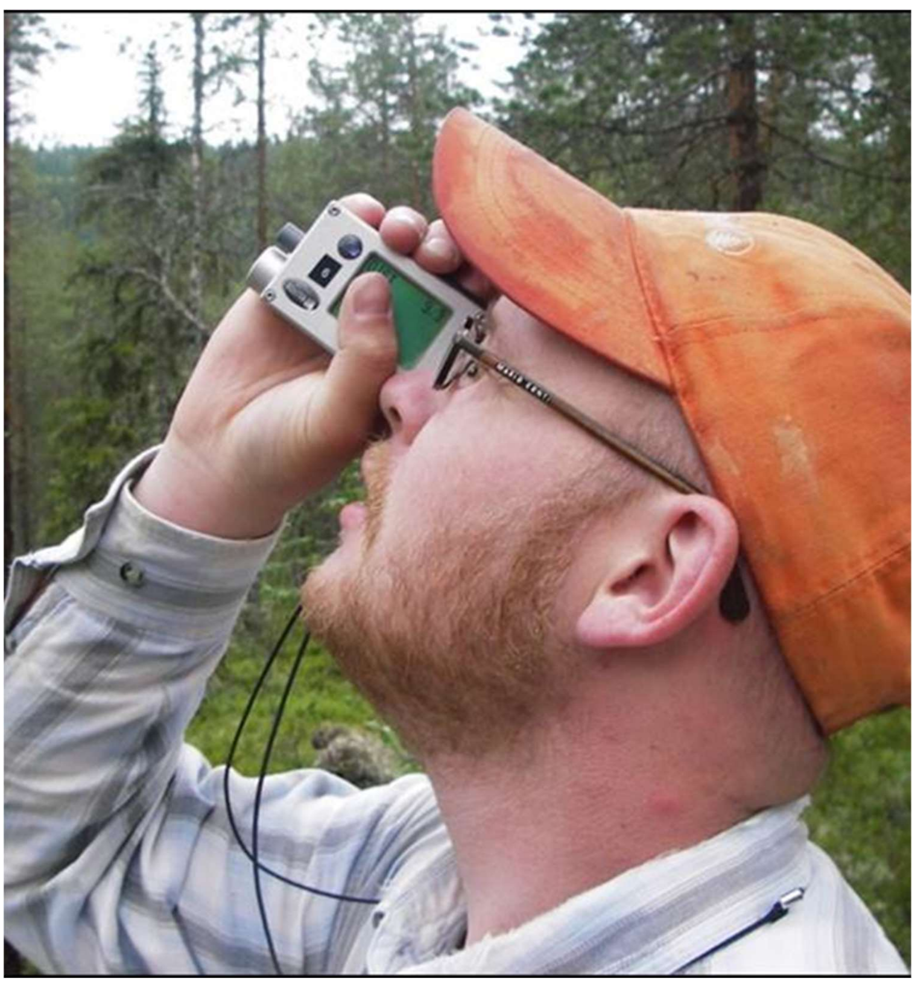
Norges miljø- og biovitenskapelige universitet
Fakultet for miljøvitenskap og naturforvaltning

2023
ISSN 2535-2806

MINA fagrapport 87

Systematic and random errors of height measurements of individual trees using Vertex hypsometer.

Ole Martin Bollandsås
Tron Haakon Eid
Elisabeth Annie Anderheim Hansen



Bollandsås, O.M., Eid, T.H. & Hansen, E.A.A. 2023. **Systematic and random errors of height measurements of individual trees using Vertex hypsometer**. – MINA fagrapport 87. 23 s.

Ås, november 2023

ISSN: 2535-2806

TILGJENGELIGHET
Åpen

PUBLISERINGSTYPE
Digitalt dokument (pdf)

KVALITETSSIKRET AV
Erik Næsset, MINA, NMBU

OPPDRAGSGIVER
Borregaard forskningsfond

FORSIDEBILDE
Endre Hofstad Hansen måler høyden på et tre med Vertex. Foto: Ole Martin Bollandsås

NØKKELORD
Vertex, trehøyde, målefeil

KEY WORDS
Vertex, tree height, measurement error

Ole Martin Bollandsås (olebo@nmbu.no), Tron Haakon Eid, Elisabeth Annie Anderheim Hansen:
Fakultet for miljøvitenskap og naturforvaltning, Norges miljø- og biovitenskapelige universitet, Ås

Summary

The aim of this study was to quantify systematic and random measurement errors of height measurements using Vertex hypsometers. The height measurements of 35 surveyors with varying levels of expertise (inexperienced, experienced and experts) were compared to reference height values of 95 trees, including both pine and spruce trees over a wide range of tree heights. The trees were located in forests with different densities. Generalized linear modelling was employed to examine the potential effects of tree species, tree height, forest density, and surveyor expertise on the magnitude of measurement errors. The results revealed a small, but yet statistically significant systematic error in the height measurements averaging at 0.3% of tree height, which was consistent across all expertise levels. The average random error was 3.75%, but when omitting the measurements of the inexperienced surveyors, the error decreased to 3.1%. If only the measurements of the experts were considered, the random error was 2.9%. Among the analysed attributes, expertise and tree species had the most pronounced impact on the magnitude of measurement errors.

Sammendrag

Formålet med denne studien var å kvantifisere systematiske og tilfeldige feil ved måling av trehøyde med Vertex høydemålere. Høydemålinger utført av 35 personer med ulikt kompetansenivå (uerfaren, erfaren, ekspert) ble sammenlignet med fasithøyder for 95 gran- og furutrær over et bredt intervall av trehøyder. Trærne var lokalisert i skog med varierende tetthet. Generaliserte lineære modeller ble brukt for å undersøke effekter av faktorer som treslag, trehøyder, tetthet og personenes kompetansenivå på størrelsen til målefeilene. Resultatene viste at det var knyttet en liten, men likevel statistisk signifikant, systematisk feil til høydemålingene på 0,3% over alle kompetansenivåene. Den gjennomsnittlige tilfeldige feilen var 3,75 %, men dersom målingene til de uerfarne personene ble utelatt, var feilen 3,1 %. Den tilfeldige feilen var 2,9 % hvis kun målinger utført av ekspertene ble brukt. For de undersøkte faktorene var det kompetansenivå og treslag som hadde den største innvirkningen på størrelsen til målefeilene.

Contents

1. Introduction 5
2. Literature review 7
3. Material and Methods 10
  3.1. Data collection 11
    3.1.1 Selection of reference trees 11
    3.1.2 Reference measurements 11
    3.1.3 Determination of reference height values 11
    3.1.4. Surveyor's measurements 12
  3.2 Calculation of differences and data reduction 12
  3.3 Data analyses 13
  3.3.1 Means and standard deviations for data subsets 13
    3.3.2 Visualization and modelling the impact of data subsets on measurement errors 14
4. Results 15
  4.1 Means and standard deviations for data subsets 15
  4.2 Visualization and modelling the impact of data subsets on measurement errors. 15
5. Discussion 19
6. Conclusions 21
References 22

1. Introduction

In forestry, effective management requires decisions concerning various silvicultural activities such as harvests, regeneration and young growth tending. These decisions involve considerations of timing, methods, and intensities of the activities, all of which have a spatial and temporal impact on subsequent decisions. Managing forests involves tackling complex optimization problems that come with multiple constraints, including the preservation of biodiversity, cultural remains, erosion control, and recreational aspects. Therefore, rational forest management decisions rely on the availability of appropriate forest information, enabling managers to achieve their goals while considering these constraints.

To accomplish this, it is crucial to have access information on current state and data on growth potential. In many cases also historical records are relevant. For instance, knowledge about the timing and intensity of previous young growth tending is relevant when determining the optimal timing and intensity of thinning operations. However, different data sources carry varying levels of uncertainties, and it is essential to estimate and understand these uncertainties to assess the reliability of the information and determine its usability for decision-making. Additionally, understanding the uncertainty associated with a data source allows for the estimation of accumulated uncertainty in estimates that rely on multiple sources of information.

Forest information is typically acquired through purposefully designed forest inventories. When the objective is to gather data for large areas like regions or entire nations, national forest inventories (NFIs) are employed. In Norway, the NFI is conducted using sample plots distributed systematically on a 3x3 km grid, and these plots are measured every five years. The data obtained from the NFI serve as a foundation for forest policy, international carbon reporting, and are also used as reference data for calibrating models reliant on remotely sensed data. On the other hand, when the goal is to obtain information at the forest stand level and effectively manage a specific forest property, forest management inventories (FMIs) are relevant. In the past two decades, FMIs have been carried out using a combination of field plots and data obtained from airborne laser scanners. Field plots provide the necessary data to develop models that establish relationships between field-measured variables such as timber volume, stand basal area, mean height, and various metrics derived from the point clouds captured by airborne laser scanners. These models are subsequently used to predict different forest attributes over the entire area in question.

In both NFIs and FMIs, single-tree measurements are fundamental. Diameters at breast height (dbh) are measured for all trees above a minimum dbh, and heights (h) are measured on sample trees. These height measurements enable the prediction of volume using allometric volume models (Brantseg, 1967; Braastad, 1966; Hansen et al., 2023; Vestjordet, 1967) or biomass using allometric biomass models (Marklund, 1988; Smith, Granhus, & Astrup, 2016; Smith et al., 2014). The predictions of single-tree volumes are then accumulated to obtain volume estimates for the respective field plots where the trees are located.

The procedures employed to calculate plot-level estimates may vary, depending on factors such as the chosen sample tree selection strategy. For instance, if there is an ample number of sample trees with measured heights available on each plot, plot-wise regression models can be constructed to predict height based on dbh, potentially including separate models for different tree species. However, if the number of sample trees are insufficient for this approach, another strategy could involve combining sample trees from multiple plots within strata defined by, for example, dominant tree species, development class, and productivity classes to construct the models. In either case, the constructed models are applied to every tree on the plot or within a stratum to predict tree heights. These predicted heights, along with the measured dbhs, are then utilized as input in the volume models to obtain predicted volumes.

An alternative strategy often employed when the number sample trees available on each plot is small is to calculate a ratio between the volume predicted using the measured height and the base volume predicted using a base height curve (Fitje & Vestjordet, 1977). To obtain a volume prediction for each tree on a plot, the base volume for every tree (requiring only dbh as a measured input) is multiplied by the mean relationship between the "true" volume and the base volume for the sample trees.

The accuracy and precision of estimating plot volume depend on the quality of input measurements, regardless of the chosen method or strategy. To comprehensively account for all sources of error and accurately estimate the propagated uncertainty (Fortin & DeBlois, 2010; Sexton et al., 2015; Vorster et al., 2020), it is essential to consider the uncertainty associated with each input used in the estimation. In the past, it has been common to only account for sampling variability and assume that field measurements are error-free when estimating uncertainty in purely field-based inventories. Furthermore, in inventories that rely on field plot values for calibration of models based on remotely sensed data, it is also important to recognize the uncertainty in the estimated plot values during model construction. The assumption of error-free response variables is prevalent in models developed using the ordinary least squares method, as well as in important models used in forest management such as site index models, growth models, and volume- and biomass models. However, it is important to acknowledge that height measurements, even in these contexts, are prone to errors. To understand the impact of such errors is also of interest (Kangas, 1996).

Despite several studies quantifying the size of height measurement errors, these studies are relatively limited in number and each of them addresses only a portion of the factors that influence error size. As such, there is a growing imperative to compile and expand upon existing knowledge to gain a more comprehensive understanding of the sources and magnitudes of errors affecting height measurements. Such an endeavour is essential to accurately estimate the total accumulated uncertainty in volume and biomass estimates, which plays a critical role in enhancing the reliability of forest management and decision-making processes. With this context in mind, the objectives of this research were twofold:

1. To conduct a comprehensive review of important scientific studies that have quantified height measurement errors.
2. To perform an independent study focused on quantifying the errors associated with field-based height measurements using Vertex IV hypsometers in Norway. This study encompassed Norway spruce and Scots pine trees and explored various biophysical tree and stand properties that might impact measurement errors. Additionally, it involved 35 surveyors with varying expertise levels, and their potential effects on error magnitudes were analysed. Reference tree heights were determined using a total station, and both systematic and random errors were reported in both absolute and relative terms.

2. Literature review

Two previous studies conducted by Fitje (1967) and Hansen (2021) stands as the only Norwegian investigations that quantified systematic and random errors in height measurements for individual trees. Fitje (1967) involved 61 forestry students from the Agricultural University of Norway (now NMBU) and encompassed measurements taken from 10 Norway spruce trees. Various types of instruments where tree heights are derived from measurements of angles and distances were tested, including the Suunto hypsometer which demonstrated the smallest levels of systematic and random errors in that particular study. As a result, the Suunto hypsometer became the standard instrument employed for field-based forest inventories in Norway for the next 25 years, until the introduction of the Vertex hypsometer. The second study, conducted by (Hansen, 2021), incorporated Vertex measurements taken by students and employees at NMBU (in total 12 surveyors) and included a larger number of reference trees compared to Fitje's study. Hansen's study encompassed 71 trees from different species (spruce and pine), varying tree heights, forest densities, and expertise levels. However, there were no professional forest surveyors among the participants. It was found that individuals with more experience obtained a random measurement error of approximately 5%. Nevertheless, after the study was completed, it was discovered that a few of the reference tree heights were likely determined erroneously, suggesting that the random error found in this study may have been overestimated. Importantly, the dataset from the Hansen study is incorporated into the current study (as described in the method section). In addition to the two studies described above, which focused on single trees, the determination of Lorey's mean height (HL) have also been investigated Nersten & Næsset (1992). In their study, 25 surveyors selected sample trees at 10 different locations in mature forests and determined HL. They found no significant systematic errors, and the random error was on average 7.3%.

The Vertex instruments, depending on the instrument model, employ either ultrasound reflected from a transponder attached to the tree or laser light reflected directly from the stem to measure distance. In Norwegian forests, the instrument model utilizing ultrasound has been more commonly employed due to the potential challenges posed by understory vegetation, which can make it difficult to find a suitable location from which to measure the tree where both the treetop and stem base are visible. The Vertex instrument is regarded as more precise and efficient compared to the Suunto due to several key factors. One advantage is that users are not required to be positioned at a specific distance from the tree. Additionally, the Vertex instrument automatically calculates the horizontal distance and angle to the transponder, which is placed at a user-defined height above the stump level. Another significant advantage of the ultrasound-based instrument is its capability to measure distances even when there is vegetation blocking a clear line of sight to the transponder.

While the recent study conducted by Hansen (2021) focused on examining the accuracy and precision of height measurements using Vertex hypsometers, and that it provides valuable insights at the national level, it is important to consider the broader body of research that has emerged internationally over the past 10 years. Numerous studies have quantified errors associated with the use of Vertex hypsometers as well as other comparable brands (Ganz et al., 2019; Holmgren, 2019; Jurjević et al., 2020; Kitahara et al., 2010; Krause et al., 2019; Luoma et al., 2017; Paudel et al., 2021; Stereńczak et al., 2019; Vasilescu, 2013; Villasante & Fernandez, 2014; Wang et al., 2019). It is important to note that while these studies all utilize the Vertex hypsometer or similar instruments and evaluate the errors associated with their use, they employ diverse approaches and study designs. As a result, direct comparisons between the studies can be challenging.

Several studies have focused exclusively on the Vertex hypsometer (Kitahara et al., 2010; Luoma et al., 2017; Paudel et al., 2021; Vasilescu, 2013), while others have included comparisons with other field-based hypsometers in addition to the Vertex (Stereńczak et al., 2019; Villasante & Fernandez, 2014). Some studies have taken a different approach by focussing on comparing measurements from Vertex hypsometer with remote sensing methods such as airborne laser scanning (ALS), terrestrial laser scanning (TLS), and/or photogrammetry (Ganz et al., 2019; Holmgren, 2019; Jurjević et al., 2020; Krause et al., 2019; Wang et al., 2019). The various approaches employed in these studies have led to differences in the number of trees measured and the number of individuals involved, as well as their experience with field-based height measurements using the Vertex hypsometer. The number of trees measured ranged from three (Vasilescu, 2013) to 1,174 (Wang et al., 2019), while the number of individuals involved varied from one (Ganz et al., 2019; Holmgren, 2019; Krause et al., 2019) to 104 (Vasilescu, 2013). Information regarding the level of experience of the individuals involved is scarce, except for studies focused on training programs for students (Kitahara et al., 2010; Paudel et al., 2021; Vasilescu, 2013). Some studies claim that the operators of the Vertex hypsometer were experienced (Krause et al., 2019; Luoma et al., 2017), but detailed information is often lacking. It is important to consider that the substantial variations in the number of trees measured, the number of individuals involved, and their skills can impact the generalizability of the results concerning error levels for practical field-based height measurements using the Vertex hypsometer.

Field-based height measurements and the associated error levels are influenced by various biophysical tree and stand properties, as well as topographic characteristics. Stereńczak et al. (2019) identified several challenges related to these properties:

- Tree species: Different tree species exhibit varying degrees of crown asymmetry and irregular shapes, leading to differing levels of difficulty in defining the treetop.
- Measurement accuracy is influenced by the height of the tree.
- Tree lean: Ensuring that the treetop is precisely located directly above the base for distance measurement can be challenging when trees are leaning.
- Forest structure: Dense or multi-layered stands pose challenges in identifying treetops due to limited visibility and a scarcity of suitable measurement locations.
- Topography: Large variations in terrain elevation, particularly when measuring from a lower elevation than the tree base, can amplify errors, especially on sloping terrain.

Stereńczak et al. (2019) also highlighted instrument errors, such as limitations dependent on precipitation and temperature, as well as the need for regular and appropriate calibration before use. Human errors, including planning failures and crew performance issues (e.g., haste, lack of attention to details, inexperience, carelessness, shaking with handheld instruments, subjectivity, and perception issues), were mentioned as additional sources of error in height measurements.

In a comprehensive study conducted in Poland, Stereńczak et al. (2019) examined the effects of various properties on height measurements, including tree species, tree age, stand structure (single-layer, two-layer, or multi-layer), stand density (canopy closure), height above sea level, and terrain slope. They also investigated the effects of different instruments (including three different hypsometers in addition to the Vertex) and five different teams. Most of the other studies reviewed focused on a limited number of properties, mainly related to individual trees (e.g., tree species, tree size, and tree crown) (Jurjević et al., 2020; Luoma et al., 2017; Vasilescu, 2013).

Determining accurate height reference values for trees in studies that quantify systematic and random errors in height measurements can be challenging. Different approaches have been used in the studies mentioned above. Some studies employed cross-checking methods, comparing different measuring techniques in pairs to establish reference values (Holmgren, 2019; Jurjević et al., 2020; Wang et al., 2019). Others used the average of all field-based measurements as reference values (Luoma et al., 2017; Vasilescu, 2013) or relied on measurements conducted by experienced surveyors or control teams (Kitahara et al., 2010; Paudel et al., 2021). Felling trees and measuring their length with a measuring tape is considered one of the most reliable methods for determining reference height values (Ganz et al., 2019; Krause et al., 2019; Stereńczak et al., 2019). However, (Stereńczak et al., 2019) highlight that even this method may introduce systematic errors due to differences between tree height and tree length. Tree height typically refers to the vertical distance from the ground to the apex of the tree (Avery & Burkhart, 2015; Kershaw et al., 2016), while tree length refers to the distance measured along the stem. Tree lean is accounted for in the definition of height, but not the straightness of the stem. However, the extent of this discrepancy varies among tree species and growing conditions, but it is generally small. It is also important to note that in some cases, such as in Norway, the definition of height may be based on the stump height rather than the ground. Nevertheless, it is essential to acknowledge that tree height is an ambiguous term (Kershaw et al., 2016) and to clarify the specific definition applied. When studying the magnitude of height measurement errors, it is crucial that the definition used during the reference measurements matches the one applied by the surveyors.

An alternative non-destructive method for determining accurate tree reference heights is the use of a total station. This approach was employed in the study conducted by Andersen, Reutebuch and McGaughey (2006), which focused on quantifying errors in height measurements. They utilized a Topcon ITS-1 total station and reported a high level of accuracy in individual tree height measurements, with an error of less than 2 cm. While very few studies have utilized total stations for determining true tree heights, there is an example from Malaysia where this method was used to validate a new tree height estimation technique for oil palm trees based on unmanned aerial vehicles (UAVs) and photogrammetry (Ramli & Tahar, 2020).

Comparing the above-mentioned studies is further complicated by the variability in how the results were reported. Different options exist for reporting errors, including systematic errors in absolute and/or relative terms, random errors in absolute and/or relative terms, and total errors in absolute and/or relative terms (Larjavaara & Muller-Landau, 2013). Among the studies mentioned, only Krause et al. (2019) encompassed all of these options, reporting systematic, random, and total errors in both absolute and relative terms. In contrast, other studies reported only systematic errors (Paudel et al., 2021; Stereńczak et al., 2019) or only random errors (Luoma et al., 2017).

3. Material and Methods

The first acquisition took place during the spring and summer of 2020 on the forest property of the Norwegian University of Life Sciences (NMBU) in Ås, Viken county. The measurements were conducted by students and employees affiliated with NMBU. The second acquisition occurred in early May 2022 in a private forest near Næroset, Innlandet County, and was carried out by employees from the Norwegian National Forest Inventory. Both locations are indicated on the map in Figure 1.

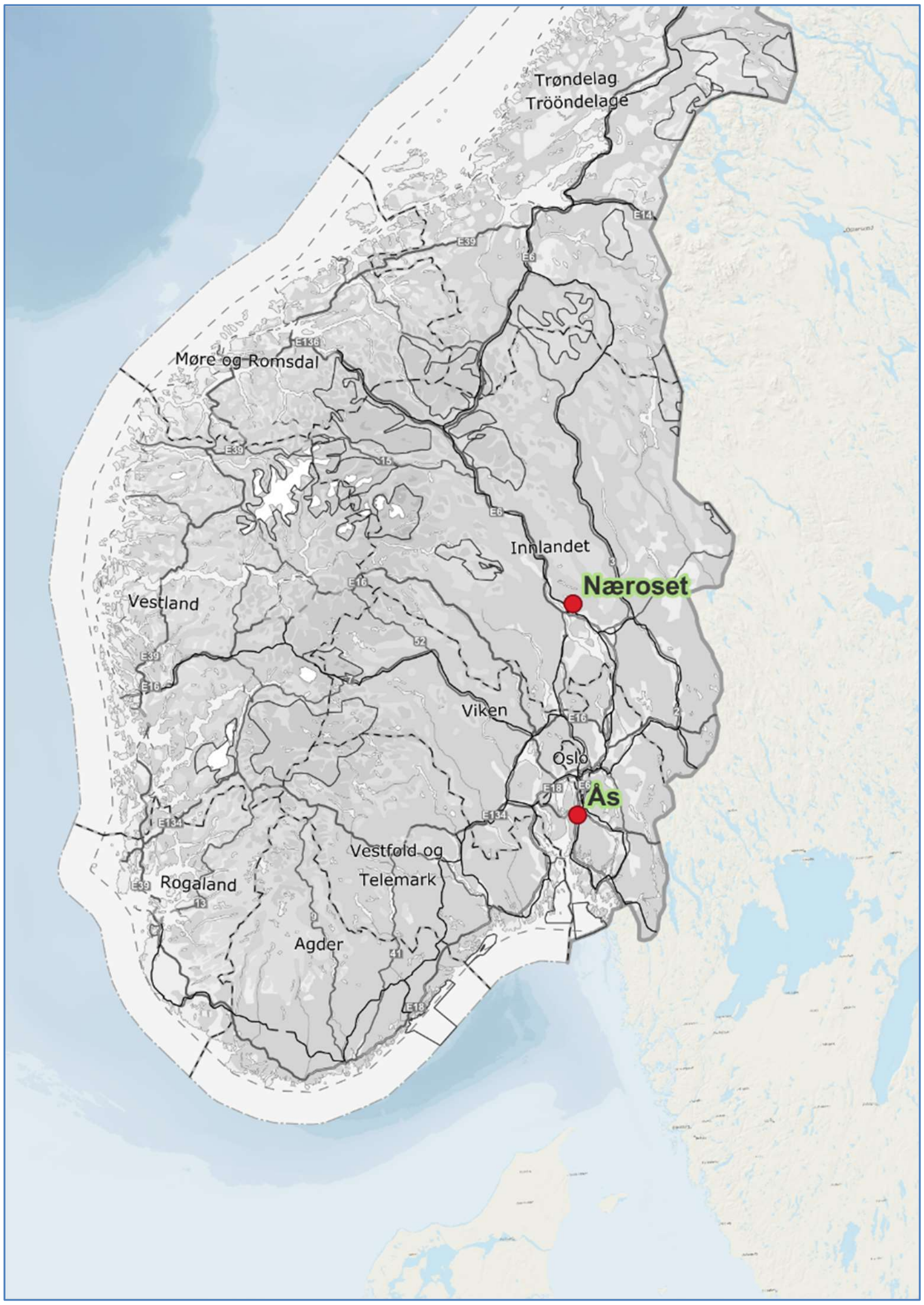


Figure 1. Map showing the locations where the data acquisitions were carried out: Ås and Næroset.

3.1. Data collection

3.1.1 Selection of reference trees

Trees were selected purposefully in both data acquisitions to ensure comprehensive representation of trees across different subsets of the final dataset (Table 1). Initially, our selection encompassed both Norway spruce and Scots pine species. Additionally, trees growing in diverse forest densities, as measured by basal area (BA), and covering a range of heights, were chosen. A relascope was used to estimate forest densities within the stands where the selected trees were located. Each tree was attributed to one of four basal area (BA) subsets: BA10 ($BA < 10\ m^2ha^{-1}$), BA15 ($10\ m^2ha^{-1} < BA < 20\ m^2ha^{-1}$), BA25 ($20\ m^2ha^{-1} < BA < 30\ m^2ha^{-1}$), or BA30 ($BA > 30\ m^2ha^{-1}$). Similarly, the trees were categorized based on their height and assigned to one of three height subsets: H15 ($h < 15\ m$), H20 ($15 < h < 25\ m$), and H25 ($h > 25\ m$). Once the trees were selected, they were individually identified by attaching ID labels, and reference heights were measured for each of them. In the 2020 and 2022 acquisitions, a total of 73 and 30 trees were chosen, respectively.

3.1.2 Reference measurements

Reference heights were measured using a Topcon total station. The total station was operated from where there were unobstructed sight lines to both the tree trunk and the treetop. To avoid measuring at steep angles, the distance between the tree and the total station was always equal or greater than the height of the tree. The total station was levelled by means of the inbuilt digital level. A paper target was placed on the tree trunk and the distance between stump height and the target ($t.height$) was measured using a surveyor's tape. After aiming at the target, both the slant distance ($slant.d$) and angle in radians (r) were registered. Then the total station was aimed at the treetop, and the angle in radians (s) was registered. If the tree was leaning, the horizontal offset ($horizontal.o$) between the treetop and the stump was visually located and measured with a surveyor's tape. In such cases, the total station was always placed perpendicular to the slanting direction.

3.1.3 Determination of reference height values

The horizontal distance ($horizontal.d$) between the total station and the tree was first calculated as

$$horizontal.d = \cos(r) \times \left(slant.d + \frac{dia}{2}\right) \tag{1}$$

where dia represents the diameter of the tree where the paper target was attached. Then the heights below ($b.height$) and above ($a.height$) the horizontal level from the total station and the paper target and the treetop, respectively, was calculated as

$$b.height = \sin(r) \times \left(slant.d + \frac{dia}{2}\right) \tag{2}$$

$$a.height = \tan(s) \times horizontal.d \tag{3}$$

and then reference tree height ($r.height$) was obtained as

$$r.height = t.height + b.height + a.height. \tag{4}$$

If a reference tree was leaning, a corrected reference height ($cr.height$) was calculated as

$$cr.height = \sqrt{(r.height^2 + horizontal.o^2)}. \tag{5}$$

3.1.4. Surveyor's measurements

In the 2020 and 2022 acquisitions, a combined total of 12 and 23 surveyors, respectively, performed individual measurements. The surveyors varied in their level of expertise, ranging from "inexperienced" students who had received basic instructions on instrument usage, to "experts" with several years of experience from the Norwegian NFI. Additionally, there was a category of surveyors labelled "experienced," consisting of employees at NMBU with prior experience from field-based forest inventories.

Each surveyor contributed a single height measurement for the reference trees they individually measured, utilizing a Vertex IV instrument. It's worth noting that none of the surveyors participated in both data acquisitions. Prior to conducting the measurements, the surveyors followed standard procedures for instrument calibration and setup. Additionally, each surveyor made independent decisions on the specific location from which a particular tree was to be measured. In total, 1565 individual measurements were carried out.

3.2 Calculation of differences and data reduction

To quantifying the errors associated with field-based height measurements, a calculation method involving four variables was employed. Firstly, for each individual measurement we calculated the deviance or the error (*d*, eq. 6) between the measured value ($\hat{h}$) and its corresponding reference height (*h*). To gain further insights into the significance of the height measurement errors, we also computed three additional variables: percentual error (*d%*, eq. 7), absolute error ($|d|$, eq. 8), and percentual absolute error ($|d|$%, eq. 9).

Analyses using the percentual error (*d%*) allows to study the relative impact of height on the magnitude of the errors. By expressing the errors in terms of the reference height, we can also better understand the underlying patterns or trends in the data, independent of the specific height values. The absolute deviances ($|d|$) are particularly useful in assessing random variations, as they can be further analysed through the mean absolute error. This metric provides a measure of the average dispersion of the data points around the reference height. Lastly, the percentual absolute error ($|d|$%) combines both the magnitude of the errors and their relative impact, providing a comprehensive perspective on the data. By employing these calculations, we aimed to comprehensively analyse the measurement data, understanding both the relative and absolute aspects of the errors. This approach allowed us to remove the potential influence of height on the differences and explore patterns of variation more effectively.

$$d = \hat{h} - h \qquad (6)$$

$$d\% = \frac{d}{y} \times 100 \qquad (7)$$

$$|d| = \left|\hat{h} - h\right| \qquad (8)$$

$$|d|\% = \frac{|d|}{\bar{y}} \times 100 \qquad (9)$$

After calculating the errors, it became evident during a preliminary screening of the data that the reference height of seven of the reference trees had been erroneously determined. All surveyors' measurements consistently indicated a substantially greater or smaller height for these trees, suggesting that the reference height measurement may have inadvertently considered a treetop other than the intended reference tree. Furthermore, during the preliminary screening, it was observed that a few measurements displayed extreme differences from the reference height, reaching as much as 13 m. Such discrepancies often occur when the surveyor is unaware that a proper distance measurement was not obtained using the Vertex instrument, leading to the instrument's default distance pre-set being used instead. To maintain the focus of the current research and exclude potential errors resulting from improper distance measurements, it was decided to omit all observations where *d* exceeded 5 m. Consequently, our dataset was reduced to data from 95 reference trees and 1459 individual measurements. Table 1 displays a data summary of reference values for the 95 remaining trees distributed among the different data subsets.

Table 1. Data summary. Number of trees (no), mean height (mean), minimum height (min) and maximum height (max) of the sample trees distributed on different data subsets.

Data subset	no	Mean (m)	Min (m)	Max (m)
All	95	19.4	6.5	34.2
Study area 1 (Ås)	65	20.2	6.5	34.2
Study area 2 (Næroset)	30	17.5	9.1	24.8
Norway spruce	44	19.6	6.5	34.2
Scots pine	51	19.2	8.2	33.2
H15[a]	33	11.6	6.5	14.9
H20[a]	41	21.0	15.3	24.9
H25[a]	21	28.3	25.0	34.2
BA10[b]	28	20.4	9.4	29.7
BA15[b]	27	17.9	7.4	33.2
BA25[b]	24	19.5	9.1	34.2
BA30[b]	16	19.8	6.5	30.1

[a]: H15 = h < 15 m, H20 = 15 m < h < 25 m, H25 = h > 25 m

[b]: BA10 = BA < 10 m^2ha^{-1}, BA15 = 10 m^2ha^{-1} < BA < 20 m^2ha^{-1}, BA25 = 20 m^2ha^{-1} < BA < 30 m^2ha^{-1}, BA30 = BA > 30 m^2ha^{-1}.

3.3 Data analyses

3.3.1 Means and standard deviations for data subsets

Within each data subset, we estimated the mean error (*ME*, eq. 10) and the percentual mean error (*ME%*, eq. 11), along with their respective standard deviations (*stdE*, eq. 12) and percentual standard deviations of the differences (*stdE%*, eq. 13). These measures of central tendency and dispersion provided valuable insights into the accuracy (*ME* and *ME%*) and precision (*stdE* and *stdE%*) of the data within each subset. The results appear in Table 2 and provide concrete results for the size of both systematic and random errors of height measurements using Vertex hypsometers. To investigate whether the random variation between the different data subsets differed, we specifically utilized the mean absolute percentual error ($|ME|\%$, eq. 14). This metric was employed as it accounts for the spread of the differences and does not necessarily average to zero, even in the absence of systematic errors.

$$ME = \frac{1}{n}\sum d \tag{10}$$

$$ME\% = \frac{1}{n}\sum d\% \tag{11}$$

$$stdE = \sqrt{\frac{1}{n}\sum d^2} \tag{12}$$

$$stdE\% = \frac{stdE}{\bar{Y}} \times 100 \tag{13}$$

$$|ME|\% = \frac{1}{n}\sum |d|\% \tag{14}$$

3.3.2 Visualization and modelling the impact of data subsets on measurement errors

To quantify and model the differences between various data subsets, our analyses specifically focused on the relative response variables, *d%* (eq. 7) and $|d|$% (eq. 9). This decision had a geometric and congruent reasoning. The accuracy of tree height measurements, determined using a hypsometer, depends on precise measurements of distances and angles. Given that these measurements are taken from a point roughly the length of one tree away, the absolute height error increases proportionally with increasing tree height. We confirmed this expectation through a preliminary analysis, where we modelled the absolute error, $|d|$ (eq. 8) as a function of tree height. The results revealed that tree height was a statistically significant predictor of the absolute error. Therefore, in subsequent analyses that aimed to investigate potential variations in the magnitude of systematic and random errors across different data subsets, we focused on examining the relative errors.

Boxplots were utilized to visually represent the distributions of *d%* for each subset of the data. To assess the combined impact of various factors on *d%* and $|d|$%, generalized linear models (GLMs) were employed. In the analysis, tree height was treated as a continuous variable, while tree species, basal area, and experience were treated as categorical factors. The GLM modelling process was conducted using the glm-function in R (R Core Team, 2020). The response variable *d*%, which exhibited a normal distribution, was modelled using a Gaussian link function. However, the response variable $|d|$% being right-skewed, required a different approach. To account for the skewness, a gamma link function was employed after adding a small constant to the response variable to ensure non-zero values. By utilizing the appropriate link functions, the GLM models were able to capture the relationships between the predictor variables and the response variables, accounting for their specific distributional characteristics.

4. Results

4.1 Means and standard deviations for data subsets

Table 2 presents the mean differences between the surveyors' measurements and reference heights in both the measurement unit (m) and percentual values, along with their corresponding standard deviations. The results displayed in the table reveal that the systematic errors of the measurements from the reference heights were relatively small, yet statistically significant. On average, there was a 0.07 m systematic error which corresponds to 0.32% relative to the reference heights. The corresponding average random errors were 0.65 m and 3.75%, respectively. Notably, the pine trees exhibited both larger systematic and random errors compared to spruce trees. When analyzing the data across different height classes, it became apparent that the systematic errors, both in meters and in relative terms, were similar. While the random error in meters increased with increasing height, its percentual equivalent showed the opposite trend. When examining the various forest density subsets, no discernible trend emerged regarding systematic errors, and the percentual random errors appeared to be relatively consistent across the subsets. It is worth mentioning that the inexperienced surveyors demonstrated both the largest systematic errors and random errors. While the systematic errors were the same for the experienced and the experts, the experts displayed the smallest random measurement errors at 2.90% followed by the experienced surveyors at 3.11%. These numbers were substantially smaller compared to the 5.26% obtained for the inexperienced surveyors.

Table 2. Mean height measurement error and corresponding standard deviation distributed over different data subsets.

Data subset	no	Mean error		Standard deviations	
		ME (m)	*ME%*	*stdE* (m)	*stdE%*
All	1459	0.07***	0.32***	0.65	3.75
Norway spruce	633	0.02	-0.02	0.56	3.11
Scots pine	826	0.11***	0.58***	0.71	4.16
H15[a]	515	0.06**	0.38**	0.51	4.63
H20[a]	695	0.06*	0.26*	0.67	3.26
H25[a]	249	0.10	0.37	0.85	2.94
BA10[b]	442	0.01	-0.06	0.74	4.48
BA15[b]	430	0.12***	0.78***	0.62	3.64
BA25[b]	396	0.05	0.24	0.47	2.40
BA30[b]	191	0.13*	0.32*	0.80	4.31
Inexperienced	381	0.21***	0.79***	0.87	5.26
Experienced	388	-0.08**	-0.61**	0.59	3.11
Expert	690	0.08***	0.59***	0.52	2.90

[a]: H15 = h < 15 m, H20 = 15 m < h < 25 m, H25 = h > 25 m

[b]: BA10 = BA < 10 m^2ha^{-1}, BA15 = 10 m^2ha^{-1} < BA < 20 m^2ha^{-1}, BA25 = 20 m^2ha^{-1} < BA < 30 m^2ha^{-1}, BA30 = BA > 30 m^2ha^{-1}.

* = 0.01 < p < 0.01, ** = 0.01 < p < 0.001, *** = p < 0.001

4.2 Visualization and modelling the impact of data subsets on measurement errors.

The distributions of *d*% for the different data subsets are displayed in Figure 2. For clarity, it is important to note that the whiskers of the boxplots extend only as far as 1.5 times the distance between the first and third quantiles, and the extreme values of the distributions are therefore displayed as numbers above and below the whiskers.

Regarding tree species (Figure 2, upper left), the graphical display suggests that both the systematic and random errors were larger for pine compared to spruce. These visual indications were confirmed by the significant model parameters for tree species in both the *d*%-model (Table 3) and the $|d|$%-model (Table 4).

For tree height (Figure 2, upper right), the boxplots suggest minor differences in the median percentual errors (*d%*) between subsets. It appears that the median error slightly decreased with increasing height, as depicted in Figure 2. The glm-modelling (Table 3), however, gave a somewhat contradictory result, and indicated a minor yet significant trend towards that the percentual errors increased with increasing tree height. On the other hand, the random percentual errors appeared to decrease as tree height increased. This observation is consistent when comparing the percentual standard deviation for different subsets (Table 2), the graphical representation in Figure 2 (upper right), and the glm-model for $|d|\%$ (Table 4).

Regarding forest density (Figure 2, lower left), represented by different subsets according to basal area, there were differences in the distribution of the percentual errors (*d*%) between some of the subsets. However, there was no clear trend suggesting that denser forests would yield larger systematic errors. The graphical display and the mean percentual errors in Table 2 indicate that subset BA15 had somewhat larger systematic error compared to the other subsets, and this finding was partially confirmed by the glm-modelling (Table 3) where the parameter estimate attributed to BA15 was associated with the highest significance level when comparing its mean error to the base subset in the model (subset BA10 in this case). There were no apparent trends in the random variation between subsets indicated by either analysis.

The analysis of the fourth factor, expertise, revealed that the measurements taken by inexperienced surveyors were more variable compared to those of the other surveyors. This is evident from the percentual standard deviations between the subsets in Table 2, which decrease with a higher level of experience. The results of the glm-modelling (Table 4) similarly show that the percentual absolute errors decreased with experience, as indicated by the significant parameter estimates comparing the difference from the base subset in the model (expert level in this case). Although there was a significant difference in the percentual errors (*d%*) between experienced and expert surveyors, the deviations from zero were small (Table 2) and smaller than those of the inexperienced subset.

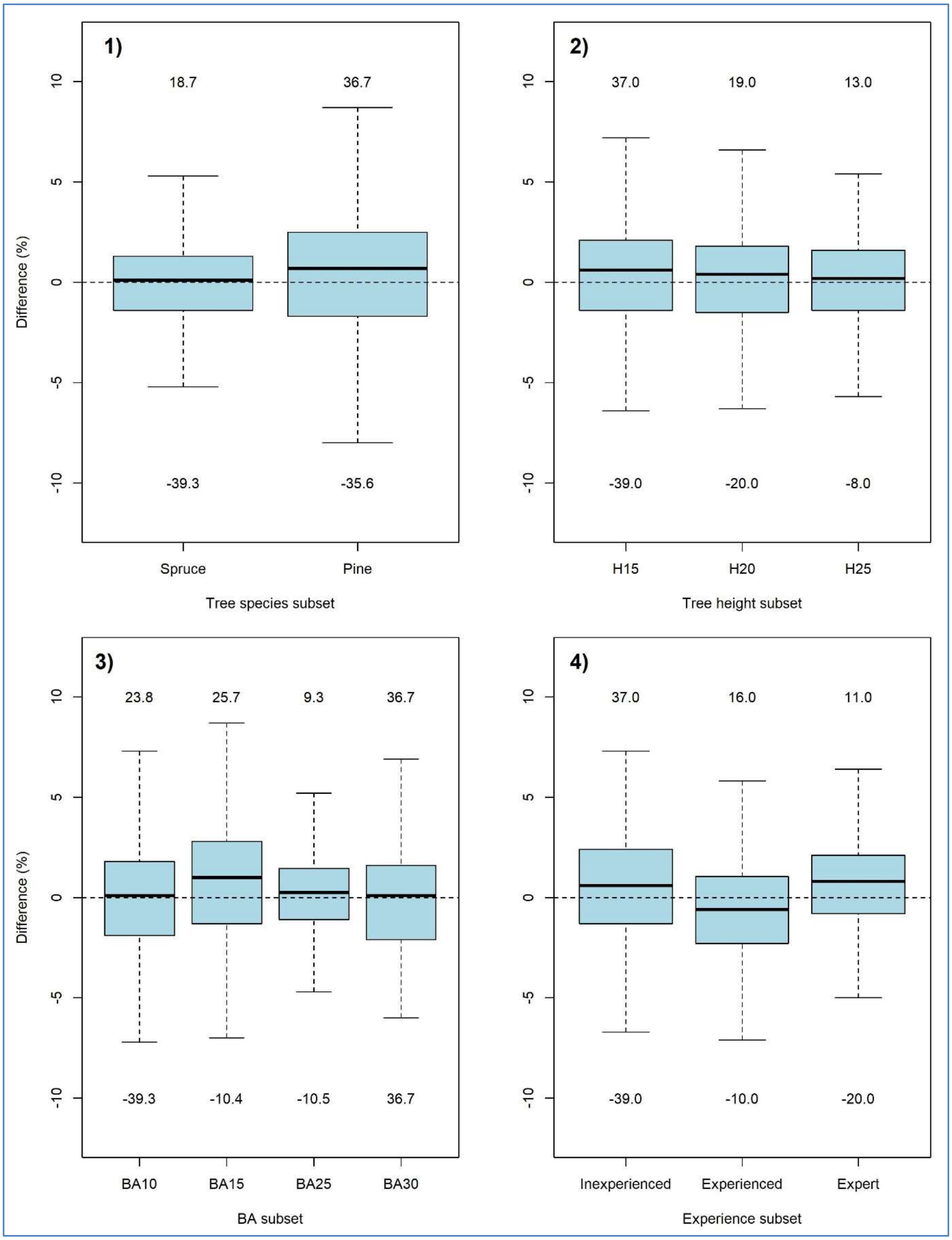


Figure 2. Distribution of percentual height meaurement errors (d%) over data subsets according to 1) tree species (spruce=1, pine=2), 2) tree height classes (H15 = h < 15 m, H20 = 15 m < h < 25 m, H25 = h > 25 m), 3) BA classes (10 = BA < 10 m^2ha^{-1}, 15 = 10 m^2ha^{-1} < BA < 20 m^2ha^{-1}, 25: 20 m^2ha^{-1} < BA < 30 m^2ha^{-1}, 30 = BA > 30 m^2ha^{-1}), and 4) experience (n = no experience, y = experienced, e = expert).

Table 3. Parameter estimates with corresponding standard deviations and significance for model of the percentual difference between measured and reference height.

Variable	Parameter estimate	St.dev	Significance
Intercept	-1.15	0.42	**
factor(Tree species)2	0.63	0.23	**
Tree height	0.04	0.02	**
factor(BA class)15	0.99	0.26	***
factor(BA class)25	0.70	0.29	*
factor(BA class)30	0.91	0.34	**
factor(Experience)n	0.13	0.26	
factor(Experience)y	-1.27	0.25	***

* = 0.01 < p < 0.01, ** = 0.01 < p < 0.001, *** = p < 0.001

Table 4. Parameter estimates with corresponding standard deviations and significance for model of the absolute percentual difference between measured and reference height.

Variable	Parameter estimate	St.dev	Significance
Intercept	0.92	0.11	***
factor(Tree species)2	0.39	0.06	***
Tree height	-0.02	0.01	***
factor(BA class)15	-0.03	0.07	
factor(BA class)25	-0.24	0.08	**
factor(BA class)30	-0.06	0.09	
factor(Experience)n	0.46	0.07	***
factor(Experience)y	0.21	0.07	**

* = 0.01 < p < 0.01, ** = 0.01 < p < 0.001, *** = p < 0.001

5. Discussion

The present research utilized a comprehensive dataset comprising spruce and pine trees, encompassing a range of tree heights, forest densities, and surveyor experience levels. Although the data were not perfectly balanced in terms of the number of reference trees and measurements within each subset, the results presented are grounded in a solid foundation. To model both response variables, *d%* and $|d|\%$, GLM was employed. GLMs are well-suited for handling unbalanced data and do not necessitate an equal number of observations in each subset or group. Moreover, appropriate link functions, Gaussian for the normally distributed d% variable and Gamma for the right-skewed $|d|\%$ variable, were incorporated into the modelling process. By employing GLMs with suitable link functions, we ensured that the statistical analysis accommodated the specific characteristics of the response variables, enabling reliable interpretation and understanding of the observed effects.

The current study demonstrated minimal systematic errors associated with height measurements using Vertex instruments. Across the entire dataset, the average systematic difference between the measured and reference tree height was 0.3% (0.07 m), which was slightly smaller than the findings reported in the studies by Ganz et al. (2019) and Krause et al. (2019), but comparable to or slightly larger than the findings in Stereńczak et al. (2019). These three aforementioned studies all relied on reference heights obtained from measurements subsequent to felling. Other studies, such as Paudel et al. (2021) and Vasilescu (2013), where reference heights were obtained by averaging the surveyors' measurements on each subject tree, reported similar systematic tendencies. In the present study, we found that the magnitude of the systematic error could, to some extent, be explained by tree species, with pine exhibiting larger errors (Table 2, Table 3). This observation aligns with Stereńczak et al. (2019), and the most likely explanatory factor is that, on average, pines have more rounded crowns compared to spruce, making it more challenging to accurately identify the apex of the tree. Aiming the Vertex at a branch that obscures the true treetop can introduce positive systematic errors. In the Stereńczak et al. (2019) study, tree age was also identified as a factor that influenced the magnitude of the error, pointing towards a similar reasoning as for tree species, where slower growth with age creates treetops that are harder to determine accurately.

Although the basis for comparison differs between studies, the systematic errors reported in all these studies, including the present one, can be considered negligible for most practical purposes. While we initially expected the systematic error not to be significantly different from zero, it is important to note that the use of Vertex instruments is sensitive to errors in the instrument settings, particularly the definition of the transponder height. In this research, we did not verify the settings of each individual instrument used, but we ensured that each surveyor was informed about the importance of correct instrument settings. However, if one or more surveyors in our study, despite the instructions of the study protocol, conducted their series of measurements with faulty instrument settings, systematic errors would arise. Therefore, our results may reflect a real-life scenario where instrument settings and errors are unavoidable. Additionally, we acknowledge the possibility of errors in the reference height measurements for certain trees, and we may not have identified and excluded all such trees in our initial investigation of height measurement errors.

The analysis of random measurement errors in the current research revealed a percentual standard deviation of 3.75% for the entire dataset. However, when excluding inexperienced surveyors, this figure decreased by approximately one percentage point. The influence of experience and training on random measurement error was also demonstrated by Kitahara et al. (2010) and Paudel et al. (2021). Experience is crucial not only for technically operating the instrument correctly, but also for making sound decisions regarding suitable locations from which to carry out the height measurement. Our modelling of the percentual absolute difference (Table 4) indicated that, compared to the experts, experienced and inexperienced surveyors had absolute errors that were 0.21 and 0.46 percentage points larger, respectively. Furthermore, random errors appeared to vary between tree species. Similar to the systematic errors, pine exhibited larger errors compared to spruce. Although not as pronounced as in the current study, Luoma et al. (2017) reported similar findings. Once again, the most likely explanation is the more variable apical crown shape of pines. The importance of apical crown shape is further highlighted by the observation that forest density does not seem to affect the magnitude of random error, as pine forests, on average, are less densely stocked, amplifying the influence of apical crown shape.

While acknowledging the need for caution in generalizing our findings to practical field scenarios, it is worth noting that the observed random measurement error of approximately 3% for experienced surveyors is supported by the findings of the aforementioned studies. Comparing to those studies, our result is even in the upper range. The relatively strong consistency of such estimates across multiple studies strengthens our confidence that a random error level of around 3% is a robust approximation for practical field-based height measurements using the Vertex hypsometer. However, it is important to consider specific environmental conditions and tree characteristics that may introduce variations in error levels, in addition to expertise. To enhance the accuracy and reliability of height measurements in different contexts, future research incorporating a wider range of tree species, diverse geographical locations, and various instrument settings would provide further insights into the generalizability of our findings and refine the understanding of error estimation for practical field applications using the Vertex hypsometer.

6. Conclusions

In conclusion, this research examined height measurement errors in spruce and pine trees, considering tree height, forest density, and surveyor experience. The results indicated minimal systematic errors when using Vertex instruments, similar to previous studies. Pines exhibited larger systematic and random errors compared to spruce, likely due to their crown shape. Random measurement errors, with a percentual standard deviation of 3.75% for the entire dataset, decreased when inexperienced surveyors were excluded. Surveyor experience and training influenced both systematic and random errors. For experienced surveyors, a random error level of approximately 3% is a reliable approximation for practical field-based height measurements with the Vertex hypsometer. However, caution should be exercised when applying these findings to different environmental conditions and tree characteristics.

Acknowledgements

We gratefully acknowledge the funding provided by Borregaard Forskningsfond for this research. We extend our heartfelt appreciation to the dedicated individuals who contributed to the data collection efforts of this study, including students and employees at the Norwegian University of Life Sciences, as well as the dedicated staff of the Norwegian NFI.

A special acknowledgment is reserved for Knut Ole Viken, for integrating data collection for this study into the already demanding schedule of the NFI's annual field worker calibration session.

References

Andersen, H.-E., Reutebuch, S., & McGaughey, R. (2006). A rigorous assessment of tree height measurements obtained using airborne LIDAR and conventional field methods. *Canadian Journal of Remote Sensing*, *32*. https://doi.org/10.5589/m06-030

Avery, T. E., & Burkhart, H. E. (2015). *Forest Measurements: Fifth Edition*. Waveland Press. https://books.google.no/books?id=IWx1CQAAQBAJ

Brantseg, A. (1967). Volume functions and tables for Scots pine: South Norway. *Meddelelser fra Det Norske Skogforsøksvesen*, *22*, 689-739.

Braastad, H. (1966). Volume tables for birch. *Meddelelser fra Det Norske Skogforsøksvesen*, *21*, 23-78.

Fitje, A. (1967). Feil ved høydemålinger på enkelttrær. *Meddelser fra Det Norske Skogforsøksvesen*, *84*(22), 107-137.

Fitje, A., & Vestjordet, E. (1977). Stand height curves and new tariff tables for Norway spruce. *Meddelelser fra Norsk Institutt for Skogforskning*, *34*(2), 23-68.

Fortin, M., & DeBlois, J. (2010). A statistical estimator to propagate height prediction errors into a general volume model. *Canadian Journal of Forest Research*, *40*(10), 1930-1939. https://doi.org/10.1139/X10-107

Ganz, S., Käber, Y., & Adler, P. (2019). Measuring Tree Height with Remote Sensing—A Comparison of Photogrammetric and LiDAR Data with Different Field Measurements. *Forests*, *10*(8). https://doi.org/10.3390/f10080694

Hansen, E., Rahlf, J., Astrup, R., & Gobakken, T. (2023). Taper, volume, and bark thickness models for spruce, pine, and birch in Norway. *Scandinavian Journal of Forest Research*, *38*(6), 413-428. https://doi.org/10.1080/02827581.2023.2243821

Hansen, E. A. (2021). Nøyaktighet og presisjon ved høydemålinger på enkelttrær. *Master thesis, NMBU*, 45 p.

Holmgren, R. (2019). Reliability of tree height measurement: A comparison of remote sensing methods and field measurement. *Bachelor Thesis, Department of Forest Ecology and Management, Swedish University of Agricultural Sciences*, 19 p.

Jurjević, L., Liang, X., Gašparović, M., & Balenović, I. (2020). Is field-measured tree height as reliable as believed – Part II, A comparison study of tree height estimates from conventional field measurement and low-cost close-range remote sensing in a deciduous forest. *ISPRS Journal of Photogrammetry and Remote Sensing*, *169*, 227-241. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2020.09.014

Kangas, A. (1996). On the bias and variance in tree volume predictions due to model and measurement errors. *Scandinavian Journal of Forest Research*, *11*(1-4), 281-290. https://doi.org/10.1080/02827589609382937

Kershaw, J. A., Ducey, M. J., Beers, T. W., & Husch, B. (2016). *Forest Mensuration*. Wiley. https://books.google.no/books?id=JDFtDQAAQBAJ

Kitahara, F., Mizoue, N., & Yoshida, S. (2010). Effects of training for inexperienced surveyors on data quality of tree diameter and height measurements. *Silva Fennica*, *44*(4), 657-667.

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

Larjavaara, M., & Muller-Landau, H. (2013). Measuring tree height: A quantitative comparison of two common field methods in a moist tropical forest. *Methods in Ecology and Evolution*, *4*, 1-9. https://doi.org/10.1111/2041-210X.12071

Luoma, V., Saarinen, N., Wulder, M. A., White, J. C., Vastaranta, M., Holopainen, M., & Hyyppä, J. (2017). Assessing Precision in Conventional Field Measurements of Individual Tree Attributes. *Forests*, *8*(2). https://doi.org/10.3390/f8020038

Marklund, L. G. (1988). Biomass functions for pine, spruce and birch in Sweden. *Rapport. Swedish Univeristy of Agricultural Sciences*, *45*, 1-73.

Nersten, S., & Næsset, E. (1992). Nøyaktighet av bestandstaksering med relaskop. *Meddelelser fra Skogforsk*, *48*(8), 1-22.

Paudel, P., Beckschäfer, P., & Kleinn, C. (2021). Impact of training on different observers in forest inventory. *Banko Janakari*, *31*(1), 12-22. https://doi.org/10.3126/banko.v31i1.37338

R Core Team. (2020). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.* In

Ramli, M. F., & Tahar, K. N. (2020). Homogeneous tree height derivation from tree crown delineation using Seeded Region Growing (SRG) segmentation. *Geo-spatial Information Science*, *23*(3), 195-208. https://doi.org/10.1080/10095020.2020.1805366

Sexton, J. O., Noojipady, P., Anand, A., Song, X.-P., McMahon, S., Huang, C., Feng, M., Channan, S., & Townshend, J. R. (2015). A model for the propagation of uncertainty from continuous estimates of tree cover to categorical forest cover and change. *Remote Sensing of Environment*, *156*, 418-425. https://doi.org/https://doi.org/10.1016/j.rse.2014.08.038

Smith, A., Granhus, A., & Astrup, R. (2016). Functions for estimating belowground and whole tree biomass of birch in Norway. *Scandinavian Journal of Forest Research*, *31*(6), 568-582. https://doi.org/10.1080/02827581.2016.1141232

Smith, A., Granhus, A., Astrup, R., Bollandsås, O. M., & Petersson, H. (2014). Functions for estimating aboveground biomass of birch in Norway. *Scandinavian Journal of Forest Research*, *29*(6), 565-578. https://doi.org/10.1080/02827581.2014.951389

Stereńczak, K., Mielcarek, M., Wertz, B., Bronisz, K., Zajączkowski, G., Jagodziński, A. M., Ochał, W., & Skorupski, M. (2019). Factors influencing the accuracy of ground-based tree-height measurements for major European tree species. *Journal of Environmental Management*, *231*, 1284-1292. https://doi.org/https://doi.org/10.1016/j.jenvman.2018.09.100

Vasilescu, M. (2013). Standard error of tree height using vertex III. *Bulletin of the Transilvania University of Brasov, Series II: Forestry, Wood Industry, Agricultural Food Engineering*, *6*, 75-80.

Vestjordet, E. (1967). Functions and tables for volume of standing trees. Norway spruce. *Meddelelser fra Det Norske Skogforsøksvesen*, *22*, 539-574.

Villasante, A., & Fernandez, C. (2014). Measurement errors in the use of smartphones as low-cost forestry hypsometers. *Silva Fennica*, *48*(5), 1-11. https://doi.org/doi:10.14214/sf.1114

Vorster, A. G., Evangelista, P. H., Stovall, A. E. L., & Ex, S. (2020). Variability and uncertainty in forest biomass estimates from the tree to landscape scale: the role of allometric equations. *Carbon Balance and Management*, *15*(1), 8. https://doi.org/10.1186/s13021-020-00143-6

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