Faculty of Science

Study programme: Environmental Science

doc. Ing. Vítězslav Moudrý, Ph.D.

Mapping of terrain and vegetation structure with spaceborne lidar

Mapování terénu a struktury vegetace satelitní laserovou altimetrií

Doctoral thesis

Supervisor: doc. Ing. Jan Wild, Ph.D.

Prague, 2024

PROHLÁŠENÍ

Prohlašuji, že jsem disertační práci napsal samostatně s využitím pouze uvedených a řádně citovaných pramenů a literatury. Tato práce nebyla využita k získání jiného nebo stejného titulu.

DECLARATION

I declare that I have written this thesis by myself, using only the information sources and literature cited. This thesis has not been submitted for the award of the same or any academic degree.

V Praze, září 2024 ……

Vítězslav Moudrý

Disertační práce byla vypracována ve spolupráci s Botanickým ústavem Akademie věd České republiky, v. v. i.

VEDOUCÍ PRÁCE

doc. Ing. Jan Wild, Ph.D.

Botanický ústav Akademie věd České republiky, v. v. i.

ACKNOWLEDGMENTS

I would like to thank my supervisor, Jan Wild, for giving me the opportunity to become a student once again. Thanks are also due to all those who participated in the preparation of the publications as they helped me learn a lot of new information and without them, the submitted papers would not be half as good as they are. My apologies are due to my wife and kids, as this work took me many weekends when I could not be with them.

LIST OF PUBLICATIONS

The dissertation is based on results presented in 4 scientific publications:

PUBLICATION I

Moudrý, V., Moudrá, L., Barták, V., Bejček, V., Gdulová, K., Hendrychová, M., Moravec, D., Musil, P., Rocchini, D., Šťastný, K., Volf, O. & Šálek, M. (2021). The role of the vegetation structure, primary productivity and senescence derived from airborne LiDAR and hyperspectral data for birds diversity and rarity on a restored site. *Landscape and Urban Planning*, 210, 104064.

MV processed data, performed the analysis, and led writing of the manuscript.

PUBLICATION II

Moudrý, V., Gdulová, K., Gábor, L., Šárovcová, E., Barták, V., Leroy, F., Špatenková, O., Rocchini, D. & 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.

MV processed data, performed the analysis, and led writing of the manuscript.

PUBLICATION III

Moudrý, V., Prošek, J., Šárovcová, E., Gdulová, K., Marselis, S., Kozhoridze, G., Torresani, M., Rocchini, D., Marešová, J., Eltner, A., Liu, X., Potůčková, M., Šedová, A., Crespo-Peremach, P., Torralba, J., Ruiz, L.A., Perrone, M., Špatenková, O. & Wild, J. How to find accurate terrain and canopy height GEDI footprints in temperate forests and grasslands? (under review in *Earth and Space Science*)

MV processed data, performed the analysis, and led writing of the manuscript.

PUBLICATION IV

Moudrý, V., Gábor, L., Marselis, S., Pracná, P., Barták, V., Prošek, J., Navrátilová, B., Novotný, J., Potůčková, M., Gdulová, K., Crespo-Peremarch, P., Komárek, J., Malavasi, M., Rocchini, D., Ruiz, L.A., Torralba, J., Torresani, M., Cazzolla Gatti, R. & Wild, J. Comparison of three global canopy height maps and their applicability to biodiversity modeling: Accuracy issues revealed. (Accepted in *Ecosphere*)

MV processed data, performed the analysis, and led writing of the manuscript.

PROHLÁŠENÍ O SPOLUPRÁCI NA ODBORNÝCH PUBLIKACÍCH

Jménem dalších spoluautorů prohlašuji, že se Vítězslav Moudrý podílel na vzniku výše uvedených publikací. Rozsah jeho podílu je uveden u jednotlivých prací.

DECLARATION OF COOPERATION ON SCIENCE PUBLICATIONS

On behalf of the other co-authors, I declare that Vítězslav Moudrý contributed to the preparation of the aforementioned publications. The extent of his participation in the individual papers is indicated in the List of publications.

V Praze, září 2024 …….

doc. Ing. Jan Wild, Ph.D.

ABSTRACT

Ecosystem structure, particularly the vertical structure of vegetation, is one of the six essential biodiversity variables, constituting an important aspect of habitat heterogeneity that influences species distribution and diversity. However, until relatively recently, measuring vegetation structure was extremely laborious and collecting such data was virtually impossible for large areas. However, this changed fundamentally with the advent of laser altimetry in the mid-1990s. In particular, airborne laser altimetry has seen an unprecedented boom in the last twenty years. As a result, we now have unique detailed data on vegetation structure in many countries around the world, particularly in Europe and North America. However, such data are lacking in the less developed countries, where most of the world's species diversity is found. Fortunately, recent years have favoured satellite-based laser altimeters that can fill in the gaps. Specifically, two devices were launched into Earth orbit in 2018: the Advanced Topographic Laser Altimeter System (ATLAS) onboard the ICESat-2 satellite and the Global Ecosystem Dynamics Investigation (GEDI) on the International Space Station. These devices provide, among other things, data on vegetation structure on a global scale, which can be used to improve our understanding of the distribution of species diversity on Earth.

This dissertation focuses on monitoring terrain and vegetation structure by laser altimetry and on using these data to model the diversity and distribution of animals. The airborne laser scanning data were used to model the diversity and rarity of birds on the Radovesická spoil heap in northern Bohemia. The study showed a strong positive effect of vegetation structure on both bird diversity and rarity. The next two studies have focused on satellite laser altimetry data, and airborne laser scanning data was used only as a reference to evaluate acquisition and environmental characteristics that may affect the quality of satellite altimeter data. While the first of those two studies focused on ICESat-2 data, the other analyzed GEDI data. Both studies identified several characteristics that affect the accuracy of vegetation and terrain height measurements and made key recommendations regarding data filtering for subsequent analyses. The last study examined the accuracy of recently published high-resolution global vegetation height maps derived from GEDI and Sentinel/Landsat data. The results of this study showed the limitations in the accuracy of the global vegetation height maps.

Typically, the height of low vegetation was overestimated and that of high vegetation was underestimated. Thus, their usefulness for modelling species diversity and distribution proved limited, distorting the true relationships between species diversity and environmental characteristics.

ABSTRAKT

Struktura ekosystému, zejména vertikální struktura vegetace, je jednou ze šesti základních EBV (Essential biodiversity variables) tříd a je důležitým aspektem heterogenity stanovišť, která ovlivňuje rozšíření druhů a jejich rozmanitost. Nicméně ještě relativně nedávno bylo měření struktury vegetace nesmírně pracné a sběr takových dat pro rozsáhlé oblasti byl prakticky nemožný. To se nicméně zásadně změnilo s příchodem laserové altimetrie v polovině devadesátých let minulého století. Zejména letecká laserová altimetrie zaznamenala v posledních dvaceti letech nebývalý rozmach a díky tomu máme nyní v mnoha státech světa, zejména v Evropě a Severní Americe, jedinečná data detailně reprezentující strukturu vegetace. Nicméně v méně rozvinutých státech světa, kde se nachází většina světové druhové diverzity, taková data chybí. Poslední léta naštěstí přála satelitním laserovým altimetrům, které mohou chybějící data doplnit. Konkrétně byla na oběžnou dráhu Země v roce 2018 vypuštěna dvě zařízení: ATLAS (Advanced Topographic Laser Altimeter System) na palubě satelitu ICESat-2 a GEDI (Global Ecosystem Dynamics Investigation) umístěný na mezinárodní vesmírnou stanici. Obě tato zařízení poskytují, mimo jiné, data o struktuře vegetace v globálním rozsahu, s jejichž pomocí je možné lépe pochopit rozložení diverzity druhů na Zemi.

Tato disertační práce se zaměřuje na sledování struktury vegetace metodou laserové altimetrie a na využití těchto dat pro modelování diverzity a distribuce živočichů. Data leteckého laserového skenování byla využita k modelování diverzity a rarity ptáků na Radovesické výsypce v severních Čechách. Studie prokázala silný pozitivní vliv struktury vegetace jak na diverzitu, tak raritu ptáků. Další dvě studie se již zaměřovaly na data ze satelitní laserové altimetrie a data leteckého laserového skenování byla využívána jen jako referenční pro vyhodnocení charakteristik měření a prostředí, které mohou ovlivňovat kvalitu dat ze satelitních altimetrů. Zatímco první z nich se zaměřila na data ze satelitu ICESat-2, druhá na data GEDI. Obě studie určily několik charakteristik ovlivňujících přesnost měření výšky vegetace a terénu a přinesly zásadní doporučení stran filtrace dat pro následné analýzy. Poslední studie zkoumala přesnost nedávno publikovaných globálních map výšky vegetace s vysokým rozlišením, které byly odvozeny z dat GEDI a Sentinel/Landsat. Výsledky této studie ukázaly, že globální mapy výšky vegetace mají velmi omezenou přesnost. Typicky docházelo

k nadhodnocování výšky nízké vegetace a podhodnocování výšky vysoké vegetace. Tudíž i jejich využitelnost pro modelování diverzity a distribuce druhů se ukázala jako omezená, snadno zkreslující skutečné vztahy mezi diverzitou druhů a charakteristikou prostředí.

TABLE OF CONTENTS

1. INTRODUCTION

The fine-grain global mapping of the canopy vertical structure and understanding its relationship with biodiversity patterns are among the main goals of ecological and biogeographical research (Herold et al. 2019). It is well recognised that the vertical structure of the vegetation canopy, including its height, is linked to the biodiversity of various taxa (Davies and Asner 2014; Cazzola Gatti et al. 2017). The observed positive association between the vegetation structure and species richness can be explained by two mechanisms with opposite directions. The first explanation assumes that higher and structurally more complex vegetation leads to a higher heterogeneity in temperature, relative humidity, and/or in the amount of light reaching the lower layers and hence provides more habitats for plants and animals to occupy (Davies and Asner 2014; Cazzolla Gatti et al. 2017). Besides, the vegetation structure affects habitat selection, movement, and other behaviours of animals (Russo et al. 2023). The other explanation assumes the opposite, i.e., that structural complexity is rather a consequence of plant and animal species diversity. Combinations of plant species with varying morphological traits, such as growth habit, size and crown shape, can lead to structurally denser and more complex canopies (Pretzsch 2014; Coverdale and Davies 2023), which results in an observed positive correlation between vegetation structure and tree species richness. The current view is that both mechanisms play a role and that structural complexity simultaneously facilitates species diversity as well as arises from it (Davies and Asner 2014; Hakkenberg et al. 2023; Russo et al. 2023). However, the causal mechanisms behind these relationships are not fully understood (Marselis et al. 2022). In addition, vegetation structure is modified by animals, which can, for example, consume and destroy the vegetation or disperse seeds (Russo et al. 2023). Furthermore, the availability of soil nutrients and disturbances, such as fire, storms, and human activities, may also influence vegetation structure (Hansen et al. 2014).

For a long time, the canopy vertical structure was difficult to measure and due to labour-intensive field inventory campaigns, such measurements were limited to small areas. This has changed over the past three decades with the significant improvement and expansion of laser altimetry, a.k.a. Light Detection And Ranging (LiDAR), an active remote sensing method for the acquisition of information on the terrain and canopy structure over large geographical extents (Lefsky et al. 2002;

Bergen et al. 2009). Airborne laser scanning (ALS), i.e., a LiDAR sensor onboard an airplane, has now become the primary method for collecting accurate terrain and vegetation structure over large areas (e.g., Wehr and Lohr 1999; Hudak et al. 2009). In addition, ALS data availability has increased continuously, particularly due to direct investments in data acquisition by international, national, or regional agencies (for example, many European countries provide their ALS data for free; see Melin et al. 2017; Kakoulaki et al. 2021). The increasing availability of LiDAR data resulted in their numerous applications in ecological research, and LiDAR has significantly contributed to our understanding of species-environment associations (see reviews by Davies and Asner 2014 and Bakx et al. 2019). However, ALS surveys are still relatively expensive, and such data are, therefore, lacking in the less developed countries where most of the world's species diversity is found. Consequently, we lack comprehensive global data on the spatial patterns of vegetation structure, which prevented us from analysing the relationship between vegetation structure and biodiversity patterns in some of the most diverse places on Earth, and little is also known about it at global and continental scales (but see Marselis et al. 2022 on global analysis).

Space-based remote sensing methods have a great potential to map forests globally (Herold et al. 2019; Marselis et al. 2022; Mulverhill et al. 2022). Until recently, however, these were mainly represented by passive sensors or radars, which are less suitable for mapping vegetation structure than LiDAR. Passive sensors could not provide information on the vertical structure of vegetation as they only record photons reflected from the top of the canopy (Hansen et al. 2013). Although radars are sensitive to the vertical vegetation structure, their backscatter signals are also influenced by other factors such as soil moisture, plant water content, and scan angle and direction, and disentangling all these factors from each other is difficult (Bae et al. 2019). In 2018, NASA launched two LiDAR sensors into the Earth's orbit: the Advanced Topographic Laser Altimeter System (ATLAS) onboard the ICESat-2 satellite and the Global Ecosystem Dynamics Investigation (GEDI) on the International Space Station (Markus et al. 2017; Dubayah et al. 2020). Both missions provide, among other things, data on vegetation structure on a global scale. Although LiDAR is a more suitable technology for mapping vertical vegetation structure than passive remote sensing or radar, the accuracy of retrieved data is still affected by many factors, such as atmospheric conditions,

solar background photons from sunlight, laser pulse energy level, canopy cover, or terrain slope. Evaluating the accuracy of ATLAS and GEDI retrievals will provide important insights into the usability of this novel data.

2. OBJECTIVES OF THE THESIS

The vegetation structure is a fundamental physical element of habitat, which is essential for ecological research. The availability of global vegetation structure data can, in many ways, be the key to protecting the world's diversity. The main aim of this thesis is to explore the quality and usability of newly available data from spaceborne laser altimeters and to show their potential for modelling species-environment relationships.

The individual goals of the dissertation were: **1)** to illustrate the role of vegetation structure derived from laser altimetry in species diversity and rarity on a reclaimed site; **2)** to assess the effects of acquisition and environmental conditions on ICESat-2 terrain and canopy height retrievals and suggest an approach for filtering out data with low accuracy; **3)** to evaluate the effects of acquisition and environmental characteristics on GEDI observations and quantify the success of various filtering criteria in eliminating inaccurate measurements; and **4)** to assess the accuracy of fine-grain global canopy height maps derived from GEDI and their usability for modelling species-environment relationships.

3. THEORETICAL BACKGROUND

3.1 REMOTE SENSING OF VEGETATION STRUCTURE

Vegetation structure can be estimated using various types of remote sensing techniques. The most common include synthetic aperture radar (SAR), digital aerial photogrammetry (DAP), or Light Detection And Ranging (LiDAR). All three types of sensors (radar, camera, or laser) can be mounted on various mapping platforms (tripod, drone, airplane, or satellite) (Bergen et al. 2009; Valbuena et al. 2020). Here, I will focus only on LiDAR. A key advantage of LiDAR (in comparison to DAP and SAR) lies in its ability to capture the terrain under the vegetation canopy (i.e., LiDAR pulses can penetrate gaps in the vegetation canopies; Lefsky et al. 2002). In addition, I will concentrate only on airborne and spaceborne LiDAR as possible platforms. Although drones and terrestrial (tripod) laser scanning provide greater detail and have been successfully used for modelling species-environment relationships (Shokirov et al. 2024), they can only cover small areas (e.g. Calders et al. 2020). Using an airplane or satellite as a laser sensor carrier provides coverage of relatively large study areas, which is crucial for modelling species-environment relationships, and most such studies have relied on them (Davies and Asner 2014; Bakx et al. 2019).

3.1.1 Laser altimetry (LiDAR)

Laser altimetry is an active remote sensing method that uses laser beams to measure distances between the sensor and a target surface and thus determine the positions of objects in three-dimensional space (Wehr and Lohr 1999). Lasers are used for measuring distances due to their unique properties, such as coherence and the ability to emit a large number of photons in a defined direction in very short pulses at a predefined wavelength (Shan and Toth 2018). The pulse length is typically around 5-10 ns (Wagner et al. 2004) and the distance from the device to the object is determined from the time difference between sending and receiving the pulse (Wehr and Lohr 1999).

The available LiDAR sensors can be distinguished based on the way the received signal is recorded (discrete return, full waveform, or single photon), platform type (e.g., airborne, or spaceborne), the size of the laser footprint (small or large) and sample scanning pattern (profiling or scanning). The most commonly used LiDAR systems have been small-footprint discrete return LiDAR and large-footprint full

waveform LIDAR. A relatively new system, single photon sensitive LiDAR, is currently gaining in popularity and competing with the two above-mentioned laser scanning systems. Its advantage lies in the potentially higher area coverage performance, which, however, comes at the price of a lower ranging accuracy and an increased outlier rate (Mandlburger et al. 2019).

Small-footprint LiDAR scanners have footprints smaller than 1 meter in diameter and record 1-5 discrete heights at peak returns of each pulse of the laser. These are typically mounted on airborne platforms (Wehr and Lohr 1999). Note, however, that modern small-footprint LiDAR can record full waveform as well. Large-footprint full waveform LiDAR sensors have footprints ranging from 10 to 100 m in diameter and record a full digitised reflected waveform (i.e. continuous height distribution of surfaces illuminated by each laser pulse). These are found mainly on spaceborne platforms (Dubayah et al. 2020). Small-footprint LiDAR systems operating on airborne platforms are characterised by high pulse repetition frequency, creating a very dense point cloud (i.e. irregular distribution of the returns in three-dimensional space). In vegetated areas, the laser beams are usually reflected by several layers of vegetation (Lefsky et al. 2002). The interaction of the laser beam with the canopy is then characterised by multiple returns from different depths in the vegetation. The first return usually comes from the vegetation canopy surface, followed by the intermediate returns from leaves and branches, with the last return indicating ideally a return from the ground (Figure 1). On the other hand, large-footprint systems operating on spaceborne platforms are characterised by a lower pulse repetition frequency. Therefore, spaceborne laser altimeters, such as ICESat, ICESat-2, or GEDI provide greater coverage, but sparse discrete measurements compared to ALS (Markus et al. 2017; Dubayah et al. 2020).

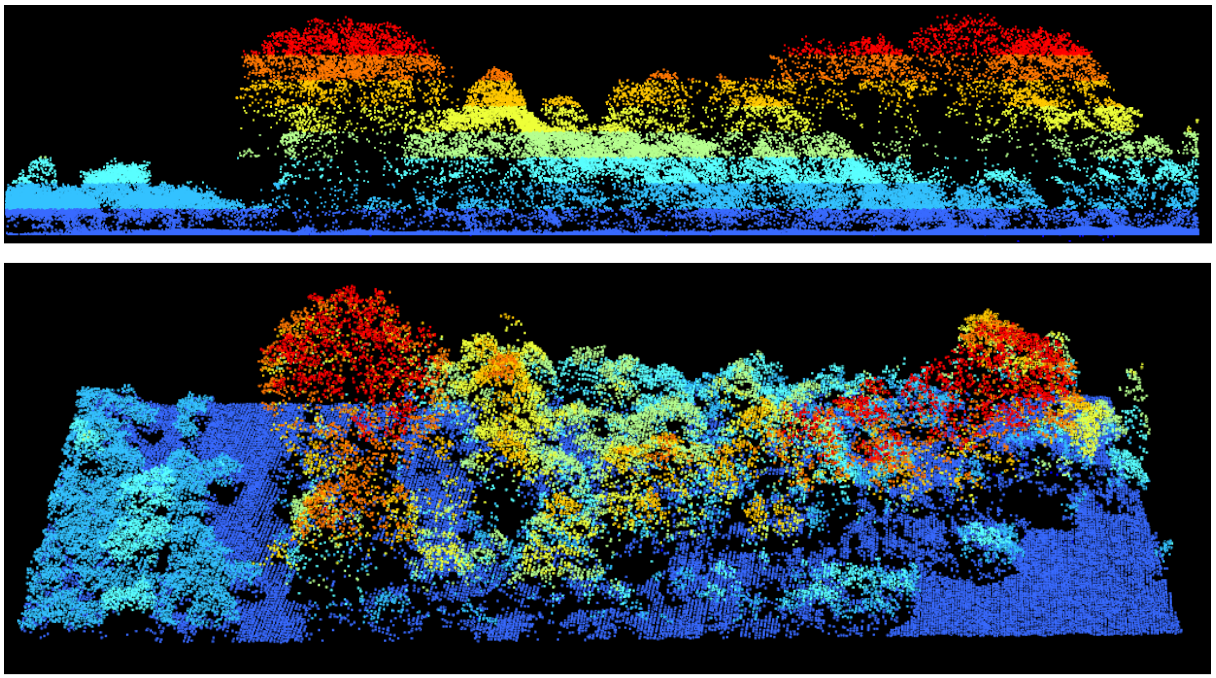


Figure 1. Example of point cloud. Profile (top) and 3D view (bottom) Dark blue indicates terrain, red indicates high vegetation.

3.1.2 Spaceborne LiDAR missions

Ice, Cloud, and land Elevation Satellite (ICESat)

The first global three-dimensional data characterising vegetation structure based on satellite laser altimeter measurements was obtained from the Geoscience Laser Altimeter System (GLAS) onboard the ICESat satellite, which was in operation between 2003 and 2009. The GLAS laser altimeter transmitted short pulses (4 nanoseconds) of infrared light (1,064 nanometres wavelength) and visible green light (532 nanometres). The laser pulses had a repetition frequency of 40 pulses per second, with footprints approximately 70 meters in diameter, spaced 170 meters apart along the Earth's surface. The horizontal accuracy of the ground footprints is about 6 m, and the vertical accuracy is below 3 cm (Abshire et al. 2005). However, the relatively low vertical accuracy in topographically complex terrain where there were problems distinguishing terrain from vegetation was a drawback of the measurements (Chen 2010). Although the main objective of the mission was to measure polar ice caps, a total of 15 datasets were produced, overlapping across scientific disciplines and also including data on ground elevation and vegetation structure (Schutz et al. 2005). The knowledge gained during the

ICESat mission was essential in the preparation of two other recently launched missions.

Ice, Cloud, and land Elevation Satellite-2 (ICESat-2)

In 2018, NASA launched the Ice, Cloud and land Elevation Satellite-2 (ICESat-2) mission. ICESat-2 operates in a 91-day exact repeat orbit, with 1,387 orbits per cycle. The ATLAS instrument onboard the ICESat-2 satellite is a single-photon counting laser altimeter that uses green laser beams (wavelength of 532 nm) to measure the distance to the surface (Markus et al. 2017). The benefit of using a single-photon approach is that the laser can operate at a much higher repetition rate. In the case of ATLAS, the pulse repetition frequency is 10 kHz (compare it with 40Hz of its predecessor GLAS), which allows along-track sampling at 0.7 m intervals, with overlapping footprints of ~12 m. The ATLAS instrument splits the output laser pulse into three pairs of beams that are arranged to produce ground tracks with a distance of 3.3 km between the pairs and of 90 m between the tracks of each pair. The pairs consist of one weak and one strong beam with an energy ratio of approximately 1:4 (Neumann et al. 2019). Theoretically, ATLAS can detect up to 16 photons (4 × 4 detector array) per outgoing shot (a negligible fraction of all emitted photons; Markus et al. 2017; Neumann et al. 2019).

The low-level data products include ATL01 and 02 (where photon times of flight are computed) and end with ATL03, which provides a latitude, longitude and elevation for each photon. The ATL03 serves as a single source of all photon data and ancillary information needed by surface-specific higher-level products such as land ice, sea ice, the atmosphere, vegetation and land, oceans and inland water. The terrain and canopy heights are provided as a part of the land and vegetation height product (ATL08; Neuenschwander et al. 2021). This information is subsequently used to produce gridded datasets, such as gridded ground surface height, canopy height and canopy cover estimates (ATL18).

Global Ecosystem Dynamics Investigation (GEDI)

The Global Ecosystem Dynamics Investigation (GEDI) mission aimed to measure forest structure and biomass in tropical and temperate regions. It was the first spaceborne LiDAR mission specifically designed for such a purpose (Dubayah et al. 2020). The GEDI started operations on March 25, 2019, and collected data until March 2023, much longer than the originally planned two years. In addition,

instead of being decommissioned and destined to burn in the Earth's atmosphere as originally planned, GEDI is stored on ISS for up to 18 months and its reinstalling at a later point to continue the vegetation mapping is planned.

The instrument was composed of three full-waveform lasers, each with a 15.6 ns pulse and a 242 Hz pulse repetition rate in the near-infrared region (1,064 nm). Two of these were used at full power (15 mJ/pulse), while the third one was split into two coverage beams of lower energy (4.5 mJ/pulse). The laser pulses were subjected to dithering, which resulted in a total of eight tracks separated by 600 m. Each pulse had a 25 m footprint, and along the track, the footprints were separated by 60 m. GEDI was limited to the area between 51.6° N and 51.6° S (Dubayah et al. 2020).

The geolocated waveforms of received energy (i.e. the number of photons) as a function of time (GEDI L1B product) are the fundamental observations made by the GEDI instrument. The returned waveform is typically multimodal with the lowest mode representing ground elevation. The accuracy of ground elevation is crucial for the estimation of canopy height metrics, as the ground serves as an elevation baseline for the canopy height estimation (Hofton and Blair 2020). The geolocated waveforms are further processed to provide products at the footprint level: the L2A and L2B data products. The L2A product provides ground elevation, canopy height, and relative height metrics representing the height at which a particular quantile of energy (i.e., the 1^{st}–100^{th} quantile) was returned, relative to the elevation of the lowest waveform mode representing the ground (Hofton and Blair 2020). The L2B product provides canopy cover and plant area index. This information is subsequently used to produce gridded datasets: the L3 (mean canopy height map and the standard deviation of canopy height map), and L4B (gridded above-ground biomass) datasets, at a 1 km^2 resolution; the best detail that has ever been produced at the near-global extent (Dubayah et al. 2020).

3.1.3 GEDI and ICESat-2 data limitations

The GEDI and ICESat-2 measurements are not without errors, and the presence of invalid or outlaying measurements has been reported for both missions (Adam et al. 2020; Tian and Shan 2021). Despite some construction differences between the two instruments (single photon vs. full waveform LiDAR), the factors affecting measurement accuracy are the same: (i) The presence of clouds affects the signal

strength and the number of photons returning from the surface. (ii) Atmospheric layers such as clouds can cause multiple scattering, which increases the photon path length (Winker 2003). Consequently, the quality of data acquired under several layers of clouds may be insufficient for terrain and vegetation characterisation (Queinnec et al. 2021). This limitation is particularly pronounced in regions with persisting cloud cover such as the tropics (Potapov et al. 2021). (iii) In temperate forests, data accuracy is affected by snow cover, which hinders the detection of the actual surface and results in underestimation of canopy height (Neuenschwander et al. 2020). (iv) The accuracy is also affected by terrain slope (Tian and Shan 2021; Liu et al. 2021), particularly in combination with (v) a high canopy cover (Malambo and Popescu 2021). (vi) Night acquisitions have higher accuracy due to the lower concentration of background noise photons from sunlight reflected off the surface (Neuenschwander et al. 2020; Adam et al. 2020). (vii) Finally, both instruments split the laser pulse into several beams of different energies, with the higher energy beams typically providing more accurate data (Liu et al. 2021). Despite the clear need for appropriate filtering of the laser altimetry data before their further use (Carabajal and Boy 2020), there is still no consensus about how such filtering should be performed to minimise the error while maximising the amount of usable data.

ICESat-2 and GEDI products provide several flags derived from auxiliary data that describe the target (e.g. surface reflectance) and acquisition (e.g. atmospheric scattering) characteristics and inform users about the conditions under which the data were acquired. We will use these auxiliary data together with external reference data (e.g., land cover, canopy cover, terrain slope) to assess lidar-environment interactions. An improved understanding of LiDAR-environment interactions will facilitate the selection of accurate data for subsequent analyses.

3.2 SPECIES DISTRIBUTION MODELLING

Understanding the interactions of species with their environment is fundamental to predicting species distribution patterns and habitat use, and thus to improving biodiversity conservation and management. Species distribution modelling (SDM) is a rapidly evolving field in biogeography and spatial ecology used for such purposes (Araújo et al. 2019). SDM encompasses two quite distinct lines of research (Ferrier et al. 2017). The first, 'explanatory modelling', aims to explain the relationships between a biodiversity-related response variable (such as the distribution of individual species) and the explanatory variables. Particularly, early studies of species–environment relationships focused on measuring environmental variables at species observation points. From this, relationships were inferred that could, however, not be used to predict species occurrence at other locations where these explanatory variables were not available. The other line of research is 'predictive modelling', which aims to predict unknown values of the biodiversity response variable based on pre-specified relationships. The use of predictive models has increased markedly with the availability of remote sensing data (e.g., Cord et al. 2013). Predictive SDM is especially useful in supporting conservation decision-making, such as in selecting protected areas; identifying critical habitats that contain essential features for endangered species conservation; or predicting the impacts of climate or land use change on biodiversity (Guisan et al. 2013).

3.2.1 Vegetation Structure as the key habitat characteristic

The selection of appropriate explanatory variables is a crucial part of SDM. The chosen variables should adequately represent the main factors affecting species' distributions, such as climate, land cover, or topography (Gardner et al. 2019; Santini et al. 2021) and should be tailored to the species' ecology and habitat requirements.

Habitat heterogeneity is one of the most important factors affecting species distributions and diversity. It is determined by the variability of environmental conditions (e.g., habitat types, species dominance and composition, vegetation density, soil types, or topographic variability). According to the habitat heterogeneity hypothesis, more complex environments can provide more niches and, thus, increase species diversity (see reviews by Tews et al. 2004; Stein et al. 2014). A fundamental physical element of habitat heterogeneity is vegetation

structure (i.e., the vertical distribution of biomass in ecosystems; MacArthur and MacArthur 1961), which is one of the 20 essential biodiversity variables (EBVs) defined by GEO BON (Group on Earth Observations Biodiversity Observation Network) and belongs under the EBV class Ecosystem Structure (https://geobon.org/ebvs/what-are-ebvs/).

Several studies have shown that vegetation structure derived from laser altimetry data can serve as a useful proxy for habitat heterogeneity (Lefsky et al. 2002; Vierling et al. 2008; Davies and Asner 2014; Vogeler and Cohen 2016; Bakx et al. 2019). The pioneering studies mainly focused on investigating (i.e., demonstrating) the effectiveness of laser altimetry-derived variables in describing species–environment associations (Hinsley et al. 2002; Hill et al. 2004; Bradbury et al. 2005; Goetz et al. 2007). Since then, the focus of exploratory studies has shifted to assessing relationships between vegetation structure and the distribution of individual species (e.g., Graf et al. 2009; Sillero and Goncalves-Seco 2014), and species diversity (e.g., Clawges et al. 2008; Lesak et al. 2011). Several studies explored differences in the applicability of laser altimetry-derived variables with respect to different functional guilds (e.g., nesting, foraging, and habitat), showing that the importance of individual variables as well as the predictability of species occurrence using vegetation structure differ between guilds (e.g., Goetz et al. 2007; Weisberg et al. 2014; Cooper et al. 2020).

3.2.2 Vegetation structure variables used in SDM

Countless variables can be derived from LiDAR data, and it would be pointless to try to introduce them all. However, I can introduce the major ones repeatedly used in many studies (Figure 2; see also reviews by Davies and Asner 2014 and Bakx et al. 2019).

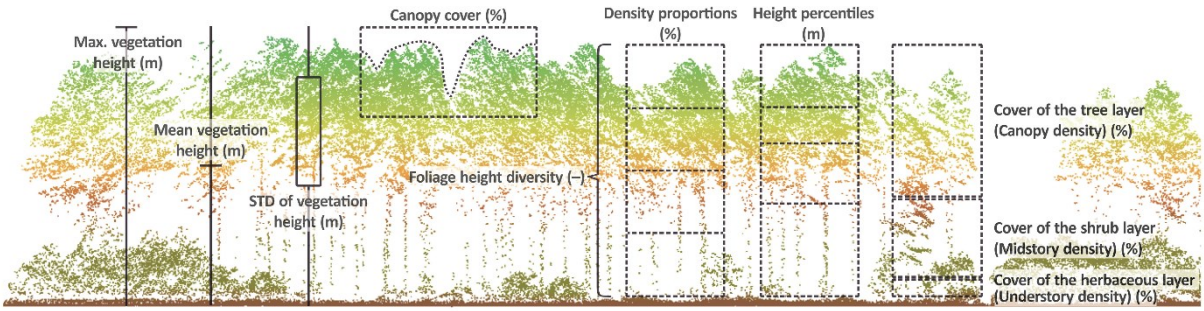


Figure 2. Example of a LiDAR point cloud profile showing common variables used to describe the vertical structure of vegetation.

The maximum, mean, and standard deviation of vegetation returns are among the simplest metrics for describing the vertical structure of vegetation. Tree height is a useful indicator of tree diameter and, to some extent, of tree age, which are important factors for species diversity. Hence, maximum height should be particularly useful in predicting the occurrence of species associated with mature, old-growth forests. For example, a positive relationship between the maximum canopy height and species richness has been reported for birds (Flaspohler et al. 2010; Lesak et al. 2011) as well as vascular plants (Mao et al. 2018). The mean height of vegetation returns is particularly useful in combination with variability in return heights. For example, Vogeler et al. (2013) found a positive relationship between the mean height, vertical variability of vegetation, and Brown Creeper (Certhia americana) occupancy. Similarly, Aguirre-Gutiérrez et al. (2017) found that butterfly diversity increased with average vegetation height and vertical variability of vegetation. Together with vegetation height, vertical height variability reflects key structural differences between land cover and habitat types and is important for their differentiation (Koma et al. 2021). For example, lower mean height and high vertical variability could indicate forests with sparse canopy and dense understorey vegetation.

Vertical height variability is often characterised by a single variable (e.g., the standard deviation of vegetation returns or foliage height diversity index based on the Shannon-Wiener index; MacArthur and MacArthur, 1961). For example, Weisberg et al. (2014) reported positive associations between bird species richness and foliage height diversity. On the other hand, Vogeler et al. (2014) found that foliage height diversity was not a strong predictor of bird species richness. This could be due to the fact that these individual variables may not fully capture the complex layering of vegetation. Characterising the vertical vegetation profile by multiple characteristics at different vegetation layers may be more useful. For example, height percentiles and density proportions describe vertical vegetation structure in more detail. Height percentiles indicate the height (in meters) below which a certain percentage of returns has been recorded (e.g., Eldegard et al. 2014). For instance, if the 70th height percentile is 10 m, it means that the lowest 70% of the vegetation returns are below 10 m. Density proportions reflect the proportion of points within a certain height bin to the total number of returns (e.g., Lesak et al. 2011). This can be calculated as fixed height bins (e.g. 5 m) between

the minimum and maximum height or within different vegetation layers (e.g., understorey, mid-storey, canopy). Such vertical stratification describes the presence of different age classes or life forms existing at certain heights (e.g., herbs, shrubs, and trees). In particular, understorey vegetation in forests (Clawges et al. 2008; Vogeler et al. 2014) and shrub vegetation such as hedgerows in agricultural landscapes (Pelletier-Guittier et al. 2020) are often considered important factors for species richness, as they provide nesting and foraging habitats, affect visibility and prey abundance, and alter the near-surface microclimate (Stickley and Fraterrigo, 2021). On the other hand, vegetation can also act as an obstacle, and it has been shown that forest-dwelling aerial insectivores, such as some bird species or bats, prefer forests without a shrub layer as an optimal foraging habitat (Lesak et al. 2011; Rauchenstein et al. 2022). Similarly, Torre et al. (2022) have shown that the diversity of Mediterranean small mammal communities is negatively affected by the structural complexity of vegetation. On the other hand, herbaceous and shrub vegetation serve as important refuge for wildlife in landscapes heavily influenced by humans, such as agricultural and urban areas (Choi et al. 2021; Melin et al. 2018). Canopy cover can serve as a proxy for light availability on the ground. Open canopy stands are often associated with dense understorey layers. Closed canopies, on the other hand, buffer microclimatic conditions such as temperature and moisture content (Davis et al. 2019; De Frenne et al. 2019).

3.3 GLOBAL CANOPY HEIGHT MAPS

An important principal limitation of spaceborne LiDAR sensors is that they collect data along discrete transects and, hence, only provide discrete samples of vegetation height and structure. For example, the GEDI mission's expected coverage of the Earth's surface is only around 4%. Relatively sparse data limit the detail at which vegetation characteristics can be provided on a global scale. Consequently, the produced global maps of vegetation structure are of coarse resolution. Just for the record, Hancock et al. (2021) estimated that with the current in-orbit technology, twelve satellites would be required to produce a continuous map of the canopy structure at a 30 m resolution every 5 years.

The density of GEDI sampling enables the production of a mean canopy height map and standard deviation of canopy height map at a 1 km resolution almost globally within -52° and 52° latitude directly from the measurements (Figure 3).

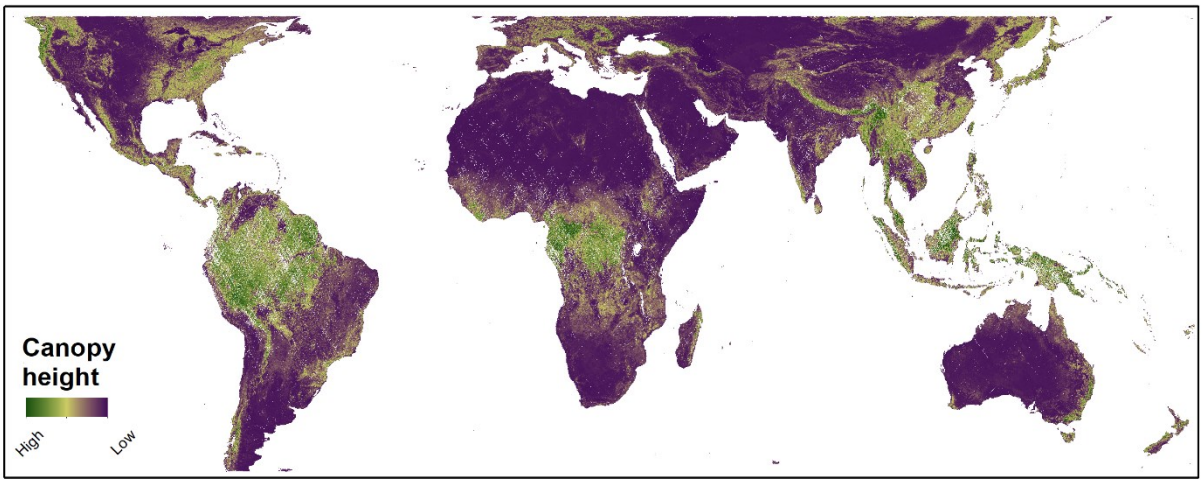


Figure 3. Near global canopy height at 1 km resolution – GEDI L3 Gridded Land Surface Metrics, Version 2.

While the spaceborne laser altimetry missions do not provide continuous coverage, their data can still be used to model canopy height in greater detail and over larger areas. Statistical techniques and machine learning algorithms can be used to interpolate and extrapolate GEDI data to create wall-to-wall estimates of canopy height and other related parameters. The currently common approach to produce continuous global maps of vegetation structure at a finer resolution (e.g. 10 m or 30 m) is to train models that combine direct measurements (e.g. ICESat, GEDI, ICESat-2) with spatially continuous ancillary data (e.g. Sentinel-2, Landsat)

enabling the estimation of the desired variable at locations not directly measured by GEDI or ICESat (Bergen et al. 2009; Lefsky, 2010). In this manner, Lefsky (2010) combined canopy heights derived from GLAS (Geoscience Laser Altimeter System instrument on board the first ICESat mission) with MODIS data to produce a global patch-based canopy height map. Similarly, Simard et al. (2011) used relationships between GLAS-derived canopy heights and multiple environmental variables (e.g., tree cover, climate, altitude) to derive a global map of canopy height at a 1 km spatial resolution. GLAS data were also used for global mapping of other important structural variables, such as canopy cover and leaf area index (Tang et al. 2014; Tang et al. 2019). More recently, Potapov et al. (2021) and Lang et al. (2023) used optical data (Landsat, Sentinel-2) to extrapolate the GEDI measurements, creating global canopy height maps at 30 m and 10 m spatial resolutions, respectively. Despite its unprecedented resolution, however, these products may suffer from low accuracy. Obviously, predictions are not as accurate as the directly measured data.

4. RESULTS AND DISCUSSION

Thanks to the continuous development in remote sensing technologies, vegetation structure data has become readily available, particularly as point clouds from airborne laser scanning data in developed countries (e.g., Europe, North America, Australia, New Zealand; Kakoulaki et al. 2021). In addition, global data from satellite laser altimetry missions, such as ICESat-2 (Markus et al. 2017) and GEDI (Dubayah et al. 2020) became available in recent years. Furthermore, recent improvements in machine and deep learning methods enabled the creation of predictive models to produce global maps of various environmental variables, including vegetation structure metrics (e.g., Lang et al. 2023). These maps are usually readily available as open data, allowing other researchers to use them as input data for calculation indicators of all kinds or as input data for mapping other variables. However, the reliability of predicted global maps is increasingly questioned, particularly with respect to the often non-realistically high accuracy estimates provided by the authors resulting, for example, from the use of inappropriate validation strategies or non-representative reference data (Duncanson et al. 2019; Meyer & Pebesma 2022).

This dissertation offers novel insights into the quality and usability of several such datasets. It includes research on the role of vegetation structure in species diversity and rarity, the accuracy of vegetation height retrievals from two recent spaceborne laser altimetry missions, and the usability of predicted global canopy height maps for modelling species diversity and distribution. Specifically, the titles of the studies are as follows:

Study I - The role of the vegetation structure, primary productivity and senescence derived from airborne LiDAR and hyperspectral data for birds diversity and rarity on a restored site

Study II - Effects of environmental conditions on ICESat-2 terrain and canopy heights retrievals in Central European mountains

Study III - How to find accurate terrain and canopy height GEDI footprints in temperate forests and grasslands?

Study IV - Comparison of three global canopy height maps and their applicability to biodiversity modeling: Accuracy issues revealed.

As full discussion is included in the respective studies, they will not be repeated here. Instead, this part aims to provide what did not fit in the articles, links among the individual presented studies, major points worth highlighting, and suggestions for future research.

4.1 Usability of vegetation structure metrics for environmental monitoring

The importance of vegetation structure for species diversity has been repeatedly assessed and acknowledged, particularly for birds (see reviews by Davies and Asner 2014 and Bakx et al. 2019). However, to the best of my knowledge, **Study I** is the first to do such an assessment at a reclaimed site, which was highlighted also by one of the reviewers. One of the most useful aspects of this study is that we combined LiDAR with field data and showed the usability of LiDAR data for assessing restoration success. In addition, we suggested that combining forest restoration with sites left to spontaneous succession appears to be the best strategy to increase bird species richness and rarity in newly restored sites after coal mining. Someone may object that such a result is not groundbreaking, and several studies have reached similar conclusions, even without sophisticated remote sensing techniques (e.g., Šálek 2012; Vojar et al. 2016; Korejs et al. 2023). However, our study's value is in identifying vegetation structure metrics that can be easily derived from ALS data and used as a more elegant and more precisely measurable alternative to simple and subjective field surveys, which can be very useful in future assessments of restoration success.

Not long ago, the use of remote sensing for restoration success assessment was considered less developed than in other disciplines (Cordell et al. 2017). Remote sensing technologies represent efficient and cost-effective sources for developing indicators relevant to large-scale decision-making. Their indispensability is particularly evident if we need to monitor the change of the environment over time and, therefore, require repeated measurements. Although a direct field survey of habitat attributes can provide valuable information, such studies are rare (but see Korejs et al. 2023) as repeated monitoring is labour-intensive. In addition, data from field surveys are typically used for a single study and then, if not lost, stored in some local computer. Although this has been improving in recent years thanks to the push for open science, and much of the data is now available through various open repositories (e.g. Powers and Hampton 2019), navigating someone else's

data, the structure of which is not standardized (i.e., not fully following FAIR principles), is often as challenging as the data collection itself.

Remote sensing data is usually available to all through online repositories and adheres to FAIR principles, providing a standardized, and permanent record of the state of the environment. Repeated data collection by means of remote sensing can provide crucial insights into the dynamics of restored ecosystems. In my future research, I would like to concentrate on monitoring the temporal changes in vegetation structure on spoil heaps and other areas of interest, such as national parks. As airborne laser scanning data availability increases, so does availability from different time points. For example, scanning of selected spoil heaps in northern Bohemia was carried out in the years 2011 and 2017 and is planned for 2024. Spaceborne laser altimetry missions may constitute another interesting input in this respect. For example, Milenković et al. (2022) used GEDI and ICESat-2 data to assess forest regrowth in the Amazon rainforest. Similarly, Guerra-Hernández and Pascual (2021) and Parra and Simard (2023) combined GEDI with ALS data to estimate forest height growth dynamics in Spain and Québec, respectively. We, however, warn against using predicted global canopy height maps at fine resolution for such purposes. In **Study IV**, we tested the accuracy and applicability of such data for species-environment relationship assessment. We found that their accuracy is low, and their use in local scale modelling leads to erroneous results, especially when horizontal habitat heterogeneity is of concern. Until these models are improved, we recommend that such studies use original measured data at relatively coarse resolution rather than predicted fine-resolution maps.

4.2 Effect of environmental and acquisition characteristics on ICESat-2 and GEDI retrievals

Studies II and **III** mainly focused on environmental and acquisition characteristics that may affect the quality of ICESat-2 and GEDI retrievals. Here, by retrievals, I mean estimates of both terrain and canopy height. Although the factors affecting the accuracy of their measurements may vary slightly, the accuracy of vegetation height determination is primarily determined by the accuracy of the ground elevation. Therefore, it is possible to discuss the factors for both variables simultaneously. Despite some construction differences between GEDI and ICESat-2 instruments (single photon vs. full waveform LiDAR), the

factors potentially affecting the measurement accuracy are the same: atmospheric conditions (i.e., the presence of clouds affecting the strength of the signal), solar background photons (i.e., sunlight reflected off the Earth's surface), and terrain slope, particularly in combination with a high canopy cover.

For both sensors, the beam strength is the most important acquisition characteristic. **Studies II and III** showed that the beams with higher energy (i.e. strong and power beams in the case of ICESat-2 and GEDI, respectively) are more likely to return, providing higher accuracy estimates of terrain and canopy cover. In addition, the results of **Study II** show that the number of returned photons from ICESat-2 declines considerably with increasing cloud cover. Furthermore, the scattering of photons in dense clouds increases the photon path length, making the surface appear lower than it actually is (Winkler 2003). We did not evaluate atmospheric effects on GEDI retrievals in **Study III**, as it was shown earlier that atmospheric conditions considerably impact GEDI data quality. For example, Fayad et al. (2021) showed that GEDI footprints acquired in the presence of clouds resulted in as much as 69% of acquired footprints being unusable. The accuracy of the retrievals also depends on whether they are taken during the daytime or at nighttime. In **Study II**, we showed that the accuracy of ICESat-2 acquisitions was higher during the nighttime than daytime, but we did not find any effect of time of acquisition (day or night) on GEDI data accuracy (**Study III**). Similarly, Liu et al. (2021) concluded that the accuracy of the GEDI data acquired during the day and at night is almost identical, whereas the accuracy of ICESat-2 data acquired during the day is lower than that of the nighttime data. This difference is due to the different technology used for the two missions. The single-photon laser used for ICESat-2 uses a Hamamatsu photomultiplier detector, which is very sensitive and, hence, also sensitive to solar background noise. The full waveform LiDAR used for GEDI uses the Si:APD detector, which has a higher noise threshold than the solar background noise and is, therefore, less sensitive to day/night differences.

The slope is another environmental parameter with a considerable impact on the accuracy of both ICESat-2 and GEDI retrievals. Results of **Study II** show that in the case of ICESat-2, the magnitude of this negative effect increases with the decreasing number of signal photons, while in the case of GEDI, this error is related to a relatively large footprint and tendency to identify a higher number of modes with increasing slope (**Study III**). In addition, horizontal displacement (i.e.,

geolocation accuracy) may further amplify the error in the accuracy of retrievals caused by the effect of slope itself. The horizontal displacement reported for ICESAT-2 tends to be 0-3 m (Neuenschwander et al. 2020; Malambo and Popescu 2021), while for GEDI, it is up to 10 m (Quiros et al. 2021; Tang et al. 2023).

The accuracy of terrain and canopy height retrievals was also affected by canopy cover. **Study II** showed that the best accuracy of the canopy height from ICESat-2 was observed for the canopy cover ranging from 40% to 60%. Higher canopy cover leads to a decreased number of ground photons; hence, the accuracy of terrain estimates is low and, consequently, so are the canopy height estimates. On the other hand, when the canopy cover is low, photons are likely reflected from its lower parts or not reflected from the canopy at all, and the canopy height is underestimated. In the case of GEDI, only footprints with sensitivities ranging from 0.9 to 1 are considered valid waveforms and used. The sensitivity of an acquired beam is a signal detection performance metric that indicates the probability that the LiDAR beam was able to penetrate through a given canopy cover (i.e. 0.9 – 1) and reach the terrain. However, in **Study III,** we showed that footprints with lower sensitivity values may also be used for areas with low canopy cover, such as grasslands.

4.3 Availability of large-scale canopy height maps

The availability of globally consistent data on vegetation structure and height, such as products from the GEDI mission, is crucial for ecological research. The availability of GEDI data greatly improves the detail of information on vertical and horizontal vegetation structure, allowing studying questions regarding species-environment relationships on a near-global scale and at much finer resolutions than has been possible to date. Integrating vegetation structure with data on biodiversity opens unprecedented opportunities for groundbreaking research on underlying mechanisms driving biodiversity patterns and forest functioning (e.g., Russo et al. 2023). Although GEDI data are still novel with few applications in biogeography, their availability has already sparked important research that showed GEDI's potential as well the fact that there is an important knowledge gap on the role of vegetation structure for biodiversity (Marselis et al. 2022). The GEDI data have been, for example, used to assess the role of forest structure in biodiversity patterns (Torresani et al. 2023), to improve models of animal-environment relationships (Smith et al. 2022), or assess protected area's

effectiveness in conserving vegetation structure and carbon stocks (Ceccherini et al. 2023; Liang et al. 2023).

My original intention was to use machine learning methods to develop a continuous canopy height model at fine resolution for the Czech Republic. However, the first experiments were not very successful and, meanwhile, several other authors developed relatively useful local or global models (e.g., Potapov et al. 2021). Therefore, rather than developing my own model, I decided to evaluate the quality and usability of existing models in **Study IV**. On the one hand, increasing data availability facilitates ecological research; on the other hand, however, the increasing availability of datasets of varying quality also brings pitfalls as their users will be confused about which of the available datasets to use. Unfortunately, most users do not have the capacity to verify the dataset accuracy independently. Even worse, many will not even think about it and use the first product they come across. As an example, Lewis et al. (2022) calculated the mean and variance of canopy height from the 2019 Global Canopy Forest Height database (Potapov et al., 2021), even though more accurate airborne laser scanning data for Georgia were available for the same period (opentopography.org). This example underlines the importance of providing independent validations. As more global products for forest height and biomass are becoming available to users, more independent evaluations are needed to assess the suitability of CHM products for given purposes (e.g., Pascual et al. 2022). In **Study IV**, we evaluated the accuracy of three global canopy height maps and assessed their suitability for modelling species distribution. Our results showed the presence of large errors in the predicted global canopy height maps and, therefore, proved that their usability for modelling species-environment relationships is limited.

Despite these criticisms, some recent studies have been able to achieve better results in predicting vegetation height at local scales. For example, Schwartz et al. (2023) developed canopy height models using GEDI and Sentinel data for France. They showed that their model considerably outperformed available global canopy height maps, which might be thanks to the improved ability of their approach to filter out unsuitable GEDI footprints. Future studies using predictive models to generate canopy height maps should combine GEDI and ICESat-2 data, and use filtering approaches proposed in **Studies II and III** to select high-quality retrievals.

Finally, I call for improved availability of accurate canopy height maps from airborne laser scanning data. While the 3DEP programme is underway in the United States (Stoker and Miller 2022) and vegetation height will soon be readily available for the whole of that country, the data in Europe are fragmented between different national providers, in different coordinate systems, and its use for the whole continent is practically impossible at the moment. To realise its full potential, data must be easily available in a form that can be accessed by users with average GIS experience.

5. CONCLUSIONS

In this thesis, I explored the accuracy of newly available data from spaceborne laser altimeters and their potential for modelling species-environment relationships. Data from ICESat-2 and GEDI missions are now intensively used for large-scale canopy height mapping, and my findings can considerably improve the selection of accurate data for such purposes. Using machine or deep learning methods that combine spaceborne laser altimeters and optical remote sensing data to produce fine-resolution canopy height maps is a popular approach. These experiments resulted, among others, in three global canopy height maps (A high-resolution canopy height model of the Earth; Global Forest Canopy Height; Global Map of Tree Canopy Height). I assessed the accuracy of these maps and showed that their accuracy is poor, considerably limiting their usability for modelling species-environment relationships. I emphasize that their use in local scale modelling and biodiversity studies may lead to erroneous results, especially when horizontal habitat heterogeneity is of concern. Nevertheless, because these models are readily available in a raster format, I am concerned that there is a risk that even in areas where more accurate ALS data are available, users will prefer global products to the tedious processing of much more accurate ALS data. It is, therefore, advisable that in states with ALS data available, authoritative institutions should take responsibility for the provision of vegetation structural variables in a raster format. However, I believe that more accurate models will replace the current ones and that vegetation structure variables will have a major positive impact on ecological modelling, just like bioclimatic variables have had in the past.

6. REFERENCES

Abshire, J. B., Sun, X., Riris, H., Sirota, J. M., McGarry, J. F., Palm, S., ... & Liiva, P. (2005). Geoscience laser altimeter system (GLAS) on the ICESat mission: on-orbit measurement performance. *Geophysical research letters*, *32*(21).

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.

Aguirre-Gutiérrez, J., WallisDeVries, M. F., Marshall, L., van't Zelfde, M., Villalobos-Arámbula, A. R., Boekelo, B., ... & Biesmeijer, J. C. (2017). Butterflies show different functional and species diversity in relationship to vegetation structure and land use. *Global Ecology and Biogeography*, *26*(10), 1126-1137.

Araújo, M. B., Anderson, R. P., Márcia Barbosa, A., Beale, C. M., Dormann, C. F., Early, R., ... & Rahbek, C. (2019). Standards for distribution models in biodiversity assessments. *Science advances*, *5*(1), eaat4858.

Bae, S., Levick, S. R., Heidrich, L., Magdon, P., Leutner, B. F., Wöllauer, S., ... & Müller, J. (2019). Radar vision in the mapping of forest biodiversity from space. *Nature Communications*, *10*(1), 4757.

Bakx, T. R., Koma, Z., Seijmonsbergen, A. C., & Kissling, W. D. (2019). Use and categorization of light detection and ranging vegetation metrics in avian diversity and species distribution research. *Diversity and Distributions*, *25*(7), 1045-1059.

Bergen, K. M., Goetz, S. J., Dubayah, R. O., Henebry, G. M., Hunsaker, C. T., Imhoff, M. L., ... & Radeloff, V. C. (2009). Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. *Journal of Geophysical Research: Biogeosciences*, *114*(G2).

Bradbury, R. B., Hill, R. A., Mason, D. C., Hinsley, S. A., Wilson, J. D., Balzter, H., ... & Bellamy, P. E. (2005). Modelling relationships between birds and vegetation structure using airborne LiDAR data: a review with case studies from agricultural and woodland environments. *Ibis*, *147*(3), 443-452.

Calders, K., Adams, J., Armston, J., Bartholomeus, H., Bauwens, S., Bentley, L. P., ... & Verbeeck, H. (2020). Terrestrial laser scanning in forest ecology: Expanding the horizon. *Remote Sensing of Environment*, *251*, 112102.

Carabajal, C. C., & Boy, J. P. (2020). ICESat-2 altimetry as geodetic control. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *43*, 1299-1306.

Cazzolla Gatti, R., Di Paola, A., Bombelli, A., Noce, S., & Valentini, R. (2017). Exploring the relationship between canopy height and terrestrial plant diversity. *Plant ecology*, *218*, 899-908.

Ceccherini, G., Girardello, M., Beck, P. S., Migliavacca, M., Duveiller, G., Dubois, G., ... & Cescatti, A. (2023). Spaceborne LiDAR reveals the effectiveness of

European Protected Areas in conserving forest height and vertical structure. *Communications Earth & Environment*, *4*(1), 97.

Chen, Q. (2010). Retrieving vegetation height of forests and woodlands over mountainous areas in the Pacific Coast region using satellite laser altimetry. *Remote Sensing of Environment*, *114*(7), 1610-1627.

Choi, H., Song, Y., Kang, W., Thorne, J. H., Song, W., & Lee, D. K. (2021). LiDAR-derived three-dimensional ecological connectivity mapping for urban bird species. *Landscape Ecology*, *36*, 581-599.

Clawges, R., Vierling, K., Vierling, L., & Rowell, E. (2008). The use of airborne lidar to assess avian species diversity, density, and occurrence in a pine/aspen forest. *Remote sensing of environment*, *112*(5), 2064-2073.

Cooper, W. J., McShea, W. J., Forrester, T., & Luther, D. A. (2020). The value of local habitat heterogeneity and productivity when estimating avian species richness and species of concern. *Ecosphere*, *11*(5), e03107.

Cord, A. F., Meentemeyer, R. K., Leitão, P. J., & Václavík, T. (2013). Modelling species distributions with remote sensing data: bridging disciplinary perspectives. *Journal of Biogeography*, *40*(12), 2226-2227.

Cordell, S., Questad, E. J., Asner, G. P., Kinney, K. M., Thaxton, J. M., Uowolo, A., ... & Chynoweth, M. W. (2017). Remote sensing for restoration planning: how the big picture can inform stakeholders. *Restoration Ecology*, *25*, S147-S154.

Coverdale, T. C., & Davies, A. B. (2023). Unravelling the relationship between plant diversity and vegetation structural complexity: A review and theoretical framework. *Journal of Ecology*, *111*(7), 1378-1395.

Davies, A. B., & Asner, G. P. (2014). Advances in animal ecology from 3D-LiDAR ecosystem mapping. *Trends in ecology & evolution*, *29*(12), 681-691.

Davis, K. T., Dobrowski, S. Z., Holden, Z. A., Higuera, P. E., & Abatzoglou, J. T. (2019). Microclimatic buffering in forests of the future: the role of local water balance. *Ecography*, *42*(1), 1-11.

De Frenne, P., Zellweger, F., Rodríguez-Sánchez, F., Scheffers, B. R., Hylander, K., Luoto, M., ... & Lenoir, J. (2019). Global buffering of temperatures under forest canopies. *Nature Ecology & Evolution*, *3*(5), 744-749.

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., ... & Silva, C. (2020). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of remote sensing*, *1*, 100002.

Duncanson, L., Armston, J., Disney, M., Avitabile, V., Barbier, N., Calders, K., ... & Williams, M. (2019). The importance of consistent global forest aboveground biomass product validation. *Surveys in geophysics*, *40*, 979-999.

Eldegard, K., Dirksen, J. W., Ørka, H. O., Halvorsen, R., Næsset, E., Gobakken, T., & Ohlson, M. (2014). Modelling bird richness and bird species presence in a boreal forest reserve using airborne laser-scanning and aerial images. *Bird Study*, *61*(2), 204-219.

Fayad, I., Baghdadi, N., & Riedi, J. (2021). Quality assessment of acquired gedi waveforms: Case study over france, tunisia and french guiana. *Remote Sensing*, *13*(16), 3144.

Ferrier, S., Jetz, W., & Scharlemann, J. (2017). Biodiversity modelling as part of an observation system. *The GEO handbook on biodiversity observation networks*, 239-257.

Flaspohler, D. J., Giardina, C. P., Asner, G. P., Hart, P., Price, J., Lyons, C. K. A., & Castaneda, X. (2010). Long-term effects of fragmentation and fragment properties on bird species richness in Hawaiian forests. *Biological conservation*, *143*(2), 280-288.

Gardner, A. S., Maclean, I. M., & Gaston, K. J. (2019). Climatic predictors of species distributions neglect biophysiologically meaningful variables. *Diversity and Distributions*, *25*(8), 1318-1333.

Goetz, S., Steinberg, D., Dubayah, R., & Blair, B. (2007). Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. *Remote Sensing of Environment*, *108*(3), 254-263.

Goetz, S. J., Steinberg, D., Betts, M. G., Holmes, R. T., Doran, P. J., Dubayah, R., & Hofton, M. (2010). Lidar remote sensing variables predict breeding habitat of a Neotropical migrant bird. *Ecology*, *91*(6), 1569-1576.

Graf, R. F., Mathys, L., & Bollmann, K. (2009). Habitat assessment for forest dwelling species using LiDAR remote sensing: Capercaillie in the Alps. *Forest Ecology and Management*, *257*(1), 160-167.

Guerra-Hernández, J., & Pascual, A. (2021). Using GEDI lidar data and airborne laser scanning to assess height growth dynamics in fast-growing species: a showcase in Spain. *Forest Ecosystems*, *8*, 1-17.

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I., ... & Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology letters*, *16*(12), 1424-1435.

Hakkenberg, C. R., Atkins, J. W., Brodie, J. F., Burns, P., Cushman, S., Jantz, P., ... & Goetz, S. J. (2023). Inferring alpha, beta, and gamma plant diversity across biomes with GEDI spaceborne lidar. *Environmental Research: Ecology*, *2*(3), 035005.

Hancock, S., McGrath, C., Lowe, C., Davenport, I., & Woodhouse, I. (2021). Requirements for a global lidar system: spaceborne lidar with wall-to-wall coverage. *Royal Society open science*, *8*(12), 211166.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., ... & Townshend, J. R. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, *342*(6160), 850-853.

Hansen, A. J., Phillips, L. B., Dubayah, R., Goetz, S., & Hofton, M. (2014). Regional-scale application of lidar: Variation in forest canopy structure across the southeastern US. Forest Ecology and Management, 329, 214-226.

Herold, M., Carter, S., Avitabile, V., Espejo, A. B., Jonckheere, I., Lucas, R., ... & De Sy, V. (2019). The role and need for space-based forest biomass-related measurements in environmental management and policy. *Surveys in Geophysics*, *40*, 757-778.

Hill, R. A., Hinsley, S. A., Gaveau, D. L. A., & Bellamy, P. E. (2004). Cover: predicting habitat quality for Great Tits (Parus major) with airborne laser scanning data.

Hinsley, S. A., Hill, R. A., Bellamy, P. E., & Balzter, H. (2006). The application of Lidar in woodland bird ecology. *Photogrammetric Engineering & Remote Sensing*, *72*(12), 1399-1406.

Hofton, M., Blair, J. B., Story, S., & Yi, D. (2020). Algorithm Theoretical Basis Document (ATBD). *Washington, DC, USA: NASA*.
Hudak, A. T., Evans, J. S., & Stuart Smith, A. M. (2009). LiDAR utility for natural resource managers. *Remote Sensing*, *1*(4), 934-951.

Kakoulaki, G., Martinez, A., & Florio P. (2021). *Non-commercial Light Detection and Ranging (LiDAR) data in Europe, EUR 30817 EN*. Publications Office of the European Union, Luxembourg. ISBN 978-92-76-41150-5, doi:https://doi.org/10.2760/212427, JRC126223.

Koma, Z., Seijmonsbergen, A. C., & Kissling, W. D. (2021). Classifying wetland-related land cover types and habitats using fine-scale lidar metrics derived from country-wide Airborne Laser Scanning. *Remote Sensing in Ecology and Conservation*, *7*(1), 80-96.

Korejs, K., Šálek, M., Bejček, V., Musil, P., Šťastný, K., Volf, O., & Riegert, J. (2024). Nine-year bird community development on Radovesická spoil heap: impacts of restoration approach and vegetation characteristics. *Landscape and Ecological Engineering*, *20*(1), 89-102.

Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2023). A high-resolution canopy height model of the Earth. *Nature Ecology & Evolution*, *7*(11), 1778-1789.

Lefsky, M. A., Cohen, W. B., Parker, G. G., & Harding, D. J. (2002). Lidar remote sensing for ecosystem studies: Lidar, an emerging remote sensing technology that directly measures the three-dimensional distribution of plant canopies, can accurately estimate vegetation structural attributes and should be of particular interest to forest, landscape, and global ecologists. *BioScience*, *52*(1), 19-30.

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).

Lesak, A. A., Radeloff, V. C., Hawbaker, T. J., Pidgeon, A. M., Gobakken, T., & Contrucci, K. (2011). Modeling forest songbird species richness using LiDAR-derived measures of forest structure. *Remote Sensing of Environment*, *115*(11), 2823-2835.

Lewis, W. B., Chandler, R. B., Delancey, C. D., Rushton, E., Wann, G. T., McConnell, M. D., & Martin, J. A. (2022). Abundance and distribution of ruffed grouse Bonasa umbellus at the southern periphery of the range. *Wildlife Biology*, *2022*(5), e01017.

Liang, M., González-Roglich, M., Roehrdanz, P., Tabor, K., Zvoleff, A., Leitold, V., ... & Duncanson, L. (2023). Assessing protected area's carbon stocks and ecological structure at regional-scale using GEDI lidar. *Global Environmental Change*, *78*, 102621.

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.

MacArthur, R. H., & MacArthur, J. W. (1961). On bird species diversity. *Ecology*, *42*(3), 594-598.

Malambo, L., & Popescu, S. C. (2021). Assessing the agreement of ICESat-2 terrain and canopy height with airborne lidar over US ecozones. *Remote Sensing of Environment*, *266*, 112711.

Mandlburger, G., Lehner, H., & Pfeifer, N. (2019). A comparison of single photon and full waveform lidar. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *4*, 397-404.

Mao, L., Dennett, J., Bater, C. W., Tompalski, P., Coops, N. C., Farr, D., ... & Nielsen, S. E. (2018). Using airborne laser scanning to predict plant species richness and assess conservation threats in the oil sands region of Alberta's boreal forest. *Forest ecology and management*, *409*, 29-37.

Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., ... & 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.

Marselis, S. M., Keil, P., Chase, J. M., & Dubayah, R. (2022). The use of GEDI canopy structure for explaining variation in tree species richness in natural forests. *Environmental Research Letters*, *17*(4), 045003.

Melin, M., Shapiro, A., & Glover-Kapfer, P. (2017). *Lidar for ecology and conservation. WWF Conservation Technology Series* 1(3), WWF-UK, Woking, United Kingdom.

Melin, M., Hinsley, S. A., Broughton, R. K., Bellamy, P., & Hill, R. A. (2018). Living on the edge: utilising lidar data to assess the importance of vegetation structure for avian diversity in fragmented woodlands and their edges. *Landscape Ecology*, *33*, 895-910.

Meyer, H., & Pebesma, E. (2022). Machine learning-based global maps of ecological variables and the challenge of assessing them. *Nature Communications*, *13*(1), 2208.

Milenković, M., Reiche, J., Armston, J., Neuenschwander, A., De Keersmaecker, W., Herold, M., & Verbesselt, J. (2022). Assessing Amazon rainforest regrowth with GEDI and ICESat-2 data. *Science of Remote Sensing*, *5*, 100051.

Mulverhill, C., Coops, N. C., Hermosilla, T., White, J. C., & Wulder, M. A. (2022). Evaluating ICESat-2 for monitoring, modeling, and update of large area forest canopy height products. *Remote Sensing of Environment*, *271*, 112919.

Neuenschwander, A., Pitts, K., Jelley, B., Robbins, J., Markel, J., Popescu, S., ... & Klotz, B. (2019). Ice, Cloud, and Land Elevation Satellite 2 (ICESat-2) algorithm theoretical basis document (ATBD) for land-vegetation along-track products (ATL08). *National Aeronautics and Space Administration: Washington, DC, USA*.

Neuenschwander, A., Guenther, E., White, J. C., Duncanson, L., & Montesano, P. (2020). Validation of ICESat-2 terrain and canopy heights in boreal forests. *Remote Sensing of Environment*, *251*, 112110.

Neumann, T. A., Martino, A. J., Markus, T., Bae, S., Bock, M. R., Brenner, A. C., ... & 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.

Parra, A., & Simard, M. (2023). Evaluation of Tree-Growth Rate in the Laurentides Wildlife Reserve Using GEDI and Airborne-LiDAR Data. *Remote Sensing*, *15*(22), 5352.

Pascual, A., Tupinambá-Simões, F., & de Conto, T. (2022). Using multi-temporal tree inventory data in eucalypt forestry to benchmark global high-resolution canopy height models. A showcase in Mato Grosso, Brazil. *Ecological Informatics*, *70*, 101748.

Pelletier-Guittier, C., Théau, J., & Dupras, J. (2020). Use of hedgerows by mammals in an intensive agricultural landscape. *Agriculture, ecosystems & environment*, *302*, 107079.

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A., ... & Hofton, M. (2021). Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sensing of Environment*, *253*, 112165.

Powers, S. M., & Hampton, S. E. (2019). Open science, reproducibility, and transparency in ecology. *Ecological applications*, *29*(1), e01822.

Pretzsch, H. (2014). Canopy space filling and tree crown morphology in mixed-species stands compared with monocultures. *Forest Ecology and Management*, *327*, 251-264.

Queinnec, M., White, J. C., & Coops, N. C. (2021). Comparing airborne and spaceborne photon-counting LiDAR canopy structural estimates across different boreal forest types. *Remote Sensing of Environment*, *262*, 112510.

Rauchenstein, K., Ecker, K., Bader, E., Ginzler, C., Düggelin, C., Bontadina, F., & Obrist, M. K. (2022). LiDAR metrics predict suitable forest foraging areas of endangered mouse-eared bats (Myotis myotis). *Forest Ecology and Management*, *515*, 120210.

Russo, N. J., Davies, A. B., Blakey, R. V., Ordway, E. M., & Smith, T. B. (2023). Feedback loops between 3D vegetation structure and ecological functions of animals. *Ecology Letters*, *26*(9), 1597-1613.

Šálek, M. (2012). Spontaneous succession on opencast mining sites: implications for bird biodiversity. *Journal of Applied Ecology*, *49*(6), 1417-1425.

Santini, L., Benítez-López, A., Maiorano, L., Čengić, M., & Huijbregts, M. A. (2021). Assessing the reliability of species distribution projections in climate change research. *Diversity and Distributions*, *27*(6), 1035-1050.

Schutz, B. E., Zwally, H. J., Shuman, C. A., Hancock, D., & DiMarzio, J. P. (2005). Overview of the ICESat mission. *Geophysical research letters*, *32*(21).

Shan, J., & Toth, C. K. (Eds.). (2018). *Topographic laser ranging and scanning: principles and processing*. CRC press.

Shokirov, S., Jucker, T., Levick, S. R., Manning, A. D., & Youngentob, K. N. (2024). Using multiplatform LiDAR to identify relationships between vegetation structure and the abundance and diversity of woodland reptiles and amphibians. *Remote Sensing in Ecology and Conservation*.

Sillero, N., & Goncalves-Seco, L. (2014). Spatial structure analysis of a reptile community with airborne LiDAR data. *International Journal of Geographical Information Science*, *28*(8), 1709-1722.

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research: Biogeosciences*, *116*(G4).

Smith, A. B., Vogeler, J. C., Bjornlie, N. L., Squires, J. R., Swayze, N. C., & Holbrook, J. D. (2022). Spaceborne LiDAR and animal-environment relationships: An assessment for forest carnivores and their prey in the Greater Yellowstone Ecosystem. *Forest Ecology and Management*, *520*, 120343.

Stein, A., Gerstner, K., & Kreft, H. (2014). Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology letters*, *17*(7), 866-880.

Stickley, S. F., & Fraterrigo, J. M. (2021). Understory vegetation contributes to microclimatic buffering of near-surface temperatures in temperate deciduous forests. *Landscape Ecology*, *36*, 1197-1213.

Stoker, J., & Miller, B. (2022). The accuracy and consistency of 3d elevation program data: A systematic analysis. *Remote Sensing*, *14*(4), 940.

Tang, H., Dubayah, R., Brolly, M., Ganguly, S., & Zhang, G. (2014). Large-scale retrieval of leaf area index and vertical foliage profile from the spaceborne waveform lidar (GLAS/ICESat). *Remote Sensing of Environment*, *154*, 8-18.

Tang, H., Armston, J., Hancock, S., Marselis, S., Goetz, S., & Dubayah, R. (2019). Characterizing global forest canopy cover distribution using spaceborne lidar. *Remote Sensing of Environment*, *231*, 111262.

Tang, H., Stoker, J., Luthcke, S., Armston, J., Lee, K., Blair, B., & Hofton, M. (2023). Evaluating and mitigating the impact of systematic geolocation error on canopy height measurement performance of GEDI. *Remote Sensing of Environment*, *291*, 113571.

Tews, J., Brose, U., Grimm, V., Tielbörger, K., Wichmann, M. C., Schwager, M., & Jeltsch, F. (2004). Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. *Journal of biogeography*, *31*(1), 79-92.

Tian, X., & Shan, J. (2021). Comprehensive evaluation of the ICESat-2 ATL08 terrain product. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(10), 8195-8209.

Torre, I., Jaime-González, C., & Díaz, M. (2022). Habitat suitability for small mammals in mediterranean landscapes: How and why shrubs matter. *Sustainability*, *14*(3), 1562.

T
orresani, M., Rocchini, D., Alberti, A., Moudrý, V., Heym, M., Thouverai, E., ... & Tomelleri, E. (2023). LiDAR GEDI derived tree canopy height heterogeneity reveals patterns of biodiversity in forest ecosystems. *Ecological Informatics*, *76*, 102082.

Valbuena, R., O'Connor, B., Zellweger, F., Simonson, W., Vihervaara, P., Maltamo, M., ... & Coops, N. C. (2020). Standardizing ecosystem morphological traits from 3D information sources. *Trends in Ecology & Evolution*, *35*(8), 656-667.

Vierling, K. T., Vierling, L. A., Gould, W. A., Martinuzzi, S., & Clawges, R. M. (2008). Lidar: shedding new light on habitat characterization and modeling. *Frontiers in Ecology and the Environment*, *6*(2), 90-98.

Vogeler, J. C., Hudak, A. T., Vierling, L. A., & Vierling, K. T. (2013). Lidar-derived canopy architecture predicts brown creeper occupancy of two western coniferous forests. *The Condor*, *115*(3), 614-622.

Vogeler, J. C., Hudak, A. T., Vierling, L. A., Evans, J., Green, P., & Vierling, K. T. (2014). Terrain and vegetation structural influences on local avian species richness in two mixed-conifer forests. *Remote Sensing of Environment*, *147*, 13-22.

Vogeler, J. C., & Cohen, W. B. (2016). A review of the role of active remote sensing and data fusion for characterizing forest in wildlife habitat models. *Revista de Teledetección*, 1-14.

Vojar, J., Doležalová, J., Solský, M., Smolová, D., Kopecký, O., Kadlec, T., & Knapp, M. (2016). Spontaneous succession on spoil banks supports amphibian diversity and abundance. *Ecological Engineering*, *90*, 278-284.

Wagner, W., Ullrich, A., Melzer, T., Briese, C., & Kraus, K. (2004). *From single-pulse to full-waveform airborne laser scanners: Potential and practical challenges* (Vol. 35). na.

Wehr, A., & Lohr, U. (1999). Airborne laser scanning—an introduction and overview. *ISPRS Journal of photogrammetry and remote sensing*, *54*(2-3), 68-82.

Weisberg, P. J., Dilts, T. E., Becker, M. E., Young, J. S., Wong-Kone, D. C., Newton, W. E., & Ammon, E. M. (2014). Guild-specific responses of avian species richness to LiDAR-derived habitat heterogeneity. *Acta Oecologica*, *59*, 72-83.

Winker, D. M. (2003, April). Accounting for multiple scattering in retrievals from space lidar. In *12th International Workshop on Lidar Multiple Scattering Experiments* (Vol. 5059, pp. 128-139). SPIE.

7. ATTACHED PUBLICATIONS

The role of the vegetation structure, primary productivity and senescence derived from airborne LiDAR and hyperspectral data for birds diversity and rarity on a restored site

Vítězslav Moudrý, Lucie Moudrá, Vojtěch Barták, Vladimír Bejček, Kateřina Gdulová, Markéta Hendrychová, David Moravec, Petr Musil, Duccio Rocchini, Karel Šťastný, Ondřej Volf, Miroslav Šálek

Landscape and Urban Planning, doi:
https://doi.org/10.1016/j.landurbplan.2021.104064

Landscape and Urban Planning 210 (2021) 104064

Contents lists available at ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

Research Paper

The role of the vegetation structure, primary productivity and senescence derived from airborne LiDAR and hyperspectral data for birds diversity and rarity on a restored site

Vítězslav Moudrý [a,b,*], Lucie Moudrá [a], Vojtěch Barták [a], Vladimír Bejček [c], Kateřina Gdulová [a], Markéta Hendrychová [d], David Moravec [a], Petr Musil [c], Duccio Rocchini [a,e], Karel Šťastný [c], Ondřej Volf [f], Miroslav Šálek [c]

[a] Department of Spatial Sciences, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Praha – Suchdol, 165 00, Czech Republic
[b] Institute for Environmental Studies, Faculty of Science, Charles University, Benátská 2, 12801 Prague 2, Czech Republic
[c] Department of Ecology, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Praha – Suchdol, 165 00, Czech Republic
[d] Department of Landscape and Urban Planning, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Praha – Suchdol, 165 00, Czech Republic
[e] Alma Mater Studiorum University of Bologna, Department of Biological, Geological and Environmental Sciences, via Irnerio 42, 40126 Bologna, Italy
[f] Spolek Ametyst, Nebílovy 37, 332 04 Nebílovy, Czech Republic

HIGHLIGHTS

- We tested the role of vegetation structure, NDVI and PSRI for bird diversity.
- Shrub cover and tree cover had strong positive effects on bird richness.
- The PSRI, shrub cover and herbaceous cover had positive effects on bird rarity.
- Heterogeneous vertical vegetation structure promotes bird richness and rarity.
- Combining forests with spontaneous succession will balance richness and rarity.

ARTICLE INFO

Keywords:
Mining
NDVI
PSRI
Senescence
Spoil heap
Vegetation structure

ABSTRACT

Management of restored areas requires ecologically meaningful spatial data providing objective measures of restoration success. Understanding relationships between species diversity on the one hand and habitat heterogeneity and productivity on the other can help establish such measures and prioritize restoration management. We used airborne LiDAR and hyperspectral data to derive characteristics of vegetation structure, primary productivity and senescent vegetation (i.e. old dead vegetation) for prediction of richness and rarity of bird communities colonizing newly available habitats restored after coal mining. In addition, we analysed, which type of restoration (i.e. agricultural, forest, or spontaneous succession) results in more favourable conditions. The boosted regression trees explained 52% and 12% of deviance of overall species richness and rarity, respectively. We found that the overall species richness was strongly affected by the variance in vegetation structure, while the rarity was also affected by the presence of senescent vegetation. The relative importance of variables differed between the richness and rarity. The shrub cover had a strong positive effect on both, while the tree cover had a strong positive effect on species richness. The herbaceous cover and presence of senescent vegetation had positive effects on species rarity. This study, therefore, supports the necessity to create a mosaic of habitats with heterogeneous vertical structure including all layers of vegetation and highlights the importance of senescent vegetation. Combination of forests restoration with sites left to spontaneous succession appears to be the best strategy to increase both bird species richness and rarity in newly restored sites after coal mining.

* Corresponding author at: Department of Spatial Sciences, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Praha – Suchdol, 165 00, Czech Republic.
E-mail address: moudry@fzp.czu.cz (V. Moudrý).

https://doi.org/10.1016/j.landurbplan.2021.104064
Received 3 June 2020; Received in revised form 25 January 2021; Accepted 30 January 2021

1. Introduction

The mining of various raw materials has expanded globally in the last few decades with the growing demands for various commodities and their increasing consumption (Kung et al., 2020; Lèbre et al., 2020). On the one side, economic benefits and wealth generated by the resource industry are substantial (Svobodova, Owen, Harris, & Worden, 2020); on the other side, areas under mining lease are subject to significant landscape changes (Worlanyo & Jiangfeng, 2020). The changes are often associated with negative environmental impacts, including irreversible land degradation and biodiversity loss (Giam, Olden, & Simberloff, 2018; Osenberg, 2018). However, under certain circumstances, mining and related activities can bring about also positive changes enhancing conservation value of the landscape (Šálek, Hendrychová, & Řehoř, 2010; Schulz & Wiegleb, 2000; Vanhée & Devigne, 2018).

Minimization of the negative effects of mining is typically ensured by ecological restoration, i.e. a process of assisting the recovery of an ecosystem that has been degraded, damaged, or destroyed (Clewell & Aronson, 2013; Martins et al., 2020). Monitoring of restored sites is required to gather ecologically meaningful data that can provide objective and quantitative measures of the restoration success. In practice, among other ecological measures, species diversity of various taxa is frequently used and many studies evaluated the effect of habitat heterogeneity on the diversity of species on restored sites (Crouzeilles, Ferreira, Chazdon, Lindenmayer, Sansevero, Monteiro, & Strassburg, 2017; Martins et al., 2020). To model relationships between the species diversity and the heterogeneity of restored sites, however, the environment is usually represented by semiquantitative or categorical measures only (e.g. by rough subjective estimates of vegetation cover). Moreover, even these measures are usually spatially and temporarily limited as field surveys traditionally used by ecologists are extremely labour-intensive, especially over larger areas (Bejček, 1988; Gould & Mackey, 2015; Hagger, Wilson, England, & Dwyer, 2019; Kolář, Tichanek, & Tropek, 2017; Vojar et al., 2016). Although a direct field survey of habitat attributes can provide valuable information, it is unsuitable for repeated monitoring due to both labour intensiveness and limited informative value of subjective estimates, especially where detailed habitat characteristics are concerned.

In this study, we aim to identify measures that can be easily derived from airborne remote sensing data and used as a more elegant and more precisely measurable alternative to simple and subjective field surveys. Airborne remote sensing data are increasingly available from national or regional scanning campagins (Melin, Shapiro, & Glover-Kapfer, 2017; Stereńczak et al., 2020). Such technologies represent efficient and cost-effective sources for developing indicators relevant to the large-scale decision making, to the understanding of continuous processes of site restoration, and developing effective management tools that will maintain high biodiversity of restored sites (e.g., Cordell et al., 2017; Laurin et al., 2020; Prošek et al., 2020; Urban, Štroner, Křemen, Braun, & Möeser, 2018). In addition, conservation strategies for post-industrial sites are highly debated in connection with restoration approaches (Hendrychová, Svobodova, & Kabrna, 2020), including the adopted restoration method (Tropek et al., 2010; Vicentini, Hendrychova, Tajovský, Pižl, & Frouz, 2020; Vymazal & Sklenicka, 2012). Therefore, we relate evaluated indicators to the individual restoration methods and provide recommendations for restoration practice.

In the next chapter, the theoretical background of the research and ecological explanation for the selection of environmental variables derived from airborne remote sensing data will be discussed and the aims of this study will be clearly expressed. Chapter 3 introduces the study area, the data and its preprocessing, the environmental variables and statistical analyses including models evaluation. Chapters 4 and 5 then present the obtained results followed by their discussion with respect to restoration goals and informing on possible limitations. Finally, the Chapter 6 contains conclusions and recommendations for resoration practice.

2. State of the art

Birds with their high dispersal ability play an important role in the early colonization of restored sites and therefore comprise one of the best indicators for the assessment of the restoration success (Bejček & Šťastný, 1984; Cardoso da Silva & Vickery, 2002; Martins et al., 2020). One of the important factors affecting bird species richness is the habitat heterogeneity. Habitat heterogeneity is determined by the variability of environmental conditions (e.g. habitat type, dominant vegetation species, soil types, topography) and it is assumed that more complex environments may provide more niches and thus increase species diversity (so-called habitat heterogeneity hypothesis; see review by Tews et al., 2004).

A common approach of indicating heterogeneity of the habitat is to use the variability in its physical structure (physiognomy) (Davies & Asner, 2014). Physiognomy of the habitat is generally determined by plants and the debate whether bird species diversity is more affected by vegetation structure or by plant composition is still ongoing (Adams & Matthews, 2019; MacArthur & MacArthur, 1961; Müller, Stadler, & Brandl, 2010). Although some studies have shown the importance of plant composition and it is clear that it should not be ignored (Adams & Matthews, 2019), the vegetation structure has been traditionally considered the primary driver of bird diversity (Müller et al., 2010). This may be partly due to intensive research addressing relationships between bird richness and vegetation structure, which has been triggered by advances in the measurement of the vegetation structure by airborne LiDAR (see reviews by Davies & Asner, 2014; Bakx, Koma, Seijmonsbergen, & Kissling, 2019).

Soon after the first studies showed the potential of LiDAR-derived vegetation structure for explaining species-environment associations, attempts begun to integrate LiDAR with variables derived using other complementary remote sensing data (e.g. multispectral or hyperspectral) assessing their relative importance and complementarity (Bae et al., 2018; Cooper, McShea, Forrester, & Luther, 2020; Goetz, Steinberg, Dubayah, & Blair, 2007; Vogeler et al., 2014). Such variables include, for example, the normalized difference vegetation index (NDVI; Tucker, 1979). The use of NDVI to model bird species richness is based on species-energy theory. According to that theory, species richness is limited by the quantity of available energy (Brown, 1981; Wright, 1983) and energy available to consumers is dependent on primary productivity (Evans, Warren, & Gaston, 2005). It is assumed that greater primary productivity of plants (i.e. biomass) supports higher animal species richness and NDVI is commonly used as a measure of vegetation productivity (Bailey et al., 2004; Hobi et al., 2017; Leyequien et al., 2007; Youngentob, Yoon, Stein, Lindenmayer, & Held, 2015).

On the other hand, however, many species including birds are specialists and/or poor competitors, preferring specific habitats with relatively lower habitat heterogeneity or low primary productivity (Reif, Hořák, Krištín, Kopsová, & Devictor, 2016). The occurrence of these species can be associated with early stages of spontaneous succession, which are rare in the cultural landscape but relatively common on restored sites (Šálek, 2012). These early successional habitats can, therefore, represent valuable refuges for rare and unique bird specialists and/or poor competitors.

For the early successional habitats in areas after coal mining, growths with old dead vegetation from previous vegetation season are typical (Hendrychová et al., 2020). This is particularly true for aquatic vegetation in areas left to spontaneous succession (e.g. *Phragmites australis* and *Typha latifolia*) but a similar representation of such old (dead) vegetation can be also observed in low steppe vegetation in agriculturally restored areas (e.g. *Calamagrostis epigejos* and *Arrhenatherum elatius*). For birds, the dead vegetation provides shelter, nesting, and foraging opportunities at the time when green vegetation is at minimal heights. It is, therefore, an important component of habitat heterogeneity not only in the autumn and winter but possibly even more so in the spring. The amount of old dead vegetation (i.e. senescent vegetation)

can be estimated using the plant senescing reflectance index (PSRI), which was developed as a quantitative measure of leaf senescence (Merzlyak, Gitelson, Chivkunova, & Rakitin, 1999) and used as a potential predictor of species occurrence in studies published previously (e.g. Soto, Pérez-Hernández, Hahn, Rodewald, & Vergara, 2017).

Determining the attributes of the habitat heterogeneity and primary productivity of restored sites and their association with species diversity and/or rarity can help to identify ecologically valuable areas on large landscape scales. Therefore, in this study, we examined how the vegetation structure, primary productivity (i.e. NDVI) and old senescent vegetation from the previous season (i.e. PSRI) derived from airborne laser scanning and hyperspectral data predict species richness and rarity of bird communities colonizing newly available (restored) habitats after coal mining. Specifically, we focused on the following questions: (i) Is there a detectable relationship between the fine-scale habitat attributes of early succession stages obtained by airborne remote sensing and the occurrence of birds? (ii) Which is of greater importance for the occurrence of birds – the primary productivity, presence of old vegetation, or vegetation structure? And (iii) which type of restoration (i.e. agricultural, forest, or spontaneous succession) results in the development of more favourable habitats for bird species richness and rarity?

3. Data and methods

3.1. Study area, type of reclamation and habitats

The study was carried out on the Radovesická spoil heap (Fig. 1) located in the North Bohemian Brown Coal Basin, Czechia, one of the largest active brown coal mining regions in Europe. The study area was subject to various methods of restoration (i.e. agricultural, forestry, and spontaneous succession; Fig. 2). Agricultural reclamations typically include the establishment of permanent grasslands with initial sowing of a species-poor grasses mixture (*Festuca, Dactylis, Phleum, Poa, Cynosurus, Agrostis*) mixed with about 10% of legumes (*Trifolium, Coronilla, Lotus, Medicago*). Such areas are mowed twice a year even after the reclamation is completed. Afforestation includes predominantly homogenous plantations of even-aged stand combining autochtonous and allochthonous trees (*Acer, Pupulus, Quercus, Fraxinus, Tilia, Carpinus, Larix*) supplemented with shrubs (*Eonymus, Padus, Lingustrum, Cornus, Symphoricarpos, Spiraea, Lonicera, Viburnum*). Successional sites in our study area are characterized by structurally diversified bare ground with sparse annuals and biennials, followed by perennials (*Tanacetum, Artemisia, Cirsium*) and grasses (especially *Calamagrostis epigejos* and *Arrhenatherum elatius*) with scattered authochtonous shrubs (*Sambucus, Rosa, Betula, Crataegus*). The tree growths on the successional sites are dominated by birch (*Betula pendula*), mixed with other deciduous trees (*Salix, Populus*). See Fig. S1 in the Supplementary material for examples of vegetation structure.

3.2. Bird survey, richness and rarity calculation

Bird data were collected in 2012 by five experienced ornithologists (co-authors of this study). Each of the 153 survey points was visited twice during the season (5–6 and 28–29 May) to increase the likelihood of detecting the earlier and later breeding species. The survey points form a grid spaced at 300 m intervals. At each survey point, all bird individuals identified by sight or sound within a 100 m distance from the survey point were recorded. The results from both visits were pooled together and the bird diversity (species richness) for each survey point was calculated as the number of species detected on the survey point. In

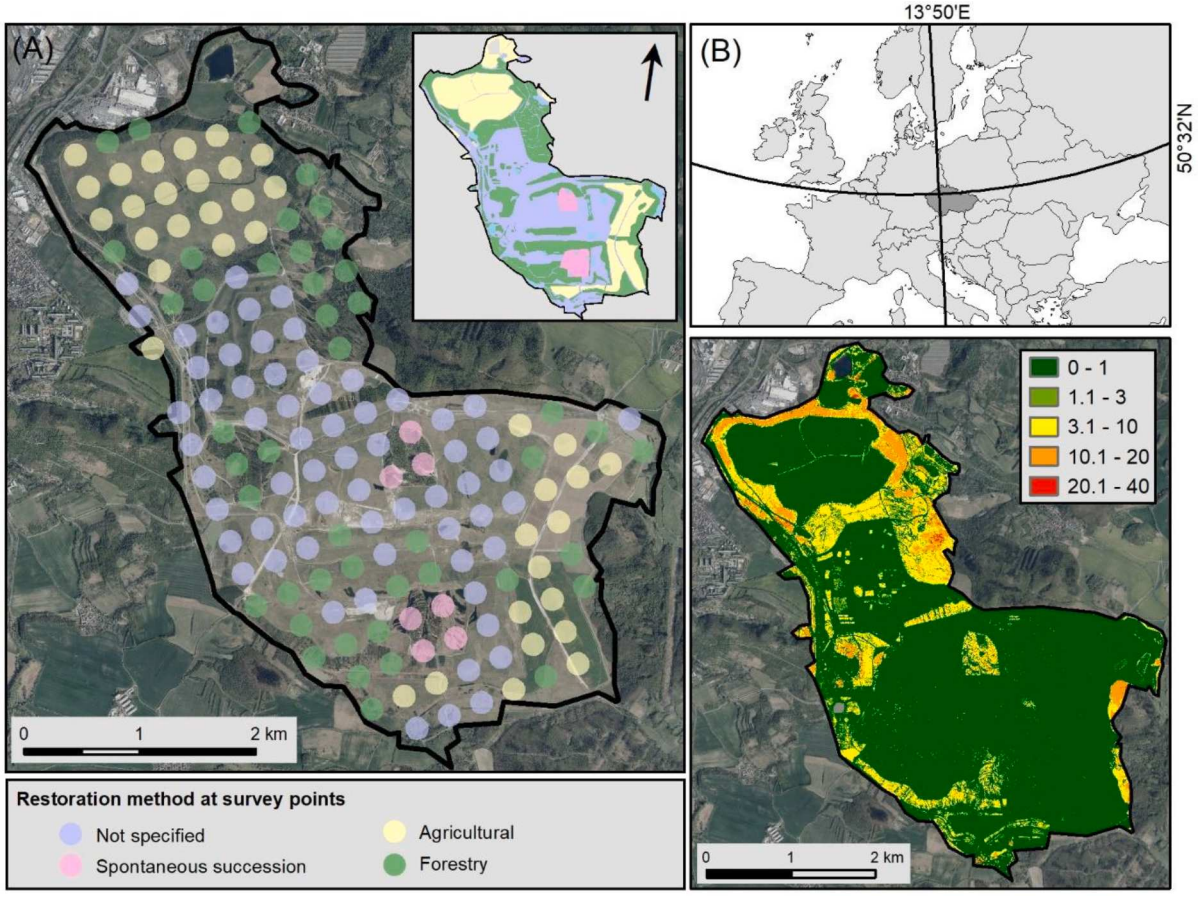


Fig. 1. Study area; (A) Location of survey points and adopted restoration method (forestry, agriculture, and naturally regenerating systems called "spontaneous succession"). The survey points represent a 100 m buffer; (B) Location of the study area in the Czech Republic; (C) Canopy height model (meters).

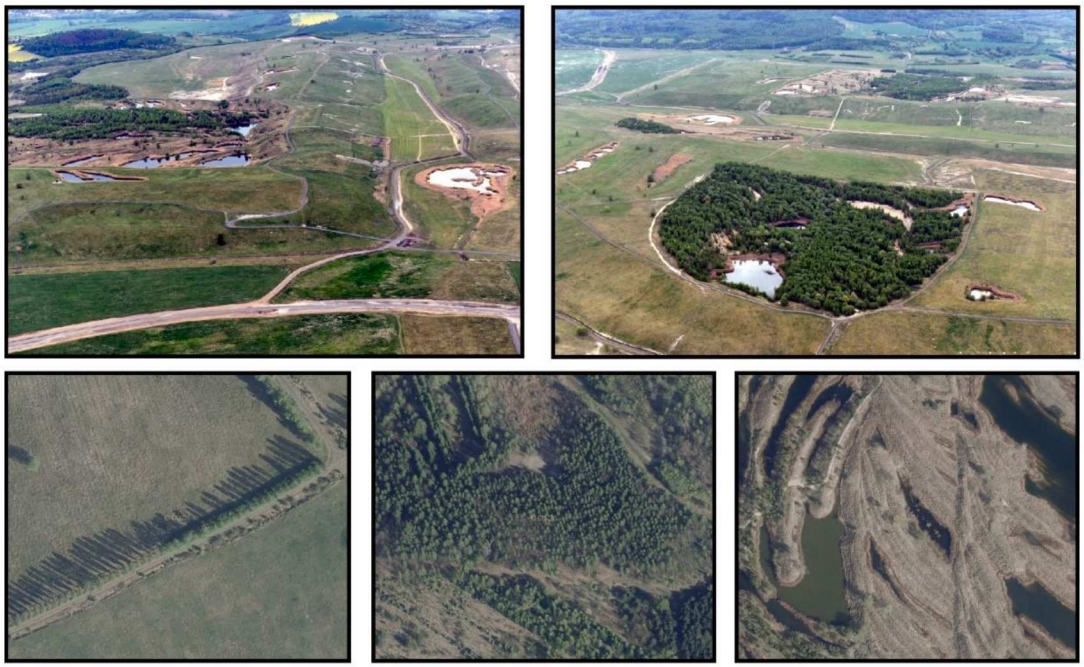


Fig. 2. Aerial photos of the study area. The two photos in the top row were taken on the 10th of May 2020 and show the two areas left to spontaneous succession surrounded mostly by agriculturally restored areas. The three bottom images were taken on 25th of April 2020 and show the areas after agricultural restoration (left), forest restoration (middle), and the area left to spontaneous succession (right). Note the presence of old dead vegetation even at this time of the year, particularly in areas left to spontaneous succession, near water bodies and other terrain depressions.

addition, we calculated an index of species rarity for each survey point reflecting the scarcity of species throughout the Czech Republic based on the large-scale quadrat mapping of birds (Šťastný, Bejček, & Hudec, 2006). For each species, the rarity index was calculated using the formula 1-*N*/628 where *N* represents the number of quadrats occupied by the species from 628 in total (Šálek, 2012; Table S2 in Supplementary material). The overall rarity for each survey point was then calculated as the sum of index values for all species recorded at a point. Individuals recorded while flying over the site or beyond the defined distance were excluded from the analysis.

3.3. Airborne data collection and pre-processing

The airborne LiDAR and hyperspectral data were acquired simultaneously on 18th May 2017 using a remote sensing platform FLIS (The Flying Laboratory of Imaging Spectroscopy; Hanuš, Fabiánek, & Fajmon, 2016). Flights for data collection were conducted at 1030 m above ground at 110 knots (ground speed). Data from the hyperspectral (Visible Near Infrared, VNIR, CASI-1500), and LiDAR sensor (Riegl LMS Q-780) were used in this study. Although there was a time lag between the field survey and LiDAR and hyperspectral data campaign, we assume that data were still useful for describing birds' habitats in terms of vegetation structure, vegetation productivity and presence of old dead vegetation as it has not changed substantially over the five years (but see more on this topic in Discussion).

3.3.1. Hyperspectral data

The hyperspectral imagery consisted of 48 bands covering the visible near-infrared range from 380 to 1050 nm (CASI-1500) with a bandwidth of 7.2 nm. Pre-processing of the hyperspectral images (i.e., radiometric correction, georeferencing and atmospheric corrections) were all carried out by the provider (CzechGlobe). Radiometric corrections were performed in the RadCorr software by converting spectral radiances to physical radiance units based on calibration parameters from the CzechGlobe spectroscopic laboratory (Hanuš et al., 2016). Radiance images were geometrically corrected, orthorectified using a digital terrain model (DTM), and georeferenced to the local Datum of Uniform Trigonometric Cadastral Network (EPSG: 5514). Data were corrected for atmospheric conditions using a radiative model MODTRAN and the BREFCOR method was used for correcting the bidirectional reflectance distribution function (BRDF) effect (ATCOR-4 software; Richter & Schläpfer, 2016).

3.3.2. LiDAR data

Airborne LiDAR data were acquired with a Riegl LMS Q-780 laser scanner. The scanner has a rotating polygon mirror and scans in parallel lines. The scan field of view is 60° and the wavelength is 1064 nm. The LiDAR data were provided in LAZ format with an average point density of 8 points per square meter. The LiDAR point cloud was processed using a proprietary software by the Global Change Research Institute CAS and referenced to the local Datum of Uniform Trigonometric Cadastral Network (EPSG: 5514) and Baltic Vertical Datum – After Adjustment (EPSG: 5705). We further processed the point cloud using LAStools (http://lastools.org) and classified the point cloud into ground and vegetation classes (Klápště et al., 2020; Moudrý et al., 2020). We divided the study area into 36 tiles and classified each tile separately to allow different settings and thus a better identification of ground and vegetation returns. In addition, we identified noise returns (e.g. returns from birds) and within the distance of 100 m from the grid survey points, we manually checked and edited point clouds for obvious errors (e.g. high voltage poles classified as vegetation). Returns other than vegetation and ground were removed from subsequent analyses. Prior to the calculation of vegetation structure variables, we height-normalized the LiDAR point cloud (i.e. the returns' height above the DTM was calculated).

3.4. Primary productivity and habitat heterogeneity variables

To investigate the importance of primary productivity and habitat heterogeneity, we derived two vegetation indices from the hyperspectral data and six variables from the LiDAR data; all these indices and variables bear a potential relevance to the bird diversity (Table 1). The two

Table 1
Overview of 8 potential explanatory variables derived from LiDAR and hyperspectral data within a 100 m vicinity of survey points. Rx denotes the reflectance at the wavelength of x nm. We excluded areas of water bodies from calculations of NDVI and PSRI. Variables in bold were used in our final models (i.e. after excluding collinear variables).

Hyperspectral and LiDAR derived metrics		Description	Category according to Bakx et al. (2020)
Primary productivity			
	Normalized Difference Vegetation Index (NDVI)	$(R_{862} - R_{662})/(R_{862} - R_{662})$; sensitive to vegetation greenness	–
Old dead vegetation from the previous vegetation season			
	Plant senescing reflectance index (PSRI)	$(R_{678} - R_{500})/R_{750}$; sensitive to senescent vegetation	–
Vegetation structure (Total vegetation)			
	Mean height	Average height of vegetation returns	Total vegetation - Height
	Standard deviation of height	Standard deviation of vegetation returns heights above 1 m	Total vegetation - Vertical variability
	Canopy cover	Number of first returns above 1 m divided by the sum of all first returns	Total vegetation - Cover
Vegetation structure (Single layers)			
	Cover of the herbaceous layer	Number of points between 0.1 m and 1 m divided by the total number of points	Single layer (Understorey) - Cover
	Cover of the shrub layer	Number of points between 1 m and 3 m divided by the total number of points	Single layer (Understorey) - Cover
	Cover of the tree layer	Number of points > 3 m divided by the total number of points	Single layer (Canopy) - Cover

calculated vegetation indices included 1) NDVI $(R_{862} - R_{662})/(R_{862} - R_{662})$, which corresponds to plant chlorophyll content and hence increases with vegetation productivity (i.e. the green component of biomass), and 2) PSRI $(R_{678} - R_{500})/R_{750}$, which is a measure of the leaf senescence and is sensitive to the carotenoid/chlorophyll ratio. PSRI is in this study used as an estimate of the amount of old dead (senescent) vegetation from the previous vegetation season, which is present in the study area even in May when both bird and airborne data were collected (see Fig. 2). The old dead vegetation from the previous season is mainly present in the aquatic vegetation (e.g. *Phragmites australis* and *Typha latifolia*) and, to a somewhat lesser extent, in low steppe vegetation (e.g. *Calamagrostis epigejos* and *Arrhenatherum elatius*).

In order to assess the effect of habitat heterogeneity on bird species richness and rarity, we described habitat heterogeneity using the variance in vegetation structure. We used vegetation structure variables adopted in the previous bird diversity studies (see Bakx et al., 2018 for the conceptual categorization of LiDAR-derived vegetation metrics). To describe the total vegetation (*sensu* Bakx et al., 2018) we used the mean, standard deviation of vegetation returns, and canopy cover (Table 1). We calculated these metrics to describe the structural variability of the vegetation directly from the point cloud (e.g. Bae et al., 2018); but note that some other studies calculated these metrics from the rasterized canopy height model (CHM) to describe a horizontal variation in the canopy cover (e.g. Müller, Moning, Baessler, Heurich, & Brandl, 2009; 2010). In addition, we used variables characterising the individual vegetation layers (single layer *sensu* Bakx et al., 2018). Three layers of vegetation are typically recognized; the herbaceous layer, the shrub layer, and the tree layer (e.g. Lesak et al., 2011; Jones, Arcese, Sharma, & Coops, 2013). The same vegetation layers are typically assessed during field inspections on postmining sites (e.g. Šálek, 2012). Therefore, we calculated the cover for three vegetation layers: first, we counted the number of points between 0.1 m and 1 m and divided this number by the sum of all points to estimate the cover of the herbaceous layer. Second, we counted the number of points between 1 m and 3 m and divided the result by the sum of all points to estimate the cover of the shrub layer. In the same way, the number of points between 3 m and 40 m was divided by the sum of all points to estimate the cover of the tree layer (in our study area, there are no trees higher than 40 m). It should be noted that the heights of these vegetation layers are selected arbitrarily based on our field experience (e.g. Šálek, 2012) and can thus greatly vary among different areas (e.g. Lesak et al., 2011). All primary productivity and habitat heterogeneity metrics were calculated within a 100 m radius of grid survey points using ENVI (version 5.5) and LAStools (version 200112), respectively.

3.5. Statistical analyses

We used boosted regression trees (BRT) implemented in the R package gbm version 2.1.5 (Greenwell, Boehmke, Cunningham, Developers, & Greenwell, 2019) and some additional features available in the package dismo version 1.1–4 (Hijmans, Phillips, Leathwick, Elith, & Hijmans, 2017) to assess how primary productivity and habitat heterogeneity were associated with species richness and rarity. First, we examined the collinearity among all variables to reduce the number of input variables (Fig. S3 in Supplementary material). Canopy cover and mean height were highly correlated with vertical vegetation structure metrics. As it has been highlighted that birds show a higher preference for the structural variability of the vegetation than for canopy cover (Davies & Asner, 2014) and as the vertical vegetation variability was of our primary concern, we retained the three vertical layers (i.e. herbaceous, shrub and tree cover) and excluded the canopy cover and mean vegetation height from further analyses. Another highly correlated pair of variables were NDVI and PSRI; however, as they represent unique components of the aboveground biomass and were essential for our study, we decided to retain both variables in the model. However, to evaluate whether retaining both NDVI and PSRI affected our results, we ran the models also individually with PSRI and NDVI, respectively (see the Supplementary material, Figs. S5–S8). Therefore, our final set of variables consisted of four variables representing the habitat heterogeneity and two variables representing the primary productivity (Table 1).

Two most important parameters that need to be specified for BRT are the tree complexity (which controls whether interactions are fitted) and learning rate (shrinkage) as they determine the number of trees required for the prediction. As a rule of thumb, a combination of tree complexity and learning rate that results in a model with at least 1000 trees is recommended. For models with less than 500 records, it is preferred to model simple trees (i.e. tree complexity 1–3) with a small learning rate to allow the model to grow enough trees. We fine-tuned the settings in preliminary testing and used models with tree complexity (the number of splits in a tree) of 1 (i.e. without interaction terms, as allowing interactions did not lead to a model improvement), shrinkage (learning rate) of 0.001, bag fraction (the proportion of data used when selecting optimal tree number) of 0.5, and the maximum number of trees of 5000. To estimate the optimal number of trees, we used 10-fold cross-validation. At each iteration, the residual deviance was calculated and the number of trees giving the best model (i.e. lowest deviance) was

identified (Elith, Leathwick, & Hastie, 2008; Hastie, Tibshirani, & Friedman, 2001; Leathwick, Elith, Francis, Hastie, & Taylor, 2006). Species richness and rarity were modelled specifying the Poisson and Gaussian error distribution, respectively. All models were fitted in R, version 3.6.0 (R Development Core Team, 2019)

3.6. Assessment of model performance

The identified best models were fitted to the entire dataset and used to produce partial dependency plots that show the effect of each variable after accounting for the average effects of all other variables (De'Ath, 2007; Elith et al., 2008). In addition, we assessed the relative importance of each variable (i.e. the contribution of each variable to the model fit scaled so that the sum adds to 100) using formulae developed by Friedman (2001) and implemented in the gbm package (Greenwell et al., 2019). The overall performance of BRT models was evaluated using the total deviance explained, which was calculated by dividing the difference between the mean total deviance and the estimated 10-fold cross-validated residual deviance by the mean total deviance. The cross-validated residual deviance is a measure of the deviance left unexplained by the model. Because results from the k-fold cross-validation can vary depending on the random selection of points for the folds, this

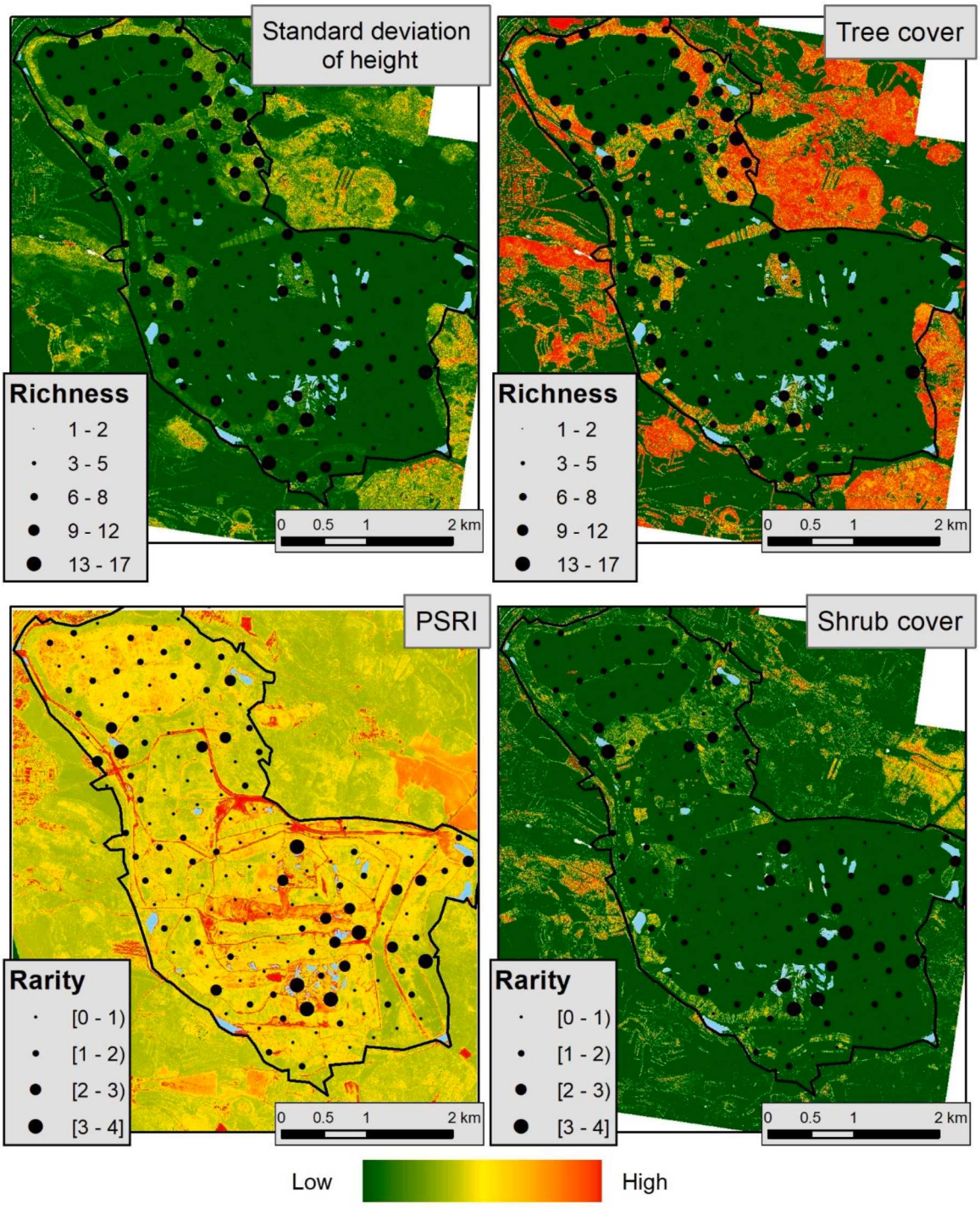


Fig. 3. Spatial distribution of species richness (upper images) and rarity (lower images) along with the four most influential explanatory variables (Standard deviation of height, Tree cover, PSRI, and Shrub cover).

procedure was repeated 5 times for each model, and overall means were calculated for the relative importance of each variable and total deviance explained by the models (Leathwick et al., 2006).

4. Results

4.1. Bird richness

We observed 83 bird species from a total of 1340 individual bird records. The overall bird species richness ranged from 1 to 17 species per survey point with a median of 6 species (Fig. 3). The BRT model of overall species richness explained 52% of deviance. The overall species richness was strongly affected by the habitat heterogeneity while primary productivity had only a minimal effect (Table S4 in Supplementary material). The effects of the midstory density and canopy density were the strongest and jointly accounted for most of the explained variability. The standard deviation of height had a moderate effect, and the effects of herbaceous cover, NDVI and PSRI were weak (Fig. 4). The partial dependency plots of the individual variables showed a rapid increase of species richness at relatively low values of the shrub cover and tree cover with only a minimal change as these variables continued to increase. The ranges of the shrub cover and herbaceous cover associated with the steepest increase in the bird richness were 0–5% and 0–20%, respectively. The standard deviation of height had a positive effect with an increase in overall richness for values between 2 m and 3 m and no effect above that value (Fig. 4).

4.2. Bird rarity

The bird rarity index ranged from 0.04 to 3.81 per survey point with a median rarity of 1 (Fig. 3). The BRT model of bird rarity explained 12% of deviance and the rarity was strongly affected by both habitat heterogeneity and primary productivity (Table S4 in Supplementary material). The shrub cover, herbaceous cover and PSRI had the strongest effects and jointly accounted for most of the explained variability. The effects of the standard deviation of height and NDVI were moderate, and the canopy density had a weak effect (Fig. 5). The partial dependency plots of the single variables showed a rapid increase in rarity for PSRI values above 0.10. A rapid increase in rarity was also shown at relatively low values of shrub cover and herbaceous cover with a minimal change as these variables continued to increase. The ranges of the shrub cover and herbaceous cover associated with the steepest increase of the bird rarity were 0–5% and 10–15%, respectively (Fig. 5).

4.3. Heterogeneity and productivity with respect to the restoration technique

Our results show a clear effect of the adopted restoration technique on habitat heterogeneity and primary productivity (Fig. 6). Most importantly, the spontaneous succession considerably differs from other sites when looking at primary productivity. The values of NDVI (i.e. vegetation greenness) and PSRI (i.e. senescent vegetation) were relatively similar for the agricultural and forest restoration but compared to them, the sites left to spontaneous succession had much lower values of NDVI and clearly higher values of PSRI (Fig. 6). The structural measures show a distribution of values commensurate with the individual habitat types. Note that the sites with unspecified restoration method were similar to agricultural restoration and mostly consisted of low vegetation.

5. Discussion

In this study, we evaluated the effect of the variance in vegetation structure, primary productivity and senescent vegetation on bird communities colonizing newly available (restored) habitats after coal mining (Fig. 3). We found a detectable relationship between fine-scale habitat attributes of early succession stages derived from airborne LiDAR and hyperspectral data and the occurrence of birds. The models with six variables representing the vegetation structure, primary productivity and the presence of senescent vegetation explained 52% and 12% of the variability in species richness and rarity, respectively, which is comparable to prior studies. The previous studies combining LiDAR data with the indices derived from passive optical sensors (e.g. NDVI) typically explained the variability in species richness between 15% and

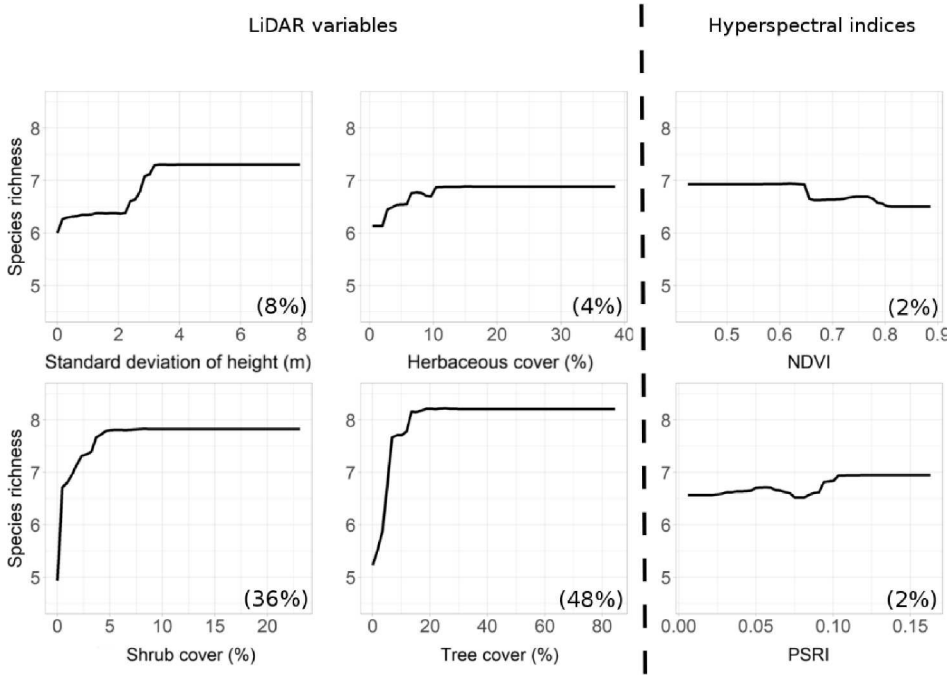


Fig. 4. Partial dependency plots for boosted tree analyses of overall species richness. The partial plots show the modelled relationships between species richness and standard deviations of height, herbaceous cover, shrub cover, tree cover, NDVI, and PSRI. The relative importance of the variable in the model is given in parentheses.

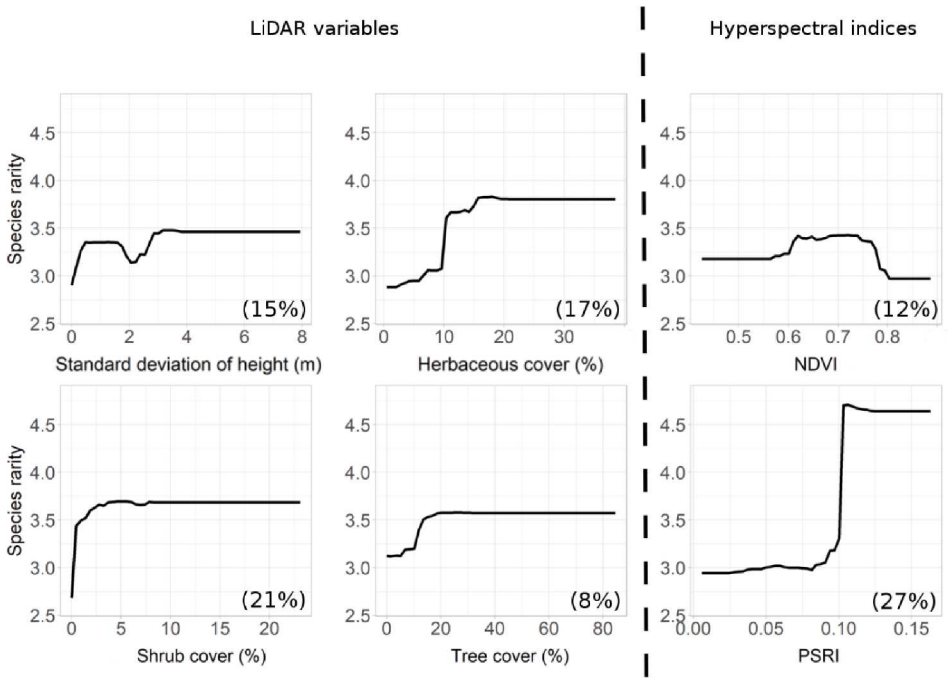


Fig. 5. Partial dependency plots for boosted tree analyses of species rarity. The partial plots show the modelled relationships between the species rarity and standard deviations of height, herbaceous cover, shrub cover, tree cover, NDVI, and PSRI. The relative importance of the variable in the model is given in parentheses.

55% (Goetz et al., 2007; Jones et al., 2013; Vogeler et al., 2014).

5.1. *Importance of vegetation structure, NDVI and PSRI for bird species richness*

Variance in vegetation structure had a positive effect on species richness and explained most of the explained variablity (Fig. 4). While NDVI and PSRI showed almost no effect on species richness, shrub cover (i.e. the cover of shrubs and saplings between 1 m and 3 m high) and tree cover (i.e. cover of vegetation >3 m high) exhibited a strong positive effect on species richness. The standard deviation of height and herbaceous cover (cover of vegetation up to 1 m high) had minor effects on species richness. Species richness climbed steeply at low values of shrub cover (0%–5%), low values of tree cover (0%–20%), and the standard deviation of height between 2 m and 3 m (Fig. 4). Such conditions are typical of forest restoration and sites left to spontaneous succession (Fig. 5). These findings are in accordance with prior studies such as Goetz et al. (2007) and Vogeler et al. (2014) who combined LiDAR with NDVI and more recently Melin, Hill, Bellamy, and Hinsley (2019) who combined LiDAR with variables derived from hyperspectral data and showed that the vegetation structure (i.e. LiDAR derived variables) is more important for the assessment of bird species richness at local scales than variables derived from passive remote sensing.

5.2. *Importance of vegetation structure, NDVI and PSRI for bird species rarity*

In contrast to species richness, our results show that combining information from LiDAR and passive optical sensors might be important when species rarity is of concern. We found a positive effect of both habitat heterogeneity and primary productivity on species rarity (Fig. 5). PSRI (i.e. old dead vegetation from the previous vegetation season) was the most important predictor with a strong positive effect on species rarity, followed by shrub and herbaceous cover. The greatest increase in rarity was associated with PSRI values higher than 0.10, above 2% for the shrub cover, and above 10% for the herbaceous cover (Fig. 5); above these values, the rarity remained more or less constant. This is likely because the senescent vegetation and relatively high herbaceous and shrub covers provide shelter and enhance the diversity of insect communities and hence food availability for birds (e.g. Müller, Bae, Röder, Chao, & Didham, 2014; Soto et al., 2017; Vergara et al., 2017).

It is, however, important to note that the model of bird rarity explained only 12% of deviance. This is likely related to the fact that typical rare species that occur in our study area are ground-nesting or foraging birds. Such species include, for example, Wheatear *Oenanthe oenanthe*, Montagús Harrier *Circus pygargus*, Bluethroat *Luscinia svecica cyanecula*, Whinchat *Saxicola rubetra*, Stonechat *Saxicola torquata*, Great Reed Warbler *Acrocephalus arundinaceus* or Meadow Pipit *Anthus pratensis*. These specialists require specific and mutually different conditions such as bare grounds (Wheatear), unmanaged grassy patches (Montagús Harrier, Whinchat, Stonechat, Meadow Pipit) or reedbeds (Bluethroat, Great Reed Warbler) that are neither adequately represented by the variance in the vegetation structure nor by NDVI and PSRI indices, respectively. Indeed, ground-nesting or foraging species are typically reported to be poorly modelled using habitat heterogeneity and productivity variables (Cooper et al., 2020; Weisberg et al., 2014). The five-year time lag between the data acquisitions and difficulties to distinguish old dead vegetation from bare surfaces might represent alternative explanations for the relatively low deviance explained by these factors (see Chapter 5.4. below).

5.3. *Which type of restoration results in the development of more favourable habitats?*

It is evident that all vertical levels of habitat heterogeneity (i.e. herbaceous, shrub and tree covers) are important either for species richness or rarity. This supports the necessity to create the mosaic of habitats with heterogeneous vertical structure during restoration to support high species richness as suggested e.g. by Harabiš, Tichanek, and Tropek (2013). However, the high variance in vegetation structure alone is not sufficient to support rare species. Indeed, management and restoration goals can change considerably depending on whether the aim is to support high species richness or rare species (Cooper et al., 2020). It seems that a high herbaceous and shrub cover combined with the presence of old dead vegetation promote rarity. However, this

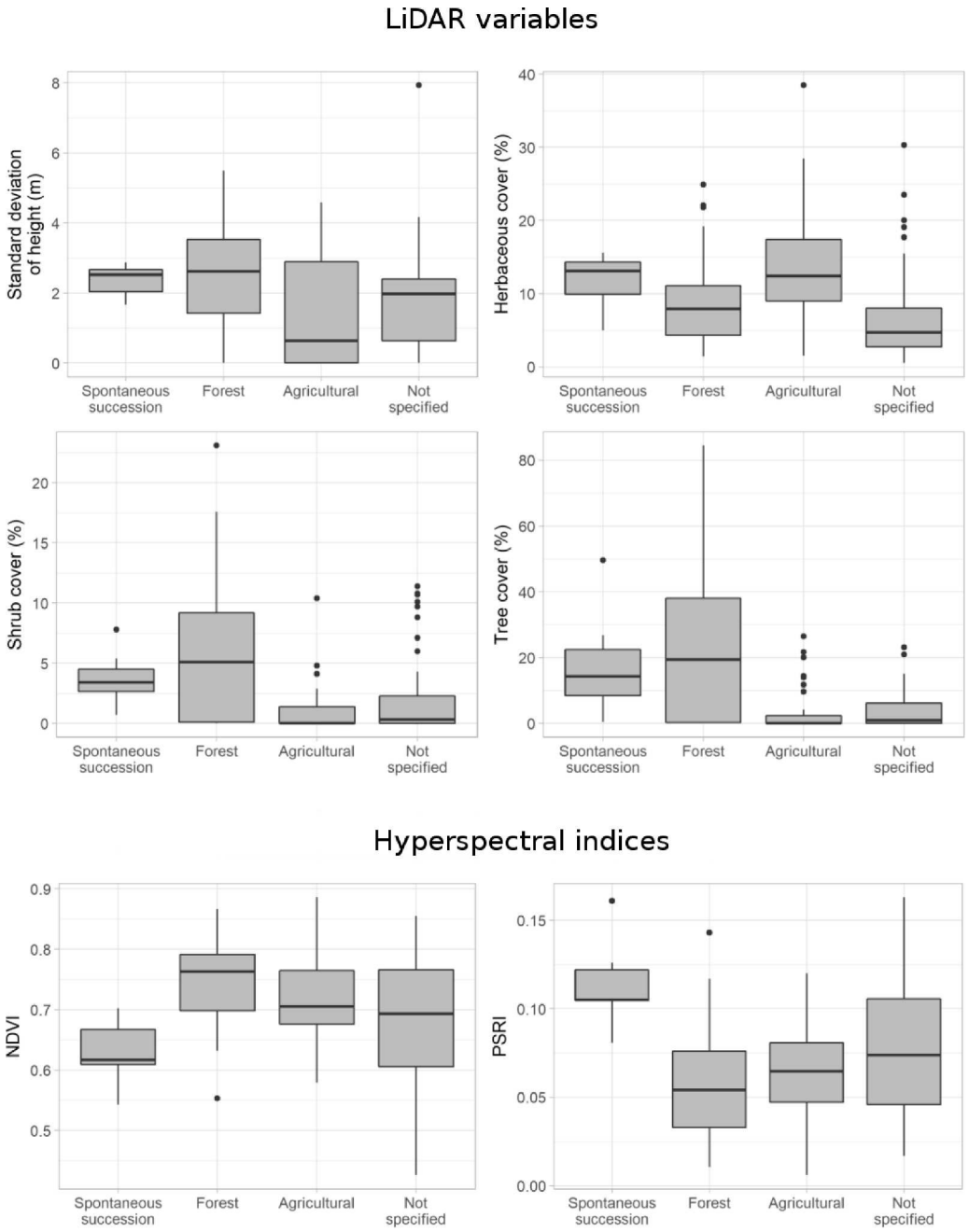


Fig. 6. Comparison of the vegetation structure (LiDAR variables), NDVI and PSRI indices depending on the adopted restoration techniques. The central horizontal line in the box marks the median. The boxes show the interquartile range (25th to 75th percentile) and the whiskers indicate 1.5 times the interquartile range.

combination is rare on technically reclaimed sites, i.e. after agricultural and forest restoration. In contrast, this study shows that such conditions are typically present on sites left to spontaneous succession (Fig. 5). This is a likely explanation for the fact that the spontaneous succession sites are particularly favoured by rare vertebrates (Šálek, 2012; Vojar et al., 2016).

5.4. Use and limitations of airborne remote sensing data in restoration practice

We have shown that the measurement of the habitat heterogeneity derived from airborne laser scanning point clouds can provide ecologically meaningful variables. As field-based estimates of vegetation structure are used as a rapid and efficient way of assessing the condition of restored sites (Gibbons & Freudenberger, 2006), LiDAR can become an alternative to such field surveys. Compared to field surveys, however, LiDAR has a higher potential for providing information that can lead to management action. As LiDAR data availability is increasing continuously due to national or regional scanning campaigns (see Melin et al., 2017; Stereńczak et al., 2020 for the list of countries and regions with LiDAR data available) and thanks to the more common adoption of open data policies (Rocchini et al., 2017), there is a high potential to use it in restoration ecology, especially for assessment of vegetation structure (Guo et al., 2017; Koska et al., 2017; Moudrý, Gdulová, et al., 2019; Moudrý, Urban, et al., 2019; Szostak, Pietrzykowski, & Likus-Cieślik, 2020). Therefore, we suggest that such data should be increasingly utilized by managing authorities for optimizing the restoration success assessment and enhancing the ecological value of reclaimed areas.

It should be, however, noted that despite the increase of LiDAR data availability, acquisitions for the same area (e.g. state) have, due to high

acquisition costs, long repetition times. Therefore, studies such as the presented one are often being done under suboptimal conditions and the time lag between field surveys and LiDAR data acquisition is common (Lesak et al., 2011). For example, Goetz et al. (2007) – 6 years; Wallis et al. (2016) – 10 years. Huber, Kienast, Ginzler, and Pasinelli (2016) – 10 years. It has been shown that several years time lag is not a fundamental source of error in mature woodland ecosystems as the changes in vegetation structure are usually relatively slow (Hill & Hinsley, 2015; Vierling, Swift, Hudak, Vogeler, & Vierling, 2014). This is, however, not the case of our study area that consists of agricultural, forest, and successional sites and the time lag between remotely sensed and bird data collection might be a potential source of error. While this is certainly not a problem for agriculturally restored sites, differences in vegetation structure might have arisen during the 5 years at sites restored through forestry and those left to spontaneous succession. On the other hand, the two sites left to spontaneous succession in our study area are 16 and 26 years old, respectively, and the current rate of changes in the vegetation structure is relatively slower than during the early stages of succession. Moreover, the spontaneous succession is often blocked by sandy soils and dense grass cover (e.g. *Calamagrostis epigejos*). Therefore, the main differences can be expected in vegetation within the 1 – 3 m height range that could have grown over 3 m.

It is even more important to minimize the time lag when hyperspectral data are used. This is especially true when indices related to the vegetation biochemistry (which can change in a matter of weeks across a vegetation season) are used, it is much preferable if the hyperspectral data collection and bird survey are performed at the same time (e.g. Melin et al., 2019). Whit this in mind, we did not use any indices related directly to vegetation biochemistry and concentrated only on indices that are positively correlated to the characteristics that should have remained relatively stable over the five years (i.e. NDVI for vegetation productivity and PSRI for the amount of old dead vegetation). Besides, the bird occurrences and remotely sensed data were both collected in May, which was particularly important for the ecological relevance as this allowed accurate estimates of the amount of old dead vegetation present during the breeding season. The old dead vegetation is typical for terrain depressions (e.g. *Phragmites australis* and *Typha latifolia*) and agriculturally restored areas (e.g. *Calamagrostis epigejos* and *Arrhenatherum elatius*) and its detectability is changing more within a year than between years due to the pronounced seasonality (see also the note above on blocked succession). The amount of old dead vegetation is actually constantly growing and such sites might have become even more favourable for rare species as Whinchat, Stonechat, Montagús Harrier or Bluethroat in 2017 (the year of remote sensing data acquisition) than they were in 2012 (the year of bird data collection).

There is, however, another potential source of error, which might possibly also have been the reason why the model of bird rarity explained only 12% of deviance, namely the fact that the PSRI is high also in areas completely without vegetation, such as roads (Fig. 2). On the other hand, however, old dead vegetation, which forms an important habitat for Whinchat or Stonechat, is often present in the ditches along the roads (e.g. *Arrhenatherum elatius*) and it is, therefore, difficult to separate their effects.

For future studies and especially for monitoring practice, it would be beneficial to agree on several LiDAR-derived metrics proven to be effective for explaining species diversity and to recommend them as standard structural indicators. To facilitate comparisons, Bakx et al. (2019) recently grouped LiDAR-derived variables into 24 classes defined by six categories of vegetation (total vegetation, single trees, canopy, understorey, other single layers, and multi-layer) and four categories of the structural type (cover, height, horizontal variability and vertical variability). Our results show that the total vegetation vertical variability (i.e. standard deviation of returns height) and canopy and understory cover (i.e. herbaceous, shrub and tree covers) are potentially relevant for the assessment of conditions on early successional restored sites.

6. Conclusions

Understanding drivers of species distributions across restored landscapes and identifying areas that have the potential for supporting high species richness, vulnerable or rare species, is important for successful management of restored sites. Overall, our results show that both habitat heterogeneity and primary productivity play an important role in bird species diversity on restored sites. Shrub cover had a strong positive effect on both species richness and rarity, while tree cover had a strong positive effect on species richness. Herbaceous cover and the presence of senescent vegetation had both positive effects on species rarity. This highlights the necessity of creating a mosaic of habitats with heterogeneous vertical structure during restoration or to design the vegetation structure in a way supporting preferred species. To support this, we suggest to reduce intensive mowing of agriculturally restored areas (e.g. by creating unmowed strips of vegetation), preserve naturally formed waterlogged areas, plant trees with different growth rates and combine them with shrub vegetation. Sites left to spontaneous succession play an important role in creating restored ecosystems of high ecological value as they represent a unique combination of vegetation density and presence of senescent vegetation that, in combination, promote high species rarity.

In our opinion, airborne remote sensing, particularly laser scanning, should constitute an integral part of restoration success assessment and should be acquired with reasonable repetition rate (e.g. 5–10 years) over the restored areas, as the information derived from such data can be more easily implemented in management actions than subjective semi-quantitative or categorical measures collected during intensive fieldwork. We suggest a wider use of vegetation structure and productivity indices derived from remotely sensed data in restoration success assessment. Our results show that the total vegetation vertical variability (i.e. standard deviation of returns height) and various vertical layers of vegetation cover (i.e. herbaceous, shrub and tree cover) in combination with senescent vegetation (i.e. PSRI) are potentially relevant for monitoring of early successional restored sites.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to three anonymous reviewers for their constructive comments that significantly improved the quality of the manuscript. This work was supported by the Internal Grant Agency of the Faculty of Environmental Sciences, Czech University of Life Sciences Prague (project No. 2020B0029), and by the Czech Science Foundation (project No. 17-17156Y).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.landurbplan.2021.104064.

References

Adams, B. T., & Matthews, S. N. (2019). Diverse temperate forest bird assemblages demonstrate closer correspondence to plant species composition than vegetation structure. *Ecography, 42*(10), 1752–1764.

Bae, S., Müller, J., Lee, D., Vierling, K. T., Vogeler, J. C., Vierling, L. A., … Thorn, S. (2018). Taxonomic, functional, and phylogenetic diversity of bird assemblages are oppositely associated to productivity and heterogeneity in temperate forests. *Remote Sensing of Environment, 215*, 145–156.

Bailey, S. A., Horner-Devine, M. C., Luck, G., Moore, L. A., Carney, K. M., Anderson, S., … Fleishman, E. (2004). Primary productivity and species richness: relationships

among functional guilds, residency groups and vagility classes at multiple spatial scales. *Ecography, 27*, 207–217.
Bakx, T. R., Koma, Z., Seijmonsbergen, A. C., & Kissling, W. D. (2019). Use and categorization of Light Detection and Ranging vegetation metrics in avian diversity and species distribution research. *Diversity and Distributions, 25*, 1045–1059.
Bejček, V., & Šťastný, K. (1984). The succession of bird communities on spoil banks after surface brown coal mining. *Ekologia Polska-Polish Journal of Ecology, 32*, 245–259.
Bejček, V. (1988). Communities of small terrestrial mammals on the spoil banks in the Most Basin. Vysoka Skola Zemedelska, Praha.
Brown, J. H. (1981). Two decades of homage to Santa Rosalia: Toward a general theory of diversity. *American Zoologist, 21*, 877–888.
Cardoso da Silva, J.M., & Vickery, P.D. (2002). Birds. In Perrow, M.R., Davy, A.J. (eds.) Handbook of Ecological Restoration, Vol. 1, Principles of Restoration, pp. 376–388. Cambridge University Press, Cambridge.
Clewell, A. F., & Aronson, J. (2013). Ecological restoration: principles, values, and structure of an emerging profession. Island Press.
Cooper, W. J., McShea, W. J., Forrester, T., & Luther, D. A. (2020). The value of local habitat heterogeneity and productivity when estimating avian species richness and species of concern. *Ecosphere, 11*(5).
Cordell, S., Questad, E. J., Asner, G. P., Kinney, K. M., Thaxton, J. M., Uowolo, A., … Chynoweth, M. W. (2017). Remote sensing for restoration planning: how the big picture can inform stakeholders. *Restoration Ecology, 25*, S147–S154.
Crouzeilles, R., Ferreira, M.S., Chazdon, R.L., Lindenmayer, D.B., Sansevero, J.B., Monteiro, L., … Strassburg, B.B. (2017). Ecological restoration success is higher for natural regeneration than for active restoration in tropical forests. Science Advances, 3, p.e1701345.
Davies, A. B., & Asner, G. P. (2014). Advances in animal ecology from 3D-LiDAR ecosystem mapping. *Trends in Ecology & Evolution, 29*, 681–691.
De'Ath, G. (2007). Boosted trees for ecological modeling and prediction. *Ecology, 88*, 243–251.
Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology, 77*, 802–813.
Evans, K. L., Warren, P. H., & Gaston, K. J. (2005). Species–energy relationships at the macroecological scale: a review of the mechanisms. *Biological Reviews, 80*, 1–25.
Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189–1232.
Giam, X., Olden, J. D., & Simberloff, D. (2018). Impact of coal mining on stream biodiversity in the US and its regulatory implications. *Nature Sustainability, 1*, 176–183.
Gibbons, P., & Freudenberger, D. (2006). An overview of methods used to assess vegetation condition at the scale of the site. *Ecological Management & Restoration, 7*, S10–S17.
Goetz, S., Steinberg, D., Dubayah, R., & Blair, B. (2007). Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. *Remote Sensing of Environment, 108*, 254–263.
Gould, S. F., & Mackey, B. G. (2015). Site vegetation characteristics are more important than landscape context in determining bird assemblages in revegetation. *Restoration Ecology, 23*, 670–680.
Greenwell, B., Boehmke, B., Cunningham, J., Developers, G. B. M., & Greenwell, M. B. (2019). Package 'gbm'. R package version, 2(5).
Guo, X., Coops, N. C., Tompalski, P., Nielsen, S. E., Bater, C. W., & Stadt, J. J. (2017). Regional mapping of vegetation structure for biodiversity monitoring using airborne lidar data. *Ecological Informatics, 38*, 50–61.
Hagger, V., Wilson, K., England, J. R., & Dwyer, J. M. (2019). Water availability drives aboveground biomass and bird richness in forest restoration plantings to achieve carbon and biodiversity cobenefits. *Ecology and Evolution, 9*, 14379–14393.
Hanuš, J., Fabiánek, T.K., Fajmon, L., 2016. Potential of airborne imaging spectroscopy at czechglobe. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLI, 15–17. DOI: 10.5194/isprsarchives-XLI-B1-15-2016.
Harabiš, F., Tichanek, F., & Tropek, R. (2013). Dragonflies of freshwater pools in lignite spoil heaps: Restoration management, habitat structure and conservation value. *Ecological Engineering, 55*, 51–61.
Hastie, T., Tibshirani, R., & Friedman, J.H. (2001). The elements of statistical learning: data mining, inference, and prediction. Springer-Verlag, New York.
Hendrychová, M., Svobodova, K., & Kabrna, M. (2020). Mine reclamation planning and management: Integrating natural habitats into post-mining land use. Resources Policy, 69, 101882.
Hijmans, R. J., Phillips, S., Leathwick, J., Elith, J., & Hijmans, M. R. J. (2017). Package 'dismo'. *Circles, 9*, 1–68.
Hill, R. A., & Hinsley, S. A. (2015). Airborne lidar for woodland habitat quality monitoring: Exploring the significance of lidar data characteristics when modelling organism-habitat relationships. *Remote Sensing, 7*, 3446–3466.
Hobi, M. L., Dubinin, M., Graham, C. H., Coops, N. C., Clayton, M. K., Pidgeon, A. M., & Radeloff, V. C. (2017). A comparison of Dynamic Habitat Indices derived from different MODIS products as predictors of avian species richness. *Remote Sensing of Environment, 195*, 142–152.
Huber, N., Kienast, F., Ginzler, C., & Pasinelli, G. (2016). Using remote-sensing data to assess habitat selection of a declining passerine at two spatial scales. *Landscape Ecology, 31*, 1919–1937.
Jones, T. G., Arcese, P., Sharma, T., & Coops, N. C. (2013). Describing avifaunal richness with functional and structural bioindicators derived from advanced airborne remotely sensed data. *International Journal of Remote Sensing, 34*, 2689–2713.
Klápště, P., Fogl, M., Barták, V., Gdulová, K., Urban, R., & Moudrý, V. (2020). Sensitivity analysis of parameters and contrasting performance of ground filtering algorithms with UAV photogrammetry-based and LiDAR point clouds. *International Journal of Digital Earth, 13*, 1672–1694.
Kolář, V., Tichanek, F., & Tropek, R. (2017). Effect of different restoration approaches on two species of newts (Amphibia: Caudata) in Central European lignite spoil heaps. *Ecological Engineering, 99*, 310–315.
Koska, Bronislav, Jirka, Vladimír, Urban, Rudolf, Křemen, Tomáš, Hesslerová, Petra, Jon, Jakub, … Fogl, Michal (2017). Suitability, characteristics, and comparison of an airship UAV with lidar for middle size area mapping. *International Journal of Remote Sensing, 38*, 2973–2990. https://doi.org/10.1080/01431161.2017.1285086.
Kung, A., Svobodova, K., Lèbre, E., Valenta, R., Kemp, D., & Owen, J. R. (2020). Governing deep sea mining in the face of uncertainty. *Journal of Environmental Management, 111593*.
Laurin, G. V., Puletti, N., Grotti, M., Stereńczak, K., Modzelewska, A., Lisiewicz, M., & Papale, D. (2020). Species dominance and above ground biomass in the Białowieża Forest, Poland, described by airborne hyperspectral and lidar data. *International Journal of Applied Earth Observation and Geoinformation, 92*.
Leathwick, J. R., Elith, J., Francis, M. P., Hastie, T., & Taylor, P. (2006). Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. *Marine Ecology Progress Series, 321*, 267–281.
Lèbre, É., Stringer, M., Svobodova, K., Owen, J. R., Kemp, D., Côte, C., & Valenta, R. K. (2020). The social and environmental complexities of extracting energy transition metals. *Nature Communications, 11*, 1–8.
Lesak, A. A., Radeloff, V. C., Hawbaker, T. J., Pidgeon, A. M., Gobakken, T., & Contrucci, K. (2011). Modeling forest songbird species richness using LiDAR-derived measures of forest structure. *Remote Sensing of Environment, 115*, 2823–2835.
Leyequien, E., Verrelst, J., Slot, M., Schaepman-Strub, G., Heitkönig, I. M., & Skidmore, A. (2007). Capturing the fugitive: Applying remote sensing to terrestrial animal distribution and diversity. *International Journal of Applied Earth Observation and Geoinformation, 9*, 1–20.
MacArthur, R. H., & MacArthur, J. W. (1961). On bird species diversity. *Ecology, 42*, 594–598.
Martins, W. B. R., Lima, M. D. R., Junior, U. D. O. B., Amorim, L. S. V. B., de Assis Oliveira, F., & Schwartz, G. (2020). Ecological methods and indicators for recovering and monitoring ecosystems after mining: A global literature review. *Ecological Engineering, 145*.
Melin, M., Shapiro, A., Glover-Kapfer, P., (2017). Lidar for ecology and conservation. WWF Conservation Technology Series 1 (3), WWF-UK, Woking, United Kingdom.
Melin, M., Hill, R. A., Bellamy, P. E., & Hinsley, S. A. (2019). On bird species diversity and Remote sensing—utilizing lidar and hyperspectral data to assess the role of vegetation structure and foliage characteristics as drivers of avian diversity. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12*, 2270–2278.
Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., & Rakitin, V. Y. (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia plantarum, 106*, 135–141.
Moudrý, V., Gdulová, K., Fogl, M., Klápště, P., Urban, R., Komárek, J., Moudrá, L., … Solský, M. (2019). Comparison of leaf-off and leaf-on combined UAV imagery and airborne LiDAR for assessment of a post-mining site terrain and vegetation structure: Prospects for monitoring hazards and restoration success. *Applied geography, 104*, 32–41.
Moudrý, V., Klápště, P., Fogl, M., Gdulová, K., Barták, V., & Urban, R. (2020). Assessment of LiDAR ground filtering algorithms for determining ground surface of non-natural terrain overgrown with forest and steppe vegetation. *Measurement, 150*.
Moudrý, V., Urban, R., Štroner, M., Komárek, J., Brouček, J., & Prošek, J. (2019). Comparison of a commercial and home-assembled fixed-wing UAV for terrain mapping of a post-mining site under leaf-off conditions. *International Journal of Remote Sensing, 40*(2), 555–572.
Müller, J., Bae, S., Röder, J., Chao, A., & Didham, R. K. (2014). Airborne LiDAR reveals context dependence in the effects of canopy architecture on arthropod diversity. *Forest Ecology and Management, 312*, 129–137.
Müller, J., Moning, C., Baessler, C., Heurich, M., & Brandl, R. (2009). Using airborne laser scanning to model potential abundance and assemblages of forest passerines. *Basic and Applied Ecology, 10*, 671–681.
Müller, J., Stadler, J., & Brandl, R. (2010). Composition versus physiognomy of vegetation as predictors of bird assemblages: The role of lidar. *Remote Sensing of Environment, 114*, 490–495.
Osenberg, C. W. (2018). No clean coal for stream animals. *Nature Sustainability, 1*, 160–161.
Prošek, J., Gdulová, K., Barták, V., Vojar, J., Solský, M., Rocchini, D., & Moudrý, V. (2020). Integration of hyperspectral and LiDAR data for mapping small water bodies. *International Journal of Applied Earth Observation and Geoinformation, 92*.
Reif, J., Hořák, D., Krištín, A., Kopsová, L., & Devictor, V. (2016). Linking habitat specialization with species' traits in European birds. *Oikos, 125*, 405–413.
Richter, R., & Schläpfer, D. (2016). ATCOR-4 User Guide. Ger. Aerosp. Center, Ger. 7.0.3, 565-01.
Rocchini, D., Petras, V., Petrasova, A., Horning, N., Furtkevicova, L., Neteler, M., … Wegmann, M. (2017). Open data and open source for remote sensing training in ecology. *Ecological Informatics, 40*, 57–61.
Šálek, M. (2012). Spontaneous succession on opencast mining sites: Implications for bird biodiversity. *Journal of Applied Ecology, 49*, 1417–1425.
Šálek, M., Hendrychová, M., & Řehoř, M. (2010). Breeding habitat of sparrowhawks, Accipiter nisus on spoil heaps after coal mining. *Acta Oecologica, 36*, 197–201.
Schulz, F., & Wiegleb, G. (2000). Development options of natural habitats in a post-mining landscape. *Land Degradation & Development, 11*, 99–110.
Soto, G. E., Pérez-Hernández, C. G., Hahn, I. J., Rodewald, A. D., & Vergara, P. M. (2017). Tree senescence as a direct measure of habitat quality: Linking red-edge vegetation indices to space use by Magellanic Woodpeckers. *Remote Sensing of Environment, 193*, 1–10.

Šťastný, K., Bejček, V., & Hudec, K. (2006). Atlas of breeding birds in the Czech Republic: 2001–2003. Aventinum, Praha.
Stereńczak, K., Laurin, G. V., Chirici, G., Coomes, D. A., Dalponte, M., Latifi, H., & Puletti, N. (2020). Global Airborne Laser Scanning Data Providers Database (GlobALS)—A New Tool for Monitoring Ecosystems and Biodiversity. *Remote Sensing, 12*(11), 1877.
Svobodova, K., Owen, J. R., Harris, J., & Worden, S. (2020). Complexities and contradictions in the global energy transition: A re-evaluation of country-level factors and dependencies. *Applied Energy, 265*.
Szostak, M., Pietrzykowski, M., & Likus-Cieślik, J. (2020). reclaimed area land cover mapping using sentinel-2 imagery and LiDAR point clouds. *Remote Sensing, 12*, 261.
Tews, J., Brose, U., Grimm, V., Tielbörger, K., Wichmann, M. C., Schwager, M., & Jeltsch, F. (2004). Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. *Journal of Biogeography, 31*, 79–92.
Tropek, R., Kadlec, T., Karesova, P., Spitzer, L., Kocarek, P., Malenovsky, I., … Konvicka, M., 2010. Spontaneous succession in limestone quarries as an effective restoration tool for endangered arthropods and plants. Journal of Applied Ecology, 47, pp.139-147.
Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation.
Urban, Rudolf, Štroner, Martin, Křemen, Tomáš, Braun, Jaroslav, & Möeser, Michael (2018). A novel approach to estimate systematic and random error of terrain derived from UAVs: a case study from a post-mining site. *Acta Montanistica Slovaca, 23*(3), 325–336.
Vanhée, B., & Devigne, C. (2018). Differences in collembola species assemblages (Arthropoda) between spoil tips and surrounding environments are dependent on vegetation development. *Scientific Reports, 8*, 1–16.
Vergara, P. M., Meneses, L. O., Grez, A. A., Quiroz, M. S., Soto, G. E., Pérez-Hernández, C. G., … Fierro, A. (2017). Occupancy pattern of a long-horned beetle in a variegated forest landscape: linkages between tree quality and forest cover across spatial scales. *Landscape Ecology, 32*, 279–293.
Vicentini, F., Hendrychova, M., Tajovský, K., Pižl, V., & Frouz, J. (2020). The effect of topography on long-term spontaneous development of soil and woody cover on graded and untreated overburden. *Forests, 11*(5), 602.
Vierling, K. T., Swift, C. E., Hudak, A. T., Vogeler, J. C., & Vierling, L. A. (2014). How much does the time lag between wildlife field-data collection and LiDAR-data acquisition matter for studies of animal distributions? A case study using bird communities. *Remote sensing letters, 5*, 185–193.
Vogeler, J. C., Hudak, A. T., Vierling, L. A., Evans, J., Green, P., & Vierling, K. T. (2014). Terrain and vegetation structural influences on local avian species richness in two mixed-conifer forests. *Remote Sensing of Environment, 147*, 13–22.
Vojar, J., Doležalová, J., Solský, M., Smolová, D., Kopecký, O., Kadlec, T., & Knapp, M. (2016). Spontaneous succession on spoil banks supports amphibian diversity and abundance. *Ecological Engineering, 90*, 278–284.
Vymazal, J., & Sklenicka, P. (2012). Restoration of areas affected by mining. *Ecological Engineering, 43*, 1–4.
Wallis, C. I., Paulsch, D., Zeilinger, J., Silva, B., Fernández, G. F. C., Brandl, R., … Bendix, J. (2016). Contrasting performance of Lidar and optical texture models in predicting avian diversity in a tropical mountain forest. *Remote Sensing of Environment, 174*, 223–232.
Weisberg, P. J., Dilts, T. E., Becker, M. E., Young, J. S., Wong-Kone, D. C., Newton, W. E., & Ammon, E. M. (2014). Guild-specific responses of avian species richness to lidar-derived habitat heterogeneity. *Acta Oecologica, 59*, 72–83.
Worlanyo, A. S., & Jiangfeng, L. (2020). Evaluating the environmental and economic impact of mining for post-mined land restoration and land-use: A review. *Journal of Environmental Management, 111623*.
Wright, D. H. (1983). Species-energy theory: an extension of species-area theory. *Oikos*, 496–506.
Youngentob, K. N., Yoon, H. J., Stein, J., Lindenmayer, D. B., & Held, A. A. (2015). Where the wild things are: using remotely sensed forest productivity to assess arboreal marsupial species richness and abundance. *Diversity and Distributions, 21*, 977–990.

Effects of environmental conditions on ICESat-2 terrain and canopy heights retrievals in Central European mountains

Vítězslav Moudrý, Kateřina Gdulová, Lukáš Gábor, Eliška Šáarovcová, Vojtěch Barták, Francois Leroy, Olga Špatenková, Duccio Rocchini, Jiří Prošek

Remote Sensing of Environment, doi:
https://doi.org/10.1016/j.rse.2022.113112

Remote Sensing of Environment 279 (2022) 113112

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Effects of environmental conditions on ICESat-2 terrain and canopy heights retrievals in Central European mountains

Vítězslav Moudrý [a,b,c,*], Kateřina Gdulová [a], Lukáš Gábor [a], Eliška Šárovcová [a], Vojtěch Barták [a], Francois Leroy [a], Olga Špatenková [a], Duccio Rocchini [a,d], Jiří Prošek [a,c]

[a] Department of Spatial Sciences, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Praha-Suchdol, 165 00, Czech Republic
[b] Institute for Environmental Studies, Faculty of Science, Charles University, Benátská 2, 12801 Prague 2, Czech Republic
[c] Institute of Botany of the Czech Academy of Sciences, Zámek 1, CZ-252 43 Průhonice, Czech Republic
[d] BIOME Lab, Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum University of Bologna, Bologna, Italy

ARTICLE INFO

Edited by Jing M. Chen

Keywords:
ATL08
Evaluation
Filtering
Forest
Ground
Laser altimeter
Radiometry
Scattering

ABSTRACT

The ICESat-2 ATL08 land and vegetation product includes several flags that can be used for the assessment of LiDAR-environment interactions and can help select data of the highest quality. However, the usability of these flags has not been sufficiently studied to date. Here, we aimed to evaluate the effects of atmospheric scattering, the presence of snow, canopy cover, terrain slope, beam strength, and solar angle on the accuracy of terrain and canopy height of the ATL08 product as well as on providing recommendations on how to filter data in order to minimize errors. We evaluated the vertical accuracy of ATL08 terrain and canopy height in European mountains by comparing them with the digital terrain model and canopy height model derived from airborne laser scanning data. Our results indicate that the assessment of atmospheric effects using the cloud confidence flag (*cloud_flag_atm*; i.e. number of cloud layers) is better than the previously used multiple scattering warning flag (*msw_flag*). Day acquisitions with more than one layer of clouds yielded a terrain elevation RMSE of 3.22 m in forests while night acquisitions with no more than a single layer of clouds resulted in RMSE of 1.73 m. The increasing atmospheric scattering effects increased the photons' path length, resulting in terrain height underestimation. The presence of snow had a strong positive effect on the number of identified ground photons, independently of the canopy cover, but resulted in an overestimation of terrain height in higher altitudes. Accordingly, the presence of snow cover resulted in a significant underestimation of canopy height in forests. The canopy height in broadleaf/mixed as well as coniferous forests was in summer underestimated on average by 2.1 m (%ME of −15.3%) and 1.2 m (%ME of −8.2%), respectively; in winter, however, the underestimation increased to 8.5 m (%ME of −56.8%) and 5.7 m (%ME of −38.3%), respectively. Canopy height estimates had better accuracy for the strong beam (RMSE of 5.09 m; %RMSE of 35.4%) than for the weak beam (RMSE of 7.03 m; %RMSE of 51.3%). Our results show that the ATL08 terrain height accuracy decreases with uneven distribution of signal photons within individual segments and further deteriorates with increasing terrain slope. Filtering out segments with poor distribution of photons, more than one layer of clouds during the day, and snow cover in high altitudes is the best approach for minimizing the error while maximizing the number of segments left for subsequent analysis.

1. Introduction

Forests cover >4.1 billion hectares of the Earth's surface and store considerable amounts of carbon, acting as an important global carbon sink (Pan et al., 2011). Forest carbon stock is an essential component of climate action plans increasingly made by many states to implement the Paris Agreement on Climate Change and the 2030 Agenda for Sustainable Development (Hein et al., 2018). However, estimates of forest carbon stocks and rates of change remain uncertain due to data limitations and availability (Goetz and Dubayah, 2011; Pugh et al., 2019). Field inventory campaigns are labor-intensive and airborne laser scanning (ALS) surveys are too expensive; in effect, both provide estimates

* Corresponding author at: Department of Spatial Sciences, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, 16500 Praha-Suchdol, Czech Republic.
E-mail address: moudry@fzp.czu.cz (V. Moudrý).

https://doi.org/10.1016/j.rse.2022.113112
Received 23 December 2021; Received in revised form 6 May 2022; Accepted 4 June 2022
Available online 17 June 2022

from limited samples in space and time (although many developed countries provide ALS point clouds free of charge, the renewal frequency is usually low). On the other hand, space-based remote sensing methods have a great potential to map forests on a global scale (Herold et al., 2019; Marselis et al., 2022; Mulverhill et al., 2022).

NASA's Ice, Cloud, and land Elevation Satellite-2 (ICESat-2) is a space-based laser altimetry mission launched in September 2018 (Markus et al., 2017). The ICESat-2 mission is primarily designed for mapping changes in Earth's cryosphere (Brunt et al., 2019; Farrell et al., 2020), but it also provides data for other geophysical applications such as mapping forest height and above-ground biomass (Nandy et al., 2021). Onboard, the ICESat-2 satellite carries the Advanced Topographic Laser Altimeter System (ATLAS), a single-photon counting laser altimeter that allows the determination of the time of flight of individual photons (Degnan, 2002; Neumann et al., 2019). The ATLAS instrument splits the output laser pulse into three pairs of beams, each of which consists of a weak and a strong beam with an energy ratio of approximately 1:4 (Neumann et al., 2019). The single-photon approach differs from the conventional discrete laser ranging systems that require the acquisition of thousands of backscattered photons to trigger a detection (Harding et al., 2011). It also differs from full-waveform systems such as GLAS (Geoscience Laser Altimeter System), ATLAS's predecessor on board the first ICESat satellite (2003–2009; Schutz et al., 2005). GLAS produced large footprints (diameter of 65 m) with poor vertical resolution causing problems when disentangling terrain and vegetation signal (i.e. forest canopy height) in mountain environment (Chen, 2010; Hilbert and Schmullius, 2012; Bolton et al., 2013). ICESat-2 allows the detection of both terrain and canopy surface even in high relief areas and can considerably improve our ability to monitor global forest biomass (Nandy et al., 2021).

ICESat-2 mission generates geophysical products over various types of surface, one of which is the land and vegetation product (ATL08). Several studies have already contributed to the validation of ATL08 accuracy. In general, studies reported higher accuracy for terrain than canopy heights. In their evaluation of ATL08 in Finland, Neuenschwander and Magruder (2019) and Neuenschwander et al. (2020) determined the terrain accuracy – in terms of root mean square error (RMSE) – to be <1 m and that of the canopy height to be approximately 3 m. Slightly worse estimates were found for the United States by Liu et al. (2021), with the RMSE of the terrain height of about 2 m and that of canopy height of approximately 7 m. In their validation over several ecozones in the United States, Malambo and Popescu (2021) reported the mean absolute deviations (MAE) between terrain elevations and canopy heights acquired from ATL08 and by ALS to be 1.2 m and 3.5 m, respectively.

The ATL08 accuracy is positively affected by the number of retrieved ground photons as the higher density of photons reflected from a surface facilitates the correct identification of signal photons. A laser beam propagating through the atmosphere is affected by atmospheric layers such as clouds that can cause multiple scattering, thus increasing photon path length (Winker, 2003). Furthermore, the presence of clouds affects the number of photon returns from the surface; in addition, returns from low clouds may be misclassified as ground photons (Smith et al., 2019). Consequently, the precision of ground detected under several layers of clouds may be limited and render ATL08 data unsuitable for terrain and vegetation characterization. Therefore, segments acquired under adverse atmospheric conditions are typically filtered out at the beginning of data processing (Neuenschwander et al., 2020; Queinnec et al., 2021).

Besides the signal photons, ICESat-2 geolocated photon data (ATL03) also contain solar background photons (i.e., noise photons) resulting from sunlight reflected off the Earth's surface (Swatantran et al., 2016; Neumann et al., 2019). It is, therefore, necessary to perform quite extensive processing to filter out the background noise before the production of any higher-level ICESat-2 products. Indeed, it has been shown that night acquisitions (with a lower amount of solar background noise photons) have RMSE of terrain height approximately 0.15 m lower than daytime acquisitions (Liu et al., 2021). The laser pulse energy level has similar effects. The strong beam results in more detected signal photons and hence simplifies the solar background noise filtering (Neuenschwander et al., 2020); however, it has been shown previously that weak beams may provide data of equivalent quality (Malambo and Popescu, 2021).

Furthermore, the accuracy of the retrieved canopy and terrain height may be affected by the presence of snow-cover. The ATLAS laser energy level is set to detect approximately 10 photons per strong beam shot on a snow-covered surface (Neuenschwander et al., 2020). However, for land and vegetation, the number of reflected photons is considerably lower due to the lower surface reflectance. Neuenschwander et al. (2020) reported detection of approximately one photon per shot from terrain for a strong beam under no-snow conditions. Therefore, the presence of snow can considerably improve the number of detected ground photons and, hence, the accuracy of terrain height estimates compared to snow-free segments (Neuenschwander et al., 2020). On the other hand, the presence of snow cover hinders the detection of the actual surface. Depending on its thickness, the snow cover may result in an overestimation of the terrain height in the ATL08 product (Deems et al., 2013). Furthermore, the ATL08 accuracy is also considerably affected by the terrain characteristics (Tian and Shan, 2021; Liu et al., 2021), particularly in combination with canopy cover characteristics (Malambo and Popescu, 2021). Dense and tall canopies reduce the number of photons that reach the ground and are reflected back to the satellite and, consequently, complicate the terrain height retrieval (e.g. Liu et al., 2021). Conversely, sparse canopy poses a challenge for the estimates of vegetation height because relatively few photons are reflected from sparse vegetation and, as a consequence, are difficult to distinguish from the solar background noise (Neuenschwander and Pitts, 2019).

Here, we use ATL08 flags derived from auxiliary data that describe the conditions under which the data were acquired and terrain and vegetation characteristics derived form ALS and Corine land cover data to study the influence of the (i) atmospheric conditions, (ii) snow cover, (iii) the distribution of detected photons along the ground track, (iv) solar background noise, (v) laser pulse energy level, (vi) canopy cover, and (vii) terrain slope, on the accuracy of the ATL08 (version 4) terrain and canopy height retrievals in the mountain environment. In addition, we evaluated the effects of some of these characteristics on the number of detected photons. Besides being useful for the assessment of LiDAR-environment interactions, these flags are expected to provide information on data usability and have great potential to facilitate data filtering. Therefore, we also aim to provide users with the best approach for the identification of problematic measurements. This should, in turn, allow the accurate detection and improve the selection of high accuracy data necessary for the generation of higher-level products such as ATL18 gridded ground surface height, canopy height, and canopy cover estimates.

2. Data and methods

2.1. Study area

The study area covers about 4500 km^2 and consists of three Central European mountain ranges and their surrounding areas situated in Germany, Czech Republic, and Poland (Fig. 1): the Ore mountains, the Giant mountains, and the Bohemian Forest. The altitudes range between 300 and 1600 m a.m.s.l. and the terrain slopes are highly variable (0°-50°). The tree line traverses the altitudinal range of 1200–1350 m a.m.s.l. and the vegetation consists of croplands and pastures, natural grasslands, spruce monocultures, and remnants of original deciduous and mixed mountain forests.

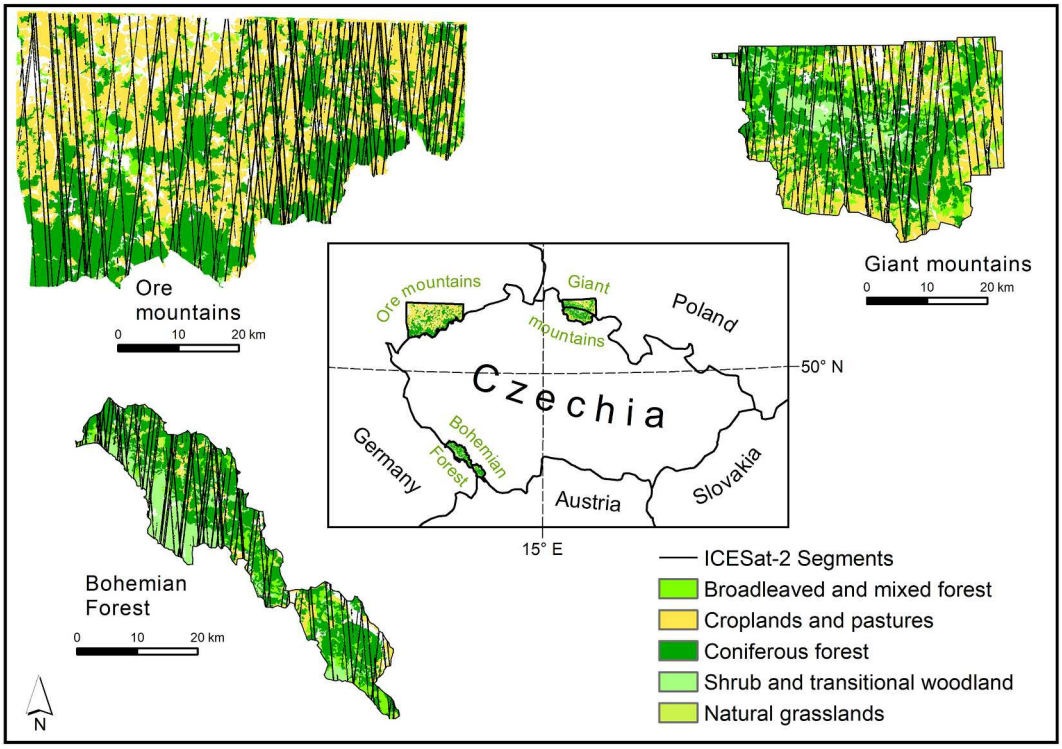


Fig. 1. Land cover derived from pan-European Corine Landcover 2018 data (Feranec et al., 2016) and location of the study areas: Ore mountains, Giant mountains, Bohemian Forest.

2.2. Reference airborne laser scanning data

The airborne LiDAR data for the study area were collected between 2012 and 2017. The data in all study areas were acquired during the leaf-on period. The LiDAR point cloud density is between 4 and 55 points per m^2 (Table 1). For the complete description of airborne laser scanning reference data processing, see Gdulová et al. (2021). The classified point clouds were processed in LAStools (version 210720). Digital terrain models (DTM) and canopy height models (CHM) for the Giant mountains and the Bohemian Forest were generated at a 1 m resolution, and the DTM and CHM for the Ore mountains were provided by "Staatsbetrieb Geobasisinformation und Vermessung Sachsen (GeoSN)" at a 2 m resolution. In order to match the horizontal datum, we used the bilinear resampling method to transform the reference ALS data into WGS84 UTM33N.

2.3. ICESat-2 ATL08 data

ICESat-2 operates in a 91-day exact repeat orbit, with 1387 orbits per cycle. The ATLAS instrument onboard the ICESat-2 satellite is a single-photon counting laser altimeter that uses green laser beams (wavelength of 532 nm) to measure the distance to the surface (Markus et al., 2017). The benefit of using a single-photon approach is that the laser can operate at a much higher repetition rate. In the case of ATLAS, the pulse repetition frequency is 10 kHz, which allows along-track sampling at 0.7 m intervals, with overlapping footprints of ~12 m. The ATLAS instrument splits the output laser pulse into three pairs of beams that are arranged to produce ground tracks with the distance of 3.3 km between the pairs and of 90 m between the tracks of each pair. The pairs consist of one weak and one strong beam with an energy ratio of approximately 1:4 (Neumann et al., 2019). Theoretically, ATLAS can detect up to 16 photons (4 × 4 detector array) per outgoing shot (a negligible fraction of all emitted photons; Markus et al., 2017; Neumann et al., 2019).

Here, we evaluated the accuracy of the terrain and canopy heights that are provided as a part of the land and vegetation height product (ATL08; Neuenschwander et al., 2021). ATL08 is derived from the global geolocated photon data product (ATL03) that serves as a single source of all photon data and ancillary information needed by higher-level products. The critical step for the development of the ATL08 product lies in the correct detection of the signal photons (i.e., photons emitted by the instrument) and filtering out the solar background noise (i.e., the photons originating from sunlight reflected off the Earth's surface; Neumann et al., 2019; Neuenschwander et al., 2021).

The filtering utilizes two signal-finding methods (i.e., histogramming approach and Differential, Regressive, and Gaussian Adaptive Nearest Neighbor) that aim to discriminate signal photons from the noise photons originating from the solar background. Subsequently, a series of iterative filters of the signal photons is adopted to capture the ground and top of canopy surface. The individual photons are then classified as ground, canopy, or noise based on their distance from the estimated canopy and ground surfaces and metrics for canopy and terrain surfaces are then provided at a segments of 100 m × 12 m (along the ground track × cross the ground track). The segment size was chosen so that a sufficient number of photons reflecting from both the terrain and canopy surfaces were available for terrain and canopy metrics estimation (Popescu et al., 2018; Neuenschwander and Pitts, 2019; Neuenschwander et al., 2021).

The terrain metrics within each segment include mean (*h_te_mean*) and median (*h_te_median*) heights of all ground photons within a

Table 1
LiDAR data characteristics.

Study area	Area (km^2)	Year	Reference system (EPSG)	Point cloud density
Bohemian Forest	680	2017	31468	55 p/m^2
Ore mountains	2640	2015–2017	25833	4 p/m^2
Giant mountains (Czechia)	480	2012	5514	5 p/m^2
Giant mountains (Poland)	720	2012	2180	4 p/m^2

segment, and also two estimates of terrain height at the location of the mid-point of a segment (*h_te_interp* and *h_te_bestfit*). The canopy metrics within each segment include minimum, mean, maximum height, and several percentile heights. Previous studies mostly concentrated on evaluating the relative canopy height at the 98th percentile (Neuenschwander et al., 2020; Liu et al., 2021); mean vegetation height, however, can provide important supplementary information characterizing vegetation structure (Chen, 2010; Malambo and Popescu, 2021). Therefore, in this study, we concentrated on the mean values (*h_te_mean* and *h_mean_canopy*).

We downloaded ATL08 version 4 containing data acquired between October 2018 and September 2021 using the Earthdata portal (accessed September 2021). The ICESat-2 mean terrain height is given as ellipsoidal height while the vertical datum of ALS data is given as normal height. Therefore, to match the heights of ICESat-2 and the reference ALS data, we used a quasigeoid of Czechia. We first interpolated the height of the quasigeoid of Czechia and surrounding areas from 53,550 positions (latitude, longitude) at a resolution of 1′ x 1.5′, and, subsequently, subtracted it from the ICESat-2 mean terrain height. Note that we used canopy heights represented as relative heights above the terrain and, thus, there was no need to match their vertical datum. In order to match the horizontal datum with ALS data, we projected the centroids of ICESat-2 ATL08 segments into WGS84 UTM33N. We did not apply any horizontal geolocation offset corrections as we are interested in the accuracy from the perspective of users who typically do not have necessary data (e.g. high-resolution ALS-DTM) to assess horizontal offsets and perform the corrections. Besides, studies that performed such assessment reported that horizontal geolocation offset tended to be relatively low (between 0 and 3 m, which is much lower than the mission requirement of 6.5 m; Neuenschwander et al., 2020; Malambo and Popescu, 2021).

The ATL08 product includes several flags that describe target (e.g. surface reflectance) and acquisition (e.g. atmospheric scattering) characteristics and inform the user on segment data quality and usability. We combined these flags with external data sources describing the target characteristics (e.g. landcover) to identify the usability of these flags as well as the causes of bias in terrain and canopy height retrievals.

2.3.1. Atmospheric effects and surface reflectance

The atmospheric layers and cloud interferences are identified using the density-dimension algorithm, and besides providing data for atmospheric sciences (as part of the ATL09 product; Herzfeld et al., 2021), they are also delivered as a part of the ATL08 product (Neuenschwander et al., 2021). There are three flags that represent atmospheric conditions: Cloud confidence flag (*cloud_flag_atm*), Multiple scattering warning flag (*msw_flag*), and layer flag (*layer_flag*). The cloud confidence flag indicates the number of aerosols or cloud layers identified in the atmosphere (i.e., 0 means no clouds, 1 corresponds to the presence of one layer of clouds, 2 corresponds to two layers, etc.). Note that the cloud confidence flag (*cloud_flag_atm*) replaced the original parameter (*cloud_flag_asr*), which was found unsuitable for cloud cover detection over dry land due to varying surface reflectance (Neuenschwander et al., 2021). The multiple scattering warning flag is estimated based on the height, thickness, and optical depth of the layer and can range from 0 to 5 where zero means no multiple scattering (i.e. no layers were detected) and 5 the greatest scattering (i.e., an atmospheric layer that touches the ground, such as fog, blowing snow, or dust storm, was detected). Finally, the layer flag simply indicates the presence or absence of clouds. Note that all above flags are observed only for strong beams, assuming the cloud condition for the corresponding weak beams is the same.

The presence of snow can significantly increase the number of detected ground photons compared to snow-free segments due to the high reflectance of the snow surface at the wavelength of 532 nm. Therefore, we also used the snow cover mask (*segment_snowcover*) that indicates a likely presence of snow within a segment (Neuenschwander et al., 2020). This flag is extracted from the National Oceanic and Atmospheric Administration (NOAA) daily snow cover product. According to the snow cover flag, snow cover was present from November through April in the study area and the median number of ground photons per 100 m segment for snow-covered surface increased more than three times compared to the snow-free surface; in non-forested areas, the values increased from 83 to 254 and in forests from 31 to 113 when snow was present.

2.3.2. Distribution of photons within a segment

Almost all negative effects affecting signal detection are bound to finally manifest themselves in the distribution of signal photons within a segment. The variability in ground photons distribution can be assessed using the subset terrain flag (*subset_te_flag*) reflecting the distribution of ground photons within each 100 m segment (each of these is divided into five 20 m sub-segments). This flag provides the user with information on whether the signal photons used to estimate the terrain height within the segment are evenly distributed or not (Neuenschwander et al., 2021).

2.3.3. Laser pulse energy, solar background noise, terrain slope, and canopy cover

We also evaluated the effect of signal strength (i.e. strong/weak beams), solar background noise (i.e., day/night acquisitions), slope, and canopy cover on terrain and canopy height retrievals. Canopy cover was estimated from ALS-CHM within each 100 m segment as the proportion of cells with canopy height greater than two meters. Note that in some analyses, we use a simple binary representation of canopy cover – forests and non-forested areas. We derived the terrain slope from an ALS DTM. We used Horn's algorithm with a 3 × 3 cell neighborhood implemented in the Slope tool of ArcGIS (version 10.8.1). To distinguish landcover categories, we used the Corine Landcover 2018 data rather than the segment landcover flag available as a part of ATL08 as the latter is based on relatively coarse MODIS land cover data.

2.4. Data pre-processing and sample selection

We limited our evaluation to the following categories of landcover: Croplands and pastures, Broadleaf and mixed forest, Coniferous forest, Transitional woodlands and shrub, and Natural grasslands (Fig. 1). We used pan-European Corine Landcover 2018 data at a 100 × 100 m resolution (Feranec et al., 2016). We used the uncertainty of ground height estimates (*h_te_uncertainty)* flag to remove invalid segments from the evaluation (Neuenschwander et al., 2021). If the number of ground photons within a segment is below the limit (50 photons), this flag shows an invalid value (3.4028E+38) for the particular segment. In addition, we noticed that some of the differences between the ATL08 terrain height and ALS-derived DTM were extremely large (i.e. outliers with height errors of hundreds of meters that may be, for example, caused by the background noise or some instrumental errors that have been incorrectly classified as ground) and should not be used for a reliable assessment. Such extremely incorrect values were often consecutive observations within a single beam, which only supports the notion of the instrumentation error being the likely cause. We limited the effect of such data by removing outliers (i.e., 1% of segments with the greatest error, segments that are in the first 0.5% and the last 0.5% of differences between ATL08 terrain height and ALS-derived DTM). This resulted in a total of 69,624 segments retained for the assessment of the terrain height (Fig. 2a). There was no further removal of segments for any of the ATL08 terrain height accuracy analyses (i.e. all analyses use all segments, which were stratified according to evaluated flags). For assessment of the canopy height, only segments associated with forest landcover types Broadleaf and mixed forest, and Coniferous forest were used. Besides, we only used segments where >5% of classified photons were classified as canopy (Neuenschwander et al., 2021). This reduced the number to 28,658 segments retained for the assessment of canopy height (Fig. 2b).

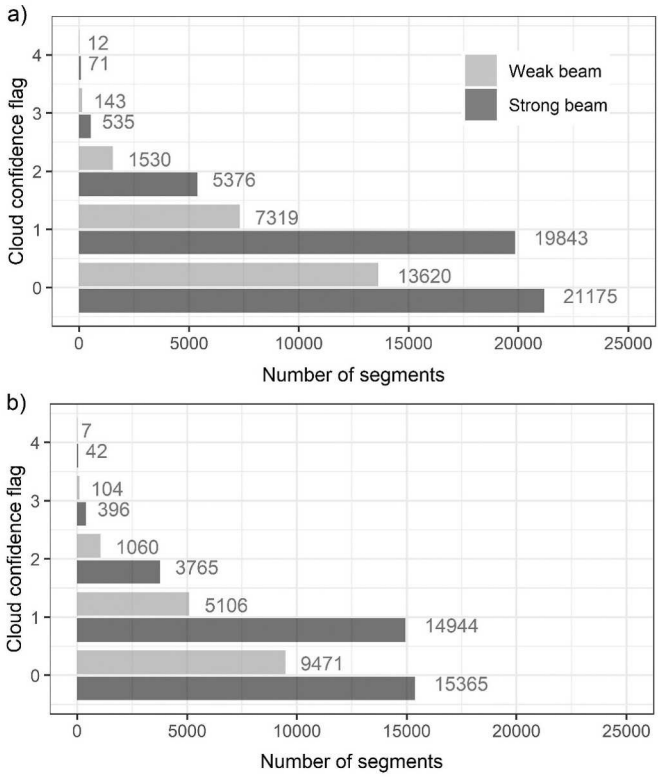


Fig. 2. Numbers of segments containing (a) terrain and (b) canopy height information grouped by beam strength (light grey – weak; dark grey – strong) and cloud confidence flag.

2.5. Assessment of accuracy and relative importance

We calculated vertical differences between the mean terrain height and mean canopy height derived from the ATL08 (*h_te_mean* and *h_mean_canopy*) and those derived from ALS-based DTM and CHM, respectively, using pairwise combinations on a segment basis. To generate segments of 100 m × 12 m, the centroids of the first and last segments were used to calculate the track inclination, and subsequently, a rectangular buffer was generated. We used the height differences to calculate several error metrics; namely, we calculated the mean error (ME) and root mean square error (RMSE), expressed as:

$$ME = \frac{1}{n}\sum_{i=1}^{n}(h_{ATL08i} - h_{REFi})$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(h_{ATL08i} - h_{REFi})^2}$$

where h_{ATL08i} is the i_{th} elevation from ATL08, h_{REFi} is the corresponding "true" elevation from ALS DTM or ALS CHM, and n is the number of samples. In addition, for the canopy height estimates, we calculated percent mean error (%ME) and percent root mean square error (% RMSE) as follows:

$$\%ME = \frac{ME}{\overline{h}_{REF}} \times 100$$

$$\%RMSE = \frac{RMSE}{\overline{h}_{REF}} \times 100$$

where $\overline{h}_{REF}$ is the mean of the reference values. We stratified the error metrics with respect to the aforementioned flags (see the previous section).

In addition, we assessed the overall importance of the flags on both the number of retrieved photons and the ATL08 height accuracy. This assessment was performed both for the ground and the canopy photons. We used flags described in Section 2.3 with binary variables (the presence of snow, beam strength, and solar angle), categorical (landcover, photons distribution in sub-segments, and cloud flag), and continuous (canopy cover, slope, and the number of terrain and canopy photons, respectively) variables. We investigated the relative contribution of each of the flags to the ME of terrain (*h_te_mean*) and canopy height (*h_mean_canopy*), respectively, and to the number of photons retrieved using Random Forest (Breiman, 2001). The hyperparameters that need to be specified for Random Forest are the minimal size for allowing node splitting (nodesize), the structure and size of the forest (i.e., the number of trees), and the number of variables considered as candidate splitting variables at each split (mtry). We used models with minimal size of a node of 5, 150 trees, and mtry set to one-third of the total number of predictors (flags). The performance of the models was evaluated by calculating the OOB coefficient of determination (R^2). To show the effect of each variable, we assessed their relative importance and produced partial dependency plots. Random Forest model provides the importance of predictors by calculating their increases in the predictive error by randomly permuting each predictor through out-of-bag observations of each tree and calculating the subsequent decrease in out-of-bag (OOB) accuracy (we scaled the relative importance of each variable so that the sum adds to 100). Random forest models were conducted in R programming language version 4.1.1. using the ranger package version 0.13.1 (Wright et al., 2020).

3. Results

We observed excellent agreement between the ATL08 mean terrain height and ALS-DTM. The overall ME of ATL08 mean terrain height with respect to ALS-DTM was −0.27 m. The RMSE of ATL08 with respect to ALS-DTM was estimated to be 1.84 m. The fact that ATL08 mean terrain height is very close to ALS-DTM is also evident from the scatterplot showing a nearly one-to-one relationship (Fig. 3a). The agreement between the ATL08 mean canopy height and ALS-CHM was considerably weaker compared to the terrain (Fig. 3b). Although most points cumulated in the vicinity of the one-to-one line, many others tend to overestimate or underestimate the mean canopy height. Overall, ATL08 underestimated ALS canopy height by an average of 2.30 m representing %ME of −16.2%. The overall RMSE was 5.67 m representing %RMSE of 39.8%.

3.1. Terrain height

3.1.1. Clouds and day/night

From the distributions, we observed a trend showing that an increasing number of cloud or aerosol layers (cloud confidence flag; *cloud_flag_atm*) resulted in larger terrain elevation errors. RMSE for daytime acquisitions increased from 1.82 m (cloud-free acquisitions) to 2.74 m (cloud confidence flag equal to two). In addition, our results show a slight increase in the underestimation of terrain with the growing cloud confidence flag. ME for daytime acquisitions increased from −0.33 m (no clouds) to −0.56 m (cloud confidence flag equal to two), likely due to the increased photon path length caused by multiple scattering in clouds. However, the number of segments available for cloud confidence flag three and, especially, four was very low. Therefore, it is difficult to infer whether such high number of cloud layers further decreases the accuracy (Table 2). The density plots of the ATL08 terrain mean error show a unimodal distribution symmetric around zero for all levels of the cloud confidence flag and both daytime and nighttime acquisitions (Fig. 4). The nighttime acquisitions have a higher accuracy than daytime acquisitions and this positive effect is evident for all levels of cloud confidence flag except for the highest level with very few segments available (Table 2).

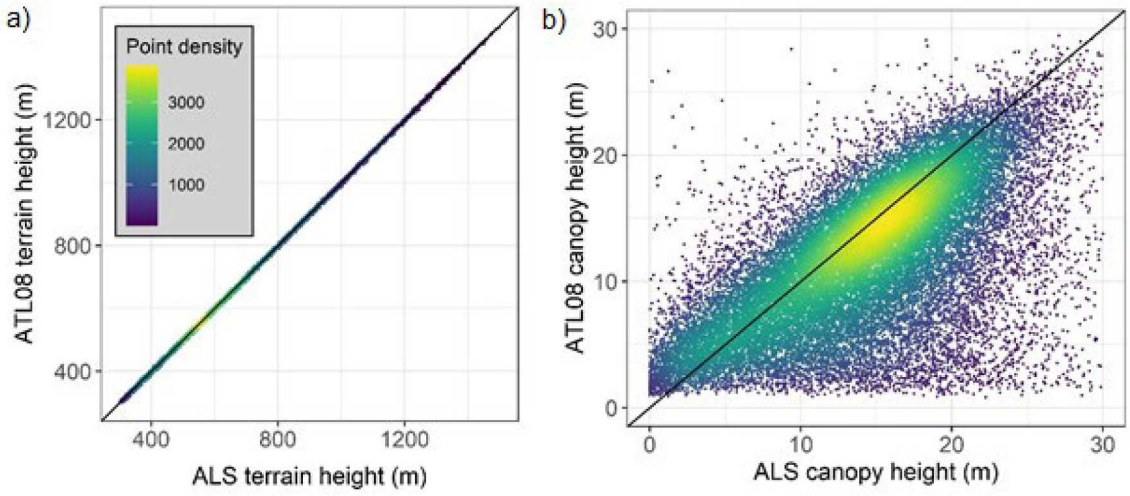


Fig. 3. Scatterplots of (a) ATL08 mean terrain height versus ALS-DTM mean terrain height and (b) ATL08 mean canopy height versus ALS-CHM mean canopy height. The solid line indicates y = x.

Table 2
The terrain height accuracy measures for increasing cloud cover (Cloud confidence flag) in relation to the solar angle (day/night).

Cloud confidence flag	Day			Night		
	Number of segments	ME (m)	RMSE (m)	Number of segments	ME (m)	RMSE (m)
0	22,555	−0.33	1.82	12,240	−0.10	1.46
1	12,274	−0.46	2.22	14,888	−0.16	1.75
2	2002	−0.56	2.74	4904	−0.17	1.59
3	216	−0.24	2.05	462	−0.22	1.82
4	4	0.33	0.51	79	−0.44	1.31

3.1.2. Snow cover and land cover

The presence of snow cover hinders the detection of the actual surface, resulting in terrain height overestimation and, consequently, in canopy height underestimation. Our results show that snow-free segments of ATL08 tend to underestimate the terrain height by lower tens of centimeters, but snow-covered segments in high altitudes overestimate it (Table 3). In low altitudes, the density plots of the ATL08 terrain mean error show a unimodal distribution symmetric around zero for both snow-covered and snow-free segments. In higher altitudes, the distribution of height differences for snow-covered segments is shifted into positive values, while for snow-free segments, it remains symmetric around zero (Fig. 5). At altitudes above 1000 m, the presence of snow resulted in an average overestimation of the terrain by almost 50 cm (Table 3). The largest terrain overestimation due to the presence of snow was observed for transitional woodlands and shrubs as well as for grasslands, which is a landcover predominating in the study area at high altitudes (Table 4; Fig. 5). On the other hand, the RMSEs for snow-free and snow-covered segments are relatively similar (Table 3), but considerably higher for forests than for grasslands and pastures (Table 4).

Table 3
Terrain accuracy measures for increasing altitudes in relation to the presence/absence of snow cover.

Elevation (m a.m.s.l.)	Snow-free segments			Snow-covered segments		
	Number of segments	ME (m)	RMSE (m)	Number of segments	ME (m)	RMSE (m)
200–500	13,560	−0.37	1.94	2424	−0.35	1.85
500–800	27,253	−0.31	1.78	6738	−0.23	1.61
800–1000	8550	−0.28	1.95	1884	0.03	2.04
1000–1200	5196	−0.24	1.90	2086	0.45	1.66
1200–1600	1334	−0.24	2.22	599	0.43	2.10

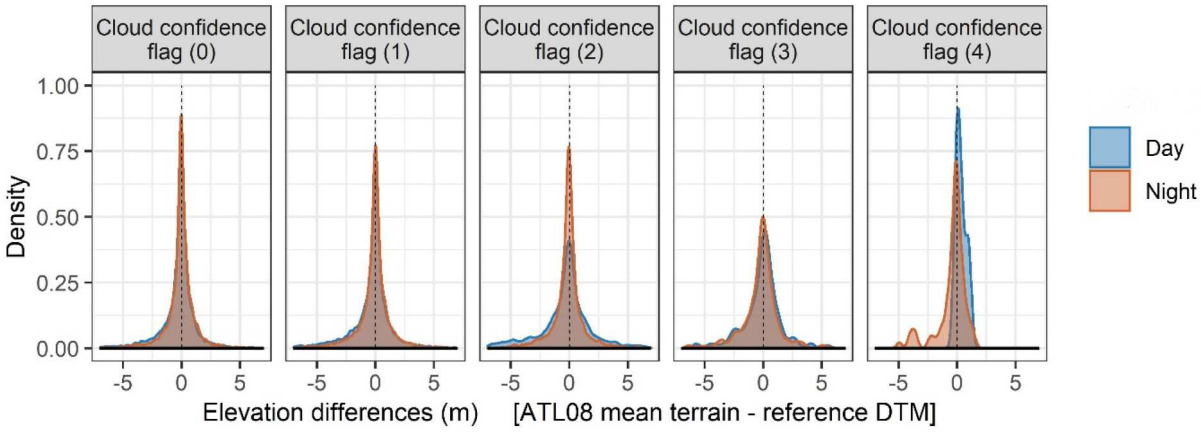


Fig. 4. Density plots illustrating the distribution of mean terrain height differences (ATL08 mean terrain height - mean DTM height), in meters, according to the cloud confidence flag grouped by the solar angle (day/night).

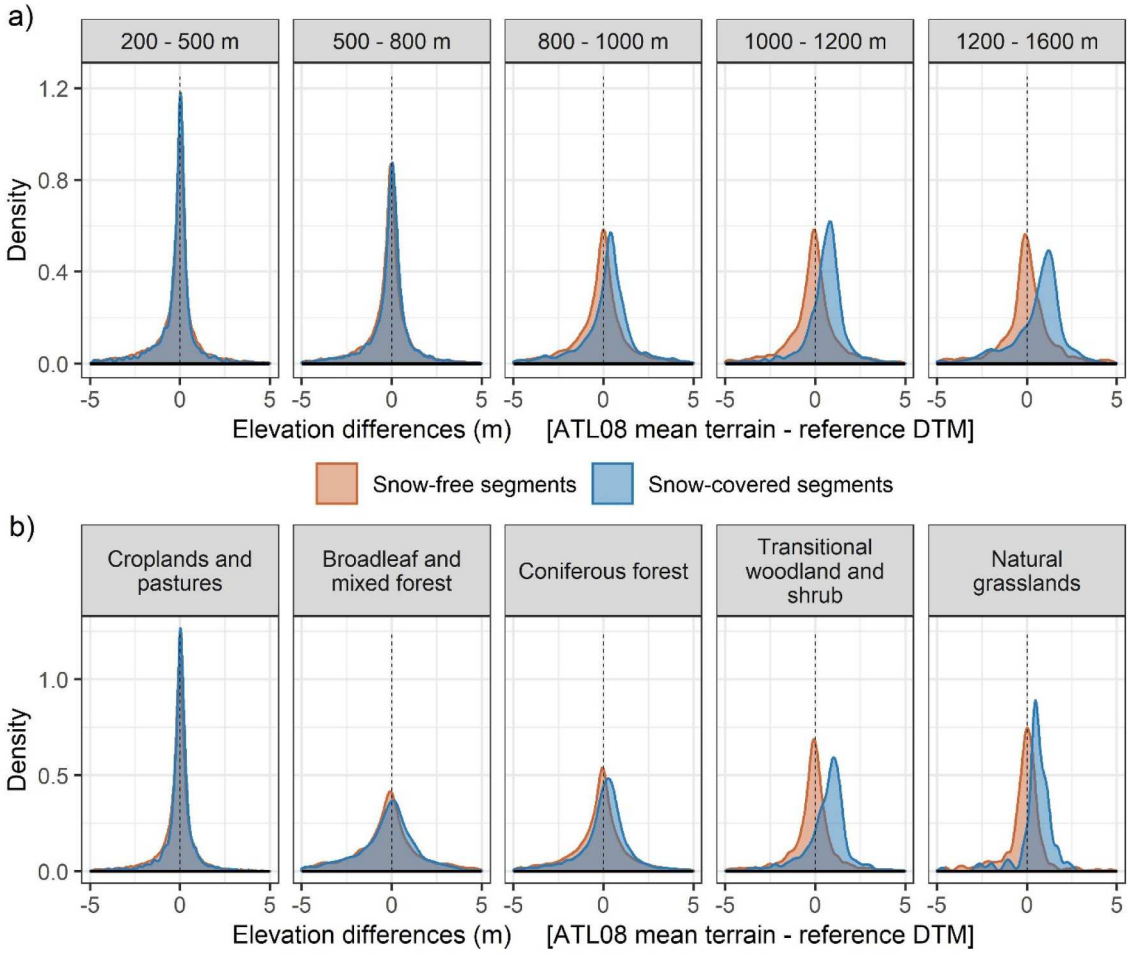


Fig. 5. Density plots illustrating the distribution of mean terrain height differences (ATL08 mean terrain height - mean DTM height), in meters, according to a) altitude and b) landcover type grouped by the snow cover flag (−covered, −free).

Table 4
Terrain height accuracy measures for main landcover types in relation to the presence/absence of snow cover.

Landcover	Snow-free segments			Snow-covered segments		
	Number of segments	ME (m)	RMSE (m)	Number of segments	ME (m)	RMSE (m)
Croplands and pastures	27,464	−0.24	1.43	5480	−0.10	1.18
Broadleaf and mixed forest	3186	−0.55	2.76	948	−0.48	2.40
Coniferous forest	20,520	−0.40	2.22	5837	−0.15	2.07
Transitional woodland and shrub	4261	−0.27	1.69	1319	0.51	1.68
Natural grasslands	462	−0.15	1.39	147	0.38	1.40

3.1.3. Interaction of the effect of clouds, snow cover, day/night, and land cover type

We examined the distribution of the vertical accuracy of ATL08 terrain height measurements grouped by cloud confidence flag for snow-covered and snow-free segments, further grouped by the solar angle (daytime vs nighttime acquisitions) and land cover type (forest/non-forest). From the distributions, we observed a trend showing that the increasing number of cloud layers resulted in considerably larger terrain elevation errors, which were particularly evident for segments with the cloud confidence flag values ≥2 (Table 5; Fig. 6); the error variability further increased for daytime acquisitions and snow- and/or forest-covered segments. In addition, weak beams appeared to be slightly more sensitive to atmospheric attenuation (Fig. 6). This shows that the cloud confidence flag (*cloud_flag_atm*) is a useful indicator of ATL08 terrain height accuracy.

3.1.4. Distribution of photons along ground track and slope

Almost all parameters influencing signal detection, both positive (snow cover, beam strength) and negative (signal attenuation by canopy or cloud cover), inevitably affect the distribution of ground returns within a segment. This, in turn, can greatly affect the estimated terrain accuracy, especially in rough terrain such as the mountain environment. Indeed, our results show that ATL08 terrain height accuracy considerably deteriorates with the lower number of 20 m sub-segments containing signal photons (Table 6; Fig. 7). In addition, the accuracy deteriorates as the slope of the terrain increases.

3.1.5. Optimal selection of segments

Based on the above analysis, we propose an optimal approach for the selection of ATL08 segments with the highest terrain accuracy as follows: (i) only data with photons in all five sub-segments should be kept (*subset_te_flag*); (ii) daytime acquisitions (*night_flag*) that have two or more layers of clouds (*cloud_flag_atm*) should be removed; and snow-covered segments (*segment_snowcover*) in high altitudes should also be removed (here, we removed snow-covered segments at altitudes above

Table 5
Root mean squared error (RMSE) of ATL08 terrain height in meters with respect to the cloud cover flag, presence of snow, landcover, and solar angle (green – lowest error; purple – highest error).

		Forest		Non-forest	
		Day	Night	Day	Night
Cloud confidence flag ≤ 1	Snow cover	2.16	1.73	1.20	1.09
Cloud confidence flag ≤ 1	Snow free	2.57	2.05	1.65	1.18
Cloud confidence flag > 1	Snow cover	2.95	2.04	2.59	2.08
Cloud confidence flag > 1	Snow free	3.22	1.96	2.09	1.14

1000 m). Such filtering removed 28% of available segments in our study area and resulted in a considerable improvement in accuracy (Fig. 8). The RMSE of removed segments was 3.02 m while the RMSE of segments left for the analysis was 1.1 m. In contrast, removal of all segments with snow cover or having cloud cover flag higher than zero (which is the commonly used approach; Neuenschwander et al., 2020; Queinnec et al., 2021) would result in the removal of 61% of data and only minimal improvement in accuracy (RMSE of 1.89 m and 1.78 of removed segments and segments left for the analysis, respectively). See Fig. 8 for density plots of the height differences between the ATL08 terrain and LiDAR ALS DTM for removed segments and segments left for the analysis.

3.2. Canopy height estimates

The error in the canopy height estimates is generally larger than that of terrain height estimates. From the distributions, we observed a trend showing that an increasing number of cloud or aerosol layers (cloud confidence flag; *cloud_flag_atm*) resulted in larger errors in canopy height estimates, especially for daytime acquisitions (Table 7). For nighttime acquisitions, the effect of cloud cover became evident for cloud confidence flag higher than two. The accuracy of nighttime acquisitions was higher than that of daytime acquisitions and this positive effect is evident for all levels of cloud confidence flag, for both strong and weak beams, and regardless of the presence/absence of snow (Table 8).

The most notable error was observed in the underestimation of the mean forest canopy height, resulting from the presence of snow on the terrain (Fig. 9). The canopy height of snow-free segments in broadleaf/mixed and in coniferous forests was on average underestimated by 2.1 m (%ME of −15.3%) and 1.2 m (%ME of −8.2%), respectively. The underestimation of snow-covered segments increased to 8.5 m (%ME of −56.8%) and 5.7 m (%ME of −38.3%), respectively; see Fig. 10 for density plots showing the distribution of the height error of ATL08 canopy height in relation to the presence/absence of snow in forests. The accuracy of the ATL08 mean canopy height also deteriorated with decreasing canopy cover and with increasing terrain slope (Fig. 11).

An examination of the agreement between ATL08 and ALS-CHM mean canopy heights by beam strength showed a lower accuracy for the weak than for the strong beam. For the weak beam, the canopy height was on average underestimated by 3.5 m representing %ME of −25.7%, while for the strong beam, the underestimation was 1.86 m, representing %ME of −12.9%. The overall RMSEs were 7.03 m (%RMSE of 51.3%) and 5.09 m (%RMSE of 35.4%) for the weak and strong beams, respectively.

3.3. Random forest variable importance

The RF models of the number of ground and canopy photons showed R^2 of 0.67 and 0.54, respectively (Fig. 12). The overall number of ground photons was strongly affected by the snow cover, beam strength, and canopy cover. The photons distribution in sub-segments, slope, cloud cover, and landcover had a moderate effect, and the effect of the solar angle was weak. Where canopy photons are concerned, the effect of slope and cloud flags on the number of photons was slightly stronger than on ground photons while the effect of canopy cover and photons distribution in sub-segments on this number was lower (Fig. 12; see Fig. S8-S9 for partial dependency plots). The RF models of ATL08 terrain and canopy height accuracy showed R^2 of 0.12 and 0.42, respectively (Fig. 12). The accuracy of ATL08 mean terrain height was particularly affected by the number of ground photons, their distribution within sub-segments, and canopy cover. The accuracy of the ATL08 mean canopy height was strongly affected by the canopy cover, presence of snow, the number of canopy photons, terrain slope, and beam strength. The effects of the remaining variables were considerably smaller (Fig. 12; see Fig. S10-S11 for partial dependency plots).

4. Discussion

4.1. Effect of atmospheric scattering on terrain and canopy height retrievals

One of the goals of our study was to examine the effect of photons attenuation in the atmosphere on the vertical accuracy of ATL08 terrain and canopy height estimates and to determine, which of the available flags is the best suited for this purpose. Our results show that cloud-free acquisitions have the best accuracy (Table 5; Fig. 6). However, the increasing number of cloud or aerosol layers alone has only a relatively low effect on the terrain and canopy height estimates accuracy. This, however, assumes that the surface was successfully detected even when cloud layers are present, which is rarely true, especially for ground photons. The increasing number of cloud or aerosol layers was associated with a decrease in the number of ground photons within individual segments (Fig. S1) as well as of the number of segments containing terrain information (Fig. 2). The number of ground photons steeply declined with cloud confidence flag values higher than one (Fig. S1).

Nighttime acquisitions over non-forested areas on snow-covered surfaces are ideal for the illustration of multiple scattering in clouds as there is no interference of other effects such as the solar background noise or canopy cover (Fig. 6). While segments acquired under the clear sky (cloud confidence flag ≤1) generally tend to overestimate the terrain height due to the presence of snow, segments acquired under a cloud cover (cloud confidence flag >1) tend, on average, to underestimate the terrain despite the presence of snow. Multiple scattering of photons in the dense clouds increases the photon path length, making the surface appear lower than it actually is. The magnitude of the error caused by the multiple-scattered signal depends on the height, thickness, and optical depth of the scattering layer (Winker, 2003).

Such effects are the reason why prior studies typically filtered out the segments acquired under adverse atmospheric conditions and did not evaluate their accuracy at all. For example, Neuenschwander et al. (2020) used only cloud-free data in their analysis. Similarly, Queinnec et al. (2021) removed all segments with multiple scattering warning flag

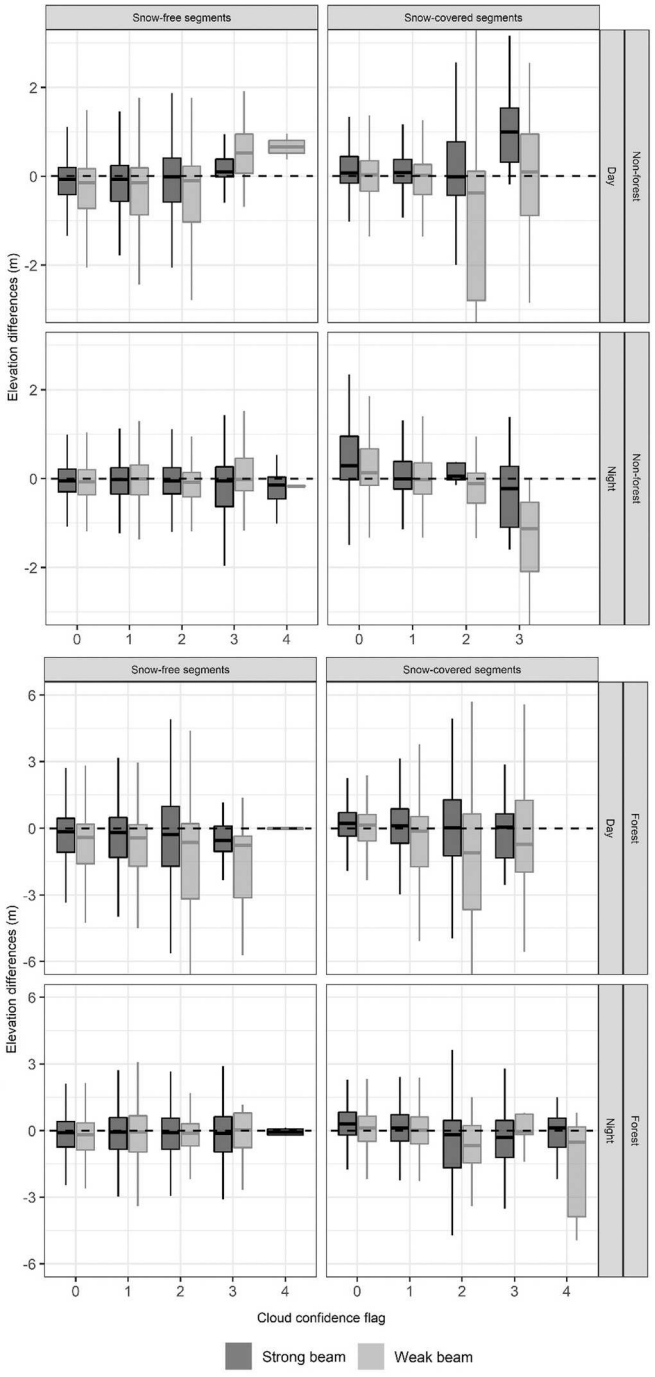


Fig. 6. Box plots showing the accuracy of ATL08 terrain height estimates (i.e., the elevation difference between ATL08 and reference DTM) considering also the cloud confidence flag (i.e. number of cloud layers), the solar angle (i.e., day and night), presence of snow (i.e., summer and winter season), landcover (i.e., forests and non-forested areas) and beam strength (strong beam; weak beam). The central horizontal line in the box indicates the median, boxes interquartile range (25th to 75th percentile), and the whiskers 1.5 times the interquartile range.

(*msw_flag*) higher than zero. However, this can lead to a significant reduction in the number of segments available for further analysis (Carabajal and Boy, 2020). For example, in our study area, such brute-force filtering would result in a 50% reduction in the number of available segments (Fig. 2), which could significantly affect the generation of higher-level products such as ATL18 gridded ground-surface height, canopy height, and canopy cover estimates. Note, however, that in other

Table 6
Root mean squared error (RMSE) of ATL08 mean terrain height in meters in relation to the number of subsegments with ground points and terrain slope (green – lowest error; purple – highest error).

Number of subsegments with ground points	Slope (0°-10°)	Slope (10°-5°)	Slope (25°-50°)
5	0.69	1.91	3.49
4	1.73	3.32	5.11
3	2.43	4.39	6.68
2	2.81	5.56	6.22
1	3.18	6.41	8.79

situations, where accuracy is much more important than the number of available segments, the removal of any potentially inaccurate segments might be a legitimate approach. In addition, it is not clear, which flag is the most useful for such filtering. Some studies use the multiple scattering warning flag (Queinnec et al., 2021) while others the cloud confidence flag (Li et al., 2021). Our results show that from three available flags (cloud confidence flag, multiple scattering warning flag, and layer flag), the cloud confidence flag is the most useful for the assessment of atmospheric effects (Fig. 6, Fig. S1 – S7). It was previously recommended to use the cloud confidence flag (*cloud_flag_atm*) only for daytime acquisitions (Palm et al., 2020), which corresponds with our results as daytime acquisitions with a high confidence flag were associated with greater error than nighttime ones (Fig. 6). For nighttime acquisitions, it was recommended to use the multiple scattering warning flag (*msw_flag*; Palm et al., 2020). However, we did not observe any patterns that would suggest that the multiple scattering warning flag is more appropriate for nighttime acquisitions; in fact, we did not observe any relationship between the number of ground photons or terrain accuracy and the multiple scattering warning flag or layer flag (*layer_flag*) at all (Fig. S1 – S7). Note, however, that we have relatively few multiple scattering warning flag records with values of 4 and, especially, 5 (i.e., the atmospheric layer that touches the ground), which, therefore, requires further exploration.

4.2. *Effects of the evaluated factors on terrain accuracy*

Besides atmospheric attenuation, the density of detected photons can be affected by landcover (canopy cover) and surface reflectance. Indeed, our results show that the decrease in the number of ground photons caused by attenuation in the atmosphere further deteriorated in forests due to the capture of photons by vegetation canopy (Fig. S1). The same effect of reduced photon penetration through the canopy cover was recently observed by Liu et al. (2021) and Malambo and Popescu (2021). In contrast, the presence of snow significantly increased the number of detected ground photons (Fig. S1). Similar seasonal variation in the average number of detected ground photons due to the presence of snow with an approximately 50% reduction in photon count when there was no snow cover was recently observed by Tian and Shan (2021). The high density of signal photons is especially important from the perspective of algorithms used for filtering as the higher density of photons reflected from a surface makes the detection of ground photons easier. Noise filtering is a critical step for accurate terrain and canopy height estimates.

It is expected that strong beams will provide products of better accuracy – given the higher density of photons, which facilitates filtering – than weak beams (Neuenschwander et al., 2020). In general, strong beams indeed resulted in more segments and more signal photons within a segment than weak beams (Fig. 2). However, we did not observe any major differences in the accuracy of segments acquired using strong and weak beams. This is in agreement with a recent study by Malambo and Popescu (2021) who have shown that weak beams may provide data of

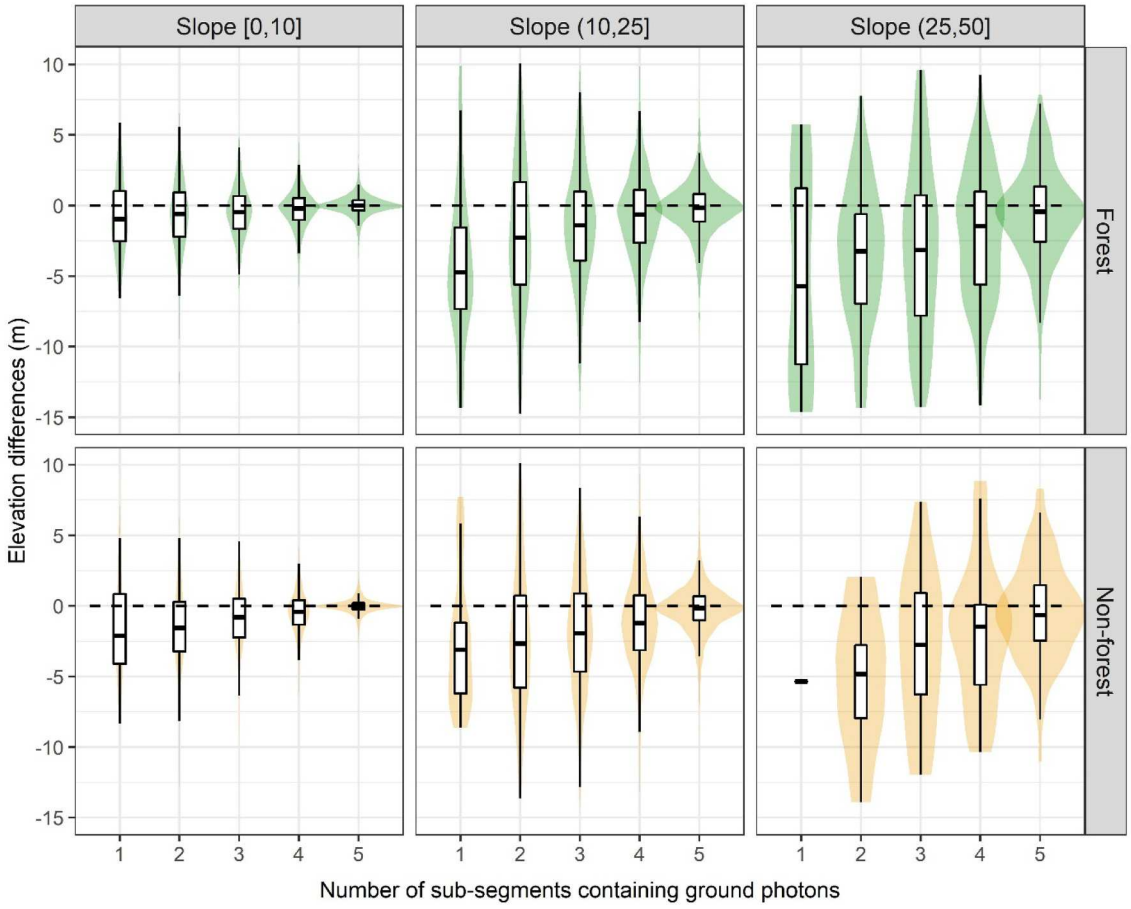


Fig. 7. Violin plots showing the ATL08 terrain height error in relation to the distribution of photons within segments (i.e. number of sub-segments containing ground photons; *subset_te_flag*), landcover (forest vs non-forest), and slope in degrees. The central horizontal line in the box indicates the median, boxes interquartile range (25th to 75th percentile), and the whiskers 1.5 times the interquartile range.

similar quality as strong beams.

The distribution of points within a segment is a better indicator of the accuracy of terrain estimates than beam strength. Almost all effects affecting signal detection, including beam strength, are reflected in the distribution of ground returns within a segment. According to Random Forest model, the accuracy of ATL08 mean terrain height was particularly affected by the number of ground photons and their distribution within sub-segments, terrain slope, and canopy cover (Fig. 12). It is, however, important to note that R^2 of this particular model was only 12%, which is relatively low and the relative importance of flags from this model has to be interpreted with caution.

Our results show that ATL08 terrain height accuracy considerably deteriorates with the decreasing number of sub-segments containing signal (ground) photons and that the magnitude of this deterioration increases with the terrain slope (Fig. 7). In flat terrain, the effect of missing ground points in sub-segments is less severe than in steep terrain as the effect of missing sub-segments on the mean terrain height estimates in the entire segment is minimal. This corresponds with the results of previous research showing that slope is the parameter with the greatest impact on the accuracy of terrain retrieval (Liu et al., 2021). The observed effect of slope may be due to the horizontal displacements of the ICESat-2 ATL08 segments relative to the reference data as the effect of horizontal displacement on vertical accuracy increases with steep slope terrain (horizontal error reported for ICESAT-2 tends to be 0–3 m; Neuenschwander et al., 2020; Malambo and Popescu, 2021).

Ground filtering is particularly complicated by solar background noise, which is especially high in daytime acquisitions (Magruder et al., 2012; Popescu et al., 2018). Indeed, our results as well as results of previous studies show that the accuracy of daytime acquisitions is worse than that of nighttime acquisitions (Neuenschwander et al., 2020; Tian and Shan, 2021). In view of this, the presence of snow cover (causing a higher density of photons reflected from the ground), can, in a way, be considered beneficial. For example, Neuenschwander et al. (2020) suggested that the presence of snow had a greater impact on terrain accuracy than beam strength or acquisition time. In addition, they suggested that a high density of ground photons from snow-covered ground can in turn improve the results of the ground filtering algorithm as it becomes less vulnerable to background noise photons. However, our results show this is true only for data acquired under favorable weather conditions. Note that Neuenschwander et al. (2020) used only cloud-free data in their analysis; in our study, the greatest error variance in the terrain quality was observed in snow-covered segments acquired during the day and under clouds (Fig. 6). It is likely that the combination of the overestimation of terrain due to the snow layer, and its underestimation due to the increased photon path length (caused by multiple scattering of photons in clouds) made the ground photon filtering even more complicated and led to the observed decrease in terrain accuracy.

On the other hand, in high altitudes, the presence of snow cover resulted in terrain overestimation and, in turn, canopy height underestimation. We observed an approximately 45 cm overestimation of the terrain height caused by the presence of snow in altitudes higher than 1000 m. For example, Neuenschwander et al. (2020) observed an average terrain overestimation of 33 cm caused by snow cover in Finland. Again, this is a reason why some studies remove snow-covered segments from their analysis (e.g. Queinnec et al., 2021). This would,

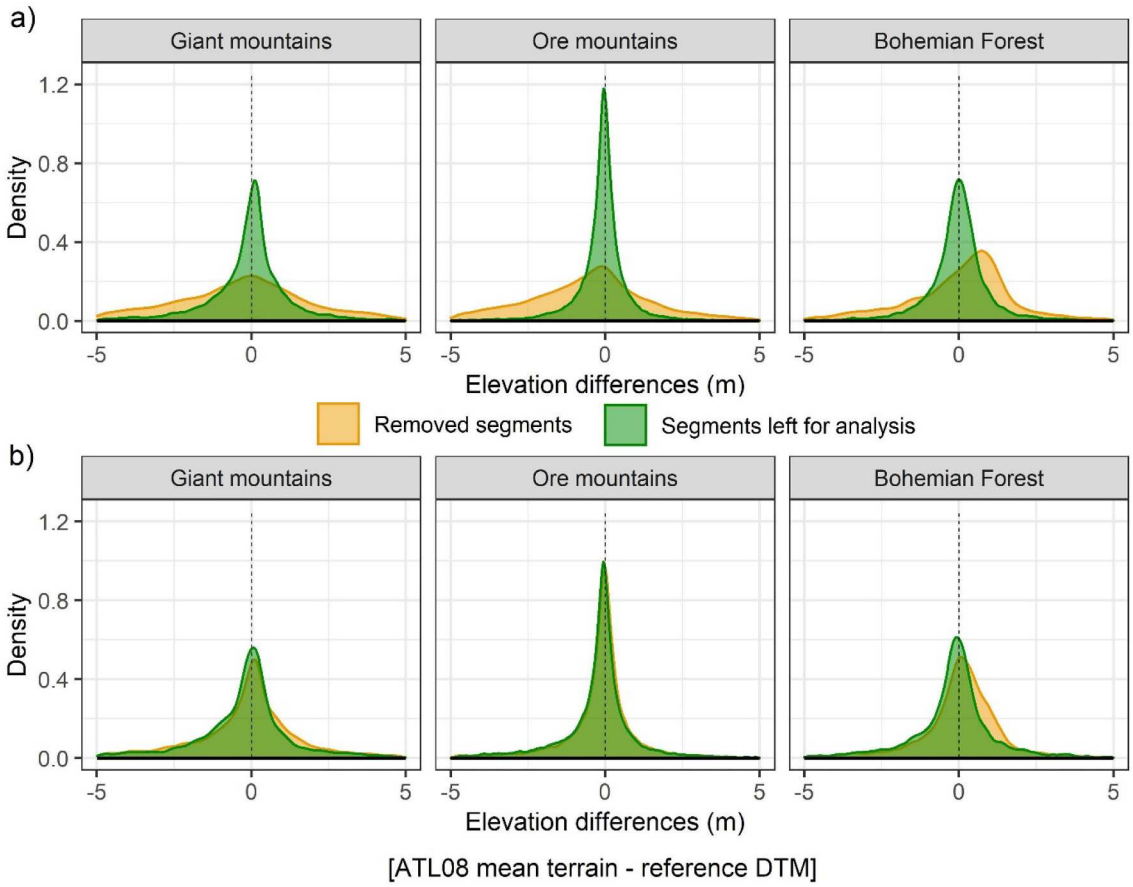


Fig. 8. Density plots illustrating the accuracy of ATL08 segments suggested for removal (orange) and segments with the highest terrain accuracy (green) left for analysis in three study areas. Two approaches for the removal of low accuracy segments are compared. The top figure (a) represents the removal of segments using the criteria suggested in this study (we removed segments with photons in less than five sub-segments; daytime acquisitions that have two or more layers of clouds; and snow-covered segments at altitudes above 1000 m ASL), while the bottom figure (b) is based on the removal of all segments with snow cover or having a cloud cover flag higher than zero, which is a common approach used in existing studies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 7
Canopy height accuracy measures for increasing cloud cover (Cloud confidence flag) in relation to solar angle (day/night).

Cloud confidence flag	Day			Night		
	Number of segments	%ME (%)	% RMSE (%)	Number of segments	%ME (%)	% RMSE (%)
0	8025	−21.8	43.2	5233	−9.8	33.9
1	4992	−23.7	49.9	6963	−11.6	31.8
2	880	−23.0	60.9	2194	−6.3	30.9
3	97	−10.0	74.9	239	−24.2	44.1
4	0			35	−3.6	17.7

however, result in a 20% reduction in the number of segments in our study area. Still, as explained above, not all snow-covered segments are necessarily erroneous and we recommend combining the flag (*segment_snowcover*) with some others (e.g., altitude, time of the year), which can serve as a proxy for the likely presence of thick snow cover.

4.3. Effects of evaluated factors on canopy accuracy

The error of canopy height estimates is considerably larger than that of terrain height and in summer, the ATL08 mean canopy height was in general underestimated by a few meters. The mean error for snow-free conditions in broadleaf/mixed forests and in coniferous forests was −2.1 m and − 1.2 m, respectively, which corresponds to the

Table 8
Percent root mean square error (%RMSE) of ATL08 canopy height with respect to the cloud cover flag, presence of snow, beam strength, and solar angle (green – lowest error; purple – highest error).

		Strong beam		Weak beam	
		Day	Night	Day	Night
Cloud confidence flag ≤ 1	Snow cover	51.5	45.7	66.1	47.6
Cloud confidence flag ≤ 1	Snow free	34.6	27.4	45.6	40.4
Cloud confidence flag > 1	Snow cover	76.5	60.0	81.5	31.5
Cloud confidence flag > 1	Snow free	40.7	29.6	68.3	32.5

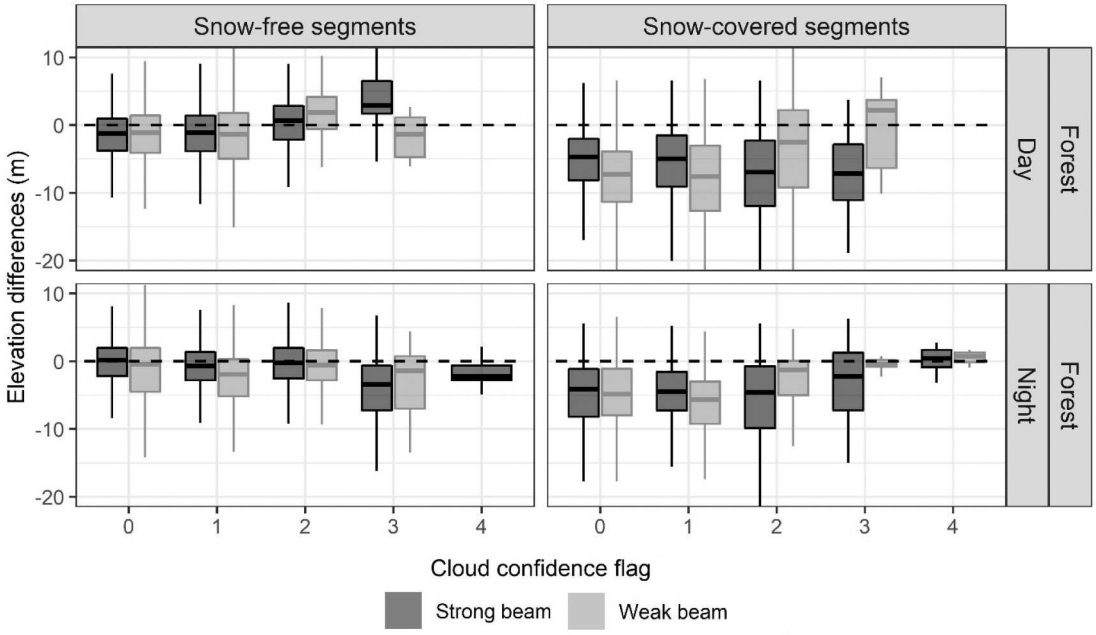


Fig. 9. Box plots showing the relationship between the accuracy of the ATL08 forest canopy height (i.e., the elevation difference between ATL08 and reference CHM) and the cloud confidence flag (i.e., number of cloud layers) with respect to the solar angle (i.e., day and night), presence of snow (i.e., summer and winter season), and beam strength (strong vs weak) beam. The central horizontal line in the box indicates the median, boxes interquartile range (25th to 75th percentile), and the whiskers 1.5 times the interquartile range.

underestimation reported by prior studies (e.g., Liu et al., 2021; Neuenschwander et al., 2020). In winter, however, the underestimation considerably increased due to the presence of snow and seasonal changes in tree foliage (Fig. 10; i.e., the reference ALS data were acquired in summer). Besides, errors in canopy height estimation were larger for the weak beam than for the strong beam. Neuenschwander et al. (2020) recommended not to use the weak beam for canopy height estimation. Our results are in line with this recommendation, particularly where daytime acquisitions are concerned (Table 8). According to the Random Forest models, the accuracy of the ATL08 mean canopy height was considerably affected by the number of retrieved canopy photons, presence of snow, and the beam strength (Fig. 12). It is, therefore, evident that the number of retrieved photons is important in estimating the canopy height. In addition, canopy cover and terrain slope also play an important role. Our results show that the accuracy of ATL08 mean canopy height depends on the extent of canopy cover and deteriorates with the increasing terrain slope (Fig. 11). This finding corresponds to the results by Liu et al. (2021) who observed deterioration in the accuracy of canopy height estimates with the extent of canopy cover and, similarly to us, also reported slope to be an important error-forming factor. The best accuracy of the mean canopy height was observed for the canopy cover ranging from 40% to 60%. Below that range, the mean height is overestimated but the error variance is relatively low. Above that range, on the other hand, canopy height is underestimated and the error variance is high, particularly on high slopes (Fig. 11). Similarly, Neuenschwander et al. (2020) reported the highest accuracy of canopy height estimates within the 40–80% canopy cover range. This is likely due to the fact that in areas with low canopy

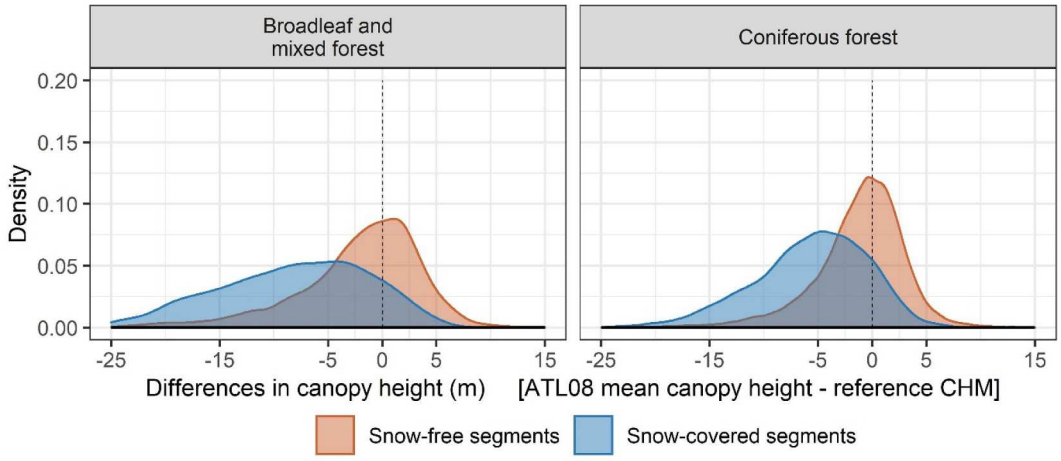


Fig. 10. Density plots showing the distribution of the absolute height error of the ATL08 mean canopy height in relation to the presence of snow according to the snow cover flag (blue and brown colors represent segments with snow and without snow, respectively) and forest type (broadleaf/mixed forest vs coniferous forest). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

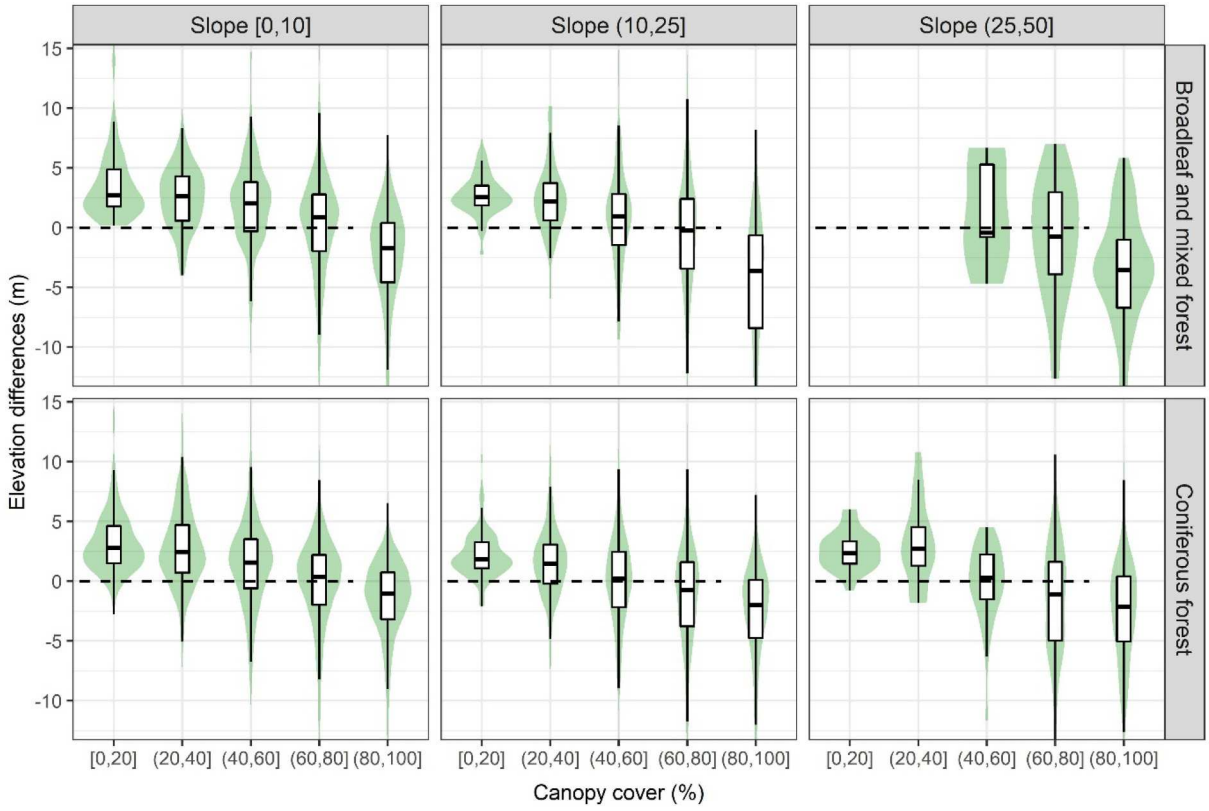


Fig. 11. Violin plots showing the ATL08 canopy height error in association with canopy cover and slope in degrees. Only segments without snow cover were used for this evaluation. The central horizontal line in the box indicates the median, boxes interquartile range (25th to 75th percentile), and the whiskers 1.5 times the interquartile range.

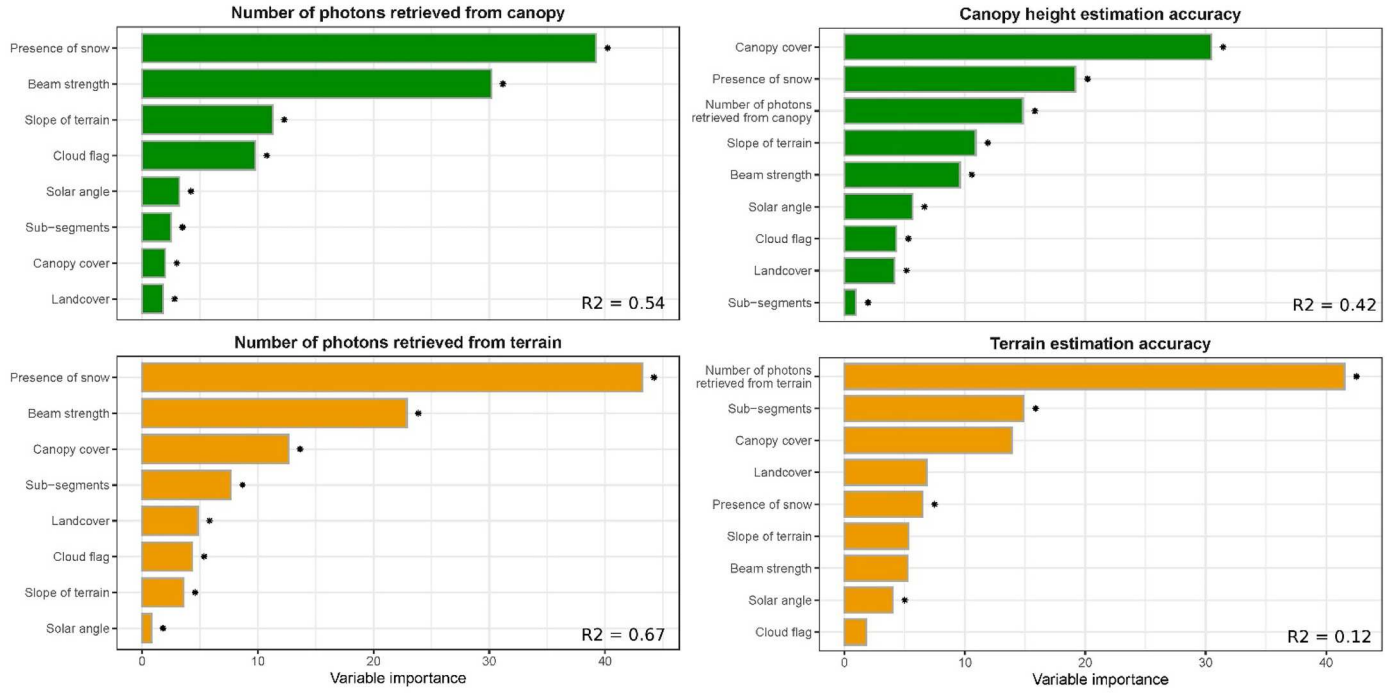


Fig. 12. Relative importance (estimated using permutation importance) of individual flags on ATL08 mean terrain and canopy height accuracy. The relative importance of each variable is scaled so that the sum adds to 100. Flags marked with an asterisk show a significant effect on the response variable. The *p*-values were computed using the Altmann et al. (2010) method with 100 permutations and the variable is considered significant if the permutation p-value is <0.05.

cover, the chance of a photon being reflected from the canopy is low. In contrast, with high canopy cover, fewer photons are returned from the ground, leading to its lower accuracy, especially on steep slopes, and consequently to poor estimates of canopy height (Pang et al., 2022).

4.4. Selection of segments with the highest accuracy

We suggest that, so far, practices for data filtering have been, on the one hand, unnecessarily strict (e.g. removing all segments with non-zero multiple scattering warning flag or all segments with snow cover), but, on the other hand, naïve and potentially resulting in keeping inaccurate records (but see Carabajal and Boy (2020) for selection of data that meet high accuracy requirements). Although this might be reasonable for exploratory studies, we need to develop more rigorous methods to minimize error while maximizing the number of segments left for subsequent analysis, such as the generation of higher-level products or the correction of global DTMs (Magruder et al., 2021). We suggest that ATL08 segments should be first filtered according to the distribution of ground photons in the segment (*subset_te_flag*; i.e., only data with photons in all five sub-segments should be kept). Although Liu et al. (2021) recently proposed that it is unnecessary to filter terrain data according to acquisition time, we suggest that it is reasonable to remove daytime acquisitions that have two or more layers of clouds (*cloud_flag_atm*) due to the synergic effect of solar background noise and increased photon path length due to atmospheric scattering. Finally, to minimize the negative effect of snow cover on terrain height estimation, we suggest removing snow-covered segments (*segment_snowcover*) in high altitudes, the combination of which serves as a substitute for the likely presence of a thick snow cover. As far as canopy height is concerned, in addition to above, we recommend removing all snow-covered segments (or even better would be removal of acquisitions in leaf-off period) and also removing weak beam acquisitions during the daytime (Table 8). This corresponds with a study by Liu et al. (2021) who recommend using strong-beam and nighttime acquisitions for canopy height estimates. Note, however, that most important influence on the accuracy of canopy height estimates has the density of canopy cover itself (Fig. 12) and the accuracy of canopy height estimates using ICESat-2 is limited (see Liu et al. (2021), for a comparison with GEDI).

5. Conclusions

We showed that atmospheric attenuation, surface reflectance, laser pulse energy level, solar background noise, canopy cover, and terrain slope are interlinked effects that affect the number of detected photons as well as the accuracy of ATL08 terrain and canopy height estimates. Generally, the error of canopy height estimates is considerably larger than that of terrain height. Results show that the accuracy of nighttime acquisitions is better than that of daytime acquisitions and that the increasing number of cloud layers causes a lower number of photons in a segment and greater error variability, especially of terrain estimates. This decrease was quite abrupt for the cloud confidence flag higher than one (i.e., data having less than two layers of clouds were of very good quality). Consequently, the accuracy of the ground detected under several layers of clouds may be limited and ATL08 segments unsuitable for terrain characterization. In general, ATL08 mean terrain estimates tend to underestimate the terrain (reference DTM) by a few tens of centimeters in summer (i.e., no snow cover), but overestimate it in winter, particularly in altitudes higher than 1000 m due to the presence of thick snow cover. Congruently, the canopy height is underestimated by a few meters in summer, and this underestimation considerably increased due to the presence of snow and seasonal loss of tree foliage in winter. The accuracy of the canopy height estimates depends on the extent of canopy cover and deteriorates with the increasing slope of the terrain. Almost all parameters affecting signal detection are reflected in the distribution of ground returns within a segment. The ATL08 terrain height accuracy deteriorates with the lower number of sub-segments containing signal photons and the magnitude of this decline increases for steep slopes. The presence of snow was associated with the strongest positive effect on the number of detected ground photons, the number of ground photons detected over snow-covered surfaces increased three times compared to snow-free surfaces. Removing segments with the poor distribution of photons, more than one layer of clouds during the day, and snow cover in high altitudes is the best approach for data filtering that minimizes errors while maximizing the number of segments left for subsequent analysis.

CRediT authorship contribution statement

Vítězslav Moudrý: Writing – original draft, Writing – review & editing, Conceptualization, Methodology, Software, Data curation, Investigation, Visualization, Supervision. **Kateřina Gdulová:** Data curation, Investigation, Writing – review & editing. **Lukáš Gábor:** Software, Investigation, Writing – review & editing. **Eliška Šárovcová:** Software, Investigation, Writing – review & editing. **Vojtěch Barták:** Software, Writing – review & editing, Formal analysis. **Francois Leroy:** Software, Writing – review & editing, Formal analysis. **Olga Špatenková:** Writing – review & editing. **Duccio Rocchini:** Writing – review & editing. **Jiří Prošek:** Writing – review & editing, Visualization, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to four anonymous reviewers for their constructive comments that significantly improved the quality of the manuscript. We are also grateful to the administrations of Šumava National Park, Krkonose Mountains National Park, and the Head Office of Geodesy and Cartography in Poland for providing the airborne laser scanning data. We also thank the “Staatsbetrieb Geobasisinformation und Vermessung Sachsen (GeoSN)” for providing access to the DTM and CHM of Saxony as open data under the dl-de/by-2-0 license. This work was supported by the Internal Grant Agency of the Faculty of Environmental Sciences, Czech University of Life Sciences Prague (project No. 2022B0035). D.R. was partially supported by the H2020 Project SHOWCASE (Grant agreement No. 862480). VM and JP were supported by long-term research development project RVO 67985939 (Czech Academy of Sciences).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2022.113112.

References

Altmann, A., Toloşi, L., Sander, O., Lengauer, T., 2010. Permutation importance: a corrected feature importance measure. Bioinformatics 26 (10), 1340–1347.
Bolton, D.K., Coops, N.C., Wulder, M.A., 2013. Investigating the agreement between global canopy height maps and airborne Lidar derived height estimates over Canada. Can. J. Remote. Sens. 39 (sup1), S139–S151.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.
Brunt, K.M., Neumann, T.A., Smith, B.E., 2019. Assessment of ICESat-2 ice sheet surface heights, based on comparisons over the interior of the Antarctic ice sheet. Geophys. Res. Lett. 46 (22), 13072–13078.
Carabajal, C.C., Boy, J.P., 2020. ICESat-2 altimetry as geodetic control. Int. Arch. Photogram. Remote Sens. Spat. Inform. Sci. 43, 1299–1306.
Chen, Q., 2010. Retrieving vegetation height of forests and woodlands over mountainous areas in the Pacific Coast region using satellite laser altimetry. Remote Sens. Environ. 114 (7), 1610–1627.

Deems, J.S., Painter, T.H., Finnegan, D.C., 2013. Lidar measurement of snow depth: a review. J. Glaciol. 59 (215), 467–479.
Degnan, J.J., 2002. Photon-counting multikilohertz microlaser altimeters for airborne and spaceborne topographic measurements. J. Geodyn. 34 (3–4), 503–549.
Farrell, S.L., Duncan, K., Buckley, E.M., Richter-Menge, J., Li, R., 2020. Mapping Sea ice surface topography in high fidelity with ICESat-2. Geophys. Res. Lett. 47 (21) e2020GL090708.
Feranec, J., Soukup, T., Hazeu, G., Jaffrain, G., (Eds.)., 2016. European Landscape Dynamics: CORINE Land Cover Data. CRC Press.
Gdulová, K., Marešová, J., Barták, V., Szostak, M., Červenka, J., Moudrý, V., 2021. Use of TanDEM-X and SRTM-C data for detection of deforestation caused by bark beetle in central European Mountains. Remote Sens. 13 (15), 3042.
Goetz, S., Dubayah, R., 2011. Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. Carbon Manag. 2 (3), 231–244.
Harding, D.J., Dabney, P.W., Valett, S., 2011. Polarimetric, two-color, photon-counting laser altimeter measurements of forest canopy structure. In: International Symposium on Lidar and Radar Mapping 2011: Technologies and applications. International Society for Optics and Photonics, p. 828629.
Hein, J., Guarin, A., Frommé, E., Pauw, P., 2018. Deforestation and the Paris climate agreement: an assessment of REDD+ in the national climate action plans. Forest Policy Econ. 90, 7–11.
Herold, M., Carter, S., Avitabile, V., Espejo, A.B., Jonckheere, I., Lucas, R., De Sy, V., 2019. The role and need for space-based forest biomass-related measurements in environmental management and policy. Surv. Geophys. 40 (4), 757–778.
Herzfeld, U., Hayes, A., Palm, S., Hancock, D., Vaughan, M., Barbieri, K., 2021. Detection and height measurement of tenuous clouds and blowing snow in ICESat-2 ATLAS data. Geophys. Res. Lett. 48 (17) e2021GL093473.
Hilbert, C., Schmullius, C., 2012. Influence of surface topography on ICESat/GLAS forest height estimation and waveform shape. Remote Sens. 4 (8), 2210–2235.
Li, B., Xie, H., Liu, S., Tong, X., Tang, H., Wang, X., 2021. A method of extracting high-accuracy elevation control points from ICESat-2 altimetry data. Photogramm. Eng. Remote. Sens. 87 (11), 821–830.
Liu, A., Cheng, X., Chen, Z., 2021. Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals. Remote Sens. Environ. 264, 112571.
Magruder, L.A., Wharton III, M.E., Stout, K.D., Neuenschwander, A.L., 2012. Noise filtering techniques for photon-counting ladar data. In: Laser Radar Technology and Applications XVII, vol. 8379. International Society for Optics and Photonics, p. 83790Q. May.
Magruder, L., Neuenschwander, A., Klotz, B., 2021. Digital terrain model elevation corrections using space-based imagery and ICESat-2 laser altimetry. Remote Sens. Environ. 264, 112621.
Malambo, L., Popescu, S.C., 2021. Assessing the agreement of ICESat-2 terrain and canopy height with airborne lidar over US ecozones. Remote Sens. Environ. 266, 112711.
Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Zwally, J., 2017. The ice, cloud, and land elevation Satellite-2 (ICESat-2): science requirements, concept, and implementation. Remote Sens. Environ. 190, 260–273.
Marselis, S.M., Keil, P., Chase, J.M., Dubayah, R.O., 2022. The use of GEDI canopy structure for explaining variation in tree species richness in natural forests. Environ. Res. Lett. 17 (4), 045003. https://doi.org/10.1088/1748-9326/ac583f.
Mulverhill, C., Coops, N.C., Hermosilla, T., White, J.C., Wulder, M.A., 2022. Evaluating ICESat-2 for monitoring, modeling, and update of large area forest canopy height products. Remote Sens. Environ. 271, 112919. https://doi.org/10.1016/j.rse.2022.112919.
Nandy, S., Srinet, R., Padalia, H., 2021. Mapping forest height and aboveground biomass by integrating ICESat-2, Sentinel-1 and Sentinel-2 data using random forest algorithm in northwest Himalayan foothills of India. Geophys. Res. Lett. 48 (14) e2021GL093799.
Neuenschwander, A.L., Magruder, L.A., 2019. Canopy and terrain height retrievals with ICESat-2: a first look. Remote Sens. 11 (14), 1721.
Neuenschwander, A., Pitts, K., 2019. The ATL08 land and vegetation product for the ICESat-2 Mission. Remote Sens. Environ. 221, 247–259.
Neuenschwander, A., Guenther, E., White, J.C., Duncanson, L., Montesano, P., 2020. Validation of ICESat-2 terrain and canopy heights in boreal forests. Remote Sens. Environ. 251, 112110.
Neuenschwander, A., Popescu, S., Nelson, R., Harding, D., Pitts, K., Robbins, J., Sheridan, R., 2021. Ice, Cloud, and Land Elevation Satellite 2 (ICESat-2) Algorithm Theoretical Basis Document (ATBD) for Land-Vegetation Along-Track Products (ATL08) release 004. National Aeronautics and Space Administration, Washington, DC, USA. Tech. Rep., Accessed: Nov. 18, 2021. [Online]. Available. https://nsidc.org/sites/nsidc.org/files/technical-references/ICESat2_ATL08_ATBD_r004.pdf.
Neumann, T.A., Martino, A.J., Markus, T., Bae, S., Bock, M.R., Brenner, A.C., 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 Sens. Environ. 233, 111325.
Palm, S., Yang, Y., Herzfeld, U., 2020. ICESat-2 Algorithm Theoretical Basis Document for the Atmosphere, Part I: Level 2 and 3 Data Products Version 8.3. National Aeronautics and Space Administration, Washington, DC, USA. Tech. Rep., Accessed: Nov. 20, 2021. [Online]. Available. https://nsidc.org/sites/nsidc.org/files/technical-references/ICESat2_ATL04_ATL09_ATBD%28part%201%29_r003.pdf.
Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Hayes, D., 2011. A large and persistent carbon sink in the world's forests. Science 333 (6045), 988–993.
Pang, S., Li, G., Jiang, X., Chen, Y., Lu, Y., Lu, D., 2022. Retrieval of forest canopy height in a mountainous region with ICESat-2 ATLAS. For. Ecosyst. 9, 100046. https://doi.org/10.1016/j.fecs.2022.100046.
Popescu, S.C., Zhou, T., Nelson, R., Neuenschwander, A., Sheridan, R., Narine, L., Walsh, K.M., 2018. Photon counting LiDAR: an adaptive ground and canopy height retrieval algorithm for ICESat-2 data. Remote Sens. Environ. 208, 154–170.
Pugh, T.A., Lindeskog, M., Smith, B., Poulter, B., Arneth, A., Haverd, V., Calle, L., 2019. Role of forest regrowth in global carbon sink dynamics. Proc. Natl. Acad. Sci. 116 (10), 4382–4387.
Queinnec, M., White, J.C., Coops, N.C., 2021. Comparing airborne and spaceborne photon-counting LiDAR canopy structural estimates across different boreal forest types. Remote Sens. Environ. 262, 112510.
Schutz, B.E., Zwally, H.J., Shuman, C.A., Hancock, D., DiMarzio, J.P., 2005. Overview of the ICESat mission. Geophys. Res. Lett. 32 (21).
Smith, B., Fricker, H.A., Holschuh, N., Gardner, A.S., Adusumilli, S., Brunt, K.M., Siegfried, M.R., 2019. Land ice height-retrieval algorithm for NASA's ICESat-2 photon-counting laser altimeter. Remote Sens. Environ. 233, 111352.
Swatantran, A., Tang, H., Barrett, T., DeCola, P., Dubayah, R., 2016. Rapid, high-resolution forest structure and terrain mapping over large areas using single photon lidar. Scientific reports 6 (1), 1–12.
Tian, X., Shan, J., 2021. Comprehensive evaluation of the ICESat-2 ATL08 terrain product. IEEE Trans. Geosci. Remote Sens. 59 (10), 8195–8209. https://doi.org/10.1109/TGRS.2021.3051086.
Winker, D.M., 2003. Accounting for multiple scattering in retrievals from space lidar. In: 12th International Workshop on Lidar Multiple Scattering Experiments, vol. 5059. International Society for Optics and Photonics, pp. 128–139. April.
Wright, M.N., Wager, S., Probst, P., 2020. Ranger: A Fast Implementation of Random Forests. R Package Version 0.13.1.

How to find accurate terrain and canopy height GEDI footprints in temperate forests and grasslands?

Vítězslav Moudrý, Jiří Prošek, Suzanne Marselis, Jana Marešová, Eliška Šárovcová, Kateřina Gdulová, Giorgi Kozhoridze, Michele Torresani, Duccio Rocchini, Anette Eltner, Xiao Liu, Markéta Potůčková, Adéla Šedová, Pablo Crespo-Peremarch, Jesús Torralba, Luis A. Ruiz, Michela Perrone, Olga Špatenková, Jan Wild

Under review in Earth and Space Science

Title: How to find accurate terrain and canopy height GEDI footprints in temperate forests and grasslands?

Authors:

Vítězslav Moudrý [a,b,c], Jiří Prošek [a,c], Suzanne Marselis [d], Jana Marešová [a], Eliška Šárovcová [a], Kateřina Gdulová [a], Giorgi Kozhoridze [a], Michele Torresani [e], Duccio Rocchini [a,f], Anette Eltner [g], Xiao Liu [g], Markéta Potůčková [h], Adéla Šedová [h], Pablo Crespo-Peremarch [i], Jesús Torralba [i], Luis A. Ruiz [i], Michela Perrone [a], Olga Špatenková [a], Jan Wild [a,c]

[a]Department of Spatial Sciences, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, 16500 Praha-Suchdol, Czech Republic.

[b]Institute for Environmental Studies, Faculty of Science, Charles University, Benátská 2, 12801 Prague 2, Czech Republic.

[c]Institute of Botany of the Czech Academy of Sciences, Zámek 1, 252 43 Průhonice, Czech Republic.

[d]Institute of Environmental Sciences, Leiden University, Einsteinweg 2, 2333 CC Leiden, the Netherlands.

[e]Free University of Bolzano/Bozen, Faculty of Science and Technology, Piazza Universitá / Universitätsplatz 1, 39100, Bolzano/Bozen, Italy.

[f]BIOME Lab, Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum University of Bologna, via Irnerio 42, 40126, Bologna, Italy.

[g]Institute of Photogrammetry and Remote Sensing, Technische Universität Dresden, 01062 Dresden, Germany

[h]Charles University, Faculty of Science, Department of Applied Geoinformatics and Cartography Albertov 6, 128 43 Praha 2, Czech Republic.

[i]Geo-Environmental Cartography and Remote Sensing Group (CGAT), Universitat Politècnica de València, Camí de Vera s/n, 46022 Valencia, Spain.

Corresponding author: Vítězslav Moudrý (moudry@fzp.czu.cz); ORCID: 0000-0002-3194-451X

Abstract

Filtering approaches on GEDI data differ considerably across existing studies and it is yet unclear which method is the most effective. We conducted an in-depth analysis of GEDI's vertical accuracy in mapping terrain and canopy heights across three study sites in temperate forests and grasslands in Spain, California, and New Zealand. We started with unfiltered data (2,081,108 footprints) and describe a workflow for data filtering using Level 2A parameters and for geolocation error mitigation. We found that retaining observations with at least one detected mode eliminates noise more effectively than sensitivity. The accuracy of terrain and canopy height observations depended considerably on the number of modes, beam sensitivity, landcover, and terrain slope. In dense forests, a minimum sensitivity of 0.9 was required, while in areas with sparse vegetation, sensitivity of 0.5 sufficed. Sensitivity greater than 0.9 resulted in an overestimation of canopy height in grasslands, especially on steep slopes, where high sensitivity led to the detection of multiple modes. We suggest excluding observations with more than five modes in grasslands. We found that the most effective strategy for filtering low-quality observations was to combine the quality flag and difference from TanDEM-X, striking an optimal balance between eliminating poor-quality data and preserving a maximum number of high-quality observations. Positional shifts improved the accuracy of GEDI terrain estimates but not of vegetation height estimates. Our findings guide users to an easy way of processing of GEDI footprints, enabling the use of the most accurate data and leading to more reliable applications.

Plain language summary

The Global Ecosystem Dynamics Investigation (GEDI) collected terrain and canopy observations using laser altimetry. The quality of terrain and canopy observations is influenced by acquisition conditions and land(cover) characteristics. Consequently, a considerable amount of GEDI observations is discarded as noise, and further filtering is necessary to retain only high-quality observations. Our objective was to assess how environmental and acquisition characteristics influence the accuracy of terrain and canopy height of GEDI observations. Although the main objective of the GEDI mission was to map forests, we also focused on grasslands. GEDI serves not only as an essential source of information on canopy height but also provides accurate terrain observations. Furthermore, it is important to know that GEDI does not overestimate the height of low vegetation as this can result in an overestimation of carbon storage. We distinguished four steps in the GEDI data processing: (1) removal of noise observations, (2) removal of low-quality data, (3) effect of additional acquisition characteristics, and (4) mitigation of geolocation error. We found that the accuracy of terrain and canopy height observations depended considerably on the number of detected modes, beam sensitivity, landcover, and terrain slope.

Keywords: error, filtering, geolocation, height, terrain, vegetation

1. Introduction

Forests play a key role in biodiversity conservation and regulation of the global carbon cycle (Mo et al., 2023; Pan et al., 2011). It is well recognized that the vertical and horizontal structure of forests, including their height, are important indicators of biodiversity (Cazzolla Gatti et al., 2017; Davies & Asner, 2014) and vegetation structure is, therefore, considered one of the six essential classes of biodiversity variables (Moudrý et al., 2023; Skidmore et al., 2021;). Furthermore, vegetation structure is directly linked to the above-ground biomass and primary productivity, both being critical components of carbon stock and flux (Fisher et al., 2019). Vegetation structure measurements are, therefore, essential for creating climate action plans, the preparation of which is underway in many countries (Grassi et al., 2017; Hoover & Smith, 2021).

However, until recently, we lacked comprehensive global data on the spatial patterns of forest structure and available estimates remained uncertain due to data limitations and availability (Herold et al., 2019). This has changed with the two recent space-based laser altimetry missions launched in 2018 that have mapped the forest structure with unprecedented accuracy on a global scale (Dubayah et al., 2022; Neuenschwander & Pits, 2019). These two missions, GEDI (Global Ecosystem Dynamics Investigation) and ICESat-2 (Ice, Cloud, and Land Elevation Satellite-2), provide direct measurements of the terrain and canopy heights, and, therefore, considerably improve our ability to monitor forests globally (Dubayah et al., 2020; Liu et al., 2021; Markus et al., 2017). The GEDI mission specifically aims to measure forest structure and biomass in tropical and temperate regions. It is the first spaceborne lidar mission specifically designed for such a purpose (Dubayah et al., 2020). The GEDI data have been, for example, used to assess the role of forest structure in biodiversity patterns (Marselis et al., 2022; Torresani et al., 2023), to improve models of animal-environment relationships (Smith et al., 2022), or to assess the effectiveness of protected areas in conserving vegetation structure and carbon stocks (Ceccherini et al., 2023; Liang et al., 2023).

The GEDI instrument was mounted on the International Space Station (ISS) and collected data between April 2019 and March 2023. GEDI fired 16 billion laser pulses at the Earth's surface each year. Compared to ICESat-2, it had a higher sampling density in temperate and tropical forests, but its spatial coverage was limited to the latitude range of 51.6°S to 51.6°N (ISS inclination angle). GEDI has a 25 m diameter footprint, which is recommended for measuring canopy height using spaceborne lidar for biodiversity and habitat science (Bergen et al., 2009). The location of each GEDI footprint is computed using the observed range combined with the instrument pointing and position (derived from star tracker and GPS sensors; Beck et al., 2021). The horizontal geolocation error of GEDI footprints is 10 ±

2 m (V002) and is expected to improve in upcoming data versions (Beck et al., 2021). An empirical vertical accuracy of GEDI canopy height estimates is approximately 2–3 m (Hakkenberg et al., 2023a), which more or less meets the recommended vertical accuracy for measuring canopy height using spaceborne lidar (Bergen et al., 2009). Such accuracy, however, assumes ideal measurement conditions, which are rarely met; the actual accuracy is, therefore, degraded due to environmental conditions during the data acquisition (e.g. Dorado-Roda et al., 2021; Li et al., 2023; Liu et al., 2021). GEDI's lasers operate at a 1,064 nm wavelength (near-infrared), thus unable to penetrate dense cloud cover (Dubayah et al., 2020). Therefore, only approximately 50% of pulses were expected to provide a useful signal, while the remaining 50% were expected to be lost or degraded (eoPortal, 2016; Dubayah et al., 2020; Fayad et al., 2021). Further, degraded accuracy may have resulted from the interaction of the laser with the Earth's atmosphere, solar background noise, and/or processes (algorithms) used to determine the height metrics (Adam et al., 2020; Liu et al., 2021; Quiros et al., 2021; Wang et al., 2022).

Due to the low accuracy of some of the GEDI observations, they must be filtered before being used to estimate terrain or vegetation structure. Indeed, the selection of accurate GEDI footprints is crucial for the preparation of downstream GEDI data products, such as gridded mean canopy height (Dubayah et al., 2021a) and above-ground biomass density (Kellner et al., 2022), or any other analysis using GEDI data (e.g. Hoffrén et al., 2023; Potapov et al., 2021). In this work, we focused on GEDI acquisitions over temperate forests. However, it is equally important to tackle the issue of the overestimation of low vegetation (such as grasslands) as this can result in an overestimation of carbon storage. Moreover, such data are important for mapping landcover types (Dwiputra et al., 2023) and the generation of global canopy height maps (i.e., both areas of high and low vegetation are needed to fit the models; Potapov et al., 2021). In addition, GEDI is not only an important source of information on canopy height, but also of accurate terrain observations that are used to validate existing and create new global digital elevation models (Narin & Gullu, 2023; Narin et al., 2024; Pronk et al., 2023). Therefore, in this study, we focus on both terrain and canopy height in both forests and grasslands.

However, identifying usable (i.e., high-quality) observations can be challenging. To select high-quality data, the GEDI user guide (Beck et al., 2021) recommends the use of power beams only for the description of dense forests, GEDI footprints acquired at night to avoid the negative effects of solar background illumination on waveform quality, and, finally, assessing footprint sensitivity. The easiest way of data filtering lies in the use of the *quality_flag* parameter as suggested by the GEDI user guide (Beck et al., 2021). Many studies using GEDI data use this flag for data filtering (see Figure 1), but it has also been suggested that the flag tends to remove usable footprints (Geremew et al., 2023). In addition, other approaches have been adopted in existing studies, including, for example, comparing the data to some existing global DEMs (e.g., TanDEM-X DEM or SRTM; Lahssini et al., 2022; Urbazaev

et al., 2022) or filtering according to the results acquired using six different algorithm setting groups and keeping only observations where all these results are similar (e.g. Potapov et al., 2021). Terrain slope is another factor contributing to the canopy height estimation error (Chen, 2010; Wang et al., 2022). It is, therefore, potentially reasonable to keep only footprints located on relatively flat terrain (e.g. Simard et al., 2011). GEDI data filtering, on the one hand, ensures the use of high-quality data, but, on the other hand, considerably reduces the number of available GEDI footprints (Figure 1). For example, in the study by Hoffrén et al. (2023), filtering resulted in the use of only 16% of all footprints available in their study area. Similarly, Cobb et al. (2023) flagged only 14% of footprints as high-quality.

For inexperienced users, it may be difficult to properly apply available filtering approaches and parameters, which can lead to an unnecessary loss of accurate data, or, vice versa, to an inclusion of inaccurate data. In this study, we evaluate three study sites in temperate forests and grasslands in Spain, California, and New Zealand as examples to show how the individual parameters used for filtering (Figure 1) complement each other. We aim (i) to evaluate the usability of the number of detected modes (i.e., of waveform maxima in the return signal) and sensitivity for removal of noise observations, (ii) to quantify the success of three filtering methods in preserving accurate measurements and eliminating inaccurate ones, (iii) to evaluate the effects of acquisition (e.g., degrade flag, beam strength, acquisition time) and environmental (landcover, slope) characteristics on GEDI observations, and (iv) to assess the geolocation error of footprints. In contrast to prior studies (Figure 1), we start with unfiltered data and outline the whole process of high-quality data selection.

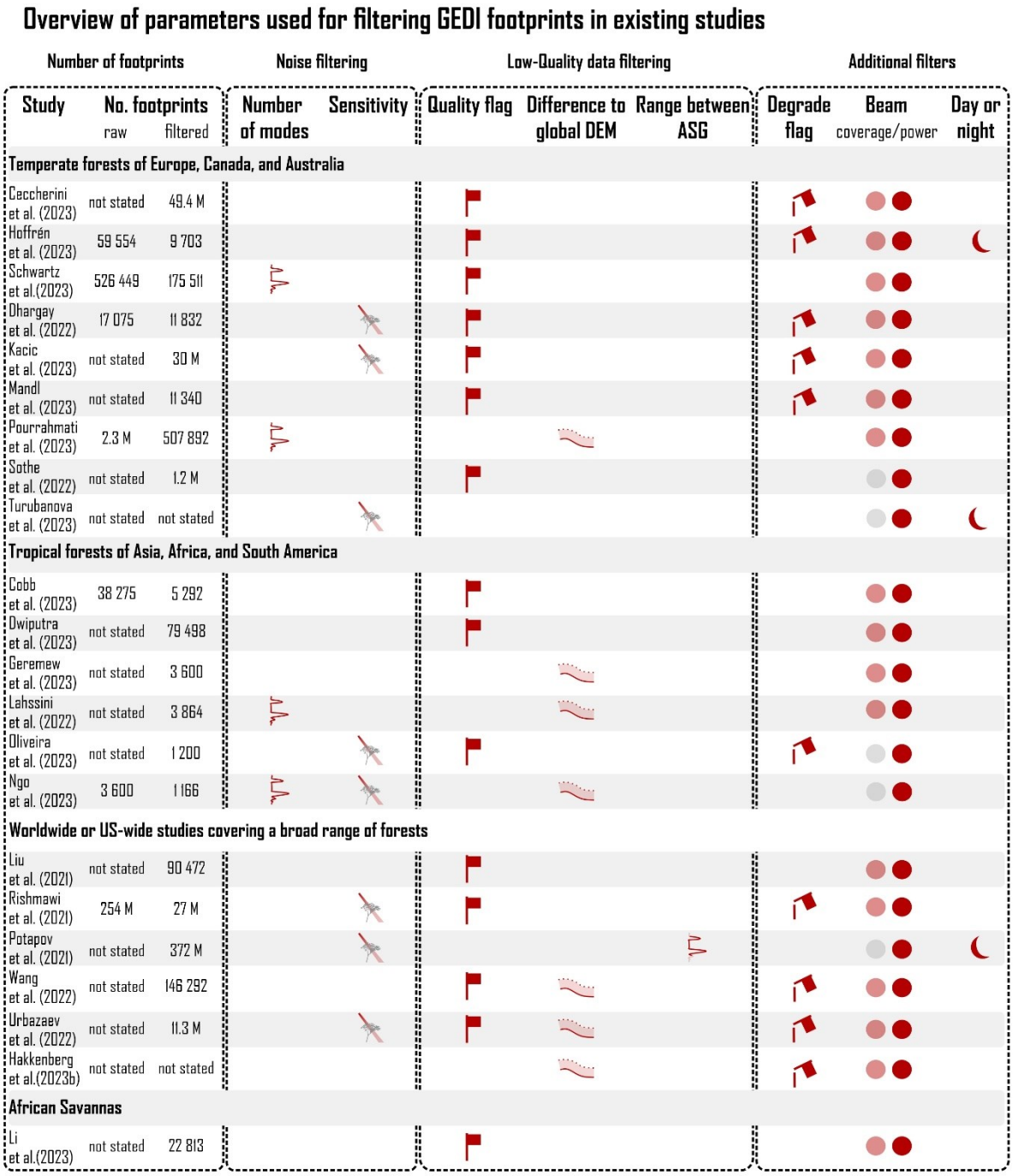


Figure 1. Overview of parameters used for filtering GEDI footprints in existing studies. ASG: algorithm setting groups; DEM: digital elevation model. M: millions.

2. Data and Methods

2.1. Study areas

We selected three study areas in the temperate biome (Olson et al., 2001) in Spain, New Zealand, and California (USA; Figure 2) based on the availability of airborne laser scanning (ALS) data from the period

overlapping with the GEDI data acquisition. We used discrete return ALS data as a reference to analyze errors in terrain elevation and maximum vegetation height of GEDI footprints.

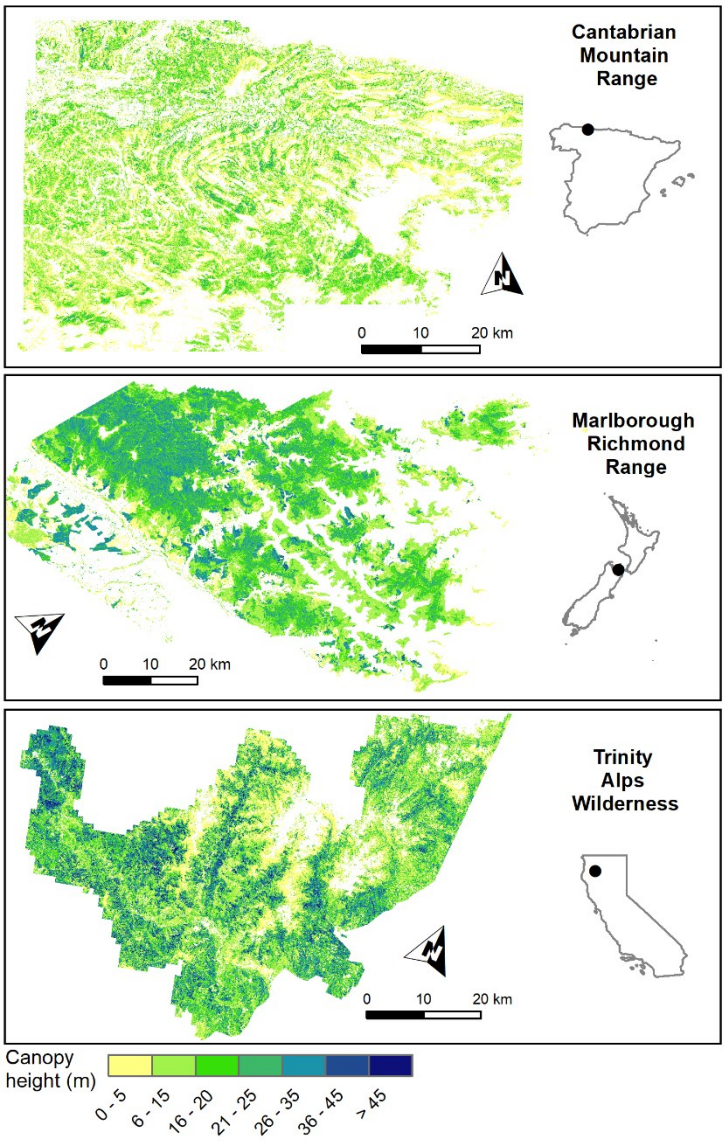


Figure 2. Location of the study areas and distribution of forest canopy height therein. Canopy height retrieved from ALS data.

Cantabrian Mountain Range (Spain)

The Cantabrian mixed forests are a transitional zone between the Eurosiberian and the Mediterranean regions of Europe. The study area comprises mainly the Principality of Asturias and small parts of Cantabria and Castilla y León (Spain), including the highest areas of the Cantabrian mountains range at the Picos de Europa National Park with elevations ranging from 0 to 2,648 m (Figure 2). It covers an area of almost 400,000 ha. The average annual temperature is 8-14°C, and average rainfall is 900-1,800 mm. The Cantabrian Range is rich in floral diversity and includes a wide range of forest types. The

forests are relatively low, with a typical height of 20 m or less (Figure 2). The lowlands are characterized by broadleaf deciduous forests with English oak (*Quercus robur*), European ash (*Fraxinus excelsior*), sweet chestnut (*Castanea sativa*), and lindens (*Tilia platyphyllos, Tilia cordata*). Forests at higher elevations are characterized by deciduous oaks (*Quercus pyrenaica, Quercus petraea*) with European beech (*Fagus sylvatica*). Subalpine plant communities, such as low shrubs and grasses, are found above the tree line.

Marlborough Richmond Range (New Zealand)

The study area comprises mainly the northern part of the Marlborough region (South Island of New Zealand), including the Mount Richmond Forest Park and the valley of the Wairau river with elevations ranging from 0 to 1,760 m (Figure 2). The average annual temperature is 7-17°C, and the average rainfall is around 650 mm distributed more or less evenly across the year. Mount Richmond Forest park, located north of Wairau river, was established in 1977 and covers an area of almost 166,000 ha. It is a mountainous landscape that consists of relatively unmodified vegetation and pastures. The height of the forests differs between native manuka (*Leptospermum scoparium*) and kanuka (*Leptospermum ericoides*) trees. Manuka can grow up to 10 m in height, whereas kanuka trees reach heights of up to 25 m (Figure 2). The higher stands, up to 45 m in height, include native beech and plantation forests (*Nothofagus spp.*). South of the Wairau river, grassland is the main vegetation cover.

Trinity Alps Wilderness (California)

The Trinity Alps Wilderness is the second largest wilderness area in California, with over 200,000 ha of land (Figure 2). The Trinity Alps are a subrange of the Klamath Mountains and are characterized by rugged subalpine topography with elevations ranging from 600 m to 2,750 m. Rainfall varies between 740 mm and 2,720 mm of precipitation annually. This area is a transitional zone between the Mediterranean climate of the south and the Northwestern coastal climate, resulting in high plant diversity (e.g., with the second highest number of conifer species worldwide). The typical height of the forest is between 20 m and 35 m, but in relatively large parts of the study area, the forest height exceeds 45 m (Figure 2). Trinity Alps are covered by mixed conifer forests (*Pinus ponderosa, Pinus contorta, Pinus lambertiana, Pseudotsuga menziesii, Abies concolor, Calocedrus decurrens*), red fir forests (*Abies magnifica, Pinus jeffreyi, Pinus monticola*), and subalpine forests (*Pinus balfouriana, Pinus albicaulis, Tsuga mertensiana*) (Ferlatte, 1974).

205 **2.2. Reference data**

206 The airborne lidar data for the study areas were collected between 2018 and 2021, with the point
207 cloud densities between 1 and 9 points per m^2 (Table 1). The ALS data for the Cantabrian Mountain
208 Range were collected as a part of the PNOA (Plan Nacional de Ortofotografía Aérea) project (2nd
209 coverage flights). The project aims to capture three-dimensional information using airborne lidar
210 sensors of the entire territory of Spain (https://pnoa.ign.es/). The reported planimetric and altimetric
211 accuracies in terms of RMSE are 30 cm and 15 cm, respectively. The ALS data for the Marlborough
212 Richmond Range were collected as a part of an acquisition campaign to obtain the New Zealand
213 National Elevation Model. The point cloud vertical accuracy at the 95% confidence level required for
214 this campaign is 20 cm in non-vegetated terrain (i.e. 95% of observations have vertical error lower than
215 20 cm, National Elevation Programme of New Zealand). The ALS data for the Trinity Alps Wilderness
216 meet the requirements of the 3D Elevation Program (3DEP), which is designed to collect high-quality
217 ALS data for the United States. The absolute vertical accuracies at the 95% confidence level required
218 for 3DEP are 20 cm and 30 cm in non-vegetated and vegetated terrain, respectively (Stoker & Miller,
219 2022).

220 The already classified ALS point clouds were used to generate digital terrain models (DTM) and canopy
221 height models (CHM) at a 2 m spatial resolution. We used points classified as ground to generate DTMs
222 and the highest returns of normalized point clouds in individual cells to generate CHMs (LAStools
223 version 230123). Noise points or points above a height threshold (60 m in Cantabria, and Marlborough,
224 and 80 m in Trinity) were considered noise and not used for CHM generation. We used the DTMs as a
225 reference for GEDI terrain validation and the CHMs as a reference for the validation of the GEDI top
226 canopy height, in line with prior studies (e.g. Liu et al., 2021; Quirós et al., 2021; Zhu et al., 2022). We
227 used landcover derived from the ESA world cover product at a 10 m resolution for the year 2020
228 (https://esa-worldcover.org/en) to distinguish forests and grasslands.

Table 1 Lidar data characteristics

Study area	Area (km^2)	Year	Horizontal datum / Projection	Vertical datum	Point density
Cantabria	4,200	2018 - 2021	ETRS 1989 UTM Zone 30N	RE-50	1-2 p/m^2
Marlborough	4,900	2020 - 2021	NZGD 2000 Transverse Mercator	NZVD2016	9 p/m^2
Trinity	2,400	2019 - 2020	NAD 1983 (2011) UTM zone 10N	NAVD 88	8 p/m^2

229

230 **2.3. GEDI instrument and data**

231 The GEDI lidar, multibeam laser altimeter was composed of three full-waveform Nd:YAG lasers, each
232 with a 15.6 ns pulse and 242 Hz pulse repetition rate in the near-infrared region (1,064 nm). Two of

these were used at full power (15 mJ/pulse), while the third one was split into two lower-energy coverage beams (4.5 mJ/pulse). The laser pulses were subjected to dithering, which resulted in a total of eight tracks separated by 600 m. Each pulse had a 25 m footprint, and along the track, the footprints were separated by 60 m (Dubayah et al., 2020).

The geolocated waveforms of received energy (i.e. number of photons) as a function of time (GEDI L1B product; Luthcke et al., 2019) are the fundamental observations made by the GEDI instrument. These are further processed to provide products at the footprint level: the L2A (ground elevation, canopy height, and relative height metrics) and L2B (canopy cover, plant area index) data products. This information is subsequently used to produce gridded datasets; the L3 (mean canopy height map and standard deviation of canopy height map), and L4B (gridded above-ground biomass) datasets, at a 1 km^2 resolution; the best detail that has ever been produced at the near-global extent.

We used version 2 of the GEDI L2A Elevation and Height Metrics product that provides ground elevation and canopy height metrics of individual footprints. The returned waveform is typically multimodal with the lowest mode (*elev_lowestmode*) representing the ground elevation. The accuracy of ground elevation is crucial for the estimation of canopy height metrics, as the ground serves as an elevation baseline for the canopy height estimation (Hofton & Blair, 2019). The GEDI L2A product provides the relative height (RH) metrics representing the height at which a particular quantile of energy (i.e., 1^{st}–100^{th} quantile) was returned, relative to the elevation of the lowest waveform mode representing the ground (Hofton & Blair, 2019). Here, we use RH98 to represent top canopy height as, for example, Milenković et al. (2022) and Li et al. (2023) regarded this as a robust metric less sensitive to noise than RH100. However, others also used RH95 (Zhu et al., 2022) or RH100 (Adam et al., 2020; Quiros et al., 2021).

2.4. Target and acquisition characteristics usable for GEDI processing

The L2A product includes several parameters that describe the target and acquisition characteristics and inform users about the footprint data usability, such as the number of modes (*num_detectedmodes*), sensitivity (*sensitivity*), quality flag (*quality_flag*), or degrade flag (*degrade_flag*). We integrated these parameters with other data sources (i.e. TanDEM-X DEM, ALS DTM, ESA world cover) outlining the target characteristics (i.e. terrain height, landcover, slope) to identify the causes of bias in terrain and canopy height observations as well as the usability of these characteristics for data filtering. We classified the data processing into four steps (Figure 3): (i) filtering based on parameters identifying noise observations (number of modes, sensitivity); (ii) filtering used to remove observations with low-quality (quality flag, difference to TanDEM-X DEM, low difference among results of the six algorithm setting groups); (iii) filtering based on additional parameters

potentially affecting observation accuracy (e.g. the degrade flag, laser pulse energy, time of acquisition, and slope); and (iv) mitigating geolocation error using a high-resolution DTM.

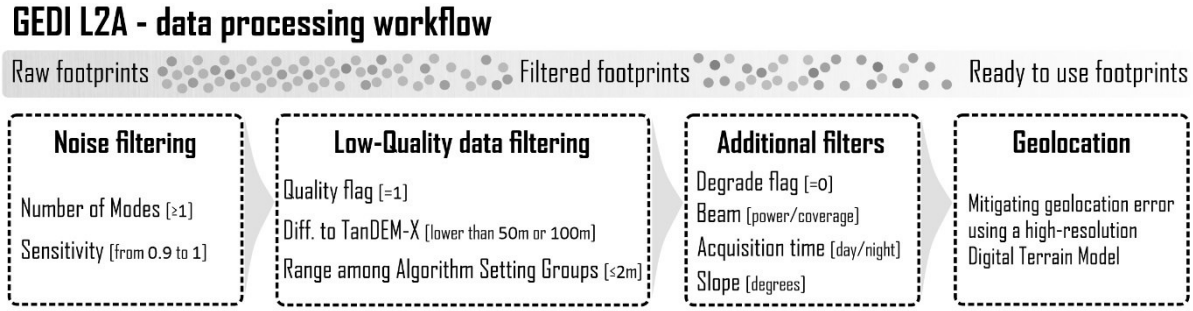


Figure 3. Processing workflow. All footprints enter the first step (*Noise filtering*), in which GEDI observations containing no useful information are identified. Here, we compared the performance of both criteria (i.e., the Number of modes and Sensitivity) for noise filtering. Subsequently, all observations that failed in at least one of the criteria were removed. The remaining observations progressed into Step 2. In Step 2 (*Low-quality data filtering*), three approaches for low-quality data filtering were employed, compared, and observations identified as low-quality were removed. The remaining observations were used in Step 3, assessing whether acquisition parameters (*Additional filters*) were needed for additional filtering or if all low-quality data were already filtered in the prior steps. Finally, the remaining observations were used in Step 4, assessing and mitigating the geolocation error of observations.

2.4.1. Noise filtering

Number of modes

The number of modes is an important parameter used for noise removal. The returned waveform can have a simple shape with only a single mode (similar to the transmitted output pulse), which is typical of bare ground, or it can be multimodal, which is typical for vegetation or rough terrain. Each mode represents a distinct reflecting surface within the laser footprint. The first mode detected above the noise is associated with the top of the canopy, the last one with the terrain (Beck et al., 2021). Waveforms without detected modes (i.e., waveforms consisting only of noise) correspond to poor observations and, as such, are typically removed from analyses (Figure 1).

Beam Sensitivity

The beam sensitivity is a signal detection performance metric that represents the maximum canopy cover that the GEDI can penetrate, correctly detecting the ground below (typically, with an estimated 90% probability of correct identification and a 5% chance of a false positive result; Hancock et al.,

2019). Beam sensitivity depends on multiple parameters (receiver efficiency, orbital altitude, telescope area, surface reflectivity, atmospheric transmission, background solar illumination, and laser pulse energy) and can be calculated using the lidar equation (Dubayah et al., 2020; Hancock et al., 2021).

To select high-quality data, the GEDI user guide (Beck et al., 2021), among other characteristics, recommends using a sensitivity threshold of 0.9 over land; under certain conditions, it has been, however, suggested that using a higher threshold can be beneficial under certain circumstances (e.g. dense forest). Therefore, only footprints with sensitivities ranging from 0.9 to 1 are considered valid and typically used as a selection criterion for the removal of noise observations (Figure 1). In addition, beam sensitivity is one of the criteria used in the quality flag (see below).

2.4.2. Low-quality data filtering

Quality flag (QF)

To facilitate filtering of the most useful high-quality observations, GEDI provides a quality flag (*quality_flag*). It is recommended only as general guidance by the GEDI user guide (Beck et al., 2021). However, for its simplicity and ease of use, it is the most commonly adopted criterion (Figure 1). The quality flag combines a set of conditions to indicate whether the waveform can be used for further analyses (1: valid waveform, 0: not valid waveform). The criteria for the quality flag include energy, amplitude, sensitivity, real-time surface tracking quality, and the difference from TanDEM-X DEM. The quality flag uses a beam sensitivity threshold of 0.9 as explained above (see section 2.4.1). Hereafter, we refer to the approach using the quality flag for filtering as QF.

TanDEM-X (TDX) difference

The GEDI L2A product includes the heights of two global digital elevation models (TanDEM-X and SRTM) that can be used for the assessment of the quality of GEDI observations. A large absolute difference between the GEDI ground elevation (*elev_lowestmode*) and elevation from an existing global digital elevation model (TanDEM-X is typically used for this purpose) can indicate erroneous observations. If the absolute difference exceeds a specified threshold, the observation is removed. The adopted thresholds differ among studies and range from 50 m (Urbazaev et al., 2022; Wang et al., 2022) to 75 m (Ngo et al., 2023), 100 m (Geremew et al., 2023; Lahssini et al., 2022; Zhu et al., 2022), or even 150 m, which is used for the generation of GEDI L3 Gridded Land Surface Metrics (Dubayah et al., 2021a). Here, we tested the thresholds of 50 m and 100 m. Hereafter, we refer to this filtering approach as TDX.

Algorithm setting groups (ASG)

The waveforms are processed using six possible algorithm setting groups to detect the ground and canopy height available in the L2A product. Each algorithm setting group uses a different combination of parameters to accommodate different acquisition scenarios (e.g. daytime, nighttime, low and high energy lowest modes; see Table 5 in Hofton & Blair, 2019). The most appropriate algorithm setting group for individual footprints is indicated in the GEDI (version 2) dataset, determined on the basis of the laser return energy, geographic region, and plant functional type (Beck et al., 2021).

We used all six algorithm setting groups to filter out the low-quality observations. A narrow range of ground elevations predicted by the six algorithm setting groups (i.e., the similarity of their results) may indicate high-accuracy observations and can be used for their filtering (e.g. Potapov et al., 2021). Following Potapov et al. (2021), we used the threshold of the maximum difference between all six algorithm setting groups of 2 m. Observations with higher values were considered low-quality. Hereafter, we refer to this filtering approach as ASG.

2.4.3. Additional filters

A zero value of the "degrade flag" indicates non-degraded conditions. Non-zero degrade flag values indicate that the shot was taken during a degraded period of positioning and/or pointing information (i.e. degraded altitude or trajectory; Beck et al., 2021), and such footprints are often removed (Figure 1). In addition, we also evaluated the effect of the signal strength (i.e. power/coverage beams) and time of acquisition (i.e., day/night acquisitions represented by solar elevation; positive values: day, negative values: night) on terrain and canopy height retrievals. These parameters are also often considered for GEDI data filtering (Figure 1). In addition, terrain slope is well known for its negative effect on space-borne lidar observations (e.g. Chen, 2010; Liu et al., 2021; Moudrý et al., 2022; Quiros et al., 2021). We derived rasters of slope from ALS DTMs and calculated the mean slope within a GEDI 25 m diameter footprint.

2.4.4 Assessment of geolocation error

To test for the effect of geolocation error, following Quirós et al. (2021) and Kutchartt et al. (2022) we shifted the footprints by 10 meters and 5 meters in eight different directions: 0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315°, relative to the ISS orbit direction. The direction was determined from the progression of shot numbers in the GEDI data. This resulted in 17 possible positions (the original position plus sixteen alternative positions). For each position, terrain and canopy height were extracted from DTM and CHM, respectively. The optimal footprint location was determined by selecting the position with the lowest error, defined as the smallest difference between *elev_lowestmode* and the terrain height derived from DTM.

2.5. GEDI pre-processing and footprints selection

We downloaded GEDI L2A version 2 data acquired between April 2019 and September 2022 using the Earthdata portal (accessed September 2022). The GEDI L2A elevation of the center of the lowest mode (*elev_lowestmode*) representing the terrain height is given as ellipsoidal height (above WGS84 ellipsoid). In contrast, the vertical datum of the ALS data is given as orthometric (Spain, California) or normal (New Zealand) heights, respectively. Therefore, to match the elevation heights of the lowest mode and the reference ALS data, we subtracted relevant reference surfaces from the GEDI L2A terrain height. We used the EGM08-REDNAP geoid model for Spain (https://datos-geodesia.ign.es/geoide/), Geoid12B for California (https://geodesy.noaa.gov/), and New Zealand Quasigeoid 2016 for New Zealand (https://data.linz.govt.nz/layer/53447-nz-quasigeoid-2016-raster/). In order to match the horizontal datum with ALS data, we projected the position of the footprints into the local coordinate system of each study area (Table 1). We limited our evaluation to two landcover categories (according to ESA world cover product): tree cover (hereafter forests) and grassland. This resulted in a total of 740,779; 695,023; and 645,306 footprints in Cantabria, Marlborough, and Trinity, respectively.

2.6. Assessment of GEDI observations filtering success and accuracy

We first concentrated on the accuracy of the GEDI terrain estimates. In the first step, we compared the use of two acquisition parameters (*number of modes* and *sensitivity*) for the identification of noise observations, which were then removed before the next step. The second step was to identify low-quality observations, which were subsequently removed, using three different approaches (QF, TDX, ASG). Finally, the remaining observations were analyzed with respect to additional characteristics that might affect GEDI observations, such as the degrade flag, laser pulse energy, time of acquisition, and slope.

We defined the vertical error in GEDI terrain estimates as their deviation from the reference surface derived from the respective ALS lidar DTM. The vertical error in GEDI terrain estimates was evaluated on pairwise vertical differences between the elevation lowest mode (*elev_lowestmode, i.e.,* the elevation estimated using the GEDI-indicated preferred algorithm setting group) of individual footprints and the lidar DTM. We used the mean elevation within a buffer with a 25 m diameter. The error distributions were summarized using two error metrics – specifically, the mean error (ME) and root mean square error (RMSE), represented as:

$$ME = \frac{1}{n}\sum_{i=1}^{n}(h_{GEDIi} - h_{REFi})$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(h_{GEDIi} - h_{REFi})^2}$$

where h_{GEDIi} is the i^{th} elevation from the GEDI L2A product, h_{REFi} is the corresponding "true" elevation from a lidar DTM, and n is the number of samples.

Our main goal was to study the effects of the acquisition and environmental characteristics on the quality of GEDI observations rather than to determine the absolute values of accuracy. We, therefore, focused on visualizing the distribution of vertical error using density plots and box plots because single-value metrics, such as ME and RMSE are a simplification and can be misleading. To investigate the effect of acquisition and environmental characteristics, we stratified the graphs and error metrics concerning the acquisition and environmental (i.e., landcover and slope) parameters (see Section 2.4. and 2.5.). In addition to removing inaccurate measurements, it is also important that the used approaches do not remove accurate measurements. Therefore, we selected high-quality observations, i.e., GEDI footprints with the error in terrain estimate lower than 3 m, and evaluated the success of the noise filtering and low-quality data filtering criteria (Figure 3) in maintaining high-quality observations. We based the threshold for high-quality observations on the theoretical vertical accuracy of GEDI canopy height estimates, which is approximately 2–3 m (Hakkenberg et al., 2023a).

Finally, we calculated pairwise vertical differences between the maximum canopy height (RH98) of individual GEDI footprints and canopy height models (CHM) derived from ALS and assessed the effect of error in GEDI terrain height estimates, slope, number of modes, and sensitivity on the accuracy of canopy height estimates.

3. Results

3.1. Accuracy of terrain estimates

3.1.1. Noise filtering (number of modes and sensitivity)

The *number of detected modes* increased with increasing *sensitivity* and had a profound effect on the observed terrain accuracy (Figure 4). The best terrain accuracy was recorded for observations with one to four modes. Further increase in the number of modes tended to underestimate the terrain; this effect was more pronounced in grasslands than in forests (Figure 4).

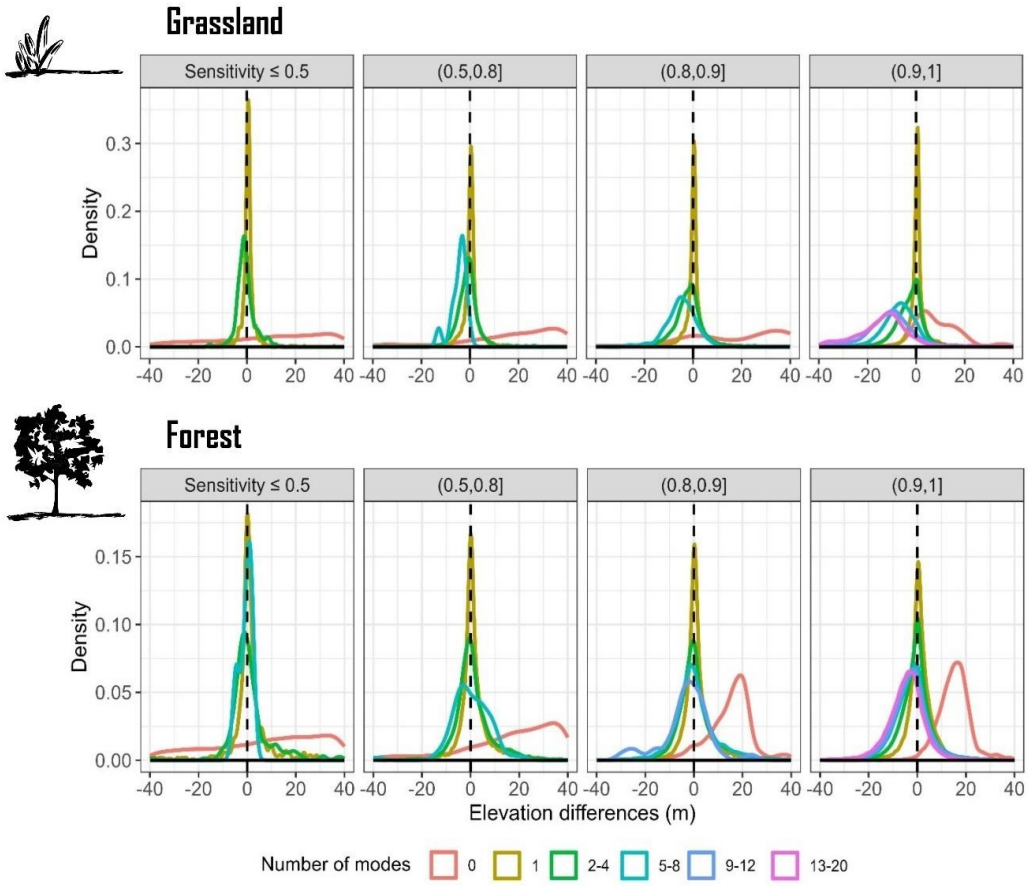


Figure 4. Density plots illustrate the distribution of elevation differences between GEDI (*elev_lowestmode*) and ALS DTM according to the number of modes classified by sensitivity and landcover type (grassland-top, forest-bottom). The number of detected modes increased with increasing sensitivity and had a profound effect on the observed terrain accuracy. Note the tendency of GEDI observations with a high number of modes in grasslands to considerably underestimate the terrain. In forests, this tendency is much lower. Observations with zero detected modes were clearly noise observations. The pattern was same for the all study areas; therefore, in this case, we merged them and the figure shows a combined result from all study areas.

In each area, the two criteria (*number of detected modes > 0* and *sensitivity between 0.9 and 1*) agreed in the identification of usable/noise observations in more than 85%. The *number of detected modes* criterion preserved more observations for the subsequent analysis than *sensitivity* (Figure 5). The sensitivity criterion tended to miss less than 1% of observations with zero detected modes. On the other hand, approximately 5% of the observations with at least one detected mode were removed due to the low sensitivity (i.e. *sensitivity* flag < 0.9). However, many such observations could be useful (Figure 6a). Observations with one detected mode in grasslands provide accurate terrain estimates. On the other hand, in forests, such observations may represent vegetation canopy and overestimate terrain height (Figure 6a).

We grouped the observations identified as noise according to landcover, time of acquisition (day/night), and beam strength to determine the characteristics most associated with the loss of signal (Table 2). The signal from the power beam was less often lost than that from the coverage beam (Table 2). Only 20% of power beam observations were identified as noise compared to approximately 30% of the coverage beam observations. The signal from the power beam showed higher sensitivity values than the coverage beam (Figure 6b). The landcover and time of acquisition (day or night) had no effect on beam sensitivity (Figure 6b) and observations removal (Table 2).

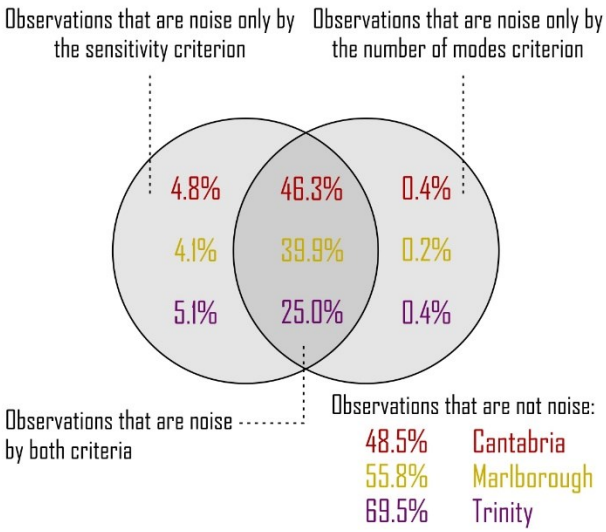


Figure 5. Venn diagram showing the percentage of observations filtered using the number of modes and sensitivity. Note that approximately 25% to 46% of GEDI footprints in our study areas are noise observations and do not provide any useful information. Between 48.5% and 69.5% of GEDI footprints passed both criteria (number of modes ≥ 1 and sensitivity ≥ 0.9).

Table 2. Percentage of GEDI observations identified as noise in forests and grasslands with respect to the time of acquisition and beam strength.

	Forest		Grassland	
	Day	Night	Day	Night
Power beam	20.0	19.7	19.6	19.6
Coverage beam	31.0	29.3	33.2	27.7

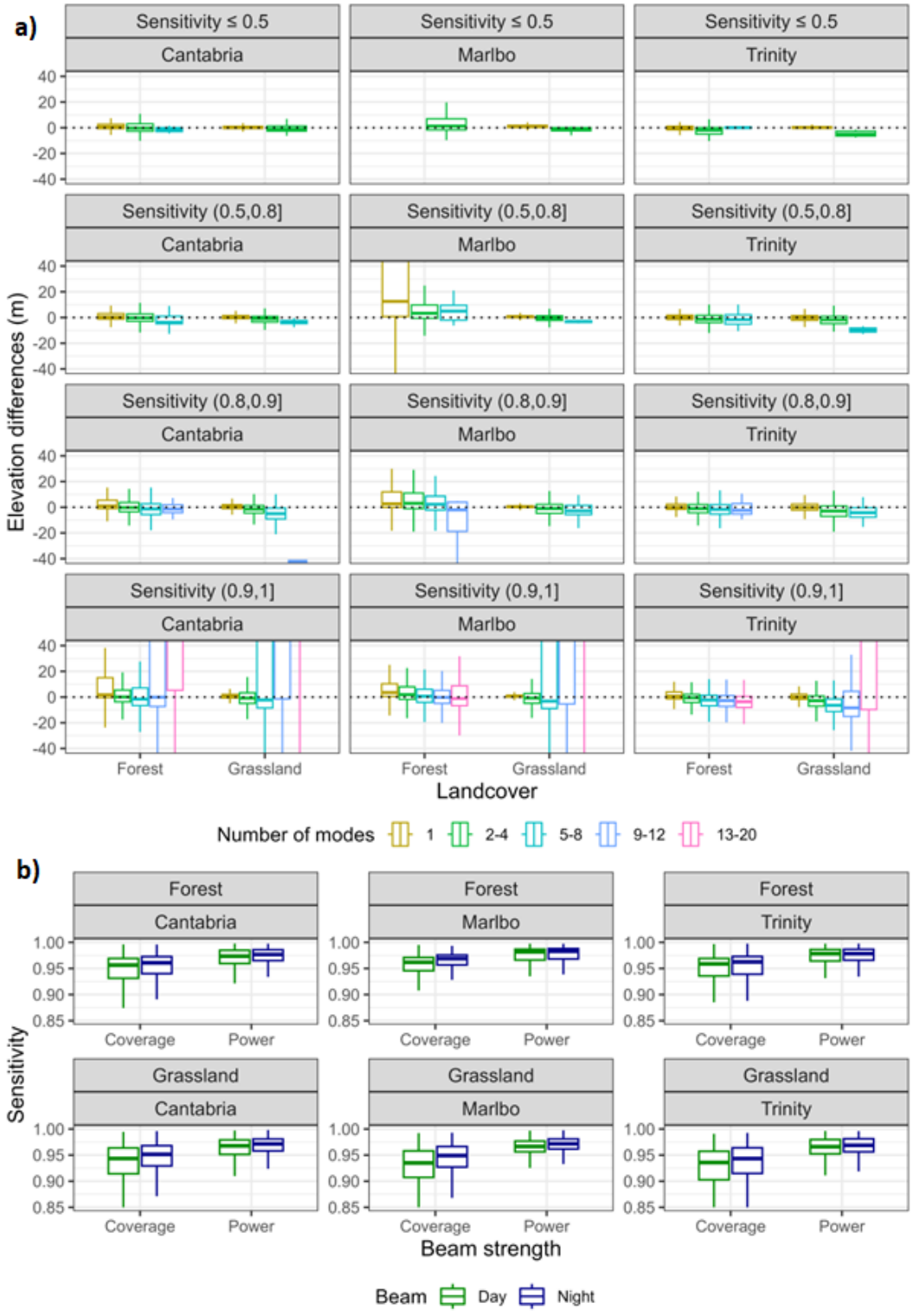


Figure 6. (a) Box plots showing the relationships between the elevation error of GEDI terrain, sensitivity, landcover (forests and grasslands), and number of modes (observations with zero detected modes were clearly noise observations, see Figure 4, and were not used in this plot). Note that observations with one detected mode, especially in grasslands, provide accurate terrain estimates even for sensitivity values lower than 0.9 (such observations are typically removed as noise). In forests, however, observations with one detected mode and low sensitivity may represent vegetation canopy

and overestimate terrain height (e.g. in Marlborough). Note that in grasslands or forests of relatively low height and density (e.g. Cantabria) a combination of the sensitivity higher than 0.9 and a high number of detected modes can lead to considerable error in terrain estimates. (b) Box plots showing the relationships between sensitivity, beam strength, landcover, and time of acquisition. Note that strong beams have higher sensitivity than coverage beams.)

3.1.2. Low-quality data filtering (QF, TDX, ASG)

In the previous section, we clearly showed that observations with zero detected modes or sensitivities between 0.1 and 0.9 were mostly noise; they were, therefore, removed and only the remaining observations (1,195,259) were used in all comparisons in the second-tier analysis presented in this section. As shown in Figure 7a, the error distributions for observations with low differences between ASGs were considerably narrower than those for measurements with high differences between ASGs. Similarly, the difference from TDX lower than 50 m and, to some degree, also the difference from TDX lower than 100 m showed a narrow distribution of error symmetric around zero, while observations with differences from TDX higher than 100 m were erroneous in two out of the three study areas (Figure 7b). The density plots in Figure 7c illustrate the effect of the quality flag. A narrow distribution symmetric around zero was observed for the quality flag equal to one, while a multimodal distribution with high positive error was evident for the quality flag equal to zero. Note, however, that the quality flag equal to 0 may also include some accurate observations (Figure 7c).

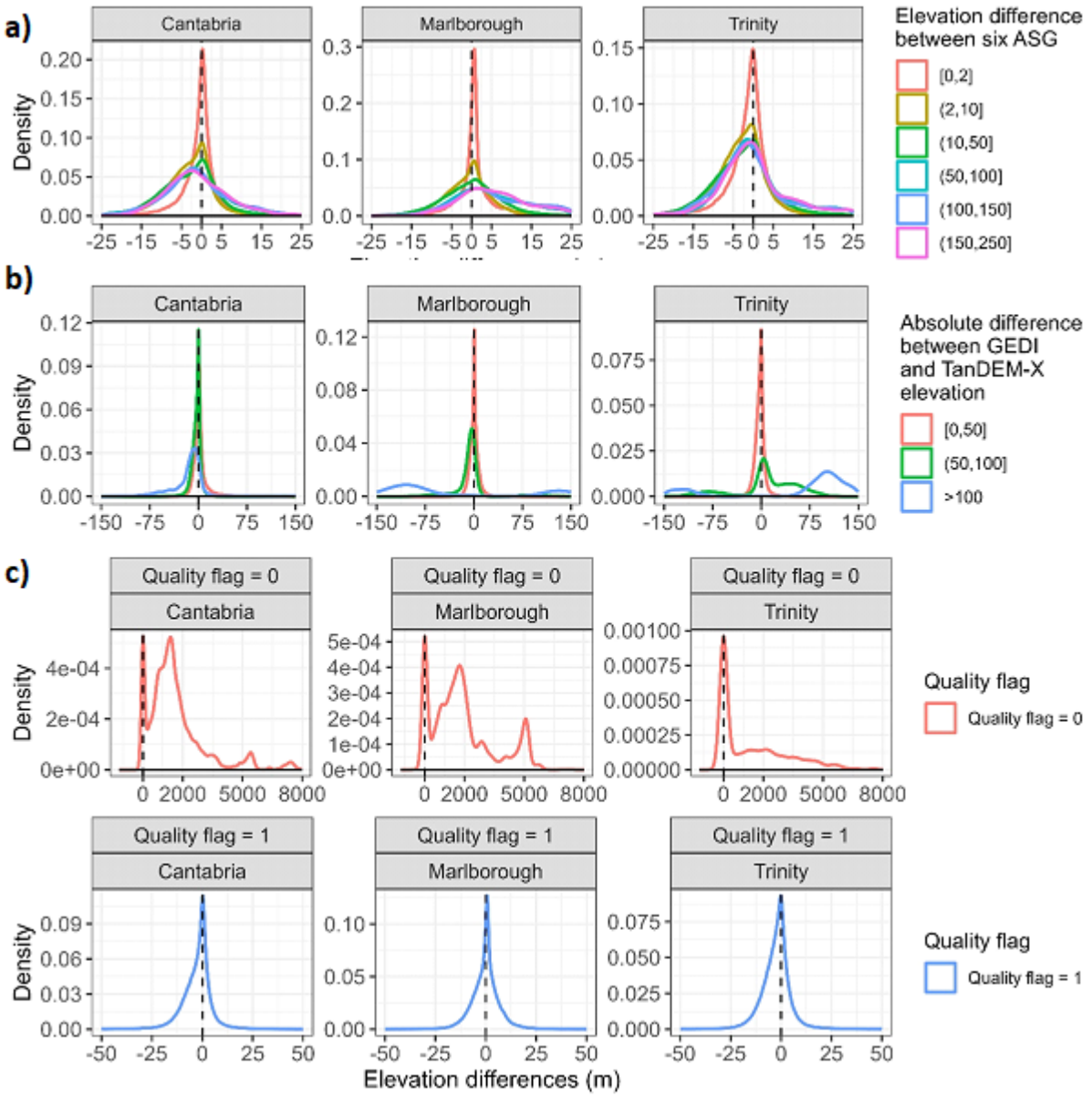


Figure 7. Density plots depicting the distribution of differences in terrain height (GEDI elevation – elevation derived from ALS DTM), in meters, according to: (a) the elevation difference between Algorithm setting groups; (b) the absolute difference between GEDI terrain elevation and TanDEM-X DEM; and (c) the quality flag (*quality_flag*).

Only 4-25% of observations (depending on the study area) were identified as low-quality by all algorithms (QF = 1, TDX < 100 m, ASG < 2 m), which indicates low agreement among them. ASG was the most restrictive approach that removed most of the observations (75-80%). The QF and TDX approaches, on the other hand, were less stringent and their combination resulted in preserving 70-90% of all non-noise observations (see Figure S1 in the Supplementary material). Combining ASG with TDX was, in terms of accuracy, the best approach (RMSEs ranging from 6.3 m to 11.1 m), followed by the combination of all methods, in which case RMSEs ranging from 7.1 m to 11.7 m were achieved

(Figure 8). Combining QF with TDX yielded slightly lower accuracy (RMSEs ranging from 11.2 m to 13.2 m); however, only 10-30% of all observations were removed (compared to 75-80 % in the ASG+TDX approach). The use of only one of the conditions, no matter which one, performed poorly and led to the selection of observations with high RMSEs (Figure 8).

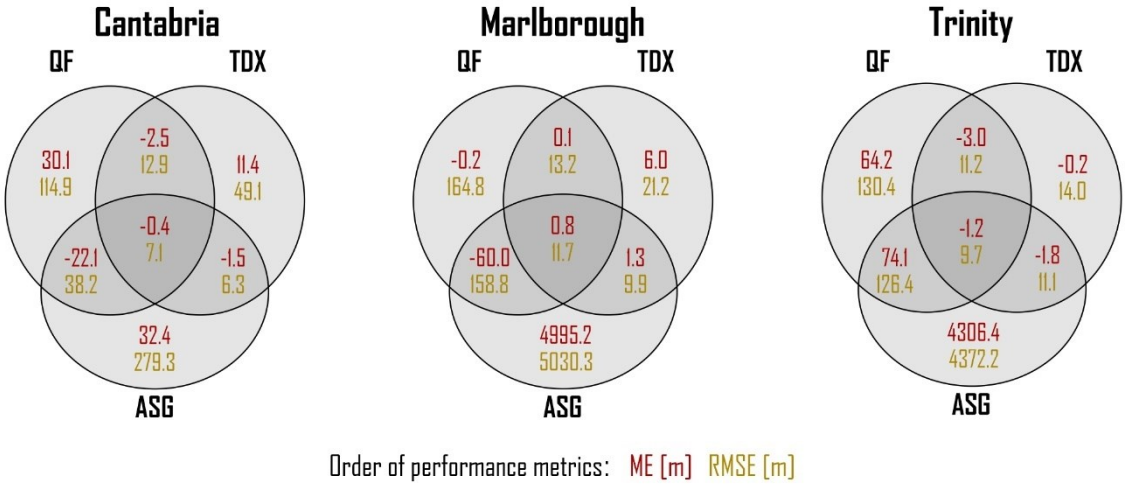


Figure 8. Venn diagrams of the performance of three approaches for filtering out low-quality data in the three study areas. The three approaches are: (QF) Quality flag equal to one; (TDX) absolute difference between the GEDI (*elev_lowestmode*) and elevation from TanDEM-X lower than 100 m; and (ASG) the range of ground elevations predicted by six algorithm setting groups is lower than 2 m. The values show performance metrics of individual or combined filtering approaches. We used the following performance metrics: Mean error (ME) and Root mean square error (RMSE). For more performance metrics see Figure S1 in the Supplementary material.

3.1.3. Success in retaining high-quality observations

We evaluated the success of the noise filtering and low-quality data filtering criteria in preserving high-quality observations (i.e., with an absolute error of terrain estimates < 3 m). Only 19.5%, 24.5%, and 30.8% of observations from the original unfiltered dataset met the high-quality criteria in Cantabria, Marlborough, and Trinity, respectively. In noise filtering, the *number of detected modes* criterion was more successful in keeping high-quality observations than *sensitivity*. The *Number of detected modes* criterion led to the removal of less than 1% of high-quality observations. In comparison, *sensitivity* removed between 9-15% of high-quality observations. In the low-quality data filtering, ASG was the most restrictive approach, removing approx. 60-65% of high-quality observations. The QF and TDX approaches, on the other hand, were less stringent and their combination resulted in the removal of less than 1% of high-quality observations (see Figure S1 in the Supplementary material).

3.1.4. Additional filters (Degrade flag, beam strength, time of acquisition, slope)

In the previous section, we showed that the combination of QF + TDX approaches led to the successful removal of low-quality observations while keeping most of the observations available for further analysis. Therefore, we removed low-quality observations using the QF and TDX approach, and only the remaining observations (991,932) were used in all comparisons in this section.

In both forests and grasslands, the accuracy of GEDI terrain estimates deteriorated with the increasing terrain slope (Figure 9a). In addition, GEDI tends to identify a higher number of modes with increasing slope (Figure 9b). Maneuvering or other ISS operations (stored in the *degrade_flag*) causing degraded positioning and/or pointing information also led to low accuracy of terrain observations. Importantly, it cannot be assumed that all problematic measurements caused by maneuvering would be removed by the low-quality filtering used in the previous section. In particular, degrade flag values 5, 50, and 85 were not removed by low-quality data filtering in our study areas (Figure 10). We did not find any effect of the time of acquisition (day or night) on GEDI terrain accuracy (Figure 10).

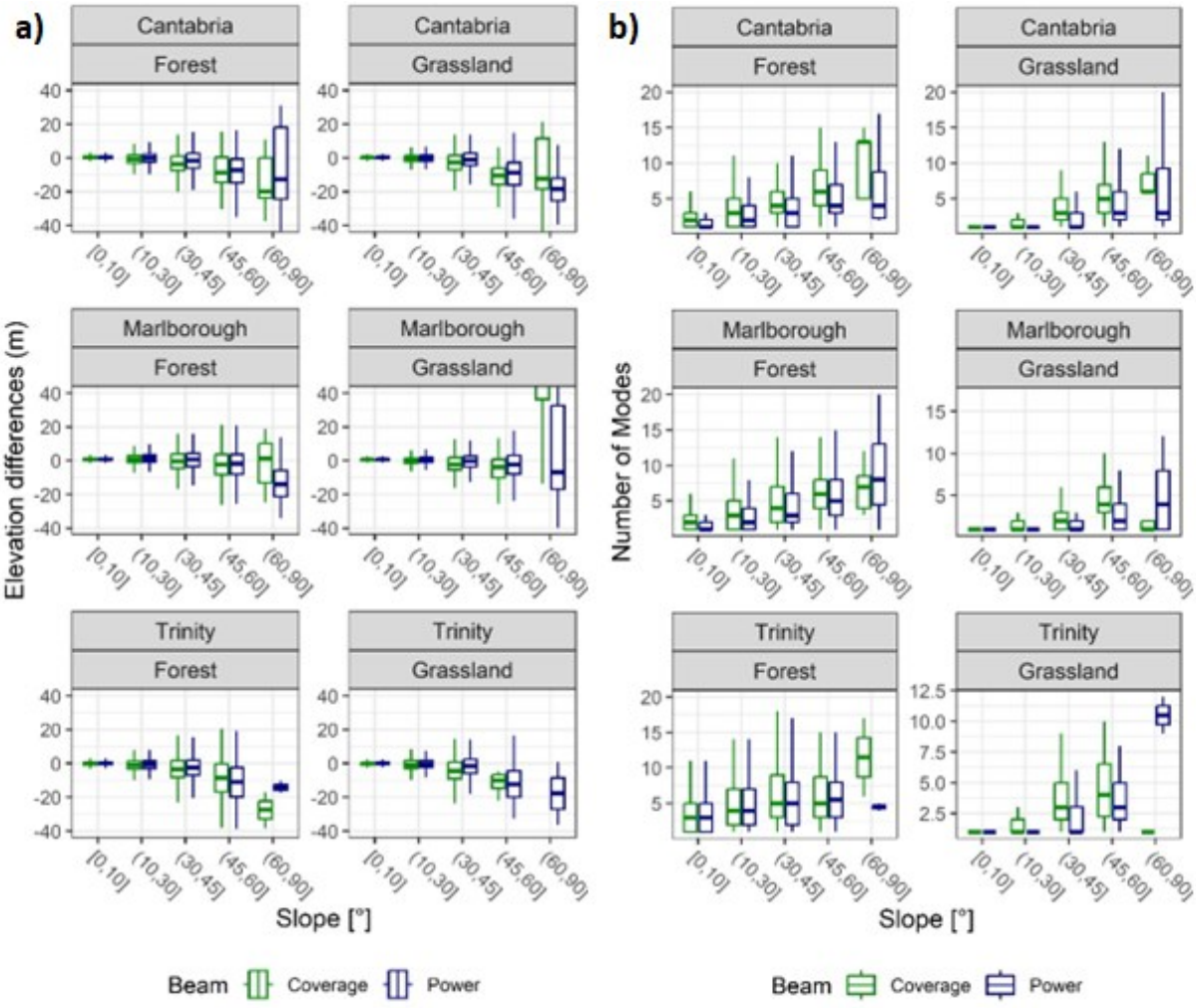


Figure 9. (a) Boxplots showing the relationship between terrain accuracy, slope, and landcover. Note that the accuracy of GEDI terrain estimates decreases with increasing slope, which can be used as an additional filter if highly accurate data are needed. (b) Boxplots show the relationship between the number of modes, terrain slope, and landcover. Note that the number of detected modes considerably increases with terrain slope in both landcover types.

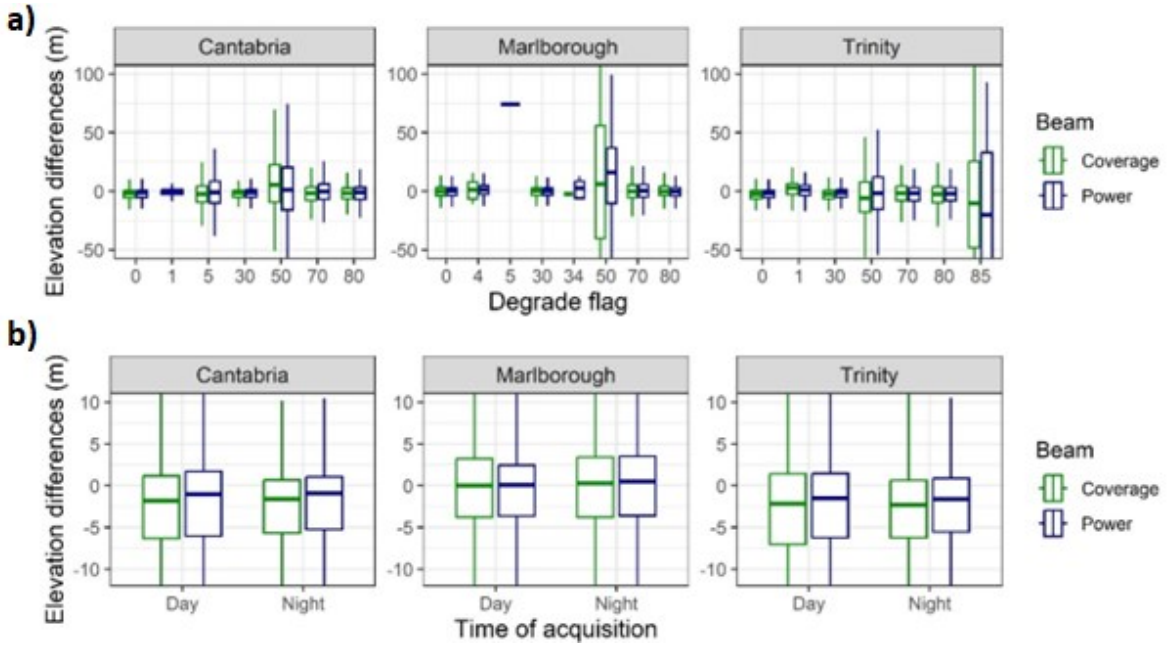


Figure 10. Boxplots showing the accuracy of GEDI terrain estimates of power and coverage beams, also considering the degrade flag (a) and time of acquisition (b). Plots are after the removal of noise and low-quality observations.

3.2. Accuracy of canopy height estimates

3.2.1. Number of modes and sensitivity

In this section, we illustrate the effect of the *number of detected modes* and *sensitivity* on the quality of canopy height estimates in forests and grasslands (we used only observations with the number of detected modes higher than zero). The observed canopy height accuracy was strongly related to the *number of detected modes*. The density plots of height differences between GEDI RH98 and lidar-derived CHM (i.e. GEDI RH98 absolute vertical error) showed different patterns with respect to the number of detected modes, landcover, and study area (Figure 11). In Cantabria, the best accuracy of forest canopy height was achieved for observations with a relatively low number of detected modes (2-4 modes); in Marlborough, the best accuracy came with a slightly higher number (9-12 modes), and in Trinity, the highest number of modes (13-20 modes) performed best. This corresponds to the canopy

height in the respective study areas – the canopy height in study areas where more modes provided better accuracy is higher (Figure 1). In grasslands, on the other hand, GEDI tended to overestimate the canopy height and the overestimation increased with the number of detected modes (Figure 11), which likely resulted from incorrect classification of noise modes below the terrain as actual terrain (Figure 4).

The *sensitivity* criterion (between 0.9 – 1.0) showed its importance for the removal of noise observations in forests where sensitivity values <0.9 led to an underestimation of canopy height (Figure 12a). In Trinity and, to some degree, also in Marlborough, canopy height was on average underestimated even for sensitivity values >0.9 (Figure 12b). In grasslands, however, observations with sensitivity values <0.9 had high accuracy and were still usable, except for Trinity (Figure 12a). Moreover, in grasslands, sensitivity values >0.9 led to an increase in error and overestimation of canopy height, particularly in Marlborough and Cantabria (Figure 12b).

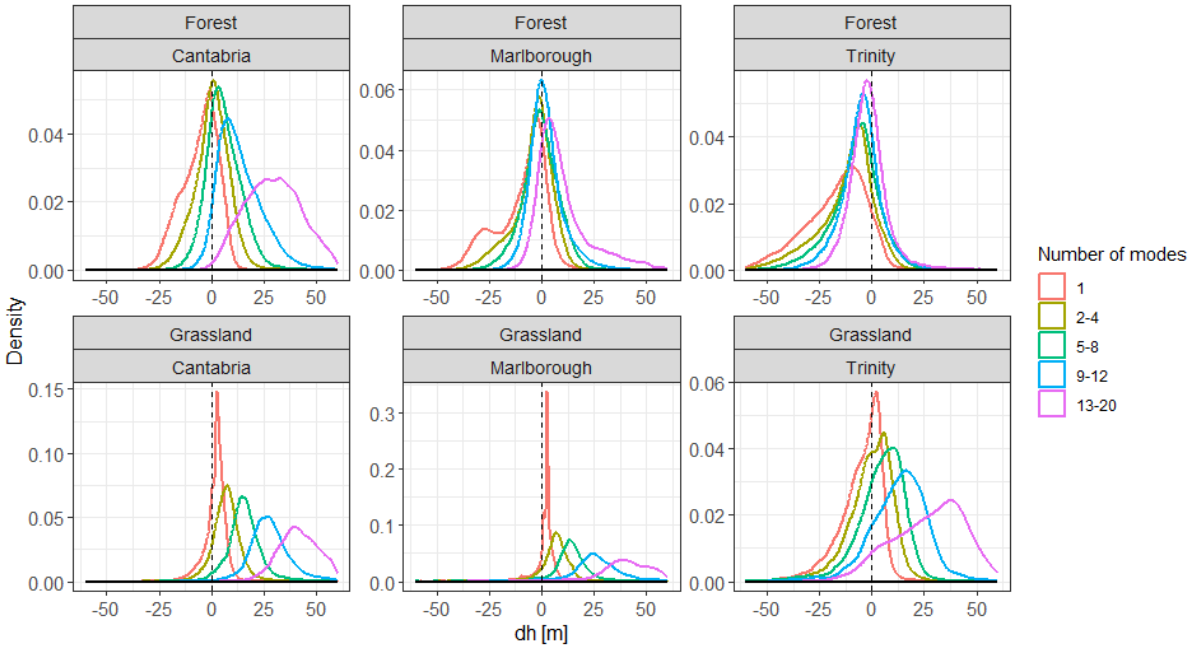


Figure 11. Density plots depicting the distribution of canopy height differences between GEDI (RH98) and ALS-derived CHM according to the number of modes grouped by landcover type. The optimal number of modes in forests differs considerably with the study area. More modes provided better accuracy in the study areas with higher canopy height (e.g. Trinity). On the other hand, in Cantabria with relatively low forest canopy height, footprints with high numbers of modes provided clearly erroneous values. Note the tendency of GEDI observations to overestimate the canopy height in grasslands with the growing number of modes.

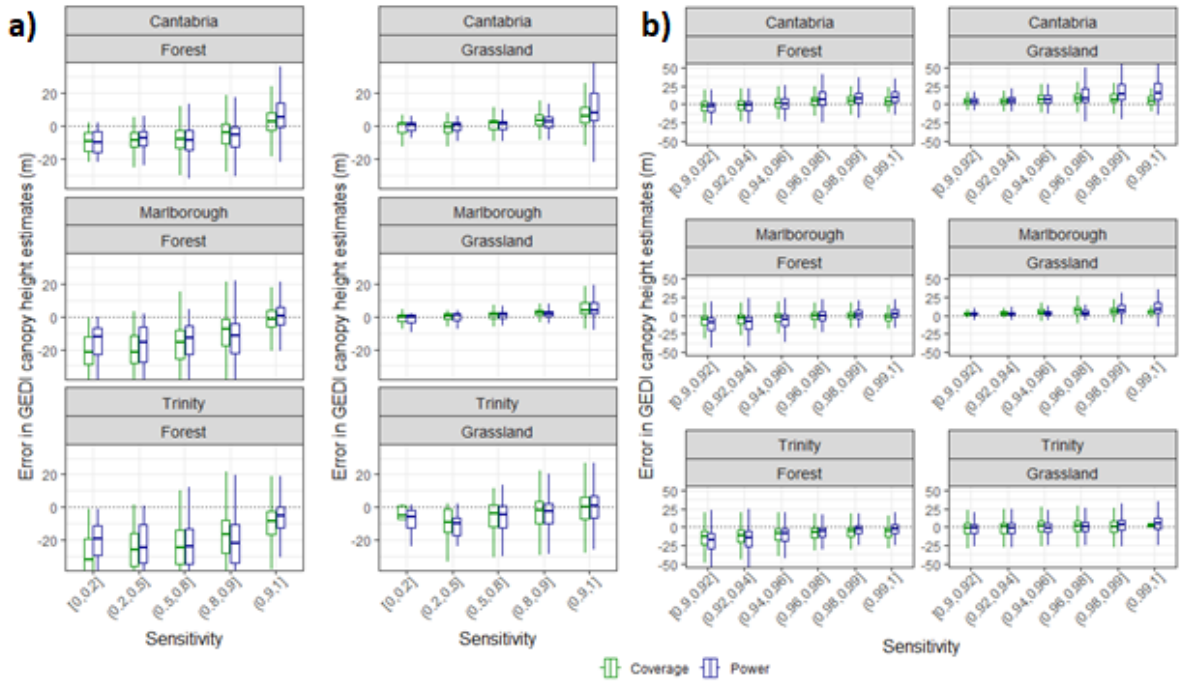


Figure 12. Box plots depicting the relationship between the error of GEDI canopy height estimates, landcover (forests and grasslands), beam strength, and sensitivity (a) from 0 to 1 and (b) only values higher than 0.9. Note the underestimation of forest canopy height for sensitivity values lower than 0.9. In dense forests (e.g. Marlborough, Trinity), sensitivity values even higher than 0.9 are needed to select high-accuracy observations. On the other hand, in grasslands, observations with high sensitivity may lead to the detection of multiple modes and, consequently, to the overestimation of canopy height. This is especially common on steep slopes (see also Figures 9 and 11).

3.2.2. Effect of the terrain on the accuracy of canopy height estimates

The accuracy of terrain estimates (see Section 3.1.) is crucial for estimates of canopy height. Therefore, we also demonstrate the role of accurate terrain observations in determining canopy height estimates. In this section, we used only observations with the *number of detected modes* higher than 0 and *sensitivity* between 0.9–1.0 (i.e. we removed noise observations). As expected, in forests, the underestimation of terrain elevation led to the overestimation of canopy height and vice versa. In grasslands, however, both overestimation and underestimation of the terrain led to the overestimation of canopy height (Figure 13a). The overestimation of canopy height increased with the increasing slope (Figure 13b). This is because GEDI tended to identify a higher number of modes with increasing slope (Figure 9b), which likely resulted in the detection of multiple modes incorrectly considered as vegetation and, therefore, resulting in an observed overestimation of canopy height (Figure 13b), particularly in grasslands (Figure 11). Observations with large terrain errors (> 50 m) were

an exception, not following these trends and causing large errors in canopy height observations (Figure 13a).

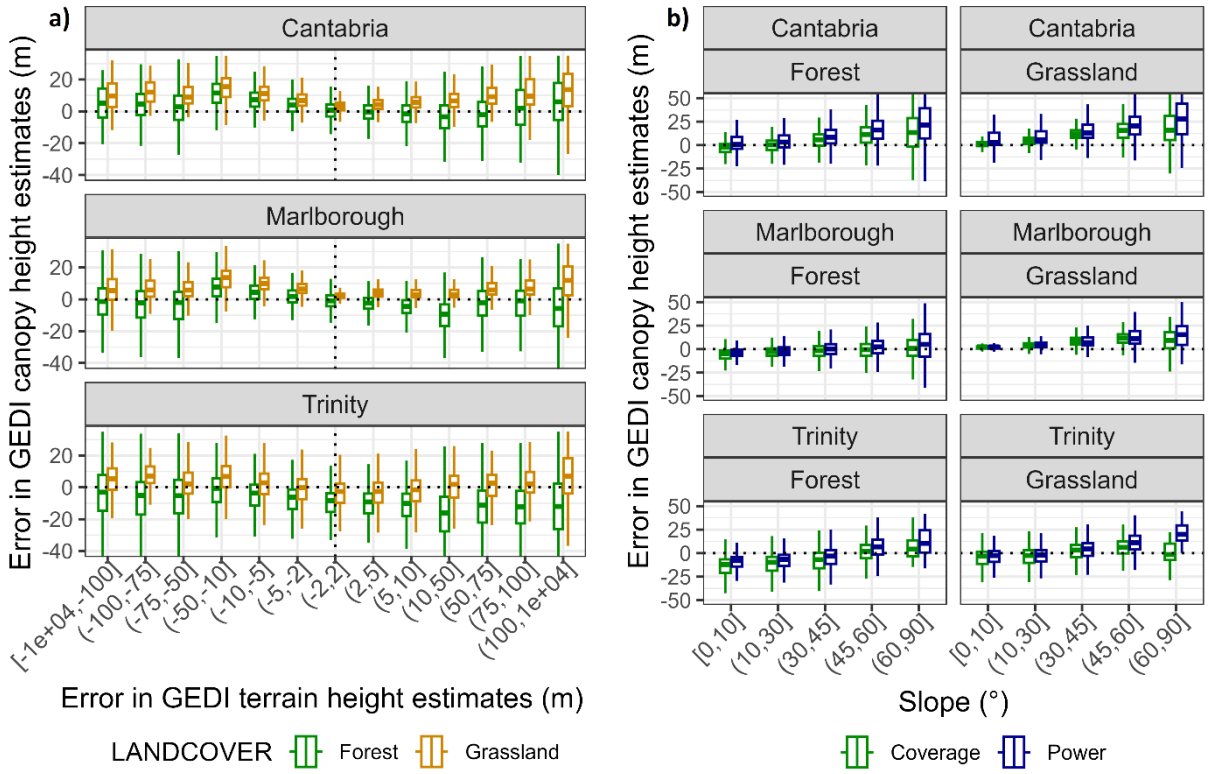


Figure 13. (a) Box plot showing the relationship between CHM accuracy and accuracy of terrain estimates. Note that in forests, an overestimation of terrain leads to the underestimation of canopy height and vice versa. (b) Box plot showing a relationship between the terrain slope and error in canopy height estimates. Note that the coverage and power beams are equally affected by the slope.

3.3. Geolocation error

We showed that maneuvering or other ISS operations causing degraded positioning and/or pointing information led to low accuracy of terrain observations. Therefore, we removed observations with a degrade flag equal to 5, 50, and 85, and only the remaining observations (963,709) were used in all comparisons in this section.

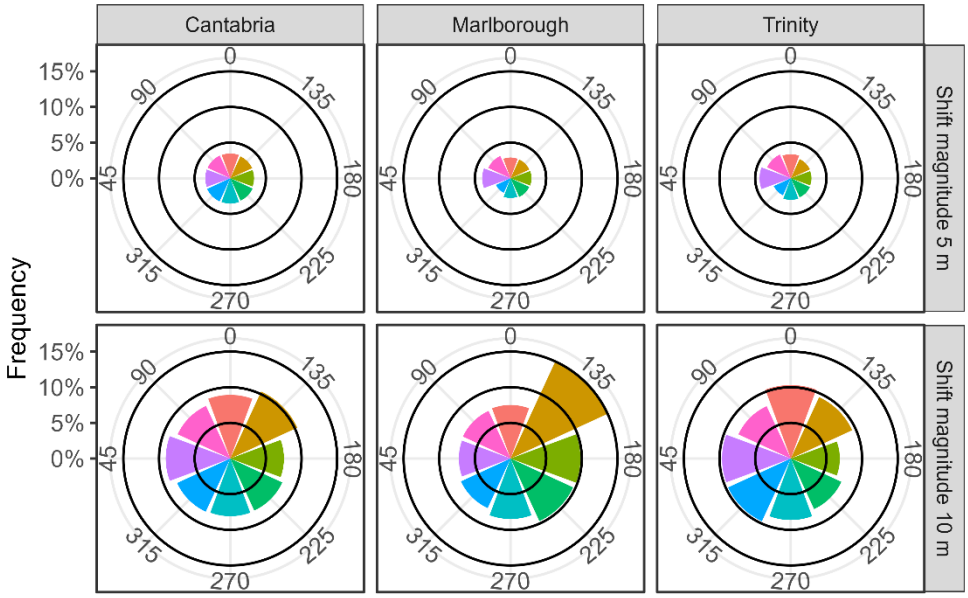
We shifted footprints in eight directions by two different distances to determine the approximate direction and magnitude of the geolocational error and its effect on the accuracy of the terrain and canopy height estimates. The shift by 10 m led to improved accuracy in 70 % of the footprints, shift by 5 m in another 26 % (no improvement after the shift was observed in the remaining 4 %, see Table 3). Neither direction was generally predominant, with the exception of Marlborough, where a 135° shift of footprints was the most common (Figure 14). The accuracy of GEDI terrain estimates improved with

609 locational shifts. However, this did not lead to a corresponding improvement in vegetation height
610 estimates (Table 3, Figure 15). In addition, the success in minimizing the geolocation error depended
611 on the number of modes and landcover (Figure 15). In forests, minimizing the geolocation error led to
612 a notable improvement in the accuracy of terrain estimates. In grasslands, on the other hand, we
613 observed much lower improvement in the accuracy of footprints with many modes, suggesting that
614 these are more severely affected by the measurement characteristics other than the geolocational
615 error.

Table 3. Frequency of shifts, terrain and canopy height accuracy measures before and after shifting.

	Frequency of shifts (%)			Terrain				Canopy			
				ME (m)		RMSE (m)		ME (m)		RMSE (m)	
	0 m	5 m	10 m	Original	Shifted	Original	Shifted	Original	Shifted	Original	Shifted
Cantabria	3.8	27.9	68.3	-2.0	-0.5	10.9	9.0	5.5	7.2	10.7	11.4
Marlborough	3.2	24.2	72.6	0.2	0.5	12.7	9.4	0.9	1.3	8.3	8.2
Trinity	3.5	25.9	70.6	-2.5	-0.9	11.3	8.3	-7.3	-6.6	13.8	13.3

616



617

618 **Figure 14.** Percentage of shifts of GEDI footprints with respect to shift magnitude and direction.

619

620

621

622

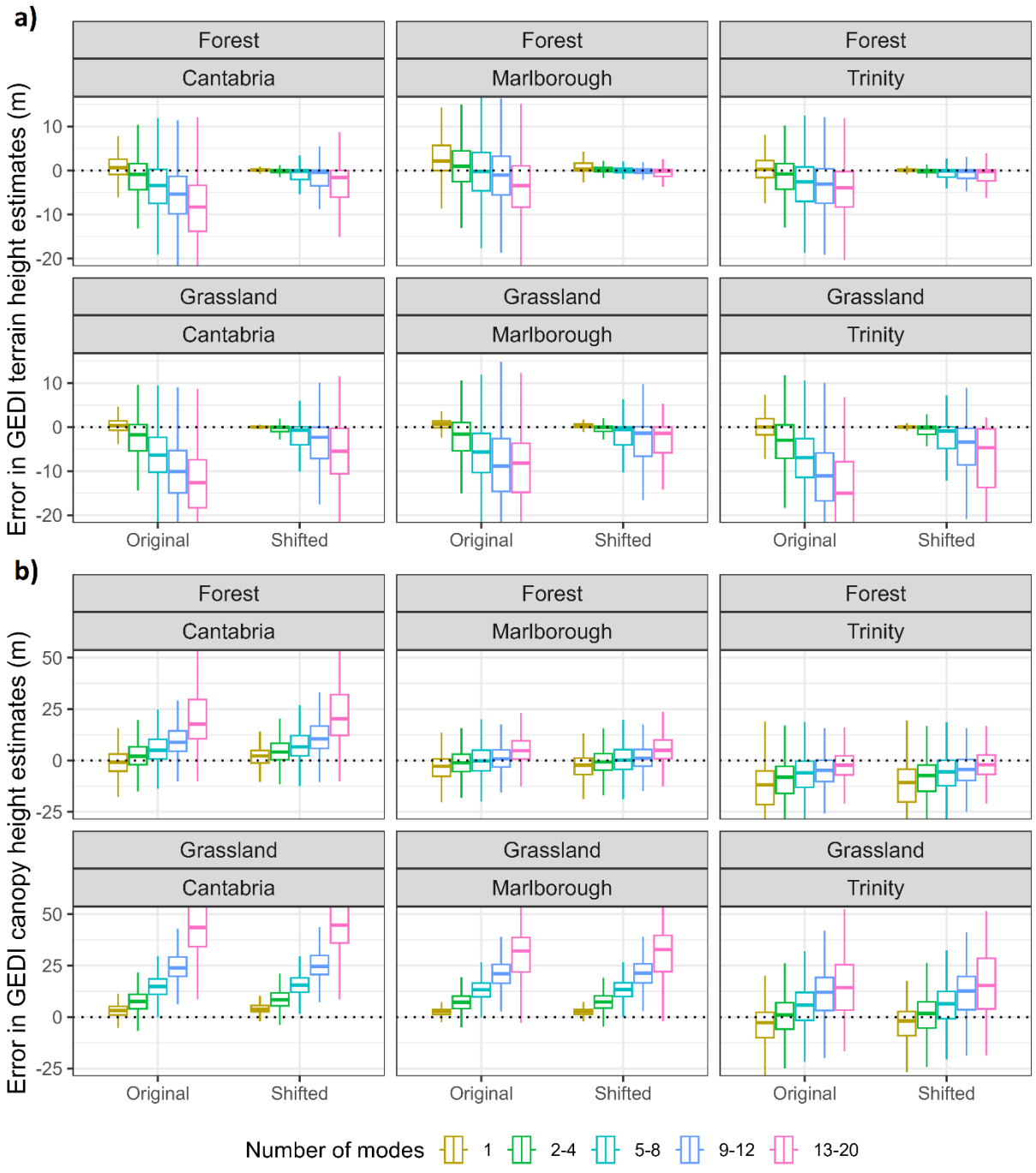


Figure 15. Box plots illustrating the height differences between GEDI and ALS reference data for terrain (a) and canopy (b) height estimates before (original) and after (shifted) correction for geolocation error, categorized by the number of modes and landcover type.

4. Discussion

In this study, we examined the effect of acquisition characteristics (such as *number of detected modes*, *sensitivity, beam strength, solar background noise, slope*) on the vertical accuracy of GEDI terrain (*elev_lowestmode*) and canopy (*RH98*) height estimates and determined their usability for identifying noise (e.g. when a beam could not reach the Earth's solid surface), and low-quality observations (e.g. when a return from canopy is identified as ground return). In this way, approximately 51.5%, 44.2%, and 30.5% of observations were identified as noise in the Cantabria, Marlborough, and Trinity regions, respectively (Figure 5). Additionally, 14.7%, 8.0%, and 6.0% of observations were categorized as low-quality in the same regions. Consequently, after the removal of noise and low-quality data, the retained data yielded percentages of 33.8%, 47.8%, and 63.5% for Cantabria, Marlborough, and Trinity, respectively.

4.1. Number of modes, sensitivity, and slope

Our results show that using a non-zero number of detected modes is a useful first filter to identify noise observations while keeping high-quality observations. In addition, the *number of detected modes* profoundly affected the observed terrain and canopy height accuracy. In forests, there was an association between the true (ALS-determined) canopy height and the optimum number of modes; for example, 9-12 modes provided the best results in Marlborough, which is characterized by tall vegetation, but only 2-4 modes in Cantabria where low forests predominate (Figure 11). In grasslands, there was a tendency to underestimate the terrain with an increasing number of modes, which, in turn, led to an overestimation of the canopy height (Figure 4; Figure 11). This problem was most evident on steep slopes. The number of modes increases with increasing slope (Figure 9b), suggesting that the waveform on steep slopes (or in a rugged terrain) takes a bimodal (or even multimodal) shape. One mode is reflected by the bottom of the slope, while the other by the top of the slope, with the latter being treated as vegetation. On slopes greater than 45°, we observed median underestimation of terrain (Figure 9) and corresponding overestimation of canopy height of up to 10 m (Figure 13b). This finding corresponds with prior studies that reported considerable effect of terrain slope and/or number of modes on GEDI terrain and canopy height estimates in both forested and non-forested landscapes (Quirós et al., 2021; Urbazaev et al., 2022; Wang et al., 2022).

Filtering using the sensitivity criterion must be applied with caution and the character of the landcover (e.g. height and density of vegetation) must be taken into account. The *sensitivity* criterion (sensitivity between 0.9 and 1) showed its importance for the removal of noise observations in forests where sensitivity values lower than 0.9 lead to an underestimation of canopy height (Figure 12a). Such a

threshold works well for forests of relatively low height and density, such as in Cantabria. However, in high and dense forests, such as Trinity and Marlborough (Figure 2), a higher sensitivity threshold, e.g. 0.95, is needed, as also recommended by the GEDI user guide (Beck et al., 2021). In addition, it was recently shown by Fayad et al. (2022) that in densely vegetated areas, beams with high sensitivity (≥ 0.98) have a higher chance of reaching the ground. On the other hand, in grasslands, a high sensitivity criterion tended to remove useful observations (Figure 6). However, observations in grasslands are equally important as those from forested areas for successful modeling of canopy height at fine resolutions (e.g. Lang et al., 2023; Potapov et al., 2021; Schwartz et al., 2023). Filtering them out using the sensitivity or quality flag criteria (Figure 1; quality flag uses a sensitivity threshold of 0.9) results in a reduced number of samples for model fitting. Approximately 5% of observations in each study area were removed this way (Figure 5). Therefore, we recommend using different sensitivity values for different landcover types to avoid unwanted removal of accurate observations. Over landcover types with low or sparse vegetation, such as grasslands, we recommend using a sensitivity higher than 0.5 (Figure 12a; which is also used over the ocean; Beck et al., 2021) but lower than 0.9. Sensitivity greater than 0.9 resulted in an overestimation of canopy height in grasslands, especially on steep slopes, where high sensitivity led to the detection of multiple modes, causing an overestimation of canopy heights. Alternatively, this issue could be addressed by excluding observations in grassland areas, on steep slopes, and those with a high number of modes (e.g. higher than 5).

4.2. Best filtering of low-quality observations (QF, ASG, TDX)

Even if observations pass the first noise filter, they may still have low accuracy, which limits their use for further analysis. Indeed, our results show that only 20-31% of observations had an absolute vertical error of terrain estimates <3 m. However, the identification of low-quality observations is challenging, and studies use different approaches for low-quality data filtering (Figure 1). Only 4-25% of observations were identified as low-quality by all algorithms, which indicates low agreement among them (Figure 8). However, none of the approaches is superior. Using only one of the approaches may lead to the selection of observations with high RMSEs (Figure 8); therefore, we recommend combining at least two approaches. ASG was the most restrictive approach, and we do not recommend using it, as it removes a large number of high-quality observations (Figure 8). On the other hand, the combination of QF and TDX approaches was less stringent, kept most of the high-quality observations, and achieved similar accuracies in terms of RMSE as when ASG was used. Therefore, the combination of these two approaches could be recommended, especially when it is important to minimize the reduction of the number of observations.

Removing observations with high absolute differences from global DEMs is a common approach (Figure 1). In our study areas, the difference between TanDEM-X DEM and GEDI terrain observations higher than 100 m indicated mostly erroneous measurements (Figure 7). The absolute vertical accuracy of global DEMs, such as TanDEM-X, depends on terrain characteristics and canopy height and, hence, considerably differs among landcover types (Gdulová et al., 2020; Hawker et al., 2019). Therefore, care must be taken when selecting the threshold, and the landcover type and forest height should be considered. Low absolute difference thresholds (e.g., 50 m) can lead to removing useful observations (Urbazaev et al., 2022). In forests, such as in our study areas, we recommend a 100 m threshold; in grasslands, it is reasonable to use a 50 m threshold. Accurate terrain models from ALS data are available in the United States, Australia, New Zealand, and many countries in Europe (e.g., Moudrý et al., 2023; Stoker & Miller, 2022). In this case, using such models in combination with a more stringent threshold is more appropriate for the selection of accurate GEDI observations.

4.3. Degrade flag, beam type, and time of acquisition

Our results show that the degrade flag is an important parameter to consider when filtering GEDI observations, as degraded observations may not be removed by noise or low-quality filters. Degrade flags 5, 50, and 85 (see Beck et al., 2021 for the explanation of these values) were particularly problematic in our study areas. On the other hand, using the degrade flag considerably lowered the number of available observations. We observed an additional 19.5% decrease in number of observations. These results correspond with the recent study by Urbazaev et al. (2022), who showed that applying the GEDI degrade flag reduced the number of GEDI observations by more than 20%. Considering individual classes of the degrade flag and removing only classes that cause large errors (e.g., by comparing to TanDEM-X DEM) might minimize losses.

The time of acquisition did not considerably affect the quality of observations in our study areas (Figure 10). Similarly, Liu et al. (2021) concluded that the accuracy of the GEDI data acquired during the day and at night is almost identical, but some studies suggested otherwise (Adam et al., 2020). The beam type did not affect the quality of retrieved terrain and canopy height observations, either (Figure 9, Figure 12), but the signal from the power beam was less often lost than that from the coverage beam (Table 2). Other studies, such as Liu et al. (2021) over multiple sites in the United States and, more recently, Rodda et al. (2023) in a tropical dry forest in India, however, typically reported that power beams were more accurate than coverage beams. Such differences might be, for example, related to forest types evaluated in individual studies. In addition, beam strength is an input for calculating sensitivity; therefore, if a sensitivity criterion is applied, coverage beams of potentially low accuracy

are removed in that step. We recommend using both power and coverage beams in temperate forests and grasslands, regardless of whether they were acquired during the day or night.

4.4 Mitigation of geolocation error

Our results show that locational shifts improved the accuracy of GEDI terrain estimates. All directions of data shifts were evenly presented, and none of them dominated (Figure 14). Although prior studies highlighted systematic shifts in locations of the first release of GEDI footprints (Quirós et al., 2021; Kutchartt et al., 2022), our results rather correspond to the more recent findings indicating that the second release of GEDI data, used in this study, achieved improved geolocation accuracy (Tang et al. 2023). However, the ability to mitigate geolocation error using high-resolution DTM decreased with the increasing number of detected modes (Figure 15), which was consistent with the observation that the same issue was reported for slope (Quirós et al., 2021) and that the number of detected modes increased with increasing slope (Figure 9b). This suggests that geolocational error should be considered only together with other characteristics, such as slope or the number of modes, the impact of which on the GEDI accuracy can be much higher than that of the geolocation error itself. Furthermore, the improvement in the accuracy of terrain height estimates did not lead to a corresponding improvement in vegetation height estimates (Table 3). Similar results were reported by other studies (Schleich et al., 2023; Tang et al., 2023). It is therefore possible that mitigating geolocation error using high-resolution DTM is insufficient and that approaches based on simulated waveforms are more appropriate (e.g., Roy et al., 2021).

4.5 Limitations & Future research

Here, we only tested filtering in temperate forests and grasslands. We do show that the filtering may need to be adjusted depending on the landcover types. Therefore, it would be important to also test filtering approaches across other biomes and landcover types. In tropical forests, for example, the optimal filtering may be different from temperate forests, as previously shown, e.g., by Fayad et al. (2022) and Lahssini et al. (2022). For example, in tropical forests, it might be reasonable to use only power beams and sensitivity higher than 0.98 (e.g., Ngo et al., 2023; Oliveira et al., 2023). Future research could focus on this knowledge gap to provide a more comprehensive overview of GEDI data filtering across all biomes (Figure 1). In addition, we considered only terrain and canopy height observations, but we did not look at the estimates of the vertical canopy structure in between that top canopy height and terrain. Future research could focus on the retrieval of such metrics as well, given that those are often used to assess forests. Here, we specifically focused on the filtering of the GEDI waveforms and we chose a value of absolute canopy height as the representative of 'high-quality

observations', we did not propagate errors of canopy height through to subsequent applications. This, and the effect of erroneous GEDI measurements on subsequent data applications (e.g. Marselis et al., 2019, 2022) and products such as Foliage height diversity index (Hirschmugl et al., 2023), Leaf area index (Wang et al., 2023), or above ground biomass (Dorado-Roda et al., 2021), could be studied in the future to further establish what a high-quality observation entails.

5. Conclusion

In this study, we evaluated the most effective ways of identifying noise and low-quality observations from GEDI, while keeping high-quality observations among the non-noise waveforms. We conclude that to identify noise observations in temperate forests and grasslands, it is best to filter by the 'number of modes' and remove all observations with less than 1 mode. The 'sensitivity' metric must be considered together with landcover. The recommended sensitivity threshold of 0.9 was sufficient in the forests of Cantabria, but higher values were required in the forests of Marlborough and Trinity. On the other hand, in areas with sparse tall vegetation, such as grasslands or low-density forests, sensitivity down to 0.5 can be considered. Importantly, sensitivity higher than 0.9 typically overestimates canopy height in grasslands. This can be avoided by removing observations with a high number of modes (e.g. higher than 5) in grasslands. We found that to filter out the low-quality observations (following the noise removal in the previous step), the combination of the quality flag (QF) and difference from TanDEM-X DEM was the most effective in terms of trade-off between removing incorrect observations and retaining as many high-quality observations as possible. Additional filters that are important to pay attention to include the degrade flag and the terrain slope. Beam type and time of acquisition (day/night) did not improve the retention of high-quality observations and are, therefore, not seen as essential in temperate forests and grasslands.

Acknowledgments

This study was funded by the Technological Grant Agency of the Czech Republic (grant no. SS02030018 DivLand) and by the Horizon Europe project EarthBridge (Grant agreement No 101079310) funded by the European Union. Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. ES was supported by the Internal Grant Agency of the Faculty of Environmental Sciences, Czech University of Life Sciences Prague (2023B0046).

Open Research

The evaluated GEDI L2A Elevation and Height Metrics Data are provided free of charge (Dubayah et al., 2021b). The original airborne laser scanning data used for the validation are sourced from the LINZ Data Service and OpenTopography, and licensed for reuse under the CC BY 4.0 license (LINZ, 2023); Spanish National Geographic Institute (SNGI, 2023); and U.S. Geological Survey (USGS, 2023).

References

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.

Beck, J., Armston, J., Hofton, M., Luthcke, S., 2021. GLOBAL Ecosystem Dynamics Investigation (GEDI) Level 02 User Guide. U.S. Geological Survey, Earth Resources Observation and Science Center, Sioux Falls, South Dakota, USA.

Bergen, K. M., Goetz, S. J., Dubayah, R. O., Henebry, G. M., Hunsaker, C. T., Imhoff, M. L., ... & Radeloff, V. C. (2009). Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. Journal of Geophysical Research: Biogeosciences, 114(G2).

Cazzolla Gatti, R., Di Paola, A., Bombelli, A., Noce, S., & Valentini, R. (2017). Exploring the relationship between canopy height and terrestrial plant diversity. Plant Ecology, 218, 899-908.

Ceccherini, G., Girardello, M., Beck, P. S., Migliavacca, M., Duveiller, G., Dubois, G., ... & Cescatti, A. (2023). Spaceborne LiDAR reveals the effectiveness of European Protected Areas in conserving forest height and vertical structure. Communications Earth & Environment, 4(1), 97.

Chen, Q. (2010). Retrieving vegetation height of forests and woodlands over mountainous areas in the Pacific Coast region using satellite laser altimetry. Remote Sensing of Environment, 114(7), 1610-1627.

Cobb, A. R., Dommain, R., Sukri, R. S., Metali, F., Bookhagen, B., Harvey, C. F., & Tang, H. (2023). Improved terrain estimation from spaceborne lidar in tropical peatlands using spatial filtering. Science of Remote Sensing, 7, 100074.

Davies, A. B., & Asner, G. P. (2014). Advances in animal ecology from 3D-LiDAR ecosystem mapping. Trends in ecology & evolution, 29(12), 681-691.

Dhargay, S., Lyell, C. S., Brown, T. P., Inbar, A., Sheridan, G. J., & Lane, P. N. (2022). Performance of GEDI space-borne lidar for quantifying structural variation in the temperate forests of south-eastern Australia. Remote Sensing, 14(15), 3615.

Dorado-Roda, I., Pascual, A., Godinho, S., Silva, C. A., Botequim, B., Rodríguez-Gonzálvez, P., ... & Guerra-Hernández, J. (2021). Assessing the accuracy of GEDI data for canopy height and aboveground biomass estimates in Mediterranean forests. Remote Sensing, 13(12), 2279.

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., ... & Silva, C. (2020). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. Science of remote sensing, 1, 100002.

Dubayah, R.O., S.B. Luthcke, T.J. Sabaka, J.B. Nicholas, S. Preaux, and M.A. Hofton. (2021a). GEDI L3 Gridded Land Surface Metrics, Version 2. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1952

Dubayah, R., Hofton, M., Blair, J., Armston, J., Tang, H., Luthcke, S. (2021b). GEDI L2A Elevation and Height Metrics Data Global Footprint Level V002 [Dataset]. NASA EOSDIS Land Processes Distributed Active Archive Center. Accessed 2024-07-23 from https://doi.org/10.5067/GEDI/GEDI02_A.002

Dubayah, R., Armston, J., Healey, S. P., Bruening, J. M., Patterson, P. L., Kellner, J. R., ... & Luthcke, S. (2022). GEDI launches a new era of biomass inference from space. Environmental Research Letters, 17(9), 095001.

Dwiputra, A., Coops, N. C., & Schwartz, N. B. (2023). GEDI waveform metrics in vegetation mapping—a case study from a heterogeneous tropical forest landscape. Environmental Research Letters, 18(1), 015007.

eoPortal 2016, ISS: GEDI (Global Ecosystems Dynamics Investigation Lidar). Available online: https://www.eoportal.org/satellite-missions/iss-gedi#laser-build-and-test-plan-overview (accessed on 29 February 2024).

Fayad, I., Baghdadi, N., & Lahssini, K. (2022). An assessment of the GEDI Lasers' capabilities in detecting canopy tops and their penetration in a densely vegetated, tropical area. Remote Sensing, 14(13), 2969.

Ferlatte, W. J. (1974). A flora or the Trinity Alps or northern Calilornia.

Fischer, R., Knapp, N., Bohn, F., Shugart, H. H., & Huth, A. (2019). The relevance of forest structure for biomass and productivity in temperate forests: New perspectives for remote sensing. Surveys in Geophysics, 40, 709-734.

Gdulová, K., Marešová, J., & Moudrý, V. (2020). Accuracy assessment of the global TanDEM-X digital elevation model in a mountain environment. Remote Sensing of Environment, 241, 111724.

Geremew, T., Gonsamo, A., Zewdie, W., & Pellikka, P. (2023). Extrapolation of canopy height and cover metrics of GEDI LiDAR in tropical montane forest ecosystem. African Geographical Review, 1-17.

Grassi, G., House, J., Dentener, F., Federici, S., den Elzen, M., & Penman, J. (2017). The key role of forests in meeting climate targets requires science for credible mitigation. Nature Climate Change, 7(3), 220-226.

Hakkenberg, C. R., Tang, H., Burns, P., & Goetz, S. J. (2023a). Canopy structure from space using GEDI lidar. Frontiers in Ecology and the Environment, 21(1), 55-56.

Hakkenberg, C. R., Atkins, J. W., Brodie, J. F., Burns, P., Cushman, S., Jantz, P., ... & Goetz, S. J. (2023b). Inferring alpha, beta, and gamma plant diversity across biomes with GEDI spaceborne lidar. Environmental Research: Ecology, 2(3), 035005.

Hancock, S., Armston, J., Hofton, M., Sun, X., Tang, H., Duncanson, L. I., ... & 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.

Hancock, S., McGrath, C., Lowe, C., Davenport, I., & Woodhouse, I. (2021). Requirements for a global lidar system: spaceborne lidar with wall-to-wall coverage. Royal Society open science, 8(12), 211166.

Hawker, L., Neal, J., & Bates, P. (2019). Accuracy assessment of the TanDEM-X 90 Digital Elevation Model for selected floodplain sites. Remote Sensing of Environment, 232, 111319.

Herold, M., Carter, S., Avitabile, V., Espejo, A. B., Jonckheere, I., Lucas, R., ... & De Sy, V. (2019). The role and need for space-based forest biomass-related measurements in environmental management and policy. Surveys in Geophysics, 40, 757-778.

Hirschmugl, M., Lippl, F., & Sobe, C. (2023). Assessing the Vertical Structure of Forests Using Airborne and Spaceborne LiDAR Data in the Austrian Alps. Remote Sensing, 15(3), 664.

Hoffrén, R., Lamelas, M. T., de la Riva, J., Domingo, D., Montealegre, A. L., García-Martín, A., & Revilla, S. (2023). Assessing GEDI-NASA system for forest fuels classification using machine learning techniques. International Journal of Applied Earth Observation and Geoinformation, 116, 103175.

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

Hoover, C. M., & Smith, J. E. (2021). Current aboveground live tree carbon stocks and annual net change in forests of conterminous United States. Carbon Balance and Management, 16(1), 1-12.

Kacic, P., Thonfeld, F., Gessner, U., & Kuenzer, C. (2023). Forest structure characterization in Germany: novel products and analysis based on GEDI, sentinel-1 and sentinel-2 data. Remote Sensing, 15(8), 1969.

Kellner, J. R., Armston, J., & Duncanson, L. (2023). Algorithm theoretical basis document for GEDI footprint aboveground biomass density. Earth and Space Science, 10(4), e2022EA002516.

Kutchartt, E., Pedron, M., & Pirotti, F. (2022). Assessment of canopy and ground height accuracy from GEDI LIDAR OVER steep mountain areas. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 3, 431-438.

Lahssini, K., Baghdadi, N., Le Maire, G., & Fayad, I. (2022). Influence of GEDI acquisition and processing parameters on canopy height estimates over tropical forests. Remote Sensing, 14(24), 6264

Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2023). A high-resolution canopy height model of the Earth. Nature Ecology & Evolution, 7(11), 1778-1789.

Li, X., Wessels, K., Armston, J., Hancock, S., Mathieu, R., Main, R., ... & Scholes, R. (2023). First validation of GEDI canopy heights in African savannas. Remote Sensing of Environment, 285, 113402.

Liang, M., González-Roglich, M., Roehrdanz, P., Tabor, K., Zvoleff, A., Leitold, V., ... & Duncanson, L. (2023). Assessing protected area's carbon stocks and ecological structure at regional-scale using GEDI lidar. Global Environmental Change, 78, 102621.

LINZ (2023). Marlborough District Council, Toitū Te Whenua Land Information New Zealand. Marlborough, New Zealand 2020-2022. Collected by Aerial Surveys, distributed by OpenTopography and LINZ [Dataset]. https://doi.org/10.5069/G97D2SB0. Accessed: 2024-07-23

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.

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

Mandl, L., Stritih, A., Seidl, R., Ginzler, C., & Senf, C. (2023). Spaceborne LiDAR for characterizing forest structure across scales in the European Alps. Remote Sensing in Ecology and Conservation, 9(5), 599-614.

Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., ... & 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.

Marselis, S. M., Tang, H., Armston, J., Abernethy, K., Alonso, A., Barbier, N., ... & Dubayah, R. (2019). Exploring the relation between remotely sensed vertical canopy structure and tree species diversity in Gabon. Environmental Research Letters, 14(9), 094013.

Marselis, S. M., Keil, P., Chase, J. M., & Dubayah, R. (2022). The use of GEDI canopy structure for explaining variation in tree species richness in natural forests. Environmental Research Letters, 17(4), 045003.

Milenković, M., Reiche, J., Armston, J., Neuenschwander, A., De Keersmaecker, W., Herold, M., & Verbesselt, J. (2022). Assessing Amazon rainforest regrowth with GEDI and ICESat-2 data. Science of Remote Sensing, 5, 100051.

Mo, L., Zohner, C. M., Reich, P. B., Liang, J., De Miguel, S., Nabuurs, G. J., ... & Ortiz-Malavasi, E. (2023). Integrated global assessment of the natural forest carbon potential. Nature, 624(7990), 92-101.

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.

Moudrý, V., Cord, A. F., Gábor, L., Laurin, G. V., Barták, V., Gdulová, K., ... & Wild, J. (2023). Vegetation structure derived from airborne laser scanning to assess species distribution and habitat suitability: The way forward. Diversity and Distributions, 29(1), 39-50.

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.

Narin, O. G., Abdikan, S., Gullu, M., Lindenbergh, R., Balik Sanli, F., & Yilmaz, I. (2024). Improving global digital elevation models using space-borne GEDI and ICESat-2 LiDAR altimetry data. International Journal of Digital Earth, 17(1), 2316113.

Neuenschwander, A., & Pitts, K. (2019). The ATL08 land and vegetation product for the ICESat-2 Mission. Remote sensing of environment, 221, 247-259.

Ngo, Y. N., Ho Tong Minh, D., Baghdadi, N., & Fayad, I. (2023). Tropical Forest Top Height by GEDI: From Sparse Coverage to Continuous Data. Remote Sensing, 15(4), 975.

943 Oliveira, P. V., Zhang, X., Peterson, B., & Ometto, J. P. (2023). Using simulated GEDI waveforms to
944 evaluate the effects of beam sensitivity and terrain slope on GEDI L2A relative height metrics over
945 the Brazilian Amazon Forest. Science of Remote Sensing, 7, 100083.

946 Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V., Underwood, E. C., ...
947 & Kassem, K. R. (2001). Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global
948 map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. BioScience,
949 51(11), 933-938.

950 Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., ... & Hayes, D. (2011). A large
951 and persistent carbon sink in the world's forests. Science, 333(6045), 988-993.

952 Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A., ... & Hofton,
953 M. (2021). Mapping global forest canopy height through integration of GEDI and Landsat data.
954 Remote Sensing of Environment, 253, 112165.

955 Pourrahmati, M. R., Baghdadi, N., & Fayad, I. (2023). Comparison of GEDI LiDAR data capability for
956 forest canopy height estimation over broadleaf and needleleaf forests. Remote Sensing, 15(6), 1522.

957 Pronk, M., Eleveld, M., & Ledoux, H. (2023). Assessing vertical accuracy and spatial coverage of
958 ICESat-2 and GEDI spaceborne lidar for creating global terrain models.

959 Quiros, E., Polo, M. E., & Fragoso-Campón, L. (2021). GEDI elevation accuracy assessment: a case
960 study of southwest Spain. IEEE Journal of Selected Topics in Applied Earth Observations and Remote
961 Sensing, 14, 5285-5299.

962 Rishmawi, K., Huang, C., & Zhan, X. (2021). Monitoring key forest structure attributes across the
963 conterminous United States by integrating GEDI LiDAR measurements and VIIRS data. Remote
964 Sensing, 13(3), 442.

965 Rodda, S. R., Nidamanuri, R. R., Fararoda, R., Mayamanikandan, T., & Rajashekar, G. (2023).
966 Evaluation of Height Metrics and Above-Ground Biomass Density from GEDI and ICESat-2 Over Indian
967 Tropical Dry Forests using Airborne LiDAR Data. Journal of the Indian Society of Remote Sensing, 1-
968 16.

969 Roy, D. P., Kashongwe, H. B., & Armston, J. (2021). The impact of geolocation uncertainty on GEDI
970 tropical forest canopy height estimation and change monitoring. Science of Remote Sensing, 4,
971 100024.

Schleich, A., Durrieu, S., & 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.

Schwartz, M., Ciais, P., Ottlé, C., De Truchis, A., Vega, C., Fayad, I., ... & Wigneron, J. P. (2023). High-resolution canopy height map in the Landes forest (France) based on GEDI, Sentinel-1, and Sentinel-2 data with a deep learning approach. International Journal of Applied Earth Observation and Geoinformation, 128, 103711.

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. Journal of Geophysical Research: Biogeosciences, 116(G4).

Skidmore, A. K., Coops, N. C., Neinavaz, E., Ali, A., Schaepman, M. E., Paganini, M., ... & Wingate, V. (2021). Priority list of biodiversity metrics to observe from space. *Nature Ecology & Evolution*, *5*(7), 896-906.

Smith, A. B., Vogeler, J. C., Bjornlie, N. L., Squires, J. R., Swayze, N. C., & Holbrook, J. D. (2022). Spaceborne LiDAR and animal-environment relationships: An assessment for forest carnivores and their prey in the Greater Yellowstone Ecosystem. Forest Ecology and Management, 520, 120343.

SNGI (2023). Spanish National Geographic Institute, Plan Nacional de Orografía Aérea (PNOA-LiDAR), LIDAR 2nd Coverage (2015–2021) [Dataset]. http://centrodedescargas.cnig.es/CentroDescargas. Accessed: 2024-07-23

Sothe, C., Gonsamo, A., Lourenço, R. B., Kurz, W. A., & Snider, J. (2022). Spatially continuous mapping of forest canopy height in Canada by combining GEDI and ICESat-2 with PALSAR and Sentinel. Remote Sensing, 14(20), 5158.

Stoker, J., & Miller, B. (2022). The accuracy and consistency of 3D elevation program data: a systematic analysis. Remote Sensing, 14(4), 940.

Tang, H., Stoker, J., Luthcke, S., Armston, J., Lee, K., Blair, B., & Hofton, M. (2023). Evaluating and mitigating the impact of systematic geolocation error on canopy height measurement performance of GEDI. Remote Sensing of Environment, 291, 113571.

Turubanova, S., Potapov, P., Hansen, M. C., Li, X., Tyukavina, A., Pickens, A. H., ... & Stolle, F. (2023). Tree canopy extent and height change in Europe, 2001–2021, quantified using Landsat data archive. Remote Sensing of Environment, 298, 113797.

Torresani, M., Rocchini, D., Alberti, A., Moudrý, V., Heym, M., Thouverai, E., ... & Tomelleri, E. (2023). LiDAR GEDI derived tree canopy height heterogeneity reveals patterns of biodiversity in forest ecosystems. Ecological Informatics, 76, 102082.

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.

USGS (2023). U.S. Geological Survey, Lidar Point Cloud - USGS National Map 3DEP Downloadable Data Collection: U.S. Geological Survey [Dataset]. https://apps.nationalmap.gov/downloader/. Accessed: 2024-07-23

Wang, C., Elmore, A. J., Numata, I., Cochrane, M. A., Shaogang, L., Huang, J., ... & Li, Y. (2022). Factors affecting relative height and ground elevation estimations of GEDI among forest types across the conterminous USA. GIScience & Remote Sensing, 59(1), 975-999.

Wang, C., Jia, D., Lei, S., Numata, I., & Tian, L. (2023). Accuracy Assessment and Impact Factor Analysis of GEDI Leaf Area Index Product in Temperate Forest. Remote Sensing, 15(6), 1535.

Zhu, X., Nie, S., Wang, C., Xi, X., Lao, J., & Li, D. (2022). Consistency analysis of forest height retrievals between GEDI and ICESat-2. Remote Sensing of Environment, 281, 113244.

SUPPLEMENTARY MATERIAL

How to find accurate terrain and canopy height GEDI footprints in temperate forests and grasslands?

Authors

Vítězslav Moudrý [a,b,c], Jiří Prošek [a,c], Suzanne Marselis [d], Kateřina Gdulová [a], Giorgi Kozhoridze [a], Eliška Šárovcová [a], Michele Torresani [e], Duccio Rocchini [a,f], Jana Marešová [a], Anette Eltner [g], Xiao Liu [g], Markéta Potůčková [h], Adéla Šedová [h], Pablo Crespo-Peremarch [i], Jesús Torralba [i], Luis A. Ruiz [i], Michela Perrone [a], Olga Špatenková [a], Jan Wild [a,c]

[a]Department of Spatial Sciences, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, 16500 Praha-Suchdol, Czech Republic.

[b]Institute for Environmental Studies, Faculty of Science, Charles University, Benátská 2, 12801 Prague 2, Czech Republic.

[c]Institute of Botany of the Czech Academy of Sciences, Zámek 1, 252 43 Průhonice, Czech Republic.

[d]Institute of Environmental Sciences, Leiden University, Einsteinweg 2, 2333 CC Leiden, the Netherlands.

[e]Free University of Bolzano/Bozen, Faculty of Science and Technology, Piazza Universitá / Universitätsplatz 1, 39100, Bolzano/Bozen, Italy.

[f]BIOME Lab, Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum University of Bologna, via Irnerio 42, 40126, Bologna, Italy.

[g]Institute of Photogrammetry and Remote Sensing, Technische Universität Dresden, 01062 Dresden, Germany

[h]Charles University, Faculty of Science, Department of Applied Geoinformatics and Cartography Albertov 6, 128 43 Praha 2, Czech Republic.

[i]Geo-Environmental Cartography and Remote Sensing Group (CGAT), Universitat Politècnica de València, Camí de Vera s/n, 46022 Valencia, Spain.

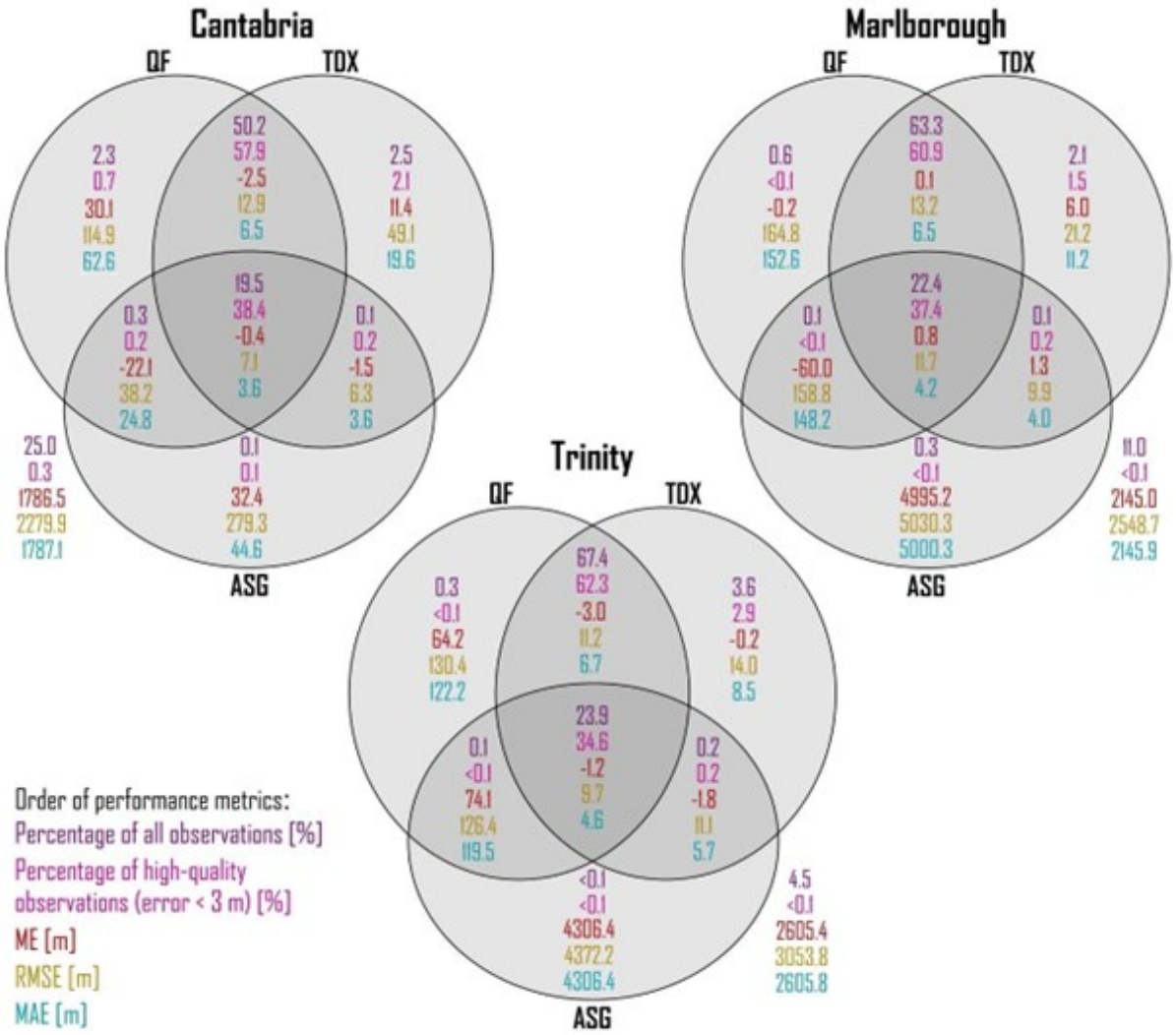


Figure S1. Venn diagrams of the performance of three approaches for filtering out low-quality data in the three study areas. The three approaches are: (QF) Quality flag equal to one; (TDX) absolute difference between the GEDI (*elev_lowestmode*) and elevation from TanDEM-X lower than 100 m; and (ASG) the range of ground elevations predicted by six algorithm setting groups is lower than 2 m. We used the following performance metrics: Percentage of all observations, Percentage of retained high-quality observations (i.e., footprints with error in terrain estimate < 3 m), mean error (ME), root mean square error (RMSE), mean absolute error (MAE). The values within the diagrams show performance metrics of individual or combined filtering approaches. In all areas, the combination of QF and TDX resulted in most preserved observations, including those of high-quality and relatively low error values. The values outside the diagrams show the performance metric of low-quality footprints removed by all three approaches. The percentage of removed observations ranged from 4.5% to 25% and had high error values (e.g. RMSE values higher than 2000 m).

Comparison of three global canopy height maps and their applicability to biodiversity modeling: Accuracy issues revealed

Vítězslav Moudrý, Lukáš Gábor, Suzanne Marselis, Petra Pracná, Vojtěch Barták, Jiří Prošek, Barbora Navrátilová, Jan Novotný, Markéta Potůčková, Kateřina Gdulová, Pablo Crespo-Peremarch, Jan Komárek, Marco Malavasi, Duccio Rocchini, Luis A. Ruiz, Jesús Torralba, Michele Torresani, Roberto Cazzolla Gatti, and Jan Wild

Accepted for publication in Ecosphere

Title: Comparison of three global canopy height maps and their applicability to biodiversity modeling: Accuracy issues revealed

Authors

Vítězslav Moudrý[1,2,3] Lukáš Gábor[1], Suzanne Marselis[4], Petra Pracná[1], Vojtěch Barták[1], Jiří Prošek[1,3], Barbora Navrátilová[5], Jan Novotný[5], Markéta Potůčková[6], Kateřina Gdulová[1], Pablo Crespo-Peremarch[7], Jan Komárek[1], Marco Malavasi[1,8], Duccio Rocchini [1,9], Luis A. Ruiz[7], Jesús Torralba[7], Michele Torresani[10], Roberto Cazzolla Gatti[9], and Jan Wild [1,3]

[1]Department of Spatial Sciences, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, 16500 Praha-Suchdol, Czech Republic.
[2]Institute for Environmental Studies, Faculty of Science, Charles University, Benátská 2, 12801 Prague 2, Czech Republic.
[3]Institute of Botany of the Czech Academy of Sciences, Zámek 1, 252 43 Průhonice, Czech Republic.
[4]Institute of Environmental Sciences, Leiden University, Einsteinweg 2, 2333 CC Leiden, the Netherlands.
[5]Global Change Research Institute of the Czech Academy of Sciences, Bělidla 4a, 60300 Brno, Czech Republic.
[6]Charles University, Faculty of Science, Department of Applied Geoinformatics and Cartography Albertov 6, 128 43 Praha 2, Czech Republic.
[7]Geo-Environmental Cartography and Remote Sensing Group (CGAT), Universitat Politècnica de València, Camí de Vera s/n, 46022 Valencia, Spain.
[8]Department of Chemistry, Physics, Mathematics and Natural Sciences, University of Sassari, Via Vienna 2, 07100 Sassari, Italy.
[9]BIOME Lab, Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum University of Bologna, via Irnerio 42, 40126, Bologna, Italy.
[10]Free University of Bolzano/Bozen, Faculty of Science and Technology, Piazza Universitá / Universitätsplatz 1, 39100, Bolzano/Bozen, Italy.

Corresponding author: Vítězslav Moudrý (moudry@fzp.czu.cz)

Open research statement:

The evaluated global canopy height maps i.e. the Global Forest Canopy Height (Potapov et al. 2021), the High-Resolution Canopy Height Model of the Earth (Lang et al. 2023), and the Global Map of Tree Canopy Height (Tolan et al. 2024) are provided free of charge, without restriction of use under Creative Commons Attribution 4.0 International License. The code used to generate virtual species and canopy height profiles is available here: https://zenodo.org/record/7332716#.Y3unguSZOUl (see also GitHub https://github.com/lukasgabor/Comparison-of-high-resolution-global-canopy-height-maps-and-their-applicability). The canopy height maps derived from airborne laser scanning are available here: https://zenodo.org/record/7697253#.ZAJswXbMJPY. The original airborne laser

scanning point clouds used for the validation are sourced from the LINZ Data Service and OpenTopography (LINZ, 2023); Federal Office of Topography swisstopo (FOTS, 2023); and U.S. Geological Survey (USGS, 2023).

Abstract

Global mapping of forest height is an extremely important task for estimating habitat quality and modeling biodiversity. Recently, three global canopy height maps have been released, the Global Forest Canopy Height Map (GFCH), the High-Resolution Canopy Height Model of the Earth (HRCH), and the Global Map of Tree Canopy Height (GMTCH). Here, we assessed their accuracy and usability for biodiversity modeling. We examined their accuracy by comparing them with the reference canopy height models derived from airborne laser scanning (ALS). Our results show considerable differences between the evaluated maps. The RMSE ranged between 10-18 m for GFCH, 9-11 m for HRCH, and 10-17 m for GMTCH respectively. GFCH and GMTCH consistently underestimated the height of all canopies regardless of their height, while HRCH tended to overestimate the height of low canopies and underestimate tall canopies. Biodiversity models using predicted global canopy height maps as input data are sufficient for estimating simple relationships between species occurrence and canopy height, but their use leads to a considerable decrease in the discrimination ability of the models and to mischaracterization of species niches where derived indices (e.g. canopy height heterogeneity) are concerned. We showed that canopy height heterogeneity is considerably underestimated in the evaluated global canopy height maps. We urge that for temperate areas rich in ALS data, activities should concentrate on harmonizing ALS canopy height maps rather than relying on modeled global products.

Keywords: Canopy height, canopy structure, habitat, heterogeneity, niche, lidar, GEDI, SDM

INTRODUCTION

Forest ecosystems cover more than 4.1 billion hectares of the Earth's surface and provide a number of services, including the regulation of the global carbon cycle and acting as reservoirs of unique species and/or high biodiversity. It is well recognized that the vertical structure of the forest, including canopy height, is directly associated with aboveground biomass (Fischer et al. 2019), primary productivity (Dănescu et al. 2016), and animal (Davies & Asner 2014) and plant diversity (Cazzolla Gatti et al. 2017). This, together with the high level of forest ecosystem threat induced by climate change and human activity, places the fine-grain global mapping of forest height among the highest-priority variables to be monitored from space (Lefsky et al. 2005; Bergen et al. 2009; Skidmore et al. 2021).

Until recently, we lacked comprehensive global data on the spatial patterns of canopy height. This changed with the rise of spaceborne laser altimetry: Light Detection And Ranging (LiDAR) missions. The Ice, Cloud, and land Elevation Satellite (ICESat) mission, operational between 2003 and 2009, collected over 250 million observations of forested areas globally (Schutz et al. 2005). In 2018, NASA launched two spaceborne LiDAR missions that have since continuously been providing data on the 3D structure of the Earth's surface. Namely, the Ice, Cloud, and land Elevation Satellite-2 (ICESat-2) and the Global Ecosystem Dynamics Investigation (GEDI), onboard the International Space Station, the latter being specifically designed for vegetation mapping (Dubayah et al. 2020; Magruder et al. 2021).

Spaceborne LiDAR sensors collect data along discrete transects. Hence, they only provide discrete samples of forest structure. Hancock et al. (2021) estimated that with the current in-orbit technology, twelve satellites would be required to produce a continuous map of canopy structure at a 30 m resolution every five years. Therefore, to produce continuous maps of global forest heights at fine resolution, the currently common approach is to train statistical or machine learning models that combine direct measurements (e.g. ICESat, GEDI, ICESat-2) with spatially continuous ancillary data (e.g. Sentinel-2, Landsat) enabling the estimation of the desired variable at locations not directly measured by GEDI or ICESat (Bergen et al. 2009; Lefsky, 2010). Simard et al. (2011) combined canopy heights collected by ICESat and multiple environmental variables (e.g., tree cover, climate, altitude) to derive a global map of canopy height at 1 km spatial resolution. That map was used e.g. to assess the association between canopy height and global water availability (Klein et al. 2015) and to test the hypothesis of a global relationship between plant species richness and canopy height (Cazzolla Gatti et al. 2017).

Biodiversity modeling is one of the important use cases for global canopy height maps as they can be used to describe habitat heterogeneity (Bergen et al. 2009; Davies & Asner, 2014; Simonson et al.

2014; LaRue et al. 2023; Fischer et al. 2024). Habitat heterogeneity has long been recognized as a key landscape characteristic determining biodiversity patterns. It is assumed that more complex environments may provide more potential niches and, thus, harbor higher species diversity (so-called habitat heterogeneity and fractal biodiversity hypotheses; Tews et al. 2004 and Cazzolla Gatti 2016). For example, taller canopies with high structural heterogeneity may harbor higher biodiversity by providing more, and more varied, resources. Indeed, canopy height has been shown to play an important role in determining the diversity and distribution of various taxa (Seavy et al. 2009; Feng et al. 2020; Cooper et al. 2020; Walter et al. 2021; Marselis et al. 2019; 2022; Tew et al. 2022; Coddington et al. 2023; Gallerani et al. 2023). The same effect was observed for horizontal heterogeneity in canopy height (Wolf et al. 2012; Fricker et al. 2021; Torresani et al. 2020). Therefore, besides the accuracy of canopy height estimates, it is also essential to know the sensitivity of global canopy height maps to canopy height changes and how well they represent canopy heterogeneity, such as edges (i.e. transitions between forest and non-forest).

According to Bergen et al. (2009), the requirements for canopy height accuracy for biodiversity and habitat science are ±2 m (but preferably ±1 m) and a resolution of 25-30 m considering the spatial scale of species' relationships with habitat characteristics (Šímová et al. 2019; Moudrý et al. 2023a). For example, Potapov et al. (2021) and Lang et al. (2023) created global canopy height maps at 30 m and 10 m spatial resolutions, respectively. More recently, Meta and World Resources Institute developed a global canopy height map at a 1 m resolution (Tolan et al. 2024). Still, despite their unprecedented spatial resolutions, modeled global canopy height maps suffer from limited accuracy related to e.g. the quality of input data and the accuracy of modeling procedures. Such errors will propagate to any subsequent analyses using these data, therefore it is important to have a good understanding of the limitations of these data products (Bolton et al. 2013; Magruder et al. 2016; Atkins et al. 2023). In addition, the reliability of predicted global maps is increasingly questioned, particularly concerning often unreliably high accuracy estimates provided by the authors, for example, because of the use of inappropriate validation strategies or non-representative reference data (Duncanson et al. 2019; Meyer & Pebesma 2022).

Therefore, in this study, we (1) evaluate the accuracy of three modeled global canopy height maps using airborne laser scanning data in Switzerland, New Zealand, and California; (2) assess the sensitivity of global canopy height maps to changes in canopy height (i.e. how well they represent canopy heterogeneity); (3) analyze the effects of environmental characteristics on the accuracy of canopy height estimates; and, finally, (4) assess the usability of the global canopy height maps for biodiversity modeling using a virtual species approach (See Figure 1 for the overall workflow of the study).

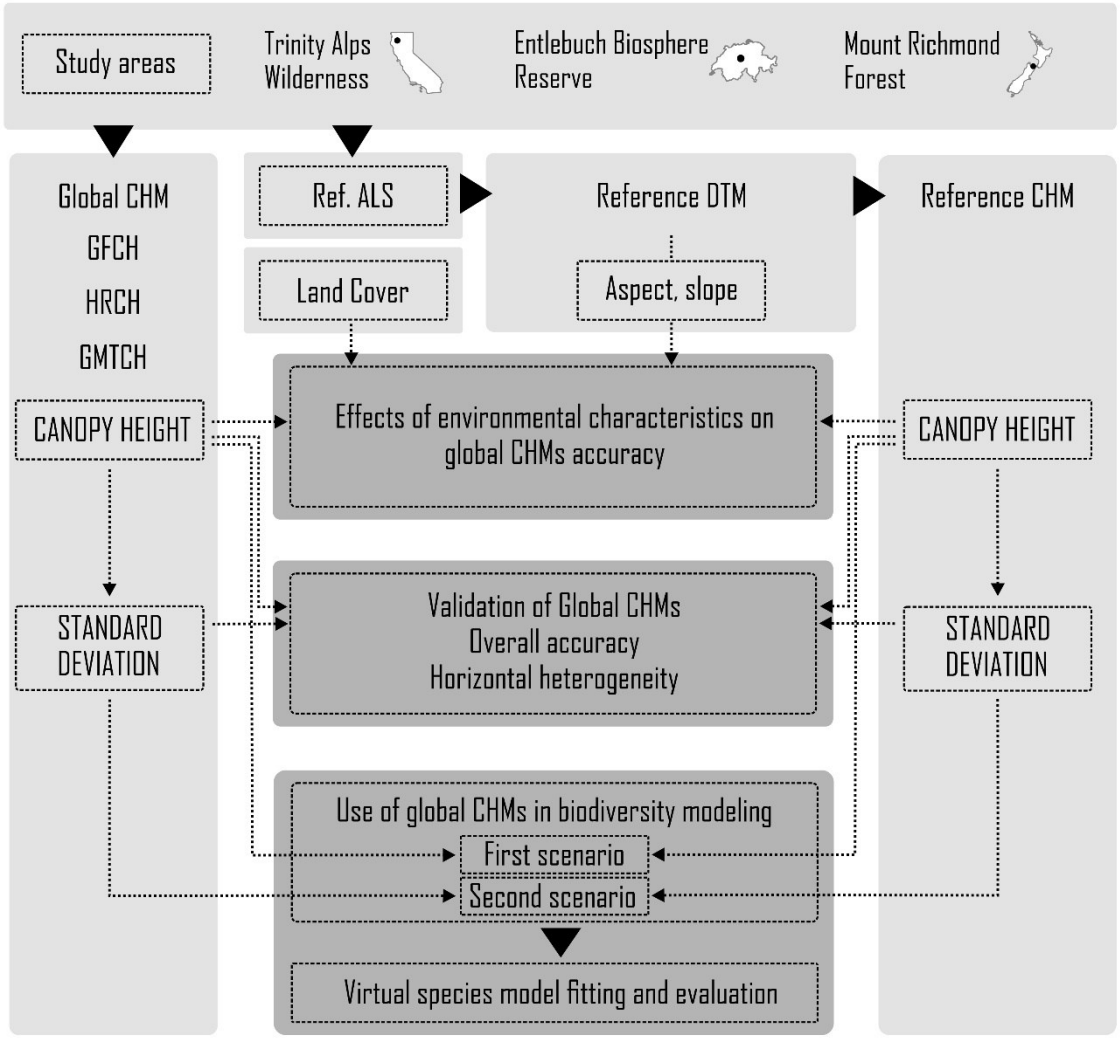


Figure 1. Workflow of the study. ALS: Airborne laser scanning, CHM: Canopy height model, DTM: Digital terrain model, GFCH: Global Forest Canopy Height (Potapov et al. 2021), HRCH: High-Resolution Canopy Height Model of the Earth (Lang et al. 2023), GMTCH: Global Map of Tree Canopy Height (Tolan et al. 2024).

MATERIALS AND METHODS

Study areas

We selected three study areas in the temperate biome (Olson et al. 2001) in California (USA), Switzerland, and New Zealand (Figure 2) based on airborne laser scanning (ALS) data availability between 2019 and 2021 (Table 1). All three sites are ecologically significant regions, each with unique ecosystems and biodiversity. Moreover, they are also renowned for their natural beauty and scenic landscapes. On the one hand, they share similarities in terms of diverse and complex topography; on

the other hand, however, these landscapes differ in several aspects of their characters. While Mount Richmond Forest includes hills and valleys, Entlebuch is a typical pre-alpine landscape with relatively gentle and undulating topography, and Trinity Alps represent a rugged mountain range with complex topography and high peaks. In terms of land cover, all these regions are forested and host a wide variety of ecosystems. Trinity Alps feature high-altitude landscapes with natural forests. This contrasts with the lower-lying terrain found in the Mount Richmond Forest, consisting of grasslands, native and plantation forests, and the cultural landscape of Entlebuch, which includes both cultivated lands, woodlands, and forests.

Table 1. Study areas and LiDAR data characteristics

Study area	Area (km^2)	Altitude (m)	LiDAR acquisition (Year)	Horizontal datum	Vertical datum	Density ($points \cdot m^{-2}$)
Trinity Alps Wilderness	2,588	600 – 2,750	2019 - 2020	NAD 1983	NAVD 88	8
Entlebuch Biosphere Reserve	803	500 – 2,350	2019	LV95	LN02	5
Mount Richmond Forest	6,267	0 – 1,760	2020 - 2021	NZGD 2000	NZVD2016	9

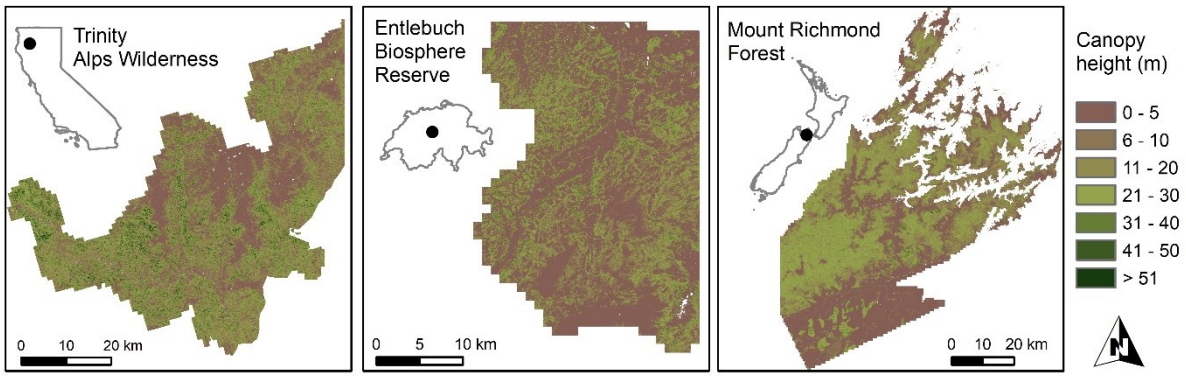


Figure 2. Location, canopy height and land cover within the study areas.

Trinity Alps Wilderness

The Trinity Alps Wilderness is the second largest wilderness area in California with over 200,000 ha of land. The Trinity Alps are a subrange of the Klamath Mountains and are characterized by rugged subalpine topography. The Trinity Alps region has the second-greatest number of conifer species of

any place in the world. They form mixed conifer forests (*Pinus ponderosa, Pinus lambertiana, Pinus contorta, Abies concolor, Pseudotsuga menziesii, Calocedrus decurrens*), red fir forests (*Abies magnifica, Pinus jeffreyi, Pinus monticola*), or subalpine forests (*Pinus balfouriana, Pinus albicaulis, Tsuga mertensiana*) (Ferlatte 1974).

Entlebuch Biosphere Reserve

The cultural landscape of Entlebuch was endorsed as a UNESCO Biosphere Reserve in 2001. The Reserve encompasses almost 40,000 ha of mostly prealpine and alpine mountain chains, forests, agriculturally utilized meadows, and small settlements in the center of Switzerland. At lower altitudes beech (*Fagus sylvatica*) is often naturally dominant, whereas in mountain forests the dominant species are spruce (*Picea abies*) and larch (*Larix decidua*).

Mount Richmond Forest

The study area comprises mainly the northern part of the Marlborough region (South Island of New Zealand), including the Mount Richmond Forest Park and the valley of the Wairau River (Figure 2). Mount Richmond Forest is a forest park established in 1977 in the South Island of New Zealand. The park covers an area of almost 166,000 ha. It is covered with relatively unmodified vegetation including native beech and plantation forests (*Nothofagus spp.*), dense native manuka (*Leptospermum scoparium*) and kanuka (*Leptospermum ericoides*) forests, and pastures. South of the Wairau river, grassland is the main vegetation cover.

Reference airborne laser scanning data

The ALS data for the Trinity Alps Wilderness meet the requirements of the 3D Elevation Program (3DEP), which is managed by the U.S. Geological Survey and designed to collect high-quality LiDAR data for the conterminous United States. The minimum absolute vertical accuracies at the 95% confidence level for 3DEP are 20 cm and 30 cm in non-vegetated and vegetated terrain, respectively (Stoker and Miller 2022). The ALS data for Entlebuch were collected as a part of an acquisition campaign to obtain Switzerland-wide LiDAR data. Their reported planimetric and altimetric accuracies are 20 cm and 10 cm (1 sigma), respectively (Federal Office of Topography in Switzerland). The ALS data for the Mount Richmond Forest were collected as a part of an acquisition campaign to obtain the New Zealand National Elevation Model. The minimum planimetric and altimetric accuracies for vegetated and non-vegetated areas are 50 cm and 10 cm, respectively (National Elevation Programme of New Zealand).

Data pre-processing

The classified ALS point clouds were processed in LAStools (version 220613; academic license). The data was divided into 2 km tiles. A buffer of 10 m was assigned to each tile to avoid edge effects when generating Digital terrain models (DTM) and Canopy height models (CHM). DTMs were generated at a 1 m resolution from points classified as ground using the blast2dem function of LAStools. Subsequently, DTMs were aggregated (using mean cell values) to the resolution of evaluated global canopy height maps (i.e., at 10 m and 30 m). The point clouds were subsequently height-normalised using the lasheight function of LAStools, which calculates the height of each point above the terrain. Canopy height models (CHM) were generated at a 1 m resolution from the normalized point cloud using the lasgrid function of LAStools. To create the CHM, we used the highest returns in individual cells; each return was represented by a small circle (10 cm) to avoid creating pits in the CHMs (Khosravipour et al. 2014). Points classified as noise (class 7 and 18), buildings (class 6), water (class 9), or bridges (class 17) were not considered during CHM generation (LAS Specification 2019). CHMs were subsequently aggregated (using maximum height values within the cell) to the resolution of evaluated global canopy height maps (i.e., at 10 m and 30 m). In addition, cells with an estimated predictive uncertainty higher than 45 m (Lang et al. 2023) as well as cells with a canopy height above 60 meters (Entlebuch, Mount Richmond Forest), and 70 meters (Trinity Alps Wilderness) were removed, as these values are higher than maximum tree height in the respective study areas and were assumed to be outliers. Out of approximately 70.5 million cells, only 421 were removed in this way.

Global canopy height maps

Global Forest Canopy Height (GFCH)

Potapov et al. (2021) published a Global Forest Canopy Height (GFCH) map for the year 2019 with a spatial resolution of 30 m (https://glad.umd.edu/dataset/gedi/). They used Landsat and Shuttle Radar Topography Mission (SRTM) data and a regression tree algorithm to extrapolate GEDI Collection 1 canopy height measurements acquired between April and October 2019. They used the 95th percentile of energy return height relative to the ground (RH95) as the estimate of canopy top height. To ensure high data quality they used only observations collected at night by GEDI's power beams (Dubayah et al. 2020), with beam sensitivity ≥0.9 and with the agreement of the predicted ground elevations among the six algorithm setting groups within ≤2 m (see Table 5 in Hofton and Blair, 2020). In addition, they excluded observations collected during the leaf-off season in temperate and boreal forests. This resulted in a sample size of 372 million GEDI footprints, 10% of which were

used for validation. More recently, an updated version for the year 2020 using GEDI Collection 2 data with better geolocation accuracy has become available (https://glad.umd.edu/dataset/GLCLUC2020; Potapov et al. 2022). Here, however, we stick to using the 2019 version (see also Potapov et al. 2022 and Table 2 for a comparison of the two versions).

High-Resolution Canopy Height Model of the Earth (HRCH)

Lang et al. (2023) created a High-Resolution Canopy Height Model of the Earth (HRCH) for the year 2020 with a spatial resolution of 10 m (https://langnico.github.io/globalcanopyheight/). They used a deep convolutional neural network to integrate GEDI data acquired between April and August of both 2019 and 2020 with Sentinel-2 images of the lowest cloud coverage between May and September 2020. They defined the canopy top height as the relative height at which 98% of the energy was returned (RH98). They filtered out GEDI observations with clouds or snow-covered images. The sample size consisted of 600 million GEDI footprints, 20% of which were used for validation.

Global Map of Tree Canopy Height (GMTCH)

Meta and World Resources Institute created a Global Map of Tree Canopy Height (GMTCH). The map was generated in three steps: (i) the extraction of features from a self-supervised model trained on high resolution (approx. 1 m) imagery from WorldView-2, WorldView-3, and Quickbird II satellites, covering the years 2009 through 2020 (80% of the imagery is from the 2018 to 2020 period). The imagery consisted of 18 million „cloud free“ satellite images of Red, Green, and Blue (RGB) bands, with approx. 0.5 m resolution, from across the globe; (ii) supervised training of a dense prediction decoder using pairs of ALS CHM and RGB satellite images; and (iii) post-processing using a convolutional network trained on 13 million GEDI observations to correct a potential bias coming from a geographically limited source of supervision (Tolan et al. 2024). The dataset creates a global baseline of tree canopy height with a spatial resolution of approximately 1 m; however, for practical reasons, especially to improve visualizations (i.e. comparability with other models) and computational time, we have aggregated the GMTCH to the resolution of 10 meters (using maximum values).

Validation of global canopy height maps

Overall accuracy

We examined the accuracy of HRCH (Lang et al. 2023), GFCH (Potapov et al. 2021), and GMTCH (Tolan et al. 2024) by comparing them with the reference ALS CHMs. The ALS CHMs are considerably more accurate compared to global canopy height maps and can thus be used as the reference dataset (true canopy height). As all global maps represent canopy top height, we used maximum canopy height at 10 m and 30 m resolutions, respectively (i.e. we aggregated ALS CHMs using focal statistics type maximum). Subsequently, we calculated vertical differences between the modeled global canopy height maps and CHMs, using pairwise combinations on a cell-by-cell basis. We used the differences to calculate error metrics for each study area; namely, we calculated the mean error (ME) and root mean square error (RMSE), expressed as:

$$ME = \frac{1}{n}\sum_{i=1}^{n}(h_{GCHMi} - h_{REFi})$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(h_{GCHMi} - h_{REFi})^2}$$

where h_{GCHMi} is the i_{th} canopy height from the global canopy height map, h_{REFi} is the corresponding "true" canopy height from the reference ALS CHM, and n is the number of samples.

Horizontal heterogeneity

Canopy height models (CHMs) provide a representation of habitat heterogeneity. We used the standard deviation of canopy height to estimate horizontal heterogeneity from the modeled global canopy height maps (HRCH, GFCH, GMTCH) and from ALS data. Selecting the appropriate resolution is an important aspect in evaluating the effects of fine-scale habitat heterogeneity on biodiversity. The resolution should be as close as possible to the ecological scale at which a focal species is expected to respond to habitat heterogeneity (see the review by Moudrý et al. 2023a). Prior studies evaluating the effects of fine-scale habitat heterogeneity on biodiversity, for example in birds, have shown that resolutions ranging from hundreds of meters to a few kilometers are appropriate for modeling (Moudrý et al. 2021; Suárez-Castro et al. 2022; Hunt et al. 2023; Roilo et al. 2023). Besides,

the resolution of all datasets is relatively fine (i.e., 1 m, 10 m and 30 m, respectively), and it is, therefore, likely that studies adopting this data will aim for fine resolution. Therefore, we evaluated the accuracy of canopy height horizontal heterogeneity at a 300 m resolution rather than a coarser one.

Use of global CHMs in biodiversity modeling

Understanding the interactions between species and their environment is crucial for predicting species distribution patterns and habitat use, which in turn enhances biodiversity conservation and management. Species distribution models (SDMs) are extensively used to assess species-environment relationships and to predict species distributions across spatial and temporal scales. However, despite their widespread use, SDMs face significant challenges associated in particular with the quality of input data (Rocchini et al. 2011; Fourcade et al. 2018; Araújo et al. 2019; Gábor et al. 2024).

Simulation of virtual species

The virtual species approach is widely used to exemplify the effect of input data (in this case global canopy height datasets) error on SDMs (e.g. Zurell et al. 2010; Meynard & Kaplan 2013; Moudrý 2015). We used a simulated virtual species to ensure the complete knowledge of species distribution in order to enable a proper assessment of model performance without confounding effects of real data, such as positional uncertainty of occurrence data, species response grain, or shape of the response curve (e.g. Zarzo-Arias et al. 2022; Gábor et al. 2022; 2023). The simulated species' niche preference is an indirect way to illustrate the potential errors that could arise when predicting species' habitat preferences using global CHMs.

We first defined the response of virtual species to the canopy characteristics using the ALS CHM. As mentioned, canopy height and horizontal heterogeneity in canopy height are important determinants of the diversity and distribution of various taxa (e.g. Fricker et al. 2015; Feng et al. 2020). Therefore, we simulated two distinct relationships (response functions) between species occurrence and an environmental variable to generate environmental suitability. In the first scenario, we defined a virtual species with specific habitat requirements in terms of canopy height. The response to canopy height was defined as the logistic response function with the following parameters: alpha = -7, beta = 20. Alpha drives the slope of the curve, while beta controls its inflection point. Hence, this virtual species might represent a forest species (e.g. bird) preferring higher canopies. As mentioned in the Introduction, habitat heterogeneity is one of the most important factors affecting species distributions and diversity. Therefore, in the second scenario, we

defined a virtual species habitat requirements in terms of horizontal heterogeneity in canopy height using the standard deviation of the canopy height. The response function was defined as Gaussian. The parameters of the response differed slightly for each study area according to canopy height heterogeneity in respective areas: Trinity Alps Wilderness (mean 10 m, standard deviation 2 m), Mount Richmond Forest (mean 7 m, standard deviation 3 m), Entlebuch Reserve (mean 7 m, standard deviation 3 m). Hence, this virtual species represents a species preferring moderate canopy height heterogeneity.

We adopted a probabilistic approach (logistic function with alpha=−0.15 and beta=0.65) to convert the environmental suitability raster into the probability of occurrence, which is subsequently used to randomly sample presences and absences. We simulated the sampling of equal numbers of presences and absences – namely, 800 samples for the first scenario and 400 for the second, respectively. We always used data of the same resolution to generate the virtual species and to fit the model. This approach assumes that each species interacts with the environment at a single scale (McGarigal et al. 2016; Moudrý et al. 2023a) defined by the resolution of canopy height models used to generate the virtual species. All simulations were undertaken in R v.3.4.3, using the virtualspecies package (Leroy et al. 2016).

Virtual species model fitting and evaluation

It is customary to fit the Gaussian response with a polynomial model. We used generalized linear models (GLMs) with binomial error distribution and logit link function. For the first scenario, we included only linear terms. For the second scenario, both linear and quadratic terms of habitat heterogeneity were included because of the known Gaussian shape of the response functions. We modeled the species distribution using the canopy heights derived from ALS data (same as used for the virtual species generation) as well as using both modeled global canopy height maps.

To evaluate the models, we split the sampled presences-absences into test (50%) and training (50%) datasets. We ran the entire process from species generation to model evaluation 25 times. Each repetition provided a different presence-absence distribution (Leroy et al., 2016). We evaluated the model calibration by plotting the estimated environmental relationships. We assessed the model discrimination capacity by the area under the curve (AUC) of the receiver operating characteristic plot (Fielding and Bell, 1997).

RESULTS

Overall accuracy

The mean error of HRCH shows a positive vertical offset when compared to the maximum canopy height derived from ALS CHMs in all areas. On the other hand, the mean error of GFCH and GMTCH show a negative vertical offset (Table 2). The RMSEs of all models are below 12 m except for GFCH and GMTCH in Trinity Alps Wilderness where the RMSE is greater than 16 m. Canopies taller than 40 m are considerably underestimated by all models, however HRCH outperforms the GFCH and GMTCH in estimating their height. GFCH and GMTCH consistently underestimate the height of all canopies regardless of their height, while HRCH tends to overestimate the height of lower canopies and underestimate taller canopies (Table 3, Figure 3). The vegetation of low height (i.e. below 10 m) is considerably overestimated by HRCH, with GFCH and GMTCH providing better accuracy (Table 3, Figure 3). The BRT models of HRCH, GFCH and GMTCH error explained between 58.2% - 73.1% of variability. All maps underestimated the horizontal heterogeneity of canopy height in all study areas (Figure 4). This coincides with an impaired ability to respond to changes in canopy height, as observed in the canopy height profiles, especially in the case of HRCH (Figure 5). Finally, the canopy height accounted for most of the variability in the global CHMs error, while the effects of slope and orientation were relatively weak (Appendix S1: Table S1). The accuracy of global canopy height maps deteriorated with the increasing slope and the relationship between the accuracy and terrain orientation seemed increasingly evident for greater slopes (Appendix S1: Figure S1).

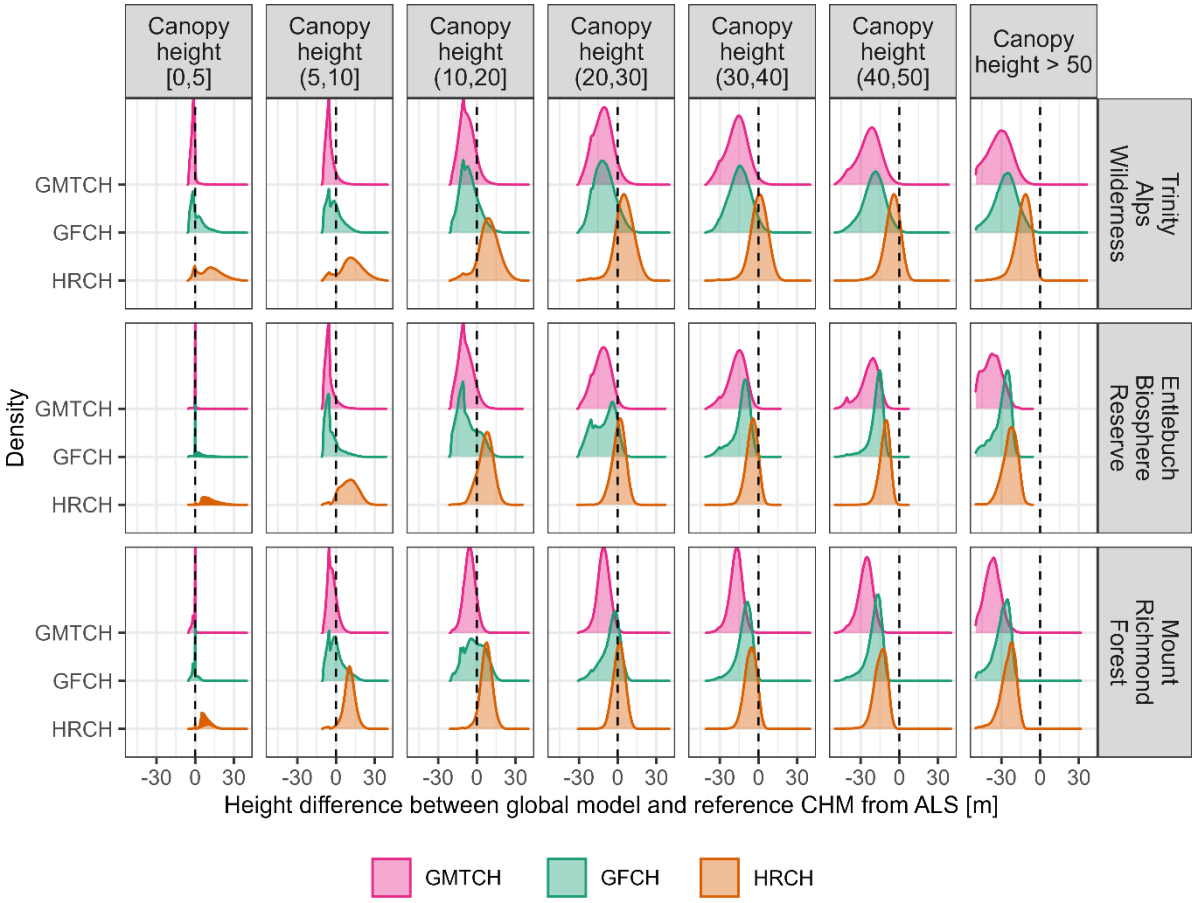
Table 2. Accuracy measures of the High-resolution canopy height model of the Earth (HRCH), the Global Forest Canopy Height (GFCH) version for the years 2019 and 2020, and the Global Map of Tree Canopy Height (GMTCH) for individual study areas based on comparison with ALS-derived CHMs

Locality	HRCH		GFCH 2019 (2020)		GMTCH	
	ME (m)	RMSE (m)	ME (m)	RMSE (m)	ME (m)	RMSE (m)
Trinity Alps Wilderness	4.2	11.2	-14.3 (-14.1)	18.3 (18.0)	-12.6	16.7
Entlebuch Biosphere Reserve	5.9	10.5	-7.1 (-7.6)	11.8 (12.2)	-6.8	10.9
Mount Richmond Forest	4.8	8.8	-4.4 (-4.6)	9.7 (9.6)	-7.2	10.2
Reported by Potapov et al. (2021)			-3.8	9.1		
Reported by Lang et al. (2023)	-4.8 – 5.9	2.8 – 9.6	-7.9 – -1.0	5.2 – 12.0		
Reported by Tolan et al. (2024)					0.6	4.4

Table 3. Accuracy measures of the High-resolution canopy height model of the Earth (HRCH), the Global Forest Canopy Height (GFCH), and the Global Map of Tree Canopy Height (GMTCH) with respect to canopy height derived from ALS reference data.

Locality	Canopy height (m)	HRCH		GFCH		GMTCH	
		ME (m)	RMSE (m)	ME (m)	RMSE (m)	ME (m)	RMSE (m)
Trinity Alps Wilderness	0–5	12	15	2	6	-2	3
	5–10	12	15	-1	6	-5	6
	10–20	9	12	-6	9	-8	10
	20–30	5	9	-10	13	-12	14
	30–40	0	7	-15	17	-17	19
	40–50	-5	8	-20	22	-24	25
	>50	-13	15	-27	29	-33	35
Entlebuch Biosphere Reserve	0–5	12	13	2	4	0	2
	5–10	10	12	-4	7	-5	6
	10–20	7	9	-7	11	-9	10
	20–30	1	5	-11	14	-12	14
	30–40	-5	7	-13	15	-17	18
	40–50	-11	12	-18	19	-24	25
	>50	-23	24	-31	32	-41	42
Mount Richmond Forest	0–5	10	11	0	4	-1	3
	5–10	11	11	0	8	-3	5
	10–20	7	9	-2	9	-6	7
	20–30	1	5	-5	9	-11	12
	30–40	-6	8	-11	13	-18	18
	40–50	-15	15	-20	21	-26	27
	>50	-25	25	-31	32	-38	38

372



373

374 **Figure 3.** Density plots illustrate the distribution of canopy height differences between global maps
375 and ALS-derived data, categorized by canopy height. GFCH: Global Forest Canopy Height (Potapov et
376 al. 2021), HRCH: High-Resolution Canopy Height Model of the Earth (Lang et al. 2023), GMTCH:
377 Global Map of Tree Canopy Height (Tolan et al. 2024).

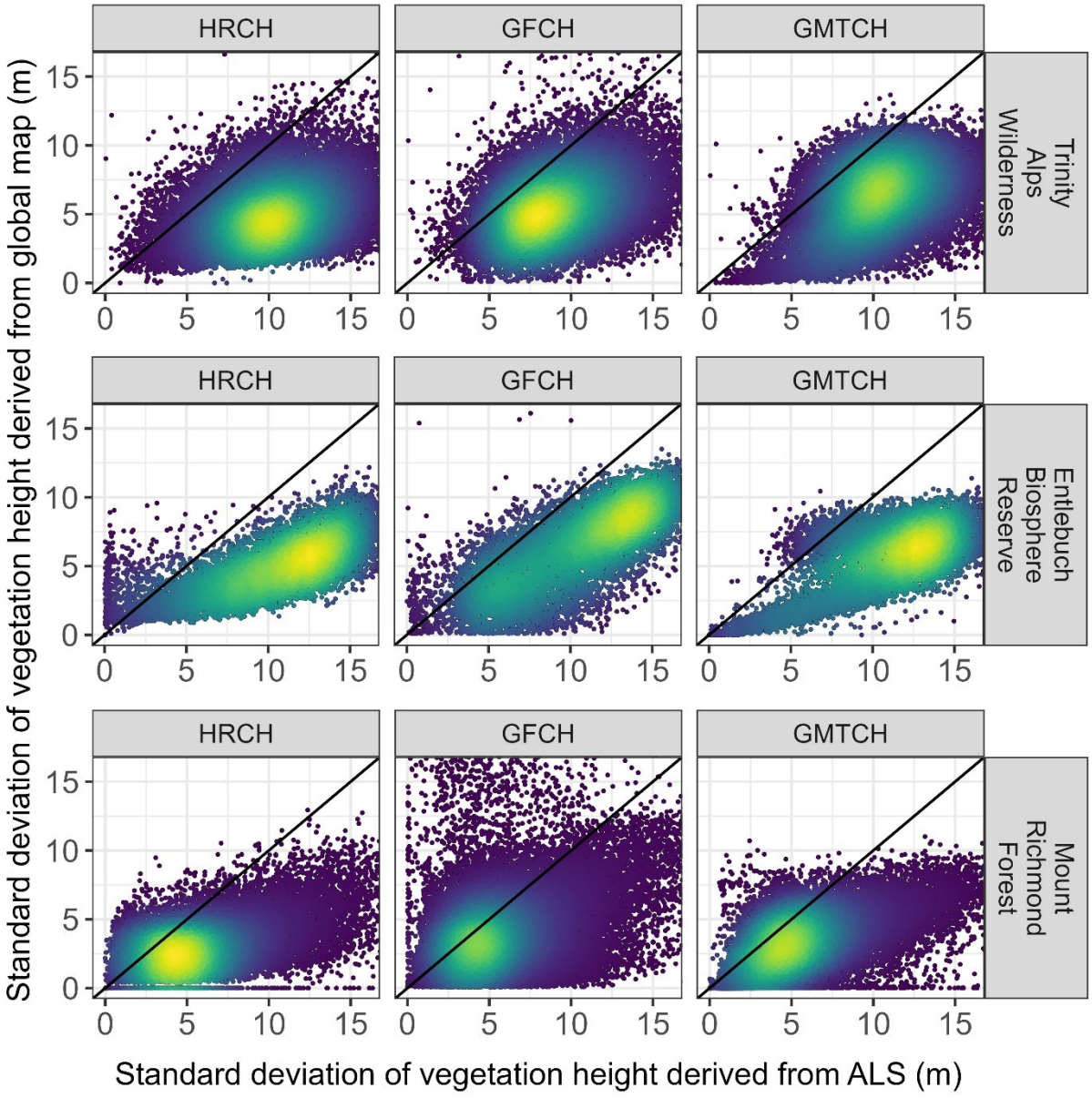


Figure 4. Standard deviation of canopy height derived from global canopy height models versus ALS-derived standard deviation of canopy height. The solid line indicates y = x. Note that all maps considerably underestimate the horizontal heterogeneity of canopy height in all study areas.

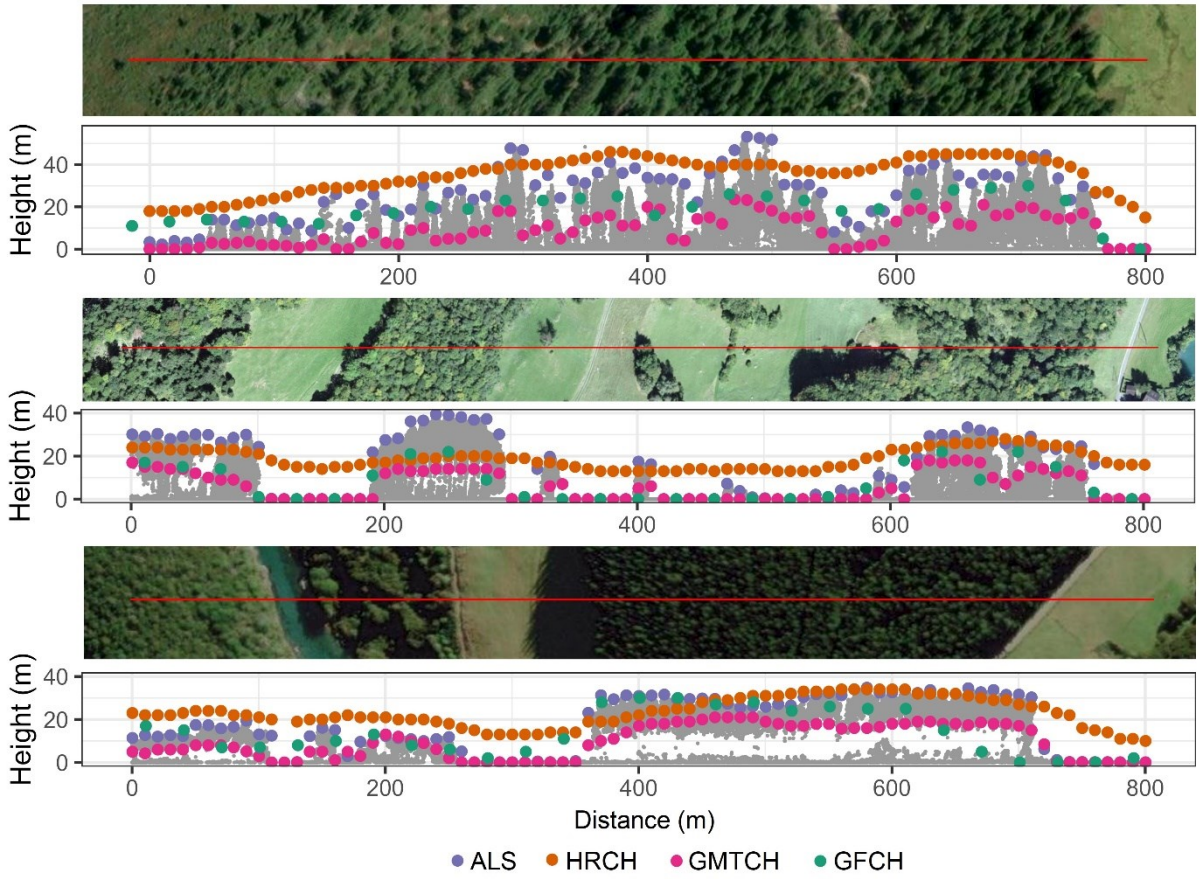


Figure 5. Representative canopy height profiles: Trinity Alp Wilderness (top); Entlebuch Biosphere Reserve (center); Mount Richmond Forest (bottom). The profiles are 10 m wide. Note the impaired ability (particularly for HRCH map) to respond to changes in the canopy height (i.e. the transition between forest and non-forest is unclear). For example, the mosaic of pastures and forests in the Entlebuch study area (bottom) looks like a forest with height ranging between 15 and 25 m in the HRCH map. In contrast, the GMTCH map performs relatively well in distinguishing between forest and non-forest areas. This is due to the considerably better resolution of the input data used to generate this map. However, the biggest underestimation of vegetation height was observed for the GMTCH. The data and code used to plot the canopy height profiles is available on Github (see the Data availability statement). ALS: Airborne laser scanning, CHM: Canopy height model, DTM: Digital terrain model, GFCH: Global Forest Canopy Height (Potapov et al. 2021), HRCH: High-Resolution Canopy Height Model of the Earth (Lang et al. 2023), GMTCH: Global Map of Tree Canopy Height (Tolan et al. 2024).

Biodiversity modeling accuracy

Our results show a decrease in the discrimination ability of the models (i.e. AUC values) when the models are fitted using the global canopy height maps compared to the reference ALS data for both scenarios (Figure 6). For the first scenario (i.e. increasing probability of occurrence with increasing canopy height), the drop was relatively low, and the global models performed well in capturing these relationships. Notably, the HRCH map demonstrated good performance (Figure 6). However, for the second scenario (i.e. species preferring moderate canopy height heterogeneity), the drop in discrimination ability was high and the global models failed to estimate the true relationship (Figure 6). In this scenario, the shape of the estimated environmental relationship was slightly closer to the virtual reality for GMTCH than for HRCH and GFCH. However, none of the models correctly identified the virtual species optimum, which led to the misidentification of the species' optimal niche by several meters.

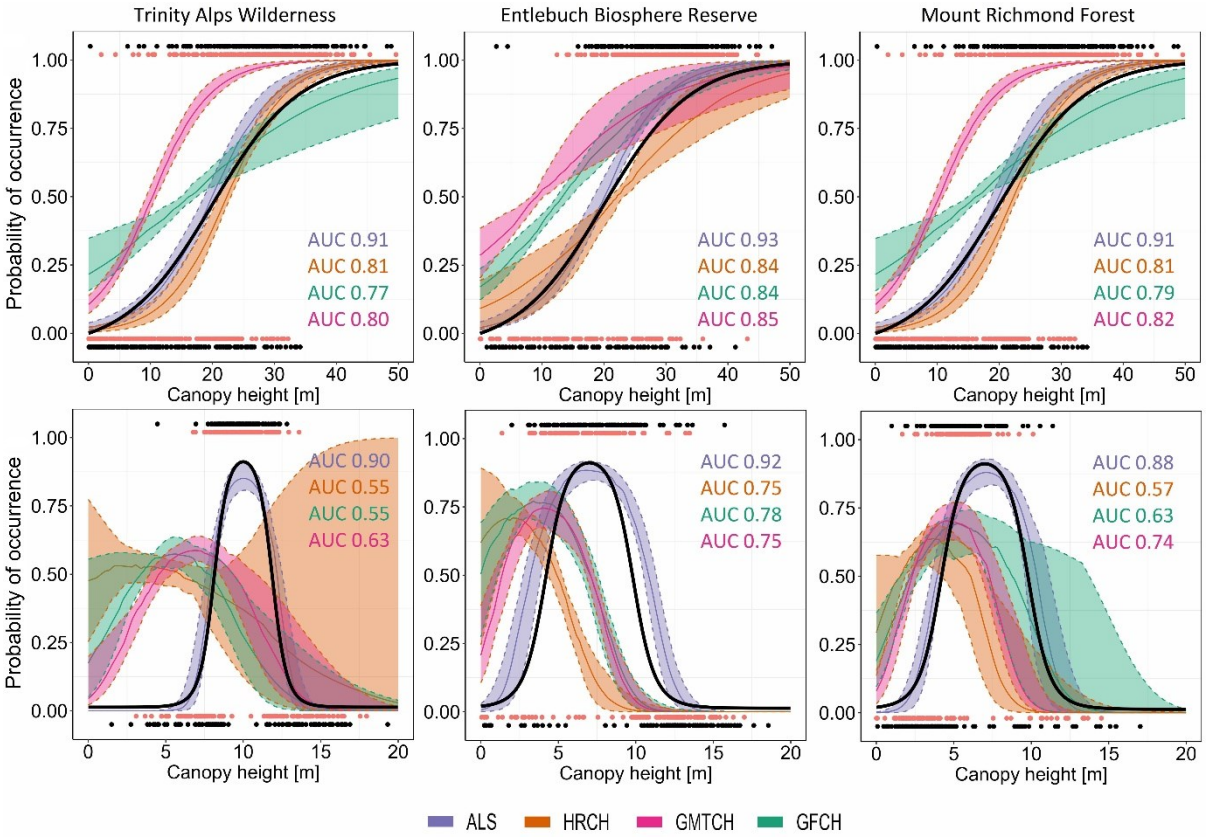


Figure 6. Modeled probability of species occurrence with respect to canopy height (top) and habitat heterogeneity (bottom) estimated with generalized linear models using ALS CHM, GFCH, GMTCH, and HRCH maps. The shaded areas represent the regions delimited by the 5th–95th percentiles of the estimated probability of occurrence obtained from 25 simulations. Black lines show the "true" relationships. Points represent presences and absences used for model training (black) and testing (red). The area under the curve (AUC) of the receiver operating characteristic plot characterizes the

discrimination ability of the models. ALS: Airborne laser scanning, CHM: Canopy height model, DTM: Digital terrain model, GFCH: Global Forest Canopy Height (Potapov et al. 2021), HRCH: High-Resolution Canopy Height Model of the Earth (Lang et al. 2023), GMTCH: Global Map of Tree Canopy Height (Tolan et al. 2024).

DISCUSSION

Global CHMs accuracy

Our results show similar RMSEs as independent validation of GFCH and HRCH provided by Lang et al. (2023) in Mount Richmond Forest (Table 2). The RMSEs in the Entlebuch Biosphere Reserve and in Trinity Alps Wilderness are slightly higher - below 12 m for both maps, with GFCH in Trinity Alps Wilderness being the only exception. In that area, the RMSE of GFCH is greater than 18 m, which is likely related to the presence of large areas with tall canopies (Figure 2). The ME values of HRCH and GFCH found by us are consistent with those reported by Lang et al. (2023) (Table 2). Our results also support the recommendation by Lang et al. (2023) to filter out inaccurate predictions using the estimated predictive uncertainty map. Indeed, the estimated predictive uncertainty correlates well with the absolute canopy height error of the HRCH map, at least as far as errors up to 12 meters are concerned, and can be used as a reliable estimate for the canopy height error (Appendix S1: Figure S2).

Estimation of canopy height using satellite images is typically hampered by the underestimation of tall canopies, with the height estimates saturating at around 30 meters (Potapov et al. 2021; Healey et al. 2020). Our results show that GMTCH and GFCH consistently underestimate the height of all canopies regardless of their height. This is consistent with independent validation of GFCH provided by Lang et al. (2023) who showed that the GFCH map tends to underestimate canopy height in areas where HRCH only slightly overestimates it (see Extended Data Table 1 in Lang et al. 2023). Indeed, HRCH outperforms the GMTCH and GFCH in estimating the height of tall canopies (taller than 30 m; Table 3), consistent with the results of Lang et al. (2023) who highlighted that their approach reduces the saturation effect. However, Lang et al. (2023) reported that the 'price they paid' for improving the performance of tall canopies was a slight overestimation of low canopy heights. According to our results, however, this 'price' is quite high, as the overestimation of low canopies in temperate biome can reach considerably high values (Table 3, Figure 3). HRCH overestimated the height of low canopies (up to 10 m high) on average by 10 m (Table 3). This is well illustrated in the canopy height profiles (Figure 5). For example, the mosaic of pastures and forests in the Entlebuch study area

(Switzerland) appears as a continuous forest in the HRCH map (i.e., the transition between forest and non-forest is unclear), while the GMTCH and GFCH maps correctly show zero canopy height in grasslands (Figure 5). The sensitivity of the GMTCH map to canopy height changes (i.e., transitions between forest and grasslands) is considerably improved compared to the HRCH and GFCH maps due to its finer resolution (Figure 5). In this context, Schwartz et al. (2023) reported that the resolution of HRCH appears to be coarser than 10 meters. This could also be partly related to the fact that the effective spatial resolution of HRCH is coarser than the nominal 10 m ground sampling distance (Lang et al. 2023). This is a consequence of the GEDI footprint (~25 m diameter) and its misalignments caused by geolocation error (Roy et al. 2021). Surprisingly, according to Lang et al. (2023), this should not severely limit the usefulness of the map. However, in our opinion, overestimating data quality increases the risk of incorrect implications derived from its use, which can in effect lead to incorrect management decisions. For this reason, we believe that it should be the producer's responsibility not to overvalue their data product.

Biodiversity modeling

Our results clearly show that the use of global tree canopy height maps for biodiversity modeling allows the estimation of simple relationships between species occurrence and canopy height (Figure 6). However, when derived indices such as canopy height heterogeneity are used, it leads to a considerable decrease in the discriminatory ability of the models and mischaracterization of species niches (but see Torresani et al. 2023 who showed the opposite). When using global canopy height maps, it is extremely important to take the accuracy of the maps and particularly of the derived indices (e.g. habitat heterogeneity and complexity) into consideration before including them in the modeling process.

In this study, we used a simplistic example that utilized a single predictor at a resolution corresponding to the known response grain of a virtual species, which was perfectly sampled without any deficiencies in occurrence data. This is, however, a very unlikely (if not impossible) combination. For example, multiple resolutions are typically tested when developing SDMs and this can considerably affect the propagation of errors from global CHMs (McGarigal et al. 2016; Gábor et al. 2022; Moudrý et al. 2018). Furthermore, deficiencies in species occurrence data, including positional accuracy (Gábor et al. 2023) and sampling bias (Rocchini et al. 2023), must also be taken into account. Failure to account for errors and their interaction can lead to misinterpretation of species-environment relationships and misidentification of areas important for species conservation. Nevertheless, the interaction of errors in species occurrences and environmental predictors remains understudied (but see Standfuß et al. 2024).

Given the poor performance of global CHMs, we attempted to replace global CHMs with Normalized Difference Vegetation Index (NDVI) in the models to evaluate whether their use would be justified at all (see Appendix S1 for details on the methods). The NDVI is generally positively correlated with canopy structure, including its height (Gamon et al. 1995; Pascual et al. 2010). However, models using NDVI performed worse than those using global CHMs in both scenarios. In the first scenario, AUC values were 0.74, 0.67 and 0.74 for Trinity Alps Wilderness, Entlebuch Biosphere Reserve and Mount Richmond Forest, respectively. This is considerably worse than the the performance of global CHMs (Figure 6). In the second scenario, AUC values were below 0.60 in all study areas, which is again worse than the performance of the best global canopy height model. This suggests that the CHMs could have merit compared to indices derived more directly from the spectral data. However, further analyses involving "real" species and a combination of multiple variables (not just one, as in our simplifying case) are needed to validate this.

We showed that vertical bias in global CHMs can prevent researchers from getting the appropriate niche estimate. Local high-accuracy CHMs, such as those derived from ALS data, which are increasingly available through governmental agencies in Europe as well as in other parts of the World, should thus be favored (e.g. Melin et al. 2017; Assmann et al. 2022; Stoker and Miller 2022; Kissling et al. 2023; Moudrý et al. 2023b).

Way forward

Estimating forest characteristics, such as canopy height, using optical images is a challenging task because optical remote sensing is not designed to capture the vertical structure of vegetation and the physical associations between spectral signatures and vertical forest structure are still poorly understood (Goetz & Dubayah 2011; Lang et al. 2023; Rodríguez-Veiga et al. 2017). Models could be improved by using large training datasets (Schwartz et al. 2022) and including additional suitable predictors that are known to affect canopy height, such as climatic and topographic variables (Klein et al. 2015; Simard et al. 2011; Tao et al. 2016; Wang et al. 2019). In addition, it might be valuable to train models with respect to land cover (Healey et al. 2020). For example, Lefsky (2010) developed models independently for broadleaf and needleleaf forests, and Simard et al. (2011) considered 12 land cover classes. Furthermore, it might be valuable to train models at a local scale. For example, Kacic et al. (2023) and Schwartz et al. (2023) developed canopy height models using GEDI and Sentinel data for Germany and France, respectively. Schwartz et al. (2023) showed that their model considerably outperforms both HRCH and GFCH. This might be thanks to the improved ability of their approach to filter out unsuitable GEDI footprints. Furthermore, the models can be improved by removing GEDI footprints on steep terrain (Simard et al. 2011; Bolton et al. 2013), which was not

considered in either HRCH or GFCH. The terrain slope is an important factor known to cause bias in canopy height estimates from space-borne LiDAR (Liu et al. 2021; Moudrý et al. 2022; Quirós et al. 2021). It is particularly problematic for full waveform large footprint LiDAR such as the Geoscience Laser Altimeter System (GLAS) onboard the ICESat satellite (e.g. Simard et al. 2011). The GEDI footprint size is considerably smaller, which reduces the effect of slope on canopy height estimates (Dubayah et al. 2020). However, our results still show an effect of slope and aspect on the accuracy of all maps (Appendix S1: Figure S1). If not accommodated for, such errors can propagate through the modeling and negatively affect canopy height predictions.

CONCLUSION

In this study, we investigated the agreement between three global canopy height maps and canopy height estimates derived from airborne LiDAR in temperate biomes. Our results show that there are large errors in the evaluated global maps. Tall canopies are underestimated by all models, however HRCH map (Lang et al. 2023) outperforms the GFCH (Potapov et al. 2021) and GMTCH (Tolan et al. 2023) maps in estimating their height. On the other hand, HRCH map appeared to be less sensitive to spatial heterogeneity in canopy height, resulting in more uniform canopy height predictions, than the GFCH and GMTCH maps which provided a better estimate of horizontal heterogeneity in canopy height.

Existing errors propagate through to subsequent analyses, which is particularly important for applications in biodiversity modeling as habitat heterogeneity is a key characteristic determining biodiversity patterns. Our virtual species experiments showed that biodiversity models using modeled global canopy height maps perform well for simple relationships, but lead to the mischaracterization of species niches when using derived indices (e.g. standard deviation of canopy height). The modeled global canopy height maps are readily available in raster format, increasing the risk that users would prefer global products to the tedious processing of more accurate ALS data. It is, therefore, particularly important that users are aware of the errors and their impact on biodiversity modeling so they can make informed decisions. We suggest that for temperate areas rich in ALS data (e.g. Europe, USA) activities should concentrate on harmonizing ALS canopy height maps rather than relying on modeled global products.

Acknowledgements

This study was funded by the Horizon Europe project EarthBridge (ID: 101079310) and by the technological grant agency of the Czech Republic (grant no. SS02030018 DivLand). Funded by the

European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them. JP and JW were supported by long-term research development project RVO 67985939 (Czech Academy of Sciences).

Author contributions

VM conceived the ideas and analyzed the data; VM and LG performed the virtual species experiment; VM led the writing with assistance from all authors.

References

Araújo, M. B., Anderson, R. P., Márcia Barbosa, A., Beale, C. M., Dormann, C. F., Early, R., ... & Rahbek, C. (2019). Standards for distribution models in biodiversity assessments. Science advances, 5(1), eaat4858.

Assmann, J. J., Moeslund, J. E., Treier, U. A., & Normand, S. (2022). EcoDes-DK15: High-resolution ecological descriptors of vegetation and terrain derived from Denmark's national airborne laser scanning data set. Earth System Science Data, 14(2), 823-844.

Atkins, J. W., Bhatt, P., Carrasco, L., Francis, E., Garabedian, J. E., Hakkenberg, C. R., ... & Krause, K. (2023). Integrating forest structural diversity measurement into ecological research. Ecosphere, 14(9), e4633.

Bergen, K. M., Goetz, S. J., Dubayah, R. O., Henebry, G. M., Hunsaker, C. T., Imhoff, M. L., ... & Radeloff, V. C. (2009). Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. Journal of Geophysical Research: Biogeosciences, 114(G2).

Bolton, D. K., Coops, N. C., & Wulder, M. A. (2013). Investigating the agreement between global canopy height maps and airborne Lidar derived height estimates over Canada. Canadian Journal of Remote Sensing, 39(sup1), S139-S151.

Cazzolla Gatti, R. (2016). A conceptual model of new hypothesis on the evolution of biodiversity. Biologia, 71(3), 343-351.

Cazzolla Gatti, R., Di Paola, A., Bombelli, A., Noce, S., & Valentini, R. (2017). Exploring the relationship between canopy height and terrestrial plant diversity. Plant ecology, 218(7), 899-908.

Coddington, C. P., Cooper, W. J., Rutt, C. L., Mokross, K., Amaral, B. R., Stouffer, P. C., & Luther, D. A. (2023). Amazonian mixed-species flocks demonstrate flexible preferences for vertical forest structure. Ecosphere, 14(12), e4720.

Cooper, W. J., McShea, W. J., Forrester, T., & Luther, D. A. (2020). The value of local habitat heterogeneity and productivity when estimating avian species richness and species of concern. Ecosphere, 11(5), e03107.

Dănescu, A., Albrecht, A. T., & Bauhus, J. (2016). Structural diversity promotes productivity of mixed, uneven-aged forests in southwestern Germany. Oecologia, 182(2), 319-333.

Davies, A. B., & Asner, G. P. (2014). Advances in animal ecology from 3D-LiDAR ecosystem mapping. Trends in ecology & evolution, 29(12), 681-691.

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., ... & Silva, C. (2020). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. Science of remote sensing, 1, 100002.

Duncanson, L., Armston, J., Disney, M., Avitabile, V., Barbier, N., Calders, K., ... & Williams, M. (2019). The importance of consistent global forest aboveground biomass product validation. Surveys in geophysics, 40(4), 979-999.

Feng, G., Zhang, J., Girardello, M., Pellissier, V., & Svenning, J. C. (2020). Forest canopy height co-determines taxonomic and functional richness, but not functional dispersion of mammals and birds globally. Global Ecology and Biogeography, 29(8), 1350-1359.

Ferlatte, W. J. (1974). A flora of the Trinity Alps of northern Calilornia.

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental conservation, 24(1), 38-49.

Fischer, F. J., Jackson, T. D., Vincent, G., & Jucker, T. (2024). Robust characterization of forest structure from airborne laser scanning-a systematic assessment and sample workflow for ecologists. *Methods in Ecology and Evolution*. 00, 1-16.

Fischer, R., Knapp, N., Bohn, F., Shugart, H. H., & Huth, A. (2019). The relevance of forest structure for biomass and productivity in temperate forests: New perspectives for remote sensing. *Surveys in Geophysics*, *40*(4), 709-734.

FOTS (2023). Federal Office of Topography swisstopo, The classified point cloud of Switzerland [Dataset]. https://www.swisstopo.admin.ch/en/height-model-swisssurface3d. Accessed: 2024-07-23

Fourcade, Y., Besnard, A. G., & Secondi, J. (2018). Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. Global Ecology and Biogeography, 27(2), 245-256.

Fricker, G. A., Crampton, L. H., Gallerani, E. M., Hite, J. M., Inman, R., & Gillespie, T. W. (2021). Application of lidar for critical endangered bird species conservation on the island of Kauai, Hawaii. *Ecosphere*, *12*(6), e03554.Gallerani, E. M., Burgett, J., Vaughn, N., Fortini, L. B., Fricker, G. A., Mounce, H., ... & Gilb, R. (2023). High resolution lidar data shed light on inter-island translocation of endangered bird species in the Hawaiian Islands. Ecological Applications, e2889.

Gábor, L., Jetz, W., Lu, M., Rocchini, D., Cord, A., Malavasi, M., ... & Moudrý, V. (2022). Positional errors in species distribution modelling are not overcome by the coarser grains of analysis. Methods in Ecology and Evolution, 13(10), 2289-2302.

Gábor, L., Jetz, W., Zarzo-Arias, A., Winner, K., Yanco, S., Pinkert, S., ... & Moudrý, V. (2023). Species distribution models affected by positional uncertainty in species occurrences can still be ecologically interpretable. Ecography, 2023(6), e06358.

Gábor, L., Cohen, J., Moudrý, V., & Jetz, W. (2024). Assessing the applicability of binary land-cover variables to species distribution models across multiple grains. Landscape Ecology, 39(3), 66.

Gamon, J. A., Field, C. B., Goulden, M. L., Griffin, K. L., Hartley, A. E., Joel, G., ... & Valentini, R. (1995). Relationships between NDVI, canopy structure, and photosynthesis in three Californian vegetation types. Ecological applications, 5(1), 28-41.

Goetz, S., & Dubayah, R. (2011). Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. Carbon Management, 2(3), 231-244.

Hancock, S., McGrath, C., Lowe, C., Davenport, I., & Woodhouse, I. (2021). Requirements for a global lidar system: spaceborne lidar with wall-to-wall coverage. Royal Society open science, 8(12), 211166.

Healey, S. P., Yang, Z., Gorelick, N., & Ilyushchenko, S. (2020). Highly local model calibration with a new GEDI LiDAR asset on Google Earth Engine reduces landsat forest height signal saturation. Remote Sensing, 12(17), 2840.

Hofton, M., Blair, J. B., Story, S., & Yi, D. (2020). Algorithm Theoretical Basis Document (ATBD).

Huber, N., Kéry, M., & Pasinelli, G. (2017). Occupancy dynamics of the Wood Warbler Phylloscopus sibilatrix assessed with habitat and remote sensing data. Ibis, 159(3), 623-637.

638 Hunt, M. L., Blackburn, G. A., Siriwardena, G. M., Carrasco, L., & Rowland, C. S. (2023). Using satellite
639 data to assess spatial drivers of bird diversity. Remote Sensing in Ecology and Conservation, 9(4),
640 483-500.

641 Kacic, P., Thonfeld, F., Gessner, U., & Kuenzer, C. (2023). Forest structure characterization in
642 Germany: novel products and analysis based on GEDI, sentinel-1 and sentinel-2 data. Remote
643 Sensing, 15(8), 1969.

644 Klein, T., Randin, C., & Körner, C. (2015). Water availability predicts forest canopy height at the global
645 scale. Ecology letters, 18(12), 1311-1320.

646 Khosravipour, A., Skidmore, A. K., Isenburg, M., Wang, T., & Hussin, Y. A. (2014). Generating pit-free
647 canopy height models from airborne lidar. Photogrammetric Engineering & Remote Sensing, 80(9),
648 863-872.

649 Kissling, W. D., Shi, Y., Koma, Z., Meijer, C., Ku, O., Nattino, F., ... & Grootes, M. W. (2023). Country-
650 wide data of ecosystem structure from the third Dutch airborne laser scanning survey. Data in Brief,
651 46, 108798.

652 Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2023). A high-resolution canopy height model of the
653 Earth. Nature Ecology & Evolution, 1-12.

654 LaRue, E. A., Fahey, R. T., Alveshere, B. C., Atkins, J. W., Bhatt, P., Buma, B., ... & Fei, S. (2023). A
655 theoretical framework for the ecological role of three-dimensional structural diversity. Frontiers in
656 Ecology and the Environment, 21(1), 4-13.

657 LAS Specification (2019). Specification 1.4–R15. The American Society for Photogrammetry & Remote
658 Sensing. Available online: http://www.asprs.org/wp-content/uploads/2019/03/LAS_1_4_r14.pdf
659 (accessed on 21 October 2023).

660 Lefsky, M. A., Harding, D. J., Keller, M., Cohen, W. B., Carabajal, C. C., Del Bom Espirito-Santo, F., ... &
661 de Oliveira Jr, R. (2005). Estimates of forest canopy height and aboveground biomass using ICESat.
662 Geophysical research letters, 32(22).

663 Lefsky, M. A. (2010). A global forest canopy height map from the Moderate Resolution Imaging
664 Spectroradiometer and the Geoscience Laser Altimeter System. Geophysical Research Letters,
665 37(15).

666 Leroy, B., Meynard, C. N., Bellard, C., & Courchamp, F. (2016). virtualspecies, an R package to
667 generate virtual species distributions. Ecography, 39(6), 599-607.

Lewis, W. B., Chandler, R. B., Delancey, C. D., Rushton, E., Wann, G. T., McConnell, M. D., & Martin, J. A. (2022). Abundance and distribution of ruffed grouse Bonasa umbellus at the southern periphery of the range. Wildlife Biology, e01017.

LINZ (2023). Marlborough District Council, Toitū Te Whenua Land Information New Zealand. Marlborough, New Zealand 2020-2022. Collected by Aerial Surveys, distributed by OpenTopography and LINZ [Dataset]. https://doi.org/10.5069/G97D2SB0. Accessed: 2024-07-23

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.

Magruder, L., Leigh, H., & Neuenschwander, A. (2016). Evaluation of terrain and canopy height products in central African tropical forests. International Journal of Remote Sensing, 37(22), 5365-5387.

Magruder, L., Neumann, T., & Kurtz, N. (2021). ICESat-2 early mission synopsis and observatory performance. Earth and Space Science, 8(5), e2020EA001555.

Marselis, S. M., Tang, H., Armston, J., Abernethy, K., Alonso, A., Barbier, N., ... & Dubayah, R. (2019). Exploring the relation between remotely sensed vertical canopy structure and tree species diversity in Gabon. Environmental Research Letters, 14(9), 094013.

Marselis, S. M., Keil, P., Chase, J. M., & Dubayah, R. (2022). The use of GEDI canopy structure for explaining variation in tree species richness in natural forests. Environmental Research Letters, 17(4), 045003.

McGarigal, K., Wan, H. Y., Zeller, K. A., Timm, B. C., & Cushman, S. A. (2016). Multi-scale habitat selection modeling: a review and outlook. Landscape ecology, 31, 1161-1175.

Melin, M., Shapiro, A., & Glover-Kapfer, P. (2017). Lidar for ecology and conservation. WWF Conservation Technology Series 1 (3), WWF-UK, Woking, United Kingdom.

Meyer, H., & Pebesma, E. (2022). Machine learning-based global maps of ecological variables and the challenge of assessing them. Nature Communications, 13(1), 1-4.

Meynard, C. N., & Kaplan, D. M. (2013). Using virtual species to study species distributions and model performance. Journal of Biogeography, 40(1), 1-8.

Moudrý, V. (2015). Modelling species distributions with simulated virtual species. Journal of Biogeography, 42(8), 1365-1366.

Moudrý, V., Lecours, V., Gdulová, K., Gábor, L., Moudrá, L., Kropáček, J., & Wild, J. (2018). On the use of global DEMs in ecological modelling and the accuracy of new bare-earth DEMs. Ecological Modelling, 383, 3-9.

Moudrý, V., Moudrá, L., Barták, V., Bejček, V., Gdulová, K., Hendrychová, M., ... & Šálek, M. (2021). The role of the vegetation structure, primary productivity and senescence derived from airborne LiDAR and hyperspectral data for birds diversity and rarity on a restored site. Landscape and Urban Planning, 210, 104064.

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.

Moudrý, V., Cord, A.F., Gábor, L., Laurin, G.V., Barták, V., Gdulová, K., Malavasi, M., Rocchini, D., Stereńczak, K., Prošek, J., Klápště, P., Wild, J. (2023b). Vegetation structure derived from airborne laser scanning to assess species distribution and habitat suitability: the way forward. Diversity and Distributions , 29(1), 39-50

Moudrý, Vítězslav, & Gábor, Lukáš. (2022). Comparison of high-resolution global canopy height maps and their applicability to biodiversity modeling in temperate biomes - dataset (1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7332572

Moudrý, V., Keil, P., Cord, A. F., Gábor, L., Lecours, V., Zarzo-Arias, A., ... & Šímová, P. (2023a). Scale mismatches between predictor and response variables in species distribution modelling: A review of practices for appropriate grain selection. Progress in Physical Geography: Earth and Environment, 03091333231156362.

Müller, J., Moning, C., Baessler, C., Heurich, M., & Brandl, R. (2009). Using airborne laser scanning to model potential abundance and assemblages of forest passerines. Basic and Applied Ecology, 10(7), 671-681

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V., Underwood, E. C., ... & Kassem, K. R. (2001). Terrestrial Ecoregions of the World: A New Map of Life on EarthA new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. BioScience, 51(11), 933-938.

Pascual, C., García-Abril, A., Cohen, W. B., & Martin-Fernandez, S. (2010). Relationship between LiDAR-derived forest canopy height and Landsat images. International Journal of Remote Sensing, 31(5), 1261-1280.

728 Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A., ... & Hofton,
729 M. (2021). Mapping global forest canopy height through integration of GEDI and Landsat data.
730 Remote Sensing of Environment, 253, 112165.

731 Potapov, P., Hansen, M. C., Pickens, A., Hernandez-Serna, A., Tyukavina, A., Turubanova, S., ... &
732 Kommareddy, A. (2022). The global 2000-2020 land cover and land use change dataset derived from
733 the Landsat archive: first results. Frontiers in Remote Sensing, 3, 856903.

734 Quirós, E., Polo, M. E., & Fragoso-Campón, L. (2021). GEDI elevation accuracy assessment: a case
735 study of southwest Spain. IEEE Journal of Selected Topics in Applied Earth Observations and Remote
736 Sensing, 14, 5285-5299.

737 Rocchini, D., Hortal, J., Lengyel, S., Lobo, J. M., Jimenez-Valverde, A., Ricotta, C., ... & Chiarucci, A.
738 (2011). Accounting for uncertainty when mapping species distributions: the need for maps of
739 ignorance. Progress in Physical Geography, 35(2), 211-226.

740 Rocchini, D., Tordoni, E., Marchetto, E., Marcantonio, M., Barbosa, A. M., Bazzichetto, M., ... &
741 Malavasi, M. (2023). A quixotic view of spatial bias in modelling the distribution of species and their
742 diversity. npj Biodiversity, 2(1), 10.

743 Rodríguez-Veiga, P., Wheeler, J., Louis, V., Tansey, K., & Balzter, H. (2017). Quantifying forest biomass
744 carbon stocks from space. Current Forestry Reports, 3(1), 1-18.

745 Roilo, S., Engler, J. O., Václavík, T., & Cord, A. F. (2023). Landscape-level heterogeneity of
746 agri-environment measures improves habitat suitability for farmland birds. Ecological Applications,
747 33(1), e2720.

748 Roy, D. P., Kashongwe, H. B., & Armston, J. (2021). The impact of geolocation uncertainty on GEDI
749 tropical forest canopy height estimation and change monitoring. Science of Remote Sensing, 4,
750 100024.

751 Schutz, B. E., Zwally, H. J., Shuman, C. A., Hancock, D., & DiMarzio, J. P. (2005). Overview of the ICESat
752 mission. Geophysical Research Letters, 32(21).

753 Schwartz, M., Ciais, P., Ottlé, C., De Truchis, A., Vega, C., Fayad, I., ... & Wigneron, J. P. (2022). High-
754 resolution canopy height map in the Landes forest (France) based on GEDI, Sentinel-1, and Sentinel-2
755 data with a deep learning approach. arXiv preprint arXiv:2212.10265.

756 Schwartz, M., Ciais, P., De Truchis, A., Chave, J., Ottlé, C., Vega, C., ... & Fayad, I. (2023). FORMS:
757 Forest Multiple Source height, wood volume, and biomass maps in France at 10 to 30 m resolution

based on Sentinel-1, Sentinel-2, and GEDI data with a deep learning approach. Earth System Science Data Discussions, 2023, 1-28.

Seavy, N. E., Viers, J. H., & Wood, J. K. (2009). Riparian bird response to vegetation structure: a multiscale analysis using LiDAR measurements of canopy height. Ecological Applications, 19(7), 1848-1857.

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. Journal of Geophysical Research: Biogeosciences, 116(G4).

Simonson, W. D., Allen, H. D., & Coomes, D. A. (2014). Applications of airborne lidar for the assessment of animal species diversity. Methods in Ecology and Evolution, 5(8), 719-729.

Šímová, P., Moudrý, V., Komárek, J., Hrach, K., & Fortin, M. J. (2019). Fine scale waterbody data improve prediction of waterbird occurrence despite coarse species data. Ecography, 42(3), 511-520.

Skidmore, A. K., Coops, N. C., Neinavaz, E., Ali, A., Schaepman, M. E., Paganini, M., ... & Wingate, V. (2021). Priority list of biodiversity metrics to observe from space. Nature ecology & evolution, 5(7), 896-906.

Standfuß, I., Geiß, C., Senaratne, H., Kerr, G., Nathan, R., Rotics, S., & Taubenböck, H. (2024). Assessing cumulative uncertainties of remote sensing time series and telemetry data in animal-environment studies. Landscape Ecology, 39(2), 7.

Stoker, J., & Miller, B. (2022). The accuracy and consistency of 3d elevation program data: a systematic analysis. Remote Sensing, 14(4), 940.

Suárez-Castro, A. F., Maron, M., Mitchell, M. G., & Rhodes, J. R. (2022). Disentangling direct and indirect effects of landscape structure on urban bird richness and functional diversity. Ecological Applications, e2713.

Tao, S., Guo, Q., Li, C., Wang, Z., & Fang, J. (2016). Global patterns and determinants of forest canopy height. Ecology, 97(12), 3265-3270.

Tew, E. R., Conway, G. J., Henderson, I. G., Milodowski, D. T., Swinfield, T., & Sutherland, W. J. (2022). Recommendations to enhance breeding bird diversity in managed plantation forests determined using LiDAR. Ecological Applications, 32(7), e2678.

Tews, J., Brose, U., Grimm, V., Tielbörger, K., Wichmann, M. C., Schwager, M., & Jeltsch, F. (2004). Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. Journal of biogeography, 31(1), 79-92.

Tolan, J., Yang, H. I., Nosarzewski, B., Couairon, G., Vo, H. V., Brandt, J., ... & Couprie, C. (2024). Very high resolution canopy height maps from RGB imagery using self-supervised vision transformer and convolutional decoder trained on aerial lidar. Remote Sensing of Environment, 300, 113888.

Torresani, M., Rocchini, D., Sonnenschein, R., Zebisch, M., Hauffe, H. C., Heym, M., ... & Tonon, G. (2020). Height variation hypothesis: A new approach for estimating forest species diversity with CHM LiDAR data. Ecological Indicators, 117, 106520.

Torresani, M., Rocchini, D., Alberti, A., Moudrý, V., Heym, M., Thouverai, E., ... & Tomelleri, E. (2023). LiDAR GEDI derived tree canopy height heterogeneity reveals patterns of biodiversity in forest ecosystems. Ecological Informatics, 76, 102082.

USGS (2023). U.S. Geological Survey, Lidar Point Cloud - USGS National Map 3DEP Downloadable Data Collection: U.S. Geological Survey [Dataset]. https://apps.nationalmap.gov/downloader/. Accessed: 2024-07-23

Walter, J. A., Stovall, A. E., & Atkins, J. W. (2021). Vegetation structural complexity and biodiversity in the Great Smoky Mountains. Ecosphere, 12(3), e03390.

Wang, Z., Li, Y., Su, X., Tao, S., Feng, X., Wang, Q., ... & Enquist, B. J. (2019). Patterns and ecological determinants of woody plant height in eastern Eurasia and its relation to primary productivity. Journal of Plant Ecology, 12(5), 791-803.

Wolf, J. A., Fricker, G. A., Meyer, V., Hubbell, S. P., Gillespie, T. W., & Saatchi, S. S. (2012). Plant species richness is associated with canopy height and topography in a neotropical forest. Remote Sensing, 4(12), 4010-4021.

Zarzo-Arias, A., Penteriani, V., Gábor, L., Šímová, P., Grattarola, F., & Moudrý, V. (2022). Importance of data selection and filtering in species distribution models: A case study on the Cantabrian brown bear. Ecosphere, 13(12), e4284.

Zurell, D., Berger, U., Cabral, J. S., Jeltsch, F., Meynard, C. N., Münkemüller, T., ... & Grimm, V. (2010). The virtual ecologist approach: simulating data and observers. Oikos, 119(4), 622-635.

Appendix S1

Ecosphere

Comparison of three high-resolution global canopy height maps and their applicability to biodiversity modeling: Accuracy issues revealed

Authors

Vítězslav Moudrý[1,2,3] Lukáš Gábor[1], Suzanne Marselis[4], Petra Pracná[1], Vojtěch Barták[1], Jiří Prošek[1,3], Barbora Navrátilová[5], Jan Novotný[5], Markéta Potůčková[6], Kateřina Gdulová[1], Pablo Crespo-Peremarch[7], Jan Komárek[1], Marco Malavasi[1,8], Duccio Rocchini [1,9], Luis A. Ruiz[7], Jesús Torralba[7], Michele Torresani[10], Roberto Cazzolla Gatti[9], and Jan Wild [1,3]

[1]Department of Spatial Sciences, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, 16500 Praha-Suchdol, Czech Republic.

[2]Institute for Environmental Studies, Faculty of Science, Charles University, Benátská 2, 12801 Prague 2, Czech Republic.

[3]Institute of Botany of the Czech Academy of Sciences, Zámek 1, 252 43 Průhonice, Czech Republic.

[4]Institute of Environmental Sciences, Leiden University, Einsteinweg 2, 2333 CC Leiden, the Netherlands.

[5]Global Change Research Institute of the Czech Academy of Sciences, Bělidla 4a, 60300 Brno, Czech Republic.

[6]Charles University, Faculty of Science, Department of Applied Geoinformatics and Cartography Albertov 6, 128 43 Praha 2, Czech Republic.

[7]Geo-Environmental Cartography and Remote Sensing Group (CGAT), Universitat Politècnica de València, Camí de Vera s/n, 46022 Valencia, Spain.

[8]Department of Chemistry, Physics, Mathematics and Natural Sciences, University of Sassari, Via Vienna 2, 07100 Sassari, Italy.

[9]BIOME Lab, Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum University of Bologna, via Irnerio 42, 40126, Bologna, Italy.

[10]Free University of Bolzano/Bozen, Faculty of Science and Technology, Piazza Universitá / Universitätsplatz 1, 39100, Bolzano/Bozen, Italy.

Effects of environmental characteristics on CHM accuracy

The accuracy of GEDI canopy heights computed from full waveforms depends, among other things, on the ground target characteristics such as terrain slope, canopy height, and land cover (Liu et al. 2021; Urbazaev et al. 2022). In addition, visual inspection of the error of predicted CHMs revealed that terrain orientation may play a role, likely due to the effects of shadows on the reflectance values of satellite images used in the prediction. The accuracy of the modeled global canopy height maps may, among other things, also depend on these characteristics. Terrain aspect and slope were derived from ALS DTMs and resampled to the resolution of the corresponding global canopy height map using Horn's algorithm with a 3 × 3 cell neighborhood implemented in ArcGIS (version 10.8.1) to calculate terrain aspect and slope (Moudrý et al. 2019). Land cover was derived from the ESA world cover product at a 10 m resolution for the year 2020 (https://esa-worldcover.org/en; Zanaga et al. 2022). Canopy height was derived from ALS and aggregated (using maximum height values within each cell) to the 10 m and 30 m resolution, respectively.

We used Boosted Regression Trees (BRT) to investigate the relative contribution of terrain slope, aspect, land cover, and canopy height to the error of the canopy height map. BRT is an advanced algorithm of machine learning that combines multiple regression tree models iteratively to enhance predictive accuracy (Elith, Leathwick, & Hastie, 2008). This method has been used extensively in various disciplines (e.g., Moudrý et al. 2021; Naghibi et al. 2022; Wu et al. 2022) and is at least equally efficient as other computationally intensive techniques like random forests (Valavi et al. 2022).

Two important parameters that have to be specified for BRT are the tree complexity (which controls whether interactions are incorporated) and shrinkage (learning rate). The shrinkage parameter is used to control the contribution of each individual tree to the final model. A smaller value (e.g., 0.01) means that the contribution of each tree is more modest, and the algorithm learns slowly. Conversely, a larger value (e.g., 0.1) allows each tree to have a more substantial effect on the model, leading to faster learning. Tree complexity and shrinkage parameters determine the number of trees needed for the prediction. As a rule of thumb, a combination leading to a model with a minimum of 1000 trees is recommended. We initialized the models with the following characteristics: a shrinkage of 0.01, a bag fraction (the proportion of data used when selecting the optimal tree number) of 0.5, and tree complexity of 2 (i.e. model with up to two-way interactions). We used a 5-fold cross-validation to estimate the optimal number of trees. At each iteration, the residual deviance was calculated and the number of trees giving the best model (i.e. lowest deviance) was identified (Elith, Leathwick, & Hastie, 2008). Canopy height error was modeled specifying the Gaussian error distribution. We assessed the relative importance of each variable using formulae developed by Friedman (2001). All models were fitted using the package gbm (version 2.1.5; Greenwell et al. 2019) and dismo (version 3.6.0; Hijmans et al. 2017) in R.

TABLE S1. The relative importance of environmental conditions for the accuracy of global canopy height models, and the variability in canopy height error explained by the models.

Locality	Model	Canopy height [%]	Landcover [%]	Aspect [%]	Slope [%]	Mean total deviance	Variability explained (%)
Trinity Alps Wilderness	HRCH	86.1	8.4	5.4	0.1	107.268	58.2
Entlebuch Biosphere Reserve		91.3	4.1	2.1	2.5	74.922	70.3
Mount Richmond Forest		94.4	2.6	2.2	0.8	53.029	72.4
Trinity Alps Wilderness	GFCH	86.1	3.5	10.2	0.2	127.994	60.0
Entlebuch Biosphere Reserve		86.8	4.6	7.2	1.5	87.14	62.6
Mount Richmond Forest		85.8	6.1	3.8	4.3	73.667	48.7
Trinity Alps Wilderness	GMTCH	97.6	1.2	0.8	0.4	119.0	61.0
Entlebuch Biosphere Reserve		98.4	1.1	0.4	0.1	72.7	70.7
Mount Richmond Forest		95.3	0.7	3.6	0.4	52.5	73.1

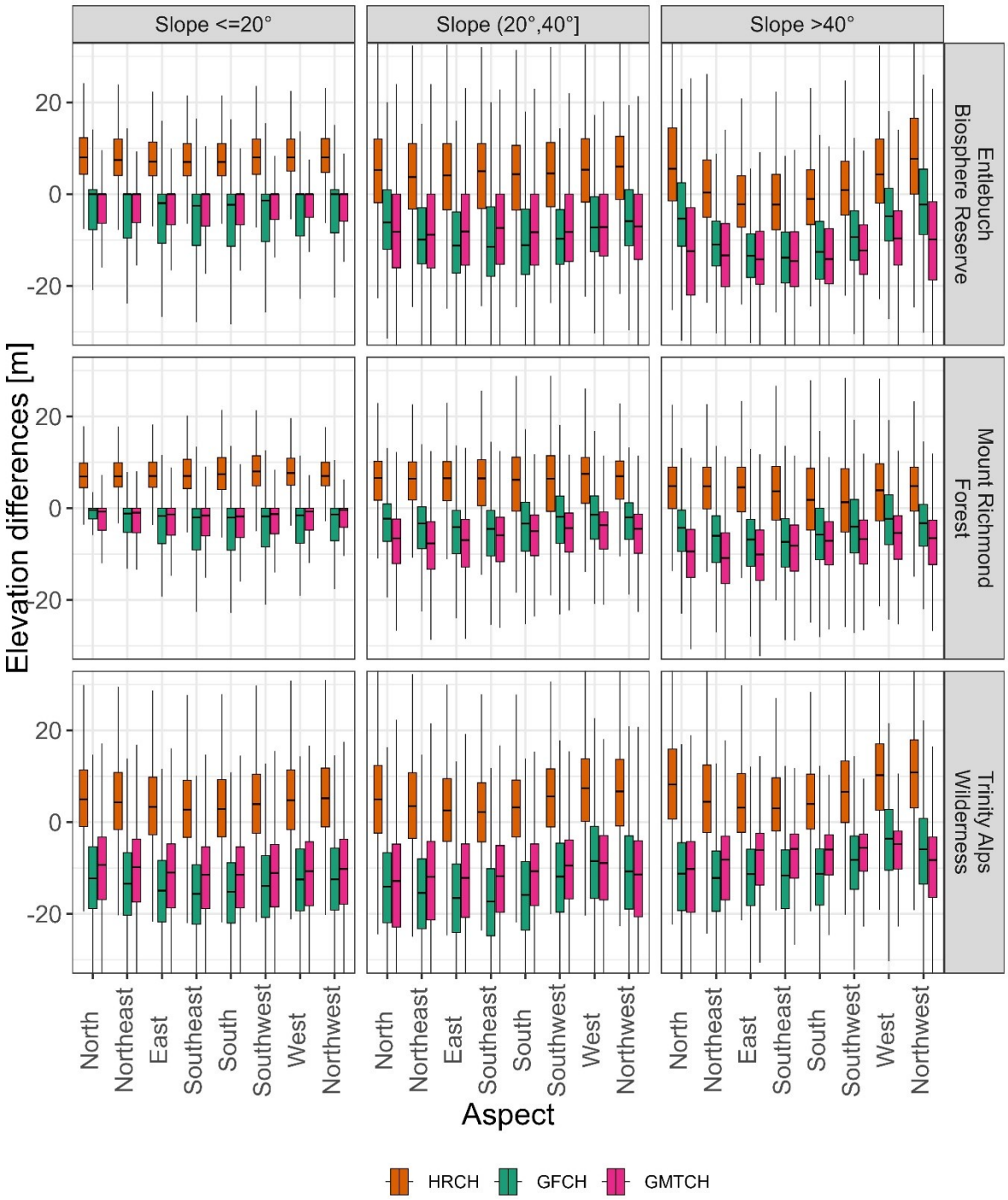


Figure S1. Boxplots showing the accuracy of the modeled global canopy height maps considering the terrain slope and orientation (aspect).

Estimates of High-Resolution Canopy Height Model of the Earth uncertainty

Lang et al. (2023) provide a map that estimates the predictive uncertainty for every cell of their High-Resolution Canopy Height Model of the Earth (HRCH). We investigated the relationship between the absolute height differences (HRCH minus LiDAR CHM) and the theoretical error represented by the

uncertainty estimates to evaluate the usability of theoretical error for filtering out inaccurate predictions (Lang et al. 2023).

The theoretical height error of HRCH, represented as the estimated predictive uncertainty, exhibits a unimodal distribution with a peak around 9 meters in all study areas. Most of the values fall within the range of approximately 6 to 14 meters (Figure S2). In the Entlebuch Biosphere Reserve and Mount Richmond Forest, the estimated predictive uncertainty accurately reflects the median of the absolute canopy height error for most of the cells. The median of the absolute canopy height error for the HRCH map increases linearly with estimated predictive uncertainty. A linear relationship is observed up to the value of approximately eight to fourteen meters, depending on the study area (Figure S2). After this value, uncertainty estimates tend to overestimate the absolute canopy height error (Figure S2). However, in the Trinity Alps Wilderness, the estimated predictive uncertainty is an unreliable estimate of the absolute canopy height error.

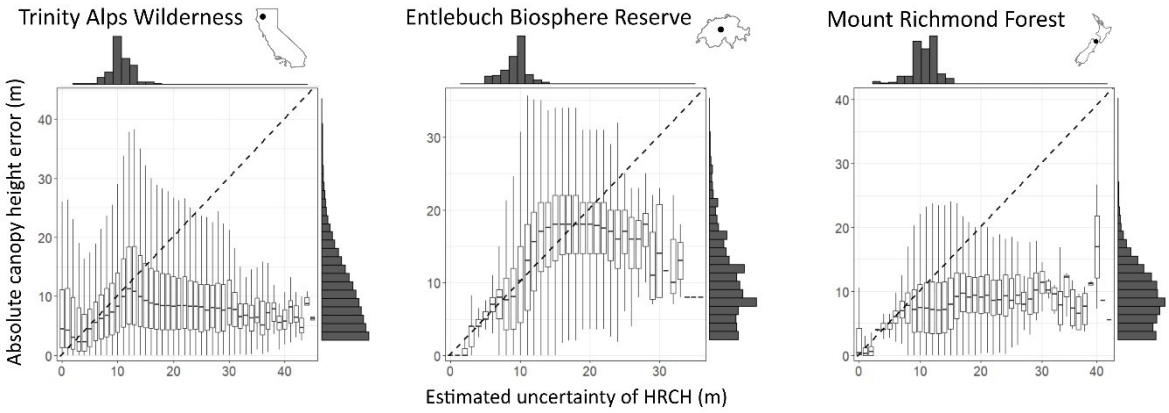


Figure S2. The theoretical height error of the HRCH map represented as the estimated predictive uncertainty versus the measured absolute canopy height error (calculated as a difference between HRCH and ALS CHM). The dashed line indicates a hypothetical perfect agreement of estimated uncertainty and observed vertical errors. Histograms show the distribution of estimated and calculated vertical error.

Calculation of Normalized Difference Vegetation Index (NDVI)

Given the limited accuracy of global CHMs, one of the reviewers suggested testing whether their use in SDMs is justified at all. Therefore, we attempted to replace global CHMs with Normalized Difference Vegetation Index (NDVI) in SDMs. We used Sentinel-2 Level-2A multispectral images to calculate (NDVI). We analyzed surface reflectance images from the Sentinel-2 satellite, covering the vegetation seasons at our study areas from 2019 to the present. For the regions in California and Switzerland, this corresponds to the months from May to September, and for New Zealand, from October to May. The Sentinel-2 platform provides spectral resolution from the visible spectrum to the shortwave infrared, with 4 visible and near-infrared bands and 2 shortwave infrared bands. The spatial resolution of the

images is 10 meters per cell for used visible and near-infrared spectrum. Level-2A processing, including native atmospheric corrections performed by the European Space Agency (ESA). Our data pre-processing included filtering for clouds, shadows, and snow. We utilized the global cloud filtering based on the Sentinel-2 Cloud Probability product with a threshold setting of 25% probability. Additionally, individual cells (pixels) were filtered to exclude cloudy, shaded, and snow-covered pixels based on the Copernicus Sentinel-2 'cloud mask' and 'QA60' quality bands. After filtration and masking, we retained 3,610 images for New Zealand, 1,080 images for California, and 1,722 images for Switzerland. We combined the individual corrected and masked satellite images into 5-year composites (seamless mosaics), based on the median values of overlapping pixels within the vegetation season. The entire remote sensing data processing was conducted using the Google Earth Engine cloud computing platform.

References

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. Journal of Animal Ecology, 77, 802–813.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189-1232.

Greenwell, B., Boehmke, B., Cunningham, J., Developers, G. B. M., & Greenwell, M. B. (2019). Package 'gbm'. R package version, 2(5).

Hijmans, R. J., Phillips, S., Leathwick, J., Elith, J., & Hijmans, M. R. J. (2017). Package 'dismo'. Circles, 9, 1–68.

Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2023). A high-resolution canopy height model of the Earth. Nature Ecology & Evolution, 1-12.

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.

Moudrý, V., Lecours, V., Malavasi, M., Misiuk, B., Gábor, L., Gdulová, K., ... & Wild, J. (2019). Potential pitfalls in rescaling digital terrain model-derived attributes for ecological studies. Ecological Informatics, 54, 100987.

Moudrý, V., Moudrá, L., Barták, V., Bejček, V., Gdulová, K., Hendrychová, M., ... & Šálek, M. (2021). The role of the vegetation structure, primary productivity and senescence derived from airborne LiDAR and hyperspectral data for birds diversity and rarity on a restored site. Landscape and Urban Planning, 210, 104064.

Naghibi, S. A., Khodaei, B., & Hashemi, H. (2022). An integrated InSAR-machine learning approach for ground deformation rate modeling in arid areas. Journal of Hydrology, 608, 127627.

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, 100067.

Valavi, R., Guillera-Arroita, G., Lahoz-Monfort, J. J., & Elith, J. (2022). Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. Ecological Monographs, 92(1), e01486.

Wu, W. B., Yu, Z. W., Ma, J., & Zhao, B. (2022). Quantifying the influence of 2D and 3D urban morphology on the thermal environment across climatic zones. Landscape and Urban Planning, 226, 104499.

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., Lesiv, M., Herold, M., Tsendbazar, N.-E., Xu, P., Ramoino, F., & Arino, O. (2022). ESA WorldCover 10 m 2021 v200 (Verze v200) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7254221