}_$Field_{all_rem}$). Finally, both photogrammetric tree height datasets containing the remaining 251 trees before and after thinning were compared to each other (PPC_{pre_rem}_PPC_{post}). Table 5 gives an explanation of the paired comparisons and Table 6 displays the overall statistical results.

Table 5: Overview of the compared dataset pairs.

	Description	***n***
Field$_{harv}$_Direct	Validation of field measurements against direct measurements of harvested trees.	34
PPC$_{pre_harv}$_Direct	Validation of the photogrammetric tree heights (February 13th, 2018) against direct measurements of harvested trees.	34
PPC$_{pre}$_Field$_{all}$	Validation of the photogrammetric tree heights (February 13th, 2018) against field measurements including trees that were subsequently harvested.	285
PPC$_{pre_rem}$_Field$_{all_rem}$	Validation of the photogrammetric tree heights (February 13th, 2018) against field measurements excluding trees that were subsequently harvested.	251
PPC$_{post}$_Field$_{all_rem}$	Validation of the photogrammetric tree heights (April 29th, 2018) against field measurements excluding harvested trees of the field-based measurements dataset.	251
PPC$_{pre_rem}$_PPC$_{post}$	Comparison of the photogrammetric tree heights before (February 13th, 2018) harvesting excluding harvested trees and after (April 29th, 2018) harvesting.	251

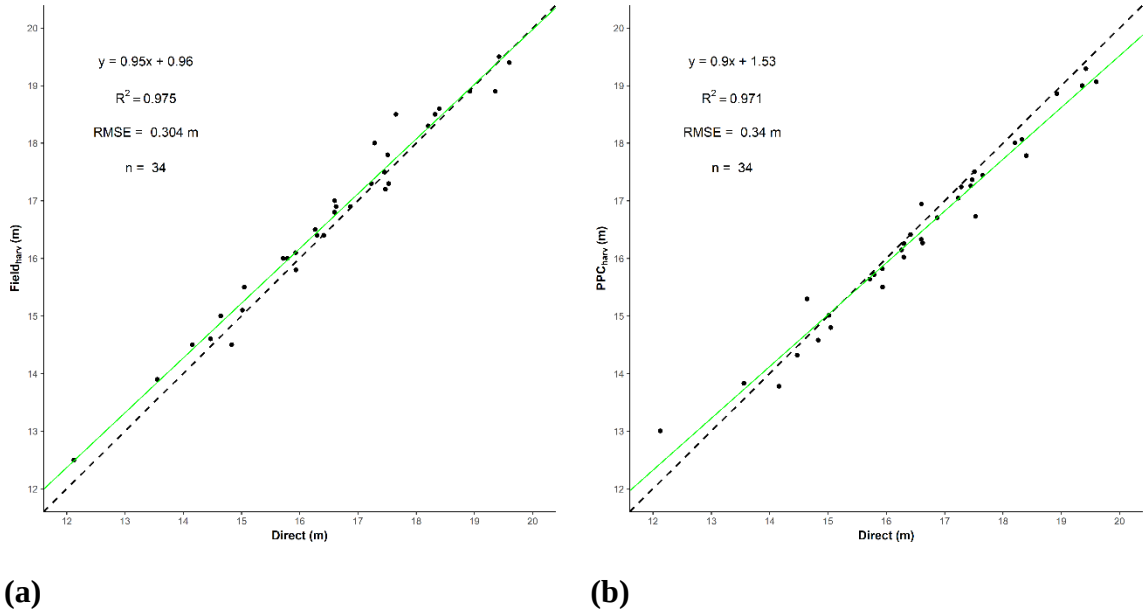
Table 6: Overall statistical results of the compared tree height datasets.

Dataset Comparison	**Total Error**		**Systematic Error (Bias)**		**Random Error**		**Correlation**		**Total**
	RMSE (m)	***RMSE (%)***	***ME (m)***	***ME (%)***	***SDE (m)***	***SDE (%)***	***R^2***	***p-value***	***n***
Field$_{harv}$_Direct	0.304	1.82	0.144	0.86	0.271	1.62	0.975	2.2e-16	34
PPC$_{pre_harv}$_Direct	0.34	2.07	-0.134	-0.81	0.317	1.93	0.971	2.2e-16	34
PPC$_{pre}$_Field$_{all}$	0.479	2.78	-0.365	-2.12	0.311	1.81	0.967	2.2e-16	285
PPC$_{pre_rem}$_Field$_{all_rem}$	0.486	2.81	-0.377	-2.18	0.308	1.78	0.967	2.2e-16	251
PPC$_{post}$_Field$_{all_rem}$	0.477	2.75	-0.369	-2.13	0.302	1.74	0.968	2.2e-16	251
PPC$_{pre_rem}$_PPC$_{post}$	0.138	0.794	0.008	0.04	0.138	0.79	0.993	2.2e-16	251

2.3.2 Validation with Direct Measurements

The direct destructive measurement method of the felled trees was assumed the most accurate measurement possible within the framework of this study and was used as a control for the field measurements as well as the photogrammetric dataset prior to harvesting. When compared to the direct measurements (Direct), the indirect field measurements (Field$_{harv}$) resulted with a RMSE

of 0.304 m (1.82 %) and a systematic error of 0.144 m (0.86 %) showing an overestimation of tree heights. In comparison with the direct measurements (Direct), the photogrammetric measurements (PPC_{pre_harv}) fared slightly worse than the field measurements with a RMSE of 0.34 m (2.07 %) and a systematic error of -0.134 m (-0.81 %) suggesting an underestimation of tree heights. In terms of random error, the field measurements showed a lower standard deviation error with 0.271 m (1.72 %) as opposed to 0.317 m (1.93 %) for the photogrammetric dataset. Both regression lines were determined significant ($p < 0.001$) however, the $Field_{harv}$ data displays a better goodness of fit particularly with the tallest trees in comparison to (see Figure 10).



(a) **(b)**

Figure 10: (a) $Field_{harv}$_Direct (b) PPC_{pre_harv}_Direct

2.3.3 Validation with Indirect Measurements

When compared to the indirect field measurements ($Field_{all}$), the photogrammetric tree heights fared similarly in terms of total error and bias. PPC_{pre} resulted with a RMSE of 0.479 m (2.78 %) against $Field_{all}$ as well as a bias of -0.365 m (-2.21 %) and random error of 0.311m (1.81 %). Figure 11 shows an underestimation of the photogrammetric tree heights (n = 285) in comparison to the field measurements.

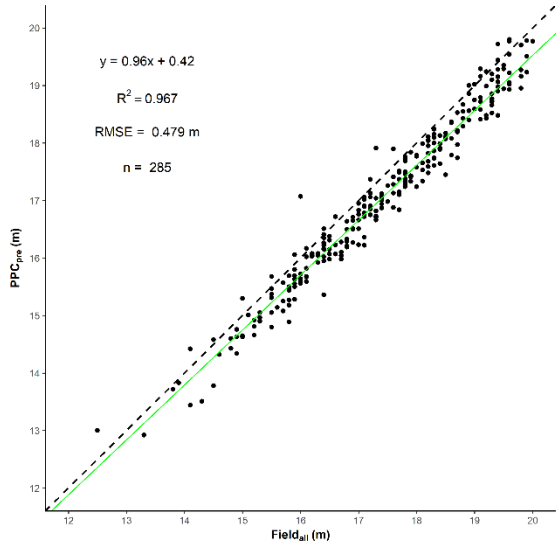


Figure 11: PPC_{pre}_$Field_{all}$

The photogrammetric datasets PPC_{pre_rem} and PPC_{post} before and after harvesting (n = 251) resulted with a RMSE of 0.486 m (2.81 %) and 0.477 m (2.75 %) respectively. Photogrammetric tree heights were also underestimated when compared to field measurements with a bias of -0.377 m (-2.18 %) for PPC_{pre_rem} and -0.369 m (-2.13 %) for PPC_{post}. As with Figure 11, Figure 12 also shows an underestimation of the photogrammetric tree heights (n = 251) in comparison to the field measurements. Important to note here is that the distribution of points below the regression line could also be partially attributable to an overestimation of field-based measurements as shown in Figure 10a.

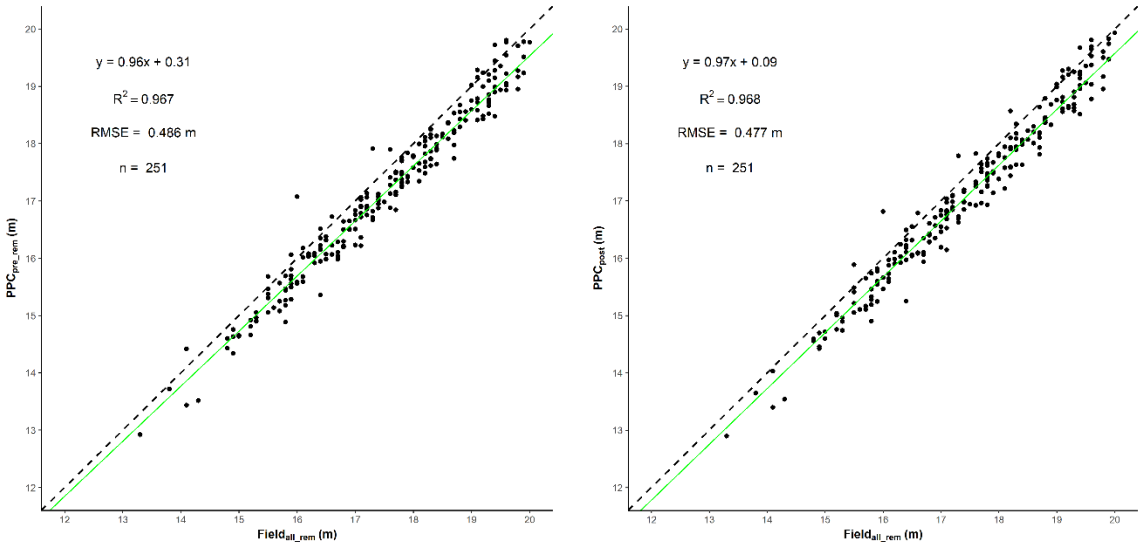


Figure 12: (a) PPC_{pre_rem}_$Field_{all_rem}$ (b) PPC_{post}_$Field_{all_rem}$

2.3.4 Comparison of Photogrammetric Datasets

Comparing both photogrammetric datasets (see Figure 13) resulted in a very low RMSE of 0.138 m (0.794 %). Systemic and Random error were also very low at 0.008 m (0.04 %) and 0.138 m (0.79 %) respectively. Additionally, the comparison showed a very high correlation between the two photogrammetric datasets (R^2 = 0.993).

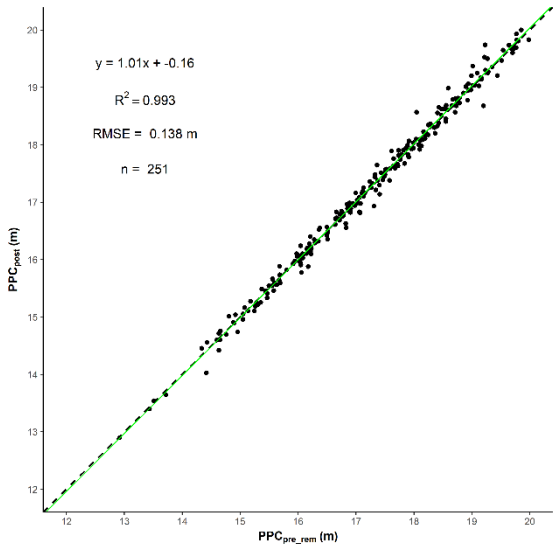


Figure 13: PPC_{pre_rem}_PPC_{post}

2.4 Discussion

Since the arrival of various modern remote sensing technologies such as ALS and UAVP, point cloud and raster derived tree heights have been typically validated against indirect field-based measurements (Apostol et al., 2016; Dempewolf et al., 2017; Falkowski et al., 2006; Gatziolis et al., 2010; Imai et al., 2004; G. Liu et al., 2018; Mielcarek et al., 2018; Panagiotidis et al., 2016; Popescu et al., 2002). Field-based tree height measurements should however be validated themselves as the measurement methods implemented alongside the experience and training of technicians could vary thus increasing uncertainty (Larjavaara & Muller-Landau, 2013). Furthermore, measurement errors in tree height can result in an increase in volume prediction error when volume functions involving diameter at breast height (*dbh*) and *h* are implemented (Arias-Rodil et al., 2017).

In this study, we showed that field-based measurements (RMSE = 0.30 m, 1.82 %) when measured by an experienced technician, can result with a similar RMSE as photogrammetric measurements (RMSE = 0.34 m, 2.07 %). Interestingly, both measurement methods showed

similar amounts of bias, however in opposing directions. The field-based measurements were overestimated with a positive bias of 0.144 m and the photogrammetric tree heights were underestimated with a negative bias of -0.134 m. Systemic error in opposing directions could be attributable to the total error of the photogrammetric datasets (PPC_{pre}, PPC_{pre_rem}, PPC_{post}) when compared to the field measurements (approx. RMSE = 0.48 m). This shows that unknown opposing biases can convolute results when uncertainty exists within validation data. According to the instruction manual of the German Federal Carbon Inventory (*Kohlenstoffinventur 2017*), error tolerance for tree height measurement is set at ±5 % for coniferous trees and ±10 % for deciduous trees and a maximum error cut-off of ±2 m for both tree types (Riedel et al., 2016). This leaves the photogrammetrically derived tree heights from this study well within the error tolerance for a typical forest inventory when direct and indirect measurements are used as validation. Important to note here is that the level of error within a forest inventory in terms of tree height is traditionally controlled by a designated technician where however uncertainty in bias could also exist.

The trees selected for harvesting were chosen with the criteria of facilitating the further development of selected "future" trees. For this reason, the selected trees for harvesting were less dominant than neighboring future trees and were often part of the lower portions of the canopy. Such positioned trees are often more challenging for photogrammetric tree height extraction and can typically result in overestimation due to occlusion from the branches of the more dominant neighboring trees (Baltsavias, 1999; St-Onge et al., 2015). Overestimation was however, not the case in this study, as underestimation of the photogrammetric tree heights was more evident. Underestimation with the photogrammetric tree heights could have been attributable to the direct measurements of trees when measured on the ground after felling. Most trees, while standing will have at least a slight leaning angle as discovered when comparing the highest pixel values in the CHM to the tree stem positions. Tree angle can also increase with less dominant trees as they will often resort to growing in the direction of openings in the canopy. When such a tree is extended out on a horizontal surface, the length of the tree could vary to the standing height as the leaning angle is somewhat eliminated. This could explain the underestimation of the photogrammetric tree height measurements caused from the extension of the tree length on the ground during direct measurements. Corrections for underestimation in such cases could be solved with a correction calculation (Riedel et al., 2016) based on the leaning angle of a tree which could be derived from the planar distance of the tree top to the stem position as well as tree height. The overestimation of field-based measurements was also reported by Wang et al. (Y. Wang et al., 2019) particularly for codominant trees.

The comparison of both photogrammetric datasets (PPC_{pre_rem}_PPC_{post}) was carried out in order to determine the repeatability and consistency among various flight missions in terms of tree height measurement. The results show an RMSE of 0.138 m between both separate flight missions which suggests a very high reliability for repeated measurements. Repeatability at this level shows a promising possibility to measure within-season tree height growth (Dempewolf et al., 2017) and could merit further study. Such studies could also prove relevant for fast growing short-rotation plantations. In terms of long-term intensive forest monitoring plots such as the Britz research station or Level II plots (Ferretti & Fischer, 2013), within-season growth could be a logical enhancement to match highly accurate tree measurement devices such as electronic dendrometers (D. M. Drew & Downes, 2009) and sap flow sensors (Meinzer et al., 2005).

The added time and cost involved with measuring GCPs and individual tree stem positions with the TS could prove feasible for long-term intensive monitoring plots. Permanent GCPs as well as tree positions are not only useful for constructing a reliable DTM for multi-temporal tree height extraction, but also for accurate crown positioning required for phenological observations. Additionally, individual tree positions remain valid throughout an individual tree's life span enabling the possibility for long-term time-series analysis with various sensors and remote sensing platforms. GCPs could also be implemented for longer periods of time depending on the durability of permanent ground markings particularly during thinning procedures. Furthermore, such sites could be implemented for the testing and calibration of remote sensing technologies for the purpose of assessment (Gerke & Przybilla, 2016; Hakala et al., 2018; Kraft et al., 2016; Z. Liang et al., 2015; Zimmermann et al., 2017) as well as sensor fusion and upscaling (Azadeh Abdollahnejad et al., 2018; Dash et al., 2018; Puliti et al., 2018).

Further studies are required in the validation of photogrammetric tree heights with direct measurement for various coniferous and deciduous tree species as well as stands of diverse structures. Additionally, for studies which involve the temporal analysis of photogrammetric datasets acquired by UAVs, factors such as for example varying solar illuminance and prevailing wind speed will need to be taken into consideration and eventually benchmarked in the case that UAVP is incorporated into national and international inventory programs.

2.5 Conclusion

The focus of this study was to determine the possibility in achieving accurate and precise tree height measurements using UAV photogrammetry at an intensive forest monitoring site. We showed that photogrammetric measurements can attain similar accuracies to that of indirect field

measurements when destructive direct measurements are used as validation. It was found that compared to direct measurements, indirect field-based measurements tend to overestimate tree heights and UAV-based photogrammetric measurements have a tendency to underestimate tree heights. Such positive and negative biases show a potential for a misinterpretation of results when photogrammetric tree heights are validated with field measurements, especially when biases are unknown. We also discovered a very high precision between photogrammetric tree height measurements derived from two separate UAV missions which could suggest a possibility to obtain tree height increment within a growth season. The authors recommend further research in the validation of UAV-based photogrammetric tree measurements for various coniferous and deciduous tree species in stands of diverse structures. Furthermore, the possibility to measure within-season tree height growth merits further study.

3 Mapping Tree Water Deficit with UAV Thermal Imaging and Meteorological Data

Abstract

The mapping of forest stands and individual trees affected by drought stress is an important step in targeted forest management with the aim of creating resilient and diverse forests. UAV-based thermal sensing is a promising method to acquire high-resolution thermal data, yet the performance of typical UAV-adapted low-cost sensors are somewhat limited in deriving accurate temperature measurements. Uncertainty is evident in the effects of internal sensor dynamics as well as environmental variables such as solar radiation intensities, relative humidity, object emissivity and wind to name a few. Furthermore, the accurate assessment of drought stress in trees is challenging to quantify, and typical research station methods can be laborious and cost-intensive and particularly challenging when carried out in the field. In this study, we explored the possibility to acquire reliable tree canopy temperature using the thermal band of the Micasense Altum multispectral sensor while examining the prospect of quantifying drought stress by implementing point dendrometers and UAV-derived tree canopy temperature to model Tree Water Deficit (TWD). In an indoor environment we showed that the usage of a limited number of pixels (<3) can result in temperature errors of over 1 K whereas increasing the spot size can reduce the mean difference to 0.02 K when using leaf temperature sensors as validation. Interestingly, leaves which were subjected to drought treatment (unwatered) resulted in a higher root mean squared error ((RMSE) (RMSE = 0.66 K and 0.73 K) than watered leaves (RMSE = 0.55 K and 0.53 K) due to most probably a lower emissivity of the dryer leaves. In a comparison of field acquisition methods, measuring the tree crown temperature of a selected tree from various incidence angles derived from typical gridded flights resulted in a mean standard deviation (SD) of 0.25 K and a maximum SD of 0.59 K (n = 12), where as a close-range hovering method resulted in a mean SD of 0.09 K and a maximum SD of 0.1 K (n = 8). Modelling the TWD from meteorological and point dendrometer data from the 2021 growth season (n = 2928) resulted with a R2 = 0.667 using a Generalised Additive Model (GAM) with the Vapor Pressure Deficit (VPD), wind speed and solar radiation as input features and a point dendrometer lag of one hour. When predicting individual tree TWD with UAV-derived tree canopy temperature, relative humidity and air temperature as input features, a RMSE of 4.92 (µm) and R2 of 0.87 was achieved with a GAM. The GAM with the Leaf-to-Air Pressure Deficit (LVPD) as an input feature resulted in a RMSE of 6.87 (µm) and a R2 of 0.71. This study presents a promising

method to acquire thermal data for the purpose of mapping TWD of beech on an individual tree basis. Further testing and development is an imperative and more drought period point dendrometer data as well as higher resolution meteorological data is required.

3.1 Introduction

Trees evolved strategies in persevering moderate drought episodes through physiological and morphological adaptations which aid in balancing cooling mechanisms in the crown while preventing excessive water loss and carbon starvation. Such adaptions involve the regulation of stomatal conductance, smaller leaf surface area and solar tracking (Mocko et al., 2017; Nobel, 2020), as well as wood traits which enable xylem to be resilient to hydraulic failure (Choat et al., 2012). Drought periods, even in a moderate form, increase the possibility for mortality (van Mantgem & Stephenson, 2007) and account for at least a reduction in growth (Brzostek et al., 2014; Nathalie et al., 2006) even in the absence of climate change (Allen et al., 2010). Recent extreme drought occurrences such as experienced in Europe in 2018 and 2019 (Leuschner, 2020; Senf & Seidl, 2021) due to climate change induced warming and shifts in precipitation patterns (Moran et al., 2017) are of raising concern in terms of amplified tree mortality and die-off regardless of climate zone (Schuldt et al., 2020).

European Beech (*Fagus sylvatica*), due to its extremely high shade tolerance, can under optimal conditions outcompete every other tree species in many parts of Europe (Kutsch et al., 2009). In recent years however, beech has been declining in growth throughout Europe (Delaporte et al., 2016; Martinez del Castillo et al., 2022; Tomasella et al., 2018; van der Werf et al., 2007) yet could gradually acclimatize to drought over time (Pretzsch et al., 2020).

Typical drought episodes can result in a reduction in carbon uptake due to a decrease in stomatal conductance, premature leaf senescence (Hartmann et al., 2021; Raspe et al., 2004) as well as a decrease in foliage the following year due to a reduction of available buds (Roloff, 1988). Prolonged drought episodes, however, can cause irreversible damage especially for anisohydric plant species in terms of xylem embolism (Tomasella et al., 2018) where permanent damage occurs to the hydraulic system (Garcia-Forner et al., 2017; Hartmann, 2011).

An approach in assessing drought-tolerant species is through the classification of a trees' hydraulic strategy in terms of the anisohydric and isohydric spectrum (Berger-Landefeldt, 1936; Hartmann et al., 2021). Despite numerous studies in the hydric behaviour of trees, there is no mathematical model describing this plant trait and is typically categorized in reference to the

relationship of stomatal conductance g_s and leaf water potential Ψ_1 (Klein, 2014). The strategy of isohydric plants is typically known for a reduction in transpiration with closing of the stomata during water shortage with the consequence of reduced CO_2 assimilation (Sade et al., 2012). On the other hand, risk-taking anisohydric plants leave the stomata open for longer periods despite water shortage making them more vulnerable to hydraulic failure yet maintain a higher CO_2 uptake during drought episodes (Burkhardt & Pariyar, 2015; Leuschner, 2020). The anisohydric strategy would in effect require more water to keep leaves cool during extreme heat and rely on substantial fluctuations of tree stem (xylem) water content with a reliance on nocturnal refilling (Yi et al., 2017). Within each species, variations in hydric behaviour can also occur depending on genetic variation in terms of drought stress tolerance (Leuschner, 2020; Moran et al., 2017) where species with high phenotypic plasticity could allow for individuals to acclimatise to changing climate conditions (Tomasella et al., 2018). Categorizing tree species and even provenances (Moran et al., 2017) into hydric behavioural classes by means of the quantification of stomatal conductance with gas exchange measurements and leaf temperature could aid in the assessment of drought stress tolerance to climate change. Such a classification should however, not be assumed, and a comprehensive holistic approach (Leuschner, 2020) in terms of whole-tree carbon balance is recommended (Garcia-Forner et al., 2017). Practically speaking however, central European species are rarely either aniso- or isohydric, but rather estimated in reference to other species. For example, *Quercus* species would be typically more anisohydric than *Fagus*, and *Pinus* would be more isohydric than *Fagus*. A better understanding of hydric behaviour among species and even individuals mapped at the regional scale could aid in focussing forest management goals.

An increased awareness of the effects of drought on tree productivity and survival (Pretzsch et al., 2020) is required if we are to select appropriate species and provenances for the purpose of improving silvicultural practices (Bolte et al., 2009) in terms of drought stress adaptation. The effects of prolonged and extreme drought conditions on forests in the future is however relatively unknown, and the ability for trees to acclimate is underestimated (Lapenis et al., 2005; Pretzsch et al., 2020; Reich et al., 2016). Pretzsch et al. (2020) showed that over a 5-year experiment of induced drought, beech acclimated faster than spruce, while spruce acclimatised faster when mixed with beech. This suggests that some species could acclimatise to extended drought stress over time within a generation providing hydraulic failure does not occur. The detection of hydric behaviour could aid in determining tree species mixing strategies as in which species co-exist well during drought as well as an assessment of hydric variability within species.

The use of thermal infrared (TIR) sensors have proven useful for non-destructive water content retrieval and stomata closure detection in plants (Cohen et al., 2005; Feller, 2016; Grant et al., 2006). Recent developments in UAV (Unmanned Aerial Vehicle)-mounted sensors provide the opportunity to acquire thermal imagery from above the crop or forest canopy. Gómez-Candón et al. (2016) used UAV-based thermal imagery to detect higher canopy temperatures in non-irrigated trees while using reference ground targets for temperature accuracy validation. For the purpose of mapping surface energy and water fluxes, Simpson et al. (2022) implemented UAS thermography and multispectral data to produce evapotranspiration maps for oak trees. The accuracy of low-cost thermal imagers can however prove challenging and various studies have been carried out to assess and develop thermal imaging acquisition methods (Acorsi et al., 2020; Aragon et al., 2020; Kelly et al., 2019; Perich et al., 2020; Ribeiro-Gomes et al., 2017; Smigaj et al., 2017; Wan et al., 2021; Zakrzewska et al., 2022). Challenges affecting thermal imaging accuracy can arise from the influence of meteorological variables such as air temperature (AT), relative humidity (RH), solar radiation (SR), and wind speed (WS) which will not only affect the tree canopy temperature but also influence the sensor itself. Furthermore, other issues exist which can affect thermal sensors such as sensor drift, internal calibration, "bad pixels", "spot-size", leaf angle due to solar tracking and the exclusion of relevant pixels through masking.

The aim of this study is to explore the possibility to acquire accurate thermal imagery at the individual tree level as would be of interest for intensive forest monitoring plots (i.e. ICP forests Level II). Here we developed a single shot method with the *Micasense Altum* sensor for the acquisition of tree crown temperature for the purpose of calculating the Leaf-to-air vapor press deficit (LVPD) and modelling TWD. Additionally, we test various acquisition and pixel extraction methods with the aim of minimizing error propagation. Using indoor and outdoor experiments we assess the dispersion of acquired thermal data while implemented leaf temperature sensors for validation. The specific aims of the study are as follows:

i. Determine the minimum number of pixels required to obtain accurate temperature and whether dry vegetation with lower emissivity will affect accuracy.

ii. Test the accuracy of tree crown temperature to upper canopy mounted leaf temperature sensors over repeated missions under varying weather conditions throughout the growth season.

iii. Evaluate the dispersion of TIR values acquired with grid type against single shot acquisition methods.

iv. Assess the possibility to model the TWD with TIR tree crown data and meteorological data.

3.2 Materials and Methods

3.2.1 Study Area

The Britz Research Station is located approximately 50 km north-east of Berlin (52.87° N 13.83 ° S, 42 m above sea level) and under management of the *Thünen Institute for Forest Ecosystems* (*www.thuenen.de*). Established in 1972, the research station was originally designed for forest hydrology research and has expanded to explore multiple aspects related to intensive forest monitoring. Since 2018 the research station has incorporated UAV technology alongside multispectral sensors for research into the applications for intensive forest monitoring (i.e. Level II), including full tree geometric measurements, phenology, growth, Leaf-Area-Index, plot mapping, drought stress assessment, and multispectral sensor validation and assessment. An overview map of the research station is available in Krause et al. (2019).

The beech stand for this study is approximately 50 years old and the nine selected trees for this study are long-term phenological observation trees mounted with six point dendrometers. Some of the trees also are equipped with analogue and digital band dendrometers as well as sap-flow sensors. Two of the selected trees for this study were equipped with leaf temperature sensors in the upper tree crown.

2.2.2 Altum Sensor

The sensor implemented for the study was the *Micasense Altum* (micasense.com). The sensor is comprised of six synchronised bands: blue, green, red, near-infrared, red-edge, and Longwave Thermal Infrared (LWIR). Table 7 displays the technical specifications of the *Altum* sensor. The radiometric-capable *Lepton* (*Lepton LWIR*) LWIR sensor from *FLIR Systems* (FLIR, 2022) has a wavelength range from 8 to 14 µm (center wavelength of 11 µm) and is integrated with thermal image processing functions such as automatic thermal environment compensation, noise filters and thermal non-uniform correction (NUC) (*Altum Integration Guide*, 2020). The NUC functions automatically recalibrates the sensor every five minutes or 2K changes in internal temperature. The thermal accuracy reported by the manufacturer is ± 5 K and thermal sensitivity

is < 50 mK (0.05K). For this study, data derived from the thermal band is named thermal infrared (TIR).

Table 7: Parameters for the Micasense Altum Sensor

Sensor	**Mega-pixel**	**Focal length [mm]**	**Pixel Size [µm]**	**Sensor Size [mm]**	**Sensor Size [pixel]**	**Aspect**	**File-type**	**GSD [cm] at 75 m**
Micasense Altum	3.2	8 (multi) 1.77 (LWIR)	4.25	7.16 x 5.35 (multi) 1.9 x 1.43 (LWIR)	2064 x 1544 160 x 120	4:3	tif	4.33

3.2.3 UAV

The UAV used for the study was a *DJI Matrice M210 RTK* with dual gimbals simultaneously carrying the *Micasense Altum* and *Zenmuse X7* (RGB) sensors. The hovering accuracy in P-mode with GNSS is reported by the manufacturer to be ± 0.5m vertical and ± 1.5 m horizontally. Maximum take-off weight is 6.14 Kg and flying time is approximately 24 minutes when also powering the *Altum* and *X7* sensors.

3.2.4 Flight planning

The *DJI Matrice M210 RTK* was controlled using the *DJI Pilot App* (www.dji.com). Flight plans were created based on the center of tree crown positions extracted from an existing Orthomoasic created from the *Zenmuse X7* and RTK-GNSS positioning. The single tree waypoint flight plan was created in *R* (R Core Team, 2022) by producing a custom *KML* (Keyhole Markup Language) file in the required *DJI* format. Flying height was set to at least 10 meters above the beech canopy which was typically a flying height of approximately 30 meters above ground. A minimum10-meter distance to the canopy was tested to be the closest possible distance where leaves where not immediately visibly affected by the downward airflow caused by the propellers. The single image waypoint flight plan (see Figure 14) required about 5-7 minutes to complete. Each waypoint was hovered over for 10 seconds enabling usually 5 images from the *Altum* sensor with an intervalometer of 2 seconds. A separate flight plan was created for the two trees with leaf temperature sensors and was usually repeated at least twice within a mission.

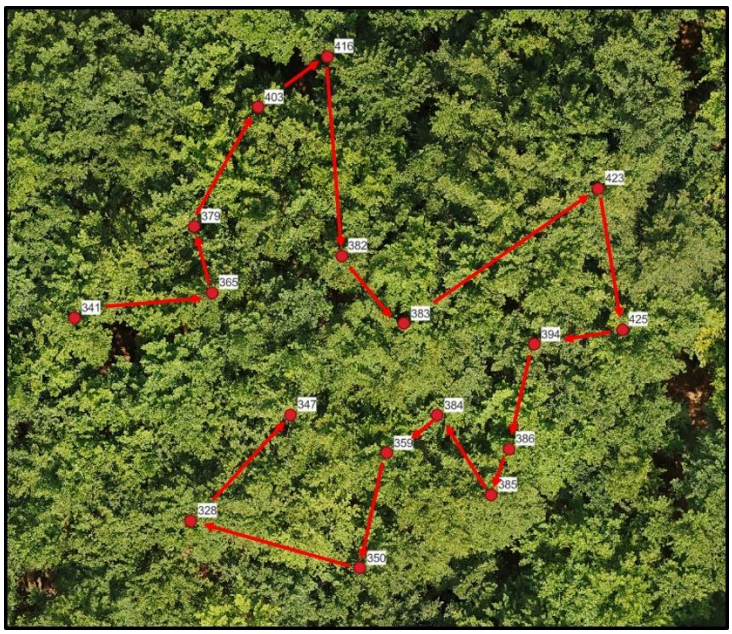


Figure 14: An example of a Waypoint flight plan for single shot images. The UAV would hover over each waypoint for 10 seconds while acquiring five thermal single shot images of the tree crown.

The gridded flight was the standard flight plan used for covering the complete Britz research station on a weekly basis. Here a forward- and side-overlap of 80 – 85 % was used at a flying height of 75 meters. Forward overlap was accomplished with a two second intervalometer (flying speed of 3 m/s) of the Altum sensor. The gridded flight plan takes approximately 30 minutes to complete for the entire research station which requires one battery change.

3.2.5 Research Station Sensors

The leaf temperature sensors used for leaf temperature validation were the *LAT-B2* sensors (Ecomatik, 2022) which measure the absolute leaf surface temperature and surrounding leaf ambient temperature (see Figure 15). According to the manufacturer, the leaf temperature sensors have an absolute accuracy of ± 0.2 K for both leaf surface and ambient temperature. The sensors are installed at the beginning of the growth season after the completion of all phenological phases for this location at the end of June.

Figure 15: The LAT-B2 (Ecomatik, 2022) leaf temperature sensor which was mounted in the upper tree crown and used for validating the TIR imagery.

The point dendrometers used in the study were of the *DR1W* type (Ecomatik, 2022) installed on six of the study trees. Point dendrometers were installed permanently before the growth season and are reported by the manufacturer to have an error of a maximum of 4.5 % of the measured value. The point dendrometers measure changes in the diameter of the tree at breast height (1.3 m) in the micrometer (µm) range and were set to record at five-minute intervals. Five of the six trees were chosen based on the completeness of data for the study.

The weather data was acquired from a *Thies* weather station (Thies, 2022) equipped to record meteorological data such as air temperature (AT), wind direction (WD), wind speed (WS), solar radiation (SR), relative humidity (RH) and air pressure (AP). The weather station is permanently positioned on an open field at two meters above ground and is approximately 180 meters from the beech study plot.

3.2.6 Feature Selection

Working with UAV-based multispectral imagery alongside near real-time meteorological data opens up many opportunities for modelling which is not typically possible when working with only pixel values alone. With the case of thermal imagery, comparing day to day raw thermal values can often be misleading. Thermal temperature values of the leaves of a plant or tree can vary within minutes especially during a hot summer day with varying cloud conditions. Incorporating meteorological variables such as for example RH, AT and SR synchronised at the time of thermal data acquisition could be incorporated to correct for various influences caused by varying weather conditions. This could not only be useful for the calibration of thermal

imagery but also for modelling tree water status such as the TWD (Drew et al., 2011; Zweifel et al., 2005).

The testing of these features in reference to their capabilities in predicting TWD is an imperative in finding the most accurate models possible with the available data. Some features such as the Vapor Pressure Deficit (VPD) derived from RH and AT, alongside soil moisture can alone be implemented to predict TWD (Zweifel et al., 2005). Incorporating UAV-based TIR imagery into TWD modelling could prove useful in enabling a spatial aspect to the model's predictions.

The use of feature engineering could also improve modelling performance by creating more effective data representations while eliminating multicollinearity and influences from irrelevant predictors (Kuhn & Johnson, 2019). A simple variation in changing the representation of a feature is that of the VPD which is shown to influence the atmospheric demand and stomata conductance (Grossiord et al., 2020; Massmann et al., 2019). Additionally, the leaf-to-air vapor pressure deficit (LVPD) incorporates the difference of the UAV-based TIR temperature and AT alongside the VPD. The LVPD can not only be a potential input feature for TWD modelling, but can also work as a stand-alone index depicting current tree crown status.

The VPD is the difference between the amount of moisture in the air and the capacity for air to hold moisture when saturated (Noh & Lee, 2022). The maximum amount of saturation is dependent on the AT, where the higher the AT, the higher the possible saturation, which is called the saturation vapor pressure (SVP) (see Equation 1). The SVP, when incorporated with RH gives the active vapor pressure (AVD) (see Equation 2) which is the actual amount of water vapor at a given AT. The VPD is then calculated by subtracting the SVP from the AVP as shown in Equation 3.

$$SVP = 0.6108 \exp\left(\frac{12.27AT}{AT+237.3}\right) \quad \text{(Equation 1)}$$

$$AVP = SVP\frac{RH}{100} \quad \text{(Equation 2)}$$

$$VPD = SVP - AVP \quad \text{(Equation 3)}$$

The LVPD is calculated by accounting for the difference between the water vapor pressure in the leaf, and the water vapor pressure of the ambient AT, (Z. Dai et al., 1992; Day, 2000;

Grossiord et al., 2020; Furtak & Nosalewicz, 2022). Here the leaf surface temperature is implemented to calculate the leaf vapor pressure (LVP) (see equation 4) instead of the ambient AT as in the AVP equation (see Equation 2). The LVPD is then calculated by subtracting the LVP from the AVP as shown in Equation 5.

$$LVP = SVP\frac{RH}{100} \quad \text{(Equation 4)}$$

$$LVPD = LVP - AVP \quad \text{(Equation 5)}$$

In this study, the leaf surface temperature was acquired via the UAV-based TIR imagery as well as leaf temperature sensors attached to the leaf directly. Table 8 shows an overview of the features and indices including abbreviations used in this study.

Table 8: List of feature and Indices with abbreviations and units.

Feature	**Abbreviation**	**Description**	**Unit**
Thermal Infrared	TIR	Thermal data acquired from the Micasense Altum	°C/K
Masked Thermal Infared	TIR_{mask}	TIR where ground and lower canopy pixels are removed	N/A
Ambient Air Temperature	AT	Weather station Britz	°C
Relative Humidity	RH	Weather station Britz	%
Solar Radiation	SR	Weather station Britz	W/m^2
Wind Speed	WS	Weather station Britz	m/s
Vapor Pressure Deficit	VPD	Derived from AT and RH	kPa
Leaf-to-air Vapor Pressure Deficit	LVPD	Derived from leaf TIR, AT and RH	kPa
Tree Water Deficit	TWD	Detrending and normalisation of point dendrometer data	µm

3.2.7 Tree Water Deficit as Drought Stress Validation

Daily diurnal stem diameter fluctuations occur due to the emptying of the stem caused by loss of water through the leaves typically during the day and the nocturnal refilling of the stem during the night. The magnitude of such stem size changes is an indication of the current tree water status and can be quantified by highly accurate point dendrometers positioned at breast height. Point dendrometers record stem size changes caused by new xylem and phloem tissues which accumulate throughout the growth season (D. Drew et al., 2011; Zweifel et al., 2005) as well as the swelling and shrinking of elastic tissues (Zweifel et al., 2005; Zweifel & Häsler, 2000). Figure 16 displays the course of daily fluctuations where stem shrinkage caused by water loss from transpiring leaves throughout the day is followed by nocturnal filling until the early morning. The red dotted line shows the time when the UAV missions for TIR acquisition was carried out.

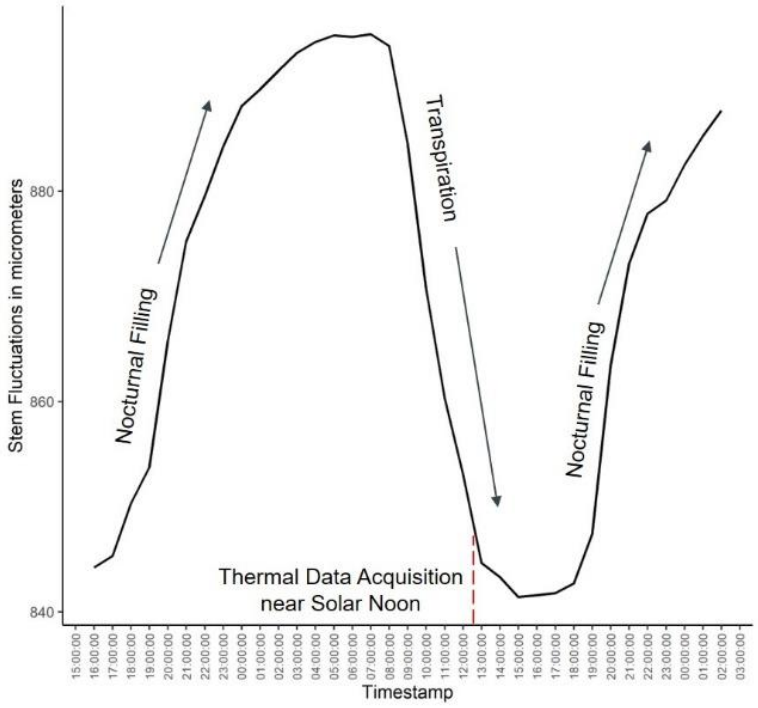


Figure 16: An example of the course of daily stem fluctuations. Nocturnal filling commences typically in the late afternoon/early evening and ends when the leaves begin transpiring in the morning. TIR data was typically acquired near solar noon which is close to when the stem is nearing its most empty state (red dotted line).

As shown in Figure 17a, and described in detail by Zweifel et al. (2016), segments between two stem radius maximum groupings are established throughout the growth season representing irreversible growth from the preceding maximum. The reversible changes in the stem ranges within this segment grouping is what is called the TWD. The growth increment is detrended, which was normalised and inversed to give values where the highest point of each curve represents the maximum stem shrinkage on a given day (see Figure 17b).

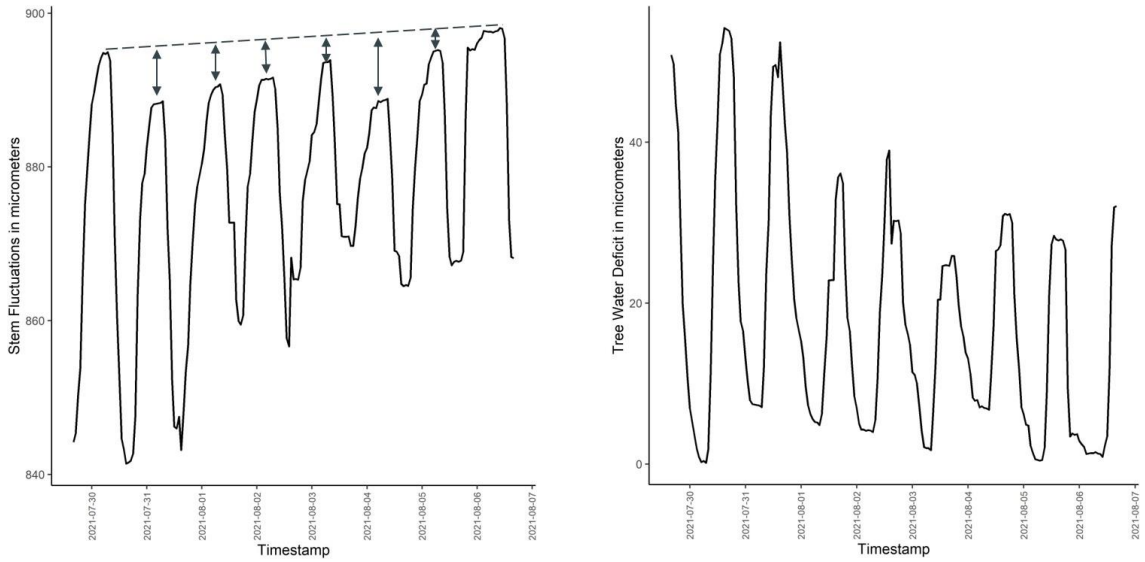


Figure 17: (a) The figure on the left shows a growth segment (dotted line) between two maximum stem radius instances. The arrows represent the amount the stem is in deficit in comparison to the growth trend. (b) The figure on the right is essentially the same segment detrended, inversed, and normalized so that the maximum shrinkage instances are shown at the top of the graph. Here the Tree Water Deficit can be determined on an hourly basis. The UAV missions were carried out typically at solar noon which would be about 3 to 6 hours before the stem has reached the maximum shrinkage for that day.

3.2.8 Image processing

The beneficial aspect of the *Micasense Altum* is the possibility to co-register the single shot TIR imagery with the synchronised multispectral bands. The non-thermal multispectral bands were converted to radiance and then radiometrically calibrated based on a Micasense calibrated reflectance panel which was sampled before and after each flight mission. The resulting images were then available with reflectance values from 0 to 1. Proceeding the radiometric calibration was the image alignment process which was comprised of unwarping the images using the built-in lens calibration, affine transformation, and the cropping and removal of extraneous pixels (micasense.com, 2022). The full process was carried out in *Python 3.6* and *Micasense* Python libraries (www.github.com/micasense/imageprocessing). A customised *Python* workflow was run in a loop for each corresponding calibration panel.

Further processing was carried out in *R* (R Core Team, 2022) where various vegetation indices were calculated for testing and masking purposes. The raw DN values of the TIR band are in centi-Kelvin directly from the sensor and converted to Celsius as recommended from the manufacturer with the following equation:

$$TIR\ °C = \left(\frac{DN}{100}\right) - 273.15 \qquad \text{Equation 6}$$

Images and derivatives were then organised in layers stacks and stored depending on mission ID numbers. Tree crowns were located manually in the single images and extracted using a rapid annotation method developed in *R* using the *Terra* package (Hijmans et al., 2022). Here the images were displayed in RGB or colour-infrared (CIR) and crown polygons were created by mouse clicking the top left of the tree crown and the bottom right of the tree crown forming the blue rectangle shown in Figure 18. A centroid of the rectangle polygon aided in establishing the length and width of the rectangle which was followed with the creation of an ellipsoid or sometimes near circle which depicted the tree crown. Due to a large number of images this process was semi-automised in a loop which would automatically move to the next image stack after user inputs. Polygons and cropped images were then stored with relevant naming conventions to insure repeatability.

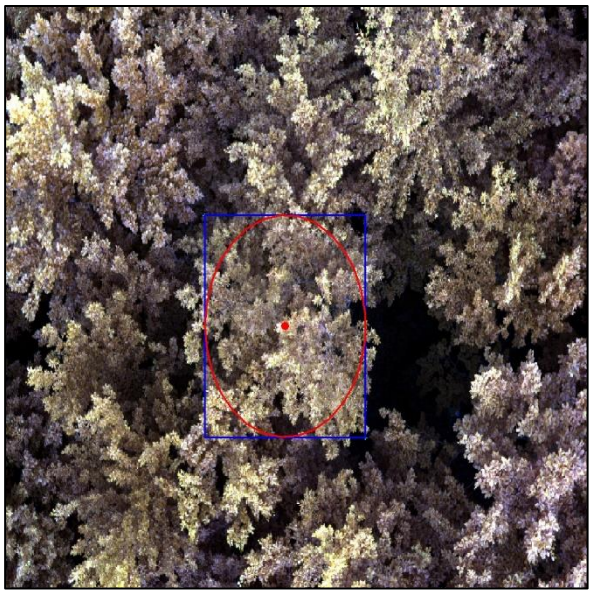


Figure 18: Rapid annotation method for the extraction of crown pixels in R (R Core Team, 2022). The resulting ellipsoid is created when the user mouse clicks on the top left and bottom right corners of the tree crown.

During the looping process, the mean extracted tree crown TIR, multispectral and vegetation index pixels values were appended to a table alongside acquisition timestamps using the *exiftoolr* package (JoshOBrien, 2021). Timestamps of the extracted pixel values were then rounded to the nearest hour and synchronised with the relevant hourly meteorological and dendrometer data from the weather station. The data from the dendrometers was originally acquired at 5-minute intervals and for this study only the calculated TWD on the hour was synchronised. Three lagging variation datasets were created in terms of TWD calculated from the dendrometer data: no lag, one-hour lag, and two-hour lag.

3.2.9 Analysis and Modelling

Statistical analysis and modelling was carried out in *R* (R Core Team, 2022). All datasets were tested for normalisation and correlation and modelling was adapted accordingly. For the correlation analysis, Pearson Correlation was implemented for the analysis of parametric data and the Spearman nonparametric measure of rank correlation was used for the data not following a normal distribution. For the purpose of averaging correlation results, correlation coefficients were transformed using the Fischer-Z method using the *fisherZ* function from the *DescTools* package (Signorell, 2022), averaged, and then converted back to correlation coefficients for further analysis (Corey et al., 1998). For the further analysis of non-parametric data, second order polynomial regression and Generalised Additive Models (GAM) were implemented. The GAMs were trained and validated using a 70/30 training and validation split and evaluated using the root-mean-squared error (RMSE), mean squared error (MAE), and R^2. Linear regression was also evaluated with RMSE, MAE, and R^2. Models were compared using the Akaike Information Criterion (AIC). Due to lacking actual drought-influenced dendrometer data, modelling was limited to simple curve regression models to enable potential extrapolation for analysis and proof of concept purposes rather than operational usage.

3.3 Results

3.3.1 Indoor Experiment

With the aim of assessing the *Altum Micasense* TIR band, extracted TIR values were compared with leaf temperature derived from mounted leaf temperature sensors. The indoor experiment involved using two indoor plants (*Epipremnum aureum*) mounted with two leaf temperature sensors per plant. The plants were chosen for practical reasons as leafed-out Beech is typically not available during the winter months in central Europe. Experiments were carried out over a range of 30-50 minutes with varying controlled timings and intensities of heat lamps and a ventilator. The sensor was set with an intervalometer at two second intervals and activated at the beginning of the experiment. The sensor was not specifically pre-warmed prior to the indoor experiments, but was in the same room for at least 30 minutes prior to operation. Heat lamps and ventilator were positioned as to not only influence the plants but also the sensor.

Thermal values were extracted from leaves in the proximity of the leaf temperature sensors (see Figure 19). Over three trials, three types of polygons were tested: all of the pixels of the full leaf were extracted, approximately 10 pixels around the leaf temperature sensors, and

approximately three pixels at the vicinity of the leaf temperature sensor. Table 9 shows the results of the three trials where the implementation of approximately 10 pixels (trial 2) to extract thermal values resulted in a mean difference of 0.02 K. Using pixels of the whole leaf and only three pixels for extraction resulted in a mean difference of 0.11 K and 1.18 K, respectively.

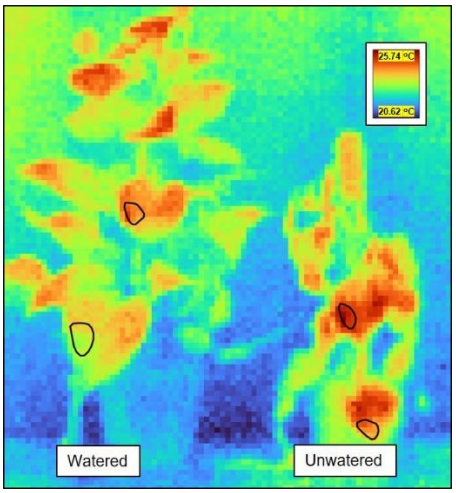


Figure 19: Thermal image from the Micasense Altum acquired during indoor experiments. The watered plant on the left side of the image contains leaf 1 (bottom) and leaf 2 (top) while the unwatered plant on the right has leaf 3 (bottom) and leaf 4 (top). For the second trial, the polygons were created in the proximity of the mounted leaf temperature sensors and are comprised of approximately 10 pixels each, which were used to extract the mean temperature at two second intervals.

In terms of accuracy assessment of the second trial, leaf one of the watered plant resulted with an RMSE of 0.55 K and MAE of 0.42 K and leaf two an RMSE of 0.53 K and MAE of 0.40 K. Both leaves of the watered plant resulted with an R^2 of 0.9. The unwatered plant resulted in a slightly higher RMSE of 0.66 K for leaf three and RMSE of 0.74 K for leaf four. A similar decrease in accuracy was found for the MAE of the unwatered plant with a MAE of 0.53 K for leaf three and a MAE of 0.59 K for leaf four. The R^2 for leaves three and four resulted in 0.88 and 0.87 respectively. Table 10 gives an overview of the RMSE, MAE, and R^2 of both plants for the second trial. Figure 20 shows the four separate leaves of the second trial modelled with polynomial regression of the second order. Similar patterns are evident among all four leaves due to analogous artificial environmental effects (i.e. heat lamps and ventilator). A wider dispersion is evident for the leaves of the unwatered plant, in particular leaf four.

Table 9: The mean difference in temperature for the three separate trials. The use of approximately 10 pixels created the best results while implementing only three pixels or less resulted in considerably more error.

Experiment	Method	Mean difference [K]
Trial 1	Whole Leaf	0.11
Trial 2	10 Pixels	0.02
Trial 3	3 Pixels	1.18

Table 10: RMSE, MAE and R^2 values for all four leaves of the two plants depicting the accuracy of temperature extraction from the thermal imagery compared to the leaf temperature sensor. The dryer leaves of the unwatered plant show increased error probably due to lower emissivity.

Object	RMSE [K]	MAE [K]	R^2	Treatment
Leaf 1	0.55	0.42	0.9	watered
Leaf 2	0.53	0.40	0.91	watered
Leaf 3	0.66	0.53	0.88	unwatered
Leaf 4	0.73	0.59	0.87	unwatered

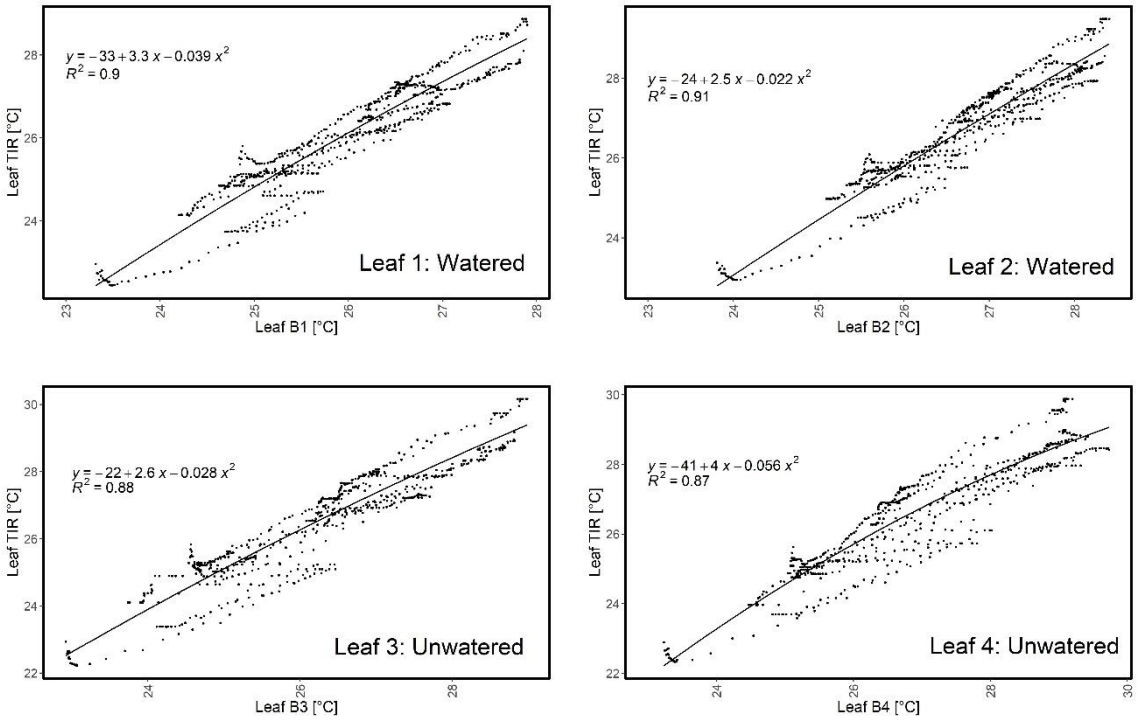


Figure 20: Results of the indoor experiment modelled using polynomial regression of the second order. Point distribution shows a lower R^2 and higher dispersion for the unwatered plant, in particular leaf 4. The variations of point clusters are due to controlled heat lamps and ventilator fluctuations as well as possible sensor drift.

The second trial lasted for approximately 50 minutes, where the plants and sensor were subjected to varying heat lamp and ventilator intensities. Figure 21, shows a visual comparison of the extracted thermal pixel values from leaf 4 in comparison to values derived from the leaf temperature sensor. The NIR band of the *Micasense Altum* sensor displays the timing of the heat lamp fluctuations. The spikes in the thermal imagery (*Leaf TIR 4)* are probably caused by the Thermal Non-Uniform Calibration (NUC) of the *Micasense Altum* sensor which as reported by the manufacturer, calibrates automatically every five minutes or when the temperature of the camera changes by 2 Kelvin (*Altum Integration Guide*, 2020).

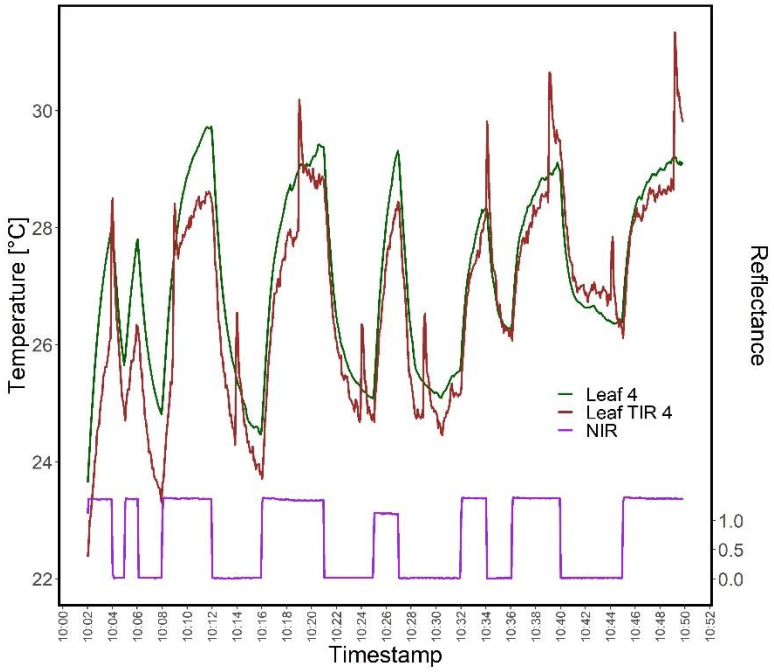


Figure 21: Visual comparison of the extracted thermal camera pixel values (Leaf TIR 4) and the leaf-temperature sensor. Spikes in the thermal dataset are most probably due to the Thermal Non-Uniform Calibration which occurs automatically at five-minute intervals.

3.3.2 Field Experiments

During the field experiments the thermal sensor was also assessed by comparing the mean crown temperature of two trees which had a leaf temperature sensor mounted in the upper tree crown. Though validation was limited to only one sensor in the upper canopy of each tree, a rough estimation of the tree's upper tree crown temperature was useful as a control in how the leaf surface was affected by varying metrological factors. Figure 22 displays the results of the comparison using linear regression (upper) as well as a time-based visual representation (lower). The average thermal values extracted from the crown of tree 328 over 13 acquisition days resulted in a RMSE of 3.31 K, a MAE of 2.95 K and a R-squared of 0.89 when compared to the

tree crown-based leaf temperature sensor. Thermal values extracted from tree 347 faired similarly with and RMSE of 3.12 K, MAE of 2.78 K and a R-squared of 0.93. Systematic error in terms of a somewhat consistent underestimation of leaf temperature is evident with the time-based visualisation for both trees.

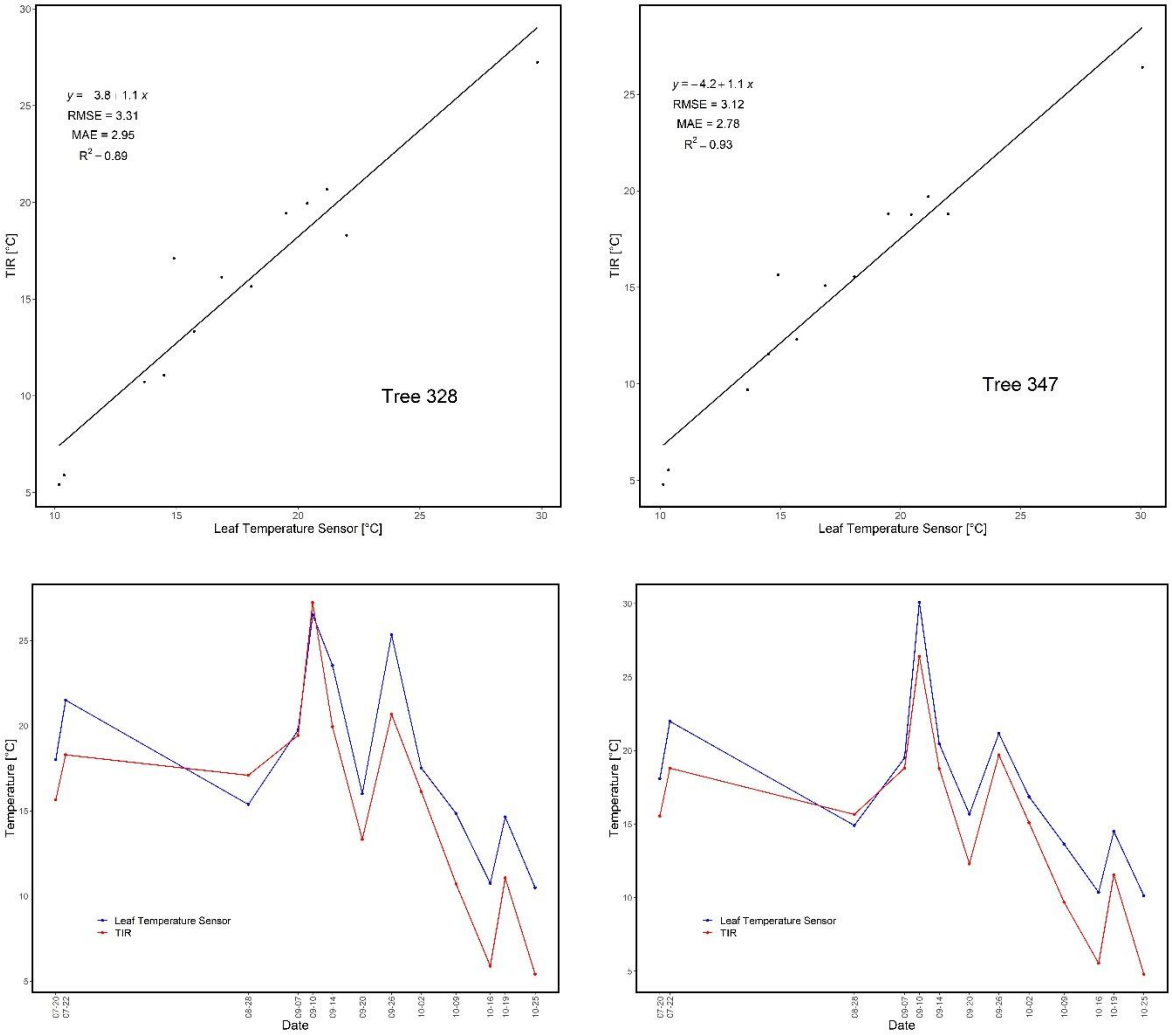


Figure 22: Linear regression and accuracy reporting (top) as well as a time-based visualisation (bottom) for tree 328 (left) and tree 347 (right)

For the purpose of assessing the accuracy requirements for a thermal sensor for tree crown temperature extraction throughout the growth season, the mean absolute difference for tree 328 and 347 was calculated across the full time series (n = 13). Tree 328 resulted with a mean absolute difference of 5.27 K while tree 347 resulted with similar mean absolute difference of 5.45 K. This shows that in order to consider the day-to-day differences at the stand level, a minimum of approximately 5 K accuracy could be necessary.

In order to assess the possibility to differentiate between extracted tree crown temperatures within a flight mission, the mean absolute difference was calculated for each epoch separately across the five trees equipped with point dendrometers resulting with a minimum of 0.7 K and a

maximum of 1.3 K. Figure 23 displays a correlation matrix for the mean absolute difference of extracted thermal values for the growth season in comparison to typically available hourly meteorological data. It can be seen that AT and VPD correlate moderately to the daily differences between tree crown temperatures while RH and WS are also considered but with minor importance. SR does not seem to be in an influencing factor this case. The masking of thermal pixels to the upper tree crown to eliminate shaded parts of the lower crown as well as ground pixels did not show any significant difference in correlation to the meteorological data except for a slight improvement for RH. On the other hand, the mean crown temperature from the thermal sensor correlates well with SR (0.68) whereas RH correlates moderately (0.49) and WS almost not at all. VPD and especially AT are highly correlated to the TIR crown temperature at 0.88 and 0.98 respectively.

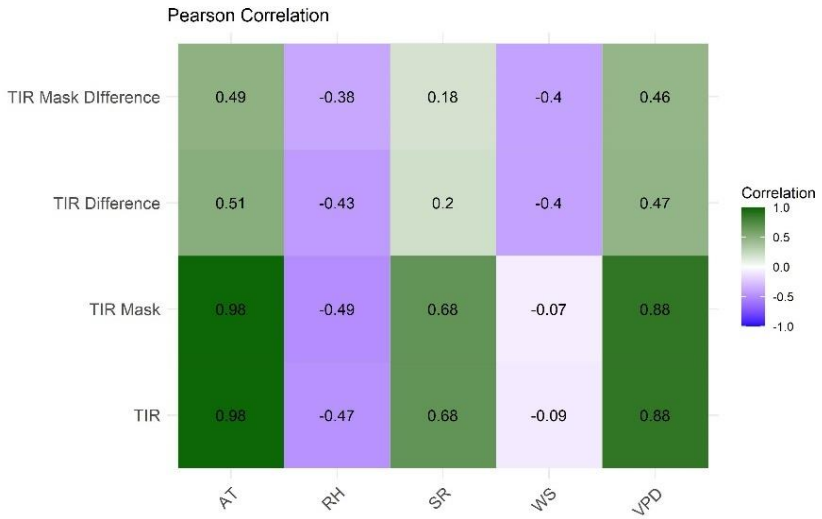


Figure 23: Pearson Correlation matrix showing the relationships between the extracted Thermal Infrared values for all of the tree crowns (n = 9) summarized for each acquisition day (n = 13). TIR = unmasked extracted TIR values, TIR Mask = masked extracted upper crown TIR values, Difference = the mean difference between extracted TIR values within an acquisition day.

3.3.3 Flight Grid Acquisition Method

Here two UAV-based acquisitions methods, flight grid and single shot, were evaluated for the purpose of assessing the effects of various viewing angles as well as timing on thermal imaging consistency. Four epochs from the growth season campaign were selected for comparison purposes. An overview of various meteorological data acquired within an hour of the TIR data acquisition is displayed in Figure 24 specific for each of the four epochs. Figure 24 also shows the relationship between the acquired TIR imagery of a sample tree crown in reference to the cloud cover on that particular day. All four days were subjected to varying cloud cover where day 240 had the most substantial cloud cover alongside low SR and high RH. The highest VPD

was evident on day 250 which coincided with the highest mean absolute difference of the thermal imagery (mdTIR) and a relatively high TWD. The lowest mdTIR was evident on day 257 which coincided with the highest SR.

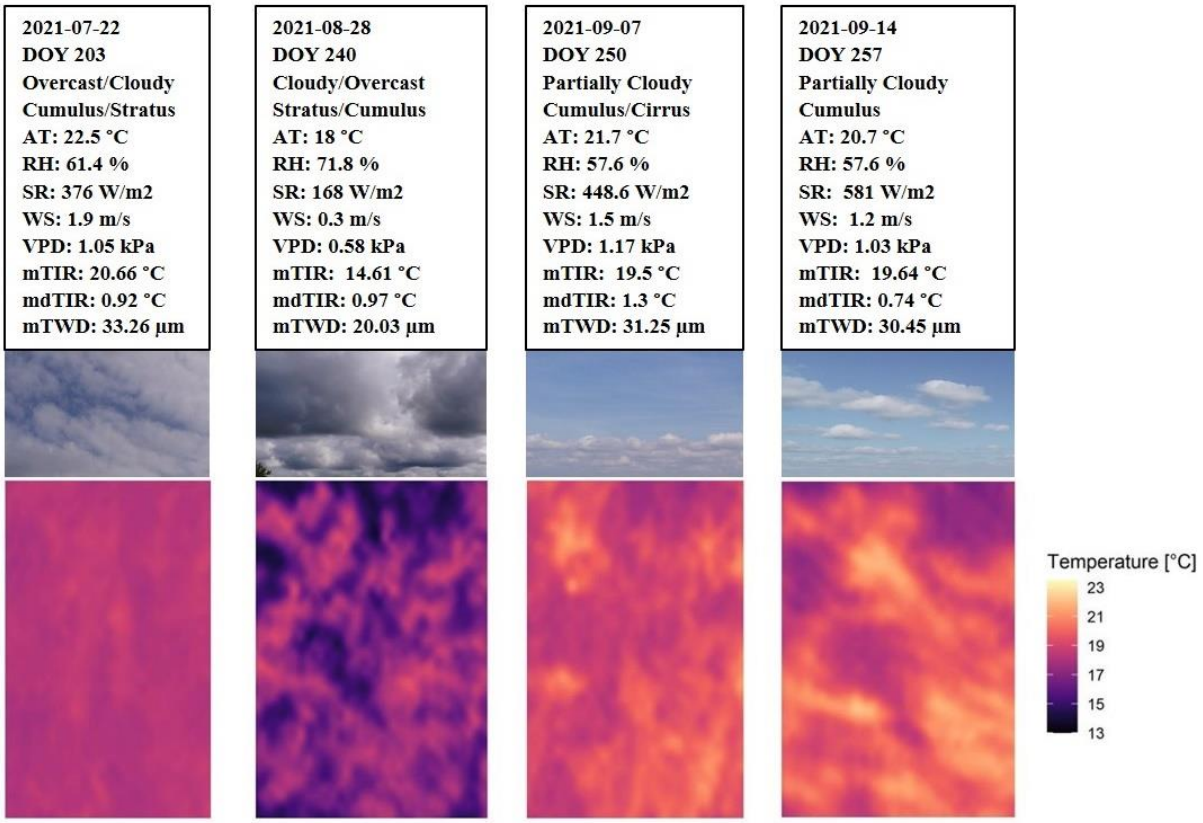


Figure 24: Four epochs from the growth season campaign were selected for comparison purposes. Overcast refers to when the sky is completely covered by clouds. The term Cloudy depicts a number of clouds which block out the sun most of the time with the possibility of short episodes which uncover the sun. Partially cloudy is when most of the sky is clear however some clouds could potentially block out the sun temporarily. Also reported is the Air Temperature (AT), Relative Humidify (RH), Solar Radiation (RH), Wind Speed (WS), Vapor Pressure Deficit (VPD), Mean TIR (mTIR), Mean absolute Difference TIR (mdTIR) and mean Tree Water Deficit (mTWD) of the 5 point dendrometer-equipped trees of the pure beech stand. DOY = Day of Year.

Concerning the flight grid mission, the tree crown temperature of a sample tree was extracted from the TIR imagery derived from three separate flight lines at a flying height of 75 meters for each of the four epochs. The distance from the sensor to the tree crown varied for each image due to changing incidence angles. The tree crown position for the first flight line was approximately +20 ° from nadir with the flight line heading in the north-east direction. The second flight line had the tree crown positioned at nadir while heading in the south-west direction. The third flight was -20 ° off nadir in the X-axis. Each flight line contained 5 images acquired at two second intervals. Table 11 displays the mean and standard deviation (SD) for each flight line for all four epochs. Day 250 showed the highest SD (0.59 K) for the flight line

with an incidence angle of 20 ° off nadir and the lowest SD was 0.06 K on Day 257 with a -20 ° angle off nadir. The overall average SD across all acquisition days was 0.25 K (n = 12). The highest SD of Day 250 corresponds with the highest mADTIR and VPD of the four epochs. Notably, the average TIR temperature over flight lines could vary almost two K.

Table 11: Comparison of the mean and standard deviation (SD) of crown temperature in ° Celsius of a selected tree from various flight lines at a flying height of 75 meters. The higher SD for the North-east direction suggests an influence due to an increased incidence angle from the sensor to the tree crown and possible sun spots reflections.

Grid	**DOY: 203**		**DOY: 240**		**DOY: 250**		**DOY: 257**	
Azimuth and angle	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
NE +20°	22.09	0.32	15.9	0.14	19.57	0.59	18.43	0.13
SW 0°	20.88	0.25	16.19	0.17	18.08	0.27	17.22	0.43
NE -20°	20.4	0.13	14.73	0.15	18.79	0.38	17.96	0.06

A depiction of the gridded flight lines is portrayed in Figure 25 where each individual tree crown extraction is shown in comparison to the tree crown temperature of a neighbouring tree equipped with a leaf temperature sensor. The leaf temperature of the neighbouring tree is not meant for a direct comparison but rather to show an approximation of how changing cloud cover is affecting the crown canopy at a given time. A difference of temperature is evident among all three incident angles which is to be expected in changing cloud cover conditions and with about a minute between flight lines. This is confirmed for the most part with similar fluctuations in the leaf temperature of the neighbouring tree. Aside from day 240, the dispersal of TIR SD is notable. Day 240 also corresponds with the lowest AT, SR, WS, VPD, mTIR and mTWD and highest RH of the four epochs.

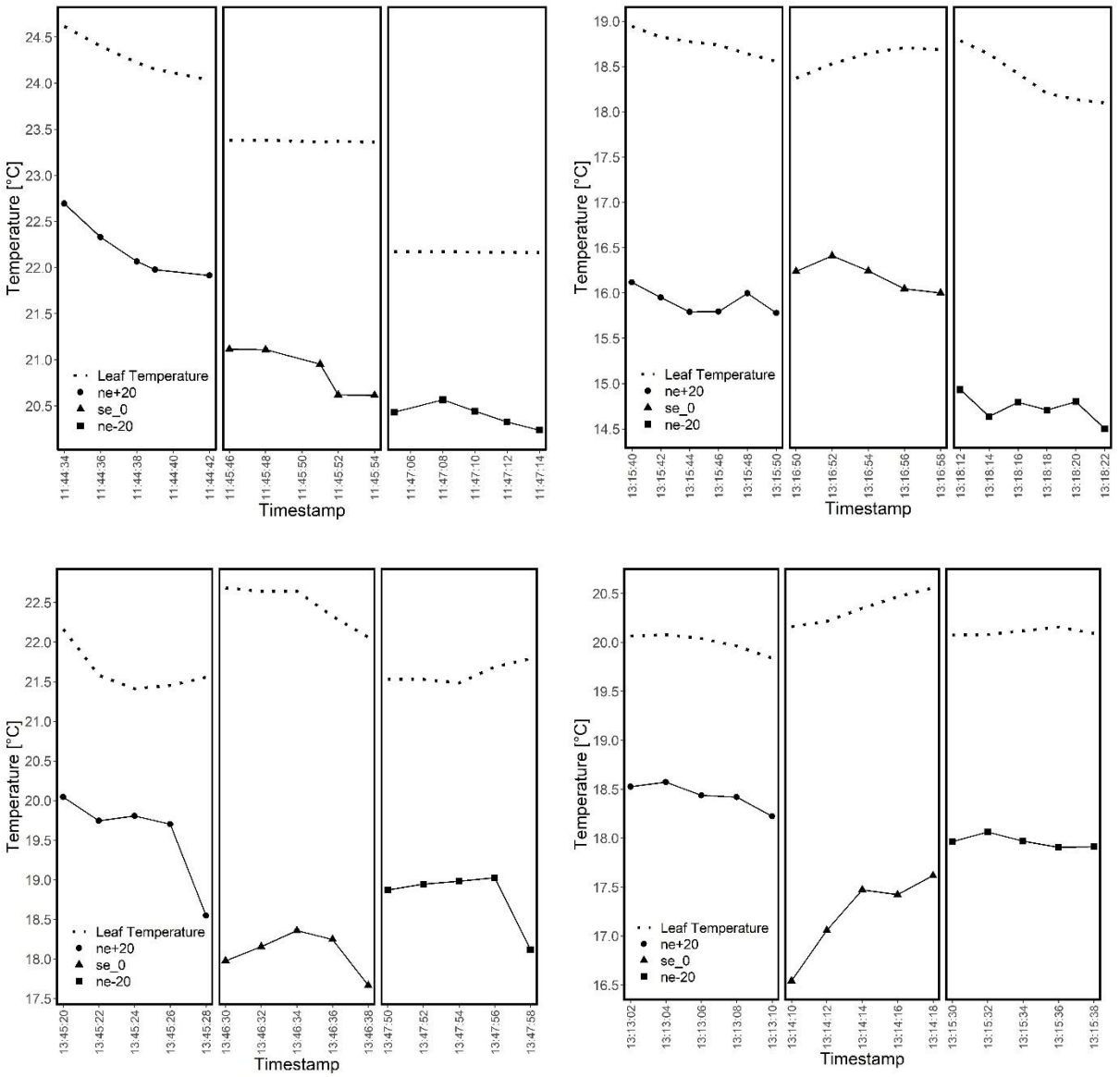


Figure 25: The extracted TIR values of the sampled tree crown from three flight lines from the gridded mission shown with the timestamp while compared to the leaf temperature of a neighboring tree. Top right: DOY 203; Top left DOY 240; Bottom left: DOY 250; Bottom right DOY 257.

3.3.4 Single Image Acquisition Method

The single tree acquisition method was accomplished by a waypoint flight planning method where the UAV would hover over each tree for 10 seconds acquiring five images via a two second intervalometer. Two passes were carried out on each acquisition day. A distance of approximately 10 meters to the tree crown was maintained for all images. The average SD across all acquisition days was 0.9 K (n = 8) with a minimum SD of 0.6 K and maximum of 0.18 K. Outside of slight variations on day 257, the low SD remained consistent throughout the datasets (see Table 12). Here, variation of mean temperature between passes was almost two K on day 203 which was also the day with the highest WS. Aside from the second pass of day 257, it can

be seen in Figure 26 that the extracted TIR values within a pass maintain a high consistency when the close-range hovering waypoint method is applied. The average temperature between passes can however vary up to two degrees during changing cloud cover and high winds.

Table 12: Results showing the mean and standard deviation of crown temperatures acquired with single shot thermal imagery at nadir and a distance of approximately 10 meters.

Single Tree	**DOY: 203**		**DOY: 240**		**DOY: 250**		**DOY: 257**	
Passes	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Pass 1	22.72	0.09	16.92	0.09	19.35	0.09	20.05	0.05
Pass 2	20.06	0.09	15.85	0.08	20.57	0.08	20.19	0.18

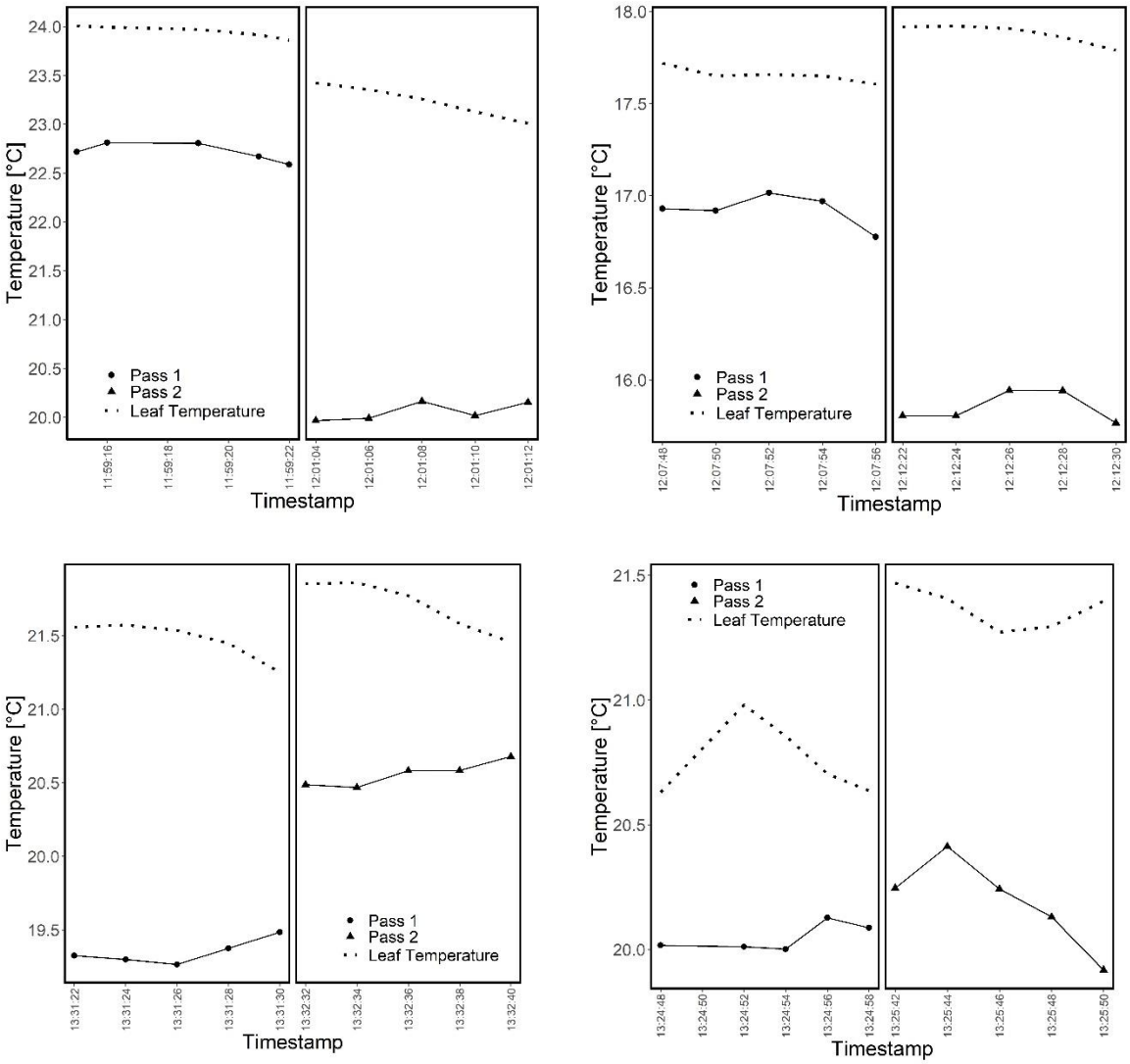


Figure 26: The UAV-derived thermal values acquired with two passes separated by a minute repeated over the four acquisition dates shown in conjunction to the direct leaf temperature values in the tree crown. Top right: DOY 203; Top left DOY 240; Bottom left: DOY 250; Bottom right DOY 257.

3.3.5 Tree Water Deficit and Meteorological Data

Of the nine point and band dendrometers available at beech study plot, five were selected for modelling and validation purposes. The TWD was calculated for the entire growth season and compared to meteorological data derived from the weather station located in an open area approximately 170 meters from the beech study plot. The hourly growth season dataset (n = 2928) was tested for normality using the Shapiro-Wilk test and the hypothesis of normality was rejected. For this reason, a GAM model was implemented to assess which meteorological features or combination of features best model the TWD estimate. Table 13 displays the Adjusted R^2 for nine different variations of feature models. Each model was also created using a one- and two-hour lag of the TWD behind the meteorological data, as well as no lag. This was done for the purpose of assessing the temporal offset in transpiration regulating variables such as VPD, WS, and SR with the timing of lower stem fluctuations expressed by the TWD. The highest R^2 scores are shown as bold text and show that a one-hour lag is for the most part the best explanation of the physiological delay between the atmospheric drivers and the TWD. The best GAM model was accomplished with a R^2 of 0.667 when using a combination of WS, WR, and VPD as input model variables. Interestingly, WS and SR fared slightly better when a two-hour lag was implemented. Figure 27 shows the ranking (left to right) of the nine models using the AIC model evaluation method. Here it can be seen that the combined WS, SR, and VPD model resulted with the lowest AIC value in comparison to all other features and feature combinations (see Figure 27).

Table 13: Overview of the Generalised Additive Models (GAMs) where various meteorological features were implemented to predict the Tree Water Deficit across the growth season in 2021 (n = 2928).

GAM Models	Adj. R2: 0 Lag	Adj. R2: 1h Lag	Adj. R2: 2h Lag
AT	0.514	**0.587**	0.586
RH	0.557	**0.622**	0.606
WS	0.102	0.146	**0.172**
SR	0.15	0.282	**0.402**
VPD	0.585	**0.659**	0.649
AT+RH	0.585	**0.657**	0.646
WS+VPD	0.591	**0.663**	0.653
SR+VPD	0.636	**0.663**	0.657
WS+SR+VPD	0.643	**0.667**	0.66

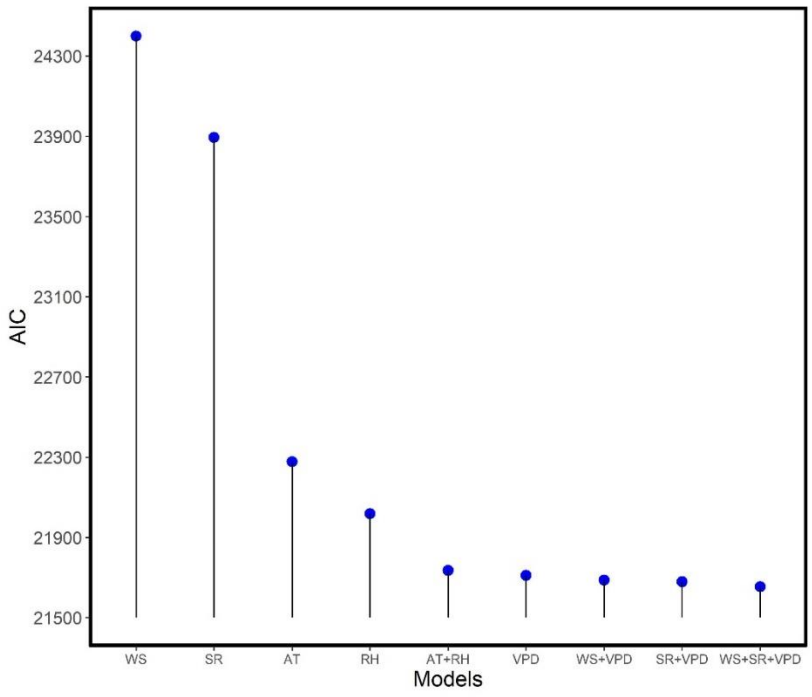


Figure 27: The nine models were compared using the Akaike Information Criterion (AIC) mathematical method to evaluate how well the model fits the data. A combination of WS, SR and the VPD produced the best results.

3.3.6 Correlation Analysis between Tree Water Deficit and Specific Features

Thermal data captured near solar noon (± 90 mins) for each of the five point dendrometer-equipped tree crowns totalled 13 flight missions during the 2021 growth season (n = 65). This dataset did not follow a normal distribution and therefore a Spearman correlation matrix was created comparing the calculated the TWD for each individual tree against the mean crown TIR temperature, meteorological data and derivatives (see Figure 28). Each matrix represents a different lag: no lag, a one-hour lag, and a two-hour lag. Here, the aim was to determine the physiological delay due to the temporal offset of current stem fluctuations and atmospheric drivers influencing transpiration and leaf temperature at the time of acquisition. Another intention of the correlation matrix analysis is to gauge the feasibility of differentiating between individual tree water status, not just on the sampled stand level amid acquisition days, but also to assess the heterogeneity among individuals on a given acquisition day. It is evident that some trees correlate less with the available features depending on the time delay. In some cases, it is also noticeable that particular features may also be an influencing factor to TWD at different time delays.

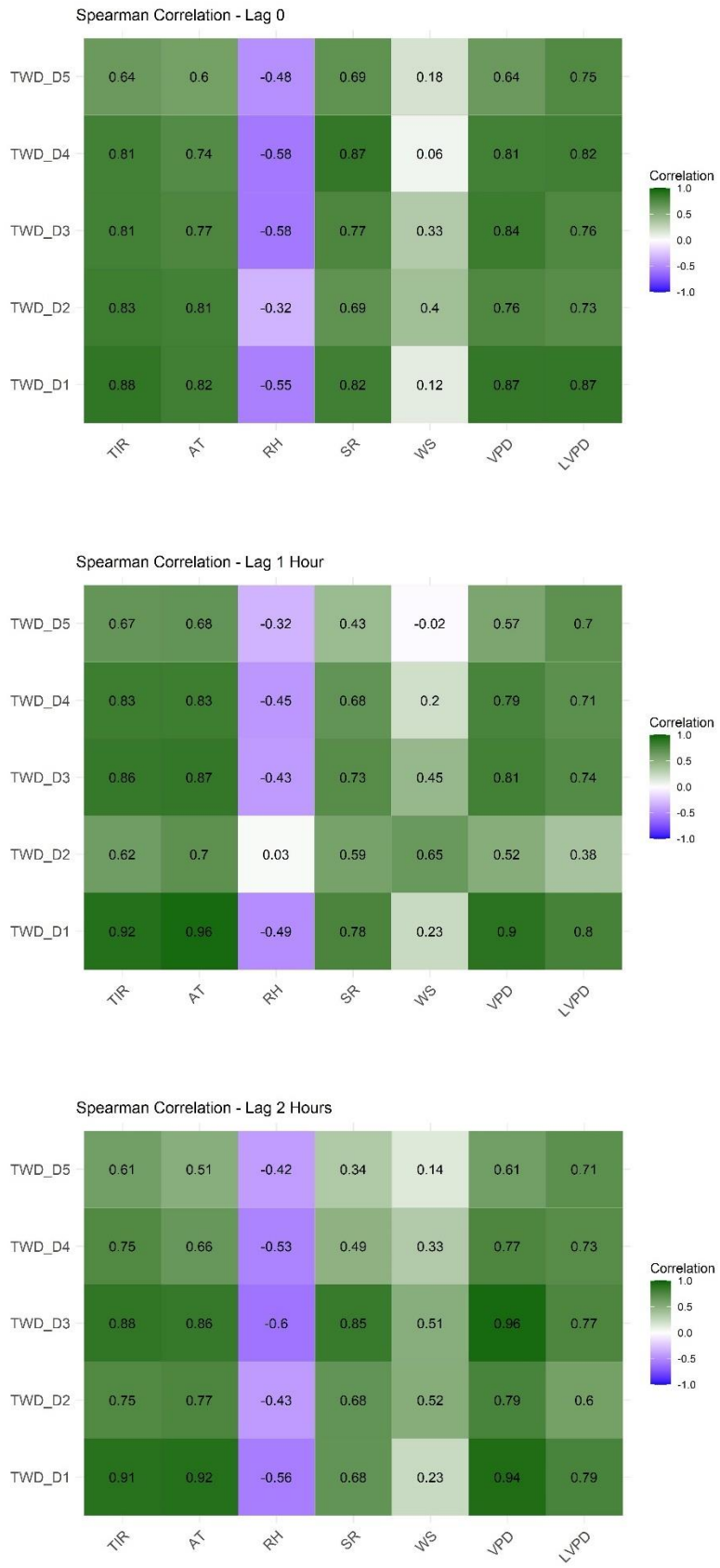


Figure 28: Spearman correlation matrix comparing the TWD calculated from 5 point dendrometers with the TIR imagery, meteorological data and derivatives. Each matrix from top to bottom coincides with no lag, a one-hour lag, and a two-hour lag.

In order to better evaluate the correlations over the various lagging times, the mean correlation was calculated for each feature over the three lagging variations (see Table 14). The values of each matrix were first subjected to the Fisher-Z transformation to mitigate against bias which can be particularly relevant for smaller datasets (Silver & Dunlap, 1987). It is evident here, that not just one lag variation is responsible for the best correlations. The TIR imagery, which is probably the most important feature of this study tends to correlate best at a one-hour lag, however an hour before and after is not a significant difference. The most noticeable difference is that SR correlates best with the TWD without any lag. Additionally, the stand-alone LVPD also correlates better with 0 lag.

Table 14: The mean correlation derived from the Pearson correlation matrix at various time delays compared to the TWD at 3 various time delays. Correlation values were first transformed to Fisher-z values in order to mitigate against bias.

Lag	TIR	AT	RH	SR	WS	VPD	LVPD	Mean
0h	0.81	0.8	-0.53	**0.74**	0.34	0.81	**0.83**	**0.54**
1h	**0.83**	**0.84**	-0.42	0.64	0.28	0.8	0.77	0.53
2h	0.82	0.81	**-0.54**	0.46	**0.38**	**0.82**	0.81	0.51

3.3.7 Modelling Tree Water Deficit

The dataset derived from the 13 flight missions of the 2021 growth season was partitioned in a 70/30 training and validation data split for usage with a GAM. Three models were trained with varying input features applying the features LVPD, TIR+RH+AT, and TIR+VPD. The model variations were then applied with the three various time delays (0h lag, 1h lag, 2h lag) (see Table 15.) The models were then used to predict the TWD for the testing datasets where the RMSE, MAE and R^2 was calculated for accuracy assessment. The models were not yet further tested on data outside of the Britz research station beech plot due to a lack of accessible point dendrometers and accompanying TIR tree crown data. In terms of R^2, the one-hour lag produced the best results across all three models where the feature combination TIR+RH+AT had the highest R^2 of 0.87 an RMSE of 4.92 and MAE of 4 (see Table 15 and Figure 29). The one-hour lag model with the features VPD+TIR also resulted in a high R^2 of 0.81, however also a high RMSE of 9.76 and MAE of 7.01. The LVPD model with a one-hour lag also shows promising results with an RMSE of 6.87 and MAE of 5.1 as well as an acceptable R^2 of 0.71. Most of the models show an almost near linear relationship except for RH (see Figure 29).

Table 15: Overview of the GAMs tested with the three lag variations.

Lag 0 (none)	**LVPD**	**TIR + RH + AT**	**VPD + TIR**
RMSE	6.21	4.56	7.53
MAE	5.1	3.95	5.72
R^2	0.48	0.75	0.26
Lag 1 Hour			
RMSE	6.87	4.92	9.76
MAE	5.1	4	7.01
R^2	0.71	0.87	0.81
Lag 2 Hours			
RMSE	6.84	7.32	6.37
MAE	5.65	6.02	4.99
R^2	0.46	0.37	0.53

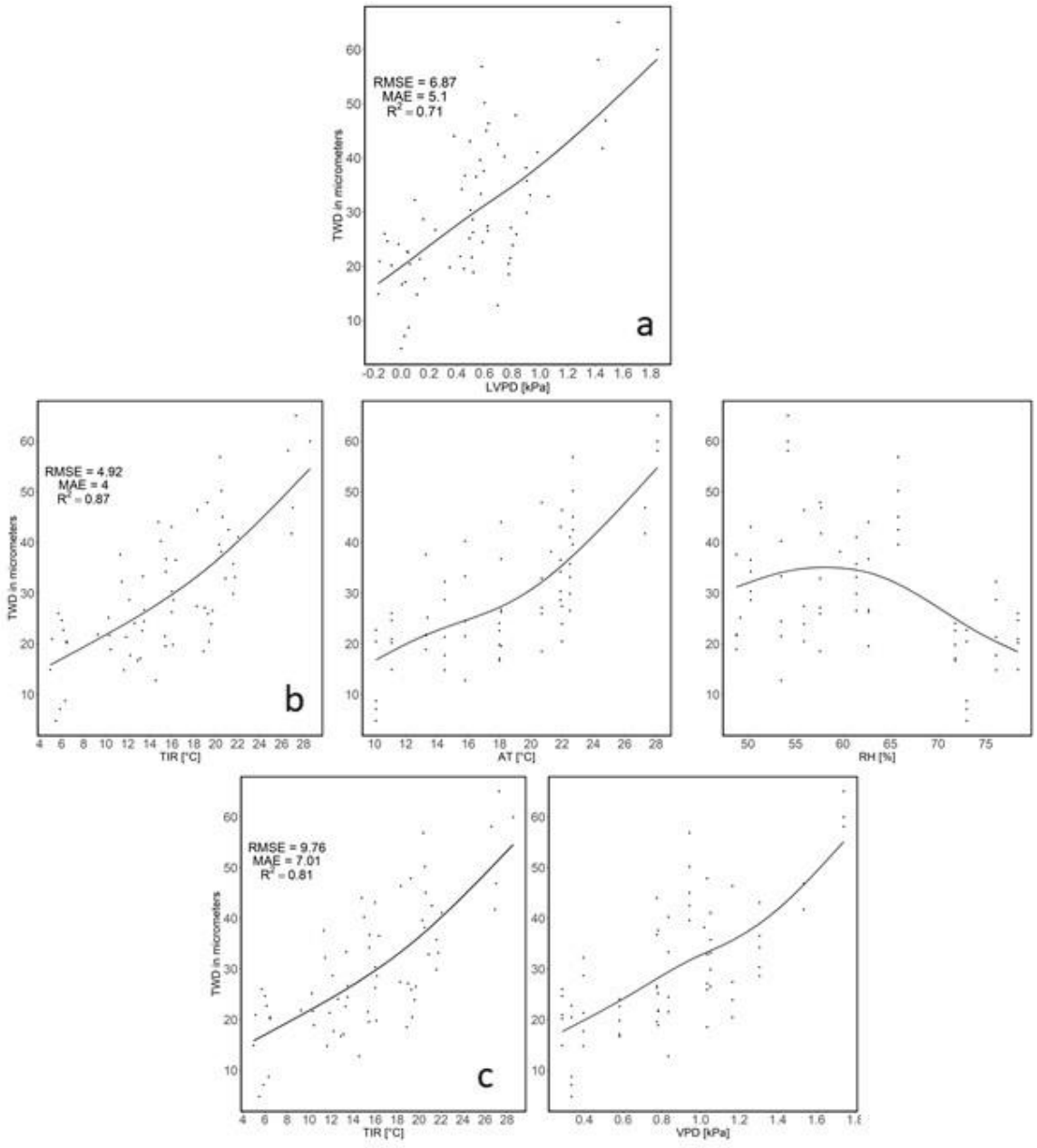


Figure 29: The three Generalised Additive Models (GAMs) tested over the lag variation of one-hour. (a) GAM using the LVPD (b) GAM using TIR, AT and RH. (c) GAM using TIR and VPD.

3.4 Discussion

3.4.1 Thermal Sensor Assessment

Results of the indoor test showed that using a limited number of pixels (<3), despite being directly at the leaf temperature sensor, can propagate errors over 1 K. However, the indoor plant experiment provides only an indication of potential error and different values could occur when tested with beech due to a different leaf morphology and physiology, however supports the concept of the importance of a minimum pixel spot size (FLIR, 2018; Playà-Montmany &

Tattersall, 2021). The spot-size effect occurs when a limited number of pixels are used (i.e. < 3) to extract the temperature of an object. Here the object can be influenced by nearby surfaces (FLIR, 2018) and also be potentially susceptible to "bad pixels". It is typically advertised that TIR array sensors will have no lower than 0.37 – 1 % bad pixels depending on reporting from the manufacturer (Holst, 2000; Vollmer & Möllmann, 2018). This would mean that a thermal sensor size of 160 X 120 could potentially have up to 192 (from 32768) faulty pixels. Therefore, it is crucial to work with a minimum spot-size or minimum number of pixels in order to have enough "good" pixels to average from. We tested a minimum of 10 pixels within an extraction polygon however FLIR (2018) recommends a spot-size of a 10 pixel diameter which would in effect result in an area more than 10 pixels.

In terms of accuracy, the indoor experiments showed an RMSE ranging from 0.55 K to 0.73 K where interestingly the dryer leaves from the unwatered plant proved to show lower accuracy. This goes along with the concept that dryer leaves will typically have a lower emissivity (see Table 16) and it is known that objects with lower emissivity can affect measurement accuracy (FLIR, 2018; Holst, 2000; Minkina & Dudzik, 2009; Vollmer & Möllmann, 2018).

Table 16: Emissivity ranges for green healthy vegetation and dry vegetation (Lillesand et al., 2015).

Material	**Ɛ over 8 – 14 µm**
Green Vegetation	0.96 – 0.99
Dry Vegetation	0.88 – 0.94

As for the accuracy of the sensor generally speaking, we showed an RMSE of below 1 K during the indoor experiments which is below the typical commercial thermal sensor accuracy ±2 K or ±2 % as reported by Vollmer & Möllmann (2018). Accuracy in the field however, proved otherwise and was more challenging to assess.

Temperature measurement accuracy assessment in the field was somewhat difficult as we had limited possibilities to determine and validate leaf temperature within the upper tree crown. Over the 13 acquisition days we reported an RMSE of 3.22 K (average of both trees) which is higher than the previously mentioned typical accuracy of 2 K. It should be mentioned however that the flight missions were carried out in varying weather conditions were high winds, extreme midday temperatures, changing cloud cover as well as the onset of autumn senescence near the end of the campaign was present. Furthermore, in both the indoor and the field experiments it was evident that the TIR sensor consistently underestimated the validation measurement where a

potential offset correction could be employed alongside further calibration with the usage of high-resolution meteorological data. Once again, however, for this study it should be stated that the limitations of only one leaf temperature sensor located in the upper crown was useful for a rough estimation however cannot be taken as a reliable validation source. It is our aim to increase the amount of upper crown leaf temperature sensors for future studies to at least four, while still remaining within practical means of maintaining a biological validation source. Important here is that the sensors are mounted at well representative positions of the upper tree crown and are capturing temperature values that the TIR imaging can also reach. The question here however, is to what extent typical shade leaves found in the lower tree crown will contribute to the estimation of the TWD. Further testing is required in this direction and could aid in an improved masking strategy for pixel extraction.

More research is required with the use of leaf temperature sensor in the tree crowns as validation and for exploring calibration methods. Here we have plans to add two validation trees at the Britz research station with each upper tree crown installed with at least four-leaf temperature sensors. Furthermore, the development of an outdoor low-cost blackbody where we can also assess sensor drift, pre-operational warming and internal periodic NUC calibration (Kelly et al., 2019; Wan et al., 2021) are also future research topics.

3.4.2 Acquisition Method Assessment

The flight mission grid acquisition method is the typical TIR image procurement method for UAVs. Here a mosaiced dataset is created of the area of interest. This is often carried out simultaneously with RGB imagery and for improved image matching accuracy and positional accuracy ground control points are also implemented (Gerke, 2018) or an onboard RTK system is used (Przybilla & Baeumker, 2020). A co-registered thermal dataset with that of a RGB or multispectral dataset can aid in the locating of individual tree crowns which can often not be discernible within aerial TIR imagery alone. The gridded method, however, can be time consuming, as one is required to know the terrain beforehand for flight planning purposes and will require longer mission times which could require battery changes and be limited in coverage due to line-of-site restrictions (*EASA*, 2022). As we showed with the results of the gridded missions the dispersal of TIR temperature can have a SD of up to 0.59 K within a single flight line. Furthermore, between flight lines and acquisition times over a minute can vary several K between flight lines. The dispersion TIR temperature values could potentially be even more

pronounced when a battery change is necessary between flight lines due to changing weather conditions.

The single shot acquisition strategy shown in this study introduces a relatively uncommon method where single images of the *Micasense Altum* sensor undergo radiometric calibration and affine transformation without the need for Structure-from-Motion processing. Additionally, multispectral bands can also aid in segmenting tree crowns as well the separation of sun and shade leaves and woody parts of the tree crown and ground pixels. The single image acquisition method also saved on processing time and reduced the length of flight missions considerably. Here the center point of the images can be used and with a footprint prediction derived from flying height, azimuth, and sensor intrinsic parameters, to estimate the ground footprint (McGlone & Lee, 2013) and used as sample datasets for ground truthing purposes.

The lower SD for the single shot acquisition method could be attributed to the closer distance to the tree crown which will reduce atmospheric transmission especially on warm and humid days were the RH is high (FLIR, 2018). Additionally, a closer distance to the object will result in a higher number of pixels within the tree crown allowing for a higher amount of pixel that are averaged and more complex masking could be possible. Furthermore, a consistent incidence angle could help reduce outliers due to variations in emissivity.

Important for further study is to explore sources of error and test which influencing factors can be reduced in the field. Minkina & Dudzik (2009) showed through error modelling, the most influential error sources are derived from a combination of Object emissivity, RH, and Camera -to-object distance (see Table 17). Technically speaking we can reduce object emissivity error by reducing incidence angles especially at extreme angles. Relative humidity error can be reduced by being closer to the object which will also increase the number of pixels which are used for averaging. This could also in turn decrease the effect of outliers propagated from bad pixels of the sensor as well as portions of the crown subjected to sun glint. The addition of near-real time local meteorological data as in for example RH, AT and WS, synchronous with thermal data acquisition could prove beneficial for calibration of thermal imagery under varying environmental conditions. The challenge here is in how to obtain such data reliably during field campaigns.

Table 17: Ranges of potential errors during simulations (adapted from Minkina & Dudzik, 2009).

Input parameter	**Object emissivity (Ɛ*ob*)**	**Ambient temperature (T*o*), K**	**Atmospheric temperature (T*atm*), K**	**Relative humidity (ω)**	**Camera-to-object distance (*d*)**
Error range	±30%	±3%	±3%	±30%	±30%

3.4.3 Drought Stress Validation

The challenge with detecting drought induced stress in beech lies not only in developing a consistent method in TIR image acquisition but also in finding a practical and reliable approach to validate to what extent a tree is experiencing drought stress. The TWD can be explained moderately well throughout the growth season with the VPD ($R^2 > 0.659$) as well as derivates RH and AT ($R^2 > 0.659$). Drew et al., (2011) found that daily variations in TWD was influenced mainly by soil water availability, however RH and AT contributed to variability over some periods. Soil moisture, though an important factor for estimating TWD was not implemented in the study due to the impracticality of acquiring soil moisture during field campaigns, however will be considered in future experiments.

The best modelling results we obtained when using a lag of one hour which was relevant except for WS and SR which showed a higher R^2 at a two-hour lag. The evidence of a lag in the TWD and VPD was also reported by Zweifel (2016) among spruce, pine and ash where modelling efficiency was maximised by shifting the VPD in steps of 30 minutes. Interestingly, this was not the case for beech which is contrary to our results. Nevertheless, this shows that it could be useful to record meteorological data such as RH and AT at a high resolution in conjunction with TIR image acquisition in order to find the appropriate lag parameters. Due to factors such as wind gusts and changing cloud cover, it would be essential to acquire TIR imagery and meteorological data at instances which are more representative for the conditions of that day. For example, a sudden increase in SR and VPD due to the clouds abruptly opening up and revealing the sun may not necessarily affect the TWD for that day or even that hour if it occurs infrequently. TIR measurement can however be highly influenced from such rapid changes. Over the 13 missions, this concept was somewhat neglected for this study and is not necessarily practical and sometimes not possible to mitigate due to time constraints and available UAV battery life. Despite these challenges however, in this study we showed that the TWD is highly correlated with the TIR Imagery ($r > 0.81$) alone regardless of which lag time is implemented and prevailing weather patterns during flight missions. This is promising for the detection of drought stress in terms of modelling TWD at the stand level, however does not necessarily mean that within stand heterogeneity for a single acquisition epoch can be obtained reliably.

With respect to the relationship between TWD and feature variables on an individual tree basis, it can be seen that some individuals correlate less at differing lag times (see Figure 28) which could be interpreted as some individuals having slower reaction times to environmental variables or the point dendrometer is positioned at a spot where the stem fluctuations are not well pronounced. The latter can be evident and problematic for larger tree stems where it is recommended to use multiple point dendrometers to consider potential differences (*Tomst*, 2022). In terms of the individual feature variables reacting at different lag times, this requires to be further researched as modelling could be subjected to overfitting especially for particular microclimates.

The resulting models in this study should be interpreted with caution as in 2021 at the Britz research station there was not any particular drought stress conditions in comparison to what was experienced from 2018 to 2020. For this reason it is important to note is that we are lacking a full range of TWD data particularly for beech where TWD values did not increase more than 60 µm for the year of 2021. Zweifel et al., (2005) reported TWD values of up to 500 µm for oak which shows that in a drought year we could have much higher values than 60 µm which in turn will broaden the range. Until then, it is somewhat difficult to evaluate our resulting RMSE values at this time, however current results when maintained with a broader dataset range could prove relevant for the purpose of establishing within stand heterogeneity. The GAM was chosen for this reason to maintain simple curves which would be more conducive to extrapolation (Wood, 2017) than for example typical decision tree machine learning algorithms. Extrapolation should however be avoided (McCartney et al., 2019), and an improvement of TWD modelling will only be possible with more data especially where data ranges attained during drought stress are available.

Special care was taken not to repeat features within a model (Kuhn & Johnson, 2019). For example, when the VPD was implemented as a feature, RH and AT would not be used as this would result in the repeated use of the feature as the VPD is a derivate of both features. Also, to consider is that when modelling TWD, it could be necessary to model features with limited feature engineering in order for ML algorithms to find particular weightings depending on that feature. Here, feature engineering as in the use of VPD or LVPD within models could miss out on such particular weightings as well as potential lag differences among features. The LVPD performed moderately well within the TWD model, however the possibility to use the LVPD as a stand-alone index could also be of interest and certainly practical if an absolute range depicting quantifiable drought stress among beech for a particular region was developed.

3.5 Conclusion

In this study we showed that single shot TIR imaging of beech tree crowns can produce promising results for the purpose of calculating LVPD and estimating TWD. Despite a dataset lacking actual drought occurrences we were able to model variations of TWD within the 2021 growth season based on close-range single shot thermal imaging and synchronous meteorological data. As opposed to typical UAV-based gridded flight plans and orthomosaic derivatives, close-range single shot thermal imaging can aid in reducing the effects of variations of RH and emissivity by decreasing incidence angles and sensor-to-object distance, as well as increasing the number of pixels for thermal data extraction. More field trials are required with the incorporation of high-resolution meteorological data synchronous to thermal imaging for calibration purposes. Furthermore, an improvement in thermal imaging accuracy validation is required in terms of an increased number of crown-based leaf temperature sensors and a field-based blackbody. This research is an important step towards the incorporation of thermal imaging for the purpose of quantifying drought stress among other applications at intensive monitoring plots.

4 European Beech Spring Phenological Phase Prediction with UAV-derived Multispectral Indices and Machine Learning Regression

Abstract

Acquiring phenological event data is crucial for studying the impacts of climate change on forest dynamics and assessing the risks associated with the early onset of young leaves. Large-scale mapping of forest phenological timing using Earth Observation (EO) data could enhance our understanding of these processes through an added spatial component. However, translating traditional ground-based phenological observations into reliable ground truthing for training and validating EO mapping applications remains challenging. This study explored the feasibility of predicting high-resolution phenological phase data for European beech (*Fagus sylvatica*) using Unoccupied Aerial Vehicle (UAV)-based multispectral indices and machine learning. Employing a comprehensive feature selection process, we identified the most effective sensors, vegetation indices, training data partitions, and machine learning models for phenological phase prediction. The model that performed best and generalized well across various sites utilized the Green Chromatic Coordinate (GCC) and Generalized Additive Model (GAM) boosting. The GCC training data, derived from the radiometrically calibrated visual bands of a multispectral sensor, were predicted using uncalibrated RGB sensor data. The final GCC/GAM boosting model demonstrated capability in predicting phenological phases on unseen datasets within a Root Mean Squared Error (RMSE) threshold of 0.5. This research highlights the potential interoperability among common UAV-mounted sensors, particularly the utility of readily available, low-cost RGB sensors. However, considerable limitations were observed with indices that implement the near-infrared (NIR) band due to oversaturation. Future work will focus on adapting models to better align with the ICP Forests phenological flushing stages.

4.1 Introduction

The concept of gathering data on the timing of leaf opening, flowering, fruiting and leaf fall alongside climatological observations “so as to show how areas differ” was first proposed by the Swedish botanist in his work *Philosphia Botinica* in 1751 (Linnaeus, 1751) and is still relevant

today (Lieth, 1974). Phenological observations have historically assisted agriculture by means of predicting the timing for cultivation practices (Zhang, 2012) and in the last 100 years emerged as a scientific discipline (Schwartz, 2013). In recent times, phenological events are recognized as a bio-indicator for climate change (Menzel, 2002). Today, phenological data is a sensitive proxy for climate change investigation (Schwartz, 2013) due to an observed relationship between phenological timing and a changing climate. In particular, spring phenology is an indicator of climate change where observations are mirrored in temperature change (Menzel et al., 2006). Understanding phenological variations at the stand level can provide insights into early spring flushing advantages and the risk of late frost damage.

The Intergovernmental Panel on Climate Change (IPCC) has reported a 1.53°C increase in average land temperature for the period 2006 - 2015 in comparison to the 1850 - 1900 period (IPCC, 2018). Warmer temperatures, alongside changing precipitation patterns, have altered the growing seasons causing increased tree mortality (IPCC, 2018) however may also lead to increased carbon storage due to longer growing seasons (M. A. White et al., 1999). The growing season has extended by approximately 10-20 days in recent decades, and projected temperature increases of 1.4-5.8°C in the next century may benefit some species while threatening others (Linderholm, 2006).

European beech (*Fagus sylvatica* L.) in central Europe faces a potential threat from earlier budbreak and leaf-unfolding with extended growing seasons. Sudden freezing temperatures, known as late frost, can damage newly unfolded leaves and affect growth (Menzel et al., 2015; Rubio-Cuadrado et al., 2021; Sakai & Larcher, 1987). Late frost damage can shorten the growth season and has been observed in younger beech and oak stands without an upper canopy layer (Sachsenforst, 2020). This highlights the need for phenological modeling on a regional scale to identify high-risk areas for adaptive forest management practices.

Temporal phenological models typically rely on the timing of phenological events recorded as the day of the year (DOY) (Zhao et al., 2013). Phenological models also consider temperature-dependent chilling days and thermal time, which can be species-specific (Menzel, 1997). However, accurate models should account for other factors such as photoperiod and precipitation (Brügger & Vasella, 2018). Accurate phenological models are essential as early warning systems for stressed forest ecosystems and for simulating phenological processes across various timeframes. These models are vital for grasping the spatial and temporal differences in forest

phenology at a regional scale, offering insights into the extent of climate change and variations in carbon fluxes (Cleland et al., 2007).

Visual phenological observations conducted by experts through long-term ground observation schemes provide valuable data on stand and individual tree levels (Menzel et al., 2006; Raspe et al., 2020; Vilhar et al., 2013). While subjective, these observations are considered highly accurate and provide information on crown condition and potential tree damage. However, visual ground observations are labor-intensive, limited in spatial coverage, and require experienced observers. Standardized observations at intensive monitoring plots, such as those conducted by ICP Forests, provide valuable information on individual tree conditions (Raspe et al., 2020).

Terrestrial observation methods using webcams and automated cameras mounted on towers have become popular for monitoring forest phenology. These methods provide quantitative multispectral data that accurately capture the timing of phenological events. However, their spatial coverage is limited, and they are often used to validate satellite observation platforms (N. Li et al., 2020) and assist in mapping and modeling phenological metrics globally (Zeng et al., 2020).

Satellite-based remote sensing has also become a valuable tool for studying phenology at different spatial scales (Kowalski et al., 2020). Open access satellite platforms like Landsat 8 and Sentinel 2a/b provide global coverage with temporal and spatial resolutions of 16 days and 30 meters (Landsat, 2022) and 5 days and 10 meters (Copernicus, 2022), respectively. The MODIS Global Land Cover Dynamics Product offers land surface phenology information on a global scale at a 500 m spatial resolution (Friedl et al., 2010; Ganguly et al., 2010). However, linking plot-level measurements to satellite-derived pixel values is challenging due to phenological heterogeneity and other factors (Klosterman et al., 2018; K. White et al., 2014).

The utilization of Unoccupied Aerial Vehicles (UAVs) has increasingly become a key tool for enhancing phenological observations in recent years, bridging the gap between terrestrial and satellite-based phenological observation systems. One major challenge is that a typical terrestrial observation plot often covers only a limited number of satellite pixels. Berra (2019) highlighted that within a single Landsat pixel, phenological events can show significant variance ($R^2 < 0.50$) when compared to UAV-derived phenometrics. Moreover, satellite pixels may not adequately account for microclimatic variability (Klosterman et al., 2018), which is especially relevant in

areas where tree phenology is not uniform. This suggests that UAV data, when trained from localized observation plots, could be scaled up to encompass larger areas, thus enhancing the training scope for more satellite pixels. Along these lines, Atkins et al. (2020) demonstrated that UAVs and terrestrial camera systems could be effectively used in conjunction to gather high-resolution phenological data.

Converting UAV data into phenological metrics presents challenges, including sensor calibration and processing workloads for analysis-ready datasets. Flight campaigns aim to determine the onset of spring leafing out and acquire training data representative of the complete phenophase range. The number of flight missions depends on the required phenophase resolution and may involve repeated missions at various observation plots (Belle & Papantonis, 2021). Automatic methods and explainable machine learning algorithms are crucial for organizing and processing the acquired data. However, expert-based qualitative methods are necessary for assessing the complexities of phenological development.

Machine learning (ML) algorithms, particularly for image classification, have gained popularity in remote sensing applications. ML algorithms, including Support Vector Machines, Random Forests, and Neural Networks, offer higher accuracy than traditional parametric methods (Maxwell et al., 2018; Schulz et al., 2018). ML techniques incorporating features such as vegetation indices and meteorological data can spatially analyze the influence of temperature and precipitation on phenological processes (Czernecki et al., 2018). UAV-based ML modeling has been applied to monitor individual tropical tree phenology using various RGB-based textures and vegetation indices Park et al. (2019).

In this study, we investigate the potential of using UAV-based multispectral data and machine learning (ML) algorithms to automate the extraction of phenological phases for European beech. Initially, we examine phenological trends from 2006 at the Britz research station to provide an overview of specific patterns. Subsequently, we analyze multispectral data and derived indices from 2019 to 2021, conducting correlation analysis and polynomial fitting with field observations for feature selection. Additionally, ML techniques in regression mode are employed to train models using data from 2019 and 2020, which are then tested against the unobserved spring phenological phases of 2021. Ultimately, the chosen model is trained with various data

subsets categorized by year of origin and thoroughly tested on new data from 2022, as well as data from older beech stands (over 50 years) and different regions.

4.2 Methods

4.2.1 Study Site

UAV and ground-based phenological observation for this study were carried out at the Britz intensive forest monitoring research station (thuenen.de) located in the lowlands of north-eastern Brandenburg, Germany. Brandenburg lies in between oceanic and continental climate zones and belongs to the young moraine landscape of the *Weichsel* glacial period. The soil at the site is locally known as a "*Finowtaler Sandbraunerde*" (Schulze & KOPP, 1998) and rated as a *Dystric Cambisol* derived from *Pleistocene* sand deposits with a *moder* organic layer (Don et al., 2019). Average yearly temperature and precipitation for the region has been recorded at approximately 8.9 °C and 570 mm (Riek, 2004) respectively. Datasets acquired from other sites, for the purpose of testing Models with unseen data, were acquired from "Kahlenberg" near the Britz research station, and the "Black Forest" located north of Freiburg in south-western Germany.

The Britz research station was initially designed in 1972 for intensive forest hydrology research with 9 large-scale lysimeters planted with a mixture of deciduous and coniferous species (Müller 2010). The site is also equipped with an array of digital and analog dendrometers, and various other sensors for sapflow and soil moisture measurement as well as meteorological data. Recently, the research station has undergone a major digitalization overhaul where sensor data is automatically synced to the cloud and individual trees are mapped at a sub-decimeter accuracy.

4.2.2 Phenological Ground Observations

Ground-based phenological observations have been carried out at the Britz research station since 2006 over all nine plots as well as at various satellite plots. UAV-based phenological missions have been implemented at the research site since 2018 with the aid of on-board multispectral sensors. As of 2021, phenological observations are additionally carried out with tower-based phenology cameras.

In this study, the spring phenology of 14 trees of an approximately 50-year-old European Beech (*Fagus sylvatica*) stand was recorded via traditional observations methods alongside UAV-based missions with an offset of a minimum of ± 3 days. After 2020, traditional observations were carried out synchronous with UAV missions.

At the Britz research station, ground-based phenological observations have been conducted since 2006 using traditional methods and satellite plots. UAV-based missions with multispectral sensors have been implemented since 2018, and tower-based phenology cameras have been used since 2021.

In this study, the spring phenology of a 50-year-old European Beech stand was observed using traditional methods and UAV missions. Traditional observations were conducted synchronously with UAV missions starting from 2020, with a minimum offset of ±3 days.

The traditional terrestrial phenological observations for beech at the Britz research station focus on spring green-up phases (0-5) and fall senescence (discoloration and foliation). The spring phases consist of five phases based on various observation techniques, emphasizing the early bud development and leaf hardening. The phases range from 0.0 to 5.0, with decimal values used to indicate progress between phases. For example, if 80% of the observed tree crown is in phase 4, it is recorded as phase 3.8. The "*Britzer*" method of spring phenological phase assessment is unique in the sense that it emphasizes the early onset of bud development with reference to the swelling of buds (phase 0.5) in early spring as well as the hardening and darkening of leaves in phase 5. Table 18 describes the various phases implemented with the "*Britzer*" method alongside other well-known techniques. The differentiation between phase 4.0 and 5.0 is implemented with the *Britzer* method and only considered with the Malisse/Schüler steps (Malaisse, 1964; Schüler, 2012). Figure 30 gives a photographic representation of the *Britzer* method phases.

Alongside the phenological phases, the *Britzer* method also takes foliation percentage into account (0-100 %) which is not directly harmonizable with the ICP Forests flushing method (Raspe et al., 2020). The ICP Forests method is shown in Table 18 for comparison purposes only and should not be taken as a direct method for conversion. Important to note is that the *Britzer* method records foliation as the coverage in % of fully developed leaves whereas the ICP Forests method of flushing records the coverage of green foliation and are not directly harmonisable.

Table 18: Overview of the various tree phenology observation methods for spring leafing out.

Forstreuter (Stages)	**Malaisse/Schüler (Steps)**	**LFE (Phases)**	**ICP Forests (Flushing)**	**Britzer (Phases)**	**Description**
0A	1	1	1	0	Buds established from previous year
0A	1	1	1	0.2	Buds in Winter dormancy
0B	2	2	1	0.5	Swelling
0C	2	2	2	0.8	First buds are bursting (ICP "flushing" = infrequent or slight)
1D	3	3	2-4	1.0	Budburst (full tree)
2E	4	4	2-4	(1.5)	
3	5	5	5	2.0	Young wrinkled leaf visible
-	6	-	-	3.0	Young leaf less wrinkled and long shoot begins to lengthen.
-	6	-	-	4.0	Long shoot fully developed and pilose. Leaves still soft and pilose.
-	7	-	-	5.0	Long shoot almost glabrous and leaves hardened as well as dark green, less pilose

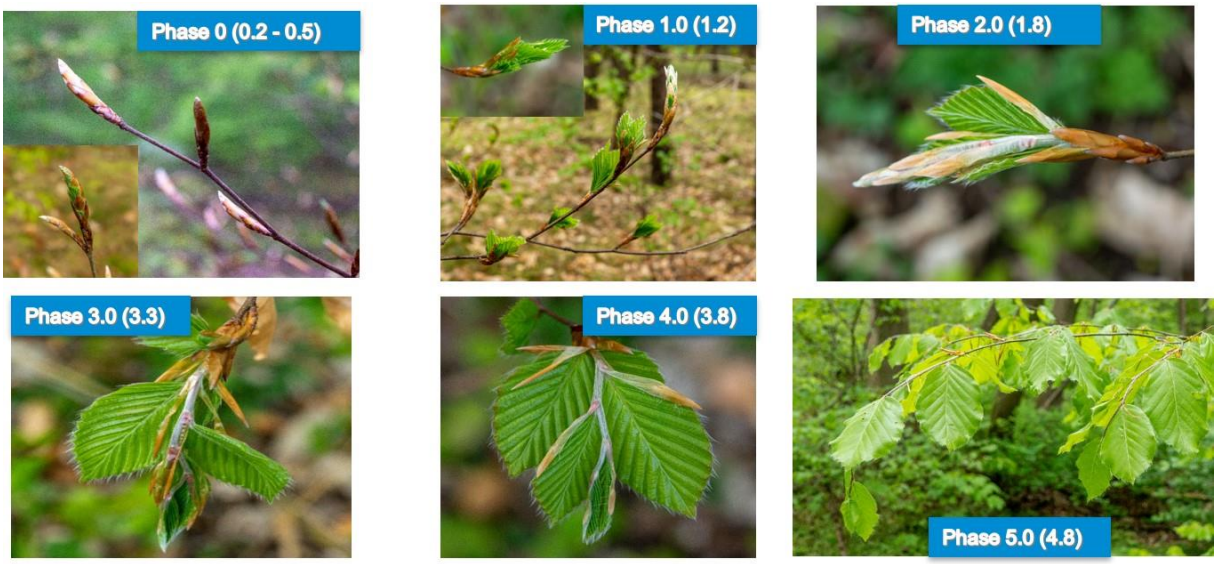


Figure 30: Photographic representation of the "Britzer" phenological phases.

4.2.3 UAV Multispectral Image Acquisition

The UAV remote sensing platform used for this study was comprised of two Unoccupied Aerial Systems (UAS) customized for dual sensor RGB and Multispectral image acquisition. The UAS implemented in 2019 was comprised of a *DJI Phantom 4 Professional Obsidian* (dji.com) with a RGB sensor (mechanical shutter) and a *Micasense Rededge-MX* (micasense.com) multispectral sensor mounted with a custom 3D-printed gimbal (droneparts.de). The UAS implemented after 2020 was a *DJI Martice 210 RTK* (dji.com) with built-in Real-Time-Kinematic capabilities which applies real-time corrections through a Network Transport of RTCM via Internet Protocol (NTRIP). Mounted on the *Matrice 210 RTK* was a *Zenmuse X7* RGB sensor alongside a *Micasense Altum* (micasense.com) multispectral sensor. Specific sensor details for the two multispectral and two RGB sensors can be found in Table 19. Table 20 displays the wavelength and bandwidth information for both multispectral sensors. The Longwave Infrared (LWIR) band was not implemented for this study.

Table 19: Overview of the sensor parameters used in this study.

Sensor	**Mega-pixel**	**Focal length [mm]**	**Pixel Size [µm]**	**Sensor Size [mm]**	**Sensor Size [pixel]**	**Aspect**	**File type**	**GSD [cm] at 75 m**
DJI Phantom 4 Pro (Hasselblad)	20	24	3.09	13.2 x 8.8	5472 x 3648	3:2	jpeg	2.64
Micasense Rededge-MX	1.2	5.4	3.75	4.8 x 3.6	1280 x 960	4:3	tif	5.27
DJI Zenmuse X7	24	24 (36mm equivalent)	3.91	23.5 x 15.7	6016 x 4008	3:2	jpeg	1.23
Micasense Altum	3.2	8 (multi)	4.25	7.16 x 5.35	2064 x 1544	4:3	tif	4.33
	0.02	1.77 (LWIR)	11.9	1.9 x 1.43		4:3		52.48

160 x 120

Table 20: Wavelength and bandwidth for the Micasense Rededge-MX and Altum multispectral sensors.

	Micasense Altum		**Micasense Rededge-MX**	
Band Name	**Center Wavelength (nm)**	**Bandwidth [nm]**	**Center Wavelength (nm)**	**Bandwidth (nm)**
Blue	475	32	475	20
Green	560	27	560	20
Red	668	14	668	10
Red Edge	717	12	717	10
Near IR	842	57	840	40
LWIR	11 µm	6 µm	-	-

Phenology-based image acquisition missions at the Britz research station were carried out every spring (since 2019) typically synchronous with ground observations. A mission is carried out shortly before budburst when buds begin swelling (0.5) and thereafter cycling in a maximum of 3 days during the fast-developing phases of 0.5 to 3. After the third phase, flights and ground observations are limited to bi-weekly acquisition days due to time constraints. Flight missions were carried out near solar noon (± 90 minutes) and calibration panel images (micasense.com, 2022) taken before and after missions as well as the acquisition of Downwelling Light Sensor (DLS) information for each individual multispectral image. After 2020, special care was taken to acquire calibration panel images during moments when the sun was the least affected by cloud cover, especially when the sun was revealed during the flight mission. In order to ensure the capture of as many phases as possible, flight missions are carried out regardless of cloud cover, refrained from during precipitation and for the most part winds over 12 km/h (3.3 ms).

Both *micasense* multispectral sensors were set to capture images with an intervalometer set at 2 seconds and automatic exposure. The RGB sensors were typically set on shutter priority with a

speed of 1/800th of a second. Due to the slow trigger speed of the *Zenmuse X7*, flight mission speed required to be reduced to 3 m/s to insure a forward overlap of at least 80 %.

4.2.4 Data Processing

After each flight campaign, images and metadata were stored on an external hard drive. Naming conventions were based upon a running ID, date, station/district, parcel number, and sensor. After the storage procedure, new similar folder names were created based on individual missions with only the selected images required for the Structure from Motion (SfM) software. These folders were then uploaded to an institute-based Linux High Performance Computing (HPC) cluster. Photogrammetric products were then produced in *Agisoft Metashape* (v1.7.5) with a semi-automated loop in *Python* where the script is interrupted to manually select Ground Control Points (GCPs) in the RGB and Multispectral imagery (see Figure 31). The GCPs insured the repeated usage of crown segmentations throughout the time-series. The script loops through all folder names per year which also implemented the naming conventions for each individual photogrammetric products such as the Point Cloud, Digital Surface Model (DSM), and Orthomosaic.

Further processing was carried out in *R* (R Core Team, 2022) where spectral information was extracted from each individual phenology tree. Each dataset was subset to an Area-of-Interest (AOI) with a minimum boundary containing all of the trees used for the ground-based phenological observations. Preceding the calculation of the vegetation indices, the datasets were resampled to 0.01 m for data derived from the RGB sensors, and 0.03 m for the multispectral sensors. This enabled layer stacking based on the DOY and sensor. For models which implemented a fusion of both RGB and multispectral data, a resolution of 0.03m was used. Shapefiles of the manually delineated tree crowns were implemented to extract the mean pixel values from all DOY layer stacks and stored in tabular format for further analysis. For the most part, the workflow was required to be automated due to copious amounts of data (e.g. 14 tree crowns from 13 – 15 epochs per year from two sensors). Figure 31 shows the full workflow from image acquisition to model evaluation.

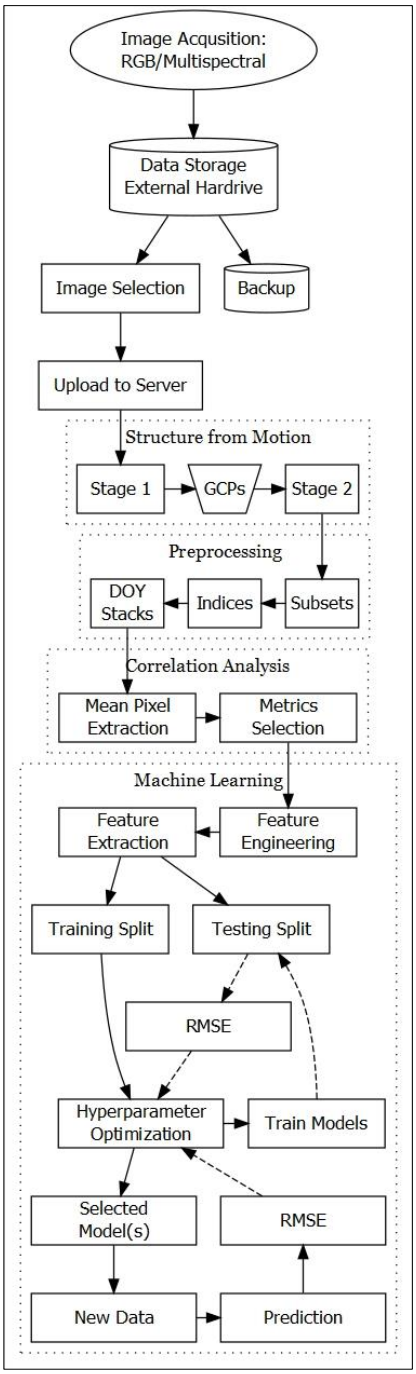


Figure 31: Workflow depicting the various steps from data acquisition to model evaluation.

4.2.5 Vegetation Indices

Vegetation indices are widely used in remote sensing and can not only aid in detecting green vegetation traits but also reduce the effects of irradiance and variations in atmospheric transmission (Jones & Vaughan, 2010; Liang & Wang, 2020). Table 21 displays the indices used in this study. The Green Chromatic Coordinate (GCC) and the Normalized Green Red Difference Index (NGRDI) are the two indices which are derived from the visible part of the electromagnetic spectrum and available for typical consumer grade RGB sensors. Indices denoted with a "UC" (e.g., GCC_UC or GNRDI_UC) depict an index derived from bands that did not undergo radiometric calibration (uncalibrated).

The Normalized Difference Red-edge Index developed originally by Gitelson and M. Merzlyak (1994) was labeled NDREI in conform with Hunt et al. (2013). The NDREI is sometimes labelled NDRE (Barnes et al., 2000), however in this study for practical reasons we use the abbreviation NDRE to depict the use of the Red-edge and Red bands rather than the Near-infrared and Red-edge bands. Another option could be to use the abbreviation ReNDVI used by Wang (2021).

The air temperature feature (AIRTEMP) was created from a summation of the daily air temperature above 0 °C ("warming days") since January 1st of a given year and implemented in this study for experimental purposes.

Table 21: Vegetation Indices (VI) used in the study. NIR = Near-infrared.

Name	VI	Formula	References
Green Chromatic Coordinate	GCC	(Green / Red + Green + Blue)	(Gillespie et al., 1987)
Normalized Green Red Difference Index	NGRDI	(Green - Red) / (Green + Red)	(Tucker, 1979)
Normalized Difference Vegetation Index	NDVI	(NIR - Red) / (NIR + Red)	(Rouse et al., 1974)
Normalized Difference Red-edge Red Index	NDRE (ReNDVI)	($Rededge_{717}$ – Red_{668}) / ($Rededge_{717}$ + Red_{668})	Wang (2021)
Normalized Difference Red-edge Index	NDREI	(NIR – Rededge / (NIR + Rededge)	(A. Gitelson & Merzlyak, 1994b)
Green Normalized Difference Vegetation Index	GNDVI	(NIR - Green) / (NIR + Green)	(A. A. Gitelson et al., 1996)
Enhanced Vegetation Index	EVI	2.5 * (NIR-Red) / (NIR + 6 * Red -7.5 * B + 1)	(Huete et al., 2002)
Normalized Difference Water Index	NDWI	(Green - NIR) / (Green + NIR)	(McFeeters, 1996)

4.2.6 Feature Selection

With the aim of selecting appropriate features for the modelling process, a correlation analysis among independent and dependent variables was implemented. The correlation coefficient is scaleless and represented with the letter *r* which is interpreted with values between -1 and 1 where -1 would represent a perfect negative correlation in that the two variables have an inverse relationship and 1 represents a perfect linear relationship. 0 would depict the situation where both variables have no linear relationship (McClave & Sincich, 2018). In this study, the non-parametric Spearman correlation matrix was implemented for each of the indices in relation to the phenological phase and foliation. A test for multicollinearity was carried out which explored the between-variable correlation among predicting features (Kuhn & Johnson, 2019) for the purpose of improving feature selection for multivariate models. The use of polynomial regression models of the first to fifth order were also used to further evaluate individual features during the selection process.

4.2.7 Statistical Analysis and Machine Learning

The ML regression models implemented in this study included Generalized Additive Models (GAMs), Boosted GAMs and Gradient Boosting Machine (GBM). Model training was conducted using *R's caret* package (Kuhn et al., 2022). The models were trained with an 80/20 split for training and validation, scaling, and 10-Fold cross-validation. The ML modelling process was divided into three main stages: 1) The training/validation split applied to the subset variations of the 2019, 2020, and 2021 datasets, followed by testing for correlation and polynomial fitting for feature selection. 2) Models were trained using the combined 2019 and 2020 datasets and then tested with the 2021 dataset. 3) Further testing involved selected models trained with subset variations of the 2019, 2020, and 2021 datasets, which were then tested against new single-epoch datasets acquired from unseen data. The rationale behind using subset variations was to identify which years contribute to error propagation.

The accuracy of the ML regression models in this study was assessed using three metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared. For interpreting RMSE and MAE values, a phase error is considered functional at values below 1.0 and below 10% for foliation. While the acceptable error margin for operational use has not been conclusively determined, an error below 0.5 for phase and 5% for foliation is deemed similar to, or better than, errors resulting from human observations.

4.3 Results and Discussion

4.3.1 Phenological Data Historical Overview

Phenological observations at the Britz research station were consistently conducted by the same technician from 2006 to 2019. The process for spring leafing out observations began typically before budburst, at the point when buds began swelling (phase 0.5). Following this initial observation, regular visits to the plots were made, often on a daily basis, with the specific aim of capturing the individual tree bud bursting events. Once leafing out was observed for all sampled trees, the frequency of observer visits was reduced to twice a week, and then to approximately once a week after phase 4.0. Observations of fruiting and flowering events were recorded only when they were present, without quantifying their magnitude. Figure 32 shows a comprehensive

overview of the yearly averaged spring phenological observations from 2006 to 2020 for the Beech plot at the Britz research station.

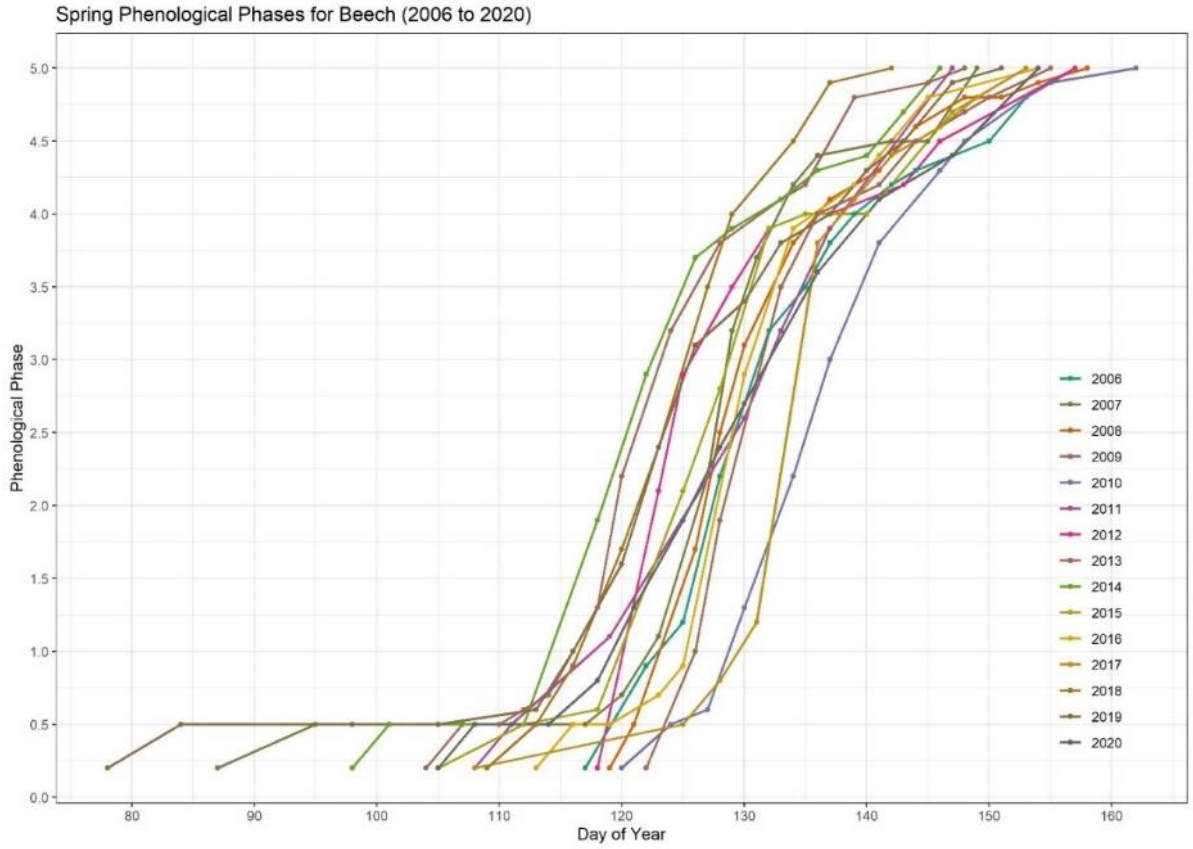


Figure 32: The phenological spring phase development for Beech at the Britz research station between 2006 and 2020.

The analysis of the duration between different phenological phases is crucial for understanding two key aspects: firstly, the timing of budburst in relation to climate change impacts, and secondly, the progression to later stages such as phase 4 and phase 5, when leaves are nearing full development. The "hardening" of leaf cell tissues, which occurs at these later stages, renders the leaves less vulnerable to late frosts, intense early spring solar radiation, and biotic pests like *Orchestes fagi*. Additionally, in early spring drought conditions, certain phases may be delayed, extending the development period from phases 1.0 to 5.0. This phenomenon was observed at the Britz research station in the years 2006, 2012, 2015, and 2019.

Figure 33 in the study visually illustrates the variability in phase duration from 2006 to 2020, which ranged from 23 to 41 days. Meanwhile, Table 22 offers a comprehensive summary with descriptive statistics for the length of time between phases. Predicting subsequent phases based on observations of a single phase would be highly beneficial, especially for determining the dates

of critical "missed" phases at external plots. The phase lengths presented in Figure 33 and Table 22 are derived from the average timings across all sampled beech trees in the phenology plot. For more accurate predictions of other phases based on a single observed phase, it might be more effective to model using data from individual trees, given the significant heterogeneity that can exist among them during the spring phenological phases. Further research in this direction is warranted to explore these possibilities.

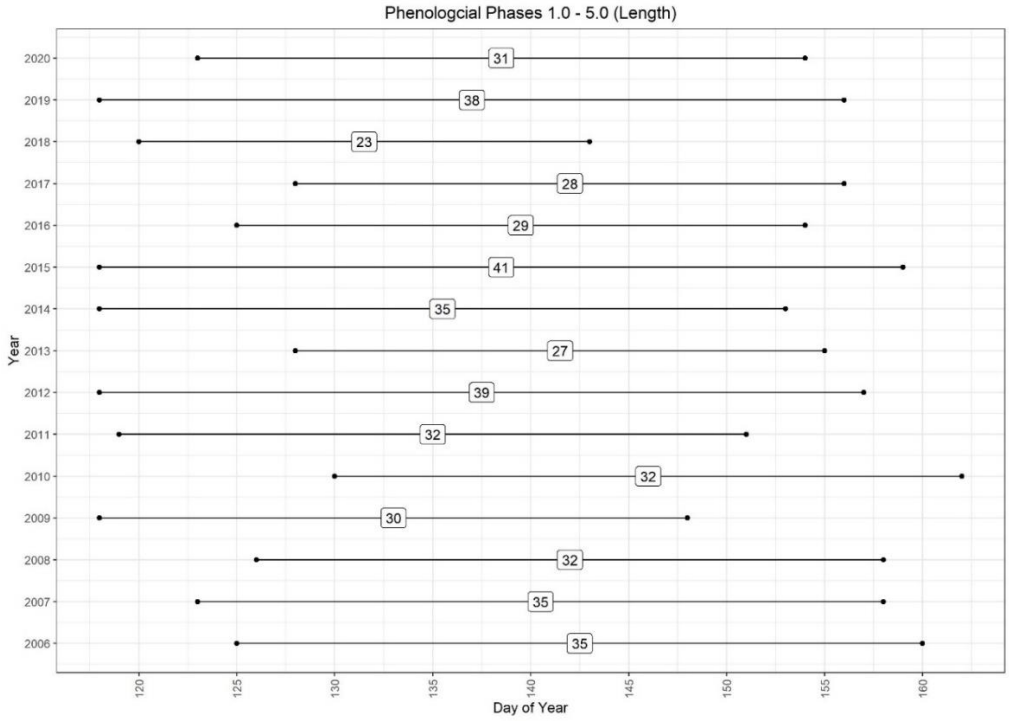


Figure 33: The average spring phenological phases at the Britz research station shown in length between phase 1 and 5 from years 2006 to 2020.

Table 22: An overview of the length between phases from 2006 to 2020. Accuracy is dependent on the temporal resolution of observations.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	Min	Max	Mean	SD
Phase 1-2	4	4	2	2	4	11	6	5	4	8	5	3	3	5	6	2	11	4.8	2.3
Phase 2-3	6	3	4	6	5	5	8	3	5	7	4	5	5	8	7	3	8	5.4	1.6
Phase 3-4	14	13	12	12	11	12	10	9	12	13	8	8	7	14	8	7	14	10.9	2.4
Phase 4-5	11	15	14	10	12	4	15	10	14	13	11	12	8	11	10	4	15	11.3	2.9
Phase 1-4	24	20	18	20	20	28	24	17	21	28	17	16	15	27	21	15	28	21.1	4.3
Phase 1-5	35	35	32	30	32	32	39	27	35	41	29	28	23	38	31	23	41	32.4	4.9

With regard to trends at the research station since 2006, it can been seen that phase 1.0 (see Figure 34; left) and phase 5.0 (see Figure 34; right) are trending towards an earlier onset. A gradual increase in average yearly air temperature (see Figure 35; left) is also evident, alongside a steady decrease in yearly precipitation (Figure 35; right).

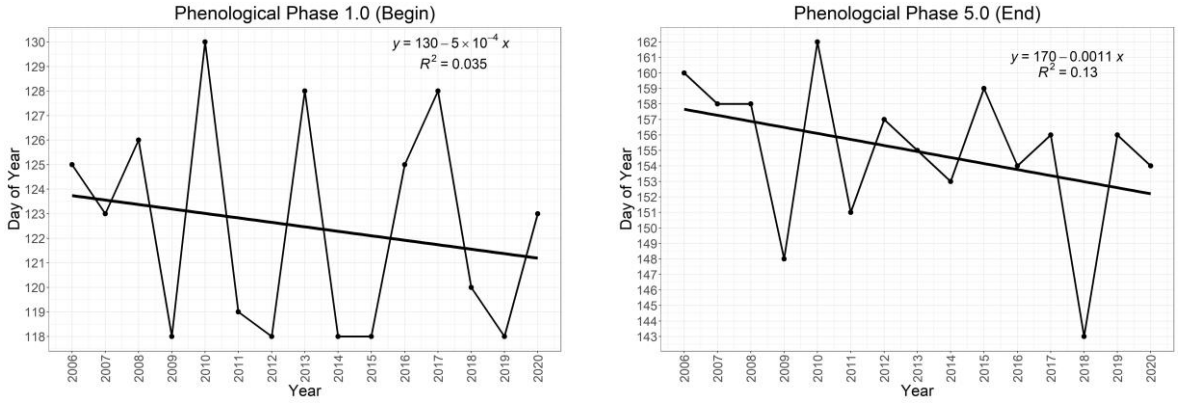


Figure 34: (left) Yearly linear trend the phenological phase 1.0; (right) Yearly linear trend of the phenological phase 5.0.

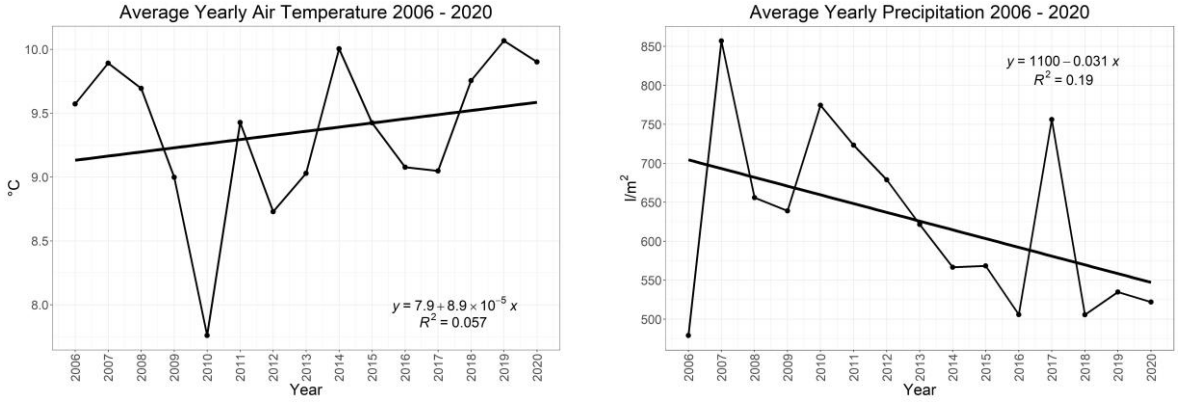


Figure 35: (left) Yearly linear trend of average air temperature between the years 2006 and 2020; (right) Yearly linear trend of average precipitation between the years 2006 and 2020. Both are results from the Britz research station.

Several of the trees used for phenological observations at the research site are equipped with electronic band dendrometers and sap flow measurement devices. Figure 36 depicts the relationship between the phenological phases and the onset of stem growth for tree number 328 during the growth season. Notably, in both 2017 and 2018, the onset of stem diameter growth in this tree coincided with the achievement of phase 3.0, which is marked by the emergence of the first fully unfolded leaves.

The dendrometer data from 2018 reveals significant fluctuations in growth deficit throughout the growth season. These fluctuations align with the prolonged drought conditions reported in that year, as documented by Schuldt et al. (2020). This correlation highlights the impact of

environmental factors, such as drought, on the growth patterns and phenological development of trees, providing valuable insights into the interplay between climatic conditions and tree physiology.

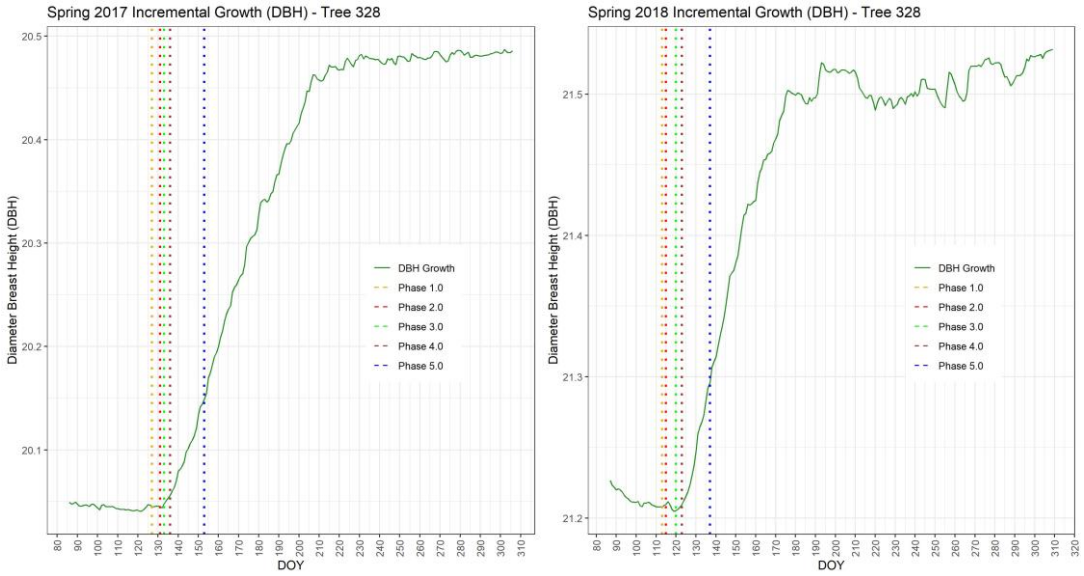


Figure 36: Spring phenological phases shown in relation to band dendrometer measurements from 2017 (left) and 2018 (right). Stem growth typically began around the arrival of phase 3.0.

The analysis of the phase and foliation datasets is further elaborated through the histograms presented in Figure 37. These histograms exhibit a distinct bimodal distribution, characterized by noticeable left and right skewed distributions on the tail ends. This pattern arises from a typical surplus of observations occurring before phase 1.0, which is primarily due to the intensified frequency of observations in anticipation of budburst. Additionally, the extended duration between phases 4.0 and 5.0 contributes to this bimodal distribution. This phenomenon highlights the uneven distribution of observations across different phenological phases, influenced by the varying rates of development and the specific focus of the observation periods.

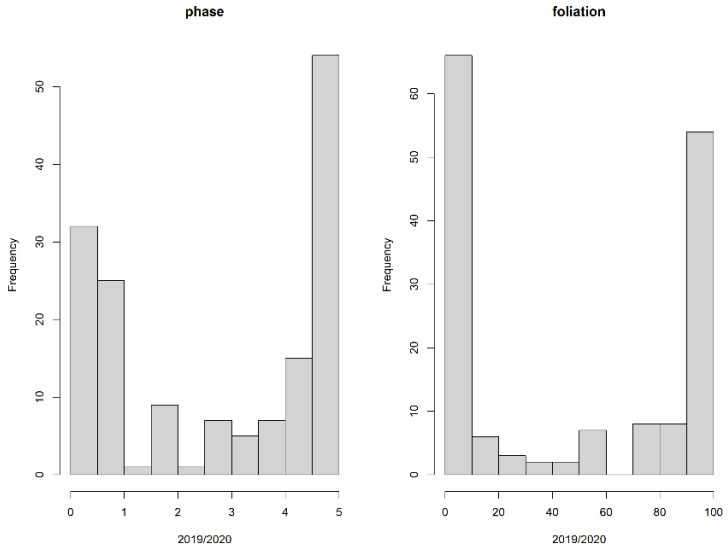


Figure 37: Histograms showing a distinct biomodial distribution of the phase and foliation ground observations from 2019 and 2020.

4.3.2 Correlation Analysis and Feature Selection

Due to the spectral reflectance characteristics of vegetation, visible bands tend to show a positive correlation among each other, whereas the NIR band a negative one (Mather & Koch, 2011). All the vegetation indices whether derived from visible or NIR bands or a combination thereof, have a positive correlation with the phase and foliation datasets except for the NDWI which has typically an inverse relationship with the phases and foliation (see Figure 38). The most consistent index throughout all datasets, whether originating from single or combined years, is evidently the NDVI with a persistent correlation of $r > 0.9$ ($p < 0.001$) over all datasets.

Indices derived from visual bands (i.e. GCC and NGRDI) showed a correlation of $r = 0.65$ ($p < 0.001$) and those uncalibrated even poorer. Interestingly, the AIRTEMP meteorological-based feature correlated very well with the ground observations ($r = 0.9$; $p < 0.001$) with a very high correlation coefficient to the phenological phases at $r = 0.95$ ($p < 0.001$).

In terms of correlation among independent features (see Figure 39), the aim was to refrain from implementing highly correlated features when multiple independent features were incorporated into the modelling process. This could be especially problematic when multiple indices are derived from the same bands (i.e. NDVI and EVI). Here we could deduce that the NDREI and

GCC, when used together for the modelling process, have a lower correlation (r = 0.73) and do not share any similar bands. Likewise, the NDRE and the NDWI do not share the same bands and have a negative correlation coefficient of r = -0.8. The NDWI and the GCC share only the green band and correlate negatively at r = -0.74.

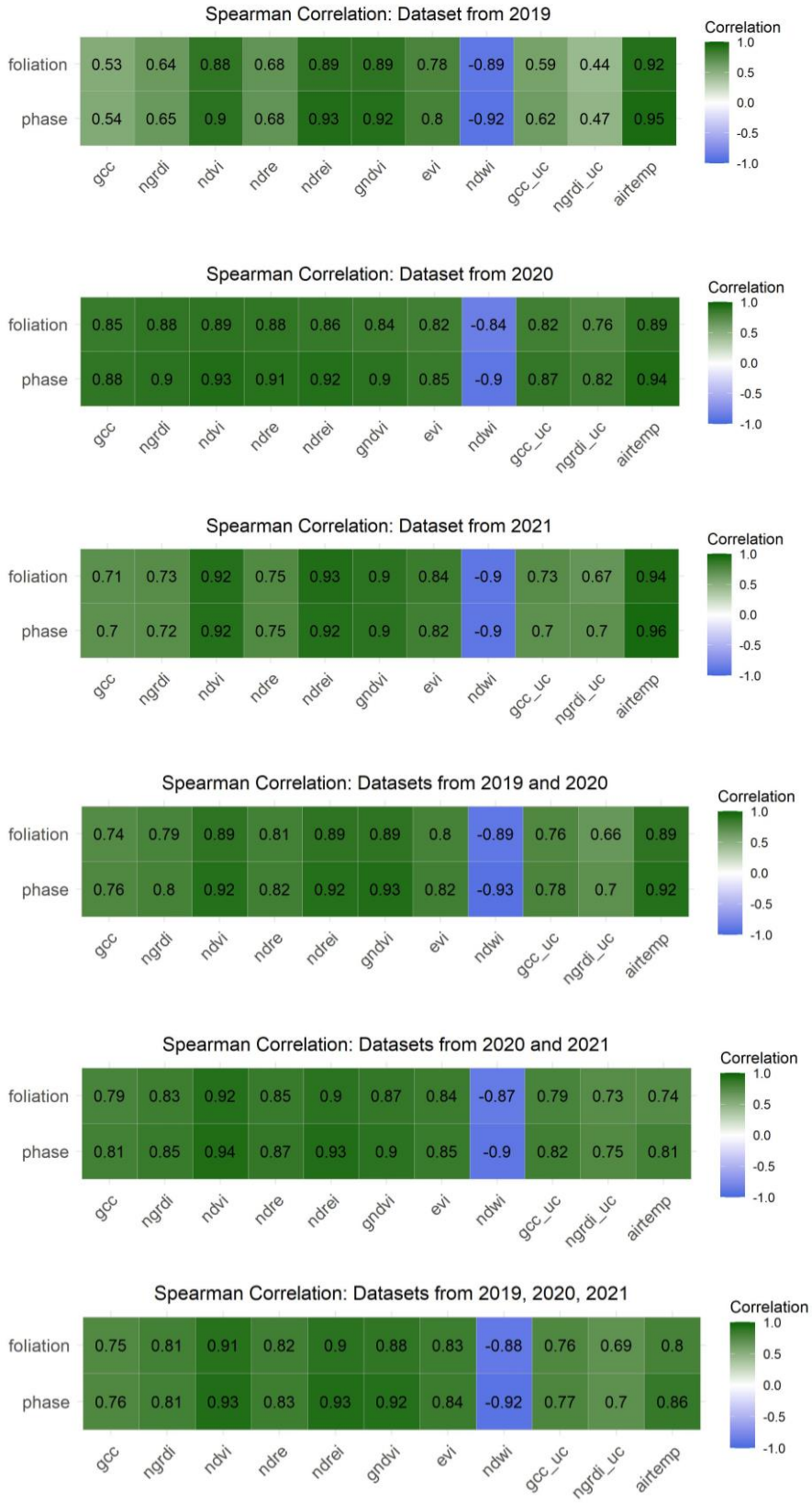


Figure 38: Spearman correlation analysis of the spectral indices derived from the 2019 and 2020 datasets in relation to the ground observations.

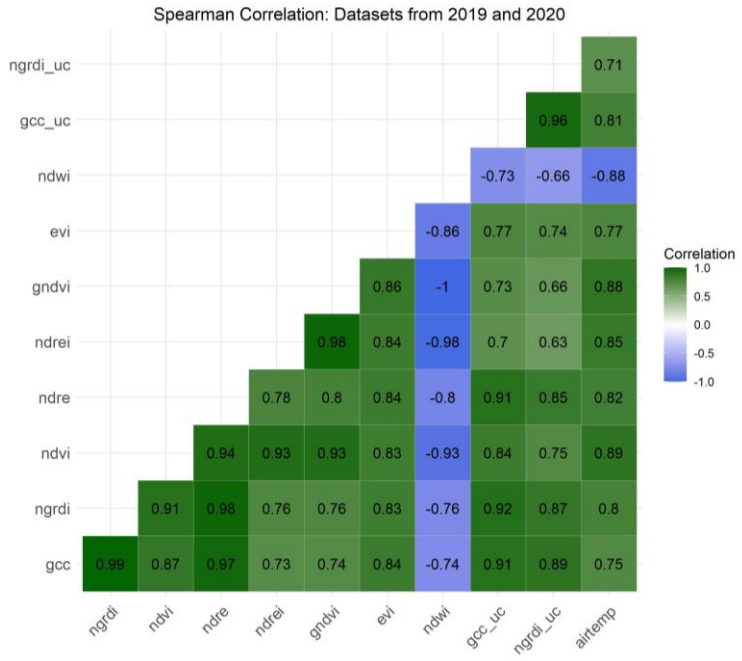


Figure 39: Between-variable Spearman correlation assessment of the 2019/2020 features.

In analyzing the use of correlation for feature selection, it's important to note that while this method is informative, particularly for evaluating multicollinearity, it can potentially be misleading. This is because correlation coefficients might be artificially high due to the bimodal influence on the dataset. The aggregation of data points at the tail ends of the distribution results in a biased similarity, caused by an oversampling of similar phases, thus leading to high correlation coefficients. Consequently, correlation filtering methods were not the sole reliance for feature selection, as outlined by Chandrashekar & Sahin (2014). This approach recognizes the limitations of using correlation analysis in isolation, especially in datasets with unique distribution characteristics like the one described here.

4.3.3 Polynomial Regression and Feature Selection

The addition of polynomial terms into regression models can aid in the characterization of nonlinear patterns (Kuhn & Johnson, 2019) and is conducive to representing phenological trends in particular that of the spring green-up phases. As the polynomial fitting may not be capable in identifying the complexities of phenology metrics in comparison to other algorithms (Rodrigues et al., 2012; Zhu et al., 2012), we used the fitting of polynomials here for the purpose of feature

selection, where the aim was to identify which features best correspond to the typical spring phenology curve. Figure 40 shows the fitting of the five polynomial orders using the example for the NDVI resulting in an RMSE 0.55, MAE of 0.41 and R-squared of 0.91. Here the third polynomial order was deemed the best choice for further analysis where the curve is not oversimplified or too complex.

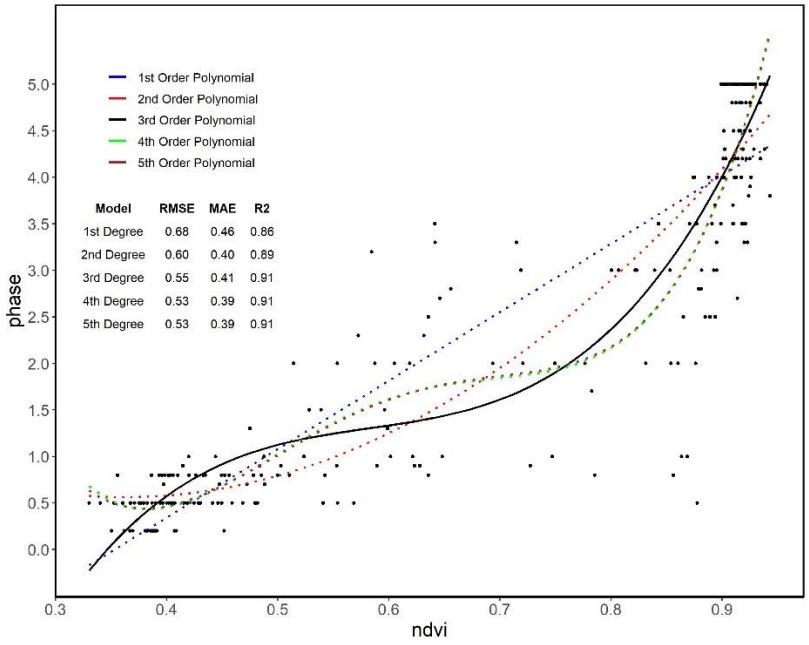


Figure 40: Modelling of the spring phenological phases (2019/2020) dataset with polynomial regression of the first to fifth order.

To follow, each of the selected individual features was tested with the 3rd order polynomial separately for the 2019/2020 and 2020/2021 datasets for both phase (Figure 41) and foliation (Figure 42). In terms of the phenological phases, the GNDVI shows quite a low dispersal of RMSE for the 2019/2020 dataset, yet the dispersal is higher for the 2020/2021 dataset. A similar result is evident for the NDVI where however less dispersal is found in the 2020/2021 dataset rather than that of the 2019/2020 dataset. The cumulative warming days (AIRTEMP) as well as the indices derived from the uncalibrated visible bands (GCC_UC and NGRDI_UC) fared poorly for both datasets. This was also the case for foliation, however AIRTEMP performed better for the 2019/2020 dataset. Regarding foliation, the NDVI also performed well for the 2020/2021 dataset as did the NDREI for both datasets.

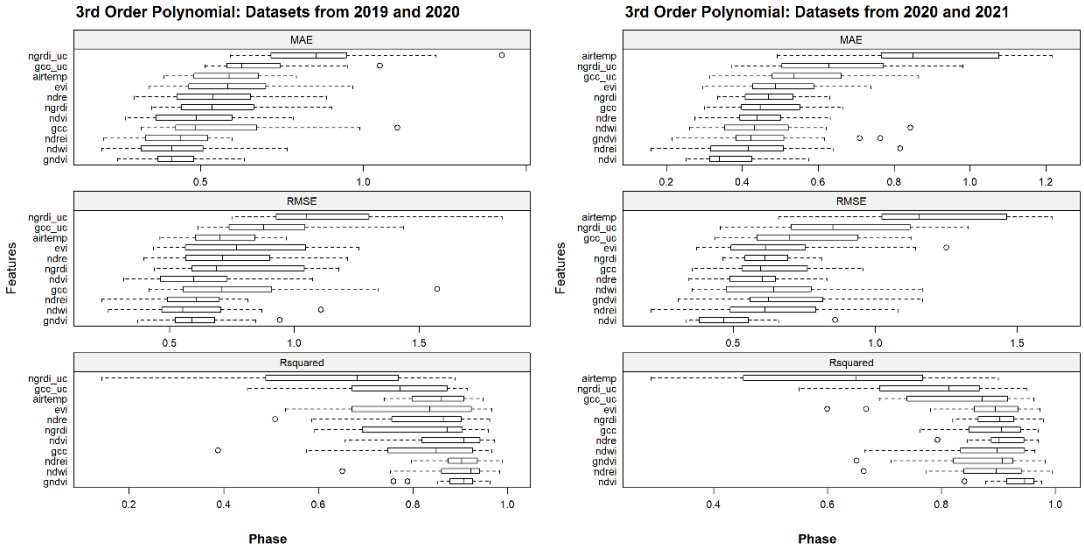


Figure 41: Overview of the spring phenological phases and indices modelled with polynomial regression of the third order for the 2019/2020 (left) and 2020/2021 (right) datasets.

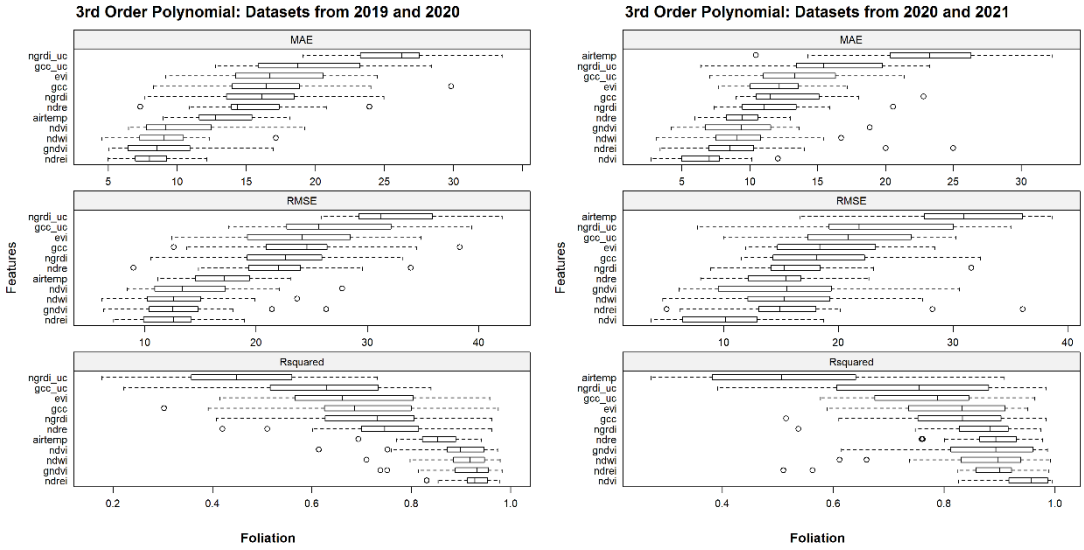


Figure 42: Overview of spring foliation and indices modelled with polynomial regression of the third order for the 2019/2020 (left) and 2020/2021 (right) datasets.

4.3.4 Machine Learning Models: 2019/2020 Datasets

Based on the results of the correlation analysis and polynomial fitting, we were able to select the most relevant features for further scrutinisation during the subsequent modelling process. Important to note here is that in the initial feature selection process using only the correlation analysis alone could have produced an unseen bias due to an aggregation of data points at the tail ends of the datasets which was evident especially for the 2019/2020 dataset. We proceeded

to build three models based on ML algorithms which in effect aided in choosing the best performing algorithms as well as features. Each of the selected individual and combined indices were modelled with each algorithm and evaluated using an 80/20 training/validation data split. This not only helped in choosing the best ML algorithm, but also assisted in a type of model-based feature selection by further narrowing down the selected features. In terms of the phenological phases, an RMSE of ≤ 0.5 (0.6) is deemed as acceptable and similar to the magnitude of potential human error. For the Britz method of foliation, a RMSE of ≤ 10 % is assumed to be acceptable, however some may argue that a RMSE of ≤ 5 % in terms of foliage observations is possible with ground observation. Here it should be reminded that the Britz method of foliation is based on the percentage of leaves which have fully opened rather than fractional cover or greening-up.

Phenological Phases

Regarding the phenological phases, the GAM Boosting algorithm showed overall the best results (see Table 23). The GAM models with the features NDREI + GCC resulted with an RMSE of 0.51, MAE of 0.33 and an R-squared of 0.95. The feature combination of NDWI + GCC resulted in a RMSE of 0.46, MAE of 0.3 and R-squared of 0.96. The top performing model was that of the GAM Boosting with the NDVI and produced a RMSE of 0.28, MAE of 0.18, and R-squared of 0.98. The second-best performing model was that of the GAM model with the NDRE + NDWI input features resulting with a RMSE of 0.44, MAE of 0.31 and R-squared of 0.96. Interestingly, the uncalibrated GCC (GCC_UC) outperformed the calibrated GCC with a RMSE of 0.73 for Gradient Boosting and the GCC_UC index as opposed to a RMSE of 0.81 for GAM Boosting and the GCC.

Table 23: Error metrics for the phase prediction of three Machine Learning algorithms. Values shown in bold font depict the best results.

Phases	**GAM**			**GAM Boosting**			**Gradient Boosting**		
	RMSE	**MAE**	**R^2**	**RMSE**	**MAE**	**R^2**	**RMSE**	**MAE**	**R^2**
GNDVI	0.81	0.5	0.87	**0.8**	**0.49**	**0.87**	0.82	0.5	0.86
NDREI	0.75	0.43	0.89	**0.75**	**0.41**	**0.89**	0.78	0.46	0.89
NDVI	0.32	0.2	0.98	**0.28**	**0.18**	**0.98**	0.38	0.3	0.98
NDRE	0.65	0.53	0.95	**0.64**	**0.53**	**0.95**	0.66	0.58	0.98
GCC	0.84	0.64	0.9	**0.81**	**0.63**	**0.90**	0.82	0.64	0.90
NDREI + GCC	0.55	0.37	0.9	**0.51**	**0.33**	**0.95**	0.6	0.39	0.93
NDWI + GCC	0.48	0.35	0.96	**0.46**	**0.3**	**0.96**	0.49	0.33	0.96
NDRE + NDWI	**0.44**	**0.31**	**0.96**	0.5	0.33	0.95	0.47	0.33	0.96
GCC_UC	0.76	0.62	0.93	0.75	0.62	0.93	**0.73**	**0.62**	**0.96**

At this stage of the modelling process, the NDVI and GAM boosting algorithm showed very good results (RMSE = 0.28) and the question is here whether the dataset is overfit for the Britz research station beech stand. At this point it is an imperative to test the models with unseen data and assess which ones are generalizable over various Beech stands especially those of increased age. In terms of the models derived from indices from the visual bands, the uncalibrated GCC performed slightly better than that of the radiometrically calibrated one and better than some of the models derived from the calibrated multispectral bands which is particularly interesting as RGB sensors are typically acquired at a much cheaper price.

Foliation

For the most part, all models failed the 10 % cutoff point except for those using the NDVI as an input feature. Both the NDVI-based GAM boosting and Gradient Boosting models obtained a RMSE of 7 %, MAE of 4 % and R-Squared of 0.98. Here overfitting could also be factor, however, will still be interesting for further model assessment of the prediction of foliation on a new dataset (2022) as well as datasets outside of the Britz research station. The worst performing models were those utilising the radiometrically calibrated GCC which acquired a RMSE of 22 %, MAE of 16 %, and R-squared of 0.92.

Table 24: Error metrics (in %) for the foliation prediction of three Machine Learning algorithms. Values shown in bold font depict the best results.

Foliation	**GAM**			**GAM Boosting**			**Gradient Boosting**		
	RMSE	**MAE**	**R^2**	**RMSE**	**MAE**	**R^2**	**RMSE**	**MAE**	**R^2**
GNDVI	**22**	**10**	**0.83**	**22**	**10**	**0.83**	22	11	0.8
NDREI	21	11	0.84	21	11	0.84	**21**	**10**	**0.85**
NDVI	**7**	**4**	**0.98**	**7**	**4**	**0.98**	9	6	0.97
NDRE	**16**	**12**	**0.97**	**16**	**12**	**0.97**	**16**	**12**	**0.97**
GCC	**22**	**16**	**0.91**	**22**	**16**	**0.91**	**22**	**16**	**0.92**
NDREI + GCC	17	9	0.90	17	9	0.91	**16**	**8**	**0.91**
NDWI + GCC	**13**	**7**	**0.94**	**13**	**7**	**0.94**	14	7	0.94
NDRE + NDWI	14	8	0.93	14	8	0.93	**14**	**7**	**0.93**
GCC_UC	**21**	**16**	**0.96**	**21**	**16**	**0.96**	21	17	0.97

4.3.5 GAM Boosting Models with Test Datasets

With the aim of testing the robustness as well as generalisability of the developed models, new data from 2022 as well as data from different forest stands (beech) was introduced. Here we tested the models on new spring phenological data from the same stand from 2022 (n = 17) as well as older beech stand in Kahlenberg (n = 10) located in the same region as the Britz research station and a beech stand in the more mountainous region of the Black Forest (n = 8) in south-western Germany. The three test datasets are limited to only one Epoch where the Kahlenberg site is comprised of mostly later phases and the Britz and Black Forest datasets have a wide range of earlier phases (< 4.0). Additionally, training datasets were divided into three different subdivisions based on the year of origin: 2019/2020, 2020/2021 and all datasets together (2019 – 2021). This was carried out for the purpose of distinguishing whether data acquisition methods from a certain year contributed to error propagation. For example, the 2019 field data was carried out by a different observer and often not recorded on the same day as flights (± 3 days) as well as low quality radiometric calibration. The models chosen for testing were those implementing GAM boosting and the RGB-derived indices GCC (Micasense Altum) and GCC_UC (Zenmuse X7) and the NDVI (Micasense Altum). Table 25 displays a list of all of the tested models with reference to the applied index, location, training data subdivision and date.

Table 25: displays the various models tested on the datasets from 2022 and/or outside of the Britz research station. The four models in bold font are those deemed operational.

Model Name	Site	Index	Training	DOY	Year
br-ndvi-19-20	Britz	ndvi	2019-2020	125	2022
br-ndvi-20-21	**Britz**	**ndvi**	**2020-2021**	**125**	**2022**
br-ndvi-19-21	Britz	ndvi	2019-2021	125	2022
br-gcc-19-20	Britz	gcc	2019-2020	125	2022
br-gcc-20-21	**Britz**	**gcc**	**2020-2021**	**125**	**2022**
br-gcc-19-21	Britz	gcc	2019-2021	125	2022
br-gcc-uc-19-20	Britz	gcc_uc	2019-2020	125	2022
br-gcc-uc-20-21	Britz	gcc_uc	2020-2021	125	2022
br-gcc-uc-19-21	Britz	gcc_uc	2019-2021	125	2022
bf-gcc-19-20	**Black Forest**	**gcc**	**2019-2020**	**115**	**2022**
bf-gcc-20-21	Black Forest	gcc	2020-2021	115	2022
bf-gcc-19-21	Black Forest	gcc	2019-2021	115	2022
bf-gcc-uc-19-20	Black Forest	gcc_uc	2019-2020	115	2022
bf-gcc-uc-20-21	Black Forest	gcc_uc	2020-2021	115	2022
bf-gcc-uc-19-21	Black Forest	gcc_uc	2019-2021	115	2022
ka-ndvi-19-20	Kahlenberg	ndvi	2019-2020	119	2020
ka-ndvi-20-21	Kahlenberg	ndvi	2020-2021	119	2020
ka-ndvi-19-21	Kahlenberg	ndvi	2019-2021	119	2020
ka-gcc-19-20	Kahlenberg	gcc	2019-2020	135	2020
ka-gcc-20-21	Kahlenberg	gcc	2020-2021	135	2020
ka-gcc-19-21	Kahlenberg	gcc	2019-2021	135	2020
ka-gcc-uc-19-20	Kahlenberg	gcc_uc	2019-2020	135	2020
ka-gcc-uc-20-21	**Kahlenberg**	**gcc_uc**	**2020-2021**	**135**	**2020**
ka-gcc-uc-19-21	Kahlenberg	gcc_uc	2019-2021	135	2020

Results of the model testing of the phenological phase prediction (see Figure 43) and foliation (see Figure 44) were ranked in order of the RMSE. Noteworthy is that all the models of the phenological phase prediction that achieved the 0.5 threshold (left of green dotted line) were those of the calibrated and uncalibrated GCC which originate from bands of the visible portion of the electromagnetic spectrum. Five of six of these models were from the Kahlenberg dataset and one from the Black Forest dataset. The best performing models were selected for each of the test sites and are mapped out in Figures 16-19. All image data acquired for the test sites with the *Zenmuse X7* lack radiometric calibration except for the Britz dataset (see Figure 48) which was acquired with the with both the *X7* and radiometrically calibrated *Micasense Altum* data.

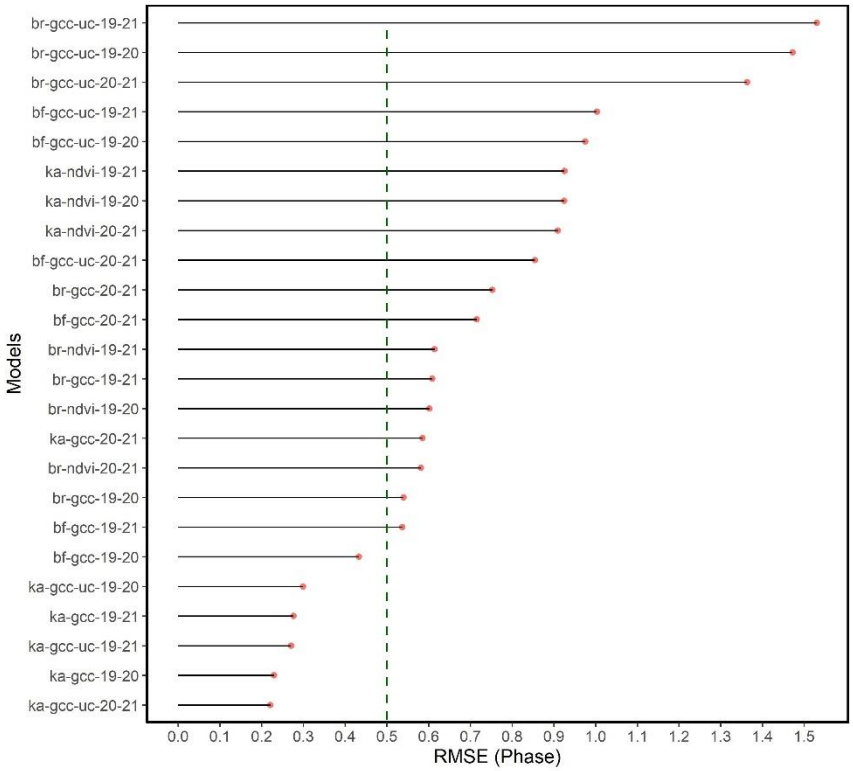


Figure 43: graph showing the RMSE for the phase prediction ranked in order from poorest to best RMSE. The green dashed line depicts the cut-off point of acceptable accuracy. Allowing an RMSE of up to 0.6 would enable the NDVI model derived from the multispectral datasets. Otherwise, only models originating from the visible bands are considered operational.

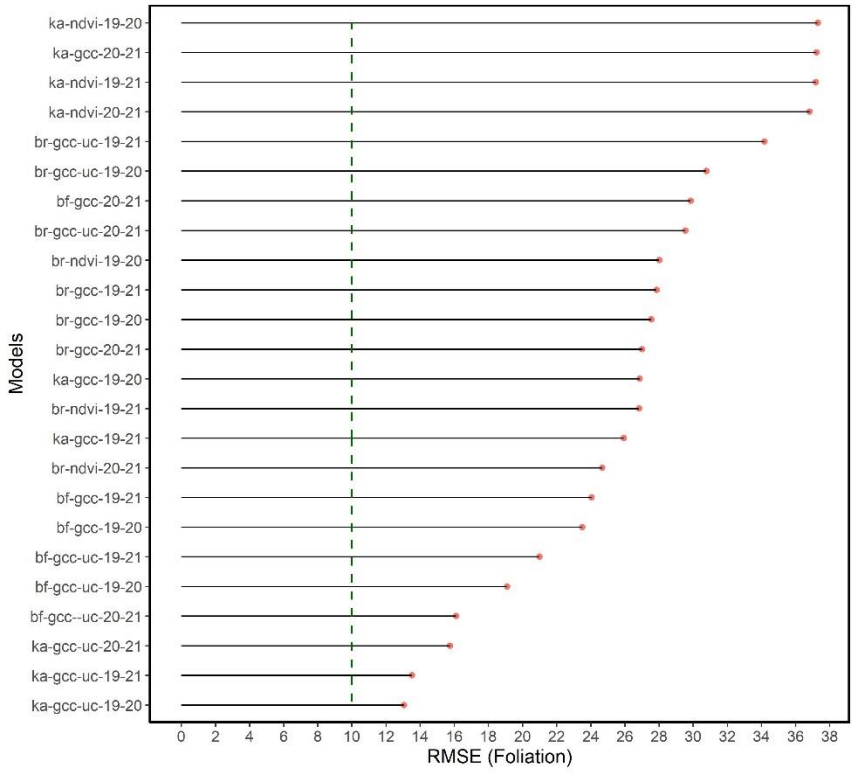


Figure 44: graph showing the RMSE for foliation prediction ranked in order from poorest to best. The green dashed line depicts the cut-off point of 10 %. None of the models for foliation prediction are considered functional.

The Kahlenberg dataset (see Figure 45) with the *gcc-uc-20-21* model resulted with a very low RMSE of 0.22, MAE of 0.16 and R-Squared of 0.08 ($n = 10$). Such a low RMSE for an uncalibrated RGB-based model is an unexpected result here and shows that the later phases, in particular phase 4.0 predict well. Phase 4.0 is a significant phase in the spring green-up as it corresponds to the completion of all leaf and shoot development. The transition to Phase 5.0 would then follow with the hardening of leaf tissue alongside a change to darker green and increased late-frost hardiness.

Sensor: RGB Zenmuse X7
Location: Kahlenberg (> 100 years)
Date: 2020-05-14
DOY : 135
Model: ka-gcc-uc-20-21

Tree ID	ML Phase	Phase	Difference
2	4.1	4.0	0.1
3	4.0	3.8	0.2
13	4.0	4.0	0.0
17	4.0	4.0	0.0
18	4.0	3.8	0.2
19	4.1	3.5	0.6
20	4.0	3.8	0.2
27	4.0	3.9	0.1
38	4.0	4.0	0.0
39	4.1	3.9	0.2
RMSE	**MAE**	**R^2**	**Mean Dif.**
0.22	**0.16**	**0.08**	**0.2**

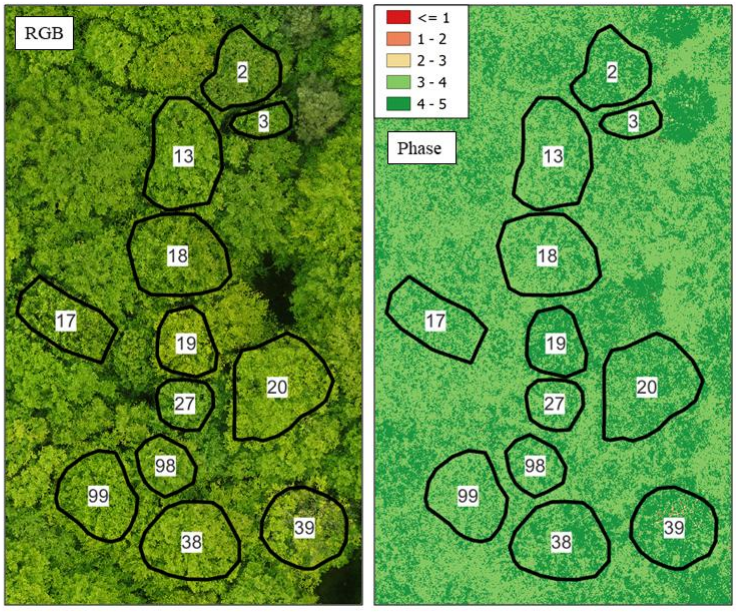


Figure 45: Phase prediction of an older Beech stand (> 100 years) utilising the model originating from the uncalibrated GCC 2020/2021 dataset. The very low RMSE of 0.22 proves a highly generalizable model however it should be noted that this is a relatively small dataset (n = 10) and comprised of only later phases (> 3.0). "ML phase" is the predicted phase and "Phase" originates from the ground-based observations.

Regarding the Black Forest dataset with the *bf-gcc-19-20* model, a RMSE of 0.43, MAE of 0.32, and R-squared of 0.02 (n = 8) was achieved (see Figure 46). Here, a scene with a wide range of phases (0.9 – 3.8) was available, and a successful phenological phase prediction was possible with the calibrated GCC model and training data from 2019 and 2020. Important to note is that the radiometrically calibrated GCC model was used to predict on the GCC which is derived from the non-calibrated *Zenmuse X7*. Significant here is that sensor mixing in terms of model training with the multispectral sensor and prediction with a consumer grade RGB sensor is attainable. We considered the low R-Squared as insignificant due to the overall low sample rate of the test datasets.

Sensor: RGB Zenmuse X7
Location: Black Forest (< 70 years)
Date: 2022-04-25
DOY : 115
Model: bf-gcc-19-20

Tree ID	ML Phase	Phase	Difference
1	2.4	2.5	-0.1
2	1.4	1.5	-0.1
3	1.5	1.3	0.2
4	3.2	3.8	-0.6
5	2.6	3.0	-0.4
6	0.8	0.9	-0.1
7	1.6	1.5	0.1
8	1.8	2.7	-0.9
RMSE	**MAE**	**R^2**	**Mean Dif.**
0.43	**0.32**	**0.02**	**-0.25**

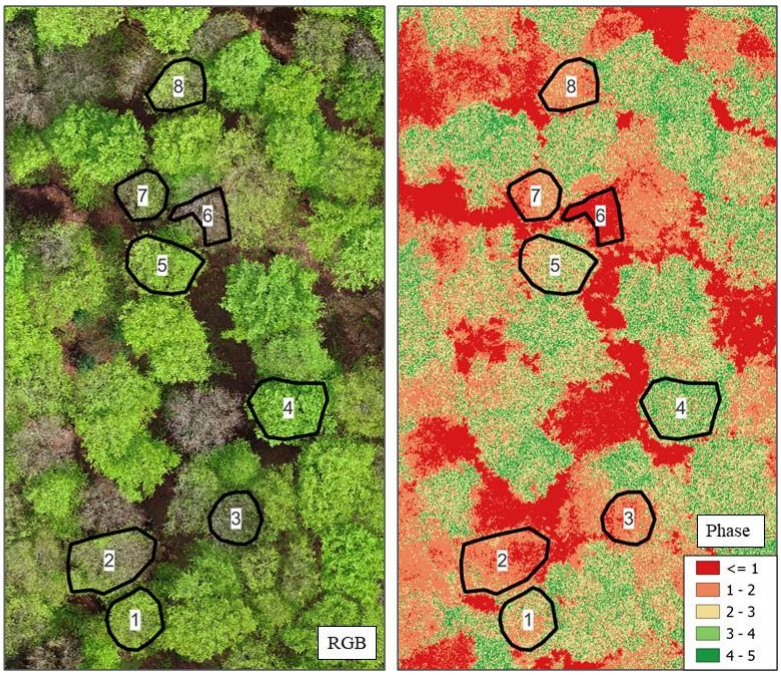


Figure 46: Phase prediction of a Beech stand (< 70 years) utilising the model originating from the calibrated GCC 2019/2020 dataset. The Black Forest dataset is a particularly challenging one as a wide range of phases are available. An RMSE of 0.43 is within the accepted error cut-off of ≤ 0.5.

The Britz dataset (see Figure 47) also implemented the GCC and 2019/2020 training model (*br-gcc-19-20)* and resulted in a RMSE of 0.54, MAE of 0.45 and R-squared of 0.65 (n = 17). Important to note is that the Britz test dataset possesses more samples than other test sites and achieved the 0.5 threshold. This test dataset, however, comprises the same trees as those in the training dataset, providing the model with an advantage at the Britz test site. It's important to note, though, that this advantage might not extend to other test sites, potentially limiting the model's ability to generalize well in different settings.

Sensor: RGB Zenmuse X7
Location: Britz Research Station (47 years)
Date: 2022-05-04
DOY: 125
Model: br-gcc-19-20

Tree ID	ML Phase	Phase	Difference
328	2.8	3.0	-0.2
341	1.5	1.8	-0.3
347	1.9	2.3	-0.4
350	1.9	2.0	-0.1
359	1.2	0.9	0.3
365	1.8	2.0	-0.2
379	1.1	0.6	0.5
382	1.9	2.2	-0.3
383	1.5	0.8	0.7
384	2.4	2.3	0.1
385	1.9	2.3	-0.4
386	1.6	0.9	0.7
394	2.2	3.3	-1.1
403	1.1	0.6	0.5
416	1.9	1.8	0.1
423	1.7	1.2	0.5
425	1.6	2.8	-1.2

RMSE	MAE	R^2	Mean Dif.
0.54	0.45	0.65	-0.04

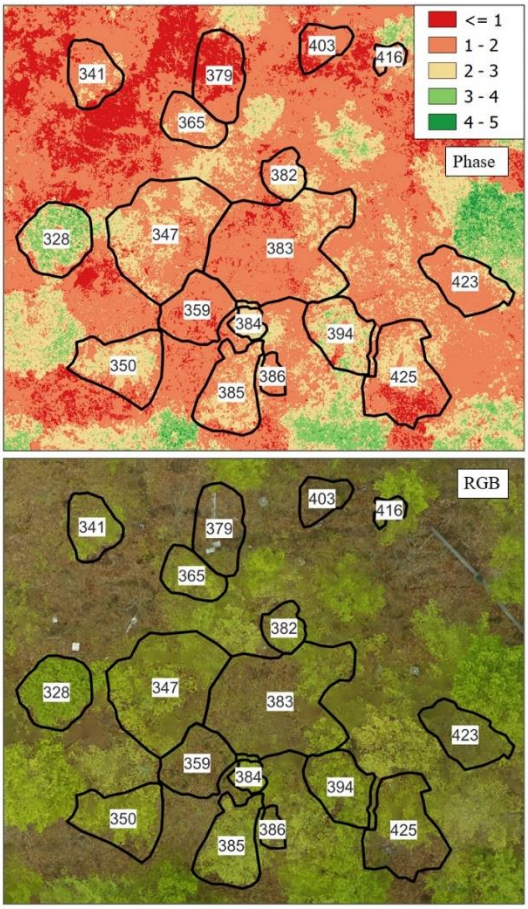


Figure 47: Phase prediction of a Beech stand (47 years) utilising the model originating from the calibrated GCC 2020/2021 dataset. Despite being a larger dataset (n = 17) in comparison to the other test sites, a RMSE of 0.54 was achieved which is can be regarded as achieving the 0.5 threshold.

With respect to the test sites involving phase prediction from the multispectral sensor (*Micasesense Altum*), only the Britz and Kahlenberg sites were available. The only NDVI-based model that was in proximity of the 0.5 threshold was Britz test dataset (*br-ndvi-20-21)* with a RMSE of 0.61, MAE of 0.52, and R-squared of 0.58 ($n = 17$). We hypothesized that the radiometric calibration methods from 2019 will influence the model accuracy, however there was only a marginal difference in RMSE of the 2019/2020 and 2020/21 datasets.

Overall, the best performing and most consistent model for the purpose of predicting the spring phenological phases was that of the calibrated GCC based on the 2019/2020 training dataset. This model could generalize well over the test sites with the highest RMSE being from the Britz 2022 test site (RMSE = 0.54).

Sensor: Multispectral Micasense Altum
Location: Britz Research Station (50 years)
Date: 2022-05-04
DOY: 125
Model: br-ndvi-20-21

Tree ID	ML Phase	Phase	Difference
328	2.7	3.0	-0.3
341	1.7	1.8	-0.1
347	1.8	2.3	-0.5
350	2.0	2.0	0.0
359	1.2	0.9	0.3
365	1.8	2.0	-0.2
379	0.9	0.6	0.3
382	1.9	2.2	-0.3
383	1.4	0.8	0.6
384	2.0	2.3	-0.3
385	1.9	2.3	-0.4
386	1.8	0.9	0.9
394	2.1	3.3	-1.2
403	1.1	0.6	0.5
416	1.6	1.8	-0.2
423	1.8	1.2	0.6
425	1.6	2.8	-1.2
RMSE	**MAE**	**R^2**	**Mean Dif.**
0.61	**0.52**	**0.58**	**0.09**

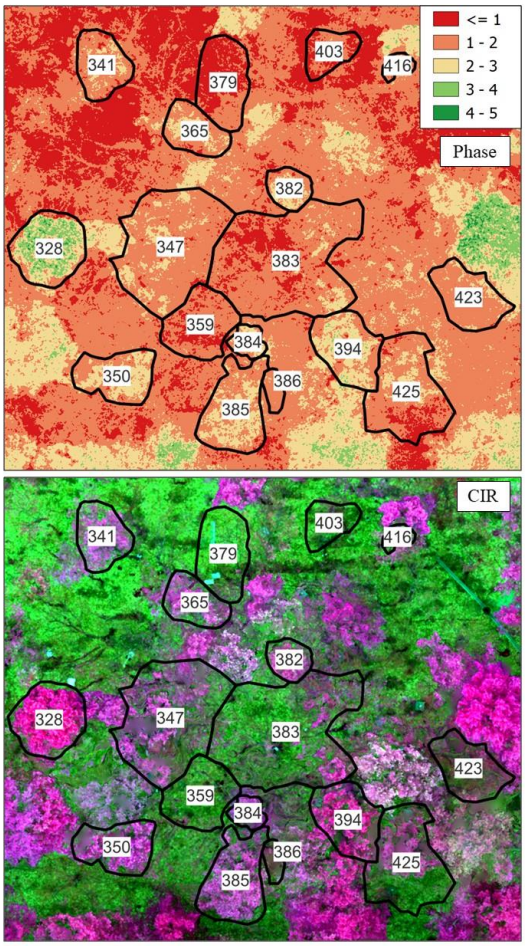


Figure 48: Phase prediction of a Beech stand (50 years) utilising the model originating from the calibrated NDVI 2020/2021 dataset. This is the only model derived from the non-visible band (NIR) which is in the proximity of the 0.5 threshold RMSE = 0.61). CIR = Color-infrared.

4.3.6 Sources of Error and Synopsis

This research highlights the challenges in obtaining radiometrically calibrated datasets over multiple growing seasons, despite pre- and post-mission calibration panel acquisition and DLS data usage. Issues arise when reflectance values bottom out, such as during the calculation of NDVI or other indices involving the NIR band, which occurs when clouds part temporarily during flight missions, exposing the terrain to direct sunlight. This issue of oversaturation in the NIR band was also reported by Wang (2021). While the DLS compensates for fluctuations in irradiance, it is effective only for global changes in lighting conditions (micasense.com). While the DLS compensates for fluctuations in irradiance, it is effective only for global changes in lighting conditions (micasense.com). The problem is exacerbated in dense forests, where

obtaining shadow-free reference panels is nearly impossible, and capturing calibration data at a different location before and after missions is impractical. This could result in time differences to the actual flight mission, during which considerable changes in solar angle might occur.

The size of reflectance panels also impacts the difficulty of radiometric calibration. Honkavaara et al., (2020) showed that larger custom-made reference panels of 1 x 1 meters calibrated better than the manufacturer's provided method. Some studies also demonstrated improved calibration methods using even larger reflectance tarps (Honkavaara et al., 2012; H. Li et al., 2015; Moran et al., 2001), however still does not alleviate the problem of acquiring calibration data in dense forests as well as previously mentioned sudden changes illumination. Therefore, further testing and development of improved field radiometric calibration strategies are imperative to more effectively utilize multispectral sensors' capabilities.

Despite the challenges with multispectral sensors, particularly in the NIR band, the utility of the RGB bands is notable. Low-cost UAV setups with RGB sensors are widely available, facilitating the collection of vast data amounts. This high data volume is crucial for developing models for various tree species at intensive monitoring plots. A key question is whether training data for models derived from visible bands need calibration from the multispectral sensor. In this case, the model trained with calibrated GCC generalized well with the uncalibrated GCC, but it remains to be seen if this holds for new datasets and other tree species.

Errors can also arise from crown segmentation in pixel value extraction. For instance, branches from a neighboring tree with earlier phenological onset could overlap into the segmented crown area of the target tree. As segmentation is typically done with a fully developed canopy (after phase 5.0), such overlapping branches are challenging to account for. Recording influential branches from neighboring trees during ground observations and excluding them from training datasets could improve the quality of training data.

The feature selection process in this research, especially partitioning training datasets by year for testing, was effective. It allowed for scrutinizing and removing training data portions that could affect model generalizability. For instance, the *br-ndvi-20-21* derived from multispectral sensors, excludes the 2019 dataset due to its lower quality radiometric calibration, time differences between observations, a slightly different multispectral sensor, and a different observer for ground observations. Conversely, the *gcc-19-20* models generalized well with the 2019 datasets

incorporated, using only bands from the visible spectrum. This suggests that the main factors in error propagation lie in the quality of radiometric calibration and sensor mixing with NIR bands, a conclusion that might not have been apparent without partitioning training by year. Interestingly, sensor mixing does not seem to be an issue with RGB imagery, which is advantageous for acquiring large data volumes.

Incorporating meteorological data, such as "warming" days (AIRTEMP), as a model feature suggests that other factors, like a dynamic start date and chilling days, should also be considered for a successful phenological model in fusion with spectral data. However, this concept is somewhat limited, as meteorological data at the individual tree level might not explain the heterogeneity of individual trees in phenological development. The fusion of meteorological and spectral data is more suited for larger scale applications, where phenological data is applied stand-wise rather than at the individual tree level.

Regarding the *Britzer* foliation method, translating ground observations into remote sensing data was not feasible. Consequently, the *Britzer* method of foliation has been abandoned at the Britz research station and replaced with the ICP flushing method. Currently, the long-term *Britzer* phase method, alongside the flushing method, is conducted with the aim of simplifying observations and enabling harmonization of Britz research station data with the ICP network at the international level.

4.4 Conclusion and Future Outlook

The research focuses on a machine learning approach for predicting spring phenological phases of European beech using UAV multispectral data. Over three years (2019–2021), synchronous ground observations and UAV-derived multisensor indices were used to train and validate a variety of machine learning models. A comprehensive feature selection method was employed, incorporating Spearman correlation, polynomial fitting, and machine learning techniques.

The models were further evaluated using unseen data, and the effectiveness of various training data partitions by year was assessed to identify potential sources of error. The most effective combination of training data partition, vegetation index, and machine learning algorithm was found to be the 2019/2020 dataset, the Green Chromatic Coordinate (GCC), and GAM boosting.

This model achieved a Root Mean Squared Error (RMSE) of 0.22 at the Kahlenberg site, 0.43 at the Black Forest site, and 0.54 with 2022 data at the Britz Research Station.

However, it was observed that the *Britzer* method of foliation could not be modeled successfully, with RMSE values significantly exceeding the 10% error threshold. The study's findings underscore the potential of a feature selection-based machine learning pipeline that leverages radiometrically calibrated visible bands. This approach is capable of predicting spring phenological phases using RGB imagery obtained from widely available, low-cost sensors. The research thus contributes to the advancement of accessible and accurate phenological modeling using UAV technology.

The *Britzer* phenological phase method is particularly valuable at intensive monitoring sites like the Britz research station, but its complexity may not be as practical for use at external plots. In such cases, the ICP Forests flushing method presents a more suitable option. The integration of UAV-based data with the ICP Forests flushing method has the potential to enhance the existing datasets by providing predictions of phenology flushing on a stand-wise basis and over larger areas.

Implementing large-scale mapping of phenological flushing could pave the way for training and validating models using data obtained from satellite platforms. This approach would enable the upscaling of phenological data, facilitating broad-scale mapping applications in forest phenology. Moreover, it would assist in creating historical phenological time-series maps, which are crucial for assessing the impacts of climate change with a spatial perspective.

Such advancements would not only improve the accuracy and scope of forest phenology monitoring but also contribute significantly to our understanding of how climate change affects forest ecosystems. This integration of various data sources and methods exemplifies the potential of combining traditional ground-based methods with modern UAV and satellite technologies for ecological research and monitoring.

Extensive research and experimentation remain necessary in this field, and several key areas are recommended for future investigation. Firstly, there should be a focus on visible bands (RGB) and testing the feasibility of sensor mixing. This involves exploring the potential and limitations of combining different sensor data for more comprehensive analyses.

Another important aspect is the acquisition of training data, where efforts should be made to ensure an even distribution of phenological phases. Achieving this requires well-coordinated and synchronous cooperation with ground observation teams. In line with this, the development of models for additional tree species present at intensive monitoring plots is crucial. This would enhance the scope and applicability of the research.

Further, adapting the ICP Forests flushing levels to UAV-based modeling on an international scale is essential. This adaptation process may involve regionalizing models to account for geographic and climatic variations. Special training data acquisition campaigns, involving ground observation experts, should be conducted to acquire more continuous flushing values (e.g., 5% levels) rather than relying solely on five code classes for model training. This approach would enable the use of regression-based machine learning models, which are often more effective than classification-based models and facilitate the collection of high-quality training data.

There is also a need for developing models that can predict other phenological phases or flushing levels from a single observation epoch. This capability is crucial for determining the accurate timing of when flushing or specific phases occur, moving beyond the limitations of statistical interpolation methods. Finally, translating observations or machine learning predictions to the stand level is significant for the development of models suitable for satellite applications. Such advancements would greatly benefit large-scale mapping and monitoring efforts, providing more precise and comprehensive insights into forest phenology.

5 General Conclusions and Outlook

In this dissertation, the possibility to apply recent technological advancements in UAV-RS was explored. Originally determined a "toy" by work colleagues during the initial stages of this research, it was clearly shown that UAV-RS with regard to intensive forest monitoring is a viable technology which is here to stay. The possibilities of implementation will most probably grow exponentially in coming years, not only in terms of UAV flight capabilities but also the capacity of sensors and prediction methods. To follow, the main research questions of this dissertation are revisited followed by a brief description of a standardization framework which could be decisive in advancing the utility of UAV-RS for intensive forest monitoring.

5.1 (i) Is UAV-based photogrammetric tree height extraction comparable to traditional methods?

It was shown that photogrammetric tree height measurements can acquire similar accuracies to traditional field methods as well as a potential increase in certainty in terms of repeatable measurements. This is significant for intensive forest monitoring purposes, as using UAS-RS methods could drastically speed up the tree height measurement process but also simultaneously deliver data on tree crown diameter and individual tree geolocations as well as tree species to name a few applications. Important here, is that with every flight mission a multipurpose dataset is created which will not only deliver reliable tree height information, but also provide additional information while exceeding time and cost expectations. The methodology shown in Section 2, should be implementable at an operational level providing accuracies are tested against traditional field methods, LiDAR or destructive methods.

5.2 (ii) Can UAV thermal imaging accurately detect drought stress in Beech?

It was also shown that UAV-based thermal imaging is a promising close-range imaging method which could potentially aid in modelling drought stress information from point dendrometers and meteorological data. Thermal imaging, is however, a challenging undertaking especially when environmental factors such as rapidly changing cloud cover and high winds are present. It was demonstrated that taking various factors such as "spot size" and distance to the object into

account during thermal imaging in the field can aid in reducing inaccuracies. Translating thermal imagery into reliable drought stress information still requires further research, however the application of meteorological data in conjunction with point dendrometers to model TWD shows potential and further research using more point dendrometer sensor data as well as higher resolution meteorological data is warranted. In terms of the LVPD as a stand-alone drought stress index, it requires more validation and the development of threshold ranges for assessment purposes.

5.3 (iii) Are UAV-derived phenological ML models accurate enough to replace ground observations?

With regard to phenological observations, it was made clear that UAV-RS prediction methods can also rival ground observations. It was shown that indices derived from the visible range of the spectrum can actually perform better than those derived from considerably more expensive multispectral sensors despite radiometric calibration. This is significant, as access to RGB sensors is very high and usage is widespread even on typical consumer drones. Additionally, it is significant that the manufacturer-prescribed radiometric calibration methods, despite being time consuming and requiring more processing expenditure, do not deliver the expected results especially over time series that span multiple growth seasons. Here, it is an imperative to develop a more customized and comprehensive calibration methods, which can account for substantially changing solar illumination conditions during flight missions.

The methods displayed in the dissertation, although optimistic, still require further extensive experimentation and validation, and there remains an array of other intensive forest monitoring parameters to explore. In order for UAV-RS to be effective however, a great deal of comparable RS data as ground truthing is required which will only be obtainable with the development of a standardized data infrastructure where large amounts of data can accrue.

5.4 Recommendations for Standardization and Outlook

5.4.1 Towards Standardisation

The standardization of UAV-RS methods for intensive forest monitoring should be based on the development of a framework which links UAV-based products and corresponding in-situ terrestrial monitoring data of intensive monitoring plots to create not only reliable upscalable ground-truthing data, but also features for modelling purposes. Here, a method for the geometric and radiometric calibration at core plots is an imperative where the accuracy and repeatability of data acquired from various UAV-based sensors can be assessed and calibrated. Calibration results and other relevant sensor parameters are then built into the metadata model which connects time-sensitive UAV datasets with high quality ground observations. A collaboration between ground-observation crews and UAV remote sensing specialists should be emphasized, with the aim of harmonizing data acquisition methods while encouraging a fusion of the two methodologies. Such a harmonization, when applied at the national or even international level, could enable unprecedented possibilities for small-scale forest status mapping. Future difficulties could arise however, in terms of the "desiloisation" of data among partners which when resolved could enable the rapid access to analysis ready "data warehouses" for modelling applications as well as "data lakes" for the development feature engineering techniques for data-driven deep learning (DL) models. A substantial challenge does however still exist in terms of ground observations and UAV-RS harmonization, and how methods are evaluated and decisions are made.

5.4.2 Expert Panels: State Partners and ICP Forests

A typical method in evaluating new methods, as well as decision making in ICP Forests is the use of Expert panels. Here, experts on a particular topic (i.e. Phenology) will gather at a workshop type of meeting or in the field, and systematically discuss and make decisions on methodology. Afterwards, decided methods are then recorded in the ICP Forests manual (Lorenz, 1995) and made available to not only project partners but also made publicly available (http://icp-forests.net/page/icp-forests-manual).

In terms of the harmonization of UAV-RS with ground observations, an example is shown in Figure 49 where autumn phenological ground observations are enhanced with the view from the

aerial perspective. The aerial perspective enables the ability to carry out assessments despite the ground view being obscured by other tree crowns which is sometimes the case in very dense stands. Additionally, the images acquired from the UAV can be incorporated into a flight plan which can be repeated various time throughout the phenological phases automatically. These images can then be used at a later date for assessment from a computer screen while enabling an overview with the full range of developmental stages. This could in turn result in more accurate assessments as well as eliminate sources of error which could arise from ground observations in terms of multiple observers of different skill levels and fatigue. Furthermore, ML methods could be deployed for the automatic prediction of crowns from the aerial images. Training data for such models could potentially be created at the computer via expert annotations.

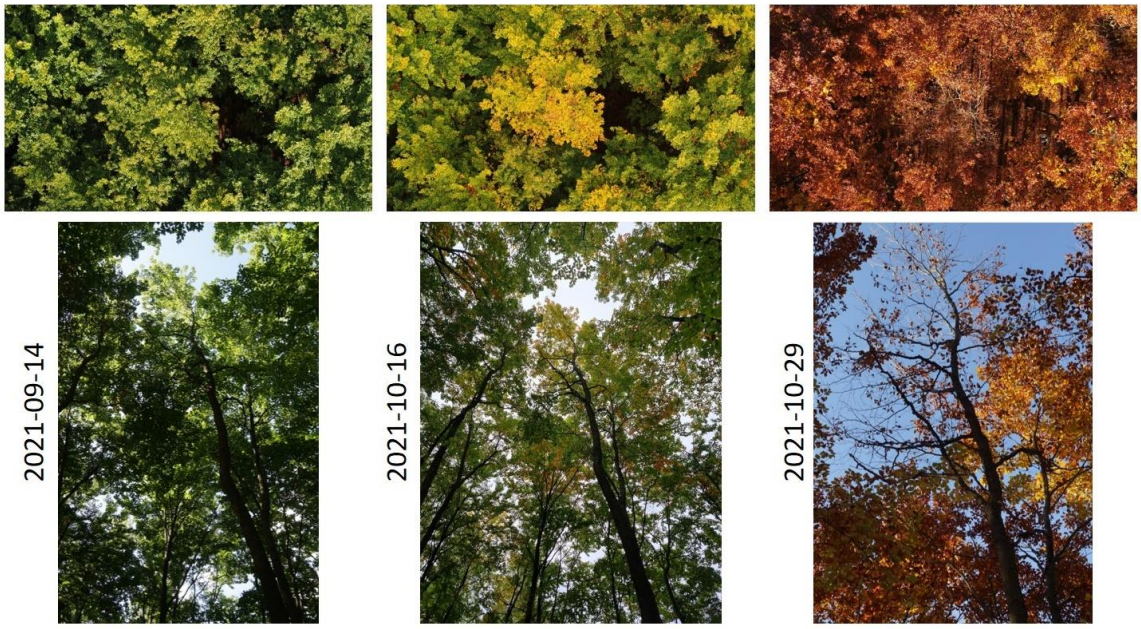


Figure 49: Example of three phenological observation epochs from the ground and simultaneous aerial view.

An evaluation of the scenarios described in the previous paragraph could be evaluated by an expert panel to assess the possibilities of only using aerial images for assessments. Here it could be of interest for an expert panel to carry out a survey among experienced ground technicians from the ground perspective and repeat the survey but using only aerial images. Important would be to assess how the assessments differ in comparison to the ground and aerial view but also the difference between observers. Furthermore, it could be assessed if a combination of both observations can enhance observations were both aerial and ground perspectives are considered

and averaged or given a specific weighting. To add, in terms of automatic ML methods, the expert panel would then also have to agree on a maximum error threshold from the resulting predictions.

With the example of phenology, one can see how an expert panel or project partners can be implemented to evaluate and arrive at decisions in how UAV-RS can be harmonized with intensive forest monitoring ground data. An important factor here is also the involvement of existing experienced ground technicians, as they are the ones who are actually practically carryong out the observations, and their input into the incorporation of UAVs could also improve field work ergonomics in the case that these “new tools” and methods are made available to them.

5.4.3 Machine Learning and Feature Engineering

The use of ML for the prediction of RS variables has become widely accepted (Maxwell et al., 2018) and is increasingly being used in forest research and management (Z. Liu et al., 2018). As opposed to traditional programming, ML models are trained or learn rather than programmed (Fergus & Chalmers, 2022) (see Figure 50) and are able to handle large datasets of high variability.

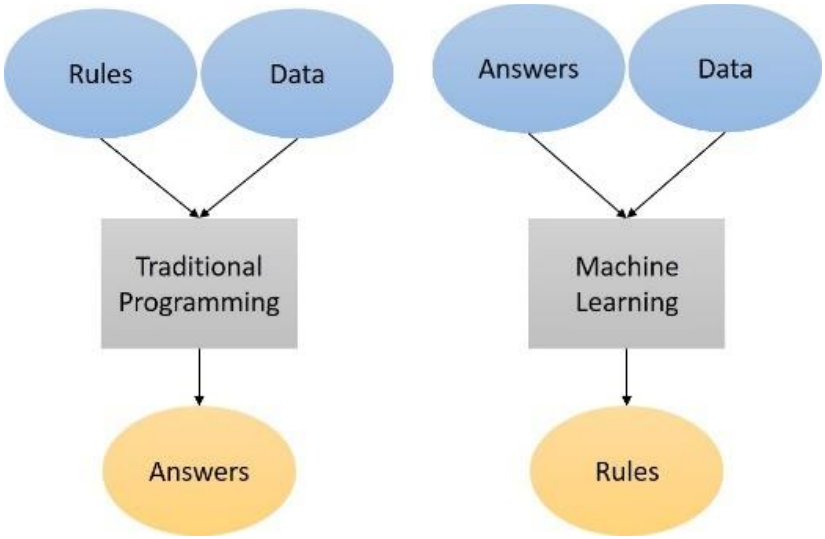


Figure 50: Traditional Programming compared to Machine Learning (adapted from Fergus & Chalmers, 2022)

In these unprecedented times of data acquisition, the term "Big Data" has in recent years been used to describe the accretion of massive amounts of multi-faceted data. The problem of dealing with "Big Data" was narrowed down to being only a problem of digital storage capabilities and we were told that ML, in particular DL will solve a lot of problems by turning raw data into useful information. Due to the hype of ML as well as terms such as DL and Artificial Intelligence (AI), many institutes and organizations are over ambitious to show that they are implementing this technology in their research and product development (Harrison & O'Neill, 2017). The reality of the fact is however that poorly acquired forest data as well as data which is not linked to a natural phenomenon with a specific level of standardisation could potentially be nothing more than unoccupied space on server or hard drive somewhere. Even with high quality data, the use of DL is still limited in terms of amount of data required to train DL algorithms. Figure 51 shows a theoretical representation of the accuracy curve of traditional ML as opposed to DL. With smaller datasets, ML tends to predict with a higher accuracy than DL, however when the amount of data increases as does the accuracy for DL models.

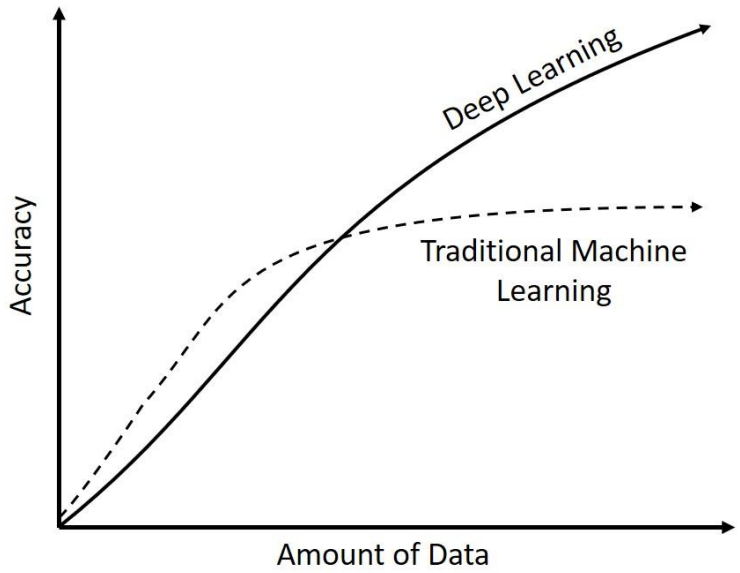


Figure 51: Graph showing the influence of amount of data for ML and DL algorithms (adapted from Aggarwal, 2018).

In reality, data managers must realize that without copious amounts of high-quality data and well-developed infrastructure, AI will remain a far-off dream. Important to understand at this point is that AI, DL, ML are sometimes used in the media interchangeably. Here, it should be noted that with for example a neural network (NN) it's origins can be either from ML and DL. When a NN is referred to as being "shallow" then this lies in the ML category. However, once a NN is built with more than two hidden layers then this can be regarded as "Deep" (Aggarwal, 2018). Shallow learning in effect can also refer to traditional ML algorithms which are known to be more "explainable" than the "Black Box" typical of DL applications. In terms of intensive forest monitoring, there is a definite need to understand natural processes and their causes and for this reason a specific emphasis on ML should be explored in the initial RS adaptation phase. DL and perhaps even AI applications can then be applied when a substantial standardized data infrastructure is in place. Figure 52 displays the "The Remote Sensing Hierarchy of Needs" which is adapted from Maslow (1943) and Rogati (2019). This graph shows how high-level ML, DL, and AI could eventually be incorporated into the adaptation of existing forest monitoring programs to RS applications. An import aspect of this development is the extensive testing with simple ML models to improve feature selection and engineering with an iterative approach. Once such a standardized data infrastructure and iterative testing pipeline is created, the deployment

of operational models is possible. A simplified version of this process was touched upon in Section 4.

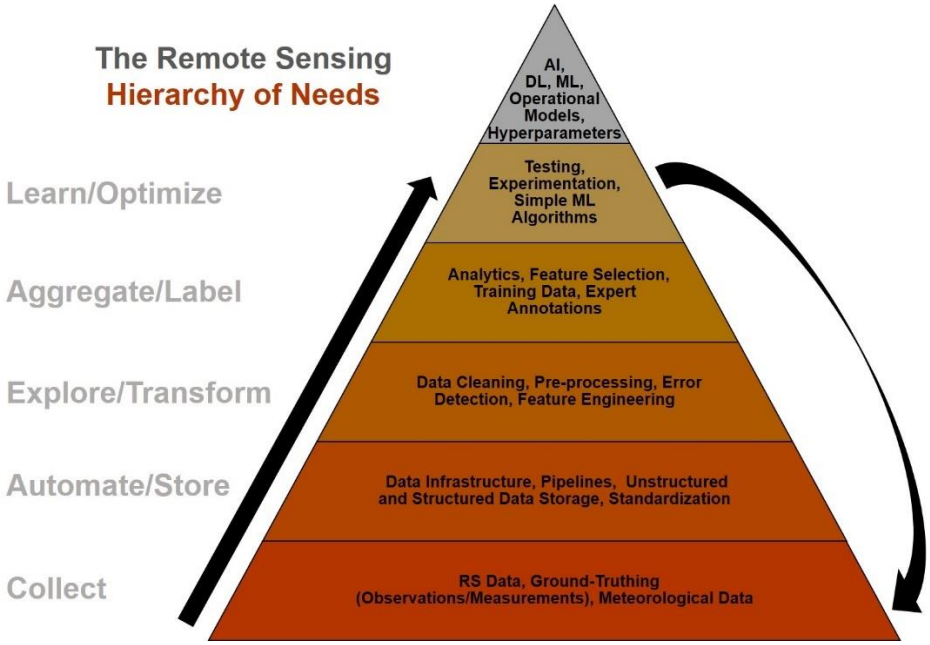


Figure 52: The Remote Sensing Hierarchy of Needs (adapted from Maslow, 1943; Rogati, 2019)

5.5 Outlook

The application of UAV-RS for intensive forest monitoring purposes will most probably become an integral part of monitoring programs in the near future. The challenges demonstrated in this research in terms of topics such as thermal imaging errors, uniform radiometric calibration and the translation of ground observations to RS data will however require further extensive exploration and experimentation. Due to the magnitude of the research required in terms of data acquisition methods, validation, annotations, feature engineering and predictive model development, a coordinated effort is an imperative amongst research institutions. Although largely a democratic process, where a healthy competitiveness in innovation among partner institutions is desired, the establishment of a governing body to maintain standardizations goals is essential.

References

Achim, A., Moreau, G., Coops, N., Axelson, J., Barrette, J., Bédard, S., Byrne, K., Caspersen, J., Dick, A., D'Orangeville, L., Drolet, G., Eskelson, B., Filipescu, C., Flamand-Hubert, M., Goodbody, T., Griess, V., Hagerman, S., Keys, K., Lafleur, B., & White, J. (2021). The changing culture of silviculture. *Forestry*. https://doi.org/10.1093/forestry/cpab047

Acorsi, M., Gimenez, L., & Martello, M. (2020). Assessing the Performance of a Low-Cost Thermal Camera in Proximal and Aerial Conditions. *Remote Sensing*, *12*, 3591. https://doi.org/10.3390/rs12213591

Adão, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., & Sousa, J. (2017). Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry. *Remote Sensing*, *9*(11), 1110. https://doi.org/10.3390/rs9111110

Aggarwal, C. C. (2018). *Neural networks and deep learning: A textbook*. Springer. https://doi.org/10.1007/978-3-319-94463-0

Ahrends, H. E., Brügger, R., Stöckli, R., Schenk, J., Michna, P., Jeanneret, F., Wanner, H., & Eugster, W. (2008). Quantitative phenological observations of a mixed beech forest in northern Switzerland with digital photography. *Journal of Geophysical Research: Biogeosciences*, *113*(G4). https://doi.org/10.1029/2007JG000650

Alberton, B., Torres, R. da S., Cancian, L. F., Borges, B. D., Almeida, J., Mariano, G. C., Santos, J. dos, & Morellato, L. P. C. (2017). Introducing digital cameras to monitor plant phenology in the tropics: Applications for conservation. *Perspectives in Ecology and Conservation*, *15*(2), 82–90. https://doi.org/10.1016/j.pecon.2017.06.004

Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D. D., Hogg, E. H. (Ted), Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J.-H., Allard, G., Running, S. W., Semerci, A., & Cobb, N. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*, *259*(4), 660–684. https://doi.org/10.1016/j.foreco.2009.09.001

Altum Integration Guide. (2020). MicaSense Knowledge Base. https://support.micasense.com/hc/en-us/articles/360010025413-Altum-Integration-Guide

Andersen, H.-E. R., Stephen E., & McGaughey, R. J. (2006). A rigorous assessment of tree height measurements obtained using airborne lidar and conventional field methods. *Canadian Journal of Remote Sensing*, *32*(5), 355–366.

Apostol, B., Lorent, A., Petrila, M., Gancz, V., & Badea, O. (2016). Height Extraction and Stand Volume Estimation Based on Fusion Airborne LiDAR Data and Terrestrial Measurements for a Norway Spruce [*Picea abies* (L.) Karst.] Test Site in Romania. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, *44*(1), 313. https://doi.org/10.15835/nbha44110155

Aragon, B., Johansen, K., Parkes, S., Malbeteau, Y., AlMashharawi, S., Al-Amoudi, T., Andrade, C. F., Turner, D., Lucieer, A., & McCabe, M. (2020). A Calibration Procedure for Field and UAV-Based Uncooled Thermal Infrared Instruments. *Sensors (Basel, Switzerland)*, *20*. https://doi.org/10.3390/s20113316

Arias-Rodil, M., Diéguez-Aranda, U., & Burkhart, H. E. (2017). Effects of Measurement Error in Total Tree Height and Upper-Stem Diameter on Stem Volume Prediction. *Forest Science*. https://doi.org/10.5849/forsci.2016-087

Atkins, J. W., Stovall, A. E. L., & Yang, X. (2020). Mapping Temperate Forest Phenology Using Tower, UAV, and Ground-Based Sensors. *Drones*, *4*(3), 56. https://doi.org/10.3390/drones4030056

Augspurger, C. K. (2009). Spring 2007 warmth and frost: Phenology, damage and refoliation in a temperate deciduous forest. *Functional Ecology*, *23*(6), 1031–1039. https://doi.org/10.1111/j.1365-2435.2009.01587.x

Awaya, Y., Tanaka, K., Kodani, E., & Nishizono, T. (2009). Responses of a beech (Fagus crenata Blume) stand to late spring frost damage in Morioka, Japan. *Forest Ecology and Management*, *257*(12), 2359–2369. https://doi.org/10.1016/j.foreco.2009.03.028

Azadeh Abdollahnejad, Dimitrios Panagiotidis, & Peter Surový. (2018). Estimation and Extrapolation of Tree Parameters Using Spectral Correlation between UAV and Pléiades Data. *Forests*, *9*(2), 85. https://doi.org/10.3390/f9020085

Baltsavias, E. P. (1999). A comparison between photogrammetry and laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *54*(2), 83–94.

Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., Waller, P., Choi, C., Riley, E., Thompson, T., Lascano, R. J., Li, H., & Moran, M. S. (2000). Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data. 16.

Baumgartner, A. (1952). Zur Phänologie von Laubhölzern und ihre Anwendung bei lokalklimatischen Untersuchungen. *Berichte Des DWD in Der US-Zone*, *42*, 69–73.

Belle, V., & Papantonis, I. (2021). Principles and Practice of Explainable Machine Learning. *Frontiers in Big Data*, *4*. https://www.frontiersin.org/articles/10.3389/fdata.2021.688969

Berger-Landefeldt, U. (1936, January 1). *Der Wasserhaushalt der Alpenpflanzen*. Schweizerbart'sche Verlagsbuchhandlung. https://www.schweizerbart.de/publications/detail/artno/144011500/Bibliotheca_Botanica_Heft_115

Berra, E. F., Gaulton, R., & Barr, S. (2019). Assessing spring phenology of a temperate woodland: A multiscale comparison of ground, unmanned aerial vehicle and Landsat satellite observations. *Remote Sensing of Environment*, *223*, 229–242. https://doi.org/10.1016/j.rse.2019.01.010

Boi, M. (2005). Efficiency of ultrasonic Vertex III hypsometer compared to the most commonly used hypsometers in Croatian forestry. *Croatian Journal of Forest Engineering*, 13.

Bolte, A., Ammer, C., Löf, M., Madsen, P., Nabuurs, G.-J., Schall, P., Spathelf, P., & Rock, J. (2009). Adaptive forest management in central Europe: Climate change impacts, strategies and integrative concept. *Scandinavian Journal of Forest Research*, *24*(6), 473–482. https://doi.org/10.1080/02827580903418224

Bragg, D. C. (2014). Accurately Measuring the Height of (Real) Forest Trees. *Journal of Forestry*, *112*(1), 51–54. https://doi.org/10.5849/jof.13-065

Brede, B., Bartholomeus, H., Barbier, N., Pimont, F., Vincent, G., & Herold, M. (2022). Peering through the thicket: Effects of UAV LiDAR scanner settings and flight planning on canopy volume discovery. *International Journal of Applied Earth Observation and Geoinformation*, *114*, 103056. https://doi.org/10.1016/j.jag.2022.103056

Brown, T. B., Hultine, K. R., Steltzer, H., Denny, E. G., Denslow, M. W., Granados, J., Henderson, S., Moore, D., Nagai, S., SanClements, M., Sánchez-Azofeifa, A., Sonnentag, O., Tazik, D., & Richardson, A. D. (2016). Using phenocams to monitor

our changing Earth: Toward a global phenocam network. *Frontiers in Ecology and the Environment*, *14*(2), 84–93. https://doi.org/10.1002/fee.1222

Brügger, R., & Vasella, A. (2018). *Pflanzen im Wandel der Jahreszeiten. Anleitung für phänologische Beobachtungen / Les plantes au cours des saisons. Guide pour observation phénologiques.* Geographica Bernensia. http://doi.org/10.4480/GB2018.N02

Brzostek, E., Dragoni, D., Schmid, H., Rahman, F., Sims, D., Wayson, C., Johnson, D., & Phillips, R. (2014). Chronic water stress reduces tree growth and the carbon sink of deciduous hardwood forests. *Global Change Biology*, *20*. https://doi.org/10.1111/gcb.12528

Burkhardt, J., & Pariyar, S. (2015). How does the VDP response of isohydric and anisohydric plants depend on leaf surface particles. *Plant Biology*, *18*. https://doi.org/10.1111/plb.12402

Butt, N., Slade, E., Thompson, J., Malhi, Y., & Riutta, T. (2013). Quantifying the sampling error in tree census measurements by volunteers and its effect on carbon stock estimates. *Ecological Applications*, *23*(4), 936–943. https://doi.org/10.1890/11-2059.1

Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, *40*(1), 16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024

Choat, B., Jansen, S., Brodribb, T. J., Cochard, H., Delzon, S., Bhaskar, R., Bucci, S. J., Feild, T. S., Gleason, S. M., Hacke, U. G., Jacobsen, A. L., Lens, F., Maherali, H., Martínez-Vilalta, J., Mayr, S., Mencuccini, M., Mitchell, P. J., Nardini, A., Pittermann, J., … Zanne, A. E. (2012). Global convergence in the vulnerability of forests to drought. *Nature*, *491*(7426), Article 7426. https://doi.org/10.1038/nature11688

Choodarathnakara, A. L., Kumar, D. T. A., Koliwad, D. S., & Patil, D. C. G. (2012). *Mixed Pixels: A Challenge in Remote Sensing Data Classification for Improving Performance*. *1*(9), 11.

Cleland, E., Chuine, I., Menzel, A., Mooney, H., & Schwartz, M. (2007). Shifting plant phenology in response to global change. *Trends in Ecology & Evolution*, *22*(7), 357–365. https://doi.org/10.1016/j.tree.2007.04.003

Clewley, D., Moghaddam, M., Lucas, R., & Bunting, P. (2013). Retrieval of forest structure and moisture from SAR data using an estimation algorithm. *2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS*, 153–156. https://doi.org/10.1109/IGARSS.2013.6721115

Cohen, Y., Alchanatis, V., Meron, M., Saranga, Y., & Tsipris, J. (2005). Estimation of leaf water potential by thermal imagery and spatial analysis. *Journal of Experimental Botany*, *56*(417), 1843–1852. https://doi.org/10.1093/jxb/eri174

Copernicus. (2022). *Sentinel Online—ESA - Sentinel Online*. https://sentinels.copernicus.eu/web/sentinel/home

Corey, D., Dunlap, W., & Burke, M. (1998). Averaging Correlations: Expected Values and Bias in Combined Pearson rs and Fisher's z Transformations. *Journal of General Psychology - J GEN PSYCHOL*, *125*, 245–261. https://doi.org/10.1080/00221309809595548

Corona, P., Chianucci, F., Quatrini, V., Civitarese, V., Clementel, F., Costa, C., Floris, A., Menesatti, P., Puletti, N., Sperandio, G., Verani, S., Turco, R., Bernardini, V., Plutino, M., & Scrinzi, G. (2017). Precision forestry: Concepts, tools and perspectives in Italy. *Forest@ - Rivista di Selvicoltura ed Ecologia Forestale*, *14*(1), 1–12. https://doi.org/10.3832/efor2285-014

Czernecki, B., Nowosad, J., & Jabłońska, K. (2018). Machine learning modeling of plant phenology based on coupling satellite and gridded meteorological dataset. *International Journal of Biometeorology*, *62*(7), 1297–1309. https://doi.org/10.1007/s00484-018-1534-2

Dai, W., Jin, H., Zhang, Y., Liu, T., & Zhou, Z. (2019). Detecting temporal changes in the temperature sensitivity of spring phenology with global warming: Application of machine learning in phenological model. *Agricultural and Forest Meteorology*, *279*, 107702. https://doi.org/10.1016/j.agrformet.2019.107702

Dai, Z., Edwards, G. E., & Ku, M. S. B. (1992). Control of Photosynthesis and Stomatal Conductance in Ricinus communis L. (Castor Bean) by Leaf to Air Vapor Pressure Deficit. *Plant Physiology*, *99*(4), 1426–1434.

Dash, J., Pearse, G., & Watt, M. (2018). UAV Multispectral Imagery Can Complement Satellite Data for Monitoring Forest Health. *Remote Sensing*, *10*(8), 1216. https://doi.org/10.3390/rs10081216

Dash, J., Pont, D., Brownlie, R., Dunningham, A., Watt, M., & Pearse, G. (2016). *Remote sensing for precision forestry*. *60*(4), 11.

Day, M. (2000). Influence of temperature and leaf-to-air vapor pressure deficit on net photosynthesis and stomatal conductance in red spruce (Picea rubens). *Tree Physiology*. https://doi.org/10.1093/TREEPHYS/20.1.57

Delaporte, A., Bazot, S., & Damesin, C. (2016). Reduced stem growth, but no reserve depletion or hydraulic impairment in beech suffering from long-term decline. *Trees*, *30*(1), 265–279. https://doi.org/10.1007/s00468-015-1299-8

Dempewolf, J., Nagol, J., Hein, S., Thiel, C., & Zimmermann, R. (2017). Measurement of Within-Season Tree Height Growth in a Mixed Forest Stand Using UAV Imagery. *Forests*, *8*(7), 231. https://doi.org/10.3390/f8070231

Dethier, B. E., Ashley, M. D., & Blair, B. (1972). *PHENOLOGY SATELLITE EXPERIMENT*. 9.

Diez, Y., Kentsch, S., Fukuda, M., López Caceres, M. L., Moritake, K., & Cabezas, M. (2021). Deep Learning in Forestry Using UAV-Acquired RGB Data: A Practical Review. *Remote Sensing*, *13*, 2837. https://doi.org/10.3390/rs13142837

Don, A., Hagen, C., Grüneberg, E., & Vos, C. (2019). Simulated wild boar bioturbation increases the stability of forest soil carbon. *Biogeosciences*, *16*(21), 4145–4155. https://doi.org/10.5194/bg-16-4145-2019

Drew, D. M., & Downes, G. M. (2009). The use of precision dendrometers in research on daily stem size and wood property variation: A review. *Dendrochronologia*, *27*(2), 159–172. https://doi.org/10.1016/j.dendro.2009.06.008

Drew, D., Richards, A., Downes, G., Cook, G., & Baker, P. (2011). The development of seasonal tree water deficit in Callitris intratropica. *Tree Physiology*, *31*, 953–964. https://doi.org/10.1093/treephys/tpr031

EASA. (2022). EASA. https://www.easa.europa.eu/domains/civil-drones

Ecomatik. (2022). *Startseite | Ecomatik GmbH*. https://ecomatik.de/

Eisenbeiß, H. (2009). *UAV photogrammetry*. ETH, Inst. für Geodäsie und Photogrammetrie.

European Commission. (2021). *Communication: New EU Forest Strategy for 2030* [Text]. European Commission. https://ec.europa.eu/info/files/communication-new-eu-forest-strategy-2030_en

Falkowski, M. J., Smith, A. M., Hudak, A. T., Gessler, P. E., Vierling, L. A., & Crookston, N. L. (2006). Automated estimation of individual conifer tree height and crown diameter

via two-dimensional spatial wavelet analysis of lidar data. *Canadian Journal of Remote Sensing*, *32*(2), 153–161.

Feller, U. (2016). Drought stress and carbon assimilation in a warming climate: Reversible and irreversible impacts. *Journal of Plant Physiology*, *203*, 84–94. https://doi.org/10.1016/j.jplph.2016.04.002

Fergus, P., & Chalmers, C. (2022). *Applied Deep Learning: Tools, Techniques, and Implementation*. Springer International Publishing. https://doi.org/10.1007/978-3-031-04420-5

Ferretti, M. (2021). New appetite for the monitoring of European forests. *Annals of Forest Science*, *78*. https://doi.org/10.1007/s13595-021-01112-w

Ferretti, M., & Fischer, R. (Eds.). (2013). *Forest monitoring: Methods for terrestrial investigations in Europe with an overview of North America and Asia* (First edition). Elsevier.

Fischer, R., Aas, W., Vries, W., Clarke, N., Cudlín, P., Leaver, D., Lundin, L., Matteucci, G., Matyssek, R., Mikkelsen, T., Mirtl, M., Oeztuerk, Y., Papale, D., Potočić, N., Simpson, D., Tuovinen, J.-P., Vesala, T., & Wieser, G. (2011). Towards a transnational system of supersites for forest monitoring and research in Europe—An overview on present state and future recommendations. *Iforest-Biogeosciences and Forestry*, *4*, 167–171. https://doi.org/10.3832/ifor0584-004

FLIR. (2018). *Radiometric Temperature Measurements: Surface characteristics and atmospheric compensation*. https://www.flir.com/globalassets/guidebooks/suas-radiometric-tech-note-en.pdf

Forstreuter, M. (2002). *Auswirkungen globaler Klimaänderungen auf das Wachstum und den Gaswechsel (CO2/H2O) von Rotbuchenbeständen (Fagus sylvatica L.)*. Techn. Univ.

Friedl, M., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 Global Land Cover: Algorithm Refinements and Characterization of new Datasets. *Remote Sensing of Environment*, *114*, 168–182. https://doi.org/10.1016/j.rse.2009.08.016

Furtak, A., & Nosalewicz, A. (2022). Leaf-to-air vapor pressure deficit differently affects barley depending on soil water availability. *South African Journal of Botany*, *146*, 497–502. https://doi.org/10.1016/j.sajb.2021.11.043

Ganguly, S., Friedl, M. A., Tan, B., Zhang, X., & Verma, M. (2010). Land surface phenology from MODIS: Characterization of the Collection 5 global land cover dynamics product. *Remote Sensing of Environment*, *114*(8), 1805–1816. https://doi.org/10.1016/j.rse.2010.04.005

Garcia-Forner, N., Biel, C., Savé, R., & Martínez-Vilalta, J. (2017). Isohydric species are not necessarily more carbon limited than anisohydric species during drought. *Tree Physiology*, *37*(4), 441–455. https://doi.org/10.1093/treephys/tpw109

Gatziolis, D., Fried, J. S., & Monleon, V. S. (2010). Challenges to estimating tree height via LiDAR in closed-canopy forests: A parable from western Oregon. *Forest Science*, *56*(2), 139–155.

Gerke, M. (2018). *Developments in UAV-Photogrammetry*. Wichmann Verlag. http://gispoint.de/fileadmin/user_upload/paper_gis_open/DLA_2018/537642028.pdf

Gerke, M., & Przybilla, H.-J. (2016). Accuracy Analysis of Photogrammetric UAV Image Blocks: Influence of Onboard RTK-GNSS and Cross Flight Patterns. *Photogrammetrie - Fernerkundung - Geoinformation*, *2016*(1), 17–30. https://doi.org/10.1127/pfg/2016/0284

Gillespie, A. R., Kahle, A. B., & Walker, R. E. (1987). Color enhancement of highly correlated images. II. Channel ratio and "chromaticity" transformation techniques. *Remote Sensing of Environment*, *22*(3), 343–365. https://doi.org/10.1016/0034-4257(87)90088-5

Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, *58*(3), 289–298. https://doi.org/10.1016/S0034-4257(96)00072-7

Gitelson, A., & Merzlyak, M. (1994a). Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *Journal of Photochemistry and Photobiology B: Biology*, *22*, 247–252. https://doi.org/10.1016/1011-1344(93)06963-4

Gitelson, A., & Merzlyak, M. N. (1994b). Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *Journal of Photochemistry and Photobiology B: Biology*, *22*(3), 247–252. https://doi.org/10.1016/1011-1344(93)06963-4

Gómez-Candón, D., Virlet, N., Labb, S., Jolivot, A., & Regnard, J.-L. (2016). Field phenotyping of water stress at tree scale by UAV-sensed imagery: New insights for thermal acquisition and calibration. *Precision Agriculture*, *17*(6), 786–800. https://doi.org/10.1007/s11119-016-9449-6

Grant, O. M., Chaves, M. M., & Jones, H. G. (2006). Optimizing thermal imaging as a technique for detecting stomatal closure induced by drought stress under greenhouse conditions. *Physiologia Plantarum*, *127*(3), 507–518. https://doi.org/10.1111/j.1399-3054.2006.00686.x

Grossiord, C., Buckley, T. N., Cernusak, L. A., Novick, K. A., Poulter, B., Siegwolf, R. T. W., Sperry, J. S., & McDowell, N. G. (2020). Plant responses to rising vapor pressure deficit. *New Phytologist*, *226*(6), 1550–1566. https://doi.org/10.1111/nph.16485

Guerra-Hernandez, J., Gonzalez-Ferreiro, E., Sarmento, A., Silva, J., Nunes, A., Correia, A. C., Fontes, L., Tomé, M., & Diaz-Varela, R. (2016). Short Communication. Using high resolution UAV imagery to estimate tree variables in Pinus pinea plantation in Portugal. *Forest Systems*, *25*(2), eSC09. https://doi.org/10.5424/fs/2016252-08895

Hakala, T., Markelin, L., Honkavaara, E., Scott, B., Theocharous, T., Nevalainen, O., Näsi, R., Suomalainen, J., Viljanen, N., Greenwell, C., & Fox, N. (2018). Direct Reflectance Measurements from Drones: Sensor Absolute Radiometric Calibration and System Tests for Forest Reflectance Characterization. *Sensors*, *18*(5), 1417. https://doi.org/10.3390/s18051417

Harrison, N., & O'Neill, D. (2017, June 7). If Your Company Isn't Good at Analytics, It's Not Ready for AI. *Harvard Business Review*. https://hbr.org/2017/06/if-your-company-isnt-good-at-analytics-its-not-ready-for-ai

Hartmann, H. (2011). Will a 385 million year-struggle for light become a struggle for water and for carbon? – How trees may cope with more frequent climate change-type drought events. *Global Change Biology*, *17*(1), 642–655. https://doi.org/10.1111/j.1365-2486.2010.02248.x

Hartmann, H., Link, R., & Schuldt, B. (2021). A whole-plant perspective of isohydry: Stem-level support for leaf-level plant water regulation. *Tree Physiology*. https://doi.org/10.1093/treephys/tpab011

Hijmans, R. J., Bivand, R., Forner, K., Ooms, J., Pebesma, E., & Sumner, M. D. (2022). *terra: Spatial Data Analysis* (1.5-34) [Computer software]. https://CRAN.R-project.org/package=terra

Holopainen, M., Vastaranta, M., & Hyyppä, J. (2014). Outlook for the Next Generation's Precision Forestry in Finland. *Forests*, *5*(7), 1682–1694. https://doi.org/10.3390/f5071682

Holst, G. C. (2000). *Common Sense Approach to Thermal Imaging*. SPIE. https://doi.org/10.1117/3.2588945

Holzwarth, S., Thonfeld, F., Abdullahi, S., Asam, S., Da Ponte Canova, E., Gessner, U., Huth, J., Kraus, T., Leutner, B., & Kuenzer, C. (2020). Earth Observation Based Monitoring of Forests in Germany: A Review. *Remote Sensing*, *12*(21), 3570. https://doi.org/10.3390/rs12213570

Honkavaara, E., Hakala, T., Markelin, L., Rosnell, T., Saari, H., & Mäkynen, J. (2012). A Process for Radiometric Correction of UAV Image Blocks. *Photogrammetrie - Fernerkundung - Geoinformation*, *2012*(2), 115–127. https://doi.org/10.1127/1432-8364/2012/0106

Honkavaara, E., Näsi, R., Alves de Oliveira, R., Viljanen, N., Suomalainen, J., Khoramshahi, E., Hakala, T., Nevalainen, O., Markelin, L., Vuorinen, M., Kankaanhuhta, V., Paivi, L.-S., & Haataja, L. (2020). Using Multitemperaol Hyper- and Multispectral UAV Imaging for Detecting Bark Beetle Infestation on Norway Spruce. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLIII-B3-2020*, 429–434. https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-429-2020

Huanmin, L., Erxue, C., Jian, C., & Xiaowen, L. (2010). *Forest height estimation methods using polarimetric SAR interferometry*. 17.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*(1–2), 195–213.

Hunt, E. R., Doraiswamy, P. C., McMurtrey, J. E., Daughtry, C. S. T., Perry, E. M., & Akhmedov, B. (2013). A visible band index for remote sensing leaf chlorophyll content at the canopy scale. *International Journal of Applied Earth Observation and Geoinformation*, *21*, 103–112. https://doi.org/10.1016/j.jag.2012.07.020

Hyyppä, J., Mielonen, T., Hyyppä, H., Maltamo, M., Yu, X., Honkavaara, E., & Kaartinen, H. (2005a). Using individual tree crown approach for forest volume extraction with aerial images and laser point clouds. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *36*(3), W19.

Hyyppä, J., Mielonen, T., Hyyppä, H., Maltamo, M., Yu, X., Honkavaara, E., & Kaartinen, H. (2005b). Using individual tree crown approach for forest volume extraction with aerial images and laser point clouds. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *36*(3), W19.

ICP-Forests. (2022). *Level II*. http://icp-forests.net/page/level-ii

Imai, Y., Setojima, M., Yamagishi, Y., & Fujiwara, N. (2004). Tree-height measuring characteristics of urban forests by LiDAR data different in resolution. . . *In Proceedings of the International Society of Photogrammetry and Remote Sensing Conference*, *4*. http://www.isprs.org/proceedings/xxxv/congress/comm7/papers/100.pdf

IPCC. (2018). *Technical Summary—Special Report on Climate Change and Land*. https://www.ipcc.ch/srccl/chapter/technical-summary/

Iwanoczko, A. (2017). *Here are the top 5 challenges working with Hyperspectral Data collected using a UAV platform*. https://www.linkedin.com/pulse/here-top-5-challenges-working-hyperspectral-data-using-iwanoczko

Jones, H. G., & Vaughan, R. A. (2010). *Remote Sensing of Vegetation: Principles, Techniques, and Applications*. Oxford University Press.

JoshOBrien. (2021). *ExifTool functionality from R* [R]. https://github.com/JoshOBrien/exiftoolr (Original work published 2018)

Jucker, T., Caspersen, J., Chave, J., Antin, C., Barbier, N., Bongers, F., Dalponte, M., van Ewijk, K. Y., Forrester, D. I., Haeni, M., Higgins, S. I., Holdaway, R. J., Iida, Y., Lorimer, C., Marshall, P. L., Momo, S., Moncrieff, G. R., Ploton, P., Poorter, L., … Coomes, D. A. (2017). Allometric equations for integrating remote sensing imagery into forest monitoring programmes. *Global Change Biology*, *23*(1), 177–190. https://doi.org/10.1111/gcb.13388

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B. M., & Wu, J.-C. (2012). An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. *Remote Sensing*, *4*(4), 950–974. https://doi.org/10.3390/rs4040950

Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *173*, 24–49. https://doi.org/10.1016/j.isprsjprs.2020.12.010

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. *Surveys in Geophysics*, *40*(4), 959–977. https://doi.org/10.1007/s10712-019-09529-9

Kelly, J., Kljun, N., Olsson, P.-O., Mihai, L., Liljeblad, B., Weslien, P., Klemedtsson, L., & Eklundh, L. (2019). Challenges and Best Practices for Deriving Temperature Data from an Uncalibrated UAV Thermal Infrared Camera. *Remote Sensing*, *11*. https://doi.org/10.3390/rs11050567

Kershaw, J. A., Ducey, M. J., Beers, T. W., & Husch, B. (2016). *Forest mensuration* (Fifth edition). John Wiley & Sons.

Klein, T. (2014). The variability of stomatal sensitivity to leaf water potential across tree species indicates a continuum between isohydric and anisohydric behaviours. *Functional Ecology*, *28*(6), 1313–1320. https://doi.org/10.1111/1365-2435.12289

Klosterman, S., Melaas, E., Wang, J. A., Martinez, A., Frederick, S., O'Keefe, J., Orwig, D. A., Wang, Z., Sun, Q., Schaaf, C., Friedl, M., & Richardson, A. D. (2018). Fine-scale perspectives on landscape phenology from unmanned aerial vehicle (UAV) photography. *Agricultural and Forest Meteorology*, *248*, 397–407. https://doi.org/10.1016/j.agrformet.2017.10.015

Klosterman, S. T., Hufkens, K., Gray, J. M., Melaas, E., Sonnentag, O., Lavine, I., Mitchell, L., Norman, R., Friedl, M. A., & Richardson, A. D. (2014). Evaluating remote sensing of deciduous forest phenology at multiple spatial scales using PhenoCam imagery. *Biogeosciences*, *11*(16), 4305–4320. https://doi.org/10.5194/bg-11-4305-2014

Kowalski, K., Senf, C., Hostert, P., & Pflugmacher, D. (2020). Characterizing spring phenology of temperate broadleaf forests using Landsat and Sentinel-2 time series. *International Journal of Applied Earth Observation and Geoinformation*, *92*, 102172. https://doi.org/10.1016/j.jag.2020.102172

Kraft, T., Geßner, M., Meißner, H., Cramer, M., Gerke, M., & Przybilla, H. J. (2016). EVALUATION OF A METRIC CAMERA SYSTEM TAILORED FOR HIGH PRECISION UAV APPLICATIONS. *ISPRS - International Archives of the*

Photogrammetry, Remote Sensing and Spatial Information Sciences, *XLI-B1*, 901–907. https://doi.org/10.5194/isprsarchives-XLI-B1-901-2016

Krause, S., Hartmann, F., & Mund, J.-P. (2016). UAV Workflow Optimization for the Acquisition of High-Quality Photogrammetric Point Clouds in Forestry. *GI_Forum*, *1*, 72–84. https://doi.org/10.1553/giscience2016_01_s72

Krause, S., & Sanders, T. G. M. (2022a). Mapping Tree Water Deficit with UAV Thermal Imaging and Meteorological Data (Manuscript submitted for publication). *Remote Sensing in Earth Systems Sciences*.

Krause, S., & Sanders, T. G. M. (2022b). The Detection of European Beech Spring Phenological Phases with UAV-derived Multispectral Indicies and Machine Learning Regression. (Manuscript submitted for publication). *Annals of Forest Science*.

Krause, S., Sanders, T. G. M., Mund, J.-P., & Greve, K. (2019). UAV-Based Photogrammetric Tree Height Measurement for Intensive Forest Monitoring. *Remote Sensing*, *11*(7), 758-. https://doi.org/10.3390/rs11070758

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of Statistical Software*, *28*(5). https://doi.org/10.18637/jss.v028.i05

Kuhn, M., & Johnson, K. (2019). *Feature Engineering and Selection: A Practical Approach for Predictive Models*. CRC Press.

Kuhn, M., Wing, J., Weston, S., & Williams, A. (2022). The caret package. *Gene Expr*.

Kutsch, W. L., Wirth, C., Kattge, J., Nöllert, S., Herbst, M., & Kappen, L. (2009). Ecophysiological Characteristics of Mature Trees and Stands—Consequences for Old-Growth Forest Productivity. In *Old-Growth Forests, 57-79 (2009)* (Vol. 207, pp. 57–79). https://doi.org/10.1007/978-3-540-92706-8_4

Laar, A. van, & Akça, A. (2007). *Forest mensuration*. Springer.

Landsat. (2022). *Landsat Science*. https://landsat.gsfc.nasa.gov/

Lapenis, A., Shvidenko, A., Shepaschenko, D., Nilsson, S., & Aiyyer, A. (2005). Acclimation of Russian forests to recent changes in climate. *Global Change Biology*, *11*(12), 2090–2102. https://doi.org/10.1111/j.1365-2486.2005.001069.x

Larjavaara, M., & Muller-Landau, H. C. (2013). Measuring tree height: A quantitative comparison of two common field methods in a moist tropical forest. *Methods in Ecology and Evolution*, *4*(9), 793–801. https://doi.org/10.1111/2041-210X.12071

Lary, D. J., Alavi, A. H., Gandomi, A. H., & Walker, A. L. (2016). Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, *7*(1), 3–10. https://doi.org/10.1016/j.gsf.2015.07.003

Lausch, A., Erasmi, S., King, D., Magdon, P., & Heurich, M. (2016). Understanding Forest Health with Remote Sensing -Part I—A Review of Spectral Traits, Processes and Remote-Sensing Characteristics. *Remote Sensing*, *8*(12), 1029. https://doi.org/10.3390/rs8121029

Lepton LWIR Micro Thermal Camera Module | Teledyne FLIR. (n.d.). Retrieved 2 July 2022, from https://www.flir.eu/products/lepton?vertical=lwir&segment=oem

Leuschner, C. (2020). Drought response of European beech (Fagus sylvatica L.)—A review. *Perspectives in Plant Ecology, Evolution and Systematics*, *47*, 125576. https://doi.org/10.1016/j.ppees.2020.125576

Li, H., Zhang, H., Chen, Z., Yang, M., & Zhang, Y. (2015). A Method Suitable for Vicarious Calibration of a UAV Hyperspectral Remote Sensor. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*, 1–15. https://doi.org/10.1109/JSTARS.2015.2416213

Li, N., Zhan, P., Pan, Y., Zhu, X., Li, M., & Zhang, D. (2020). Comparison of Remote Sensing Time-Series Smoothing Methods for Grassland Spring Phenology Extraction on the Qinghai–Tibetan Plateau. *Remote Sensing, 12*(20), 3383. https://doi.org/10.3390/rs12203383

Liang, S., & Wang, J. (2020). *Advanced remote sensing: Terrestrial information extraction and applications*.

Liang, X., Kukko, A., Balenović, I., Ninni, S., Junttila, S., Kankare, V., Holopainen, M., Mokros, M., Surovy, P., Kaartinen, H., Jurjevic, L., Honkavaara, E., Nasi, R., Jingbin, L., Hollaus, M., Tian, J., Yu, X., Jie, P., Shangshu, C., & Hyyppä, J. (2022). Close-Range Remote Sensing of Forests: The State of the Art, Challenges, and Opportunities for Systems and Data Acquisitions. *IEEE Geoscience and Remote Sensing Magazine, 10*, 2–41. https://doi.org/10.1109/MGRS.2022.3168135

Liang, Z., Hanfeng, L., Dingjie, W., Yanqing, H., & Jie, W. (2015). Asynchronous RTK precise DGNSS positioning method for deriving a low-latency high-rate output. *Journal of Geodesy*, *89*(7), 641–653. https://doi.org/10.1007/s00190-015-0803-7

Lieth, H. (Ed.). (1974). *Phenology and Seasonality Modeling* (Vol. 8). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-51863-8

Lillesand, T., Kiefer, R. W., & Chipman, J. (2015). *Remote sensing and image interpretation*. John Wiley & Sons.

Linderholm, H. W. (2006). Growing season changes in the last century. *Agricultural and Forest Meteorology*, *137*(1–2), 1–14. https://doi.org/10.1016/j.agrformet.2006.03.006

Linnaeus, C. (1751). *Philosophia botanica: In qua explicantur fundamenta botanica cum definitionibus partium, exemplis terminorum, observationibus rariorum, adjectis figuris aeneis*. apud Godofr. Kiesewetter.

Liu, G., Wang, J., Dong, P., Chen, Y., & Liu, Z. (2018). Estimating Individual Tree Height and Diameter at Breast Height (DBH) from Terrestrial Laser Scanning (TLS) Data at Plot Level. *Forests*, *9*(7), 398. https://doi.org/10.3390/f9070398

Liu, Z., Peng, C., Work, T., Candau, J.-N., DesRochers, A., & Kneeshaw, D. (2018). Application of machine-learning methods in forest ecology: Recent progress and future challenges. *Environmental Reviews*, *26*(4), 339–350. https://doi.org/10.1139/er-2018-0034

Lorenz, M. (1995). International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests-ICP Forests-. *Water, Air, & Soil Pollution*, *85*(3), 1221–1226. https://doi.org/10.1007/BF00477148

Malaisse, F. (1964). CONTRIBUTION A L'ÉTUDE DES HÊTRAIES D'EUROPE OCCIDENTALE: Note 4: Quelques observations phénologiques de hêtraies en 1963. *Bulletin de La Société Royale de Botanique de Belgique / Bulletin van de Koninklijke Belgische Botanische Vereniging, 97*, 85–97. JSTOR.

Maltamo, M., Næsset, E., & Vauhkonen, J. (Eds.). (2014). *Forestry Applications of Airborne Laser Scanning* (Vol. 27). Springer Netherlands. http://link.springer.com/10.1007/978-94-017-8663-8

Martinez del Castillo, E., Zang, C. S., Buras, A., Hacket-Pain, A., Esper, J., Serrano-Notivoli, R., Hartl, C., Weigel, R., Klesse, S., Resco de Dios, V., Scharnweber, T., Dorado-Liñán, I., van der Maaten-Theunissen, M., van der Maaten, E., Jump, A., Mikac, S., Banzragch, B.-E., Beck, W., Cavin, L., … de Luis, M. (2022). Climate-change-driven growth decline of European beech forests. *Communications Biology*, *5*(1), Article 1. https://doi.org/10.1038/s42003-022-03107-3

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*, 370–396. https://doi.org/10.1037/h0054346

Massmann, A., Gentine, P., & Lin, C. (2019). When Does Vapor Pressure Deficit Drive or Reduce Evapotranspiration? *Journal of Advances in Modeling Earth Systems*, *11*. https://doi.org/10.1029/2019MS001790

Matasci, G., Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., Hobart, G. W., & Zald, H. S. J. (2018). Large-area mapping of Canadian boreal forest cover, height, biomass and other structural attributes using Landsat composites and lidar plots. *Remote Sensing of Environment*, *209*, 90–106. https://doi.org/10.1016/j.rse.2017.12.020

Mather, P. M., & Koch, M. (2011). *Computer processing of remotely-sensed images: An introduction* (4th ed). Wiley-Blackwell.

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, *39*(9), 2784–2817. https://doi.org/10.1080/01431161.2018.1433343

McCartney, M., Häringer, M., & Polifke, W. (2019). Comparison of Machine Learning Techniques in the Interpolation and Extrapolation of Flame Describing Functions. *Journal of Engineering for Gas Turbines and Power*, *142*. https://doi.org/10.1115/1.4045516

McClave, J. T., & Sincich, T. T. (2018). *Statistics, Global Edition* (13. Edition). Pearson Education Limited.

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, *17*(7), 1425–1432. https://doi.org/10.1080/01431169608948714

McGaughey, R. J. (2018). *FUSION/LDV: Software for LIDAR data analysis and visualization* (V3.80) [Computer software]. US Department of Agriculture, Forest Service. http://forsys.cfr.washington.edu/fusion/FUSION_manual.pdf

McGlone, J. C., & Lee, G. Y. G. (2013). *Manual of photogrammetry* (6th ed.). American Society for Photogrammetry and Remote Sensing. http://www.worldcat.org/oclc/856568615

Meinzer, F. C., Bond, B. J., Warren, J. M., & Woodruff, D. R. (2005). Does water transport scale universally with tree size? *Functional Ecology*, *19*(4), 558–565. https://doi.org/10.1111/j.1365-2435.2005.01017.x

Menzel, A. (1997). *Phänologie von Waldbäumen unter sich ändernden Klimabedingungen: Auswertung der Beobachtungen in den internationalen phänologischen Gärten und Möglichkeiten der Modellierung von Phänodaten*. Frank.

Menzel, A. (2002). Phenology: Its Importance to the Global Change Community. *Climatic Change*, *54*(4), 379–385. https://doi.org/10.1023/A:1016125215496

Menzel, A., Helm, R., & Zang, C. (2015). Patterns of late spring frost leaf damage and recovery in a European beech (Fagus sylvatica L.) stand in south-eastern Germany based on repeated digital photographs. *Frontiers in Plant Science*, *6*. https://doi.org/10.3389/fpls.2015.00110

Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aasa, A., Ahas, R., Alm-Kübler, K., Bissolli, P., Braslavská, O., Briede, A., Chmielewski, F. M., Crepinsek, Z., Curnel, Y., Dahl, Å., Defila, C., Donnelly, A., Filella, Y., Jatczak, K., Måge, F., … Zust, A. (2006). European phenological response to climate change matches the warming pattern. *Global Change Biology*, *12*(10), 1969–1976. https://doi.org/10.1111/j.1365-2486.2006.01193.x

micasense.com. (2022). *Micasense.com*. https://micasense.com/

MicaSense—Drone sensors for vegetation mapping. (n.d.). Micasense. Retrieved 2 July 2022, from https://micasense.com/

Mielcarek, M., Stereńczak, K., & Khosravipour, A. (2018). Testing and evaluating different LiDAR-derived canopy height model generation methods for tree height estimation. *International Journal of Applied Earth Observation and Geoinformation*, *71*, 132–143. https://doi.org/10.1016/j.jag.2018.05.002

Minkina, W., & Dudzik, S. (2009). *Infrared Thermography: Errors and Uncertainties*. John Wiley & Sons.

Mocko, K., Nicotra, A., & Jones, C. (2017). Extent of Solar Tracking Differs between Two Co-occurring Congeneric Geophytes That Differ in Leaf Shape. *International Journal of Plant Sciences*, *179*, 000–000. https://doi.org/10.1086/695698

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*(9), 340. https://doi.org/10.3390/f8090340

Moran, E., Lauder, J., Musser, C., Stathos, A., & Shu, M. (2017). The genetics of drought tolerance in conifers. *New Phytologist*, *216*. https://doi.org/10.1111/nph.14774

Moran, M. S., Bryant, R. B., Clarke, T. R., & Qi, J. (2001). Deployment and calibration of reference reflectance tarps for use with airborne imaging sensors. *Photogrammetric Engineering and Remote Sensing*, *67*(3), 273–286.

Moskal, L. M., Erdody, T., Kato, A., Richardson, J., Zheng, G., & Briggs, D. (2018). *Lidar Applications in Precision Forestry*. 11.

Müller, J., Bolte, A., & others. (2009). The use of lysimeters in forest hydrology research in north-east Germany. *Landbauforschung (VTI Agric. For. Res.)*, *59*, 1–10.

MüllerA, J. (2010). Forest hydrology research with lysimeter in the northeast German lowlands special methods and results for the forest management. *Soil Solutions for a Changing World: Proceedings of the 19th World Congress of Soil Science, Edited by RJ Gilkes and N. Prakongkep*, 28–31. https://www.researchgate.net/profile/Andreas_Bolte/publication/265980128_Forest_hydrology_research_with_lysimeter_in_the_northeast_German_lowlands_special_methods_and_results_for_the_forest_management/links/54d2097d0cf28370d0e199cb.pdf

Næsset, E., Gobakken, T., Holmgren, J., Hyyppä, H., Hyyppä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, Å., & Söderman, U. (2004). Laser scanning of forest resources: The nordic experience. *Scandinavian Journal of Forest Research*, *19*(6), 482–499. https://doi.org/10.1080/02827580410019553

Nathalie, B., Huc, R., Granier, A., & Dreyer, E. (2006). Temperate forest trees and stands under severe drought: A review of ecophysiological responses, adaptation processes and long-term consequences. *Http://Dx.Doi.Org/10.1051/Forest:2006042*, *63*. https://doi.org/10.1051/forest:2006042

Nobel, P. (2020). *Physicochemical and environmental plant physiology*. Elsevier.

Noh, H., & Lee, J. (2022). The Effect of Vapor Pressure Deficit Regulation on the Growth of Tomato Plants Grown in Different Planting Environments. *Applied Sciences*, *12*(7), 3667. https://doi.org/10.3390/app12073667

Panagiotidis, D., Abdollahnejad, A., Surový, P., & Chiteculo, V. (2016). Determining tree height and crown diameter from high-resolution UAV imagery. *International Journal of Remote Sensing*, 1–19. https://doi.org/10.1080/01431161.2016.1264028

Pandey, E. P. C., & Arellano, P. (2023). *Advances in Remote Sensing for Forest Monitoring*. 400.

Park, J. Y., Muller-Landau, H. C., Lichstein, J. W., Rifai, S. W., Dandois, J. P., & Bohlman, S. A. (2019). Quantifying Leaf Phenology of Individual Trees and Species in a Tropical Forest Using Unmanned Aerial Vehicle (UAV) Images. *Remote Sensing*, *11*(13), 1534. https://doi.org/10.3390/rs11131534

Perich, G., Hund, A., Anderegg, J., Roth, L., Boer, M., Walter, A., Liebisch, F., & Aasen, H. (2020). Assessment of Multi-Image Unmanned Aerial Vehicle Based High-Throughput Field Phenotyping of Canopy Temperature. *Frontiers in Plant Science*, *11*, 150. https://doi.org/10.3389/fpls.2020.00150

Persson, H. J., Ekström, M., & Ståhl, G. (2022). Quantify and account for field reference errors in forest remote sensing studies. *Remote Sensing of Environment*, *283*, 113302. https://doi.org/10.1016/j.rse.2022.113302

Picard, N., Saint-André, L., & Henry, M. (2012). *Manual for building tree volume and biomass allometric equations from filed measurement to prediction*. Food and Agriculture Organization of the United Nations (FA0) [u.a. http://www.fao.org/docrep/018/i3058e/i3058e.pdf

Playà-Montmany, N., & Tattersall, G. J. (2021). Spot size, distance and emissivity errors in field applications of infrared thermography. *Methods in Ecology and Evolution*, *12*(5), 828–840. https://doi.org/10.1111/2041-210X.13563

Pohl, C., & Loong, C. K. (2015). In-situ data collection for oil palm tree height determination using synthetic aperture radar. *IOP Conference Series: Earth and Environmental Science*, *34*, 012027. https://doi.org/10.1088/1755-1315/34/1/012027

Popescu, S. C., Wynne, R. H., & Nelson, R. F. (2002). Estimating plot-level tree heights with lidar: Local filtering with a canopy-height based variable window size. *Computers and Electronics in Agriculture*, *37*(1–3), 71–95. https://doi.org/10.1016/S0168-1699(02)00121-7

Pretzsch, H., Grams, T., Häberle, K. H., Pritsch, K., Bauerle, T., & Rötzer, T. (2020). Growth and mortality of Norway spruce and European beech in monospecific and mixed-species stands under natural episodic and experimentally extended drought. Results of the KROOF throughfall exclusion experiment. *Trees*, *34*(4), 957–970. https://doi.org/10.1007/s00468-020-01973-0

Proctor, C., & He, Y. (2015). WORKFLOW FOR BUILDING A HYPERSPECTRAL UAV: CHALLENGES AND OPPORTUNITIES. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XL-1/W4*, 415–419. https://doi.org/10.5194/isprsarchives-XL-1-W4-415-2015

Przybilla, H.-J., & Baeumker, M. (2020). *RTK and PPK: GNSS-Technologies for direct georeferencing of UAV image flights*.

Puliti, S., Saarela, S., Gobakken, T., Ståhl, G., & Næsset, E. (2018). Combining UAV and Sentinel-2 auxiliary data for forest growing stock volume estimation through hierarchical model-based inference. *Remote Sensing of Environment*, *204*, 485–497. https://doi.org/10.1016/j.rse.2017.10.007

Pulliainen, J., Hari, P., Hallikainen, M., Patrikainen, N., Peramaki, M., & Kolari, P. (2004). Monitoring of soil moisture and vegetation water content variations in boreal forest from C-band SAR data. *IEEE International IEEE International IEEE International Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004*, *2*, 1013–1016. https://doi.org/10.1109/IGARSS.2004.1368582

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Raspe, S., Fleck, S., Beuker, E., Bastrup-Birk, A., & Preuhsler, T. (2020). *Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests.* Thünen Institute of Forest Ecosystems, Eberswalde, Germany. http://www.icp-forests.org/Manual.htm

Raspe, S., Schulz, C., & Kroll, F. (2004). Wenn schon im Sommer tonnenweise Blätter fallen. Baumentwicklung und Streufall. *LWF Aktuell*, *43*, 11–13.

Reich, P. B., Sendall, K. M., Stefanski, A., Wei, X., Rich, R. L., & Montgomery, R. A. (2016). Boreal and temperate trees show strong acclimation of respiration to warming. *Nature*, *531*(7596), 633–636. https://doi.org/10.1038/nature17142

Ribeiro-Gomes, K., Hernández-López, D., Ortega, J. F., Ballesteros, R., Poblete, T., & Moreno, M. A. (2017). Uncooled Thermal Camera Calibration and Optimization of the Photogrammetry Process for UAV Applications in Agriculture. *Sensors*, *17*(10), Article 10. https://doi.org/10.3390/s17102173

Riedel, T., Demnant, B., Hennig, P., Neubauer, M., Polley, H., & Schwitzgebel, F. (2016). *Aufnahmeanweisung für die Kohlenstoffinventur 2017 (Version 1.2): Vol. 1. Auflage*. Bundesministerium für Ernährung und Landwirtschaft (BMEL).

Riek, W. (2004). *Eigenschaften typischer Waldböden im Nordostdeutschen Tiefland unter besonderer Berücksichtigung des Landes Brandenburg: Vol. Bd. 19*. Ministerium für Landwirtschaft, Umweltschutz und Raumordnung des Landes Brandenburg, Presse- und Öffentlichkeitsarbeit [u.a.]. http://www.worldcat.org/oclc/64657313

Rodrigues, A., Marcal, A. R. S., & Cunha, M. (2012). Phenology parameter extraction from time-series of satellite vegetation index data using phenosat. *2012 IEEE International Geoscience and Remote Sensing Symposium*, 4926–4929. https://doi.org/10.1109/IGARSS.2012.6352507

Rogati, M. (2019, May 22). *The AI Hierarchy of Needs*. Medium. https://medium.com/hackernoon/the-ai-hierarchy-of-needs-18f111fcc007

Rogers, E. J. (1947). *Estimating tree heights from shadows on vertical aerial photographs*. https://www.nrs.fs.fed.us/pubs/3142

Roloff, A. (1988). Morphologie der Kronenentwicklung von Fagus sylvatica L. (Rotbuche) unter besonderer Berücksichtigung neuartiger Veränderungen: II. Strategie der Luftraumeroberung und Veränderungen durch Umwelteinflüsse. *Flora*, *180*(3), 297–338. https://doi.org/10.1016/S0367-2530(17)30325-0

Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., & Harlan, J. C. (1974). Monitoring the vernal advancements and retrogradation. *Texas, Texas A & M University*.

Roussel, J.-R., & Auty, D. (2018). *LidR: Airborne LiDAR Data Manipulation and Visualization for Forestry Applications* (R package version 1.4.1) [Computer software]. https://CRAN.R-project.org/package=lidR

Rubio-Cuadrado, Á., Camarero, J. J., Rodríguez-Calcerrada, J., Perea, R., Gómez, C., Montes, F., & Gil, L. (2021). Impact of successive spring frosts on leaf phenology and radial growth in three deciduous tree species with contrasting climate requirements in central Spain. *Tree Physiology*, *41*(12), 2279–2292. https://doi.org/10.1093/treephys/tpab076

Sachsenforst. (2020, June 25). *Frost trifft Forst*. https://medienservice.sachsen.de/medien/news/238026

Sade, N., Gebremedhin, A., & Moshelion, M. (2012). Risk-taking plants. *Plant Signaling & Behavior*, *7*(7), 767–770. https://doi.org/10.4161/psb.20505

Sakai, A., & Larcher, W. (1987). *Frost Survival of Plants* (Vol. 62). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-71745-1

Sanders, T. G. M., Krüger, I., & Holzhausen, M. (2020). *Das intensive forstliche Monitoring—Level II*. Johann Heinrich von Thünen-Institut. https://doi.org/10.3220/PB1608106763000

Schuldt, B., Buras, A., Arend, M., Vitasse, Y., Beierkuhnlein, C., Damm, A., Gharun, M., Grams, T. E. E., Hauck, M., Hajek, P., Hartmann, H., Hiltbrunner, E., Hoch, G., Holloway-Phillips, M., Körner, C., Larysch, E., Lübbe, T., Nelson, D. B., Rammig, A., … Kahmen, A. (2020). A first assessment of the impact of the extreme 2018 summer drought on Central European forests. *Basic and Applied Ecology*, *45*, 86–103. https://doi.org/10.1016/j.baae.2020.04.003

Schüler, S. (2012). *Genetische Variation und Plastizität des Blattaustriebs von Herkünften der Rot-Buche*. 10.

Schulz, K., Hänsch, R., & Sörgel, U. (2018). Machine learning methods for remote sensing applications: An overview. *Earth Resources and Environmental Remote Sensing/GIS Applications IX*, *10790*, 1079002. https://doi.org/10.1117/12.2503653

Schulze, G., & KOPP, D. (1998). Anleitung für die forstliche Standortserkundung im nordostdeutschen Tiefland (Standortserkundungsanleitung) SEA 95, Teil C–Forstliche Auswertung. *Bodenformen-Katalog. Merkmalsübersichten Und-Tabellen Für Haupt-Und Feinbodenformen. Unter Mitarbeit von D. Kopp*, *3*.

Schwartz, M. D. (Ed.). (2013). *Phenology: An Integrative Environmental Science*. Springer Netherlands. https://doi.org/10.1007/978-94-007-6925-0

Senf, C., & Seidl, R. (2021). Persistent impacts of the 2018 drought on forest disturbance regimes in Europe. *Biogeosciences*, *18*(18), 5223–5230. https://doi.org/10.5194/bg-18-5223-2021

Sibona, E., Vitali, A., Meloni, F., Caffo, L., Dotta, A., Lingua, E., Motta, R., & Garbarino, M. (2016). Direct Measurement of Tree Height Provides Different Results on the Assessment of LiDAR Accuracy. *Forests*, *8*(1), 7. https://doi.org/10.3390/f8010007

Siebert, S., & Teizer, J. (2014). Mobile 3D mapping for surveying earthwork projects using an Unmanned Aerial Vehicle (UAV) system. *Automation in Construction*, *41*, 1–14. https://doi.org/10.1016/j.autcon.2014.01.004

Signorell, A. (2022). *DescTools: Tools for Descriptive Statistics version 0.99.45 from CRAN*. https://rdrr.io/cran/DescTools/

Silva, C. A., Hudak, A., & Vierling, L. A. (2015). *RLiDAR: An R Package for Reading, Processing and Visualizing LiDAR (Light Detection and Ranging) Data*. https://cran.r-project.org/web/packages/rLiDAR/rLiDAR.pdf

Silver, N. C., & Dunlap, W. P. (1987). Averaging correlation coefficients: Should Fisher's z transformation be used? *Journal of Applied Psychology*, *72*(1), 146–148. https://doi.org/10.1037/0021-9010.72.1.146

Simpson, J. E., Holman, F. H., Nieto, H., El-Madany, T. S., Migliavacca, M., Martin, M. P., Burchard-Levine, V., Cararra, A., Blöcher, S., Fiener, P., & Kaplan, J. O. (2022). UAS-based high resolution mapping of evapotranspiration in a Mediterranean tree-grass ecosystem. *Agricultural and Forest Meteorology*, *321*, 108981. https://doi.org/10.1016/j.agrformet.2022.108981

Smaltschinski, T., & Becker, G. (2009). *Precision Forestry und forstliche Wertschöpfungskette*. 4.

Smigaj, M., Gaulton, R., Suarez Minguez, J., & Barr, S. (2017). Use of Miniature Thermal Cameras for Detection of Physiological Stress in Conifers. *Remote Sensing*, *9*, 957. https://doi.org/10.3390/rs9090957

Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A. M., Friedl, M., Braswell, B. H., Milliman, T., O'Keefe, J., & Richardson, A. D. (2012). Digital repeat photography for phenological research in forest ecosystems. *Agricultural and Forest Meteorology*, *152*, 159–177. https://doi.org/10.1016/j.agrformet.2011.09.009

Spathelf, P., Stanturf, J., Kleine, M., Jandl, R., Chiatante, D., & Bolte, A. (2018). Adaptive measures: Integrating adaptive forest management and forest landscape restoration. *Annals of Forest Science*, *75*(2). https://doi.org/10.1007/s13595-018-0736-4

Spurr, S. H. (1960). *Photogrammetry and Photo-Interpretation. With a section on applications to forestry. Second edition.* Ronald Press.

St-Onge, B., Audet, F.-A., & Bégin, J. (2015). Characterizing the Height Structure and Composition of a Boreal Forest Using an Individual Tree Crown Approach Applied to Photogrammetric Point Clouds. *Forests*, *6*(11), 3899–3922. https://doi.org/10.3390/f6113899

Straub, C., Stepper, C., Seitz, R., & Waser, L. T. (2013). Potential of UltraCamX stereo images for estimating timber volume and basal area at the plot level in mixed European forests. *Canadian Journal of Forest Research*, *43*(8), 731–741. https://doi.org/10.1139/cjfr-2013-0125

Šumarstvo, P. (2010). *Precision Forestry – Definition and Technologies*. 10.

Taylor, S. E., Veal, M. W., Grift, T. E., McDonald, T. P., & Corley, F. W. (2002). Precision Forestry: Operational Tactics For Today And Tomorrow. *In Proceedings of the International Meeting of the Council on Forest Engineering*, 7.

Thenkabail, P. S. (2015). *Land Resources Monitoring, Modeling, and Mapping with Remote Sensing*. CRC Press. https://doi.org/10.1201/b19322

Thies. (2022). *thiesclima.com—Produkte*. https://www.thiesclima.com/de/Produkte/Sonstige-Geraete-Wetterstationen/

Tomasella, M., Beikircher, B., Häberle, K.-H., Hesse, B., Kallenbach, C., Matyssek, R., & Mayr, S. (2018). Acclimation of branch and leaf hydraulics in adult Fagus sylvatica and Picea abies in a forest through-fall exclusion experiment. *Tree Physiology*, *38*(2), 198–211. https://doi.org/10.1093/treephys/tpx140

Tomaštík, J., Mokroš, M., Surový, P., Grznárová, A., & Merganič, J. (2019). UAV RTK/PPK Method—An Optimal Solution for Mapping Inaccessible Forested Areas? *Remote Sensing*, *11*(6), 721. https://doi.org/10.3390/rs11060721

Tomst. (2022). https://tomst.com/web/en/

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127–150.

van der Werf, G. W., Sass-Klaassen, U. G. W., & Mohren, G. M. J. (2007). The impact of the 2003 summer drought on the intra-annual growth pattern of beech (Fagus sylvatica L.) and oak (Quercus robur L.) on a dry site in the Netherlands. *Dendrochronologia*, *25*(2), 103–112. https://doi.org/10.1016/j.dendro.2007.03.004

van Mantgem, P. J., & Stephenson, N. L. (2007). Apparent climatically induced increase of tree mortality rates in a temperate forest. *Ecology Letters*, *10*(10), 909–916. https://doi.org/10.1111/j.1461-0248.2007.01080.x

Vastaranta, M., Niemi, M., Karjalainen, M., Peuhkurinen, J., Kankare, V., Hyyppä, J., & Holopainen, M. (2014). Prediction of Forest Stand Attributes Using TerraSAR-X Stereo Imagery. *Remote Sensing*, *6*(4), 3227–3246. https://doi.org/10.3390/rs6043227

Vilhar, U., Beuker, E., Mizunuma, T., Skudnik, M., Lebourgeois, F., Soudani, K., & Wilkinson, M. (2013). Tree Phenology. In *Developments in Environmental Science* (Vol. 12, pp. 169–182). Elsevier. https://doi.org/10.1016/B978-0-08-098222-9.00009-1

Vollmer, M., & Möllmann, K.-P. (2018). *Infrared Thermal Imaging: Fundamentals, Research and Applications*. John Wiley & Sons.

Wallace, L., Lucieer, A., Malenovský, Z., Turner, D., & Vopěnka, P. (2016). Assessment of Forest Structure Using Two UAV Techniques: A Comparison of Airborne Laser Scanning and Structure from Motion (SfM) Point Clouds. *Forests*, *7*(12), 62. https://doi.org/10.3390/f7030062

Wan, Q., Brede, B., Smigaj, M., & Kooistra, L. (2021). Factors Influencing Temperature Measurements from Miniaturized Thermal Infrared (TIR) Cameras: A Laboratory-Based Approach. *Sensors*, *21*, 8466. https://doi.org/10.3390/s21248466

Wang, C. (2021). At-Sensor Radiometric Correction of a Multispectral Camera (RedEdge) for sUAS Vegetation Mapping. *Sensors*, *21*(24), Article 24. https://doi.org/10.3390/s21248224

Wang, Y., Lehtomäki, M., Liang, X., Pyörälä, J., Kukko, A., Jaakkola, A., Liu, J., Feng, Z., Chen, R., & Hyyppä, J. (2019). Is field-measured tree height as reliable as believed – A comparison study of tree height estimates from field measurement, airborne laser scanning and terrestrial laser scanning in a boreal forest. *ISPRS Journal of Photogrammetry and Remote Sensing*, *147*, 132–145. https://doi.org/10.1016/j.isprsjprs.2018.11.008

Watts, S. B., & Tolland, L. (2005). *Forestry handbook for British Columbia, Part 2*. Forestry Undergraduate Society, University of British Columbia.

West, P. W. (2015). *Tree and Forest Measurement*. Springer International Publishing. https://doi.org/10.1007/978-3-319-14708-6

White, J., Coops, N., Wulder, M., Vastaranta, M., Hilker, T., & Tompalski, P. (2016). Remote Sensing Technologies for Enhancing Forest Inventories: A Review. *Canadian Journal of Remote Sensing*, 1–23. https://doi.org/10.1080/07038992.2016.1207484

White, J., Stepper, C., Tompalski, P., Coops, N., & Wulder, M. (2015). Comparing ALS and Image-Based Point Cloud Metrics and Modelled Forest Inventory Attributes in a Complex Coastal Forest Environment. *Forests*, *6*, 3704–3732. https://doi.org/10.3390/f6103704

White, K., Pontius, J., & Schaberg, P. (2014). Remote sensing of spring phenology in northeastern forests: A comparison of methods, field metrics and sources of uncertainty. *Remote Sensing of Environment*, *148*, 97–107. https://doi.org/10.1016/j.rse.2014.03.017

White, M. A., Running, S. W., & Thornton, P. E. (1999). The impact of growing-season length variability on carbon assimilation and evapotranspiration over 88 years in the eastern US deciduous forest. *International Journal of Biometeorology*, *42*(3), 139–145. https://doi.org/10.1007/s004840050097

Wood, S. N. (2017). *Generalized Additive Models*. 497.

Xu, Z., & Zhu, D. (2018). High-resolution miniature UAV SAR imaging based on GPU Architecture. *Journal of Physics: Conference Series*, *1074*(1), 012122. https://doi.org/10.1088/1742-6596/1074/1/012122

Yi, K., Dragoni, D., Phillips, R. P., Roman, D. T., & Novick, K. A. (2017). Dynamics of stem water uptake among isohydric and anisohydric species experiencing a severe drought. *Tree Physiology*, treephys;tpw126v1. https://doi.org/10.1093/treephys/tpw126

Zakrzewska, A., Kopeć, D., Krajewski, K., & Charyton, J. (2022). Canopy temperatures of selected tree species growing in the forest and outside the forest using aerial thermal infrared (3.6–4.9 µm) data. *European Journal of Remote Sensing*, *55*(1), 313–325. https://doi.org/10.1080/22797254.2022.2062055

Zeng, L., Wardlow, B. D., Xiang, D., Hu, S., & Li, D. (2020). A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sensing of Environment*, *237*, 111511. https://doi.org/10.1016/j.rse.2019.111511

Zhang, X. (2012). *Phenology and Climate Change*. https://doi.org/10.5772/2146

Zhao, M., Peng, C., Xiang, W., Deng, X., Tian, D., Zhou, X., Yu, G., he, H., & Zhao, Z. (2013). Plant phenological modeling and its application in global climate change research: Overview and future challenges. *Environmental Reviews*, *21*. https://doi.org/10.1139/er-2012-0036

Zhu, W., Mou, M., Wang, L., & Jiang, N. (2012). Evaluation of phenology extracting methods from vegetation index time series. *2012 IEEE International Geoscience and Remote Sensing Symposium*, 1158–1161. https://doi.org/10.1109/IGARSS.2012.6351342

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine*, *5*(4), 8–36. https://doi.org/10.1109/MGRS.2017.2762307

Zimmermann, F., Eling, C., Klingbeil, L., & Kuhlmann, H. (2017). Precise Positioning of UAVs – Dealing With Challenging RTK-GPS Measurement Conditions During Automated UAV Flights. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *IV-2/W3*, 95–102. https://doi.org/10.5194/isprs-annals-IV-2-W3-95-2017

Zweifel, R. (2016). Radial stem variations—A source of tree physiological information not fully exploited yet. *Plant, Cell & Environment*. https://doi.org/10.1111/pce.12613

Zweifel, R., Haeni, M., Buchmann, N., & Eugster, W. (2016). Are trees able to grow in periods of stem shrinkage? *New Phytologist*, *211*. https://doi.org/10.1111/nph.13995

Zweifel, R., & Häsler, R. (2000). Frost-induced reversible shrinkage of bark of mature subalpine conifers. *Agricultural and Forest Meteorology*, *102*, 213–222. https://doi.org/10.1016/S0168-1923(00)00135-0

Zweifel, R., Zimmermann, L., & Newbery, D. M. (2005). Modeling tree water deficit from microclimate: An approach to quantifying drought stress. *Tree Physiology*, *25*(2), 147–156. https://doi.org/10.1093/treephys/25.2.147