

ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE
FAKULTA ŽIVOTNÍHO PROSTŘEDÍ
KATEDRA PROSTOROVÝCH VĚD

**Kvalita a hustota výškových dat
družicového laserového skenování na
území ČR**
BAKALÁŘSKÁ PRÁCE

Vedoucí práce: Ing. Vojtěch Barták, Ph. D.

Bakalant: Ondřej Krameš

2024

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Ondřej Krameš

Geografické informační systémy a dálkový průzkum Země v životním prostředí

Název práce

Kvalita a hustota výškových dat družicové laserové altimetrie na území ČR

Název anglicky

Quality and density of the elevation data from space-born laser altimetry in the Czech Republic

Cíle práce

Posoudit kvalitu a hustotu dat družicové laserové altimetrie na území ČR pro účely digitálního modelování terénu.

Metodika

Práce se zaměří na posouzení kvality a hustoty dat družicové laserové altimetrie na území ČR pro účely digitálního modelování terénu. Budou použita a porovnána data z misí GEDI a ICESat-2, se zaměřením na posouzení různých přístupů k filtraci chybných měření, a to jak na základě údajů o kvalitě jednotlivých segmentů dat, která jsou publikována společně s daty, tak na základě externích proměnných (např. svažitost). Přesnost filtrovaných modelů bude posouzena pomocí referenčních výškových dat z leteckého laserového skenování.

Doporučený rozsah práce

30 – 50 stran

Klíčová slova

DTM, IceSat, GEDI, Space-born laser scanning

Doporučené zdroje informací

- Adam, M., Urbazaev, M., Dubois, C., & Schmullius, C. (2020). Accuracy assessment of GEDI terrain elevation and canopy height estimates in European temperate forests: Influence of environmental and acquisition parameters. *Remote Sensing*, 12(23), 3948.
- Moudrý, V., Gdulová, K., Gábor, L., Šárovcová, E., Barták, V., Leroy, F., ... & Prošek, J. (2022). Effects of environmental conditions on ICESat-2 terrain and canopy heights retrievals in Central European mountains. *Remote Sensing of Environment*, 279, 113112.
- Urbazaev, M., Hess, L. L., Hancock, S., Sato, L. Y., Ometto, J. P., Thiel, C., ... & Schmullius, C. (2022). Assessment of terrain elevation estimates from ICESat-2 and GEDI spaceborne LiDAR missions across different land cover and forest types. *Science of Remote Sensing*, 6, 100067.

Předběžný termín obhajoby

2023/24 LS – FŽP

Vedoucí práce

Ing. Vojtěch Barták, Ph.D.

Garantující pracoviště

Katedra prostorových věd

Elektronicky schváleno dne 18. 3. 2024

prof. Ing. Petra Šímová, Ph.D.

Vedoucí katedry

Elektronicky schváleno dne 19. 3. 2024

prof. RNDr. Michael Komárek, Ph.D.

Děkan

V Praze dne 28. 03. 2024

Čestné prohlášení

Prohlašuji, že jsem bakalářskou práci na téma: Kvalita a hustota výškových dat družicového laserového skenování na území ČR vypracoval samostatně a citoval jsem všechny informační zdroje, které jsem v práci použil a které jsem rovněž uvedl na konci práce v seznamu použitých zdrojů.

Jsem si vědom, že na moji bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů, ve znění pozdějších předpisů, především ustanovení § 35 odst. 3 tohoto zákona, tj. o užití tohoto díla.

Jsem si vědom, že odevzdáním bakalářské práce souhlasím s jejím zveřejněním podle zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů, ve znění pozdějších předpisů, a to i bez ohledu na výsledek její obhajoby.

Svým podpisem rovněž prohlašuji, že elektronická verze práce je totožná s verzí tištěnou a že s údaji uvedenými v práci bylo nakládáno v souvislosti s GDPR.

V Praze dne

.....

Poděkování

Rád bych tímto poděkoval svému vedoucímu Ing. Vojtěchu Bartákovi, Ph.D. za jeho odborné vedení, cenné rady a velkou trpělivost.

Dále bych rád poděkoval své přítelkyni za pomoc a rady při tvorbě této práce.
V neposlední řadě bych rád poděkoval své rodině a přátelům za veškerou oporu během studia.

Abstrakt

Digitální výškové modely jsou v dnešní době stále více využívány v široké škále aplikací a oborů, mezi které primárně patří lesnictví, archeologie, stavebnictví a enviromentální vědy. Pro většinu těchto oborů je důležitá přesnost a rozsah modelů. Mise ICESat – 2 a GEDI využívající vesmírné LiDARové technologie a již od roku 2019 nepřetržitě skenují povrch Země. Data z těchto misí by mohla být použita k tvorbě digitálních výškových modelů téměř kdekoliv na planetě Zemi. Vytváření těchto modelů ovšem představuje značnou řadu problémů. Cílem této práce je zhodnotit a navrhnout filtrace těchto dat za účelem vytvoření digitálního modelu terénu na území České republiky. Pro každou misi byl navrhnout vhodný filtr z poskytovaných atributů. Dále byla využita metoda filtrování, v níž byly vypočteny mediány výšek nejbližších okolních bodů a podle velikosti odchylky od měřené výšky byly záznamy vyřazeny. Výsledky byly porovnány s digitálním modelem reliéfu 5. generace (DMR5g). Mise GEDI měla původní RMSE 2595,79 m. Po filtraci se RMSE snížilo na 8,52 m s tím, že počet bodů klesl o 76 %. U mise ICESat – 2 bylo původní RMSE 1,47 m a po filtraci 1,23 m. Počet bodů poklesl o 19 %. Data GEDI byla výrazně méně přesnější, navzdory tomu obsahovala petinásobný počet záznamů oproti datům ICESat – 2.

Klíčová slova: DTM, IceSat, GEDI, vesmírné laserové skenování

Abstract

Digital elevation models are increasingly being used in a wide range of applications and fields today, including forestry, archaeology, construction, and environmental sciences. For most of these fields, the accuracy and range of the models are important. The ICESat-2 and GEDI missions, utilizing spaceborne LiDAR technologies, have been continuously scanning the Earth's surface since 2019. Data from these missions could be used to create digital elevation models almost anywhere on the planet Earth. However, creating these models poses a considerable number of challenges. The goal of this work is to evaluate and propose data filtering for the purpose of creating a digital terrain model in the Czech Republic. A suitable filter was designed for each mission from the provided attributes. Furthermore, a filtering method was used, in which the medians of the heights of the nearest surrounding points were calculated, and records were discarded based on the deviation size from the measured height. The results were compared with the 5th generation digital terrain model (DMR5g). The GEDI mission had an original RMSE (Root Mean Square Error) of 2595.79 m. After filtering, the RMSE decreased to 8.52 m, with the number of points dropping by 76%. For the ICESat-2 mission, the original RMSE was 1.47 m, and after filtering, it was 1.23 m. The number of points dropped by 19%. GEDI data was significantly less accurate, despite having five times the number of records compared to ICESat-2 data.

Key words: DEM, IceSat, GEDI, Space-born laser scanning

1. Úvod.....	1
2. Cíle práce	2
3. Literární rešerše.....	3
3.1 Laser.....	3
3.2 LiDAR.....	3
3.3 Global Ecosystem Dynamics Investigation (GEDI)	4
3.4 ICESat - 2.....	7
4. Charakteristika studovaného území	9
5. Metodika	10
5.1 Získání dat.....	10
5.2 HDF5 formát	10
5.3 Zvolené atributy GEDI.....	10
5.3.1 Quality flag.....	10
5.3.2 Number of modes.....	11
5.3.3 Waveform fidelity flag	12
5.3.4 Elevation lowestmode	12
5.3.5 Lon lowestmode a Lat lowestmode	13
5.4 Zvolené atributy ICESat – 2.....	13
5.4.1 Cloud Flag.....	13
5.4.2 Number of terrain photons	13
5.4.3 Spacecraft Orientation.....	13
5.4.4 Beam	14
5.4.4.1 Strong Beam a Weak Beam.....	14
5.5 Hodnocení přesnosti.....	15
5.6 Filtrace podle atributů	16
5.7 Filtrace pomocí mediánu výšek nejbližších bodů	16
5.8 Počet bodů na čtverec.....	17
6. Výsledky	18
6.1 GEDI	18
6.2 ICESat – 2	22
7. Diskuse.....	27
8. Závěr	28
9. Přehled literatury a použitých zdrojů	29

9.1 Odborné publikace	29
9.2 Internetové zdroje.....	34
10. Přílohy	35

1. Úvod

Ve světě neustále rostoucí digitalizace a technologického pokroku se kvalita a hustota výškových dat stává klíčovou pro řadu aplikací, od plánování infrastruktury a ochrany před přírodními katastrofami až po environmentální monitorování a řízení vodních zdrojů. Družicová laserová altimetrie, jakožto jedna z nejpřesnějších metod získávání výškových dat na velké plochy, hraje zásadní roli v tvorbě digitálního modelu terénu (DMT) (Yang et al., 2014). Přestože jsou výsledky těchto technologií obecně považovány za přesné, rozdíly v kvalitě a hustotě dat mezi jednotlivými misemi mohou mít značný dopad na spolehlivost a užitečnost vytvořených modelů terénu, zejména na území s komplexním terénem (Adam et al., 2020; A. Liu et al., 2021).

V České republice, zemi s různorodým reliéfem od nížin po horské oblasti, je zajištění vysoké přesnosti a hustoty výškových dat pro účely DMT zásadní. Přestože existují studie zaměřené na hodnocení a porovnání výškových dat získaných z družicové laserové altimetrie (Liu et al., 2021; Narin & Gullu, 2023; Urbazaev et al., 2021), žádná z nich se detailně nevěnuje specifikům družicové laserové altimetrie na území ČR. Mise GEDI a ICESat – 2 přináší nové možnosti, jak získávat výšková data, ale současně představují výzvy spojené s efektivní filtrací chybných měření a interpretací dat (Li et al., 2021; Liu et al., 2021).

2. Cíle práce

Tato práce si klade za cíl posoudit kvalitu a hustotu dat z družicové laserové altimetrie na území Čech a jejich využití pro digitální modelování terénu. Zaměří se na analýzu dat z misí GEDI a ICESat-2 s cílem identifikovat a porovnat různé přístupy k filtraci chybných měření. Analýza bude zahrnovat hodnocení údajů o kvalitě poskytovaných spolu s daty, jakož i využití externích proměnných, například svazitosti, pro optimalizaci filtrací. Klíčovým výstupem práce je porovnání přesnosti filtrovaných modelů s referenčními výškovými daty získanými leteckým laserovým skenováním, což poskytne důležité poznatky pro zlepšení využití družicové laserové altimetrie v praxi.

3. Literární rešerše

3.1 Laser

Laser, akronym pro Light Amplification by Stimulated Emission of Radiation, funguje stimulací atomů nebo molekul k vydávání světla na specifických vlnových délkách a jeho následném zesílení. Tento proces vytváří vysoce soustředěný, koherentní paprsek světla s unikátními vlastnostmi, které umožňují jeho využití v mnoha oborech.

Lasery generují světlo procesem stimulované emise. Když jsou atomy v laserovém médiu, (které může být plyn, kapalina nebo pevná látka) excitovány do vyšší energetické úrovně, mohou při návratu do nižší energetické úrovně emitovat fotony (částice světla). Pokud tyto fotony stimulují další vzrušené atomy k emisi více fotonů stejné energie a fáze, dochází k řetězové reakci, která zesiluje světlo do koherentního paprsku (Majumdar & Manna, 2011). Laser se typicky sestává ze tří hlavních komponentů: zdroje energie, který vzrušuje atomy v laserovém médiu, samotného laserového média, kde dochází k zesílení světla, a optického rezonátoru, který umožňuje světlu oscilovat a zvyšovat intenzitu. Optický rezonátor se obvykle skládá ze dvou zrcadel, z nichž je jedno částečně průhledné, aby umožnilo únik části světla v podobě laserového paprsku (Haley, 2023).

Existuje mnoho typů laserů, které se kategorizují na základě stavu laserového média (pevná látka, plyn, kapalina nebo polovodič). Každý typ má unikátní vlastnosti a aplikace. Například CO₂ lasery se široce používají v průmyslovém řezání a gravírování, zatímco Nd:YAG lasery najdou uplatnění v zpracování materiálů a lékařství. Polovodičové lasery, například diodové lasery, se používají v komunikačních zařízeních a čtečkách čárových kódů (Gagliano et al., 1969).

3.2 LiDAR

LiDAR, což je zkratka pro Light Detection and Ranging, je metoda sběru dat dálkového průzkumu, která používá laserové světlo k měření vzdáleností. Principem LiDARu je vysílání laserových paprsků k objektu a následné zachycení odraženého světla pomocí senzoru. Doba, kterou laserový paprsek stráví cestou k objektu a zpět, se používá k výpočtu vzdálenosti mezi senzorem a objektem. Tato metoda umožňuje

vytvářet přesné 3D modely povrchů, objektů a prostředí (Collis, 1969; Spinhirne, 1993). Výhodou LiDARu je možnost získávání dat ve dne i v noci, pokud podmínky nejsou ovlivněny silnou mlhou, dýmem či extrémní vlhkostí, což zahrnuje situace jako je déšť, sníh či mraky, které by mohly blokovat laserové paprsky mezi zařízením a cílem. Klíčovým atributem LiDARu je také schopnost laseru proniknout skrz mezery v korunách stromů, což umožňuje nejenom mapování struktury a výšky korun, ale i měření výšky terénu pod lesním porostem (G. Liu et al., 2018).

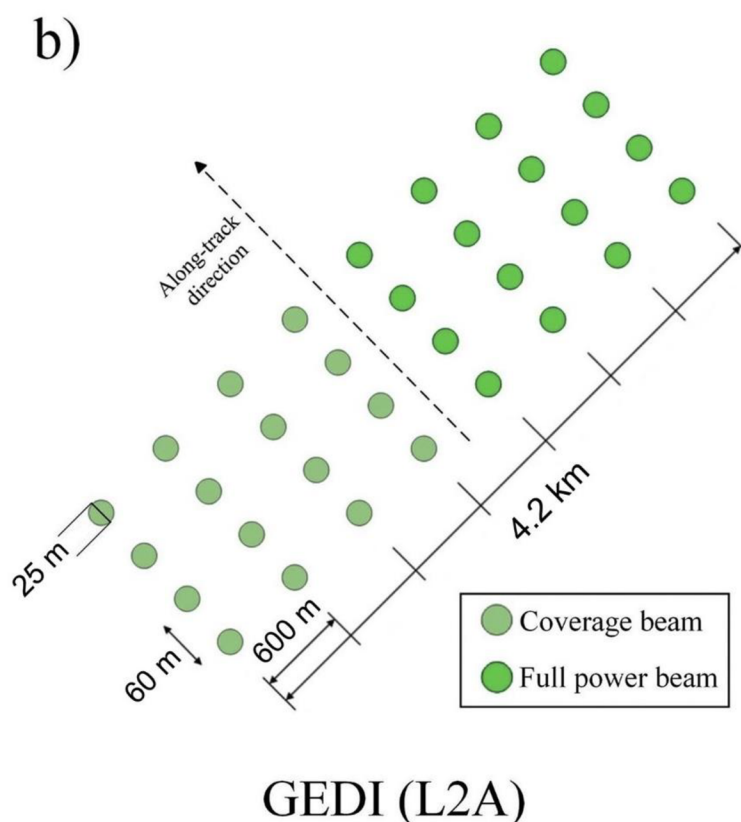
LiDAR se využívá v řadě aplikací, od geodézie, archeologie, lesnictví přes stavebnictví až po autonomní vozidla. V geografických informačních systémech (GIS) se LiDAR používá k vytváření podrobných digitálních modelů terénu (DMT) (Lohani & Ghosh, 2017).

3.3 Global Ecosystem Dynamics Investigation (GEDI)

Global Ecosystem Dynamics Investigation (GEDI) je vesmírný LiDARový systém vypuštěn agenturou NASA v roce 2018 a instalovaný na Mezinárodní vesmírné stanici (ISS) (Dubayah et al., 2020). Od března roku 2019 GEDI zaznamenává kruhové plochy s průměrem 25 metrů na zemském povrchu v rozpětí mezi 51,6° severní a jižní šířky. Geolokace paprsků GEDI je odvozena od vlastní měřicí jednotky GEDI, globálních polohových systémů (GPS) a senzorů hvězdné sledovací soustavy na palubě ISS (Dubayah et al., 2020; Luthcke et al., 2019). Nízká oběžná dráha, velikost a tvar ISS však vedou k větším mechanickým vibracím a možným změnám v orientaci a nadmořské výšce než u tradičních družic dálkového průzkumu Země (Dou et al., 2014). V důsledku toho se očekávala přesnost horizontální polohy kruhových ploch po kalibraci na 10 metrů (Dubayah et al., 2020). U první verze produktů GEDI (v1), vydané před kalibrací za letu, dosáhla průměrná chyba horizontální geolokace hodnoty 23,8 m. Po kalibračním procesu zohledňujícím chyby geolokace byla v dubnu 2021 vydána druhá verze dat (v2) s odhadovanou chybou polohování 10,2 m a konečnou cílovou přesností 8 m (Dubayah et al., 2021).

Přístroj GEDI se skládá ze tří laserů, které vysílají 14 ns dlouhé pulzy blízkého infračerveného laseru s vysokou frekvencí (242 Hz). Jeden laser je rozdělen na dva pokrývací (coverage) paprsky, zatímco další dva lasery vytvářejí dva paprsky

plného výkonu (full power). Každý paprsek je každým druhým výstřelem odkloněn pomocí jednotky pro změnu směru paprsku, což vede k osmi paralelním kruhovým plochám. Paprsky jsou při dopadu na Zemi vzdáleny mezi sebou 600 m v řadě kolmo ve směru letu. Průměry paprsků jsou 25 m a další řady jsou vzdáleny 60 m od středu paprsku, jak je znázorněno na obrázku č. 1. (Dubayah et al., 2020). Vlny odražené a od Země jsou zaznamenány a zpracovány tak, aby byly poskytovány pro každý kruhový paprsek (Hofton et al., 2019).



Obrázek č. 1 – Schéma paprsků z GEDI při dopadu na zem. Schéma není v měřítku a slouží pouze jako ilustrace. Zdroj: (A. Liu et al., 2021)

Zpracované vlny jsou poskytovány v mnoha produktech. Například produkt L1B obsahuje geolokalizované vlny, L2B vegetační strukturu nebo L4B hustotu biomasy (Dubayah et al., 2020). V produktu L2A jsou odvozeny nadmořské výšky terénu, vrcholy korun stromů a metriky RH (relativní výšky) z geolokalizovaných vln (produkt L1B). RH odpovídají kumulativní energii vlny od spodku (0 %) po vrchol (100 %) výšek v 1 % intervalu (RH0 až RH100) (Hofton et al., 2019).

Nedávné studie hodnotily kvalitu dat GEDI dle odhadu nadmořské výšky terénu, výšky korun stromů a hustotu nadzemní biomasy (aboveground biomass) v porovnáním s daty z leteckého LiDARového skenování (Duncanson et al., 2020, 2022; Lang et al., 2022). Bylo zjištěno, že GEDI poskytuje přesné údaje o nadmořské výšce terénu a výšce korun stromů, ačkoliv s chybami dosahujícími až několika metrů (Adam et al., 2020; A. Liu et al., 2021; Urbazaev et al., 2021). Významná část chyb byla připsána nízké horizontální přesnosti (Adam et al., 2020; Duncanson et al., 2020; Lang et al., 2022; Potapov et al., 2021; Schleich et al., 2023; Urbazaev et al., 2021). Ve studii Quiros et al., (2021) dosáhli RMSE 6,13 m porovnáním s leteckým LiDARovým skenování na 11 lokalitách ve Španělsku. Na základě simulací dat GEDI Milenkovič et al. (2017) ukázali, že chyby odhadu nadzemní biomasy se zvyšují s rostoucí chybou geolokace. Chyba geolokace má větší vliv v heterogenních lesích a na členitých plochách než v homogenních lesích (Milenkovič et al., 2017; Schleich et al., 2023). Sklon a hustota vegetačního pokryvu prokázaly vliv na přesnost GEDI (Adam et al., 2020; A. Liu et al., 2021; Quiros et al., 2021), ale v těchto studiích nebyla testována souvislost s dopadem chyby geolokace. Nicméně, vzhledem k tomu, že chyby geolokace v souřadnicích GEDI na svažitém terénu mohou vést k větším rozdílům v nadmořské výšce mezi skutečnými a poskytnutými souřadnicemi, než na rovinném terénu je zřejmé, že svažité terény budou chybami geolokace ovlivněny více než rovinné.

Zlepšení geolokace je důležité a vyžaduje specifické přístupy. Nejrozšířenější metodou zlepšení geolokace je použití dat z leteckého LiDARového skenování k simulaci vln podobných GEDI v okolí původní polohy kruhové plochy (Blair & Hofton, 1999; Hancock et al., 2019; Lang et al., 2022). Metoda postupuje po sobě jdoucími shluky kruhových ploch a korektní geolokace je přiřazena tam, kde je maximalizována korelace mezi simulovanými a skutečnými vlnami GEDI (Hancock et al., 2019; Lang et al., 2022). Metoda prezentovaná v Hancock et al., 2019 byla primárně a úspěšně použita k vylepšení georeferencování GEDI. Nicméně, vyžaduje data z leteckého laserového skenování a je proto omezena na oblasti, kde se taková data nachází, ideálně v době blízké se získáním dat GEDI.

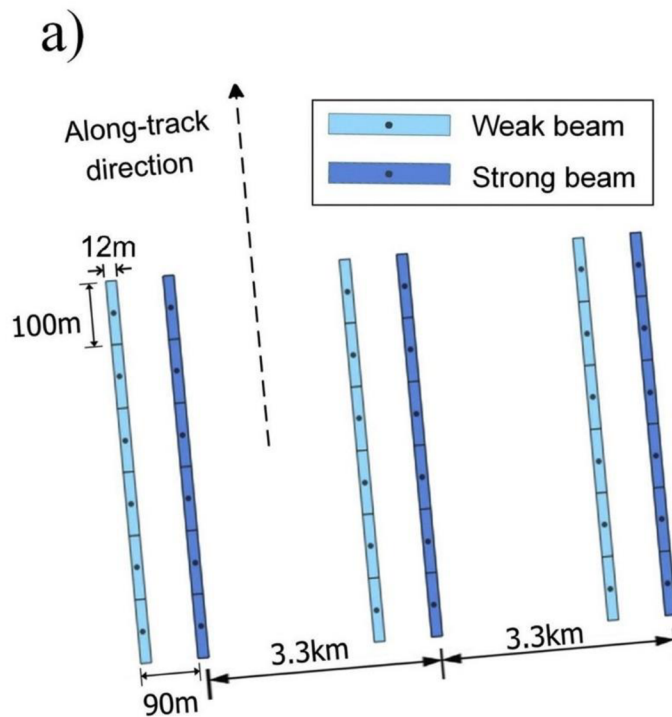
3.4 ICESat - 2

Ice, Cloud, and Land Elevation Satellite 2 (ICESat-2) je novou generací družicového LiDARového výškoměru. Stejně jako jeho předchůdce ICESat-1, má za cíl kvantifikovat příspěvek polárních ledovců ke vzestupu hladiny moře a dopad změny klimatu (Markus et al., 2017). ICESat-1 byl vypuštěn v roce 2003 a vyřazen z provozu v roce 2009 (Abdalati et al., 2010). ICESat-1 je první vesmírný LiDARový přístroj navržený pro globální pozorování. Data z něj byla úspěšně použita k odhadu vertikální struktury terénu lesa, včetně topografie terénu a výšky korun (Harding & Carabajal, 2005; Lefsky et al., 2005). ICESat-2, spuštěný v roce 2019, v této misi pokračuje s vylepšeným LiDARovým zařízením Advanced Topographic Laser Altimeter System - ATLAS. (Markus et al., 2017)

ICESat-2 se vznáší na polární dráze s úhlem sklonu 92 stupňů, která shromažďuje data od 88° severní šířky do 88° jižní šířky s přesnou opakovací dobou 91 dnů (Markus et al., 2017). Laserový přístroj ATLAS o vlnové délce 532 nm (zelená barva viditelného spektra) vysílá jeden puls, který se rozdělí do tří párů paprsků, přičemž každý pár má jeden silný paprsek, který je čtyřikrát silnější než odpovídající slabý paprsek. Vzájemná vzdálenost tří párů paprsků je přibližně 3,3 km na příčném směru dráhy a vzdálenost paprsků uvnitř jednoho páru je přibližně 90 m, jak je znázorněno na obrázku č. 2, přičemž přesná geometrie je řízena orientací družice (Neumann et al., 2019). Každý bod z těchto paprsků má v ideálním případě pokryv o průměru přibližně 17 metrů (Neuenschwander & Pitts, 2019).

ICESat-2 pracuje s frekvencí pulsů 10 kHz, což znamená, že laser vystřeluje 10 000krát za sekundu. Hustota bodů závisí na zeměpisné šířce dráhy. Například v polárních oblastech s vysokými zeměpisnými šířkami jsou pozemní dráhy úzce rozmístěny, což vede k vyšší hustotě měřených bodů, zatímco v oblastech středních zeměpisných šířek a na rovníku jsou výsledky rozmístěny s větším odstupem, což snižuje hustotu měřených bodů (Neumann et al., 2019).

Stejně jako mise GEDI, ICESat – 2 svá data poskytuje ve více produktech. Například pevninský led (ATL06), mořský led (ATL07), terén/vegetace (ATL08), atmosféra (ATL09), oceány (ATL12) a vnitrozemské vody (ATL13) (Markus et al., 2017).



ICESat-2 (ATL08)

Obrázek č. 2 – Schéma paprsků z družice ICESat – 2 při dopadu na zem. Silný (strong) a slabý (weak) paprsek se mění podle toho v jaké orientaci je družice. Schéma není v měřítku a slouží pouze jako ilustrace. Zdroj: (Liu et al., 2021)

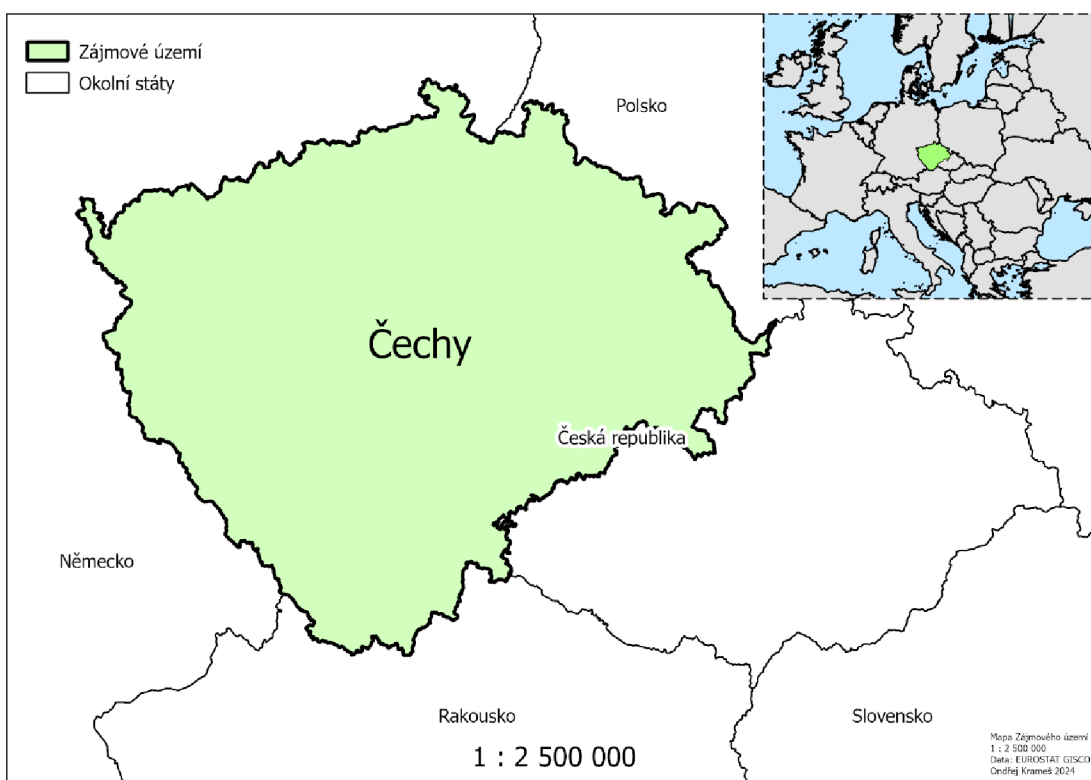
I když hlavním vědeckým cílem mise ICESat-2 je pozorování polárních oblastí Země (Markus et al., 2017), data se dají využít i v celé řadě jiných praktických a vědeckých využití. Různorodost využití dokládají úspěšné studie (Kropáček et al., 2014; Kurtz and Markus, 2012; Lefsky, 2010), které vzešly z původní mise ICESat.

V předchozích výzkumech Neuenschwander et al. (2019) představili kvantitativní hodnocení výšky terénu produktu ICESat-2 ATL-08 ve srovnání s daty z leteckého laserového skenování (Neuenschwander & Magruder, 2019). Počáteční výsledek ověření nejlepšího odhadu výšky terénu ve Finsku ukazuje, že výška terénu má průměrnou absolutní chybu 0,5 m, zatímco RMSE je 0,82 m.

4. Charakteristika studovaného území

Česká republika je stát, který se nachází ve střední Evropě, jak je znázorněno na obrázku č. 3. Čechy jsou region na západě území České republiky. Jejich rozloha je 52 065 km² což představuje 66% rozlohy celé ČR. Hranice jsou definovány na severozápadě pohořím Krušné hory, na severovýchodě Krkonošemi, na jihovýchodě pak hranicí s Moravou a na jihozápadě pohořím Šumavou. Západní hranice je tvořena státní hranicí se Spolkovou republikou Německo.

Čechy nabízejí širokou škálu topografických prvků, od nížinných oblastí v České tabuli přes pahorkatiny až po horské řetězce na hranicích. Nejvyšší bod Čech, Sněžka v Krkonoších, dosahuje výšky 1 603 m n. m. Střední nadmořská výška je 465 m n. m. a 66 % její výšky se nachází do 500 m n. m.



Obrázek č. 3 – Mapa zobrazující zájmové území v měřítku 1 : 2 500 000. Zdroj: v tiráži obrázku

5. Metodika

5.1 Získání dat

Jak mise GEDI, tak ICESat – 2 jsou pod záštitou vesmírné agentury NASA. Většinu všech volně dostupných dat NASA nabízí na svém datovém portálu search.earthdata.nasa.gov. Na tomto portálu jsou k nalezení i data z obou misí. Pro GEDI byla zvolena datová sada L2A, jelikož obsahuje výšky terénu. Pro družici ICESat – 2 byl zvolen datový produkt s názvem ATL08, protože obsahuje také výšky terénu. Jelikož bylo potřeba co nejvíce záznamů dat s největší možnou hustotou, byla použita veškerá dostupná data. Konkrétněji data GEDI do 18. 4. 2019 do 1. 11. 2023 a ICESat 14. 10.2018 až 15.10.2023. Portál umožňuje nechat si předem oříznout data na zvolené území. Obě datové sady byly tedy oříznuty podle hranic Čech.

5.2 HDF5 formát

Obě družice využívají k distribuci dat formát Hierarchical Data Format (HDF). Toto je otevřený formát používaný k efektivnímu ukládání a organizování objemných dat. Je podporovaný v mnoha programovacích jazycích a je primárně využíván ve vědeckých aplikacích. K jeho otevření a prozkoumání byl použit program HDF5View od neziskové organizace HDFgroup. Data bylo potřeba vyexportovat do formátu Comma-separated values (CSV), aby se s nimi dalo lépe pracovat v pozdější fázi. Na to byl použit statistický jazyk R, ve kterém byl využit balíček rhdf5. Díky tomu byly vyexportovány kompletní záznamy pro každý paprsek (beam) do CSV tabulek. Skript na vyexportování dat může být nalezen v příloze číslo 1.

5.3 Zvolené atributy GEDI

5.3.1 Quality flag

Tento údaj nám indikuje, zda se vlna dá považovat za použitelnou pro další analýzy. Zahrnuje soubor podmínek, které dohromady naznačují celkovou platnost záznamu pro měření povrchové struktury. Podmínky pro označení vlny jako kvalitní (1) jsou tyto:

```
stale_return_flag = 0
rx_rangewindow_clip_front = 0
rx_rangewindow_clip_back = 0
rx_clipflag = 0
rx_rxwindow_limit = 0
rx_rxwindow_exist ≠ 0
rx_rxwindow_clip_front = 0
rx_rxwindow_clip_back = 0
rx_lbinwaveform_flag = 0
rx_pulseflag ≠ 0
```

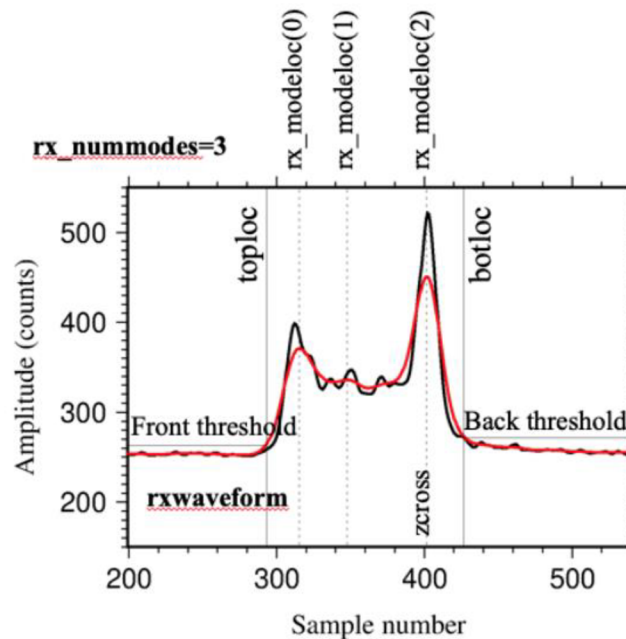
Stale return flag indikuje, zda algoritmus, hledající v 10km okně, nezachytil vracející se vlnu v reálném čase. *Rx rangewindow clip front* indikuje, jestli se vlna nachází na vrcholu 10km okna a *Rx rangewindow clip back* jestli se nachází na spodu hledajícího okna. *Rx clipflag* označuje, zda vlna překročila hranici maximální amplitudy vlny. *Rx rxwindow limit*, *rx rxwindow exist*, *rx rxwindow clip front* a *rx rx window clip back* obecně označují různé problémy, které mohly nastat během letu nebo příjmu vlny. *Rx pulseflag* indikuje, že žádná vlna nebyla zachycena.

V případě, že jedna z podmínek není splněna, je vlna označena jako nekvalitní hodnotou 0 (Hofton et al., 2019).

5.3.2 Number of modes

Tato proměnná udává počet módů, které byly detekovány v daném profilu LiDARu. Profil LiDARu je vertikální sloupec dat, který reprezentuje odrazy laserového paprsku od Země. Počet módů v profilu LiDARu může být použit k odhadu struktury vegetace. Například profil LiDARu s jedním módem pravděpodobně reprezentuje oblast s nízkou vegetací, zatímco profil LiDAR s více

mody pravděpodobně reprezentuje oblast s vysokou vegetací. Obrázek č. 4 reprezentuje příklad pohybu amplitudy u jednoho paprsku.



Obrázek č. 4 – Příklad GEDI rxwaveform před (černě) a po (červeně) konvoluci s Gaussovou funkcí šířky 6,5 ns. Algoritmus prohledává pouze část rxwaveform mezi signal_start a signal_end pro odražené mody. Nastavení prahu definuje „nejvyšší“ (nejvíce vlevo, toploc) a „nejnižší“ (nejvíce vpravo, botloc). V tomto příkladu byly algoritmem detekovány tři odlišné mody (rx_nummodes=3). (Hofton et al., 2019)

5.3.3 Waveform fidelity flag

Waveform fidelity flag (rx_assess_flag) je soubor více atributů dohromady. Tento indikátor ukazuje, zda mohl nastat nějaký problém při příjmu vlny. Pokud se hodnota rovná 0, je bez problému. V opačném případě je konkrétní chyba označena příslušným číslem.

5.3.4 Elevation lowestmode

Údaj "elevation lowestmode" představuje nejnižší výšku nad referenčním elipsoidem detekovanou laserovým paprskem v rámci jednoho laserového záběru.

Při analýze laserových odrazů se může laser odrazit od různých objektů, jako je třeba listů v koruně stromu nebo samotný zemský povrch. Údaj "elevation lowestmode" konkrétně udává výšku nejnižšího odrazu zachyceného v rámci daného laserového záběru. To bývá obvykle zemský povrch, ale v hustě zalesněných oblastech může jít i o spodní část vegetace.

5.3.5 Lon lowestmode a Lat lowestmode

Atributy „Lon lowestmode“ a „Lat lowestmode“ obsahují souřadnice v souřadnicovém systému WGS84 (World Geodetic System 1984). Lon lowestmode obsahuje zeměpisnou délku a Lat lowestmode zeměpisnou šířku.

5.4 Zvolené atributy ICESat – 2

5.4.1 Cloud Flag

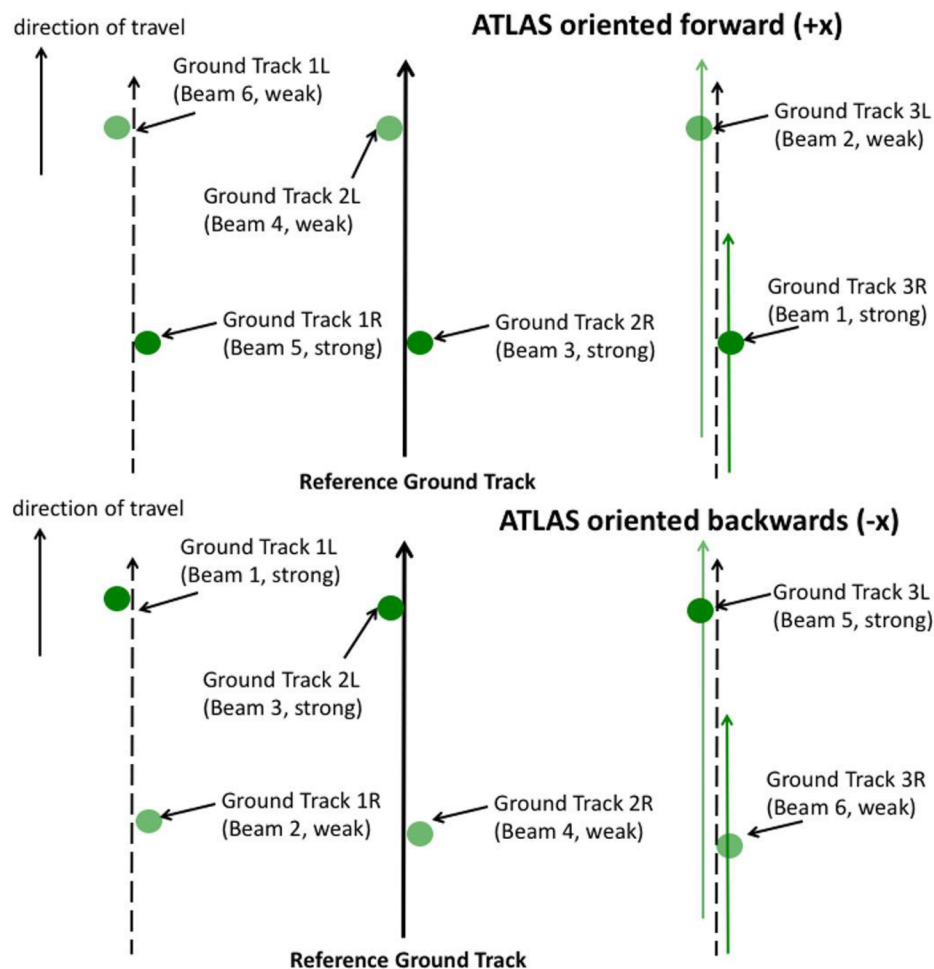
Cloud Flag (cloud_flag_atm) udává s jakou jistotou se v každém 25 Hz atmosférickém profilu vyskytují oblaka či vrstvy aerosolu. Pokud je hodnota větší než 0 může obsahovat oblaka či vrstvy aerosolu. Nabývá hodnot od 0 do 10.

5.4.2 Number of terrain photons

Tento atribut (n_te_phot) udává, kolik fotonů bylo v segmentu klasifikováno jako terénních.

5.4.3 Spacecraft Orientation

Parametr (sc_orient) nabývá hodnot 0, 1 či 2. Určuje orientaci družice mezi vpřed, vzad a přechod mezi orientacemi. Družice ICESat – 2 letí vpřed když weak beam vede strong beam. Vzad letí, pokud strong beam vede weak beam, což je znázorněno na obrázku č. 5. Když je hodnota 0, tedy přechází mezi vpřed a vzad, dochází k potencionálnímu zhoršení kvality. Orientace družice se mění zhruba dvakrát ročně z důvodu maximalizování osvětlení solárních panelů (Neumann et al., 2022).



Obrázek č. 5 – Obrázek ukazuje, jak se mění pozice Strong a Weak paprsku při orientaci vzad (+x) a vpřed (-x). (Neumann et al., 2019)

5.4.4 Beam

Atribut označující, o jaký paprsek se jedná. Paprsky se označují 1l, 1r, 2l, 2r, 3l a 3r.

5.4.4.1 Strong Beam a Weak Beam

Paprsek (beam) se rozděluje na silný (strong) a slabý (weak). Tato vlastnost se dá zjistit podle toho, v jaké orientaci se družice nachází (atribut Spacecraft Orientation) a jaký má beam název. Podmínka vypadá takto:

- Orientace je vpřed a název paprsku končí na písmeno l, je paprsek slabý
- Orientace je vpřed a název paprsku končí na písmeno r, je paprsek silný
- Orientace je vzad a název paprsku končí na písmeno l, je paprsek silný
- Orientace je vzad a název paprsku končí na písmeno r, je paprsek slabý

Takto definovaná podmínka ve statistickém programu R je vidět na obrázku č. 6

```
library(dplyr)

data_frame <- data_frame %>%
  mutate(beam = case_when(
    beam %in% c('1l', '2l', '3l') & sc_orient == 1 ~ 'weak Beam',
    beam %in% c('1r', '2r', '3r') & sc_orient == 1 ~ 'Strong Beam',
    beam %in% c('1l', '2l', '3l') & sc_orient == 0 ~ 'Strong Beam',
    beam %in% c('1r', '2r', '3r') & sc_orient == 0 ~ 'weak Beam',
    TRUE ~ beam
  ))
```

Obrázek č. 6 – Podmínka pro určení, zda je paprsek (beam) silný (strong) či slabý (weak) v programu R Studio. Zdroj: vlastní

5.5 Hodnocení přesnosti

Ke zjištění přesnosti dat bylo třeba data validovat. Ke zjištění rozdílů výšek dat od skutečnosti byl využit digitální model reliéfu 5. generace (DMR-5g), což je Zeměměřickým úřadem spravovaný model reliéfu s úplnou střední chybou výšky 0,18 m na odkrytém terénu a 0,3 m na zalesněném terénu. Model vznikl leteckým LiDARovým skenováním mezi lety 2009 až 2013. Jeho rastrová podoba byla poskytnuta Katedrou prostorových věd na Fakultě životního prostředí.

Výšky z tohoto modelu byly do dat extrahovány funkcí Extract Multi Values to Point v prostředí ArcGIS Pro. Tato funkce byla zvolena kvůli její dobré optimalizaci na objemná data. Pro každý záznam z dat misí GEDI a ICESat – 2 bylo třeba extrahovat výšky kvazigeoidu, jelikož v současnosti se pro Českou a Slovenskou republiku používají výšky normální (Moloděnského), jenž jsou vztaženy ke kvazigeoidu. Podle terminologického slovníku zeměměřictví a katastru nemovitostí se kvazigeoid definuje jako „referenční plocha normálních výšek blízka geoidu (na rozdíl od geoidu však nemá fyzikální význam); model kvazigeoidu se využívá jako transformační plocha pro vzájemný převod normálních výšek a výšek elipsoidických“ (Kostecký J., 2019). Tento model je ke stažení na stránkách katastru nemovitostí v textovém formátu.

Kvazigeoid byl importován do programu ArcGIS Pro funkcí XY Table to Point a funkcí spline interpolován na raster. Rastrová podoba kvazigeoidu byla pak použita ve funkci Extract Multi Values to Points pro získání hodnot kvazigeoidu pro jednotlivé body.

Hodnoty z DMR-5g a kvazigeoidu byly sečteny pro získání finální výšky, které bude možné porovnat s měřenými výškami z družic. Díky rozdílu výšek je možné zjistit hodnotu Root-mean-square error (RMSE) před a po filtracích dat.

Root Mean Square Error (RMSE) je střední kvadratická chyba která kvantifikuje rozdíl mezi hodnotami predikovanými modely a skutečně pozorovanými hodnotami. Vypočítává se jako druhá odmocnina průměru čtverců rozdílů mezi predikovanými a skutečnými hodnotami. Nízká hodnota RMSE naznačuje, že model má dobrou přesnost předpovědi, zatímco vysoká hodnota RMSE ukazuje, že mezi predikovanými a skutečnými hodnotami existují velké rozdíly. RMSE je obzvláště užitečný, protože penalizuje velké chyby více než malé, což znamená, že modely s velkými výkyvy v předpovědích budou mít vyšší RMSE. Vzoreček použitý pro výpočet RMSE je:

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

Vzorec č. 1 – Vzorec RMSE, kde n je počet měření, y_i je skutečná hodnota a \hat{y}_i je predikovaná hodnota.

5.6 Filtrace podle atributů

Vybrané atributy byly navzájem kombinovány a výsledky byly zaznamenány do tabulek. Byla vytvořena tabulka pro výsledné RMSE v porovnání měřených výšek družicemi a DMR5g, a také tabulka pro počet bodů po filtraci. U dat GEDI byly kombinovány atributy Waveform fidelity flag (WFF), quality flag a number of modes. U družice ICESat-2 byly kombinovány atributy jako number of terrain photons, weak/strong beam a cloud flag. Tyto kombinace byly provedeny v programu R Studio, a na obrázku č. 7 je vidět příklad jedné z kombinací u dat GEDI.

```
library(dplyr)
filtered_data_frame <- data_frame %>%
  filter((quality_flag == TRUE & rx_assess_flag == 0 & num_detectedmodes <= 4))
```

Obrázek č. 7 –

Na obrázku je ukázka filtru kombinací atributů quality flag (quality_flag), waveform fidelity flag (rx_assess_flag) a number of detected modes (num_detectedmodes). Obrázek pochází z programu R Studio a je využit balíček dplyr. Zdroj: vlastní

5.7 Filtrace pomocí mediánu výšek nejbližších bodů

Filtrace pomocí mediánu výšek nejbližších bodů funguje pro body s výškovými hodnotami a známými souřadnicemi. V této filtraci se nastavují dva parametry, počet sousedících bodů a hranice filtrování. U počtu sousedících bodů se zadává, z kolika

nejbližších bodů bude medián spočítán a hranice filtrování určuje, jak moc se od sebe může medián a výška bodu lišit, než bude bod odstraněn. V příloze 2 je tento filtr vidět napsaný ve skriptu v programu R. Tato metoda filtrace byla provedena pro obě mise.

5.8 Počet bodů na čtverec

Aby bylo možné vypočítat medián výšek bodů na čtverec bylo třeba vytvořit čtvercovou síť v zájmovém území. Čtvercové sítě byly vytvořeny v programu ArcGIS Pro nástrojem Create Fishnet, ve kterém byl nastaven rozměr čtverců a jejich rozsah. Rozsah byl nastaven tak, aby překrýval celé studované území a rozměr čtverců 0,5 km, 1 km, 2 km, 5 km a 10 km. Každému čtverci byl přidělen unikátní identifikátor (ID). Čtvercové sítě byly vyexportovány ve formátu SHP (Shapefile).

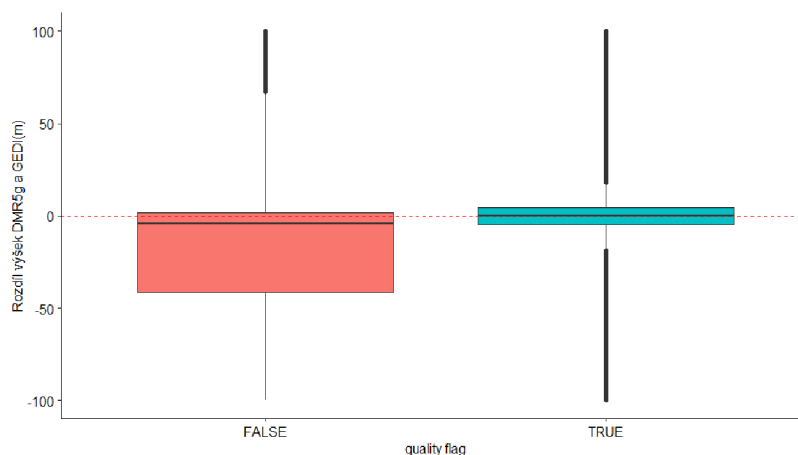
V programu R Studio bylo zjištěno, kolik bodů z již vyfiltrovaných dat se nachází v jednotlivých čtvercích. Toto bylo provedeno pro všechny rozměry čtvercových sítí pomocí funkce intersect z balíčku Terra. Následně bylo zjištěno, kolik čtverců obsahuje nulový počet bodů a jaký je medián počtu bodů ve čtverci. Výsledky byly zaznamenány do tabulek. Skript filtrací a výpočtů na čtverec může být nalezen v příloze číslo 2.

6. Výsledky

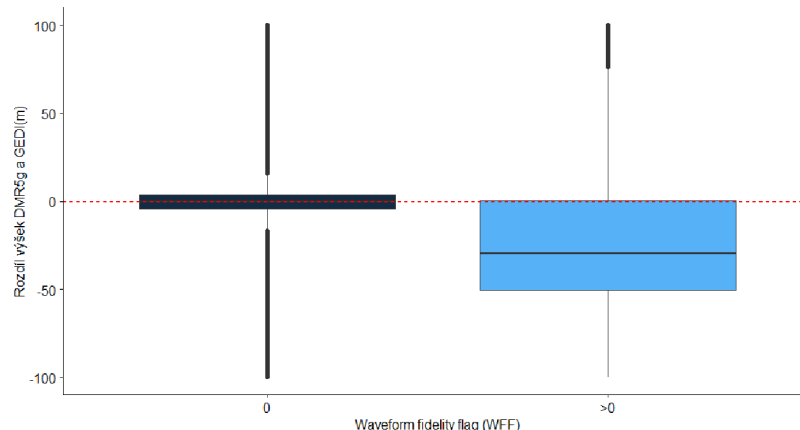
Výsledky analýzy a následné filtrace byly posuzovány dle hodnot RMSE, celkového počtu bodů a mediánu počtu bodů na čtverec. Hodnota RMSE byla vypočtena z rozdílu výšek dat laserového družicového skenování a digitálního modelu reliéfu 5. generace (DMR – 5g).

6.1 GEDI

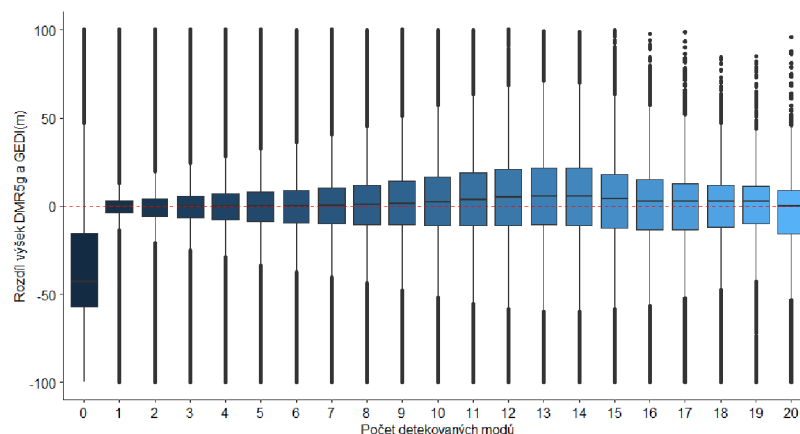
Po oříznutí obsahovala data z družice GEDI na území Čech 36 857 568 bodů s RMSE 2595,79 m. Takto velké RMSE je nevhodné k tvorbě digitálního modelu terénu, proto bylo potřeba odstranit hodnoty s největším výškovým rozdílem. To bylo možné pomocí filtrace z atributů. Zde byly využity atributy „Quality flag“, „Waveform fidelity flag - WFF“ a „Number of modes - NoM“. Největší impakt na RMSE měl atribut Quality flag = 1 (obrázek číslo 8). Tento atribut snížil RMSE na 24,80 m a počet bodů na 12 022 879. Výsledky byly také zlepšeny díky atributu Waveform fidelity flag (WFF). Po vyfiltrovaní nenulových hodnot atributu WFF bylo RMSE 195,37 m a počet bodů byl roven 15 290 321 (obrázek číslo 9). U atributu 'Number of modes' (NoM) jsou hodnoty v rozmezí 0 – 20 avšak 91 procent bodů se nacházelo pouze v intervalu od 0 do 4. Výsledky potvrdily předpoklad (obrázek číslo 10), že se výsledky s nejnižšími odchylkami budou nacházet v hodnotách 1 – 4. RMSE pro hodnoty NoM v rozmezí 1 až 4 dosáhlo 690,24 m, zatímco pro ostatní hodnoty NoM RMSE vzrostlo na 3430,57 m.



Obrázek č. 8 – Krabicový graf atributu quality flag. Na ose X jsou hodnoty TRUE (kvalitní) a FALSE (nekvalitní). Na ose Y je rozdíl výšek měřených (GEDI) a referenčních (DMR5g). Graf je omezen -100 a 100 m na ose Y. Zdroj: vlastní



Obrázek č. 9 – Krabicový graf atributu waveform fidelity flag (WFF). Na ose X jsou hodnoty 0 (bez erroru) a >0 (s errorem). Na ose Y je rozdíl výšek měřených (GEDI) a referenčních (DMR5g). Graf je omezen -100 a 100 m na ose Y. Zdroj: vlastní



Obrázek č. 10 – Krabicový graf atributu number of modes (počet módů). Na ose X jsou hodnoty 0 – 20 zaznamenaných módů. Na ose Y je rozdíl výšek měřených (GEDI) a referenčních (DMR5g). Graf je omezen -100 a 100 m na ose Y. Zdroj: vlastní

Podmínka pro další filtraci byla zjednodušena, a to na hodnoty větší než 4 a hodnoty menší nebo rovné 4. Tento krok byl proveden, jelikož kombinace s „quality flag“ prokázala, že $NoM = 0$ ve 100 % případů odpovídá $quality\ flag = 0$. Při kombinaci atributů k filtraci nejlépe vyšla kombinace $WFF = 0$, $quality\ flag = 0$ a $number\ of\ modes \leq 4$ s RMSE 14,14 m a 10 359 430 bodů, což je vidět na tabulkách 1 a 2.

RMSE	WFF = 0	WFF ≠ 0	WFF = 0	WFF ≠ 0
	number of modes > 4		number of modes ≤ 4	
quality flag = 1	27,18	102,17	14,14	126,54
quality flag = 0	1586,83	2169,52	280,58	3503,60

Tabulka č.1 – Tabulka ukazuje RMSE při kombinaci atributů Quality flag, Waveform fidelity flag (WFF) a number of modes.

Počet bodů	WFF = 0	WFF ≠ 0	WFF = 0	WFF ≠ 0
	number of modes > 4		number of modes ≤ 4	
quality flag = 1	1 312 350	227 273	10 359 430	122 880
quality flag = 0	121 250	1 682 636	3 496 595	19 538 270

Tabulka č.2 – Tabulka ukazuje počet bodů při kombinaci atributů Quality flag, Waveform fidelity flag (WFF) a number of modes.

K dalšímu procesu byla použita nejvhodnější filtrace, která vznikla kombinací atributů WFF = 0, quality flag = 0 a number of modes ≤ 4. V dalším kroku byl využit filtr mediánu výšek nejbližších bodů. Medián okolních bodů byl vypočten z 9 nejbližších k filtrovanému bodu a odchylka mediánu výšek od měřených byla nastavena 5 metrů. Počet bodů po této filtraci byl 8 958 558 a RMSE se snížilo na 8,52 m (tabulka číslo 15).

K porovnání efektivity filtrací byla data proložena čtvercovou sítí. Velikost čtverců byla v 5 rozměrech a to 0,5 km, 1 km, 2 km, 5 km, 10 km. Data byla vyfiltrována dle velikosti rozdílů výšek mezi DMR5g a měřených výšek GEDI. Hranice rozdílů výšek byla nastavena na menší než 10, 15, 20 a 30 metrů. Počet bodů pro hranice rozdílů výšek v tabulce číslo 7 poklesl po filtraci o 41–47 %. Pro každý čtverec byl zjištěn počet bodů a kolik čtverců se nachází bez jediného bodu. Toto bylo provedeno pro data před a po filtracích. V tabulkách 3 a 4 je uvedený medián počtu bodů ve čtverci před a po filtracích. Rozdíl mediánu před a po filtracích se pohybuje mezi 41,6 % až 48,1 %.

Dle tabulek číslo 5 a 6 čtverce bez bodů se objevují u čtvercové síti s rozměrem 1 km a nižší. Počet čtverců bez bodů o rozměru 1 km je v rozmezí 70 až 75 čtverců před filtrací a 75 až 113 čtverců po filtraci. To představuje podíl 0,13 – 0,22 % z celkových 51 077 čtverců.

Data GEDI před filtrací					
medián počtu bodů		rozdíl [m] (GEDI – DMR5g)			
		<10	<15	<20	<30
velikost strany čtverce	0,5 km	61	69	74	79
	1 km	244	279	298	319
	2 km	981	1121	1198	1282
	5 km	6223	7061	7602	8107
	10 km	25340	28636	30643	32750

Tabulka č. 3 – Tabulka ukazující medián počtu bodů ve čtvercích pro data GEDI před filtrací o velikostech od 0,5 – 10 km. Každý čtverec ukazuje medián pro odchylky výšek 10, 15, 20 a 30 metrů od DMR5g

Data GEDI po filtraci					
medián počtu bodů		rozdíl [m] (GEDI – DMR5g)			
		<10	<15	<20	<30
velikost čtverce	0,5 km	35	39	40	41
	1 km	142	157	163	167
	2 km	573	631	654	670
	5 km	3641	3961	4130	4218
	10 km	14787	16151	16754	17095

Tabulka č. 4 – Tabulka ukazující medián počtu bodů ve čtvercích pro data GEDI po filtraci o velikostech od 0,5 – 10 km. Každý čtverec ukazuje medián pro odchylky výšek 10, 15, 20 a 30 metrů od DMR5g

Data GEDI před filtrací					
počet čtverců s 0 body		rozdíl [m] (GEDI – DMR5g)			
		<10	<15	<20	<30
velikost čtverce	0,5 km	1365	1143	1075	1037
	1 km	75	71	70	70
	2 km	0	0	0	0
	5 km	0	0	0	0
	10 km	0	0	0	0

Tabulka č. 5 – Tabulka ukazující počet čtverců, ve kterých se nenachází ani jeden bod před filtrací u dat GEDI.

Data GEDI po filtraci					
počet čtverců s nulou		rozdíl [m] (GEDI - DMR5g)			
		<10	<15	<20	<30
velikost čtverce	0,5 km	3940	2418	1812	1421
	1 km	113	85	79	75
	2 km	0	0	0	0
	5 km	0	0	0	0
	10 km	0	0	0	0

Tabulka č. 6 – Tabulka ukazující počet čtverců, ve kterých se nenachází ani jeden bod po filtracích u dat GEDI.

Celkový počet bodů – GEDI				
	rozdíl [m] (GEDI – DMR5g)			
	<10	<15	<20	<30
před filtrací	12 909 166	14 684 870	15 688 232	16 823 344
po filtraci	7 529 485	8 288 879	8 595 055	8 815 028

Tabulka č. 7 – Počet bodů před a po filtraci omezených rozdílem výšek mezi daty GEDI a DMR5g.

6.2 ICESat – 2

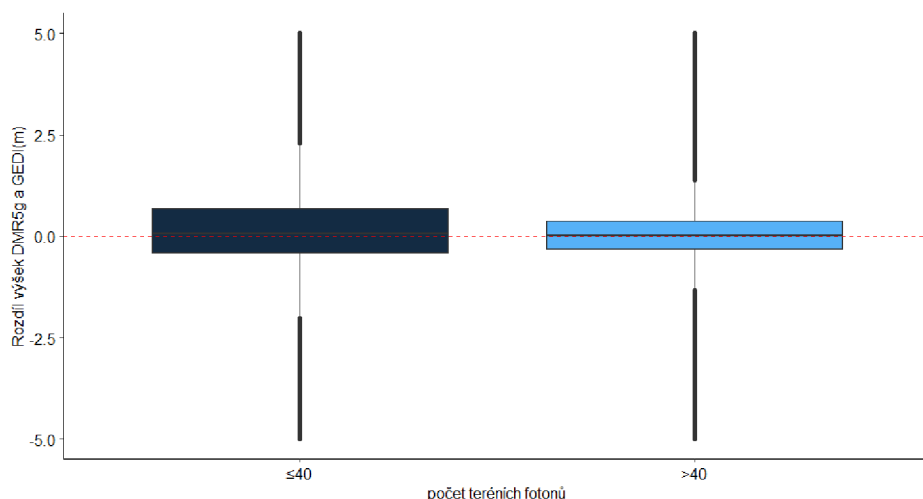
Vstupní data družice ICESat – 2 měla 6 860 345 bodů a hodnotu RMSE 1,47 m. Oproti vstupním datům družice GEDI je RMSE výrazně menší a odchylky od DMR5g mají menší variabilitu. Tento fakt kompenzuje, že data mají oproti GEDI menší počet bodů. K filtraci byly zvoleny atributy number terrain photons, cloud flag a beam type, který byl vytvořen pomocí atributů beam a spacecraft orientation. Nejužitečnějším atributem se stal number terrain photons > 40, který snížil RMSE na 1,36 m a počet bodů snížil jen o 20 % (obrázek číslo 11). Kombinace atributů s nejnižším RMSE vznikla použitím kombinace atributů number terrain photons > 40, cloud flag > 1 a Strong Beam. Tato kombinace zredukovala body na pouhých 615 108. Takto velký úbytek není vhodný a pro další filtraci, proto bylo třeba využít méně razantní kombinaci. Nejvhodnější kombinace byla zvolena použitím kombinace atributů terrain photons > 40 a Strong beam nebo Weak Beam. Zvolena byla z důvodu potřeby zachovat co nejvíce bodů i přes zvýšení hodnoty RMSE. Tato kombinace snížila počet bodů o 9,6 % a RMSE na 1,39 m a byla použita pro další filtraci.

<i>RMSE</i>	terrain photons > 40	terrain photons ≤ 40	terrain photons > 40	terrain photons ≤ 40
	Strong Beam		Weak beam	
cloud flag > 1	1,17	2,20	1,43	1,60
cloud flag ≤ 1	1,38	2,06	1,36	1,65

Tabulka č.8 – Tabulka ukazuje RMSE při kombinaci atributů terrain photons, Weak/Strong beam a cloud flag u dat ICESat – 2.

<i>Počet bodů</i>	terrain photons > 40	terrain photons ≤ 40	terrain photons > 40	terrain photons ≤ 40
	Strong Beam		Weak beam	
cloud flag > 1	615 108	106 272	135 187	81 197
cloud flag ≤ 1	3 640 156	555 459	1 120 377	602 555

Tabulka č.9 – Tabulka ukazuje počet bodů při kombinaci atributů terrain photons, Weak/Strong beam a cloud flag u dat ICESat – 2.



Obrázek č. 11 - Krabicový graf atributu number of terrain photons (počet terénních fotonů). Na ose X jsou hodnoty ≤40 a >40 terénních fotonů. Na ose Y je rozdíl výšek měřených (GEDI) a referenčních (DMR5g). Graf je omezen -100 a 100 m na ose X. Zdroj: vlastní

Metoda filtrace mediánu výšek nejbližších bodů byla aplikována i zde pro dodatečné snížení RMSE. Pro výpočet mediánu výšek bylo zvoleno devět nejbližších bodů a práh pro odstranění byl nastaven na 2 metry. Hodnota prahu byla zvolena tak, aby odstranila co nejméně bodů a zároveň RMSE co nejvíce snížila. Výpočet s těmito parametry snížil počet bodů o 10 % na 5 554 456 a hodnotu RMSE na 1,23 m (tabulka číslo 15).

Stejně jako u dat GEDI k porovnání efektivity filtrací byla data proložena čtvercová síť. Velikost čtverců byla ve stejných rozměrech a to 0,5 km 1 km, 2 km, 5 km, 10 km. Data byla vyfiltrována dle velikosti rozdílů výšek mezi DMR5g a měřených výšek GEDI. Hranice rozdílů výšek byla nastavena na menší než 0,25 m, 0,5 m, 1 m, 2 m a 5 metrů. Oproti datům GEDI byly hranice nižší vzhledem k menším odchylkám ve výškách. Počet bodů pro hranice rozdílů výšek v tabulce číslo 14 poklesl po filtraci o 11 – 22 %. Pro každý čtverec byl zjištěn počet bodů a kolik čtverců se nachází bez jediného bodu. V tabulkách 10 a 11 je uvedený medián počtu bodů ve čtverci před a po filtracích. Rozdíl mediánu před a po filtracích se pohybuje mezi 0 % až 19 %.

Dle tabulek číslo 12 a 13 čtverce bez bodů se objevují u čtvercové síti s rozměrem 1 km a nižší. Počet čtverců bez bodů u čtverců o rozměru 1 km je v rozmezí 70 až 75 čtverců před filtrací a 24 až 109 čtverců po filtraci. To představuje podíl 0,04 – 0,21 % z celkových 51 077 čtverců.

Data ICESat – 2 před filtrací						
medián počtu bodů		rozdíl [m] (ICESat-2 – DMR5g)				
		<0,25	<0,5	<1	<2	<5
velikost strany čtverce	0,5 km	12	20	28	33	35
	1 km	37	64	91	108	118
	2 km	156	268	381	452	480
	5 km	1094	1845	2564	2991	3152
	10 km	4653	7902	10568	12127	12942

Tabulka č. 10 – Tabulka ukazující medián počtu bodů ve čtvercích pro data ICESat – 2 před filtrací o velikostech od 0,5 – 10 km. Každý čtverec ukazuje medián pro odchylky výšek 0,25, 0,5, 1, 2 a 5 metrů od DMR5g

Data ICESat – 2 po filtraci						
medián počtu bodů		rozdíl [m] (ICESat-2 – DMR5g)				
		<0,25	<0,5	<1	<2	<5
velikost strany čtverce	0,5 km	12	19	26	29	30
	1 km	33	56	77	88	116
	2 km	135	230	319	364	375
	5 km	960	1615	2193	2469	2534
	10 km	4586	6906	9552	10592	10622

Tabulka č.11 – Tabulka ukazující medián počtu bodů ve čtvercích pro data ICESat – 2 po filtraci o velikostech od 0,5 – 10 km. Každý čtverec ukazuje medián pro odchylky výšek 0,25, 0,5, 1, 2 a 5 metrů od DMR5g

Data ICESat – 2 před filtraci						
počet čtverců s 0 body		rozdíl [m] (ICESat-2 – DMR5g)				
		<0,25	<0,5	<1	<2	<5
velikost strany čtverce	0,5 km	71465	64817	60889	59112	584480
	1 km	4848	4247	3957	3821	3774
	2 km	32	29	25	24	24
	5 km	0	0	0	0	0
	10 km	0	0	0	0	0

Tabulka č. 12 – Tabulka ukazující počet čtverců, ve kterých se nenachází ani jeden bod před filtrací u dat ICESat – 2.

Data ICESat – 2 po filtraci						
počet čtverců s 0 body		rozdíl [m] (ICESat-2 – DMR5g)				
		<0,25	<0,5	<1	<2	<5
velikost strany čtverce	0,5 km	85517	78302	73341	70860	70088
	1 km	6946	6034	5471	5183	5078
	2 km	109	76	56	50	48
	5 km	0	0	0	0	0
	10 km	0	0	0	0	0

Tabulka č. 13 – Tabulka ukazující počet čtverců, ve kterých se nenachází ani jeden bod po filtraci u dat ICESat – 2.

Celkový počet bodů - ICESat – 2					
	rozdíl [m] (ICESat-2 – DMR5g)				
	<0,25	<0,5	<1	<2	<5
Před filtrací	2 575 071	4 134 219	5 590 270	6 458 481	6 809 646
Po filtraci	2 290 280	3 620 689	4 794 010	5 024 160	5 498 443

Tabulka č. 14 – Počet bodů před a po filtraci omezených rozdílem výšek mezi daty GEDI a DMR5g

Výsledky filtrací u dat GEDI a ICESat - 2				
	GEDI		ICESat	
	počet bodů	RMSE [m]	počet bodů	RMSE [m]
počáteční data	36 857 568	2595,79	6 860 345	1,47
filtrace z atributů	10 359 430	14,14	6 198 614	1,39
filtrace mediánu výšek okolních bodů	8 958 558	8,52	5 554 456	1,23

Tabulka č. 15 – Tabulka zobrazuje RMSE a počet bodů filtrací a počátečních dat u misí GEDI a ICESat – 2.

7. Diskuse

Předchozí výzkumy se shodují, že data získaná z GEDI a ICESat-2 poskytují lepší výsledky po provedení filtrace z různých příčin (Quirós et al. 2021; Adam et al. 2020; Liu et al. 2021). V rámci studie od Adama et al. (2020) bylo zjištěno, že mezi daty GEDI se nachází 9 – 13 % odlehlých hodnot. V mé bakalářské práci bylo identifikováno, že odlehlé hodnoty tvořily u dat GEDI 47 % a u dat z ICESat-2 přibližně 6 %. Co se týče dat GEDI, moje zjištění nekorrespondují se studií od Adama et al. (2020). U dat z ICESat-2 se moje výsledky blíží těm, které zaznamenali Narin & Gullu (2023), s podílem odlehlých hodnot v rozmezí 2–7 %.

U dat ICESat – 2 Neuenschwander a Pitts (2022) popisuje negativní dopad na měření při nízkém počtu signálních fotonů, ze kterých se určují terénní fotony. Tento dopad je možný vidět i u mých dat, kdy odchylka výškových dat byla vyšší, pokud byl počet terénních fotonů menší než 40.

Při porovnáním výsledků od Quiros et al., (2021), kde dosáhli 6,13 m RMSE u družice GEDI, se stejně jako v této práci podařili dostat pod ± 10 metrovou odchylku. Narin & Gullu, (2023) ve své studii porovnávali data z GEDI a ICESat-2 s dalšími digitálními modely terénu. Mé výsledky se shodují s jejich zjištěním, že ICESat – 2 poskytuje lepší výsledky z hlediska RMSE. Nicméně, po vytvoření digitálního modelu terénu se RMSE hodnoty dat z ICESat-2 výrazně zvyšují, což autoři připisují nízké hustotě měřených bodů.

8. Závěr

Výsledky této práce u dat GEDI a ICESat – 2 dosáhly přesnosti podobné předchozím studiím. Při tvorbě DMT na území České republiky z dat ICESat – 2 se vertikální přesnost modelu může očekávat s RMSE 2 m a rozlišení modelu v rozmezí 750 m až 2 km. U dat GEDI je vertikální přesnost po filtracích 8 m RMSE. Vzhledem k větší hustotě dat v mírných pásmech by pro Čechy mohlo být zvoleno větší rozlišení pohybující se v rozmezí 500 m až 1 km.

K dosažení co nejvyšší vertikálních přesnosti pro tvorbu DMT se osvědčily atributy Quality flag a WFF u dat GEDI a number of terrain photons u družice ICESat – 2. Pro obě mise je vhodná filtrace mediánu nejbližších výšek, která dobře filtruje extrémně odlehlé, ale je třeba medián počítat alespoň ze 6 sousedních bodů.

Z výsledků je zřejmé, že data z družice ICESat- 2 vychází ohledně RMSE mnohem lépe než data z GEDI. Již bez filtrace byly výškové hodnoty se znatelně menší odchylkou. Družice ICESat – 2 obsahuje na území Čech méně bodů než GEDI a její hustota je tedy o 81% řidší. Filtrace dat GEDI smaže okolo tří čtvrtin bodů, což jejich počet dostane na úroveň ICESat – 2, kde filtrace smaže jen 20 %.

Aby se dosáhlo dalšího zlepšení, možnost spojení obou družicových dat k filtraci pomocí mediánu nejbližších bodů byla zvažována, ale neuskutečněna kvůli časovým a výpočetním omezením. Pro tvorbu modelů terénu s vyšší vertikální přesností by mohlo spojení dat napomoc s hustotou dat.

9. Přehled literatury a použitých zdrojů

9.1 Odborné publikace

Abdalati, W., Zwally, H. J., Bindschadler, R., Csatho, B., Farrell, S. L., Fricker, H. A., Harding, D., Kwok, R., Lefsky, M., Markus, T., Marshak, A., Neumann, T., Palm, S., Schutz, B., Smith, B., Spinhirne, J., & Webb, C. (2010). The ICESat-2 laser altimetry mission. *Proceedings of the IEEE*, 98(5), 735–751. <https://doi.org/10.1109/JPROC.2009.2034765>

Adam, M., Urbazaev, M., Dubois, C., & Schmullius, C. (2020). Accuracy assessment of gedi terrain elevation and canopy height estimates in European temperate forests: Influence of environmental and acquisition parameters. *Remote Sensing*, 12(23), 1–28. <https://doi.org/10.3390/rs12233948>

Blair, J. B., & Hofton, M. A. (1999). Modeling laser altimeter return waveforms over complex vegetation using high-resolution elevation data. *Geophysical Research Letters*, 26(16), 2509–2512. <https://doi.org/10.1029/1999GL010484>

Collis, R. T. H. (1969). lidar for routine meteorological observations. *Bulletin of the American Meteorological Society*, 50(9), 688–695. <https://doi.org/10.1175/1520-0477-50.9.688>

Dou, C., Zhang, X., Guo, H., Han, C., & Liu, M. (2014). Improving the Geolocation Algorithm for Sensors Onboard the ISS: Effect of Drift Angle. *Remote Sensing* 2014, Vol. 6, Pages 4647-4659, 6(6), 4647–4659. <https://doi.org/10.3390/RS6064647>

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P. L., Qi, W., & Silva, C. (2020a). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*, 1, 100002. <https://doi.org/10.1016/J.SRS.2020.100002>

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P. L., Qi, W., & Silva, C. (2020b). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*, 1, 100002. <https://doi.org/10.1016/J.SRS.2020.100002>

Duncanson, L., Kellner, J. R., Armston, J., Dubayah, R., Minor, D. M., Hancock, S., Healey, S. P., Patterson, P. L., Saarela, S., Marselis, S., Silva, C. E., Bruening, J., Goetz, S. J., Tang, H., Hofton, M., Blair, B., Luthcke, S., Fatoyinbo, L., Abernethy, K., ... Zraggen, C. (2022). Aboveground biomass density models for NASA's Global Ecosystem Dynamics Investigation (GEDI) lidar mission. *Remote Sensing of Environment*, 270. <https://doi.org/10.1016/j.rse.2021.112845>

Duncanson, L., Neuenschwander, A., Hancock, S., Thomas, N., Fatoyinbo, T., Simard, M., Silva, C. A., Armston, J., Luthcke, S. B., Hofton, M., Kellner, J. R., & Dubayah, R. (2020). Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County, California. *Remote Sensing of Environment*, 242, 111779. <https://doi.org/10.1016/J.RSE.2020.111779>

Dutta Majumdar, J., & Manna, I. (2011). Laser material processing. *International Materials Reviews*, 56(5–6), 341–388. <https://doi.org/10.1179/1743280411Y.0000000003>

Gagliano, F. P., Lumley, R. M., & Watkins, L. S. (1969). Lasers in Industry. In *PROCEEDINGS OF THE IEEE* (Vol. 57, Issue 2).

Haley, D. (2023). Basic principles of lasers. *Anaesthesia and Intensive Care Medicine*, 22(3), 194–196. <https://doi.org/10.1016/j.mpaic.2021.01.004>

Hancock, S., Armston, J., Hofton, M., Sun, X., Tang, H., Duncanson, L. I., Kellner, J. R., & Dubayah, R. (2019). The GEDI Simulator: A Large-Footprint Waveform Lidar Simulator for Calibration and Validation of Spaceborne Missions. *Earth and Space Science*, 6(2), 294–310. <https://doi.org/10.1029/2018EA000506>

Harding, D. J., & Carabajal, C. C. (2005). ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure. *Geophysical Research Letters*, 32(21), 1–4. <https://doi.org/10.1029/2005GL023471>

Hofton, M., Blair, B., Story, S., & Yi, D. (2019). Algorithm Theoretical Basis Document (ATBD) for GEDI Transmit and Receive Waveform Processing for L1 and L2 Products.

Kostecký J.: Globální polohové souřadnicové systémy, vydala Česká technika – nakladatelství ČVUT, 2019 (skripta)

Kropáček, J., Neckel, N., & Bauder, A. (2014). Estimation of mass balance of the grosser Aletschgletscher, swiss alps, from ICESat laser altimetry data and digital elevation models. *Remote Sensing*, 6(6), 5614–5632. <https://doi.org/10.3390/RS6065614>

Kurtz, N. T., & Markus, T. (2012). Satellite observations of Antarctic sea ice thickness and volume. *Journal of Geophysical Research: Oceans*, 117(C8), 8025. <https://doi.org/10.1029/2012JC008141>

Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., & Wegner, J. D. (2022). Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. *Remote Sensing of Environment*, 268, 112760. <https://doi.org/10.1016/J.RSE.2021.112760>

Lefsky, M. A. (2010). A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. *Geophysical Research Letters*, 37(15). <https://doi.org/10.1029/2010GL043622>

Lefsky, M. A., Harding, D. J., Keller, M., Cohen, W. B., Carabajal, C. C., del Bom Espirito-Santo, F., Hunter, M. O., & de Oliveira, R. (2005). Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical Research Letters*, 32(22), 1–4. <https://doi.org/10.1029/2005GL023971>

Li, Y., Fu, H., Zhu, J., & Wang, C. (2021). A Filtering Method for ICESat-2 Photon Point Cloud Data Based on Relative Neighboring Relationship and Local Weighted Distance Statistics. *IEEE Geoscience and Remote Sensing Letters*, 18(11), 1891–1895. <https://doi.org/10.1109/LGRS.2020.3011215>

Liu, A., Cheng, X., & Chen, Z. (2021). Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals. *Remote Sensing of Environment*, 264, 112571. <https://doi.org/10.1016/j.rse.2021.112571>

Liu, G., Wang, J., Dong, P., Chen, Y., & Liu, Z. (2018). Estimating individual tree height and diameter at breast height (DBH) from terrestrial laser scanning (TLS) data at plot level. *Forests*, 8(2). <https://doi.org/10.3390/f9070398>

Lohani, B., & Ghosh, S. (2017). Airborne LiDAR Technology: A Review of Data Collection and Processing Systems. *Proceedings of the National Academy of Sciences India Section A - Physical Sciences*, 87(4), 567–579.
<https://doi.org/10.1007/S40010-017-0435-9/TABLES/4>

Luthcke, S. B., Rebold, T., Thomas, T., & Pennington, T. (2019). Algorithm Theoretical Basis Document (ATBD) for GEDI Waveform Geolocation for L1 and L2 Products.

Dutta Majumdar, J., & Manna, I. (2011). Laser material processing. *International Materials Reviews*, 56(5–6), 341–388.
<https://doi.org/10.1179/1743280411Y.0000000003>

Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Farrell, S., Fricker, H., Gardner, A., Harding, D., Jasinski, M., Kwok, R., Magruder, L., Lubin, D., Luthcke, S., Morison, J., Nelson, R., Neuenschwander, A., Palm, S., ... Zwally, J. (2017). The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation. *Remote Sensing of Environment*, 190, 260–273. <https://doi.org/10.1016/J.RSE.2016.12.029>

Milenković, M., Schnell, S., Holmgren, J., Ressl, C., Lindberg, E., Hollaus, M., Pfeifer, N., & Olsson, H. (2017). Influence of footprint size and geolocation error on the precision of forest biomass estimates from space-borne waveform LiDAR. *Remote Sensing of Environment*, 200, 74–88.
<https://doi.org/10.1016/J.RSE.2017.08.014>

Narin, O. G., & Gullu, M. (2023). A comparison of vertical accuracy of global DEMs and DEMs produced by GEDI, ICESat-2. *Earth Science Informatics*, 16(3), 2693–2707. <https://doi.org/10.1007/s12145-023-01060-1>

Neuenschwander, A. L., & Magruder, L. A. (2019). Canopy and terrain height retrievals with ICESat-2: A first look. *Remote Sensing*, 11(14).
<https://doi.org/10.3390/rs11141721>

Neuenschwander, A., & Pitts, K. (2019). The ATL08 land and vegetation product for the ICESat-2 Mission. *Remote Sensing of Environment*, 221, 247–259.
<https://doi.org/10.1016/J.RSE.2018.11.005>

Neumann, T. A., Martino, A. J., Markus, T., Bae, S., Bock, M. R., Brenner, A. C., Brunt, K. M., Cavanaugh, J., Fernandes, S. T., Hancock, D. W., Harbeck, K., Lee, J., Kurtz, N. T., Luers, P. J., Luthcke, S. B., Magruder, L., Pennington, T. A., Ramos-Izquierdo, L., Rebold, T., Thomas, T. C. (2019). The Ice, Cloud, and Land Elevation Satellite – 2 mission: A global geolocated photon product derived from the Advanced Topographic Laser Altimeter System. *Remote Sensing of Environment*, 233, 111325. <https://doi.org/10.1016/J.RSE.2019.111325>

Neumann, T. A., Brenner, A., Hancock, D., Robbins, J., Gibbons, A., Lee, J., Harbeck, K., Saba, J., Luthcke, S., & Rebold, T. (2022). ICE, CLOUD, and Land Elevation Satellite-2 (ICESat-2) Project Algorithm Theoretical Basis Document (ATBD) for Global Geolocated Photons ATL03. <https://doi.org/10.5067/GA5KCLJT7LOT>

Neumann, T. A., Martino, A. J., Markus, T., Bae, S., Bock, M. R., Brenner, A. C., Brunt, K. M., Cavanaugh, J., Fernandes, S. T., Hancock, D. W., Harbeck, K., Lee, J., Kurtz, N. T., Luers, P. J., Luthcke, S. B., Magruder, L., Pennington, T. A., Ramos-Izquierdo, L., Rebold, T., ... Thomas, T. C. (2019). The Ice, Cloud, and Land Elevation Satellite – 2 mission: A global geolocated photon product derived from the Advanced Topographic Laser Altimeter System. *Remote Sensing of Environment*, 233, 111325. <https://doi.org/10.1016/J.RSE.2019.111325>

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C. E., Armston, J., Dubayah, R., Blair, J. B., & Hofton, M. (2021). Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sensing of Environment*, 253, 112165. <https://doi.org/10.1016/J.RSE.2020.112165>

Quiros, E., Polo, M. E., & Fragoso-Campon, L. (2021). GEDI Elevation Accuracy Assessment: A Case Study of Southwest Spain. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 5285–5299. <https://doi.org/10.1109/JSTARS.2021.3080711>

Schleich, A., Durrieu, S., Soma, M., & Vega, C. (2023). Improving GEDI Footprint Geolocation Using a High-Resolution Digital Elevation Model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16, 7718–7732. <https://doi.org/10.1109/JSTARS.2023.3298991>

Spinhirne, J. D. (1993). Micro Pulse Lidar. *IEEE Transactions on Geoscience and Remote Sensing*, 31(1), 48–55. <https://doi.org/10.1109/36.210443>

Urbazaev, M., Hess, L. L., Hancock, S., Sato, L. Y., Ometto, J. P., Thiel, C., Dubois, C., Heckel, K., Urban, M., Adam, M., & Schullius, C. (2022). Assessment of terrain elevation estimates from ICESat-2 and GEDI spaceborne LiDAR missions across different land cover and forest types. *Science of Remote Sensing*, 6, 100067. <https://doi.org/10.1016/J.SRS.2022.100067>

Urbazaev, M., Hess, L. L., Sato, L., Ometto, J. P., Thiel, C., Dubois, C., Adam, M., & Schullius, C. (2021). Accuracy assessment of terrain and canopy height estimates from ICESat-2 and GEDI LiDAR missions in temperate and tropical forests: first results. <https://doi.org/10.34726/wim.1861>

Yang, F., He, Y., Chen, W., & Zhan, Y. (2014). Laser altimeter based on random code phase modulation and heterodyne detection. *IEEE Photonics Technology Letters*, 26(23), 2337–2340. <https://doi.org/10.1109/LPT.2014.2356333>

9.2 Internetové zdroje

Hierarchical Data Format. Online. In: Wikipedia: the free encyclopedia. San Francisco (CA): Wikimedia Foundation, 2001-, 2024-02-22. Dostupné z: https://en.wikipedia.org/wiki/Hierarchical_Data_Format. [cit. 2024-03-20].

The HDF Group. Online. The HDF5 Library and File Format. 2006, 2024-03-08. Dostupné z: <https://www.hdfgroup.org/solutions/hdf5>. [cit. 2024-03-20].

The HDF Group. Online. HDF View. 2006, 2024-03-08. Dostupné z: <https://www.hdfgroup.org/downloads/hdfview/>. [cit. 2024-03-20].

R Project [online]. b. r. [cit. 2024-03-28]. Dostupné z: <https://www.r-project.org/>

VÝZKUMNÝ ÚSTAV GEODETICKÝ, TOPOGRAFICKÝ A KARTOGRAFICKÝ, V.V.I. [VÚGTK]. *Terminologický slovník zeměměřictví a katastru nemovitostí*. Online. C2005-2020, 2020. Dostupné z: <https://www.slovníkcuzk.eu/>. [cit. 2024-03-28].

GLEN, Sephanie. *Statistics How To*. Online. RMSE: Root Mean Square Error. C2024. Dostupné z: <https://www.statisticshowto.com/probability-and-statistics/regression-analysis/rmse-root-mean-square-error/>. [cit. 2024-03-28]

10. Přílohy

Příloha číslo 1 – skript v jazyce R sloužící k otevření HDF5 formátu a převedení do textového formátu.

```
##if (!require("BiocManager", quietly = TRUE))
## install.packages("BiocManager")

#BiocManager::install("rhdf5")
library(rhdf5)
library(readr)

setwd("D:/data_GEDI/export2")

filles <- list.files(path="D:/data_GEDI/unzip",
                    full.names = T, recursive = T, pattern = ".h5", include.dirs = T)

beams = c("BEAM0000", "BEAM0001", "BEAM0010", "BEAM0011", "BEAM0101", "BEAM0110", "BEAM1000", "BEAM1011")

#x = "GEDI02_B_2022298050102_021903_02_T08276_02_003_01_V002.h5"

for (b in beams) {

  # The L2B algorithm is run if this flag is set to 1 indicating data have sufficient waveform fidelity for L2B to run
  algorithmrun_flag = vector()

  # Beam identifier
  beam = vector()
  FLY = vector()
  lat_lowestmode = vector() #Latitude of center of lowest mode
  lon_lowestmode = vector() #Longitude of center of lowest mode
  shot_number= vector() #Unique shot ID.
  quality_flag= vector()
  degrade_flag= vector()
  sensitivity= vector()
  elev_lowestmode= vector()
  elev_lowestreturn_a1= vector()
  elev_highestreturn= vector()
  delta_time= vector()
  solar_elevation= vector()
  surface_flag= vector()
  landsat_treecover= vector()
  selected_mode_flag= vector()
  stale_return_flag= vector()
  rx_gflag= vector()
  rx_assess_flag= vector()
  energy_total= vector()
  num_detectedmodes= as.numeric(vector())

  for (x in filles) {

    h5f = H5Fopen(x, "H5F_ACC_RDONLY")

    value <- tryCatch(
      as.vector(h5read(h5f, paste(b, "/geolocation/lat_lowestmode_a1", sep="")),
        error = function(c) {NA}
      )
    )
    lat_lowestmode <- c(lat_lowestmode, value)

    value <- tryCatch(
      as.vector(h5read(h5f, paste(b, "/geolocation/lon_lowestmode_a1", sep="")),
        error = function(c) {NA}
      )
    )
    LENGTH=length(value) #Počet radku, abych v dalsim kroku tusil kolik jich je treba vyhat
    lon_lowestmode <- c(lon_lowestmode, value)
    print(length(lon_lowestmode))

    value <- tryCatch(
      as.vector(replicate(n = LENGTH, b)),
      error = function(c) {NA}
    )
    beam <- c(beam, value)

    value <- tryCatch(
      as.vector(h5read(h5f, paste(b, "/quality_flag", sep="")),
        error = function(c) {NA}
      )
    )
    quality_flag <- c(quality_flag, value)

    value <- tryCatch(
      as.vector(h5read(h5f, paste(b, "/geolocation/sensitivity_a1", sep="")),
        error = function(c) {NA}
      )
    )
    sensitivity <- c(sensitivity, value)

    value <- tryCatch(
      as.vector(h5read(h5f, paste(b, "/elev_lowestmode", sep="")),
        error = function(c) {NA}
      )
    )
    elev_lowestmode <- c(elev_lowestmode, value)

    value <- tryCatch(
      as.vector(h5read(h5f, paste(b, "/elev_highestreturn", sep="")),
        error = function(c) {NA}
      )
    )
    elev_highestreturn <- c(elev_highestreturn, value)

  }

}
```

```

value <- tryCatch(
  as.vector(h5read(h5f, paste(b, "/geolocation/elev_lowestreturn_a1", sep="")),
    error = function(c) {NA})
)
elev_lowestreturn_a1 <- c(elev_lowestreturn_a1, value)

value <- tryCatch(
  as.vector(h5read(h5f, paste(b, "/degrade_flag", sep="")),
    error = function(c) {NA})
)
degrade_flag <- c(degrade_flag, value)

value <- tryCatch(
  as.vector(h5read(h5f, paste(b, "/delta_time", sep="")),
    error = function(c) {NA})
)
delta_time <- c(delta_time, value)

value <- tryCatch(
  as.vector(h5read(h5f, paste(b, "/solar_elevation", sep="")),
    error = function(c) {NA})
)
solar_elevation <- c(solar_elevation, value)

value <- tryCatch(
  as.vector(h5read(h5f, paste(b, "/surface_flag", sep="")),
    error = function(c) {NA})
)
surface_flag <- c(surface_flag, value)

value <- tryCatch(
  as.vector(h5read(h5f, paste(b, "/land_cover_data/landsat_treecover", sep="")),
    error = function(c) {NA})
)
landsat_treecover <- c(landsat_treecover, value)

value <- tryCatch(
  as.vector(h5read(h5f, paste(b, "/selected_mode_flag", sep="")),
    error = function(c) {NA})
)
selected_mode_flag <- c(selected_mode_flag, value)

value <- tryCatch(
  as.vector(h5read(h5f, paste(b, "/geolocation/stale_return_flag", sep="")),
    error = function(c) {NA})
)
stale_return_flag <- c(stale_return_flag, value)

value <- tryCatch(
  as.vector(h5read(h5f, paste(b, "/rx_1gaussfit/rx_gflag", sep="")),
    error = function(c) {NA})
)
rx_gflag <- c(rx_gflag, value)

value <- tryCatch(
  as.vector(h5read(h5f, paste(b, "/rx_assess/rx_assess_flag", sep="")),
    error = function(c) {NA})
)
rx_assess_flag <- c(rx_assess_flag, value)

value <- tryCatch(
  as.vector(h5read(h5f, paste(b, "/energy_total", sep="")),
    error = function(c) {NA})
)
energy_total <- c(energy_total, value)

value <- tryCatch(
  as.vector(h5read(h5f, paste(b, "/num_detectedmodes", sep="")),
    error = function(c) {NA})
)
num_detectedmodes <- c(num_detectedmodes, value)

#beam = c(beam, as.vector(replicate(n = length(lon_lowestmode), b))) #Priradi kazdemu radku kterej je to BEAM

value <- tryCatch(
  as.vector(h5read(h5f, paste(b, "/shot_number", sep=""), bit64conversion='double')),
    error = function(c) {NA})
)
shot_number <- c(shot_number, value)

file_name = strsplit(basename(x), "\\.[1][1] # odst?ihnu koncovku
fly_id = strsplit(file_name, "[1][4] # vyt?hnu ?as

value <- tryCatch(
  as.vector(replicate(n = LENGTH, fly_id)),
    error = function(c) {NA})
)
FLY <- c(FLY, value)
H5Fclose(h5f)
}
TEMP <- data.frame( lat_lowestmode, lon_lowestmode, beam, shot_number, FLY, quality_flag, degrade_flag,
  sensitivity, elev_lowestmode, elev_highestreturn, elev_lowestreturn_a1, delta_time,
  solar_elevation, surface_flag, landsat_treecover, selected_mode_flag, stale_return_flag,
  rx_gflag, rx_assess_flag, energy_total, num_detectedmodes)

target <- paste(basename(b), ".csv", sep = "")
print(target)
write_csv(TEMP, target)

```

Příloha číslo 2 – Skript v jazyce R sloužící k filtraci pomocí atributů, filtrací pomocí mediánu nejbližších bodů a rozdělení bodů na čtverce.

```

library(readr)
library(dplyr)
library(sf)
library(rgdal)
library(terra)
#####
#data

velky <- read_csv("D:/data_GEDI/export_clip_z/velky/finale_lite.csv")

polygons1km <- vect('D:/data_GEDI/mriz/M1km.shp')
polygons500m <- vect('D:/data_GEDI/mriz/M500m.shp')
polygons250m <- vect('D:/data_GEDI/mriz/M250m.shp')
#settings

neighbors <- c(9)
thresholds <- c(5)

# filter nastav & num_detectedmodes <= 4
model <- velky %>%
  filter((quality_flag == TRUE & rx_assess_flag == 0 & num_detectedmodes <= 4))

#####

polygons1km <- project(polygons1km, "EPSG:4326")
polygons500m <- project(polygons500m, "EPSG:4326")
polygons250m <- project(polygons250m, "EPSG:4326")
model <- st_as_sf(model, coords = c("XX", "YY"), crs = 4326)
coords <- st_coordinates(model)
results_df <- data.frame(neighbor = integer(),
                        threshold = integer(),
                        RMSE = numeric(),
                        count = integer())

for (n in neighbors) {

  # Find 10 nearest neighbors (including the point itself, so k = 11 for 10 others)
  nn <- RANN::nn2(coords, coords, k = n)

  # Calculate median height for the neighbors
  model$median_height_neighbors <- apply(nn$nn.idx, 1, function(idx) {
    median(model$elev_lowestmode[idx])
  })

  for (t in thresholds) {

    model1 <- model %>%
      filter(abs(elev_lowestmode - median_height_neighbors) <= t)

    squared_differences <- model1$diff^2
    mean_squared_error <- mean(squared_differences)
    rmse <- sqrt(mean_squared_error)

    count <- nrow(model1)
    results_df <- rbind(results_df, data.frame(neighbor = n, threshold = t, RMSE = rmse, count = count))
    print(t)
    print(n)
    print(sys.time())
    print(rmse)
  }
}
model1 <- model1[model1$diff < 30, ]
coords <- st_coordinates(model1)
model1 <- data.frame(model1, X = coords[,1], Y = coords[,2])
model1$geometry <- NULL
model1 <- vect(model1, geom=c("X", "Y"), crs="EPSG:4326")
##250m
modelint <- intersect(model1, polygons250m)
modelint <- as.data.frame(modelint, geom="XY")
point_counts <- modelint %>%
  group_by(Id) %>%
  summarise(count = n())
polygons <- as.data.frame(polygons250m)
joined <- full_join(polygons, point_counts, by = "Id")

print(count(modelint))
print(summary(joined))
print("250m")
##500m
modelint <- intersect(model1, polygons500m)
modelint <- as.data.frame(modelint, geom="XY")
point_counts <- modelint %>%
  group_by(Id) %>%
  summarise(count = n())
polygons <- as.data.frame(polygons500m)
joined <- full_join(polygons, point_counts, by = "Id")

print(count(modelint))
print(summary(joined))
print("500m")
##1km
modelint <- intersect(model1, polygons1km)
modelint <- as.data.frame(modelint, geom="XY")
point_counts <- modelint %>%
  group_by(Id) %>%
  summarise(count = n())
polygons <- as.data.frame(polygons1km)
joined <- full_join(polygons, point_counts, by = "Id")

print(count(modelint))
print(summary(joined))
print("1km")

```