




RESEARCH PAPER

Open Access



# An analysis of stand-level size distributions of decay-affected Norway spruce trees based on harvester data

Janne Räty<sup>1,2\*</sup> , Ari M. Hietala<sup>3</sup>, Johannes Breidenbach<sup>2\*</sup> and Rasmus Astrup<sup>2</sup>

## Abstract

**Key message** We studied size distributions of decay-affected Norway spruce trees using cut-to-length harvester data. The harvester data comprised tree-level decay and decay severity recordings from 101 final felling stands, which enabled to analyze relationships between size distributions of all and decay-affected trees. Distribution matching technique was used to transfer the size distribution of all trees into the diameter at breast height (DBH) distribution of decay-affected trees.

**Context** Stem decay of Norway spruce (*Picea abies* [L.] Karst.) results in large economic losses in timber production in the northern hemisphere. Forest management planning typically requires information on tree size distributions. However, size distributions of decay-affected trees generally remain unknown impeding decision-making in forest management planning.

**Aims** Our aim was to analyze and model relationships between size distributions of all and decay-affected Norway spruce trees at the level of forest stands.

**Methods** Cut-to-length harvester data of 93,456 trees were collected from 101 final felling stands in Norway. For each Norway spruce tree (94% of trees), the presence and severity of stem decay (incipient and advanced) were recorded. The stand-level size distributions (diameter at breast height, DBH; height, H) of all and decay-affected trees were described using the Weibull distribution. We proposed distribution matching (DM) models that transform either the DBH or H distribution of all trees into DBH distributions of decay-affected trees. We compared the predictive performance of DMs with a null-model that refers to a global Weibull distribution estimated based on DBHs of all harvested decay-affected trees.

**Results** The harvester data showed that an average-sized decay-affected tree is larger and taller compared with an average-sized tree in a forest stand, while trees with advanced decay were generally shorter and thinner compared with trees having incipient decay. DBH distributions of decay-affected trees can be matched with smaller error index (EI) values using DBH ( $EI = 0.14$ ) than H distributions ( $EI = 0.31$ ). DM clearly outperformed the null model that resulted in an EI of 0.32.

Handling editor: John Lhotka

\*Correspondence:

Janne Räty

janne.raty@luke.fi;

Johannes Breidenbach

johannes.breidenbach@nibio.no

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

**Conclusions** The harvester data analysis showed a relationship between size distributions of all and decay-affected trees that can be explained by the spread biology of decay fungi and modeled using the DM technique.

**Keywords** Root and butt rot, *Heterobasidion* spp., *Armillaria* spp., Cut-to-length harvester, Forest management and planning

## 1 Introduction

Norway spruce (*Picea abies* [L.] Karst.) dominates the boreal forests in Northern Europe and the subalpine areas of the Alps and Carpathian Mountains. Owing to its good performance in different site conditions, it was also planted outside its natural distribution on lower elevations in more temperate forests (Caudullo et al. 2016). Wood of Norway spruce has a low extractive content, which makes it more prone to decay compared to Scots pine (*Pinus sylvestris* [L.]), the other widely distributed conifer in Eurasia. Fungi in the genera of *Heterobasidion* and *Armillaria*, the so-called white-rot fungi capable of decomposing all structural polymers in wood, are the most important root and stem wood-decaying fungi of Norway spruce. In Norway, a nationwide survey of stumps in clear-cut forests concluded that 26.8% of the Norway spruce trees had stem decay that was most often (71% of the decay-affected stumps) associated with the *Heterobasidion* species (Huse et al. 1994). The situation is similar in the other Nordic countries, where 20% or more of the trees in managed Norway spruce forests show decay caused by *Heterobasidion* species by the time of final harvest (Bendz-Hellgren and Stenlid 1995).

In particular, the *Heterobasidion* species have benefited from the current forest management practices that involve logging operations year-round. Primary infection of forest stands by these fungi takes place through the colonization of fresh stumps cut during the active sporulation of the fungus. After spore infection of a fresh stump, the fungal mycelia rapidly colonize the dead root system. Once the next-generation plants establish root contact with the pathogen-colonized stump roots, the infection is transferred to the next tree generation (Piri 1996; Stenlid and Redfern 1998). A similar transfer process takes place between infected mature Norway spruce trees and advanced regeneration (Piri and Korhonen 2001). Owing to active tree defense, the spread of *Heterobasidion* mycelia is slower in living roots than in roots of stumps. In general, the roots of the next-generation saplings can become infected by the fungus after around 10 years of growth (Piri 2003). Stem colonization usually initiates only after stem heartwood has started to develop, which in Norway spruce begins between ages of 25 and 40 (Korhonen and Stenlid 1998). Besides the infection routes through root contacts of neighboring trees, wounds on roots

created by mechanical logging during snowless and frost-free periods are potential infection routes as well (Metsläid et al. 2018). The risk of infection transfer between neighboring trees via root contacts is higher in pure species forest stands in comparison with admixed stands (Lindén and Vollbrecht 2002; Möykkynen and Pukkala 2010).

While *Heterobasidion* species have benefited from summertime loggings, the species of *Armillaria* typically infect weakened trees, such as drought-stressed Norway spruce trees. From infected trees, or stumps of infected trees, *Armillaria* species spread to neighboring trees with the aid of rhizomorphs that grow freely in the soil. A single individual of an *Armillaria* species can occupy Norway spruce trees at a distance of tens of meters (Prospero et al. 2003).

Wood decay causes considerable economic losses in timber production because decay arises normally from the roots and impairs the quality of the most valuable part of tree stem. While decay caused by *Armillaria* normally reaches a height of 1–2 m in a stem, the heartwood decay column caused by *Heterobasidion* species can reach a height of 10–12 m in stems of mature Norway spruce trees at a late stage of infection. Regardless of stem decay, decay-affected trees can remain alive even for several decades and do not necessarily show any obvious external signs of decay. The decay-affected part of the stem does not fulfill the quality criteria required for sawlogs but is instead used as pulpwood or energy wood, depending on the degree of decay. Within the European Union, the annual losses attributed to *Heterobasidion* spp. in timber production were estimated to be approximately 800 million € (Woodward et al. 1998). Besides the losses due to timber wood quality, decay negatively affects tree growth and carbon sequestration due to investment of energy resources to tree defense instead of growth (Bendz-Hellgren and Stenlid 1995; Oliva et al. 2012). In addition, decay-affected trees are also prone to other forest damages such as stem breakage due to wind or snow. Considering the spread biology of *Heterobasidion*, there are numerous reasons to anticipate that the amount of decay and the inflicted economic losses will increase along with climate warming (Müller et al. 2014). The anticipated increase of dry summers along the progression of climate change is likely to increase the occurrence of decay caused by species of *Armillaria*.

Since decay fungi are generally confined to the physiologically inactive heartwood in Norway spruce, the presence of decay in a tree and decay frequency at stand level cannot be reliably determined without destructive sampling, e.g., drilling (Vollbrecht and Agestam 1995). Unless one has prior knowledge of decay history at the stand level, i.e., documented records of decay in the previous tree generation or upon selective harvests (Müller et al. 2018), it is difficult to take decay into account when assessing the economic value of timber volume and when considering alternative forest management scenarios at the level of forest stands.

National Forest Inventories (NFI) typically collect information on the occurrence of stem decay, and NFI data have been used to model the risk of decay in forests (Thor et al. 2005; Mattila and Nuutinen 2007; Hysten and Granhus 2018). Some previous studies have proposed mechanistic models for simulation of the spread of decay at the level of forest stands (Pukkala et al. 2005; Honkaniemi et al. 2014). The applicability of such models as decision tools in practical forest management planning would ideally require knowledge of a large group of parameters, such as localization and frequency of decay in the prior tree generation. Currently, the most effective way to collect tree-level decay information for complete forest stands is to utilize data collected by cut-to-length harvesters, but this approach requires that the operator of a harvester observes and records the presence of decay during harvesting operation. Räty et al. (2021) utilized such harvester data and showed that stand-level timber volume affected by decay can be mapped by means of airborne laser scanning data and environmental variables at satisfactory error levels, given that harvester data are available close to the target forests.

The planning of stand-level forest management treatments typically requires simulations that use a group of models to account for growth dynamics of forests. In order to apply tree-level models in growth simulators, the description of tree sizes in a forest stand is a prerequisite. The distribution of tree sizes is also needed in the calculation of timber assortment volumes, which form the basis of the economic value of a forest stand. The most common way to describe the size distribution of all trees is to utilize a theoretical diameter at breast height (DBH) distribution model, such as a two-parameter Weibull distribution, which can be estimated for forest stands by means of relatively rapid field measurements (Siipilehto and Mehtätalo 2013) or by coupling a sample of field plots and remotely sensed information (Gobakken and Næsset 2004). Tree size information of decay-affected trees, combined with information on the number of decay-affected trees, would facilitate the assessment of

the economic value of timber resources and the growth potential of forests at the level of forest stands.

While the spread mechanisms of decay fungi are relatively well understood, stand-level information on the size distribution of decay-affected trees is rarely available to support decision-making in forest management planning. Harvester data with tree-level decay recordings enable to observe size distributions of all and decay-affected trees at the level of entire forest stands. Stand-level data of the presence and extension of decay in each tree are rare, since such data are too laborious to collect using traditional field measurements. If such data are accessible, the data allow to study the relationship between the size distribution of decay-affected trees and that of all trees at a forest stand. To this end, we used the distribution matching (DM) technique (González and Woods 2002) to model the relationship between size distributions of all and decay-affected trees. The objectives of this study were as follows:

- To compare DBH and height (H) distributions of all trees with corresponding distributions of decay-affected trees at the level of harvested stands. Our aim was to scrutinize whether the size distributions of all and decay-affected trees differ at stand level.
- To analyze and compare size distributions of trees with incipient and advanced stem decay. Our main goal was to find out to what extent the degree of decay influences the shape and location of tree size distributions.
- To propose DM models that transfer DBH or H distributions of all trees into DBH distribution of decay-affected trees.

## 2 Materials and methods

### 2.1 Study area

The forest stands from which the harvester data were collected are located between the latitudes 59° and 65° in Norway. Because harvests are typically carried out in forests used for commercial timber production, the forest stands considered here are dominated by the boreal coniferous tree species Norway spruce or Scots pine. Broadleaved species, mostly birch (*Betula* spp. [L.]), may occur as mixtures in these conifer-dominated forests.

### 2.2 Harvester data

Harvester data were collected using five different machines between 2019 and 2021. The harvest operations were planned and operated by different companies, and the harvested sites are heterogeneous units consisting of several forest stands. The harvested sites were final-felled forests, which means that all merchantable trees, with the

exception possible retention and seed trees, were cut and recorded by the harvesters.

The harvesters were equipped with GNSS (global navigation satellite system) devices, which recorded an XY location when a tree was felled. Three of the harvesters determined the position of the boom tip with sensors that measure crane length and orientation. The rest of the harvesters only recorded the XY location of machine. The machine-positioned trees followed the machine's driving routes and required post-processing prior to stand delineation. The machine-positioned trees were distributed by adding random deviations of 8 m to the x and y coordinates (Räty et al. 2021). The positioning errors of the harvested trees are expected to vary between 5 and 20 m.

The XY locations of the harvested trees were overlaid with segments provided by the Norwegian forest resource map SR16 (Astrup et al. 2019; Hauglin et al. 2021). The SR16 segments, based on numerous forest attributes derived from remotely sensed data, aim to mimic actual forest stands. The harvested sites did not usually follow the boundaries of the SR16 segments. Therefore, the XY locations of trees were used to delineate harvested sites by creating two-dimensional alpha shapes ( $\alpha = 25$  m) using the *alphahull* package (Pateiro-Lopez and Rodriguez-Casal 2019) in the R environment (R Core Team 2022). For each alpha shape, a buffer of 2 m was added to approximately account for the distance between crown edge and stem location. Finally, each SR16 segment was cropped with the corresponding alpha shape polygon to establish a harvested segment. More information on the generation of harvested segments can be found in Räty et al. (2021). We only selected harvested segments that were larger than 0.5 ha and had a spruce volume proportion  $\geq 90$  %. Since our aim was to model the size distribution of decay-affected trees using theoretical models, we selected harvested segments with a sufficient number of decay-affected trees (at minimum 20 trees per segment). We wanted to focus on harvested segments in which the frequency of decay-affected trees was comparable to that found in prior studies in Norway (Huse et al. 1994). Therefore, we selected harvested segments with more than 15% of trees showing stem decay, which resulted in an average proportion of segment-level decay-affected trees of 22%. A total of 101 harvested segments ( $n = 101$ ) containing 93,456 harvested trees (94% Norway spruce trees) were utilized in this study (Fig. 1).

The harvesters registered diameter in 10 cm intervals along each stem, and these measurements were used to estimate DBH for each tree. The harvesters also registered merchantable tree length, but they were not able to register total tree length (Nordström and Hemmingsson 2018). Thus, tree height was predicted using species-specific height-diameter curves (Eide 1954; Strand

1967; Blingsmo 1985) that were calibrated according to the harvester-based diameter and corresponding height measurements along stem (Hauglin et al. 2018; Räty et al. 2021).

As part of a research project, the harvester operators recorded the presence of visually observable decay at crosscut surfaces during the harvest operations. The operators also evaluated the severity of decay based on the width of the decay column. The harvested trees with decay were categorized into one of the following decay classes: (1) *incipient* decay, trees with decay columns covering less than 50% of stem diameter at all crosscuttings, and (2) *advanced* decay, trees with decay columns covering 50% or more of stem diameter at least at a single crosscut. The harvester data were stored in the Standard for Forest machine Data and communication (Stand-ForD2010) format (Arlinger et al. 2012). Forest attributes associated with the harvested segments are shown in Table 1.

Figure 2 shows the DBH and H distribution of healthy and decay-affected trees in the harvester data. Figure 9 in the Appendix shows proportions of decay-affected trees by DBH and H classes.

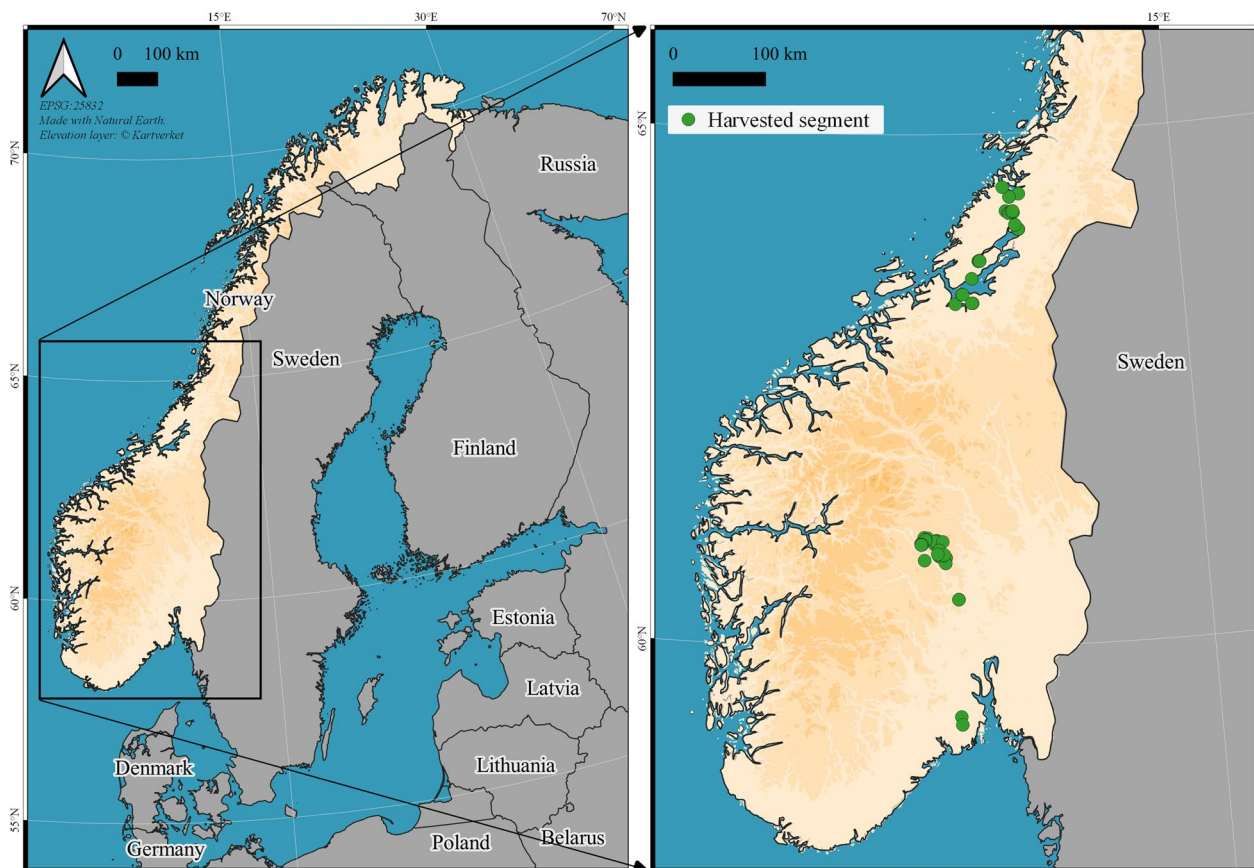
## 2.3 Analyzing tree size distributions

### 2.3.1 Assessing the similarity of the observed tree size distributions

The two-sample Kolmogorov-Smirnov (KS) test was used to statistically test the similarity of the observed tree size distributions. The KS test was carried out to test the similarity of size distributions associated with (1) healthy and decay-affected trees, (2) all and decay-affected trees, and (3) trees with incipient and advanced decay. The comparisons were carried out separately for each harvested segment. The null hypothesis (H0) is that the two tree size distributions come from the same continuous distribution. The statistical significance level of 5% was used to test the null hypothesis. The KS test was carried out using the *stats* package in the R environment.

### 2.3.2 Weibull distribution

A Weibull probability density function (pdf) was fitted separately for both DBHs and Hs by maximizing the likelihood in each harvested segment for the following groups of trees: (1) *all trees*, (2) *decay-affected trees*, (3) *trees with incipient decay*, and (4) *trees with advanced decay*. The group *all trees* contains healthy and decay-affected trees in a harvested segment, while the group of *decay-affected trees* contains trees with decay (independent on the degree of decay). The group *trees with incipient decay* contains trees that have minor decay, whereas the group *trees with advanced decay* includes trees with severe decay.



**Fig. 1** Map of the study area and locations of harvested sites

**Table 1** Means, standard deviations (SD), minimums (min), and maximums (max) of attributes associated with the harvested segments ( $n = 101$ )

	Mean	SD	Min	Max
Merchantable volume ( $\text{m}^3 \cdot \text{ha}^{-1}$ )	229.62	104.72	90.74	775.87
Lorey's height (m)	18.30	2.45	13.13	24.53
Stem frequency ( $\text{stems} \cdot \text{ha}^{-1}$ )	714	214	228	1485
Basal area ( $\text{m}^2 \cdot \text{ha}^{-1}$ )	29.88	9.86	13.07	82.7
Quadratic mean diameter (cm)	23.25	2.84	18.27	32.25
Proportion of decay-affected trees (%)	22.07	7.88	11.59	53.92
Elevation above sea level (m)	430.34	304.31	7.92	931.57
Area (ha)	1.32	0.87	0.62	5.11

The two parameters of the Weibull function were estimated for each harvested segment by maximizing the likelihood function. The two-parameter Weibull pdf is as follows:

$$f(x | c, b) = \frac{c}{b} \left(\frac{x}{b}\right)^{c-1} e^{-(x/b)^c} \quad (1)$$

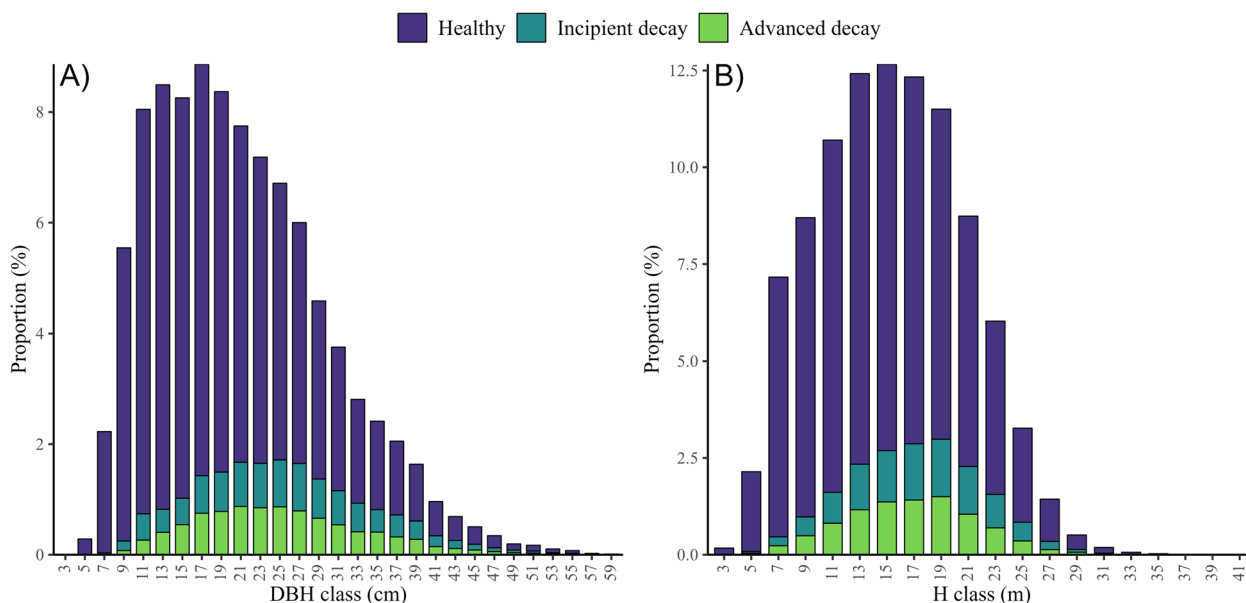
where  $c$  is the shape parameter,  $b$  is the scale parameter, and  $x$  is a harvester measured DBH or H.  $c, b, x > 0$ .

The likelihood function to be maximized is as follows:

$$l(\theta_i | x_i) = \sum_{j=1}^k \log f(x_{ij} | \theta_i) \quad (2)$$

where  $\theta_i$  is the vector of Weibull parameters  $b$  and  $c$  in harvested segment  $i$  ( $i = 1, \dots, n$ ) that are to be estimated,  $x_i$  refers to the vector of harvester-based tree measurements (DBH or H) in harvested segment  $i$ ,  $k$  is the number of trees in harvested segment  $i$ ,  $x_{ij}$  is tree measurement  $j$  in harvested segment  $i$  ( $j = 1, \dots, k$ ), and  $f(\cdot)$  is the Weibull pdf.

The Weibull parameters were estimated using the maximum likelihood estimation (MLE) implemented in the *ForestFit* package (Teimouri 2021) in the R environment. No transformations were applied for the Weibull parameters during the estimation.



**Fig. 2** The diameter (A DBH) and height (B H) distribution of harvested healthy and decay-affected trees in the harvester data. The decay-affected trees were categorized into incipient and advanced decay classes

### 2.3.3 Comparing tree size distributions using deciles of fitted Weibull distributions

We compared tree size distributions of all and decay-affected trees in terms of their shape and horizontal position using deciles associated with the fitted Weibull pdfs. We used Weibull pdfs, instead of raw distributions, in the comparison of tree size distributions of all and decay-affected trees for two reasons. First, the focus was to investigate the general relationships, such as location and width, between all and decay-affected trees, and the continuous pdfs are more suitable for that purpose than the discrete raw distribution. Second, we wanted to keep the comparison part of this paper in line with our ultimate goal to propose a smooth transformation function to match size distributions of all and decay-affected trees (Section 2.4).

We compared (1) deciles of the DBH distribution of all trees with those of decay-affected trees and (2) deciles of the H distribution of all trees with those of decay-affected trees. To consider the relationship between decay severity and tree size, we also compared the size distributions of trees with incipient and advanced decay.

The comparison of distributions was carried out by calculating differences between deciles of two fitted Weibull distributions ( $\Delta D_{10}$ ,  $\Delta D_{20}$ , ...,  $\Delta D_{90}$ ) for each harvested segment (Fig. 3). Finally,  $\Delta D_{10}$ , ..., and  $\Delta D_{90}$  were visually presented using box plots that show relationships between the compared density distributions.

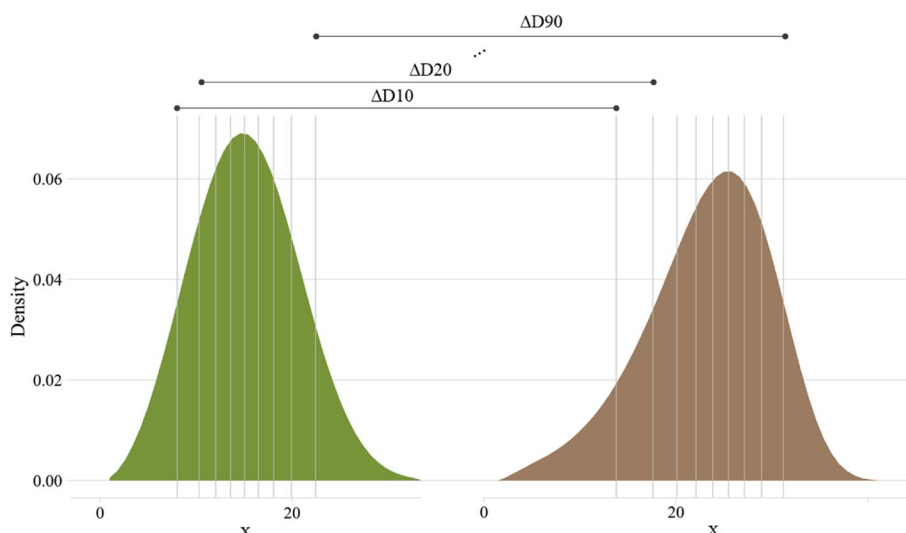
### 2.4 Predicting DBH distributions of decay-affected trees

#### 2.4.1 Null model

A straightforward approach to estimate DBH distributions of decay-affected trees at the level of harvested segments is to use a global estimate for all harvested segments. The global estimate was constructed by fitting a Weibull pdf using MLE and all harvested decay-affected trees in the study area. We call this approach “null model.” The null model is used as a reference approach in the evaluation of the performance of DM for the prediction of DBH distributions of decay-affected trees.

#### 2.4.2 Distribution matching

DM is a well-known method in the field of digital image processing, and DM has also been applied in various remote sensing-based forestry applications, such as the calibration of predicted forest attribute maps (Baffetta et al. 2012) and transformation of crown-radii distributions to DBH distributions (Vauhkonen and Mehtätalo 2015). We used DM to transform DBH or H distributions of all trees (initial distribution) to DBH distributions of decay-affected trees (target distribution). The initial distributions are fitted Weibull pdfs (Section 2.3.2). We considered H distributions here because they may be easier to obtain than DBH distributions using remotely sensed data in the future. DM pursues to predict the shape and location of the DBH distribution of decay-affected trees, which means that the actual number of decay-affected trees was not predicted.



**Fig. 3** An illustration on how two tree size distributions of a harvested segment were compared using the observed decile values (D10, ..., D90) associated with the fitted Weibull distributions. *x*, height or diameter at breast height

The transformation of initial distributions to target distributions was carried out using a linear mixed-effects model and distribution percentiles:

$$y_{il} = \beta^{(1)} + \beta^{(2)}q_i + \beta^{(3)}r_{il} + \beta^{(4)}r_{il}^2 + z_i^{(1)} + z_i^{(2)}r_{il} + \varepsilon_{il} \tag{3}$$

where  $y_{il}$  is the  $l$ th percentile of the target distribution ( $l = 1, 2, \dots, 99$ ) in harvested segment  $i$ ,  $r_{il}$  is the corresponding percentile of the initial distribution in harvested segment  $i$ ,  $q_i$  refers to the ratio of the interquartile range to the median of the initial distribution of all trees,  $\beta^{(l)}$  are the fixed model parameters,  $z_i^{(1)}$  and  $z_i^{(2)}$  are random intercept and slope parameters in harvested segment  $i$ , and  $\varepsilon_{il}$  is the residual error for the  $l$ th percentile of the target distribution in harvested segment  $i$ . The random parameters were set to zero in the prediction phase.

We established two different DM setups to predict the DBH distribution of decay-affected trees: (1) using the DBH distribution of all trees as an initial distribution ( $DM_{DBH}$ ), and (2) using the H distribution of all trees as an initial distribution ( $DM_H$ ). The parameters of the DM models were estimated in the R environment using the *nlme* package (Pinheiro et al. 2020).

### 2.4.3 Performance assessments

The performances of DM and null model were assessed in a leave-one-out cross-validation (LCV), where target and neighboring harvested segments with a center closer than 500 m were omitted from the training data. The distance limit was used to avoid modeling with data that include potentially very similar forests compared to

the target segment. The predictive performances of the DM and null models were evaluated against the observed DBH distribution of decay-affected trees using the error index (EI, Eq. 4). The EI is a variant of the well-known Reynold’s index (Reynolds et al. 1988).

$$EI_i = \sum_{m=1}^p |f_m - \hat{f}_m| \tag{4}$$

where  $f_m$  and  $\hat{f}_m$  refer to the observed and predicted density of DBH class  $m$  in harvested segment  $i$ , respectively, and  $p$  is the number of DBH classes. The observed DBH distribution of decay-affected trees refers to the Weibull pdf fit. A bin width of 2 cm and  $p = 30$  were chosen, which resulted in the bin midpoints  $m = 1, 3, \dots, 59$  cm. An  $EI_i$  value of 2 refers to a complete disagreement of the distributions compared, whereas an EI value of 0 refers to a complete match of the compared distributions. Mean, minimum, standard deviation, and maximum of the  $EI_i$  values over all harvested segments were reported. Furthermore, the mean error of the mean DBH of decay-affected trees ( $ME_{MDBH}$ , Eq. 5) and the root mean squared error of the mean DBH of decay-affected trees ( $RMSE_{MDBH}$ , Eq. 6) were calculated.

$$ME_{MDBH} = \frac{\sum_{i=1}^n (MDBH_i - \widehat{MDBH}_i)}{n} \tag{5}$$

$$RMSE_{MDBH} = \sqrt{\frac{\sum_{i=1}^n (MDBH_i - \widehat{MDBH}_i)^2}{n}} \tag{6}$$

where  $MDBH_i$  and  $\widehat{MDBH}_i$  refer to the observed and predicted mean DBH of decay-affected trees in harvested segment  $i$ , respectively, and  $n$  is the number of harvested segments. The  $MDBH_i$  and  $\widehat{MDBH}_i$  values were calculated based on the midpoints of the 2 cm bins and corresponding densities.

### 3 Results

#### 3.1 Testing the similarity of the observed size distributions

The KS test rejected the null hypothesis (5% statistical significance level) in 98% and 90% of the harvested segments, when the DBH and H distributions of healthy and decay-affected trees were compared, respectively. Correspondingly, the null hypothesis was rejected in 92% and 78% of the harvested segments, when DBH and H distributions of all and decay-affected trees were compared, respectively. The results of the KS test strongly indicated that the size distributions of decay-affected trees were different compared with those of healthy and all (all category includes decay-affected trees) trees.

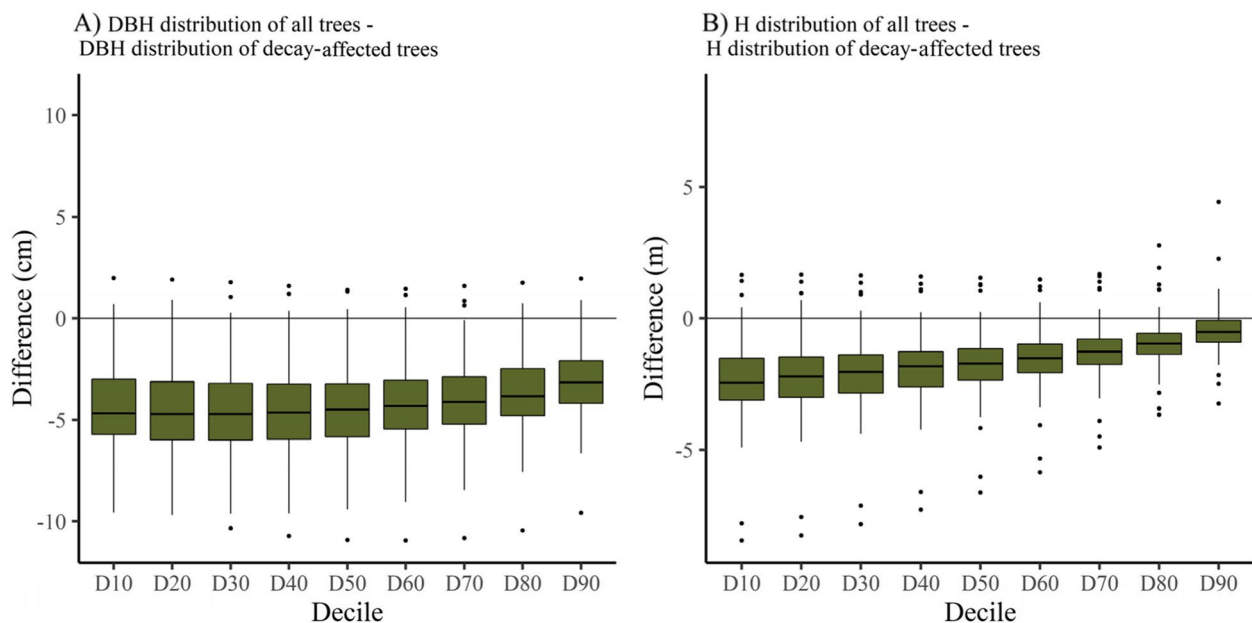
The KS rejected the null hypothesis of no difference (5% statistical significance level) in 67% and 45% of the harvested segments when DBH and H distributions of trees with incipient or advanced decay were compared, respectively. This means that there was a larger difference between the DBH distributions of trees with incipient and advanced decay compared with the corresponding H distributions.

#### 3.2 Comparing size distributions of all and decay-affected trees

Trees with decay showed a right-shifted DBH distribution in comparison with all trees (Fig. 4). This observation implies that stem decay is more likely to be present in larger trees than in smaller trees. Our data showed that the median DBH of decay-affected trees is roughly 4 cm larger compared with the median DBH of all trees. In addition, Fig. 4A shows the general trend that the decile values associated with DBH distributions of all trees were smaller compared with those of decay-affected trees. A similar observation can also be made in the case of H distributions (Fig. 4B), although the difference between the H distribution of decay-affected trees and that of all trees is not as large as in the case of DBH distributions. In the case of DBH distributions, the relationship between decile values of all trees and those of decay-affected trees is rather stable, a slight trend of decreasing difference being visible for deciles 60–90 (Fig. 4A). The corresponding differences associated with H distributions steadily decreased towards large deciles, which indicates different shapes of the compared distributions (Fig. 4B).

#### 3.3 Comparing size distributions of trees with incipient and advanced decay

Trees with advanced decay were in general smaller in terms of DBH and H than trees with incipient decay. This



**Fig. 4** Differences in decile values of 101 harvested segments between **A** DBH distributions of all trees and DBH distributions of decay-affected trees and between **B** H distributions of all trees and H distributions of decay-affected trees in the harvested segments. DBH, diameter at breast height; H, height



trend was clearer in the case of H distributions in comparison with the DBH distributions (Fig. 5A and B).

### 3.4 Distribution matching

EI,  $ME_{MDBH}$ , and  $RMSE_{MDBH}$  values associated with the null model and DMs are presented in Table 2. The model parameters of  $DM_{DBH}$ , the best-performing model, are in Table 3 in the Appendix. DM outperformed the null model when the initial distribution of the transformation was the DBH distribution of all trees, whereas DM using the H distribution of all trees as an initial distribution did not generally outperform the null model.

$DM_{DBH}$  outperformed  $DM_H$  in terms of mean EI,  $ME_{MDBH}$ , and  $RMSE_{MDBH}$ . Figures 6, 7, and 8 show examples of the harvested segments with a small, average, or large EI value associated with  $DM_{DBH}$ , respectively. To avoid showing extreme cases, the examples were selected based on the cumulative distribution of the EI values using 5%, 50%, and 95% cumulative percentages for small, average, and large EI values, respectively.

## 4 Discussion

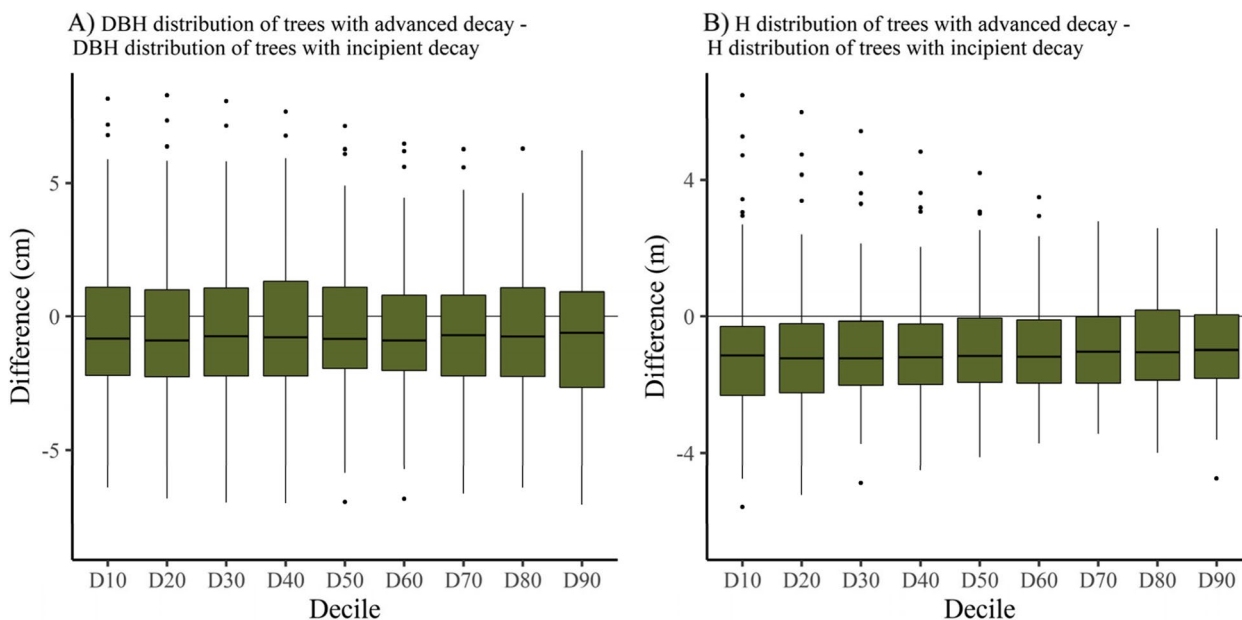
Traditional field measurements are often carried out using field sample plots, because a census of DBHs for whole forest stands would be too laborious. On the contrary, harvesters automatically collect tree-level information from all merchantable trees in forest stands, and stem decay can be simultaneously registered without any considerable reduction in cost efficiency. Previous

**Table 2** Performance assessments of the null model and different distribution matching (DM) setups to predict DBH distributions of decay-affected trees

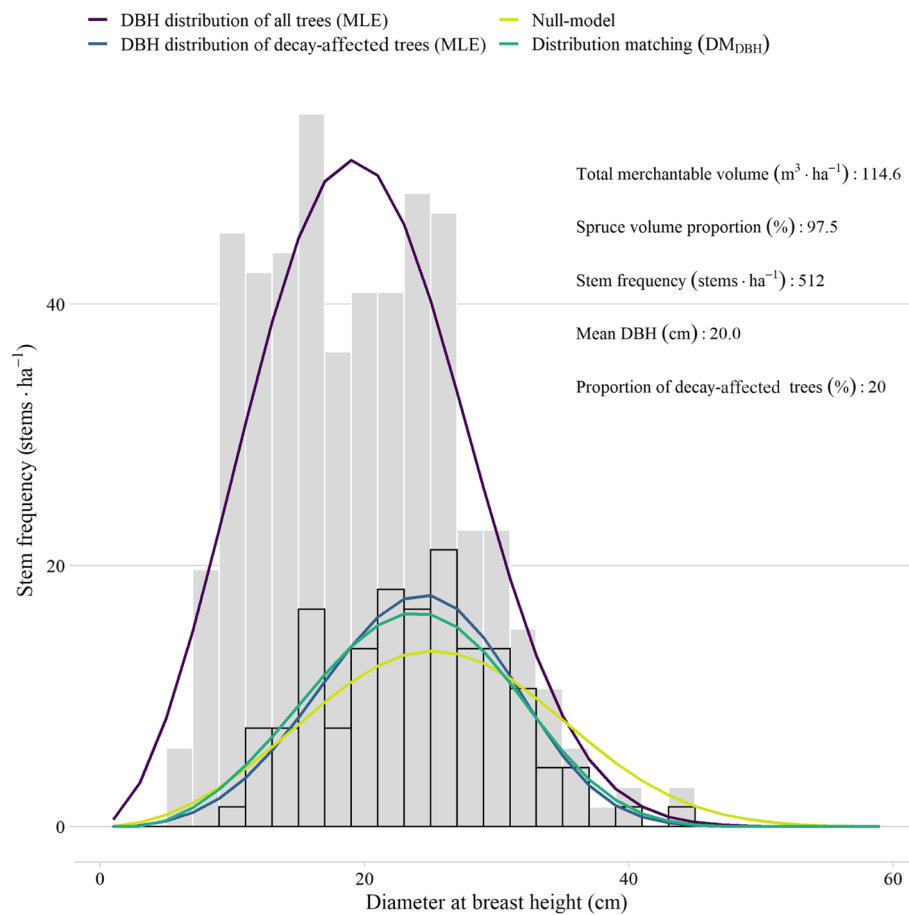
Model	Mean EI (min, SD, max)	$ME_{MDBH}$ (cm)	$RMSE_{MDBH}$ (cm)
Null model	0.32 (0.03, 0.16, 0.94)	0.03	3.46
$DM_{DBH}$	0.14 (0.02, 0.10, 0.68)	0.04	1.53
$DM_H$	0.31 (0.02, 0.20, 1.02)	-0.09	3.62

DBH diameter at breast height, H height, SD standard deviation, EI error index,  $ME_{MDBH}$  mean error of mean DBH of decay-affected trees,  $RMSE_{MDBH}$  root-mean-square error of mean DBH of decay-affected trees

large-scale studies on forest decay have typically relied on NFI data, where the presence of decay is assessed using increment core samples taken at breast height (Thor et al. 2005; Mattila and Nuutinen 2007; Hysten and Granhus 2018). While it would, in principle, be possible to drill all trees of a forest in some circumstances, increment core samples tend to underestimate the frequency of decay-affected trees (Hysten and Granhus 2018). The underestimation results from the fact that short decay columns cannot be observed at breast height (Tamminen 1985). Short decay columns are typical at an early phase of stem decay. In addition, some wood decay fungi do not cause high stem decay columns. Previous studies on the decay presence in Norway spruce stands have also been based on assessment of stumps after clear cutting (Huse et al. 1994). The harvester-based collection of decay



**Fig. 5** **A** Differences in decile values associated with diameter at breast height (DBH) distributions of trees with advanced and incipient decay in the harvested segments and **B** a corresponding comparison in terms of height (H) distributions



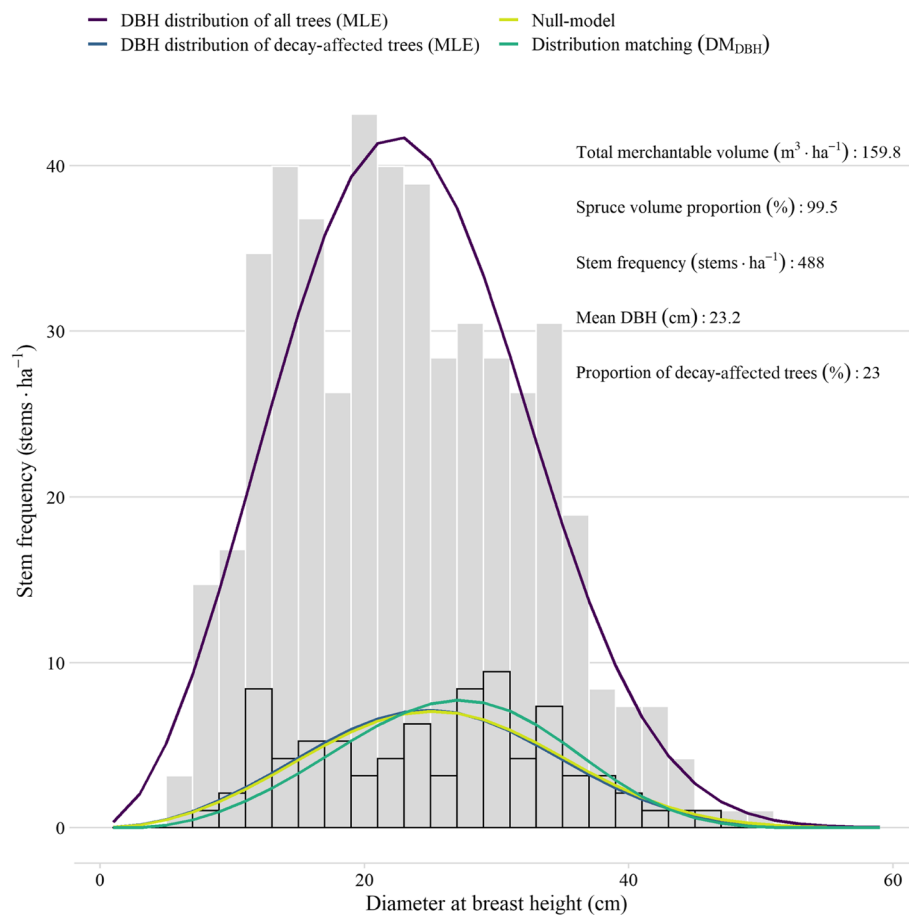
**Fig. 6** Example of the harvested segment that produced a small error index ( $EI = 0.08$ ) value associated with the distribution matching (DM) of diameter at breast height (DBH) distribution of decay-affected trees. The gray bars show the DBH distribution of all trees, whereas the bars with black borders refer to the DBH distribution of decay-affected trees. Note that density distributions were scaled using observed stem frequencies. MLE, maximum likelihood estimation

information resembles stump surveying, but as a benefit, the decay observations can be linked to the corresponding harvester-based stem measurements. As a drawback, harvesters only operate in productive forests which introduce a selection bias in the data collection. In addition, the smallest trees are typically underrepresented ( $DBH < 10$  cm) in the harvested trees compared with the factual status in the forest, since the smallest trees may be too time-consuming to harvest in terms of their commercial value or may not fulfill dimension requirements of merchantable logs.

Our data showed that stem decay was absent or rare in trees below 9 cm in DBH, while the proportion of decay-affected trees increased along with DBH classes, reaching a level of up to 40% in DBH classes  $\geq 45$  cm. As a result, the median DBH of decay-affected trees was roughly 4 cm larger than that of all trees. We did not aim to determine the causative agents of decay in this study, but the findings reflect the biology of the most common wood

decay pathogens in Norway spruce, namely white-rot fungi in the genera *Heterobasidion* and *Armillaria*, which are considered to cause more than 90% of the decay in Norway spruce in Norway (Huse et al. 1994). Stem decay caused by *Heterobasidion* and *Armillaria* generally originates from root infection, but due to active tree defense, the spread of decay fungi is slow in living roots of young trees (Bendz-Hellgren et al. 1999). Therefore, it is reasonable to assume that also many of the small trees had infection by wood decay fungi, although the decay was still confined to the roots, and thus not yet visible at the stem base during harvesting.

The observed increase in the proportion of decay-affected trees along with tree size is in line with the findings of Hysten and Granhus (2018). This finding reflects the spread biology of the most common decay fungi of Norway spruce. The older the next-generation trees are, the more root contacts they have established with stumps of decay-affected previous-generation trees.

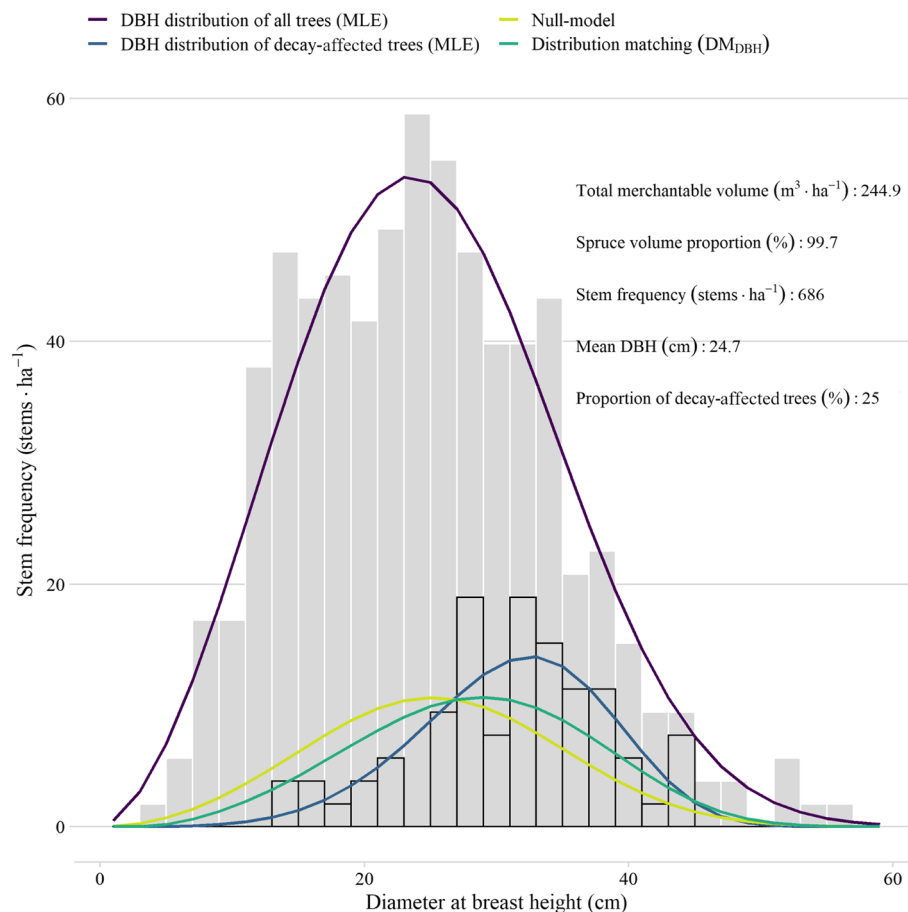


**Fig. 7** Example of the harvested segment that produced an average error index ( $EI = 0.17$ ) value associated with the distribution matching (DM) of diameter at breast height (DBH) distribution of decay-affected trees. The gray bars show the DBH distribution of all trees, whereas the bars with black borders refer to the DBH distribution of decay-affected trees. Note that density distributions were scaled using observed stem frequencies. MLE, maximum likelihood estimation

The established root contacts increase the probability of infection transmission (Pukkala et al. 2005). Trees receiving infection via root contacts in turn spread the infection further to neighboring trees once the trees establish root contacts or grafts. It is also worth noting that mycelial spread of decay fungi is faster in roots of stumps than in living roots of trees owing to the absence of defense responses in dead roots (Schönhar 1978).

Based on the current understanding, the recorded trees with incipient stem decay represent more recent infections than the trees with advanced stem decay. Our findings indicate that trees with incipient decay are on average larger and taller compared with trees with advanced decay. This is expected, since the proportion of water-conductive stem sapwood decreases along with advancement of decay, while at the same time, the affected trees allocate more carbon to tree defense phenolics at the interface between pathogen-colonized wood and inner sapwood (Bendz-Hellgren and Stenlid 1995;

Oliva et al. 2012; Nagy et al. 2022). Because the heights of the harvested trees were predicted using conventional taper curves, the effect of decay on the growth rates was neglected, which may result in overpredictions of heights associated with decay-affected trees. We are not aware of any study that would have considered whether, and to what extent, the heartwood decay of Norway spruce affects the relation between diametric and height growth. An additional factor is that trees with advanced stem decay likely also have a less functioning root system than trees with incipient decay, but this aspect is not well established in previous studies. When considering the timing of final felling in a rotation forestry, decay may have implications on optimal rotation time from the perspective of carbon sequestration or economy as suggested by Möykkynen and Pukkala (2010). The characterization of the DBH distribution of decay-affected trees allows to better account for economic loss associated with timber quality resulting from decay and even



**Fig. 8** Example of the harvested segment that produced a large error index ( $EI = 0.36$ ) value associated with the distribution matching (DM) of diameter at breast height (DBH) distribution of decay-affected trees. The gray bars show the DBH distribution of all trees, whereas the bars with black borders refer to the DBH distribution of decay-affected trees. Note that density distributions were scaled using observed stem frequencies. MLE, maximum likelihood estimation

reduced biomass growth in the simulations of different management scenarios.

DM, i.e., distribution matching using the observable DBH distribution of all trees in a stand as input, outperformed the null model assuming a fixed distribution in the prediction of the DBH distribution of the decay-affected trees. This suggests a predictable relationship between the size distribution of all trees and that of decay-affected trees. DM predicts densities of DBH classes, which means that the number of trees with decay must be derived from other sources. Potential data sources for the estimation of the number of decay-affected trees are, for example, nation-wide stump surveys (Huse et al. 1994), NFI data-based models combined with climatic or environmental data (Hysten and Granhus 2018), or a combination of harvester and remotely sensed data (Räty et al. 2021).

The practical applicability of our current DM approach is limited to certain forest types and requires that the DBH

(or H) distribution of merchantable trees in a forest stand is known. Our data consisted of spruce-dominated stands, and therefore, we were not able to study the effect of species mixtures on the size distribution of decay-affected trees. Möykkynen and Pukkala (2010), among many others, concluded that the spread rate of *Heterobasidion* decay is significantly slower in admixed forests compared with pure spruce forests. It is noteworthy that this scenario does not apply to *Armillaria*, as the rhizomorphs of *Armillaria* species causing decay in Norway spruce use also broadleaved trees or their stumps as a food base (Keča and Solheim 2011). It should also be noted that our harvester data were collected from mature forests that are used for commercial timber production. Therefore, our DM may not be applicable in young forests or forests that are not managed for the purposes of timber production. Regarding the silvicultural activity, a critical question associated with the spread of decay fungi is the history of silvicultural treatments, such as thinnings. Thinnings are not routinely carried out

in most forests in Norway, which may have a mitigating effect on the spread of decay compared with thinned forest stands (Metslaid et al. 2018). Thinnings typically affect the width and shape of DBH distribution of all trees, which may change the relationships between size distributions of all and decay-affected trees. We did not have data about the previous management history of the stands included in our study, and a history of thinning could be one reason for the poor predictive performance of our method in some stands (Fig. 8).

Future studies could supplement this study by accounting for local differences in the prediction of the DBH distribution of decay-affected trees, for example, by using various remotely sensed data and environmental, climatic, and historical data sources. The methodology presented here could also be adapted to stand-level forest inventories, where the initial DBH distribution is predicted, for example, by means of airborne laser scanning data or rapid field measurements. In addition, the potential of harvester data for the modeling of decay column properties, such as height and width, should be investigated. Our study indicates that harvester data have a high potential to increase

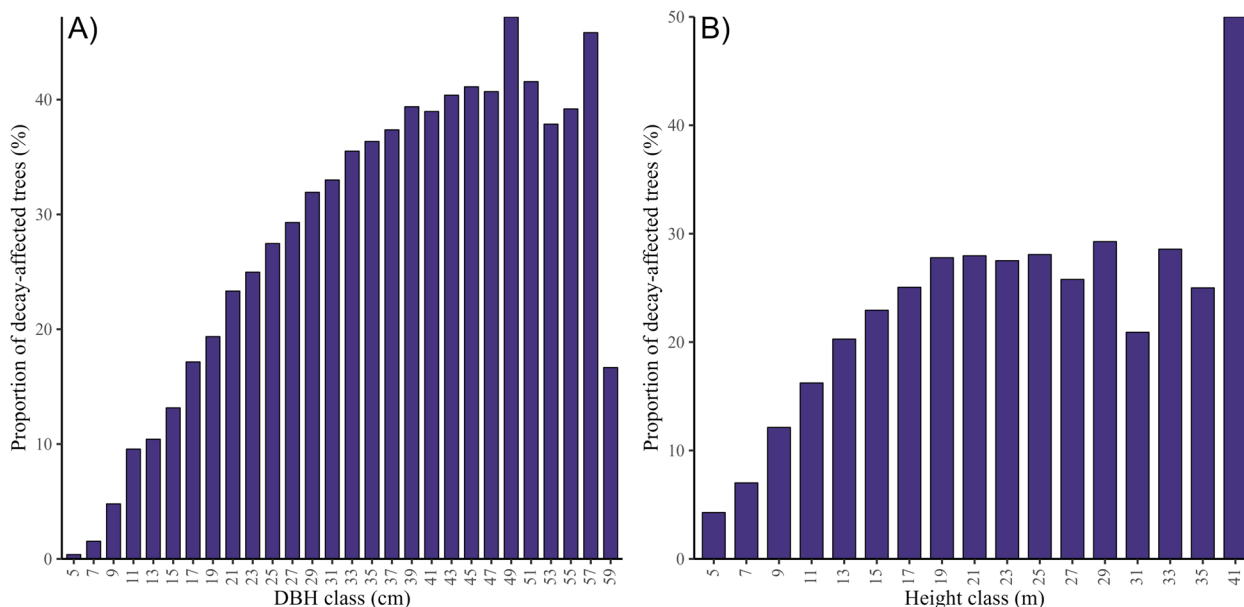
knowledge on stand-level timber decay damages in mature forests.

### 5 Conclusions

We draw the following conclusions based on our analyses carried out in mature spruce-dominated forests using cut-to-length harvester data:

- (1) An average-sized decay-affected tree is larger and taller compared with an average-sized tree in a forest stand.
- (2) Trees with incipient decay are on average larger and taller compared with trees with advanced decay.
- (3) The abovementioned findings reflect a relationship between the DBH distribution of all trees and the DBH distribution of decay-affected trees.
- (4) Distribution matching that transforms the DBH distribution of all trees into the DBH distribution of decay-affected trees can be used to learn the relationship between all and decay-affected trees and to predict the DBH distribution of decay-affected trees for forest stands.

### Appendix



**Fig. 9** Proportions of decay-affected trees by DBH (A) and height (B) classes in the harvester data. DBH, diameter at breast height

**Table 3** Parameter estimates of the model associated with the distribution matching ( $DM_{DBH}$ ). For more information on model formulation and parameters, please refer to Section 2.4

Predictors	Estimates	Std. error	p
Intercept	-7.885	1.254	< 0.001
$r$	1.176	0.007	< 0.001
$r^2$	-0.005	< 0.001	< 0.001
$q$	19.248	2.170	< 0.001
<b>Random effects</b>			
$\text{var}(z_i^{(1)})$	2.315 <sup>2</sup>		
$\text{var}(z_i^{(2)})$	0.066 <sup>2</sup>		
$\text{cor}(z_i^{(1)}, z_i^{(2)})$	-0.76		
$\sigma^2$	0.231 <sup>2</sup>		

**Acknowledgements**

We thank Simon Berg for technical support with the harvester data.

**Code availability**

The codes used to analyze the data were not disclosed.

**Authors' contributions**

Conceptualization, JR, AMH, JB, and RA; methodology, JR, JB, and AMH; writing — original draft preparation, JR and AMH; writing — review and editing, JR, AMH, JB, and RA; and funding acquisition, RA. The authors read and approved the final manuscript.

**Funding**

This work was supported by the Norwegian Research Council through the PRECISION project — precision forestry for improved resource use and reduced wood decay in Norwegian forests (NFR 281140).

**Availability of data and materials**

The dataset supporting the conclusions of this article is not publicly available due to the private ownership of the harvester datasets.

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup>Natural Resources Institute Finland (LUKE), Yliopistokatu 6, 80100 Joensuu, Finland. <sup>2</sup>Norwegian Institute of Bioeconomy Research (NIBIO), Høgskoleveien 8, 1433 Ås, Norway. <sup>3</sup>Norwegian Institute of Bioeconomy Research (NIBIO), Skolegata 22, Building P1, 7713 Steinkjer, Norway.

Received: 5 July 2022 Accepted: 9 December 2022

Published online: 05 January 2023

**References**

Arlinger J, Nordström M, Möller J (2012) StandForD 2010: Modern kommunikation med skogsmaskiner  
Astrup R, Rahlf J, Bjørkelo K et al (2019) Forest information at multiple scales: development, evaluation and application of the Norwegian forest

resources map SR16. *Scand J For Res* 34:484–496. <https://doi.org/10.1080/02827581.2019.1588989>

Baffetta F, Corona P, Fattorini L (2012) A matching procedure to improve k-NN estimation of forest attribute maps. *For Ecol Manag* 272:35–50. <https://doi.org/10.1016/j.foreco.2011.06.037>

Bendz-Hellgren M, Brandtberg P-O, Johansson M et al (1999) Growth rate of *Heterobasidion annosum* in *Picea abies* established on forest land and arable land. *Scand J For Res* 14:402–407. <https://doi.org/10.1080/02827589950154104>

Bendz-Hellgren M, Stenlid J (1995) Long-term reduction in the diameter growth of butt rot affected Norway spruce, *Picea abies*. *For Ecol Manag* 74:239–243

Blingsmo K (1985) Avsmalingsfunksjoner og -tabeller for bjørk. Rapport fra Norsk institutt for skogforskning 10:35

Caudullo G, Tinner W, de Rigo D (2016) *Picea abies* in Europe: distribution, habitat, usage and threats. In: San-Miguel-Ayanz J, de Rigo D, Caudullo G, Houston Durrant T, Mauri A (eds) *European Atlas of Forest Tree Species*. Publ. Off. EU, Luxembourg

Eide E (1954) Avsmalingstabell for granskog. *Utvidet og revidert 1954*, p 63

Gobakken T, Næsset E (2004) Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. *Scand J For Res* 19:529–542. <https://doi.org/10.1080/02827580410019454>

González RC, Woods RE (2002) *Digital image processing*, 3. ed., internat. ed. Prentice-Hall, Upper Saddle River

Hauglin M, Hansen E, Sørngård E et al (2018) Utilizing accurately positioned harvester data: modelling forest volume with airborne laser scanning. *Can J For Res*. <https://doi.org/10.1139/cjfr-2017-0467>

Hauglin M, Rahlf J, Schumacher J et al (2021) Large scale mapping of forest attributes using heterogeneous sets of airborne laser scanning and National Forest Inventory data. *For Ecosyst* 8:65. <https://doi.org/10.1186/s40663-021-00338-4>

Honkaniemi J, Ojansuu R, Piri T et al (2014) Hmodel, a *Heterobasidion annosum* model for even-aged Norway spruce stands. *Can J For Res* 44:796–809. <https://doi.org/10.1139/cjfr-2014-0011>

Huse K, Solheim H, Venn K (1994) Råte i gran registrert på stubber etter hogst vinteren 1992 [Stump inventory of root and butt rots in Norway spruce cut in 1992]. *Rapp Skogforsk* 23:1–26

Hyllen G, Granhus A (2018) A probability model for root and butt rot in *Picea abies* derived from Norwegian national forest inventory data. *Scand J For Res* 33:657–667. <https://doi.org/10.1080/02827581.2018.1487074>

Keča N, Solheim H (2011) Ecology and distribution of *Armillaria* species in Norway. *For Pathol* 41:120–132. <https://doi.org/10.1111/j.1439-0329.2010.00644.x>

Korhonen K, Stenlid J (1998) Biology of *Heterobasidion annosum*. In: Woodward S, Stenlid J, Huttermann A, Karjalainen R (eds) *Heterobasidion annosum: biology, ecology, impact, and control*. Cab International, Wallingford

Lindén M, Vollbrecht G (2002) Sensitivity of *Picea abies* to butt rot in pure stands and in mixed stands with *Pinus sylvestris* in southern Sweden. *Silva Fenn* 36:767–778

Mattila U, Nuutinen T (2007) Assessing the incidence of butt rot in Norway spruce in southern Finland. *Silva Fenn* 41. <https://doi.org/10.14214/sf.473>

Metsläid M, Granhus A, Scholten J et al (2018) Long-term effects of single-tree selection on the frequency and population structure of root and butt rot in uneven-sized Norway spruce stands. *For Ecol Manag* 409:509–517. <https://doi.org/10.1016/j.foreco.2017.11.050>

Möykkynen T, Pukkala T (2010) Optimizing the management of Norway spruce and Scots pine mixtures on a site infected by *Heterobasidion coll.* *Scand J For Res* 25:127–137. <https://doi.org/10.1080/02827581003667322>

Müller MM, Kaitera J, Henttonen HM (2018) Butt rot incidence in the northernmost distribution area of *Heterobasidion* in Finland. *For Ecol Manag* 425:154–163. <https://doi.org/10.1016/j.foreco.2018.05.036>

Müller MM, Sievänen R, Beuker E et al (2014) Predicting the activity of *Heterobasidion parviporum* on Norway spruce in warming climate from its respiration rate at different temperatures. *For Pathol* 44:325–336. <https://doi.org/10.1111/efp.12104>

Nagy NE, Norli HR, Fongen M et al (2022) Patterns and roles of lignan and terpenoid accumulation in the reaction zone compartmentalizing pathogen-infected heartwood of Norway spruce. *Planta* 255:1–16

Nordström M, Hemmingsson J (2018) Measure up! A Skogforsk guide to harvester measurement. Skogforsk, Uppsala

- Oliva J, Camarero JJ, Stenlid J (2012) Understanding the role of sapwood loss and reaction zone formation on radial growth of Norway spruce (*Picea abies*) trees decayed by *Heterobasidion annosum* s.l. *For Ecol Manag* 274:201–209. <https://doi.org/10.1016/j.foreco.2012.02.026>
- Pateiro-Lopez B, Rodriguez-Casal A (2019) alphahull: generalization of the convex hull of a sample of points in the plane
- Pinheiro J, Bates D, DebRoy S, et al (2020) nlme: linear and nonlinear mixed effects models
- Piri T (1996) The spreading of the S type of *Heterobasidion annosum* from Norway spruce stumps to the subsequent tree stand. *Eur J For Pathol* 26:193–204. <https://doi.org/10.1111/j.1439-0329.1996.tb00839.x>
- Piri T (2003) Early development of root rot in young Norway spruce planted on sites infected by *Heterobasidion* in southern Finland. *Can J For Res* 33:604–611. <https://doi.org/10.1139/x02-200>
- Piri T, Korhonen K (2001) Infection of advance regeneration of Norway spruce by *Heterobasidion parviporum*. *Can J For Res* 31:937–942. <https://doi.org/10.1139/cjfr-31-6-937>
- Prospero S, Rigling D, Holdenrieder O (2003) Population structure of *Armillaria* species in managed Norway spruce stands in the Alps. *New Phytol* 158:365–373. <https://doi.org/10.1046/j.1469-8137.2003.00731.x>
- Pukkala T, Möykkynen T, Thor M et al (2005) Modeling infection and spread of *Heterobasidion annosum* in even-aged Fennoscandian conifer stands. *Can J For Res* 35:74–84. <https://doi.org/10.1139/x04-150>
- R Core Team (2022) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Räty J, Breidenbach J, Hauglin M, Astrup R (2021) Prediction of butt rot volume in Norway spruce forest stands using harvester, remotely sensed and environmental data. *Int J Appl Earth Obs Geoinf* 105:102624. <https://doi.org/10.1016/j.jag.2021.102624>
- Reynolds M, Burk T, Huang W (1988) Goodness-of-fit tests and model selection procedures for diameter distribution models. *For Sci* 34:373–399
- Schönhar S (1978) Infektionsversuche mit *Fomes annosus*-Sporen an vitalen und absterbenden Fichtenwurzeln. *Allgemeine Forst- und Jagdzeitung* 149:73–74
- Siipilehto J, Mehtätalo L (2013) Parameter recovery vs. parameter prediction for the Weibull distribution validated for Scots pine stands in Finland. *Silva Fenn* 47. <https://doi.org/10.14214/sf.1057>
- Stenlid J, Redfern D (1998). In: Woodward S, Stenlid J, Karjalainen R, Hüttermann A (eds) Spread within the tree and stand, w: *Heterobasidion annosum*, biology, ecology, impact and control. Cab International, Wallingford, pp 125–141
- Strand L (1967) Avsmalingstabeller for furu sønnafjells. *Medd norske Skogforsves* 84:431–482
- Tamminen P (1985) Butt rot in Norway spruce in southern Finland. In: *Communicationes Instituti Forestalis Fenniae*, p 52
- Teimouri M (2021) ForestFit: statistical modelling for plant size distributions
- Thor M, Ståhl G, Stenlid J (2005) Modelling root rot incidence in Sweden using tree, site and stand variables. *Scand J For Res* 20:165–176. <https://doi.org/10.1080/02827580510008347>
- Vauhkonen J, Mehtätalo L (2015) Matching remotely sensed and field-measured tree size distributions. *Can J For Res* 45:353–363. <https://doi.org/10.1139/cjfr-2014-0285>
- Vollbrecht G, Agestam E (1995) Identifying butt rotted Norway spruce trees from external signs. *Forest and Landscape Research*, Denmark
- Woodward S, Stenlid J, Hüttermann A, Karjalainen R (1998) *Heterobasidion annosum*: biology, ecology, impact, and control. Cab International, Wallingford

## Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

