

**Česká zemědělská univerzita v Praze**

**Fakulta lesnická a dřevařská**

**Katedra hospodářské úpravy lesa**



**Popis rozdílu mezi borovicí lesní a smrkem ztepilým na  
úrovni bodových mračen individuálních stromů pomocí  
matematicko-statistických metod**

**Diplomová práce**

**Autor: Bc. Tomáš Cimr**

**Vedoucí práce: Ing. Martin Slavík, Ph.D.**

**© 2022 ČZU v Praze**

# ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE

Fakulta lesnická a dřevařská

## ZADÁNÍ DIPLOMOVÉ PRÁCE

Bc. Tomáš Címr

Lesní inženýrství

Lesní inženýrství

Název práce

**Popis rozdílu mezi borovicí lesní a smrkem ztepilým na úrovni bodových mračen individuálních stromů pomocí matematicko-statistických metod**

Název anglicky

**Describing the difference of scots pine and norwegian spruce on individual tree point cloud level based on mathematical-statistical methods**

---

### Cíle práce

Cílem práce je popsat rozdíl dvou vybraných druhů lesních dřevin, konkrétně borovice lesní a smrku ztepilého, na úrovni UAV bodových mračen individuálních stromů pomocí matematicko-statistických metod.

### Metodika

- 1) Na vybraných plochách provést sběr dat pomocí UAV s technologií LIDAR
- 2) Na odvozených bodových mračnách reprezentujících jednotlivé plochy detekovat individuální stromy a jejich bodová mračna selektovat
- 3) Ze selektovaných bodových mračen individuálních stromů odvodit jednotlivé metricky
- 4) Popsat rozdíl mezi dvěma druhy lesních dřevin na základě jednotlivých metrik vybraných pomocí zvolené metody
- 5) Zpracování dat provést pomocí vybraného programovacího jazyka

Harmonogram:

LS 2020/2021 – Zpracování bodových mračen (detekce individuálních stromů)

ZS 2021/2022 – Zpracování skriptu pro výpočet metrik

LS 2021/2022 – Odvození rozdílů mezi vybranými dřevinami, finální kompilace práce

**Doporučený rozsah práce**

40-50 stran

**Klíčová slova**

LIDAR, druhová klasifikace, UAV

**Doporučené zdroje informací**

- HOLMGREN, Johan a Åsa PERSSON, 2004. Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment* [online]. 90(4), 415–423. ISSN 00344257. Dostupné z: doi:10.1016/S0034-4257(03)00140-8
- KAMIŇSKA, Agnieszka, Maciej LISIEWICZ, Krzysztof STERENĆZAK, Bartłomiej KRASZEWSKI a Rafal SADKOWSKI, 2018. Species-related single dead tree detection using multi-temporal ALS data and CIR imagery. *Remote Sensing of Environment* [online]. 219(September), 31–43. ISSN 00344257. Dostupné z: doi:10.1016/j.rse.2018.10.005
- KUŽELKA, Karel, Martin SLAVÍK a Peter SUROVÝ, 2020. Very High Density Point Clouds from UAV Laser Scanning for Automatic Tree Stem Detection and Direct Diameter Measurement. *Remote Sensing* [online]. Dostupné z: <https://www.mdpi.com/2072-4292/12/8/1236>
- NÆSSET, Erik a Terje GOBAKKEN, 2005. Estimating forest growth using canopy metrics derived from airborne laser scanner data. *Remote Sensing of Environment* [online]. 96(3–4), 453–465. ISSN 00344257. Dostupné z: doi:10.1016/j.rse.2005.04.001
- PEKÁR, Stanislav; BRABEC, Marek. Modern analysis of biological data: generalized linear models in R. Masarykova univerzita, 2016.
- SUROVÝ, Peter, Karel KUŽELKA, Azadeh ABDOLAHNEJAD, Kateřina SIROTKOVÁ, Zlatica MELICHOVÁ a Martin SLAVÍK, 2019. Aplikace dálkového průzkumu Země v lesnictví. B.m.: ČZU v Praze, Fakulta Lesnická a Dřevařská. ISBN 8001015971.
- ZHAO, Kaiguang, Sorin POPESCU, Xuellian MENG, Yong PANG a Muge AGCA, 2011. Characterizing forest canopy structure with lidar composite metrics and machine learning. *Remote Sensing of Environment* [online]. 115(8), 1978–1996. ISSN 00344257. Dostupné z: doi:10.1016/j.rse.2011.04.001

**Předběžný termín obhajoby**

2021/22 LS – FLD

**Vedoucí práce**

Ing. Martin Slavík

**Garantující pracoviště**

Katedra hospodářské úpravy lesů

Elektronicky schváleno dne 29. 3. 2022

**doc. Ing. Peter Surový, PhD.**

Vedoucí katedry

Elektronicky schváleno dne 29. 3. 2022

**prof. Ing. Róbert Marušák, PhD.**

Děkan

V Praze dne 10. 04. 2022

## **Čestné prohlášení**

Prohlašuji, že jsem svou bakalářskou práci "Popis rozdílu mezi borovicí lesní a smrkem ztepilým na úrovni bodových mračen individuálních stromů pomocí matematicko-statistických metod" vypracoval samostatně pod vedením Ing. Martina Slavíka, Ph.D. a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu literatury na konci práce. Jako autor uvedené diplomové práce dále prohlašuji a jsem si vědom, že zveřejněním diplomové práce souhlasím s jejím zveřejněním dle zákona č. 111/1998 Sb. o vysokých školách v platném znění, a to bez ohledu na výsledek její obhajoby

V Praze dne 10. 4. 2022

Bc. Tomáš Cimr

## **Poděkování**

Rád bych touto cestou poděkoval svému vedoucímu Ing. Martinu Slavíkovi, Ph.D., za umožnění zabývat se tímto tématem a také za mimořádně ochotný, odborný přístup při konzultacích. Dále chci poděkovat své rodině za pomoc a podporu.

# Popis rozdílu mezi borovicí lesní a smrkem ztepilým na úrovni bodových mračen individuálních stromů pomocí matematicko-statistických metod

## Abstrakt

Rychlý vývoj technologie bezpilotních leteckých prostředků (UAV) s technologickým pokrokem v oblasti strojového učení v kombinaci se zvyšujícími se požadavky na monitorování a sběr dat z lesních porostů, přináší nové postupy pro zpracování dat a jejich následné uplatnění pro lesní management.

Tato práce popisuje, zda je možné klasifikovat druhy, konkrétně borovici lesní (*Pinus sylvestris* L.) a smrk ztepilý (*Picea abies* L.), pomocí rozdílů v matematicko-statistických veličinách mezi bodovými mračky, pořízenými pomocí technologie LiDAR na platformě UAV. Na čtyřech plochách (dvě pro každou dřevinu) byl proveden sběr dat laserovým skenováním. Z bodových mračen byl vytvořen digitální model terénu (DTM) a model výšky zápoje (CHM) v prostředí ArcGIS 10.6. Na tento model byl aplikován algoritmus Inverse Watershed Segmentation (IWS), který vytvořil horizontální průměty korun jednotlivých stromů. Pomocí těchto ploch byla vyselektována a následně vyříznuta bodová mračna pro jednotlivé stromy. Pro tato data byly v prostředí R studio vypočítány konkrétní statistické veličiny. Statistická analýza byla provedena zobecněným lineárním modelem, který byl zpřesněn pomocí krokové regrese, a následně byly porovnány rozdíly mezi daty jednotlivých druhů stromů. Celková přesnost byla vypočítána pomocí chybové matice s výsledkem úspěšností klasifikace 95,6 % u borovice lesní a 89,8 % u smrku ztepilého.

**Klíčová slova:** LiDAR, druhová klasifikace, UAV

# Describing the difference of scots pine and norwegian spruce on individual tree point cloud level based on mathematical-statistical methods

## Abstract

The rapid development of Unmanned Aerial Vehicles (UAV) technology with technological advances in machine learning, combined with the increasing demand for monitoring and data collection from forest stands, brings new processes for data processing and their subsequent application to forest management.

This work describes whether it is possible to classify species, namely *Pinus sylvestris* and *Picea abies*, by using differences in mathematical-statistical variables between point clouds, obtained by using LiDAR technology on the UAV platform. Data collection by laser scanning was carried out on four areas (two for each tree species). A Canopy Height Model (CHM) was created from point clouds in the ArcGIS software. The Inverse Watershed Segmentation (IWS) algorithm was applied to this model and polygons were created bounding the areas of individual trees. Using these areas, point clouds for individual trees were selected and then cut out. For this data, specific statistical quantities were calculated in the R studio software. The statistical analysis was performed by a generalized linear model, which was refined by step regression and then the differences between the data of individual tree species were compared. The overall accuracy was calculated using a confusion matrix with a classification success rate of 95.6 % for pine and 89.8 % for spruce.

**Keywords:** LiDAR, species classification, UAV

# Obsah

1 Úvod.....	14
2 Cíl práce.....	15
3 Literární rešerše .....	16
3.1 Dálkový průzkum země .....	17
3.1.1 UAV .....	17
3.1.2 Senzory.....	18
3.1.3 Fotogrammetrie (letecké snímkování) .....	19
3.1.4 Letecké laserové skenování .....	20
3.2 Habitus smrku ztepilého ( <i>Picea abies</i> ) .....	21
3.3 Habitus borovice lesní ( <i>Pinus sylvestris</i> ).....	21
3.4 Metody klasifikace dat .....	22
3.4.1 Random forest .....	23
3.4.2 Neuronová síť (Neural network, Convolutional neural network) .....	25
3.4.3 K-Mean a ISODATA .....	26
3.4.4 Metoda podpůrných vektorů (support vector machine) .....	26
3.4.5 Zobecněný lineární model .....	27
3.4.6 Algoritmus Inverse Watershed Segmentation .....	28
3.5 Použití uvedených klasifikačních metod .....	28
4 Metodika .....	33
4.1 Sběr dat pomocí UAV s technologií LiDAR.....	33
4.2 Zpracování bodových mračen v prostředí ArcGIS 10.6. pro detekování jednotlivých stromů a selektování jejich bodových mračen.....	39
4.3 Zpracování selektovaných bodových mračen v prostředí R studio a odvození jednotlivých metrik .....	44
4.4 Určení rozdílů mezi dvěma druhy lesních dřevin na základě jednotlivých metrik vybraných pomocí metody zobecněného lineárního modelu.....	48
4.5 Kompletní skript v prostředí R studio .....	48
5 Výsledky .....	51
5.1 Detekování a dovození výšek jednotlivých stromů v prostředí ArcGIS 10.6. ...	51
5.2 Vybrané metriky pomocí zobecněného lineárního modelu pro určení rozdílů mezi borovicí lesní a smrkem ztepilým .....	52
5.3 Rozdíly hodnot jednotlivých metrik mezi smrkem ztepilým a borovicí lesní ..	55
6 Diskuse .....	58
6.1 Zpracování bodových mračen pro detekování a následné vyřiznutí jednotlivých stromů .....	58



6.2	Význam počítaných matematicko-statistických metrik.....	60
6.3	Srovnání přesnosti klasifikace druhů jinými metodami .....	61
7	Závěr .....	63
8	Literatura.....	64

## Seznam tabulek, obrázků a grafů

### Obrázky:

Obrázek 1 – habitus smrku ztepilého ( <i>Picea abies</i> )(Málek & Roháček, n.d.) .....	21
Obrázek 2 – habitus borovice lesní ( <i>Pinus sylvestris</i> )(Málek & Roháček, n.d.) .....	22
Obrázek 3 – dataset ( <a href="https://victorzhou.com/blog/intro-to-random-forests">https://victorzhou.com/blog/intro-to-random-forests</a> ) .....	23
Obrázek 4 – Rozhodovací stromy ( <a href="https://victorzhou.com/blog/intro-to-random-forests">https://victorzhou.com/blog/intro-to-random-forests</a> ) .....	24
Obrázek 5 – SVM rozdělení, (Andreola, 2009).....	27
Obrázek 6 - schéma hierarchické klasifikace (Bilgilioglu et al., 2017). .....	31
Obrázek 7 – RiCOPTER od firmy REIGL (foto Ing. Martin Slavík, Ph.D.) .....	33
Obrázek 8- laserový skener VUX-1UAV, (RIEGL Laser Measurement Systems, 2020) ...	35
Obrázek 9 – plocha BO1 a BO2, Liberecký kraj, ( <a href="http://www.mapy.cz">www.mapy.cz</a> ).....	37
Obrázek 10 – plocha SM1, Středočeský kraj, ( <a href="http://www.mapy.cz">www.mapy.cz</a> ) .....	37
Obrázek 11 – plocha SM2, Středočeský kraj, ( <a href="http://www.mapy.cz">www.mapy.cz</a> ) .....	38
Obrázek 12 - model výšky zápoje (CHM), výstup LAS Height Metric As Raster, ArcGIS 10.6. ....	40
Obrázek 13 – identifikované vrcholové body, výstup Canopy Peaks, ArcGIS 10.6.....	41
Obrázek 14 – výsledek použití filtru Focal Statistics, ArcGIS 10.6. ....	42
Obrázek 15 – výstup funkce Tree Crown Extend, ArcGIS 10.6. ....	43
Obrázek 16 – výsledné body ohraničující oblast pro selektování bodových mračen, ArcGIS 10.6.....	44
Obrázek 18 – rozdíl v detekci vrcholových bodů při použití funkce Canopy peaks, ArcGIS 10.6. ....	59

## **Tabulky:**

Tabulka 1 - specifikace dronu RICOPTER, (RIEGL Laser Measurement Systems, 2020) .	34
Tabulka 2 - počty detekovaných stromů .....	51
Tabulka 3 – odvozené průměrné výšky s variabilitou.....	52
Tabulka 4 – výsledky zobecněného lineárního modelu.....	52
Tabulka 5 – výsledky zobecněného lineárního modelu s krokovou regresí.....	53
Tabulka 6 – výsledky klasifikace.....	54

## **Grafy:**

Graf 1 – Histogram rozdělení predikcí, prostředí R studio .....	55
Graf 2 – Krabicové grafy pro vybrané metriky v prostředí R studio .....	57

## Seznam použitých zkratk a symbolů

ALS – Aerial Laser Scanning

CHM – Canopy Height Model

DTM – Digital Terrain Model

GPS – Global Position System

GNSS – Global Navigation Satellite System

IWS – Inverse Watershed Segmentation

LiDAR – Light Detection and Ranging

RTK – Real Time Kinematic

SfM – Structure from Motion

UAS – Unmanned Aerial System

UAV – Unmanned Aerial Vehicle

EFI – Enhanced forest inventory

DPZ – Dálkový průzkum země

SVM – Support vector machine

DSM – Digital surface model

CNN – Convolutional neural network

NIR – Blízká infračervená spektroskopie

ULS – Unmanned laser scanning

# 1 Úvod

Důležitým aspektem při studiu lesních ekosystémů je diverzita druhů dřevin. Aplikace v oblasti ochrany a udržitelného hospodaření s lesy, jako je inventarizace lesů, monitorování biologické rozmanitosti, modelování stanovišť volně žijících živočichů, řízení rizik a studie změny klimatu, jsou z velké části založeny na klasifikaci druhů stromů. Většina stávajících metod pro klasifikaci je omezena přenositelností, omezena na konkrétní druhy, nebo vyžaduje velké množství dat pro adaptaci na nové místo, což omezuje jejich použitelnost a činí je nákladově náročnými (Fassnacht et al., 2016). Vzhledem k flexibilitě získávání dat kdykoli a kdekoli se v poslední době bezpilotní letecké prostředky (UAV) stávají nezbytným nástrojem získávání dat pro podrobnou charakterizaci zápojů porostů na rozdíl od jiných platforem dálkového průzkumu země (Nevalainen et al., 2017). Většina výzkumů pro klasifikaci druhů dřevin využívá UAV jako platformu pro nesení různých typů kamer pro sběr dat v podobě snímků a následného zpracování pomocí fotogrammetrie. Technologie LiDAR (Light Detection and Ranging) se v těchto případech používá jen jako doprovodná, pro vytváření modelů, nebo ověřování dat. Nicméně jak říká (Kuželka, Slavík, & Surový, 2020), bodová mračna získaná technologií LiDAR nesené na UAV, s hustotou tisíců bodů na metr čtvereční a s dostačující přesností, dělá z UAV laserového skenování (ULS) velmi vhodný nástroj pro detailní mapování lesního porostu. LiDAR technologie by tedy mohla být sama o sobě přínosná pro získávání dat i pro složitější informace z lesního prostředí jako je právě klasifikace druhů.

## 2 Cíl práce

Cílem této práce je vyzkoušení jednoho z možných postupů klasifikace druhů stromů z dat pořízených pomocí technologie LiDAR na platformě UAV, tedy digitálních bodových mračen. Konkrétně se jedná o zjištění rozdílů v matematicko-statistických veličinách mezi bodovým mračnem z dvou porostů borovice lesní a dvou porostů smrku ztepilého. Pro tento účel bude nejprve z dat vytvořen digitální model terénu (DTM) a model výšky zápoje (CHM) v prostředí ArcGIS 10.6. Na tento model bude aplikován algoritmus Inverse Watershed Segmentation (IWS) a vytvořeny horizontální průměty korun jednotlivých stromů. Pomocí těchto ploch budou vyselektovány a následně vyříznuty z bodového mračna body pro jednotlivé stromy. Pro tato data budou v prostředí R studio vypočítány konkrétní statistické veličiny. Statistická analýza bude provedena zobecněným lineárním modelem, který bude zpřesněn pomocí krokové regrese a následně budou porovnány rozdíly mezi daty jednotlivých druhů stromů.

### 3 Literární rešerše

Základem pro běžné operace v lesích je sběr dat, který vyžaduje velkou spotřebu práce, času a stále se zvětšující ekonomické náklady na jejich pořízení. Mimo to je běžný sběr dat limitován na malé plochy, proto je omezen počet operací, které mohou být v rozumném čase zvládnuty (Goodbody, Coops, & White, 2019). S adoptováním lesnických praktik pro získání přesných dat je klíčovým požadavkem rychlost. Speciálně to platí při rychle se měnících strukturách porostů z důvodu biotických i nebiotických činitelů (Iglhaut et al., 2019).

Poskytování přesných a aktuálních dat o lesích je nezbytné pro usnadnění postupů hospodaření s lesy založených na datech a pro formulování efektivní politiky v oblasti lesnictví. Pro kvantitativní analýzu struktury a dynamiky lesa, vyhodnocování poškození a sledování regenerace lesa jsou nezbytné detailní informace o charakteristikách stromů, jako jsou počty stromů, výšky stromů, výšky korun a průměr v prsní výšce (Mohan et al., 2017). Ačkoli tradiční metody sběru dat mohou být použity k získávání podrobných informací týkající se těchto charakteristik, mohou se stát neekonomickými, časově náročnými a vyčerpávajícími, a proto nejsou ideální pro pravidelný sběr dat (Tang & Shao, 2015). Vysoké počáteční náklady inventarizace, složitost získávání dat a přetrvávající nejistota ohledně budoucího stavu lesů v důsledku změny klimatu jsou hlavními motivátory pro zlepšení a modernizaci způsobů inventarizace lesů v celosvětovém měřítku. Stejně jako v jiných oblastech hospodaření se zdroji se poptávka po obsahu a kvalitě inventarizace znásobila. Přirozená složitost lesních ekosystémů podněcuje argument, že rutinní získávání dat k aktualizaci inventarizace je nutné k zachycení a začlenění těchto změn s cílem zvýšit znalosti o dynamice lesů, zlepšit správu lesů a v konečném důsledku poskytnout daty podložená zdůvodnění pro politiku v oblasti lesnictví a životního prostředí (Goodbody et al., 2019).

Proto, aby byla vylepšena inventarizace a pro začlenění strukturálních charakteristik lesů se stále více začínají používat technologie jako je letecké laserové skenování (ALS). Tyto technologické inovace byly využity ke zvýšení hodnoty inventarizace lesa prostřednictvím zlepšení účinnosti měření, předpovídání vývoje, efektivnosti nákladů a zajištění stále se zvětšující různorodosti dat. Dále také umožnili vytváření modelů vývoje a mapových podkladů lesních porostů. Metody, které nám

umožňují získat takováto data, lze označit jako vylepšené inventarizace lesů (EFI) (White et al., 2016).

Oproti tradičním metodám sběru dat poskytují EFI nadbytek výhodných, netradičních informací, jako jsou strukturální charakteristiky lesů, které lze využít k lepšímu obhospodařování lesů. Získávání údajů ALS se záměrem generovat EFI se stalo globálně běžnější v důsledku zlepšení specifikací senzorů, kvality souborů dat a inovativního výzkumu obhospodařování lesů. Soubory dat z metod ALS jsou stále více přijímány a využívány v průmyslovém obhospodařování lesů jako metoda pro zvýšení obsahu inventarizace, jakož i překlenutí mezer mezi strategickou, taktickou, a provozní úrovní (White et al., 2016).

### **3.1 Dálkový průzkum země**

Jak již bylo zmíněno platformy pro dálkový průzkum země, jako jsou UAS, satelity a letadla s příslušnými senzory (ALS) se stávají čím dál více používanými. Stále jsou rozvíjeny pro plnou optimalizaci, mimo jiné i pro management lesů a jejich relevantnost v tomto oboru roste. Pro aplikaci na lesnictví je ovšem často nutné mít data s velmi přesným rozlišením. Protože tradiční platformy jako satelity nesplňují tyto požadavky a letecké platformy jsou velmi nákladné, nejlepší možností jsou drony neboli UAV (Pádua et al., 2017).

#### **3.1.1 UAV**

Díky flexibilitě použití se UAV (česky drony) stávají vzrůstajícím technologickým nástrojem s velkou perspektivou a také se zvyšující se použitelností. Jejich nedávný vývoj spolu s vývojem technologií jako je umělá inteligence a samoučící se systémy, vytvořili celou řadu nových možností pro využití v lesnické praxi. Usnadněním a zlepšením sběru dat v rámci časové a prostorové přesnosti s možností vytvoření upravených dat přesně podle specifické potřeby, mají potenciál nahradit stávající metody (Surovỳ & Kuželka, 2019).



UAV byly původně vyvinuty počátkem minulého století a po 50. letech jejich hlavní využití spočívalo ve vojenské oblasti prostřednictvím průzkumu nebo sledování. V posledním desetiletí UAV různých velikosti, tvarů a schopnosti získaly uplatnění i v civilních oblastech. Jedná se především o zemědělství, lesnictví, výzkum volně žijících živočichů, obhospodařování půdy, sledování dopravy a mnoho dalších (Tiberiu Paul Banu, Gheorghe Florian Borlea, & Constantin Banu, 2016).

Nejčastější klasifikace UAV je provedena dle typu vzletu a přistání.

- vzlet a přistání horizontálně, typické pro bezpilotní letadla s pevnými křídly (letadla)
- vzlet a přistání vertikálně, typické pro drony s rotačními křídly (vrtulníky, kvadrokoptéry, hexakoptéry atd.)

Stabilita UAV a pokrytá plocha při jednom přeletu jsou klíčové prvky v aplikacích dálkového průzkumu. První kategorie UAV má výhodu většího pokrytí plochy za jeden let a druhá kategorie má výhodu lepší stability, která poskytuje vyšší prostorové rozlišení, ale se zmenšenou plochou pokryvu (Tiberiu Paul Banu et al., 2016).

Další klasifikace bezpilotních letounů je založena na jejich napájení a přímo ovlivňuje maximální dobu letu. V této kategorii lze bezpilotní letouny rozdělit na dva různé typy: elektrické a spalovací. UAV s elektrickým napájením se doporučují pro aplikace dálkového snímání ve srovnání s těmi s vnitřním spalováním, které nejsou tak úsporné a mají vyšší vibrace (Tiberiu Paul Banu et al., 2016).

Dále můžeme rozdělit UAV podle ovládání. Buď mohou létat autonomně nebo jsou dálkově ovládané pomocí dálkového ovládání. Autonomní let je naplánován dříve a je vhodný pro systematické mapování (Tiberiu Paul Banu et al., 2016).

### **3.1.2 Senzory**

Senzory, které mohou být použity v DPZ, se rozdělují na aktivní nebo pasivní. Aktivní senzory jsou zodpovědné za poskytování energie potřebné k detekci objektů. Senzor vystřelí paprsek záření směrem k objektu, který je následně odražen od povrchu objektu a zpětně detekován a měřen senzorem. Laserové výškoměry, LiDARy a RADARy patří mezi nejpoužívanější aktivní senzory v DPZ. Pokud jde o pasivní senzory,

ty mohou detekovat přirozené záření, které je emitováno nebo odraženo objekty. Většina těchto senzorů pracuje ve viditelné, infračervené části elektromagnetického spektra. RGB, multispektrální, hyperspektrální a tepelné senzory jsou některé příklady pasivních senzorů. Zatímco LiDAR senzory mohou přímo získat 3D vlastnosti v podobě bodových mračen, pro dosažení stejného výsledku s některými pasivními senzory lze použít fotogrammetrické zpracování (Guimarães et al., 2020).

### **3.1.3 Fotogrammetrie (letecké snímkování)**

Jednou z hlavních metod, které využívají pasivní senzory je fotogrammetrie. Úkolem fotogrammetrie je rekonstrukce tvarů, měření objektů a určování polohy objektu na snímcích. Dnes se nejběžněji používá automatizovaný proces rekonstrukce 3D objektů z dvourozměrných fotografií. Tato technika se nazývá Structure from motion (SfM) a využívá sekvenci fotografií téhož objektu z různých pozic. Pomocí algoritmů jsou na snímcích automaticky detekovány klíčové body a z jejich vzájemného posunu na dalších snímcích vypočítána relativní poloha. Pro výpočet pozic a rekonstrukci objektu je potřeba, aby daný objekt byl statický a neměnný. Všechny vypočítané pozice jsou pouze relativní, nemají měřítko ani orientaci v prostoru. Pro jejich lokalizování se standardně používá metoda vlíčovacích bodů, které se umístí do prostoru foceného objektu tak, aby byly dobře pozorovatelné na snímcích. Tyto body mají zpravidla určené souřadnice geodetickými metodami. Díky tomu je celý systém detekovaných bodů a pozic kamer transformován do daného souřadnicového systému (Surový & Kuželka, 2019).

Pro zahájení procesu rekonstrukce SfM jsou všechny snímky zpracovány pomocí automatického algoritmu detekce a popisu funkcí nazvaného SIFT (Scale Invariant Feature Transform) (Clemons, 2007). Tento algoritmus je založený na generování klíčových bodů objektů v digitálním obraze a jejich následné identifikaci na různých snímcích. Tyto body následně mohou být detekovány i v případě posunu, rotace, či změny měřítka obrazu. Každý takový bod je převeden na vektor popisující jeho pozici. Současně se vypočítávají 3D souřadnice těchto bodů. Tisíce takovýchto vektorů pak reprezentuje celý obraz. Při spojování více obrazů jsou pak klíčové body detekovány a spojovány do skupinek (clusterů). Každý klíčový bod na různých obrazech si musí být

podobný a zároveň zachovávat svou relativní pozici vůči sousedním klíčovým bodům (Surový & Kuželka, 2019). Dalším algoritmem je SURF (Speeded Up Robust Feature). Tento algoritmus je postaven na stejném principu jako SIFT, jsou zde ale v každém kroku upraveny postupy. Díky tomu je tento algoritmus několikanásobně rychlejší.

(James & Robson, 2012) říká, že získávání odpovídajících dat většinou vyžaduje drahé nástroje a odbornost při získávání i zpracování. Nicméně ukazuje, že pomocí běžných digitálních kamer a zdarma dostupného softwaru lze díky fotogrametrii získat 3D modely terénu. Takovéto modely jsou dostačující pro aplikaci ve velké řadě vědních oborů. Další výhodou, kterou si přístupy založené na kamerách zachovávají oproti laserovému skenování jsou jejich náklady a nízká váha, proto se dají namontovat na lehké bezpilotní letecké plošiny jako jsou draci a modely vrtulníků.

#### **3.1.4 Letecké laserové skenování**

Airborne Laser Scanning (ALS) je rychlá a vysoce přesná metoda měření, která se používá k měření povrchu země pomocí světelných impulzů a jejich odrazů. Tyto světelné impulzy jsou vyzařovány z laserového skeneru (LiDAR), který je namontován na vzdušném dopravním prostředku, jako letadla a vrtulníky (Smreček, 2012).

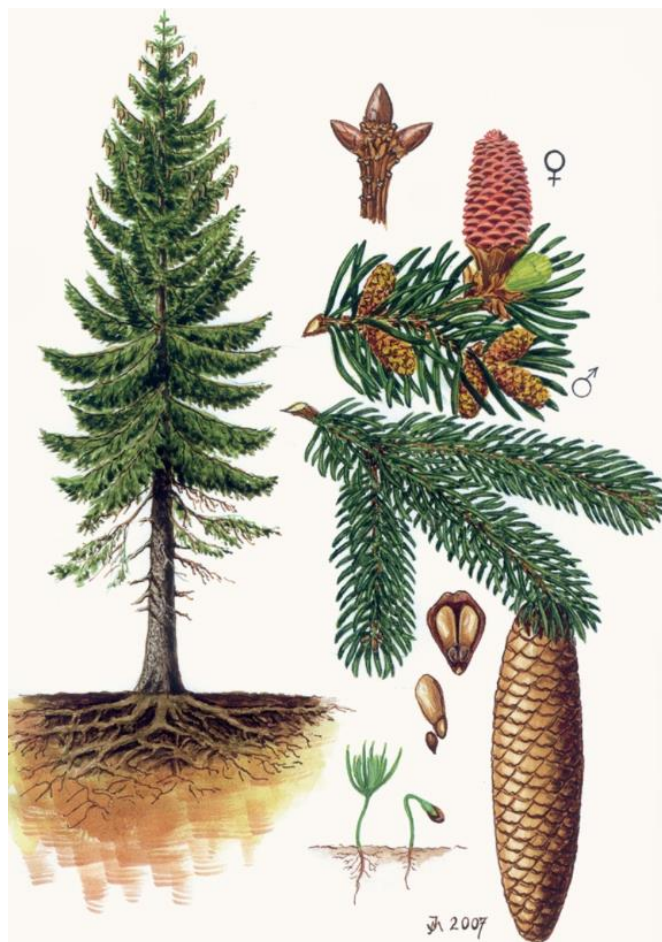
Velkou výhodou leteckého laserového skenování je schopnost poskytovat data jak z terénu, tak i z oblasti stromového zápoje. Za hlavní využití vzdušného laserového skenování lze považovat vytváření digitálních terénních a povrchových modelů. Letecké laserové skenování se používá především při geodetickém průzkumu, archeologii, urbanistickém plánování, řízení životního prostředí a výzkumu jako lesnictví, geologie, geomorfologie a seismologie (Smreček, 2012).

Základem laserového skenování je měření vzdáleností pomocí laserového skeneru. Ten vystřelí světelný paprsek, obvykle v infračerveném spektru, směrem k objektu, od něhož se následně odrazí a je zpětně zachycen senzorem. Existují dva základní typy LiDARu. První z nich je pulzní, který velmi přesně změří čas mezi vysláním paprsku a jeho zachycením. Vzdálenost je následně vypočítána jako dráha, kterou urazí světlo za daný časový úsek. Druhým typem je fázový, kdy je paprsek světla vyslán konstantně. Senzor poté zaznamenává, v jaké fázi byl paprsek zachycen a z tohoto posunu dále vypočítá vzdálenost. Při každém vyslání paprsku je také zaznamenán

vertikální a horizontální úhel. To umožňuje přesné určení 3D pozice bodů vůči skeneru (Surový & Kuželka, 2019).

### 3.2 Habitus smrku ztepilého (*Picea abies*)

Smrk ztepilý je rychle rostoucí vždyzelený jehličnatý strom s rovným, až do vrcholu koruny průběžným kmenem, kuželovitou korunou a špičatým vrcholem, větvemi v pravidelných přeslenech, mírně převisající. Jeho habitus je lehce závojovitý. Je schopen dorůst výšky 50 až 70 m a průměru kmene až 2 m a je tak nejvyšším domácím stromem v Evropě (Musil & Hamerník, 2003).



Obrázek 1 – habitus smrku ztepilého (*Picea abies*)(Málek & Roháček, n.d.)

### 3.3 Habitus borovice lesní (*Pinus sylvestris*)

Vždyzelený jehličnatý strom, je to druh poměrně variabilní, u kterého je možné rozlišit několik zeměpisných poddruhů odlišných podle místa původu, proto má i habitus proměnlivý. V mládí vytváří kuželovitou korunu, ve stáří až deštníkovitou, dosahuje

vysokého stáří (500–1000 let), má tmavý přímý kmen a poměrně hustou tmavou korunu, dorůstá výšky 30 m, na optimálním stanovišti v domovině až 50 m, kořenový systém je mohutně vyvinutý, s dlouhým křivým kořenem, kmen 100–180 cm tlustý (výčetní tloušťka), letorosty silné, zelenohnědé, lysé, kůra šedá, šedohnědá nebo tmavohnědá, zimní pupeny vejcovité nebo cylindricky vejcovité, slabě pryskyřičnaté; jehlice po 2 ve svazku mohou být až 19 cm dlouhé, tmavozelené, vytrvávají asi 4 –7 let (Musil & Hamerník, 2003).



Obrázek 2 – habitus borovice lesní (*Pinus sylvestris*)(Málek & Roháček, n.d.)

### 3.4 Metody klasifikace dat

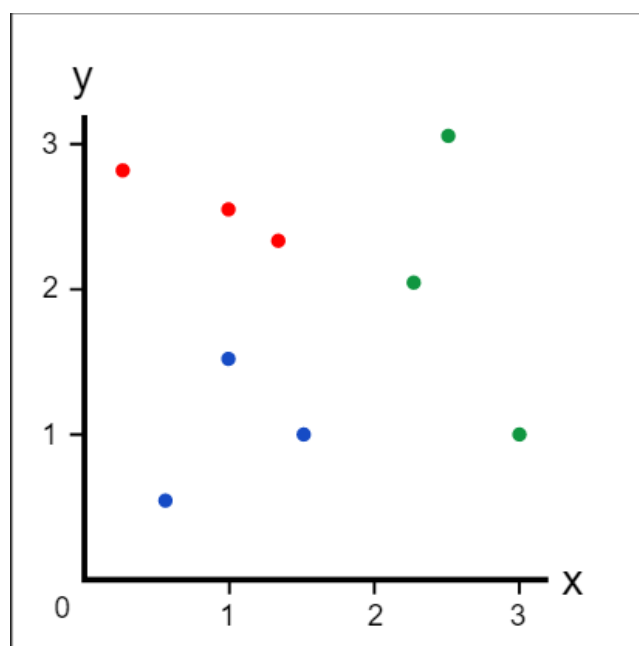
Klasifikace je jednou z nejvíce používaných analýz pro data z dálkového průzkumu země. V literatuře je popsáno několik metod, které mohou být rozděleny podle různých kritérií, například klasifikace podle pixelů, subpixelů nebo objektů. V průběhu minulých desetiletí, vzhledem k hrubému rozlišení satelitních snímků, byla analýza ve většině klasifikačních technik na úrovni pixelů. Tyto techniky byly založeny pouze na spektrální

informaci jednotlivých pixelů. Jedná se o Minimum Distance Classifier, Mahalanobis Distance Classifier, Parallelepiped, Maximum-Likelihood Classifier, Support Vector Machine, Decision Tree, Random Forest a Artificial Neural Network (Amini, Homayouni, Safari, & Darvishsefat, 2018).

### 3.4.1 Random forest

Pro využití moderních datových souborů velkých velikostí, potřebujeme nyní samoučící algoritmy, které dokážou držet tempo s objemem informací a zároveň zachovávají dostatečnou statistickou efektivitu. Náhodné lesy, které vymyslel Breiman na počátku 21. století (Jin et al., 2020), jsou na seznamu nejúspěšnějších metod, které jsou v současnosti k dispozici pro zpracování takovýchto dat. Tento proces strojového učení pod dohledem, ovlivněný ranou prací (Amit, Geman, & Wilder, 1997), (Ho, 1998) a (Dietterich, 2000), funguje podle jednoduchého, ale účinného principu „rozděl a panuj“: tedy rozdělit data na malé vzorky, dále na každém malém kousku dat vytvořit tzv. „Decision tree“ (prediktor), podle kterého bude algoritmus rozdělovat data, následně sloučit (agregovat) tyto prediktory dohromady (Biau & Scornet, 2016).

Prvním krokem při použití této metody je naučit algoritmus na malém vzorku dat podle jakých vstupních hodnot bude data rozdělovat. Pro přiblížení principu na zjednodušeném příkladu, předpokládejme že vzorek dat se skládá ze tří datových tříd, kde každá z nich má určený parametr, souřadnice  $x$ ,  $y$ .



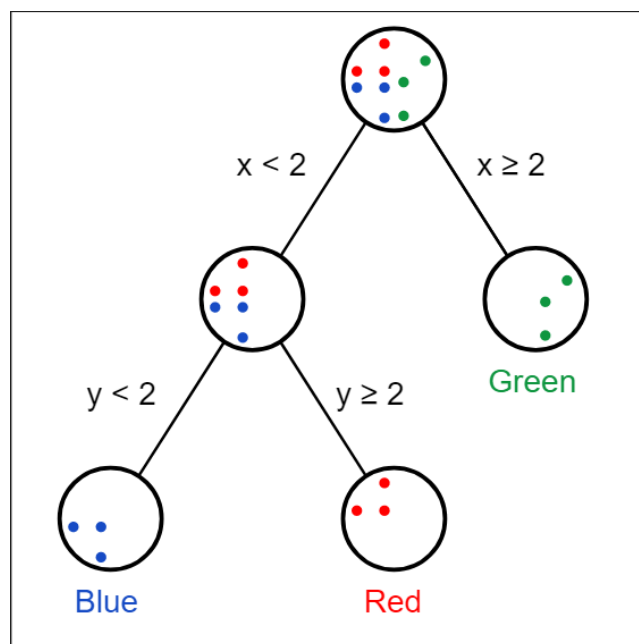
Obrázek 3 – dataset (<https://victorzhou.com/blog/intro-to-random-forests>)

Cílem je, aby rozdělení dat bylo co nejpřesnější, proto je nutné nejprve vypočítat v jakém úseku rozdělení provedeme. K tomu slouží například GINI index, pomocí kterého je vypočítána pravděpodobnost nesprávného zařazení náhodně vybraného prvku v datovém souboru, pokud by byl náhodně označen podle rozdělení tříd v datovém souboru. Vypočítá se jako:

$$G = 1 - \sum C p(i)^2$$

kde  $C$  je počet tříd a  $p(i)$  je pravděpodobnost náhodného vybrání prvku z třídy  $i$ . V tomto příkladě vyšlo jako nejlepší rozdělení na hodnotě  $x = 2$  a  $y = 2$ .

Následně se stanoví, že první rozhodovací strom bude testovat zda  $x > 2$ . Pokud naše datová třída testem projde je označena zeleně. V případě rozhodovacího stromu  $x < 2$  nám vyjdou dvě datové třídy, které je potřeba následně znovu rozdělit. To bude v našem případě stanoveno rozhodovacím stromem  $y < 2$ , bude tedy označena modře a pro  $y > 2$  označena červeně. Výsledné rozdělení vidíme na obrázku č.2.



Obrázek 4 – Rozhodovací stromy (<https://victorzhou.com/blog/intro-to-random-forests>)

Jakmile jsou všechny prvky z datového vzorku rozděleny, je vytvořen již zmiňovaný prediktor (Decision tree), neboli jeden strom.

Pro vytvoření celého náhodného lesa je potřeba tyto prediktory sloučit. Na to se používá technika zvaná „bagging“, znamená to agregování všech vytvořených prediktorů. Nicméně tyto prediktory mají jen jeden parametr, např. počet stromů.

Náhodný les má ještě druhý parametr, který řídí kolik funkcí je třeba vyzkoušet při hledání nejlepšího rozdělení. V tomto příkladě byly jenom dva parametry, ale většina datových souborů bude mít stovky až tisíce parametrů. Místo toho, aby se při každém vytvoření prediktoru zkoušely všechny parametry, zkouší se pouze podmnožina. Dělá se to proto, aby byla vytvořena nahodilost, která jednotlivé predikátory dělá unikátní a snižuje tak korelaci mezi nimi (Biau & Scornet, 2016).

### **3.4.2 Neuronová síť (Neural network, Convolutional neural network)**

Neuronová síť je řada algoritmů, které rozpoznávají základní vztahy v souboru dat prostřednictvím procesu, který napodobuje způsob, jakým pracuje lidský mozek. Tento koncept se silně podobá statistickým metodám jako například regresní analýza, ale má kořeny v umělé inteligenci. Základní jednotkou této sítě je Neuron, v zásadě je to matematická funkce, která shromažďuje a třídí informace podle konkrétní předem definované architektury. Spojením více Neuronů vzniká síť, která má několik vrstev vzájemně propojených uzlů. Hlavními vrstvami jsou vstupní pro zadání dat a výstupní, která provádí klasifikaci a udává výsledek, mezi nimi jsou tzv. „hidden“ vrstvy. Neuronová síť může mít jakýkoliv počet vrstev s jakýmkoliv počtem neuronů. Tato metoda se používá například pro finanční operace, plánování, analýzy, v předpovědi počasí atd. Pro nás je však důležité využití pro analýzu a identifikaci dat jako jsou digitální snímky (<https://victorzhou.com/blog/intro-to-cnns-part-1/>), (O'Shea & Nash, 2015).

Pro vytvoření neuronové sítě je potřeba nejprve nakódovat každý neuron, a protože dnešní snímky mají vysoké rozlišení, bylo by nutné nakódovat pro každý neuron každý pixel, což je nemožné. Proto se v případě snímků používá Konvoluční neuronová síť. V této síti jsou vrstvy založené na matematické operaci konvoluce, lze si to představit jako dvourozměrná pole čísel. Při zpracování snímků nejsou důležité samostatné pixely, ale spíše jejich souvislost s okolními pixely. Tato metoda tedy sjednocuje a přepočítává hodnoty několika okolních pixelů dohromady podle daného filtru, například filtr 3x3 sjednotí a přepočítá hodnotu 9 pixelů do jednoho. Tento filtr se konkrétně nazývá Sobel operátor a používá se pro detekování hran objektů



na snímcích (<https://victorzhou.com/blog/intro-to-cnns-part-1/>), (O'Shea & Nash, 2015).

### **3.4.3 K-Mean a ISODATA**

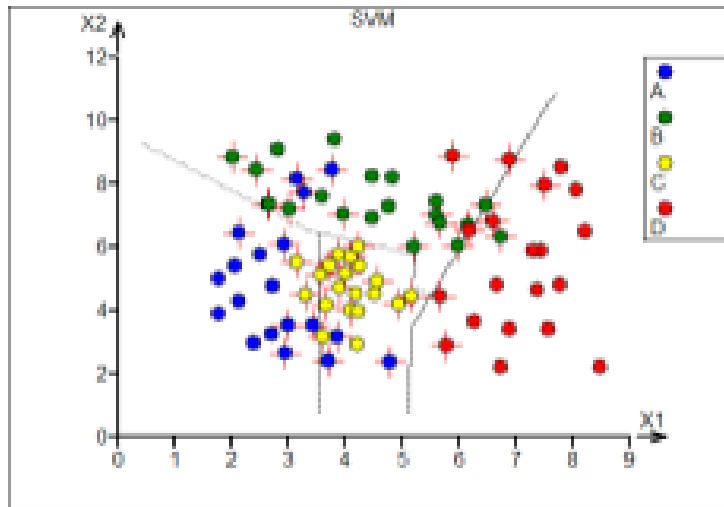
K-Means clustering je jeden z nejpůlárnějších algoritů z kategorie unsupervised machine learning. Princip tohoto algoritmu spoívá v seskupení podobných datových bodů a objevení vzorců mezi nimi. K dosažení tohoto cíle definuje pevný počet ( $k$ ) shluků neboli clusterů, které odkazují na počet tzv. centroidů. Tyto body představují imaginární, nebo reálné umístění středu clusteru. Následně je každý bod z datové řady přidělen k nejbližšímu clusteru, přičemž ponechává centroidy co nejmenší. Tímto způsobem rozdělí datový soubor do skupin se stejnými nebo podobnými vlastnostmi (Waseem et al., 2016).

Zatímco v K-Means přístupu zůstává počet clusterů ( $K$ ) stejný po celou dobu opakování, i když se později může ukázat, že více nebo méně clusterů by lépe odpovídalo datům. Tato nevýhoda může být překonána v algoritmu ISODATA, který umožňuje automatické nastavení počtu shluků. Sloučením podobných clusterů a rozdělením clusterů s velkými směrodatnými odchylkami tak vytváří přesnější rozdělení dat (Waseem et al., 2016).

### **3.4.4 Metoda podpůrných vektorů (support vector machine)**

Support vector machine (SVM), neboli Metoda podpůrných vektorů, je další progresivní metoda z oblasti strojového učení. Lze ji využít především pro klasifikační úlohy, rovněž ale nachází uplatnění v regresním modelování a neparametrických odhadech hustoty. Modely SVM dokáží minimalizovat vhodně definovanou chybu, například špatnou klasifikaci u klasifikačních modelů. Principem je nalezení přímky, která rozděljuje (diskriminuje) obě třídy tak, že zachová maximální vzdálenost od experimentálních dat a tím za obecných předpokladů minimalizuje riziko špatné klasifikace v predikci z nových dat. Model SVM pak může mimo jiné sloužit k tomu, abychom na základě nových hodnot nezávislé proměnné uměli odhadnout (včetně

pravděpodobností) do které třídy příslušný objekt spadá. Tato metoda může být použita i pro robustní regrese nebo jádrové transformace (Andreola, 2009).



Obrázek 5 – SVM rozdělení, (Andreola, 2009)

### 3.4.5 Zobecněný lineární model

Zobecněný lineární model slouží ke zkoumání vztahu mezi dvěma a více proměnnými. Skládá se z prediktoru ( $x$ ), neboli veličin známých reálných hodnot, a z veličin, které se snaží určit, tedy výsledku ( $y$ ). Tyto dvě veličiny představují ideální vztah, který je na reálných datech těžko dosažitelný. Proto se používá vztah, který v sobě zahrnuje náhodnou veličinu tzv. reziduum ( $\varepsilon$ ), reprezentující odchylku od ideálního stavu. Takovýto obecný vztah pak při počtu měření  $i$  může vypadat takto:

$$y_i = \alpha + \beta x_i + \varepsilon_i$$

(Pekár & Brabec, 2016)

V publikaci (Kuželka & Surový, 2018) píše, že zobecněný lineární model představuje flexibilnější možnost tvorby modelů oproti obecnému lineárnímu modelu, z důvodu možnosti proměnné mít nelineární vztah k prediktorům. Dále také nemusí mít normální rozdělení, ale může se použít rozdělení například binomické. Tyto dva faktory otevírají dveře daleko širšímu využití oproti obecnému modelu. Dalším rozdílem je použitá metoda k hledání řešení, zatímco u obecného lineárního modelu se používá metoda nejmenších čtverců, u zobecněného lineárního modelu je to metoda maximální věrohodnosti (maximum likelihood).

### 3.4.6 Algoritmus Inverse Watershed Segmentation

Jedním ze skupiny segmentačních modelů založených na detekování obrysů objektu v obraze je algoritmus Inverse Watershed Segmentation (IWS). Tento algoritmus je jednou z nejméně používaných technik pro segmentaci modelu výšky zápoje (CHM), protože je intuitivní brát každou konkávní korunu stromu v obráceném CHM jako mísu. Díky detekování lokálního maxima a nejbližších minimálních hodnot, dokáže tento algoritmus rozčlenit povrch rastru do jednotlivých misek. Na základě tohoto přístupu dokáže IWS segmentovat odlišné stromy s výškovými hodnotami a průměry korun (Edson & Wing, 2011).

Tato metoda předpokládá, že místní maxima přítomné v CHM jsou vrcholy stromů, nicméně je nutné dát si pozor při aplikování lokálního maxima v komplexních strukturách porostů. Někdy může dojít například k identifikování několika lokálních maxim v rámci jedné koruny kvůli nepravidelnosti a částečně i náhodným chybám spojených se vznikem modelu CHM. Možným řešením je použití vyhlazovacích filtrů, zde je ale nutná opatrnost, protože silné filtry mohou eliminovat menší stromy (Panagiotidis, Abdollahnejad, Surový, & Chiteculo, 2017).

### 3.5 Použití uvedených klasifikačních metod

Ve studii (Nevalainen et al., 2017) zkoumali účinnost fotogrammetrie a hyperspektrálních dat získaných z UAV pro detekci jednotlivých stromů a klasifikaci druhů v boreálních lesích. Klasifikace byla provedena několika metodami (k-Nearest Neighbors, Naive Bayes, Decision Trees, Multilayer Perception, Random Forest). Celková přesnost datového souboru se všemi funkcemi byla nejvyšší právě u metody Random forest a Multilayer Perception. Přesnost identifikace jednotlivých stromů se pohybovala mezi 40 % a 95 % v závislosti na charakteristice oblasti. Klasifikace bez normalizace dat přinesla mírně horší výsledky ve srovnání s výsledky dosaženými s normalizovanými daty. Bez normalizovaných dat poskytl Random Forest nejlepší výsledek s celkovou přesností 93 %.

Dalším příkladem může být studie (Michez, Piégay, Lisein, Claessens, & Lejeune, 2016). Tato studie navrhuje metodický rámec pro rozlišení druhů a jejich zdravotního

stavu v lesích kolem povodí z časových řad snímků získaných během vegetativního období. Pomocí Agisoft Photoscan bylo provedeno fotogrammetrické zpracování do digitálního modelu povrchu (DSM). Pro určení výškové polohy zde byl také použit digitální model terénu (DTM) získaný z leteckého laserového skenování. Kombinací těchto dvou modelů získali výškový model zápoje (CHM), který byl použit pro následnou segmentaci a klasifikaci. Pro identifikaci, která proměnná je pro rozlišení druhů nebo zdravích stromů nejrelevantnější, využili algoritmu Random forest. Pro implementaci tohoto algoritmu byl využit balíček BVSURF v prostředí R (Allaire, 2011). Přesnost klasifikace se lišila podle rozsahu analýzy, což naznačuje, že celková přesnost se zlepšovala se zmenšující se velikostí analyzovaných objektů. Jako nejvhodnější proměnné pro účel této studie byly identifikovány tzv. Band ratios ze spektrálního zobrazení, následované proměnnými vztahující se k vertikální struktuře lesa. Celkové výsledky klasifikace ukázali přesnost od 79,5 % do 84,1 % pro rozlišení druhového složení. Pro klasifikaci týkající se zdravotního stavu bylo dosaženo nejlepší přesnosti 90,6 %.

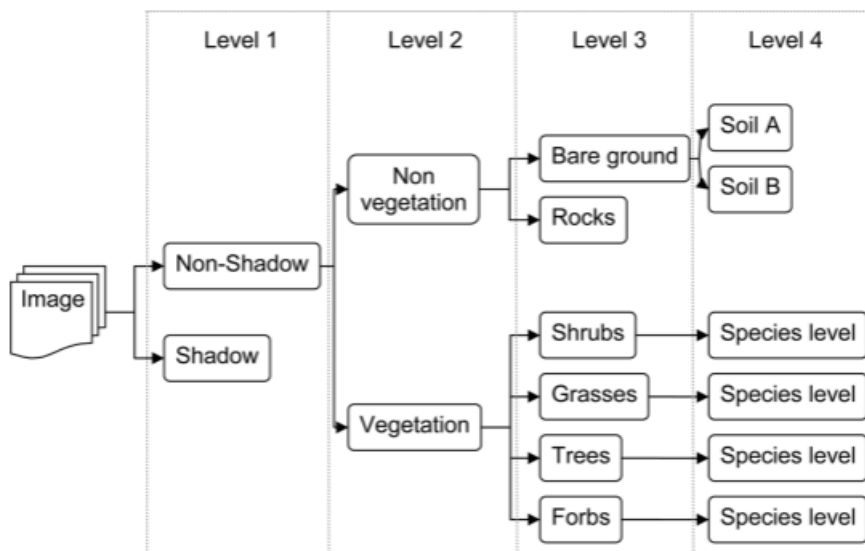
Metoda konvolučních neuronových sítí byla použita například ve studii (Morales et al., 2018), pro segmentování specifického druhu palmy z leteckých snímků pořízených pomocí RGB kamer namontovaných na UAV. Navrhnutá byla sémantická segmentace pomocí CNN v programu Deeplab, s architekturou skládající se ze tří hlavních modulů: enkodér, ASPP (Atrous Spatial Pyramid Pooling) a dekodér. Zjednodušený princip této architektury spočívá v aplikování několika různých konvolucí pro snížení velikosti obrazu, následně dalších konvolucí pro analyzování rysů, jako jsou hrany nebo tvary obrazu, v různých měřítcích a převedení pomocí Sigmoidovy funkce na binární obraz. Nakonec je obraz převeden zpátky na původní velikost snímku. Hlavním rozdílem oproti běžným strukturám bylo použití tzv. Atrozních oddělitelných konvolucí. Pro porovnání výsledků a výkonu této navržené sítě s rozdílným přístupem segmentování, navrhly další čtyři sítě založené na struktuře U-NET. Pro statistickou analýzu chování jejich sítě vůči ostatním spočítali čtyři metriky: přesnost vůči celkovému pozorování (ACC), přesnost pozorování mezi sebou (PREC), citlivost (SN) a specifčnost (SP). Při srovnání pozorovali, že jejich metoda dosahovala nejlepších výsledků ve všech statistických parametrech. Dále také měla jejich metoda lepší výkon v předvídání nových vzorků mimo trénovací sadu. Použitím Atrozních oddělitelných konvolucí se také výrazně snížil počet výpočtů

a celková doba provedení. Celková přesnost této metody, co se týče segmentace druhu palmy, byla 98 % ACC, 96 % PREC, 95 % SN a 96 % SP. Kromě dospělých a oddělených jedinců byla tato metoda schopna detekovat také palmy v mladém věku, nebo palmy částečně zakryté jinými druhy vegetace (Morales et al., 2018).

Ve studii (Gini, Passoni, Pinto, & Sona, 2014) zkoumali rozdíly klasifikace druhů stromů mezi kontrolovanými a nekontrolovanými algoritmy. Sběr dat byl proveden pomocí UAS a ze snímků z kamer RGB a NIR byl vytvořen digitální model povrchu zkoumané plochy. Poté byly ze snímků generovány ortofoto mapy a napasovány pomocí bodů tak, aby odpovídaly modelu povrchu. Tyto vrstvy modelů byly nejprve klasifikovány pomocí nekontrolovaného algoritmu ISODATA (pozn. kontrolovaný algoritmus znamená, že je nejprve vytrénován na malém vzorku dat). Výsledky této klasifikace nebyly dostačující téměř pro žádnou kategorii povrchu, kromě kategorií, které byly jednoznačně rozpoznatelné, jako například betonový povrch. Kategorie vegetace byly přeházené a sloučené do sebe. Nicméně jak tato studie uvádí, i takováto klasifikace může být užitečná pro předběžnou analýzu, jako krok před aplikováním kontrolovaného algoritmu. Výsledky byly vyhodnoceny pomocí chybové matice, která ukazuje, kolik pixelů z každé kategorie bylo správně zařazeno. Celková přesnost této metody byla 50 %.

Druhou metodou byla zvolena klasifikace pomocí algoritmu Maximum Likelihood, u kterého bylo nejprve provedeno trénování na vzorku dat. Do tohoto algoritmu byla vložena data z předchozí klasifikace. Kvalita klasifikace byla vyhodnocena znovu chybovou maticí a porovnána s předchozími výsledky. Z celkové analýzy výsledků vyplývá, že klasifikace druhů vegetace nebyla přesná. U kategorií, které nespádají do vegetace, byla přesnost 76 %. Pro druhovou klasifikaci vegetace byla celková přesnost největší u dvou druhů stromů (Pajasan 76 % a Akát 66 %), nicméně ostatní druhy byly většinou nesprávně zařazeny do ostatních kategorií a tím snížily celkovou přesnost. Na závěr je zde zmíněno, že použitím komplementárních dat mohou být některé nedostatky této metody vyloučeny, například použitím výškového modelu zápoje by se zamezilo špatné klasifikaci vegetace s výškovými rozdíly (Gini et al., 2014).

Využitím hierarchické klasifikace obrazu se zabývali ve studii (Bilgilioglu, Ozturk, Sariturk, & Seker, 2017). Cílem této studie bylo otestovat proveditelnost přístupu k hodnocení a monitorování krajiny na bázi UAV pro účely kvantifikace standardních ukazatelů vegetace rostlinných společenstev v americkém Idahu. Konkrétními cíli bylo porovnání měření na bázi obrazu s pozemním měřením z hlediska přesnosti a spotřeby času. Ke sběru dat byly využity dva senzory, video kamera a digitální RGB kamera. Zpracování obrazu se sestávalo z operací ortorektifikace a sestavení mozaiky snímků, oříznutí mozaiky do 50 x50 m ploch (které byly zároveň měřeny na zemi) a následné klasifikace obrazu. Pro klasifikaci obrazu byl použit program OBIA (object-based), který je vhodnější pro účely klasifikace snímků s vysokým rozlišením. Obraz byl segmentován pomocí segmentace spektrálních rozdílů, která vytváří větší obrazové objekty s podobnými spektrálními odezvami (velké holé plochy), při zachování malých objektů uvnitř (keře). Následně bylo použito schéma hierarchické klasifikace (obrázek č. 6), kdy byl snímek nejprve klasifikován do stínových/nestínových oblastí, poté následovalo další oddělení do skupin struktury vegetace a poté, pokud to bylo možné, na úroveň druhu.



Obrázek 6 - schéma hierarchické klasifikace (Bilgilioglu et al., 2017).

Přístup hierarchické klasifikace má podle této studie několik výhod. První z nich je, že snadněji definovatelné třídy jsou na vrcholu hierarchie. Další výhodou může být aplikace stejného algoritmu analýzy na snímky různých zkoumaných ploch jen změnou

prahových hodnot pro třídy. Tato metoda také umožňuje porovnání výsledků z různých míst a za různých podmínek (Bilgilioglu et al., 2017).

Porovnání pozemního měření s daty získaných ze snímků bylo provedeno určením korelačních koeficientů a pomocí párového t-testu. Pro posouzení přesnosti obrazové mozaiky použili stratifikovaný náhodný bodový vzorkovací přístup s 600 body pro pět tříd a vytvořili chybovou matici pro určení celkové, produkční a uživatelské klasifikace přesnosti a statistiky Kappa. Korelace mezi dvěma metodami pro odhad procentuálního pokrytí vedla k hodnotám  $r$ -kvadrát v rozmezí od 0,86 do 0,98. Celková přesnost klasifikace pro mozaiky obrazů byla 83 až 88 %. Odhadovaná spotřeba času byla výrazně lepší pro metodu snímkování (Bilgilioglu et al., 2017).

(Bilgilioglu et al., 2017) došli tedy k závěru, že hierarchický klasifikační přístup je vhodný pro aplikování klasifikace rostlinných společenstev a výsledky jsou použitelné pro různé klasifikační úrovně (tj. vegetační/nevegetační, na úrovni struktur nebo druhů) a dále mohou být využity například pro manažerská rozhodnutí.

## 4 Metodika

### 4.1 Sběr dat pomocí UAV s technologií LiDAR

#### Použité vybavení

Pro sběr dat byl použit bezpilotní systém (UAS) VUX-SYS s dronem RiCOPTER od společnosti REIGL (viz obrázek č. 7).



Obrázek 7 – RiCOPTER od firmy REIGL (foto Ing. Martin Slavík, Ph.D.)

Jedná se o bateriově poháněný octacopter, tedy osm vrtulí, uspořádaných na čtyřech ramenech po dvou. Konstrukce ramen a krytu se skládá z karbonových vláken pro co největší odlehčení. Samotný dron (bez baterií) váží do 11 kg s tím, že může nést zátěž až 8 kg. Takto se dostává na celkovou váhu do 25 kg, díky tomu může operovat ve většině evropských zemích v kategorii lehkých regulací (viz tabulka 1). Podvozek je antivibrační, což umožňuje stabilní let. Platforma je mechanicky i elektricky vybavena pro připojení a nesení sensorových systémů.

Pro větší bezpečnost má dvě řídicí jednotky. V případě poruchy hlavní řídicí jednotky nastupuje záložní. Zároveň je v řídicí jednotce zabudovaný autopilot systém.



Ten umožňuje létání po předem definované dráze, díky tomu můžeme plně využít délku letu, která je omezená výdrží baterií. Také díky tomu můžeme opakovat lety nad stejným územím se stejnou dráhou letu, například při sběru časových dat. Další výhodou autopilota je, že dron letí rovně a stabilně (RIEGL Laser Measurement Systems, 2020). Specifikace a výkon dronu udává následující tabulka č. 1.

Tabulka 1 - specifikace dronu RICOPTER, (RIEGL Laser Measurement Systems, 2020)

<b>Rozměry</b>	
Plně rozložený	1 920 mm x 1 820 mm x 470 mm
Složený pro transport	624 mm x 986 mm x 470 mm
<b>Maximální váha při letu</b>	25 kg
<b>Maximální váha připojených senzorů</b>	Do 6,5 kg
<b>Váha bez baterií senzorů</b>	11 kg
<b>Maximální testovaná a povolená nadmořská výška (při dodržení podmínek)</b>	Až do 4000 m.n.m
<b>Maximální délka letu</b>	30 min
<b>Rychlost letu</b>	6-8 m/sec
<b>Typ vzletu/ přistání</b>	vertikální

Maximální povolená rychlost větru: 8 m/sec

Minimální teplota vzduchu: -5 °C

Maximální teplota vzduchu: +40 °C

Na dron byl namontován laserový skenovací systém VUX-SYS, značky REIGL, s těmito komponenty:

Laserový skener VUX-1UAV

Inerciální jednotka APX-20 UAV

Kamera Sony Alpha 7R III

## Laserový skener VUX-1UAV



Obrázek 8- laserový skener VUX-1UAV, (RIEGL Laser Measurement Systems, 2020)

Na obrázku č. 8 vidíme kompaktní, odlehčený laserový skener s možností více odrazů. Se zorným polem 330° je navržen tak, aby mohl být připevněn na platformu jakýmkoliv směrem, podle požadavků daného letu. Je napájený pomocí jednoho kabelu se spotřebou 60 W. Poskytuje velmi rychlé získávání dat pomocí úzkého infračerveného paprsku a rychlého, skenovacího line mechanismu. Tento mechanismus porovnává body v jedné horizontální nebo vertikální rovině najednou, díky tomu může určit hodnoty v rovině přilehlých bodů najednou. Data jsou ukládána na SSD disk s kapacitou 1TB a/nebo živě přenášena do pozemní stanice pomocí integrovaného LAN-TCP/IP připojení (RIEGL Laser Measurement Systems, 2020).

**Frekvence laseru:** 50-550 kHz

**Dosah laseru:**

minimální dosah: 3 m

Maximální dosah: až 1000 m. Liší se v závislosti na frekvenci laseru (PRR-pulse repetition rate), podmínkách slunečního záření a výšce skenovaného cíle.

**Přesnost:** 10 mm

**Váha:** okolo 3,75 kg

**Rozměry:** 227 x 209 x 129 mm

### **Inerciální jednotka APX-20 UAV**

Inerciální jednotka (IMU/GNSS) zodpovídá za GPS orientaci celého systému. Připojením přes GPS anténu a pomocí akcelerometrů, gyroskopů a magnetometrů udává směr, náklon a rychlost pohybu UAV. Umožňuje tedy určit velmi přesnou GPS pozici skenovacího zařízení v reálném čase, což je extrémně důležité pro dodržení přesné dráhy letu autopilotem a také pro přesnou orientaci bodových mračen.

Nezbytnou součástí vedle systému RiCOPTER je také pozemní stanice pro plánování letu a následnou koordinaci během letu.

### **Příprava vybavení**

Aby v terénu všechno fungovalo, je nutné předem zajistit řádnou přípravu technického vybavení. V první řadě je to vyzkoušení funkčnosti všech částí samotného UAV, tedy motorů, vrtulí, zapojení kabelů, kontrola případného vnějšího poškození, dále pak dobití baterií u UAV, kamery, dálkového ovládače a počítače.

Nesmíme také zapomenout na kompatibilitu softwaru. Dále bylo potřeba vyzkoušet spojení signálu mezi UAV a pozemní stanicí pro plánování letu. Pro skener bylo nutné zajistit dostatečnou kapacitu paměti pro ukládání dat.

Důležité bylo také předběžné naplánování a nastavení softwaru pro trasování letu UGCS, abychom zajistili co nejefektivnější sběr dat. To znamená, abychom pokryli celé zájmové území při co nejkratší době letu. Tato trasa se samozřejmě může změnit podle samotných podmínek v terénu.

### **Sběr dat**

Sběr dat byl proveden celkem na čtyřech plochách, dvě plochy borovice lesní (*Pinus sylvestris*) a dvě plochy smrku ztepilého (*Picea abies*).

První dvě plochy označeny jako BO1 a BO2 byly situovány na školním pozemku u obce Doksy, okres Česká Lípa v Libereckém kraji, v rámci Katedry hospodářská úprava lesa na ČZU viz. obrázek č. 9.



Obrázek 9 – plocha BO1 a BO2, Liberecký kraj, (www.mapy.cz)

Třetí plocha označená jako SM1 se nacházela jižně od obce Oplany, v okrese Praha-východ ve Středočeském kraji (obrázek č. 10).



Obrázek 10 – plocha SM1, Středočeský kraj, (www.mapy.cz)

Čtvrtá plocha SM2 se nacházela jižně od obce Jevany, v okrese Praha-východ ve Středočeském kraji (obrázek č. 11)



Obrázek 11 – plocha SM2, Středočeský kraj, (www.mapy.cz)

Po příjezdu na místo vždy bylo zkontrolováno, zda podmínky, se kterými se počítalo v před přípravném plánování, vyhovují. Bylo ověřeno, zdali je vhodné počasí a zda nebyly provedeny změny na místech, které by nám znemožňovaly použití letového plánu. Jednalo se hlavně o umístění pozemní stanice a nouzových přistávacích ploch.

Následně bylo vyjmuto vybavení z auta. Na odletovém místě byla roztažena a upevněna plachta, aby se zamezilo víření prachu a listí, které by nám mohlo ohrozit bezpečný start. Na plachtu byl rozložen samotný dron a pečlivě zkontrolováno upevnění všech částí dronu a veškerého vybavení. Mezitím byla na známý bod ustavena a zapnuta pozemní RTK GPS stanice pro přesnou orientaci. Pro orientaci byl na všech plochách zvolen mezinárodní souřadnicový systém WGS 84/ UTM 33N, protože při použití souřadnicového systému pro Českou republiku S-JTSK dochází při vyhodnocování dat k problému s úhlovým vyrovnáním. Dále byla připravena pozemní stanice. Byl tedy zapnut notebook se softwarem UGCS na plánování letu a s datalinkem přes anténu byl připojen dron.

Po připojení potřebného softwaru pozemní stanice bylo UAV uvedeno do chodu a před zahájením letu se čekalo na kalibraci inerciální jednotky. Poté pilot zahájil manuální vzlet a ve vzduchu obletěl dvakrát osmičku pro další kalibraci. Následně s dronem vystoupal do požadované výšky a doletěl na počáteční bod naplánovaného letu. Od této chvíle byl zapnut autopilot, který převzal řízení dronu. Ten létal po předem naplánované trajektorii vzoru (zig-zag pattern). Skenování probíhalo jenom v rovině, tedy operátor ho musel v zatáčkách manuálně vypínat. Po dokončení skenování byla opět převzata kontrola řízení pilotem, ten opět obletěl dvě

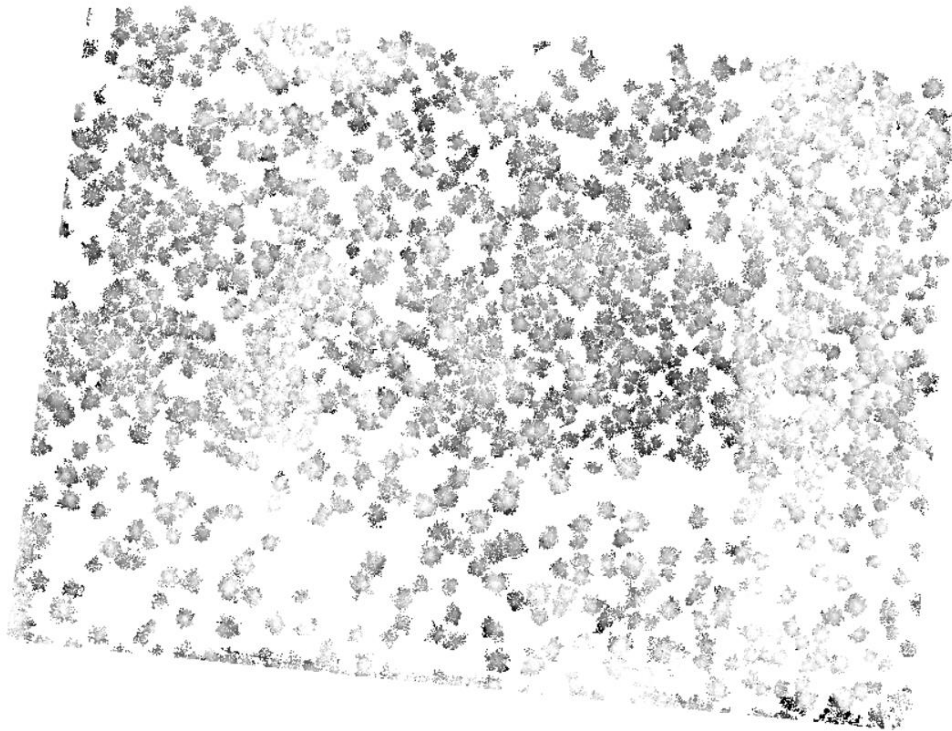
osmičky, následně přistál, a ještě pět minut probíhala kalibrace dronu po skenování. Veškerá data byla stažena z disku dronu a uložena do počítače. Poté byl dron vypnut a složen zpět pro převoz.

## **4.2 Zpracování bodových mračen v prostředí ArcGIS 10.6. pro detekování jednotlivých stromů a selektování jejich bodových mračen**

Prvním krokem zpracování bylo získaná data upravit a vypočítat souřadnice veškerých bodů z bodového mračna do souřadnicového systému. Úprava dat spočívala v odstranění bodů v prostoru, které nám nerepresentují žádný objekt. Tyto úpravy byly provedeny v programu RiPROCESS a následně vyexportovány do souborů ve formátu LAS. Celkem nám takto vznikly čtyři datové soubory, každý z jedné zkoumané plochy.

Z tohoto formátu bylo pro naši klasifikaci potřeba rozdělit bodové mračno do skupin bodů podle jednotlivých stromů, tak abychom byli schopni tyto body pro jednotlivé stromy vyříznout. Tato operace byla provedena v prostředí ArcGIS 10.6. a přidaného nástroje 3D Sample Tools.

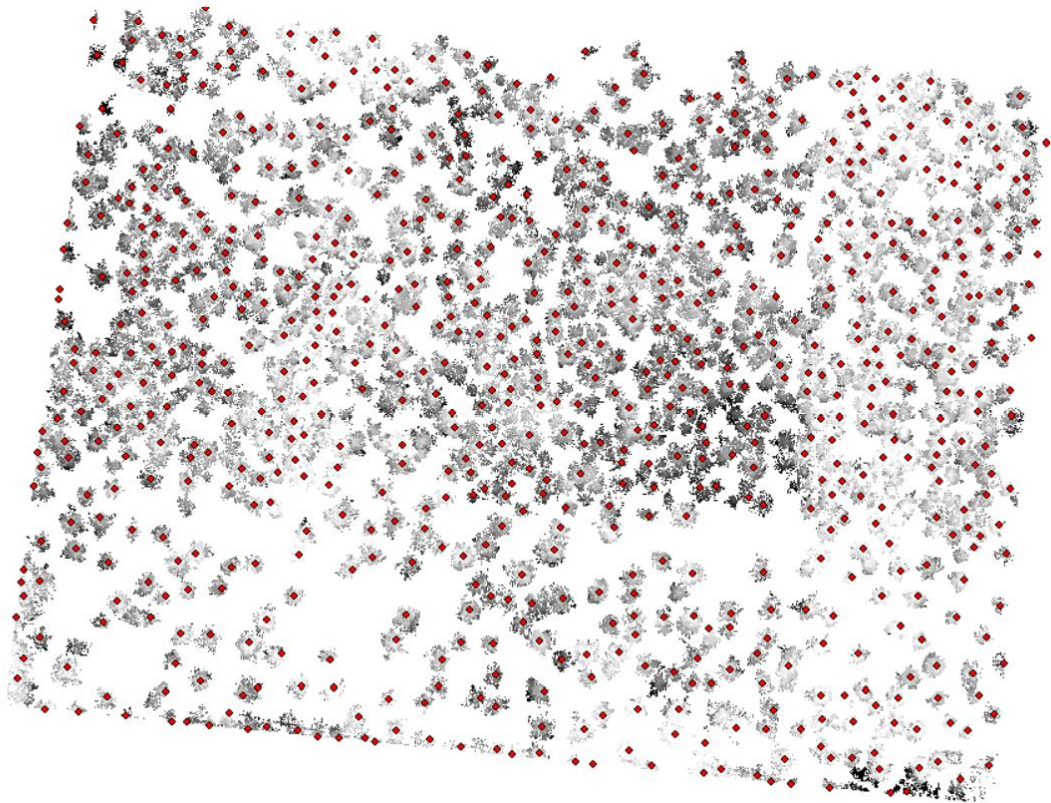
Po nahrání datového souboru bylo dalším krokem převést data, přesněji hodnoty výšek, do rastrového formátu pomocí nástroje LAS Height Metric As Raster. V tomto nástroji bylo upřesněno několik kritérií: velikost buňky/pixelu rastrového formátu na 0,05 m, minimální výška 3 metrů pro vyloučení bodů nad zemí, které pro naše účely nejsou potřebné, a dále minimální počet čtyř bodů v jedné buňce. Následným výstupem je model výšky zápoje (CHM), který je zobrazen na obrázků č. 12.



Obrázek 12 - model výšky zápoje (CHM), výstup LAS Height Metric As Raster, ArcGIS 10.6.

Následně bylo potřeba určit co nejpřesněji vrcholové body každého jednotlivého stromu v CHM. Pro tuto operaci byla využita funkce Canopy Peaks. Tato funkce vyhledala lokální maxima, tedy body s největší souřadnicí „z“ podle nastavených konstant A, B, C a velikosti okna, pro který vyhledá jeden bod. Pro vyhledání vrcholových bodů pro co nejvíce stromů bylo potřeba parametry pro tuto funkci upravovat do doby, než bylo patrné identifikování většiny stromů. V některých případech byly vyhledány body v prostoru, kde viditelně žádný strom nebyl, takovéto body se následně manuální editací smazaly v pozdějším kroku. Výsledné body vidíme na obrázku č. 13.

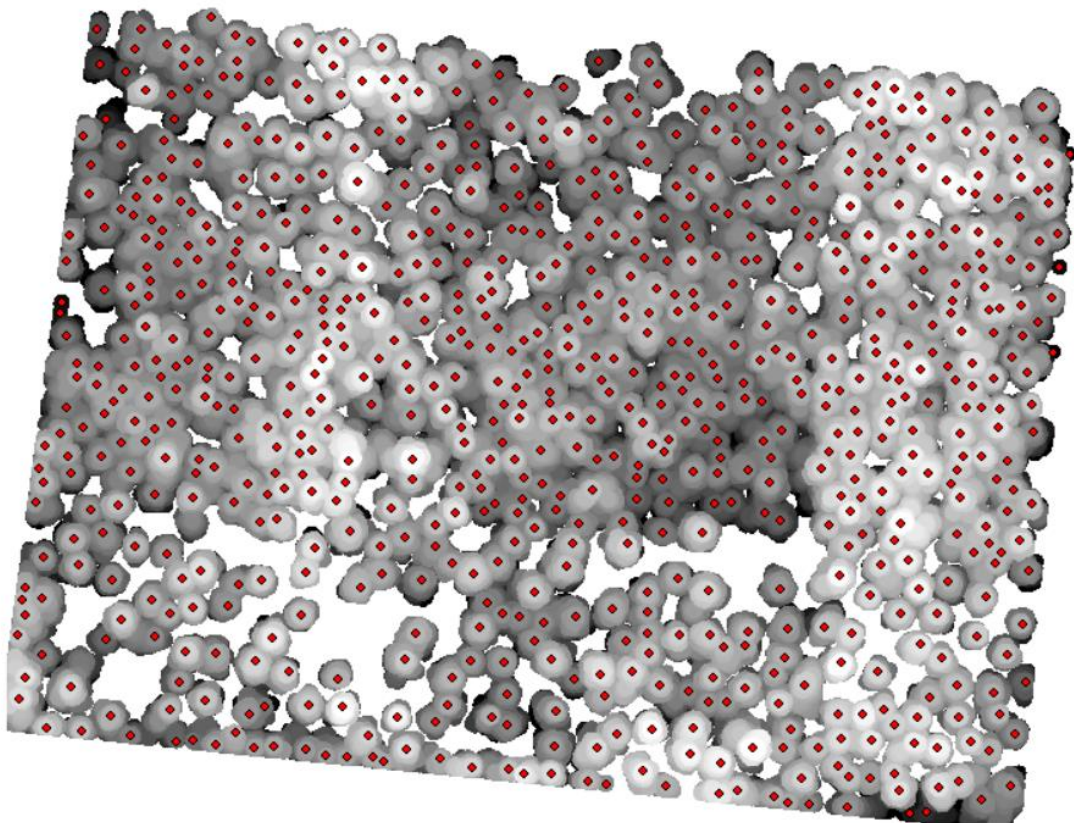
V tomto kroku byly také z CHM vypočítány výšky detekovaných stromů a z nich spočítána průměrná výška pro jednotlivé plochy.



Obrázek 13 – identifikované vrcholové body, výstup Canopy Peaks, ArcGIS 10.6.

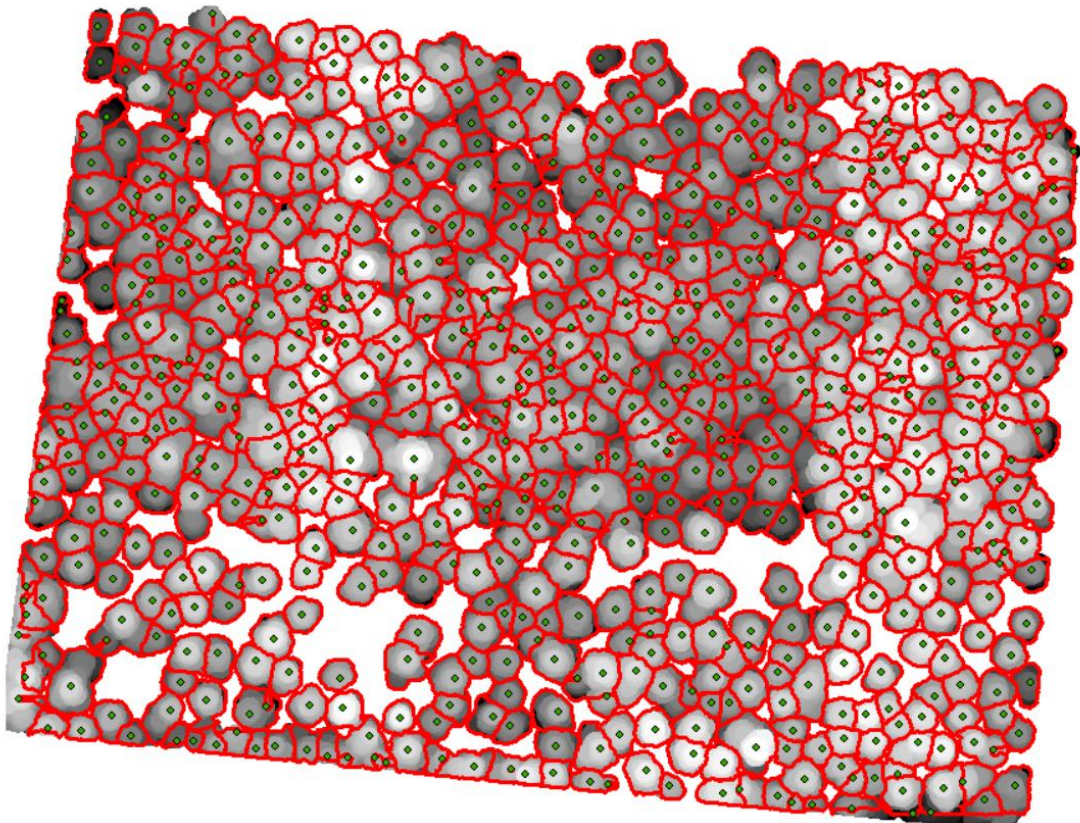
Dalším krokem pro identifikování jednotlivých stromů bylo použití filtru (funkce Focal Statistics) pro vyrovnání hodnot na kruhových plochách kolem vrcholových bodů. Tento filtr vypočítal výstupní hodnoty buněk v dané oblasti (kruhu) pomocí funkce hodnot všech vstupních buněk. Pro to, abychom eliminovali zdvojená lokální maxima a zlepšili tak následné vytváření polygonů, byla tato statistická funkce zvolena jako maximální. Dalším vstupním parametrem, který bylo potřeba nastavit, byla velikost poloměru kruhu. Tento parametr rovněž ovlivnil následné vytváření polygonů. V některých případech bylo zahrnuto do jedné plochy více stromů, nebo byl jeden strom rozdělen do více ploch. Proto bylo potřeba vyzkoušet více hodnot poloměru. I když se veškeré chyby nepodařilo úplně vyloučit, zvolena byla taková hodnota, u níž byla viditelně nejmenší chybovost. Pro každou plochu byla hodnota poloměru jiná z důvodu rozdílnosti porostních ploch. Výsledný výstup vidíme na obrázku č. 14.





Obrázek 14 – výsledek použití filtru Focal Statistics, ArcGIS 10.6.

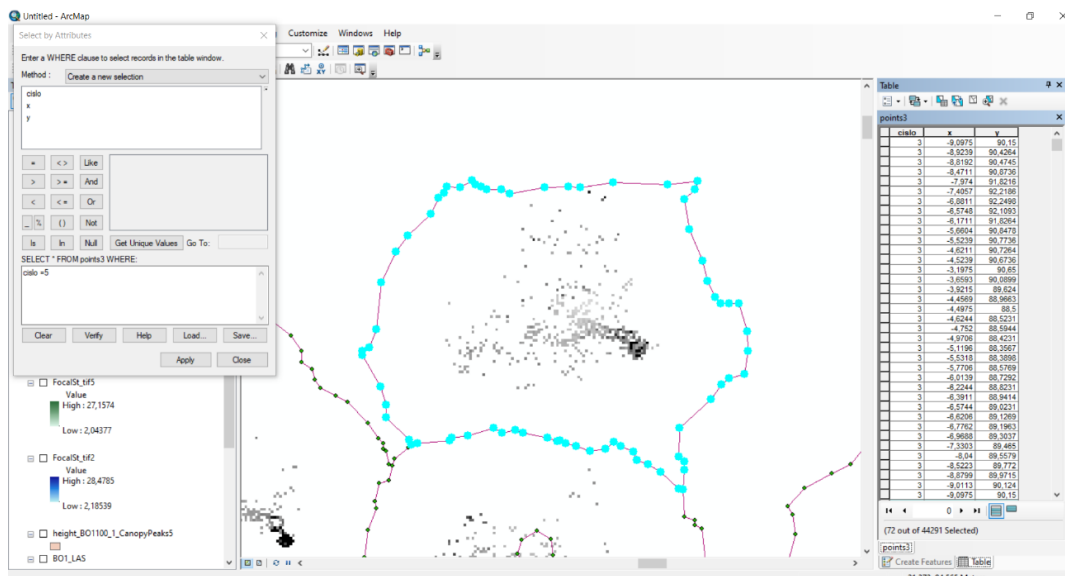
Pro dosažení úplné segmentace stromů a jejich korun byl následně využit algoritmus Inverse Watershed Segmentation pomocí funkce Tree Crown Extend. Tento algoritmus na základě hodnot buněk vyhledal linie, které ohraničují koruny stromů, a vytvořil tak polygony kolem vrcholových bodů. Jak již bylo zmíněno, tato funkce je značně ovlivněna strukturou porostu a předchozími kroky, a proto výsledné polygony ne vždy odpovídají reálným korunám, jak můžeme vidět na obrázku č. 15.



Obrázek 15 – výstup funkce Tree Crown Extend, ArcGIS 10.6.

Pro následné zpracování byly tyto výstupy ještě ručně upraveny. Pomocí editace došlo k odmazání některých polygonů, například pokud byly uvnitř jiného polygonu.

Posledním krokem bylo takto ohraničené segmenty vyříznout z bodového mračna. Pro takovou operaci bylo potřeba identifikovat hranice každého polygonu a každý z nich musel být číselně označen. To bylo provedeno editací tabulky, kdy byl přidán sloupec s čísly označující pořadí polygonu. Funkcí Polygons to lines byly převedeny polygony na linie. Linie byly následně převedeny funkcí Vertices to points na body. V tomto kroku bylo potřeba zkontrolovat, zda se nám číselné označení polygonů, které jsme předtím editovali, přepsalo do tabulky u bodů. Výsledné body reprezentující polygony i s výslednou tabulkou vidíme na obrázku č. 16. V tabulce vidíme označení bodů podle reprezentujícího polygonu i s vypočítanými souřadnicemi (x, y). Tato tabulka byla následně vyexportována do textového souboru. Vyříznutí bodů poté proběhlo na základě polygonů v prostředí R studio.



Obrázek 16 – výsledné body ohraničující oblast pro selektování bodových mračen, ArcGIS 10.6.

### 4.3 Zpracování selektovaných bodových mračen v prostředí R studio a odvození jednotlivých metrik

Vyříznutím bodů, které představují jednotlivé stromy, jsme získali data ve formě textových souborů, uspořádané podle jednotlivých ploch do samostatných složek. Každý textový soubor obsahuje informace o bodech tvořící právě jeden strom. Zde bylo nutné některé textové soubory vyloučit pro nedostatek bodů, protože by při následném zpracování docházelo k chybám při výpočtech. Tento nedostatek bodů je způsoben několika důvody: chybovostí zpracování v prostředí ArcGIS 10.6. v předchozích krocích, nedostatečným počtem bodů pro reprezentaci celého stromu, chybným určením stromu (např. u terénní nerovnosti, kterou software určil jako strom). Proto byly soubory s menší velikostí než 5 kb vyřazeny z následných výpočtů.

Následné zpracování a vypočítání statistických veličin bylo provedeno v prostředí R studio. Pro správné fungování všech příkazů, bylo nutné instalování balíčků tzv. library. Doinstalovány byly Tidyverse, Rlas, Plyr a writexl. Prvním krokem bylo vytvoření skriptu, který nahrál textové soubory ze složky, tedy nastavení přístupové cesty pomocí příkazu: `setwd("C:/Users/cimrt/Desktop/stromy/bo1")`

Dalším krokem bylo vytvoření tabulky pro vypočítání metrik. Abychom vytvořili tabulku, která bude mít odpovídající počet řádků, bylo nejprve potřeba definovat,

pro kolik stromů na dané ploše máme data, tedy textové soubory. Počet textových souborů byl definován příkazem `num_files` a následná tabulka vytvořena pomocí příkazu `df`:

```
num_files <- length(list.files("C:/Users/cimrt/Desktop/stromy/bo1"))
df <- setNames(data.frame(matrix(0,num_files,1)), "mean")
```

Po vytvoření tabulky "df" s počtem řádků odpovídající počtu textových souborů a jedním sloupcem bylo potřeba vytvořit příkaz tak, aby nám vypočítal a do této tabulky zapsal námi zvolené metriky pro každý strom. Pro tuto operaci jsme zvolili funkci for loop. Jedná se prakticky o smyčku, která nahraje první strom, vypočítá dané veličiny, zapíše do tabulky a následně nahraje druhý strom. Takováto operace je provedena pro všechny stromy v dané složce. Tímto způsobem dostaneme tabulku s vypočítanými veličinami v jednotlivých řádcích pro každý strom. Zjednodušený příkaz pro tuto funkci vypadá takto:

```
for (tree in 1 : num_files) {
  filename <- paste ("tree (",as.character(tree), ").txt", sep = "")
  T <- read.table(filename, header = TRUE, dec = ".", sep = ",")
  names(T)
}
```

Zjednodušený je proto, že do takovéhoho příkazu bylo potřeba ještě vepsat právě ty metriky, které jsme si zvolili a chceme spočítat. Jelikož naše data s body měla absolutní výšky, bylo nutné přepočítat tyto výšky na normalizované a dále počítat s nimi. Tento výpočet byl první částí kódu, který byl vložen do již zmiňovaného for loopu:

```
T$h <- T$Z - min(T$Z)
T$hn <- (T$h/max(T$h))*100
```

Následně byly skriptem definovány tyto statistické veličiny:

**Aritmetický průměr:** který vezme sumu všech hodnot a vydělí ji celkovým počtem hodnot (Lepš & Šmilauer, 2016).

Obecný vzorec:  $X = \frac{X_1+X_1+\dots+X_n}{n}$  Skript: `df$mean <- mean(T$hn)`

**Kvadratický průměr:** představuje druhou odmocninu aritmetického průměru druhých mocnin daných hodnot. Umocnění hodnot má za následek větší váhu hodnot vzdálenějších od nuly. Dále je také vždy nezáporný a větší nebo roven aritmetickému průměru (Lepš & Šmilauer, 2016).

Obecný vzorec:  $K = \sqrt{\frac{X_1^2 + X_2^2 + \dots + X_n^2}{n}}$  Skript: `df$kvad <- sqrt(mean(T$hn^2))`

**Kubický průměr:** představuje třetí odmocninu aritmetického průměru třetích mocnin daných hodnot. Má stejné vlastnosti jako Kvadratický průměr (Lepš & Šmilauer, 2016).

Obecný vzorec:  $K = \sqrt[3]{\frac{X_1^3 + X_2^3 + \dots + X_n^3}{n}}$  Skript: `df$klub <- sqrt(mean(T$hn^3))`

**Průměrná absolutní odchylka:** jedná se o průměr z absolutních odchylek od centrálního bodu. Centrálním bodem může být průměr, nebo medián, nebo jakýkoliv náhodný bod související s danou sadou dat. Vypočítá se součtem hodnot rozdílů mezi body od střední hodnoty a vydělením celkovým počtem bodů (Lepš & Šmilauer, 2016).

Obecný vzorec:  $d = \frac{\frac{1}{n} \sum_{i=1}^n |x_i - m(X)|}{n}$  Skript: `df$pd <- mean(abs(T$hn - mean(T$hn)))`

**Směrodatná odchylka:** jedná se odmocninu z rozptylu náhodné veličiny. Vypovídá o tom, nakolik se od sebe navzájem liší jednotlivé případy v souboru zkoumaných hodnot. Pokud je malá, jsou si prvky souboru navzájem podobné a naopak. Nejčastěji se používá jako míra statistické variability (Lepš & Šmilauer, 2016).

Obecný vzorec:  $\sigma = \sqrt{E((X - E(X))^2)}$  (X-veličina; E(X) její střední hodnota)

Skript: `df$sd <- sd(rnorm(T$hn))`

**Rozptyl:** jedná se o charakteristiku variability rozdělení pravděpodobnosti náhodné veličiny, která vyjadřuje variabilitu rozdělení souboru náhodných hodnot její střední hodnoty. Je to tedy průměr druhých mocnin odchylek od aritmetického průměru (Lepš & Šmilauer, 2016).

Obecný vzorec:  $Var(X) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$  Skript: `df$var <- var(rnorm(T$hn))`

**Koeficient šikmosti:** jedná se o charakteristiku rozdělení náhodné veličiny, která popisuje jeho nesymetrii. Pokud je její hodnota nulová, značí to rovnoměrné rozdělení na obě strany od střední hodnoty, tedy normální rozdělení. V případě kladné hodnoty jsou vpravo od průměru odlehlejší hodnoty nežli vlevo (rozdělení má tzv. pravý ocas) a většina hodnot se nachází blízko vlevo od průměru. U záporné šikmosti je tomu naopak (Lepš & Šmilauer, 2016).

Obecný vzorec:  $\gamma = \frac{E(X-E(X))^3}{(var X)^{3/2}}$  (E(X) – střední hodnota; var X – rozptyl)

Skript: `df$ks <- psych::skew(T$hn)`

**Kvantil:** dělí seřazený soubor hodnot na několik zhruba stejně velkých částí. Je to tedy číslo, které udává míru polohy rozdělení pravděpodobností náhodné veličiny. Do této skupiny spadá například medián, tercil, kvartil, percentil (Lepš & Šmilauer, 2016).

Skript: `df$quant5 <- quantile(T$hn,0.05)`

**Kvartil:** jedná se o typ kvantilu. Jsou to tři body, které rozdělují data do čtyř stejných skupin, z nich každá představuje čtvrtinu vzorku dat. První kvartil se tedy rovná 25 % dat seřazených postupně (oddělí nejmenších 25 % dat od největších 75 %) (Lepš & Šmilauer, 2016).

Skript: `df$quart1 <- diff(quantile(T$hn, c(0, 0.25)))`

Poslední sloupec v tabulce identifikoval označení druhu dřeviny, abychom následně mohli porovnávat vypočítané metriky mezi sebou. Pro jednoduchost byla označena borovice číslem 0 a smrk číslem 1. To bylo provedeno skriptem: `df$druh <- 1`.

Kompletní skript for loopu byl poté spuštěn pro každou plochu zvlášť, výstupem tedy byly čtyři tabulky obsahující vypočítané metriky pro každý jednotlivý strom. Tyto tabulky byly vyexportovány do formátu csv pomocí skriptu `write.csv(df, "přístupová cesta",row.names = FALSE)` a následně spojeny do jednoho kompletního datasetu "Komplet.csv".

## 4.4 Určení rozdílů mezi dvěma druhy lesních dřevin na základě jednotlivých metrik vybraných pomocí metody zobecněného lineárního modelu

Protože se v našem případě jedná o porovnání pouze dvou druhů dřevin, byl pro analýzu dat zvolen zobecněný lineární model, který byl vypočítán opět v prostředí R studio. Pro tento výpočet byl do softwaru nahrán kompletní dataset: `(data <- read.table("Komplet.csv", header = TRUE, sep = ";"))`. Odvození zobecněného lineárního modelu bylo definováno pomocí skriptu: `full.model <- glm(druh ~ ., data = data, family = binomial)`, kdy všechny metriky byly vázány na druh dřeviny a bylo použito binomické rozdělení, na základě dvou druhů dřevin. Poté bylo potřeba vypočítaný model zobrazit pomocí: `summary(full.model)`. Následně pro zpřesnění zobecněného lineárního modelu byla použita kroková regrese. Ta provedla výpočet lineární regrese a vyhodnotila, která metrika není signifikantní, a ta byla z výpočtu odstraněna. Tento proces se několikrát opakoval do doby, než se lineární model přestal zpřesňovat. Cílem bylo nalezení metrik, které jsou nejvýznamnější z hlediska rozdílů mezi dřevinami. Tento krok byl definován skriptem: `step <- step(full.model, direction = "both")`. Kroková regrese byla definována tak, aby proběhla oběma směry pro větší přesnost. Následně opět pro zobrazení: `summary(step)`.

Posledním krokem bylo vytvoření krabicových grafů z vypočítaných metrik pro lepší přehlednost a srovnání rozdílů mezi dřevinami. Příklad skriptu:

```
boxplot(mean ~ druh, data = data, main = "Aritmeticky prumer", ylab = "hodnota metriky", xlab = "drevina", outline = FALSE)
```

## 4.5 Kompletní skript v prostředí R studio

```
#nastavení přístupové složky
```

```
setwd("C:/Users/cimrt/Desktop/stromy/bo1")
```

```
#smazání mezipaměti
```

```
rm(list = ls())
```

```

#definování počtu souborů
    num_files <- length(list.files("C:/Users/cimrt/Desktop/stromy/bo1"))
#založení tabulky
    df <- setNames(data.frame(matrix(0,num_files,1)), "mean")
#smyčka for loop s normalizováním výšky a se statistickými metrikami
    for (tree in 1 : num_files) {
        filename <- paste ("tree (",as.character(tree), ").txt", sep = "")
        T <- read.table(filename, header = TRUE, dec = ".", sep = ";")
        names(T)
        T$h <- T$Z - min(T$Z)
        T$hn <- (T$h/max(T$h))*100
        df$mean[tree] <- mean(T$hn)
        df$kvad[tree] <- sqrt(mean(T$hn^2))
        df$skub[tree] <- sqrt(mean(T$hn^3))
        df$spd[tree] <- mean(abs(T$hn - mean(T$hn)))
        df$sd[tree] <- sd(rnorm(T$hn))
        df$var[tree] <- var(rnorm(T$hn))
        df$ks[tree] <- psych::skew(T$hn)
        df$quant5[tree] <- quantile(T$hn,0.05)
        df$quant10[tree] <- quantile(T$hn,0.10)
        df$quant15[tree] <- quantile(T$hn,0.15)
        df$quant20[tree] <- quantile(T$hn,0.20)
        df$quant25[tree] <- quantile(T$hn,0.25)
        df$quant30[tree] <- quantile(T$hn,0.30)
        df$quant35[tree] <- quantile(T$hn,0.35)
        df$quant40[tree] <- quantile(T$hn,0.40)
        df$quant45[tree] <- quantile(T$hn,0.45)
        df$quant50[tree] <- quantile(T$hn,0.50)
        df$quant55[tree] <- quantile(T$hn,0.55)
        df$quant60[tree] <- quantile(T$hn,0.60)
        df$quant65[tree] <- quantile(T$hn,0.65)
        df$quant70[tree] <- quantile(T$hn,0.70)
        df$quant75[tree] <- quantile(T$hn,0.75)
        df$quant80[tree] <- quantile(T$hn,0.80)
        df$quant85[tree] <- quantile(T$hn,0.85)
        df$quant90[tree] <- quantile(T$hn,0.90)
        df$quant95[tree] <- quantile(T$hn,0.95)
        df$druh[tree] <- 0
    }
#načtení kompletní tabulky
    data <- read.table ("Komplet.csv", header = TRUE, sep = ";")

```



*#zobecněný lineární model a zobrazení*

```
full.model <- glm(druh ~ ., data = data, family = binomial)  
summary(full.model)
```

*#kroková regrese a zobrazení*

```
step <- step(full.model, direction = "both")  
summary(step)
```

## 5 Výsledky

### 5.1 Detekování a dovození výšek jednotlivých stromů v prostředí ArcGIS 10.6.

Detekováním jednotlivých stromů z bodového mračna pomocí výše popsané metody v prostředí ArcGIS 10.6. bylo vyřezáno 2 947 jednotlivých stromů na čtyřech plochách dohromady (tabulka č. 2). Na první ploše borovice lesní označené jako BO1 bylo detekováno 729 stromů, na druhé BO2 bylo detekováno 1 016 stromů. Na první ploše smrku ztepilého označené jako SM1 bylo detekováno 776 stromů a na druhé ploše SM2 425 stromů. Následně byly z každé plochy vyřazeny chybně detekované stromy, u kterých byl nedostatečný počet bodů. U BO1 to bylo 66, u BO2 114, u SM1 61 a u SM2 28 chybně detekovaných stromů. Počty stromů vstupujících do matematicko-statistické analýzy jsou 663 stromů pro BO1, 902 stromů pro BO2, 715 stromů pro SM1 a 397 stromů pro SM2. Celkový počet je tedy 2 677 stromů, pro které budou vypočítány dané metriky. Reálný počet stromů není znám, z tohoto důvodu nelze určit procentuální úspěšnost detekování.

Tabulka 2 - počty detekovaných stromů

Plocha	Počet detekovaných stromů	Počet vyřazených stromů	Počet stromů vstupujících do analýzy
BO1	729	66	663
BO2	1016	114	902
SM1	776	61	715
SM2	425	28	397
Celkem	2946	269	2677

Z odvozených výšek stromů z CHM a následného výpočtu byly získány průměrné výšky pro jednotlivé plochy. Nejmenší průměrná výška 22,58 m byla na první ploše BO1, následně 23,28 m na ploše BO2, dále 32,71 m na ploše SM1 a 37,24 m na ploše SM2. Průměrné výšky pro jednotlivé plochy s jejich variabilitou můžeme vidět v tabulce č. 3.

Tabulka 3 – odvozené průměrné výšky s variabilitou

Plocha	Průměrná výška	Průměrná absolutní odchylka	Směrodatná odchylka	Rozptyl
BO1	22,58	2,1488	0,9977	0,9924
BO2	23,28	1,6724	0,9811	0,9893
SM1	32,71	1,8352	0,975	0,956
SM2	37,24	2,0272	0,996	1,0689

## 5.2 Vybrané metriky pomocí zobecněného lineárního modelu pro určení rozdílů mezi borovicí lesní a smrkem ztepilým

Prvotní výsledky zobecněného lineárního modelu můžeme vidět v tabulce č. 4.

Tabulka 4 – výsledky zobecněného lineárního modelu

	Estimate	Std. Error	z value	Pr(> z )
<b>(Intercept)</b>	75.559951	20.107596	3.758	0.000171
<b>mean</b>	4.973319	0.639612	7.776	7.51e-15
<b>kquad</b>	-10.401366	1.253579	-8.297	< 2e-16
<b>kub</b>	0.311989	0.052325	5.962	2.48e-09
<b>pd</b>	1.856456	0.307260	6.042	1.52e-09
<b>sd</b>	-17.452105	14.407232	-1.211	0.225764
<b>var</b>	-7.709926	4.481854	-1.720	0.085386
<b>ks</b>	-13.384814	2.940729	-4.552	5.33e-06
<b>quant5</b>	-0.025894	0.044834	-0.578	0.563566
<b>quant10</b>	-0.035811	0.029557	-1.212	0.225669
<b>quant15</b>	0.020979	0.022982	0.913	0.361336
<b>quant20</b>	0.043777	0.015943	2.746	0.006037
<b>quant25</b>	0.048431	0.014527	3.334	0.000857
<b>quant30</b>	0.042431	0.013809	3.073	0.002121
<b>quant35</b>	0.020053	0.014417	1.391	0.164256
<b>quant40</b>	0.012480	0.017600	0.709	0.478276
<b>quant45</b>	-0.011111	0.022231	-0.500	0.617228
<b>quant50</b>	-0.011432	0.031803	-0.359	0.719247
<b>quant55</b>	-0.068015	0.030295	-2.245	0.024761
<b>quant60</b>	0.171461	0.096954	1.768	0.076981
<b>quant65</b>	0.121338	0.147835	0.821	0.411781
<b>quant70</b>	-0.009538	0.160081	-0.060	0.952491
<b>quant75</b>	-0.249253	0.175785	-1.418	0.156207
<b>quant80</b>	0.382491	0.183685	2.082	0.037313
<b>quant85</b>	-0.006016	0.181409	-0.033	0.973543
<b>quant90</b>	0.100253	0.159685	0.628	0.530126
<b>quant95</b>	0.712685	0.139080	5.124	2.99e-07

V prvním sloupci "Estimate" vidíme odhadovanou hodnotu, kterou daná metrika nabývá. Druhý sloupec "Std. Error" udává standardní chybu předchozí odhadované hodnoty. Následně "value" vyjadřuje hodnotu testovací statistiky. Nejdůležitější údaj vyjadřuje poslední sloupec "Pr (>|z|)", který nám říká, jaká je pravděpodobnost chyby prvního druhu. Se zmenšováním této chyby roste významnost dané metriky pro určení rozdílu mezi druhy. Proto, aby byla metrika považována za signifikantní, musí být tato hodnota alespoň 0,05.

Zde bychom mohli analyzovat dané signifikantní metriky, nicméně jak bylo dříve popsáno, bylo provedeno zpřesnění zobecněného lineárního modelu pomocí krokové regrese. Celkově v tomto modelu proběhl výpočet desetkrát a výsledné optimální řešení modelu můžeme vidět v tabulce č. 5.

Tabulka 5 – výsledky zobecněného lineárního modelu s krokovou regresí

	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt; z )</b>
<b>(Intercept)</b>	82.495233	18.432431	4.476	7.62e-06
<b>Aritmetický průměr</b>	4.918346	0.518602	9.484	< 2e-16
<b>Kvadratický průměr</b>	-10.557600	1.194271	-8.840	< 2e-16
<b>Kubický průměr</b>	0.328257	0.048232	6.806	1.00e-11
<b>Absolutní odchylka</b>	1.763843	0.272984	6.461	1.04e-10
<b>Směrodatná odchylka</b>	-20.276919	13.592966	-1.492	0.135772
<b>Rozptyl</b>	-8.365843	4.487973	-1.864	0.062314
<b>Koeficient šikmosti</b>	-14.040478	2.695116	-5.210	1.89e-07
<b>Kvantil 10</b>	-0.034523	0.022177	-1.557	0.119543
<b>Kvantil 20</b>	0.043957	0.013113	3.352	0.000801
<b>Kvantil 25</b>	0.043486	0.009523	4.566	4.97e-06
<b>Kvantil 30</b>	0.038713	0.008153	4.748	2.05e-06
<b>Kvantil 35</b>	0.021027	0.010084	2.085	0.037054
<b>Kvantil 55</b>	-0.081297	0.022907	-3.549	0.000387
<b>Kvantil 60</b>	0.254377	0.063855	3.984	6.79e-05
<b>Kvantil 75</b>	-0.185219	0.123555	-1.499	0.133854
<b>Kvantil 80</b>	0.404740	0.130346	3.105	0.001902
<b>Kvantil 95</b>	0.785440	0.097018	8.096	5.69e-16

Jak můžeme vidět, bylo stanoveno 17 signifikantních metrik z původního počtu 26 metrik. Z toho je 11 s hodnotou  $p < 0,001$ . Jsou to v první řadě průměry (Aritmetický, Kvadratický a Kubický), dále Průměrná odchylka, Koeficient šikmosti a Kvantily (20, 25, 30, 55, 60 a 95). Následuje Kvantil80 s hodnotou  $p < 0,01$  a Kvantil35 s hodnotou  $p < 0,05$ . Zbylé čtyři metriky (Směrodatná odchylka, Rozptyl, Kvantil10

a Kvantil 75) mají hodnotu  $p > 0,05$ , nicméně byly zahrnuty do výpočtu, protože nesnižují přesnost výsledného řešení.

Zhodnocení celkové přesnosti bylo provedeno pomocí chybové matice v prostředí R studio. Výsledné hodnoty pro klasifikaci druhů jsou zobrazeny v tabulce č. 6.

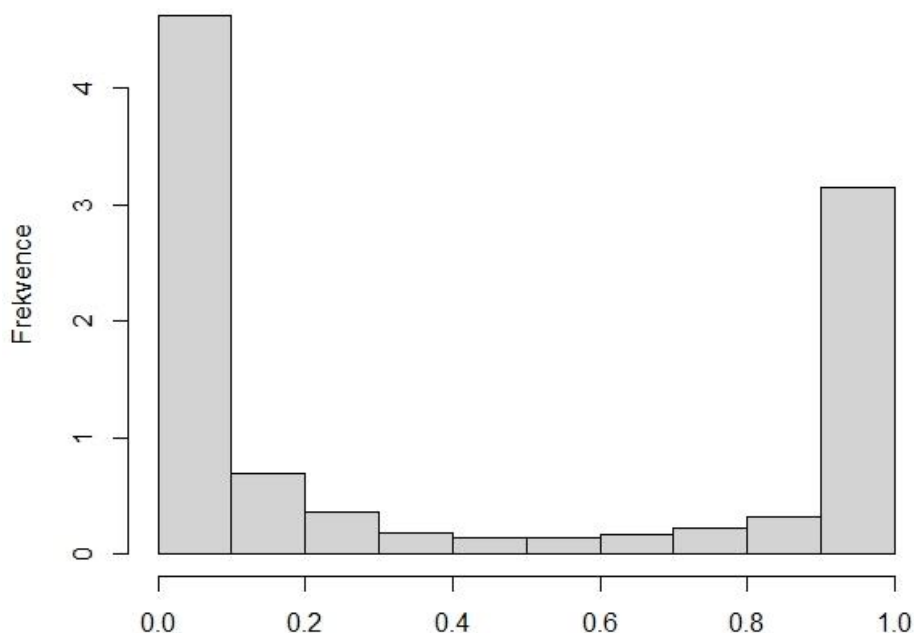
Tabulka 6 – výsledky klasifikace

Dřevina	Celkový počet	Správně určeno	Chybně určeno	Úspěšnost v %
BO	1565	1496	69	95,59
SM	1112	999	113	89,84
Celkem	2677	2495	182	92,71

Pro dřevinu BO z celkového počtu 1565 stromů bylo určeno správně 1496 stromů a chybně 69 stromů. Úspěšnost klasifikace pro tuto dřevinu je 95,59 %. U dřeviny SM s celkovým počtem 1112 stromů bylo správně zařazeno 999 stromů a chybně 113 stromů, s úspěšností klasifikace 89,84 %. Celková úspěšnost této metody pro klasifikaci dřevin byla tedy 92,71 %.

Grafické rozdělení úspěšnosti klasifikace můžeme vidět na Histogramu predikcí (graf č. 1). Požadovanými výsledky je predikce blízko hodnotě 0 nebo 1, tím indikují první nebo druhou dřevinu. Pokud je hodnota predikce mezi těmito hodnotami, udává to, na kolik procent se predikce přiklání k jedné, nebo druhé dřevině. V našem případě toto grafické rozdělení predikcí odpovídá vypočítané přesnosti klasifikace, kdy většina predikcí se nachází právě na hodnotě 1 respektive 0 a směrem ke středu hodnota významně klesá.

### Histogram predikcí



Graf 1 – Histogram rozdělení predikcí, prostředí R studio

### 5.3 Rozdíly hodnot jednotlivých metrik mezi smrkem ztepilým a borovicí lesní

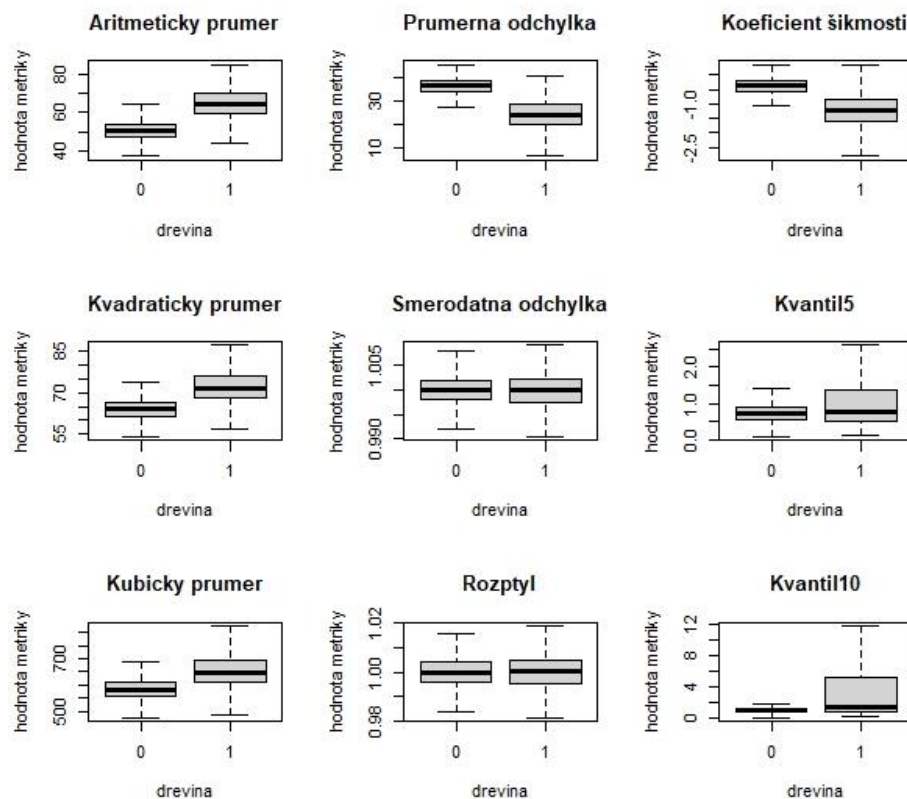
První metrikou byl Aritmetický průměr, jejíž rozdíl mezi dřevinami můžeme vidět na prvním krabicovém grafu (graf č. 2), který ukazuje větší hodnotu pro smrk ztepilý, tzn. že u tohoto druhu bylo v průměru naměřeno více bodů. Podobné rozdíly můžeme vidět u následujících dvou grafů pro Kvadratický a Kubický průměr, jelikož se jedná o metriky počítané z Aritmetického průměru. S rostoucí váhou pro vzdálenější body se od Aritmetického průměru úměrně zmenšuje rozdíl mezi hodnotami. U Kvadratického průměru se rozdíl hodnot mezi dřevinami mírně zmenšil a tento trend pokračuje u Kubického průměru, kdy je rozdíl ještě menší.

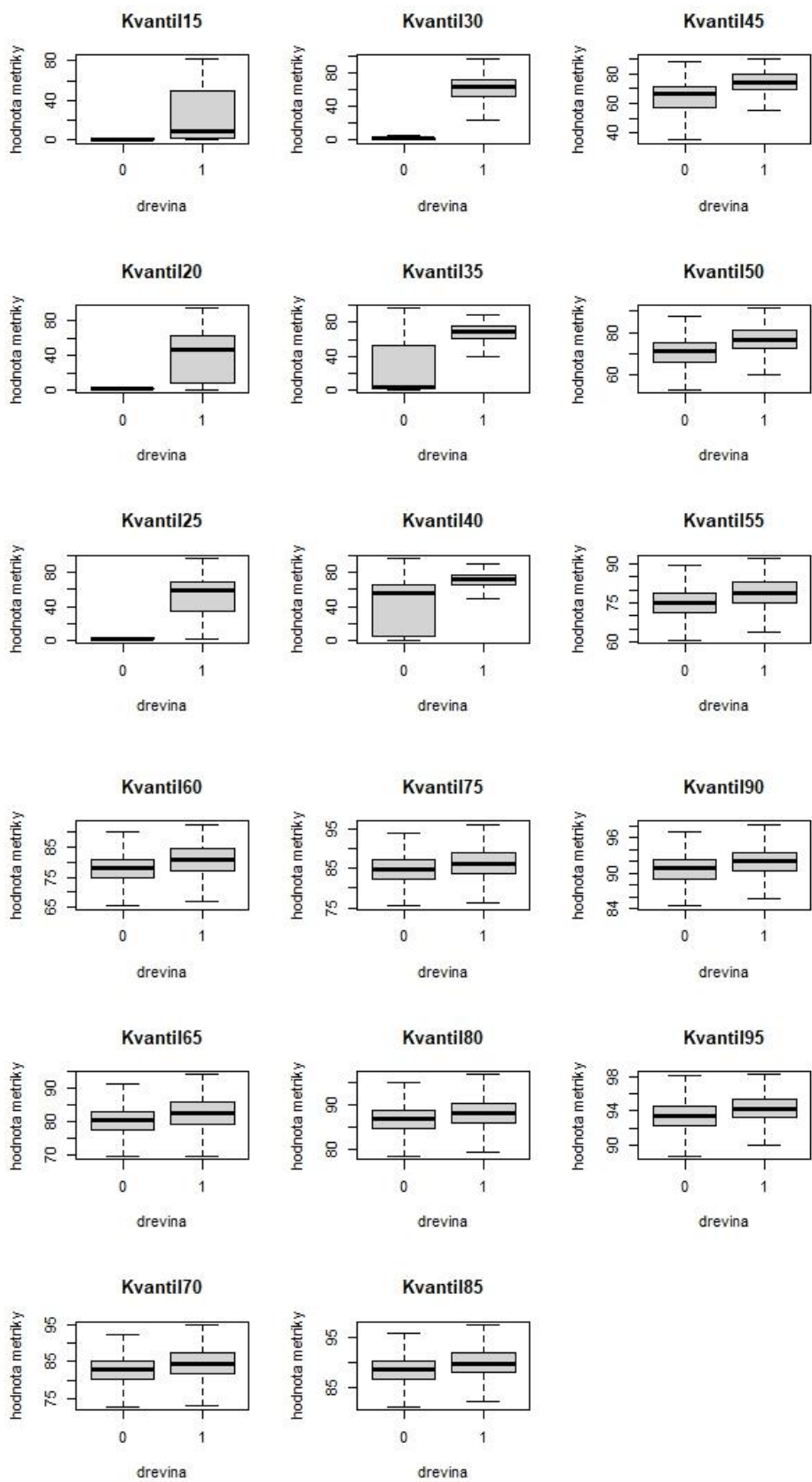
Další počítanou metrikou byla Průměrná absolutní odchylka, která udává větší hodnotu pro borovici lesní, u této dřeviny tedy bylo více bodů vzdáleno od průměru. Následuje Směrodatná odchylka a Rozptyl, hodnoty u těchto metrik nevykazují téměř

žádné rozdíly. Na grafu můžeme pozorovat, že medián je u obou metrik pro obě dřeviny stejný. Znamená to tedy, že body obou druhů dřevin byly rozptýleny stejně.

Následující metrikou byl Koeficient šikmosti, který vyšel v obou případech záporný, tzn. většina hodnot se nachází vpravo od průměru. Pro druh borovice bylo rozdělení blíže normálnímu rozdělení, u smrku bylo více levostranně asymetrické.

Posledními metrikami bylo rozdělení výšek kvantilů po 5 %. U tohoto rozdělení ve všech jeho částech vyšla vyšší hodnota bodů pro smrk. U prvního kvantilu (Kvantil5) je medián u obou dřevin téměř na stejné hodnotě. V následujících sekcích ale začíná významně převyšovat u smrku. Zároveň od Kvantil10 po Kvantil35 je u borovice hodnota metriky blízká nule, tudíž je zde jen malý počet bodů. Hodnota metriky se u borovice začíná zvyšovat od Kvantil40 a postupně roste až po poslední Kvantil95, nicméně je zde vždy nějaký rozdíl mezi druhy, s tím že u smrku je vždy více bodů, což souhlasí s výsledky metrik pro průměry.





Graf 2 – Krabicové grafy pro vybrané metricky v prostředí R studio



## 6 Diskuse

Získávání dat z bodových mračen za pomoci technologie UAV pro lesnické účely v posledních letech výrazně pokročilo. Dalo by se říct, že v tomto sektoru překonalo dosavadní způsoby sběru dat, jak časově, tak i po ekonomické stránce. Diskutabilní může být přesnost, záleží pro jaké účely a jaké veličiny měříme. Nicméně tyto nové metody se zatím do široké praxe moc nedostávají. Z mého pohledu to může být způsobeno konzervativním přístupem v tomto odvětví, zmiňovanou přesností, ale zejména potřebou odborného přístupu při zpracování datových výstupů. Dalším faktorem můžou být velké vstupní náklady pro technologii LiDAR. Co se týká fotogrammetrie, bylo provedeno několik studií, např. (Koh & Wich, 2012), kdy i s drony vybavenými kamerami pro spotřebitele byli schopni získat data pro měření korun stromů, odhad počtu stromů a výšky stromů. Výrazně tak snížili náklady a urychlili výkon práce ve srovnání s tradičními metodami.

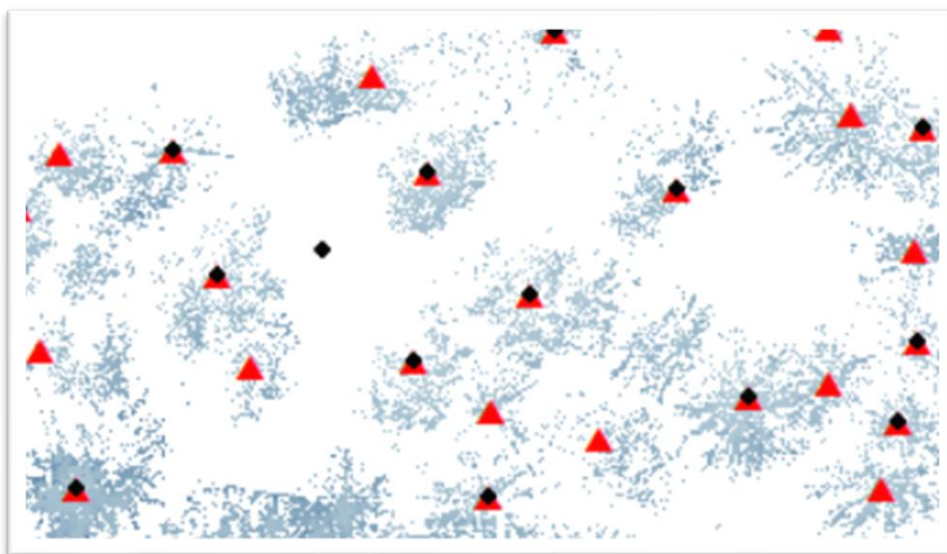
Stále je ale potřeba hledat, zjednodušovat a automatizovat nové postupy zpracování dat, což se také poměrně velkou rychlostí děje, díky práci vědeckých týmů z celého světa.

### 6.1 Zpracování bodových mračen pro detekování a následné vyřiznutí jednotlivých stromů

Jak již bylo zmíněno, je snaha o automatizování postupů zpracování dat, to je i případ této práce. Nicméně právě detekování jednotlivých stromů z bodových mračen pořízených pomocí UAV v prostředí ArcGIS je značně náročné automatizovat. Protože každý lesní porost je unikátní, je zde potřeba objektivního zhodnocení a občasné manuální zpracování. Prvním příkladem může být právě chybné detekování vrcholových bodů stromů pomocí funkce Canopy peaks, kdy je pak potřeba některé body ručně odmazávat.

Pro předcházení tohoto problému je nutné zadání správné kombinace atributů, které do této funkce vstupují. Tato kombinace se může lišit u každé plochy a zejména pro různé dřeviny. Na obrázku č. 18 můžeme vidět značný rozdíl přesnosti detekce jednotlivých vrcholových bodů stromů funkcí Canopy peaks, při změně vstupních

atributů. Černé tečky znázorňují detekované vrcholové body při použití přednastavených atributů, červené trojúhelníky značí vrcholové body při použití upravených atributů.



Obrázek 17 – rozdíl v detekci vrcholových bodů při použití funkce Canopy peaks, ArcGIS 10.6.

Dalším problémem jsou takzvaná zdvojená lokální maxima, kdy je špatnou detekcí, nebo chybami v modelu CHM, u jednoho stromu detekováno více vrcholových bodů. Toto řeší použití takzvaných filtrů, v našem případě to byla funkce Focal statistics.

Jak říká (Panagiotidis et al., 2017), zde je potřeba dát si pozor a zadat parametry tak, aby byly odfiltrována zdvojená maxima, ale zároveň nedošlo k odmazání menších stromů. Pro automatizování těchto kroků bychom mohli přeskočit ruční úpravy a pracovat s přednastavenými parametry, to nám ale ovlivní přesnost delineaace stromů při použití následného algoritmu IWS. Proto je velmi důležité si nejprve uvědomit, za jakým účelem zpracování dat provádíme, a zda se spokojíme s tím, že ne všechny stromy budou detekovány.

Na druhou stranu zpracování v prostředí ArcGIS 10.6. je jeden z méně složitých postupů pro zpracování takovýchto dat. Je potřeba jen základních znalostí tohoto softwaru a vědomost, kdy a jaké funkce použít pro daný krok. Není zde potřeba znalostí jako programování, například v softwaru R studio, nebo navíc ještě principy strojového učení jako je tomu u složitějších metod klasifikace (např. Random forest).

## 6.2 Význam počítaných matematicko-statistických metrik

Účelem této práce bylo zjistit, zda se dají klasifikovat vybrané druhy stromů na základě rozdílů mezi vybranými metrikami. První metrikou, která byla určena jako signifikantní, byl aritmetický průměr, který ukazuje větší hodnotu pro smrk ztepilý, tedy bylo u něho naměřeno více bodů. To může být pravděpodobně způsobeno nižším nasazením větví a celkově větším objemem koruny než u borovice lesní. Stejným způsobem můžeme vyvodit výsledky u průměru Kubického a Kvadratického, ty ukazují podobné, avšak zmenšující se, rozdíly mezi počtem bodů. Tyto metriky dávají větší váhu bodům, které jsou vzdálenější od průměru, tzn. že u borovice lesní bylo více vzdálenějších bodů než u smrku ztepilého, to může být způsobeno větší šířkou koruny, nebo jednotlivých větví. Z těchto metrik je tedy zřejmé, že smrk ztepilý má delší, užší korunu oproti borovici lesní, která má kratší, ale širší korunu, což odpovídá jejich habitům.

Další metrikou, která vykazovala značný rozdíl mezi druhy s hodnotou, je Průměrná odchylka. Ta byla větší u borovice lesní, tzn. že u této dřeviny bylo více bodů rozmístěno ve větší vzdálenosti od průměru bodů. To opět může souviset s délkou koruny/větví, která je u této dřeviny zpravidla širší.

Následuje Koeficient šikmosti, který pro obě dřeviny vyšel záporný. Většina hodnot se tedy nachází vpravo od průměru, tzn. rozdělení je asymetrické.

Poslední z kategorie metrik, které byly určeny jako signifikantní, jsou některé kvantily. Prvním z nich je Kvantil20, u kterého byl značný rozdíl v hodnotách, protože je zde u borovice lesní malý počet bodů. Může to souviset s větším zakmeněním na plochách borovice, kdy z tohoto důvodu bylo změřeno malé množství bodů v dolních částech stromů. Nicméně tento rozdíl spíše odpovídá habitu smrku ztepilého, kdy je u něho typické nižší nasazení korun (zhruba 20 % výšky). Stejný rozdíl hodnot ukazují i Kvantil25 a Kvantil30. Dalším signifikantním Kvantilem je Kvantil55, zde se rozdíl hodnot značně snížil, protože borovice lesní už v tomto úseku má také koruny. Pořád ale platí větší hodnota pro smrk ztepilý. Do kategorie signifikantních náleží ještě Kvantil60 a Kvantil95, nicméně rozdíl mezi hodnotami dřevin už se od Kvantilu55 téměř nemění.

Zde je potřeba ještě zmínit Kvantil35 a Kvantil40, které sice měly hodnotu  $p > 0,001$ , ale zároveň jsou důležité z důvodu pravděpodobného nasazení koruny u borovice lesní. V těchto dvou úsecích se počet bodů u borovice lesní zvedl z hodnoty blízké nule a téměř vyrovnal rozdíl hodnot smrku ztepilého.

### 6.3 Srovnání přesnosti klasifikace druhů jinými metodami

Ve studii (Nevalainen et al., 2017) zkoumali účinnost fotogrammetrie a hyperspektrální dat získaných z UAV pro detekci jednotlivých stromů a klasifikaci druhů v boreálních lesích. Přesnost klasifikace jednotlivých stromů se pohybovala mezi 40 % a 95 % v závislosti na charakteristice oblasti. Klasifikace bez normalizace dat přinesla mírně horší výsledky ve srovnání s výsledky dosaženými s normalizovanými daty. Bez normalizovaných dat poskytl Random Forest nejlepší výsledek s celkovou přesností 93 %.

Využití algoritmu Random forest pro rozlišení druhů také zkoumali ve studii (Michez et al., 2016). Přesnost klasifikace se lišila podle rozsahu analýzy, což naznačuje, že celková přesnost se zlepšovala se zmenšující se velikostí analyzovaných objektů. Výsledky klasifikace ukázali přesnost od 79,5 % do 84,1 % pro rozlišení druhového složení.

Metoda konvolučních neuronových sítí byla použita ve studii (Morales et al., 2018) pro segmentování specifického druhu palmy z leteckých snímků. Celková přesnost této metody, co se týče segmentace druhu palmy, byla 98 %. Kromě dospělých a oddělených jedinců byla tato metoda schopna detekovat také palmy v mladém věku, nebo palmy částečně zakryté jinými druhy vegetace.

Ve studii (Gini et al., 2014) zkoumali rozdíly klasifikace druhů stromů mezi kontrolovanými a nekontrolovanými algoritmy. Výsledky klasifikace pomocí nekontrolovaného algoritmu ISODATA nebyly dostačující s přesností 50 %. Druhá metoda byla klasifikace pomocí algoritmu Maximum Likelihood, tedy kontrolovaným algoritmem. Zde bylo dosaženo největší přesnosti u dvou druhů stromů (Pajasan 76 % a Akát 66 %), nicméně ostatní druhy byly většinou nesprávně zařazeny do ostatních kategorií a tím snížily celkovou přesnost.

Hierarchickou klasifikací obrazu se zabývali ve studii (Bilgilioglu et al., 2017). Došli k závěru, že hierarchický klasifikační přístup je vhodný pro aplikování klasifikace rostlinných společenstev. Výsledky, s celkovou přesností mozaiky obrazů 83 až 88 %, jsou použitelné pro různé klasifikační úrovně (tj. vegetační/nevegetační, na úrovni struktur nebo druhů).

Při srovnání přesností klasifikace druhů různými metodami vidíme, že metoda popsána v této práci vykazovala velice dobrou přesnost (92,71 %). Nicméně je důležité zmínit, že všechny výše popsané metody byly prováděny na datech získaných ze snímků zpracovaných fotogrammetrií. Zatímco metoda v této práci byla provedena s daty z LiDARu, které jsou přesnější. Dalším aspektem může být také počet druhů vstupujících do klasifikace. V této práci byly klasifikovány pouze dva druhy, které navíc vykazují poměrně značné rozdíly na úrovni habitu. Proto s klasifikací více druhů, nebo méně rozdílných druhů, může přesnost této metody klesat. Tyto problémy bude nutné překonat v budoucím výzkumu.

## 7 Závěr

V této práci bylo testováno, zda je možné rozlišit a určit druhy lesních dřevin na základě matematicko-statistických metrik, vypočítaných z bodových mračen. Pro tento účel byly změřeny pomocí technologie UAV dvě plochy borovice lesní a dvě plochy smrku ztepilého a z nich získána bodová mračna. Výše popsanou metodou byl vytvořen DTM a z něho odvozen CHM. Dále došlo k detekování jednotlivých stromů a delineaci bodových mračen pomocí algoritmu IWS. Pro selektovanou bodovou mračna bylo v prostředí R studio vypočítáno 27 metrik, ze kterých bylo pomocí zobecněného lineárního modelu s krokovou regresí vybráno 17 metrik. Tyto metriky byly označeny jako signifikantní pro určení rozdílů v bodových mračnách jednotlivých dřevin. V tomto případě určení těchto 17 metrik a jejich rozdílů mezi borovicí lesní a smrkem ztepilým ukázalo, že je možné rozlišit tyto dva druhy na základě statistické analýzy s přesností 92,71 %. Celkové zhodnocení ukázalo, že pro zachycení statistických rozdílů při klasifikaci druhů z dat získaných z ALS jsou nejdůležitější metriky, které přímo, nebo nepřímo souvisejí s tvarem a nasazením koruny.

Otázkou zůstává, jak by tato metoda byla úspěšná v rozlišování jiných dřevin, nebo dřevin s podobným habitem, například smrku ztepilého s jedlí bělokorou. Pro potenciální využití této metody by bylo potřebné znát rozdíly na statistické úrovni mezi všemi dřevinami, pro které by byla prováděna klasifikace. Hlavní výhodou je, že je to supervizovaná metoda a může tedy probíhat autonomně, jen s drobnými lidskými zásahy. Tím vzniká potenciál pro její automatizaci, například v podobě sofistikovaného softwaru.

## 8 Literatura

- Allaire, J. (2011). The R User Conference 2011. *R Studio: Intergrated Development Environment For R*, 770, 14.
- Amini, S., Homayouni, S., Safari, A., & Darvishsefat, A. A. (2018). Object-based classification of hyperspectral data using Random Forest algorithm. *Geo-Spatial Information Science*, 21(2), 127–138. <https://doi.org/10.1080/10095020.2017.1399674>
- Amit, Y., Geman, D., & Wilder, K. (1997). Joint induction of shape features and tree classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(11), 1300–1305. <https://doi.org/10.1109/34.632990>
- Andreola, R. (2009). Support Vector Machines Na Support Vector Machines Na. *지능정보연구*, 16(1), 130.
- Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, 25(2), 197–227. <https://doi.org/10.1007/s11749-016-0481-7>
- Bilgilioglu, B. B., Ozturk, O., Sariturk, B., & Seker, D. Z. (2017). Object-based classification of unmanned aerial vehicle (UAV) imagery for forest fires monitoring. *38th Asian Conference on Remote Sensing - Space Applications: Touching Human Lives, ACRS 2017, 2017-Octob(6)*, 661–672.
- Clemons, J. (2007). Sift : Scale Invariant Feature Transform. *Transform*, 2, 91–110. Retrieved from [http://www.eecs.umich.edu/~silvio/teaching/EECS598/lectures/lecture10\\_1.pdf](http://www.eecs.umich.edu/~silvio/teaching/EECS598/lectures/lecture10_1.pdf)
- Dietterich, T. G. (2000). Ensemble Methods in Machine Learning. *International Workshop on Multiple Classifier Systems*, 1–15. Retrieved from <http://www.cs.orst.edu/~tgd>
- Edson, C., & Wing, M. G. (2011). *Airborne light detection and ranging (LiDAR) for individual tree stem location, height, and biomass measurements. Remote Sensing (Vol. 3)*. <https://doi.org/10.3390/rs3112494>
- Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., ... Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, 186, 64–87. <https://doi.org/10.1016/j.rse.2016.08.013>

- Gini, R., Passoni, D., Pinto, L., & Sona, G. (2014). Use of unmanned aerial systems for multispectral survey and tree classification: A test in a park area of northern Italy. *European Journal of Remote Sensing*, 47(1), 251–269. <https://doi.org/10.5721/EuJRS20144716>
- Goodbody, T. R. H., Coops, N. C., & White, J. C. (2019). Digital Aerial Photogrammetry for Updating Area-Based Forest Inventories: A Review of Opportunities, Challenges, and Future Directions. *Current Forestry Reports*, 55–75. <https://doi.org/10.1007/s40725-019-00087-2>
- Guimarães, N., Pádua, L., Marques, P., Silva, N., Peres, E., & Sousa, J. J. (2020). Forestry remote sensing from unmanned aerial vehicles: A review focusing on the data, processing and potentialities. *Remote Sensing*, 12(6). <https://doi.org/10.3390/rs12061046>
- Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8), 832–844. <https://doi.org/10.1109/34.709601>
- Iglhaut, J., Cabo, C., Puliti, S., Piermattei, L., O'Connor, J., & Rosette, J. (2019). Structure from Motion Photogrammetry in Forestry: a Review. *Current Forestry Reports*, 5(3), 155–168. <https://doi.org/10.1007/s40725-019-00094-3>
- James, M. R., & Robson, S. (2012). Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. *Journal of Geophysical Research: Earth Surface*, 117(3), 1–17. <https://doi.org/10.1029/2011JF002289>
- Jin, Z., Shang, J., Zhu, Q., Ling, C., Xie, W., & Qiang, B. (2020). RFRSF: Employee Turnover Prediction Based on Random Forests and Survival Analysis. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12343 LNCS, 503–515. [https://doi.org/10.1007/978-3-030-62008-0\\_35](https://doi.org/10.1007/978-3-030-62008-0_35)
- Koh, L. P., & Wich, S. A. (2012). Dawn of drone ecology: Low-cost autonomous aerial vehicles for conservation. *Tropical Conservation Science*, 5(2), 121–132. <https://doi.org/10.1177/194008291200500202>
- Kuželka, K., Slavík, M., & Surový, P. (2020). Very high density point clouds from UAV laser scanning for automatic tree stem detection and direct diameter measurement.



- Remote Sensing*, 12(8). <https://doi.org/10.3390/RS12081236>
- Kuželka, K., & Surový, P. (2018). *Statistika v R. Zpracování dat závěrečných prací pro lesnické obory* (první). Praha: Česká zemědělská univerzita v Praze.
- Lepš, J., & Šmilauer, P. (2016). *Biostatistika*. (R. Folprechtová, Ed.) (1.). České Budějovice: Jihočeská univerzita v Českých Budějovicích.
- Málek, J., & Roháček, A. (n.d.). VLS\_Brevir\_stromy (p. 24). Vojenské lesy a statky ČR, s.p. Retrieved from [www.vls.cz](http://www.vls.cz)
- Michez, A., Piégay, H., Lisein, J., Claessens, H., & Lejeune, P. (2016). Classification of riparian forest species and health condition using multi-temporal and hyperspatial imagery from unmanned aerial system. *Environmental Monitoring and Assessment*, 188(3), 1–19. <https://doi.org/10.1007/s10661-015-4996-2>
- Mohan, M., Silva, C. A., Klauberg, C., Jat, P., Catts, G., Cardil, A., ... Dia, M. (2017). Individual tree detection from unmanned aerial vehicle (UAV) derived canopy height model in an open canopy mixed conifer forest. *Forests*. <https://doi.org/10.3390/f8090340>
- Morales, G., Kemper, G., Sevillano, G., Arteaga, D., Ortega, I., & Telles, J. (2018). Automatic segmentation of *Mauritia flexuosa* in unmanned aerial vehicle (UAV) imagery using deep learning. *Forests*, 9(12). <https://doi.org/10.3390/f9120736>
- Musil, I., & Hamerník, J. (2003). *Lesnická dendrologie 1. Jehličnaté dřeviny. Lesnická dendrologie 1. Jehličnaté dřeviny*.
- Nevalainen, O., Honkavaara, E., Tuominen, S., Viljanen, N., Hakala, T., Yu, X., ... Tommaselli, A. M. G. (2017). Individual tree detection and classification with UAV-Based photogrammetric point clouds and hyperspectral imaging. *Remote Sensing*. <https://doi.org/10.3390/rs9030185>
- O'Shea, K., & Nash, R. (2015). An Introduction to Convolutional Neural Networks, 1–11. Retrieved from <http://arxiv.org/abs/1511.08458>
- Pádua, L., Vanko, J., Hruška, J., Adão, T., Sousa, J. J., Peres, E., & Morais, R. (2017). UAS, sensors, and data processing in agroforestry: a review towards practical applications. *International Journal of Remote Sensing*, 38(8–10), 2349–2391. <https://doi.org/10.1080/01431161.2017.1297548>
- Panagiotidis, D., Abdollahnejad, A., Surový, P., & Chiteculo, V. (2017). Determining tree height and crown diameter from high-resolution UAV imagery. *International*

- Journal of Remote Sensing*, 38(8–10), 2392–2410.  
<https://doi.org/10.1080/01431161.2016.1264028>
- Pekár, S., & Brabec, M. (2016). *Modern Analysis of Biological Data: Generalized Linear Models in R*. Masarykova univerzita. Retrieved from <https://books.google.cz/books?id=vHfkDwAAQBAJ>
- RIEGL Laser Measurement Systems. (2020). RiCOPTER for Multi-Purpose Applications.
- Smreček, R. (2012). Utilization of ALS Data for Forestry Purposes. 2012) *GI\_Forum*, (May), 365–376. Retrieved from [http://www.gispoint.de/fileadmin/user\\_upload/paper\\_gis\\_open/537521048.pdf](http://www.gispoint.de/fileadmin/user_upload/paper_gis_open/537521048.pdf)
- Surový, P., & Kuželka, K. (2019). Acquisition of forest attributes for decision support at the forest enterprise level using remote-sensing techniques-a review. *Forests*, 10(3). <https://doi.org/10.3390/f10030273>
- Surový, P., & Kuželka, K. a kol. (2019). *Aplikace dálkového průzkumu Země v lesnictví* (první). Praha: Česká zemědělská univerzita v Praze.
- Tang, L., & Shao, G. (2015). Drone remote sensing for forestry research and practices. *Journal of Forestry Research*, 26(4), 791–797. <https://doi.org/10.1007/s11676-015-0088-y>
- Tiberiu Paul Banu, Gheorghe Florian Borlea, & Constantin Banu. (2016). The Use of Drones in Forestry. *Journal of Environmental Science and Engineering B*, 5(11), 557–562. <https://doi.org/10.17265/2162-5263/2016.11.007>
- Waseem, A., Khyber, A., Ahmad, N., Abdur, S., Abid, R., Akbar, M., & Khan, A. (2016). K-Means and ISODATA Clustering Algorithms for Landcover Classification Using Remote Sensing. *Sindh University Research Journal (Science Series)*, 48(2), 315–318. Retrieved from <https://www.researchgate.net/publication/303971825>
- White, J. C., Coops, N. C., Wulder, M. A., Vastaranta, M., Hilker, T., & Tompalski, P. (2016). Remote Sensing Technologies for Enhancing Forest Inventories: A Review. *Canadian Journal of Remote Sensing*, 42(5), 619–641. <https://doi.org/10.1080/07038992.2016.1207484>
- ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE, Inc., 2020. *ArcMap | Documentation* [online]. Dostupné z: <https://desktop.arcgis.com/en/arcmap/>
- ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE, Inc., Clayton CRAWFORD, Khalid DURI a Jie CHANG, 2020. *3D Sample Tools* [online] [vid. 2020-

03 17]. Dostupné z: <https://www.arcgis.com/home/item.html?id=fe221371b77940749ff96e90f2de3d10>

SEZNAM.CZ, 2020. *Mapy.cz* [online] [vid. 2020-03-20]. Dostupné z: <https://mapy.cz/zakladni?x=13.6414619&y=50.6844609&z=16&base=ophoto>

RStudio Team (2021). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA URL <http://www.rstudio.com/>.