Earth Syst. Sci. Data, 14, 5695–5716, 2022
https://doi.org/10.5194/essd-14-5695-2022

Open Access Earth System Science Data

Forest structure and individual tree inventories of northeastern Siberia along climatic gradients

Timon Miesner[1,2], Ulrike Herzschuh[1,2,3], Luidmila A. Pestryakova[4], Mareike Wieczorek[1], Evgenii S. Zakharov[5,4], Alexei I. Kolmogorov[4], Paraskovya V. Davydova[4], and Stefan Kruse[1]

[1]Alfred Wegener Institute (AWI) Helmholtz Centre for Polar and Marine Research, Research Unit Potsdam, Potsdam, Germany
[2]Institute of Environmental Sciences and Geography, University of Potsdam, Potsdam-Golm, Germany
[3]Institute of Biochemistry and Biology, University of Potsdam, Potsdam-Golm, Germany
[4]Institute of Natural Sciences, North-Eastern Federal University, Yakutsk, Russia
[5]Institute for Biological Problems of Cryolithozone, Russian Academy of Sciences, Siberian Branch, Yakutsk, Russia

Correspondence: Timon Miesner (timon.miesner@awi.de)

Received: 16 May 2022 – Discussion started: 24 May 2022
Revised: 2 November 2022 – Accepted: 17 November 2022 – Published: 22 December 2022

Abstract. We compile a data set of forest surveys from expeditions to the northeast of the Russian Federation, in Krasnoyarsk Krai, the Republic of Sakha (Yakutia), and the Chukotka Autonomous Okrug (59–73° N, 97–169° E), performed between the years 2011 and 2021. The region is characterized by permafrost soils and forests dominated by larch (*Larix gmelinii* Rupr. and *Larix cajanderi* Mayr).

Our data set consists of a plot database describing 226 georeferenced vegetation survey plots and a tree database with information about all the trees on these plots. The tree database, consisting of two tables with the same column names, contains information on the height, species, and vitality of 40 289 trees. A subset of the trees was subject to a more detailed inventory, which recorded the stem diameter at base and at breast height, crown diameter, and height of the beginning of the crown.

We recorded heights up to 28.5 m (median 2.5 m) and stand densities up to 120 000 trees per hectare (median 1197 ha^{-1}), with both values tending to be higher in the more southerly areas. Observed taxa include *Larix* Mill., *Pinus* L., *Picea* A. Dietr., *Abies* Mill., *Salix* L., *Betula* L., *Populus* L., *Alnus* Mill., and *Ulmus* L.

In this study, we present the forest inventory data aggregated per plot. Additionally, we connect the data with different remote sensing data products to find out how accurately forest structure can be predicted from such products. Allometries were calculated to obtain the diameter from height measurements for every species group. For *Larix*, the most frequent of 10 species groups, allometries depended also on the stand density, as denser stands are characterized by thinner trees, relative to height. The remote sensing products used to compare against the inventory data include climate, forest biomass, canopy height, and forest loss or disturbance. We find that the forest metrics measured in the field can only be reconstructed from the remote sensing data to a limited extent, as they depend on local properties. This illustrates the need for ground inventories like those data we present here.

The data can be used for studying the forest structure of northeastern Siberia and for the calibration and validation of remotely sensed data. They are available at https://doi.org/10.1594/PANGAEA.943547 (Miesner et al., 2022).

1 Introduction

In total, 20 % of the world's forests are located in Russia (FAO, 2020), with the majority of these forests located in the sparsely populated north and east of the country. As the high latitudes are warming at a much faster rate than the global average, these forests are experiencing, and will face, further massive, abrupt changes (Scheffer et al., 2012). The threat of feedback loops to the global climate system (Bonan, 2008), possibly through the thawing of permafrost (Schuur et al., 2015) or changes in biosphere and soil carbon stocks (Walker et al., 2019), make it crucial to understand these ecosystems.

While the major portion of the world's boreal forests are made up of evergreen coniferous forest, northeastern Asia is dominated by summergreen coniferous trees of the species *Larix gmelinii* and *Larix cajanderi* (Abaimov, 2010). This vegetation type covers an area of several million square kilometres and stretches from the northern China in the south and the central Siberian Plateau in the west, where mixed stands occur with evergreen coniferous trees, to the northern treeline near the Arctic Ocean, where sparse forest tundra and stunted growth forms prevail (Wieczorek et al., 2017; Kruse et al., 2020a). Much of the geographical range is underlain by continuous permafrost (Osawa and Zyryanova, 2010). Recurrent forest fires also play a vital role in the ecosystem (Payette, 1992).

There has been no comprehensive forest inventory and planning in Russia in the post-Soviet era, and thus estimations on the volume of wood in the nation's forests vary widely (Schepaschenko et al., 2021). A national forest inventory, conducted between 2006 and 2020, aimed to shed light on this, but no definite results have been published as of May 2022. There are several studies that deal explicitly with larch-dominated ecosystems in Russia, for example, Kharuk et al. (2019) and Dolman et al. (2004) and the comprehensive volume by Osawa and Zyryanova (2010), but there are only a few that come with forest inventory data. The range of *Larix gmelinii* extends into the northernmost provinces of China, where it is used for afforestation. In this area, there has been much research on this species, e.g. Jia and Zhou (2018), Widagdo et al. (2020), and Xiao et al. (2020), but the properties of the species – and thus the ecosystems formed – vary widely, depending on growing conditions, which are a lot harsher in the northern parts of its range (Wang et al., 2005).

Remote sensing data can give insight into many forest-related parameters, such as aboveground biomass, growing stock volume, or canopy height (Simard et al., 2011; Santoro et al., 2018), and in the past decade, there has been a massive increase in detailed, freely available remote sensing data products. The ground-truthing that is necessary for such products tends to have a bias towards more accessible forest areas where previous forest surveys have been conducted (e.g. Yang and Kondoh, 2020). Another issue is that sparsely forested ecosystems at the tundra–taiga ecotone are often not understood as forests, e.g. by the influential Food and Agriculture Organization (FAO) definition (FAO, 2000), and therefore, they may be excluded from such data. Other aspects, such as the compositional complexity of forest in terms of height, age, and species distribution, are still difficult to capture from space, meaning that it is necessary to take on-site measurements in order to understand these ecosystems.

To meet this demand, joint Russian–German expeditions to Siberia have been conducted since 2011 to the Russian Federation subjects of Krasnoyarsk Krai, the Republic of Sakha (Yakutia), and the Chukotka Autonomous Okrug. In this study, we present the collected forest measurement data of the combined expeditions, both at the level of single trees and at the plot level, which can potentially be further upscaled. The central questions that motivate this study are as follows: what are the patterns of forest composition in northeast Asian larch ecosystems? How much growing stock of wood do they hold? How strong is the role of climate as a driver for these variables? How well do available remote sensing products describe what we see on the ground?

2 Methodology

2.1 Area of interest

The areas of interest are the larch-dominated forests in northeastern Asia, including the transition zones to the tundra and to evergreen deciduous forest (see Fig. 1). The area is characterized by permafrost soils and strongly continental climate (Kajimoto et al., 1999). Precipitation is generally below 300 $mm\,yr^{-1}$, although this is sometimes exceeded towards the boundaries of the area. Winter temperatures are mostly below −30 °C, while the warmest months average between 20 °C in central Yakutia to 8 °C near the Arctic Ocean (Fig. 2). The forests of the region are sparse and slow-growing. Recurring fires are an important driver for this ecological system (Kharuk et al., 2011).

2.2 Forest inventories

Eight summer expeditions were led to different destinations in the Russian Federation, i.e. to the tundra–taiga transition zone in 2011, 2012, 2013, 2014, 2016, and 2018, to the mountainous tundra treeline in 2016 and 2018, and to the boreal forest in 2018 and 2021 (Overduin et al., 2017; Kruse et al., 2019). The main goals varied between the expeditions but all included forest inventories using the same methodology. The expedition plots are not evenly distributed across the area, as the focus was on transition zones, especially the tundra–taiga ecotone at the northern limit of the range of *Larix*, and the transition to evergreen forests in the southwest of its range.

The sites at which the surveys were performed were chosen beforehand, with the consideration of remote sensing

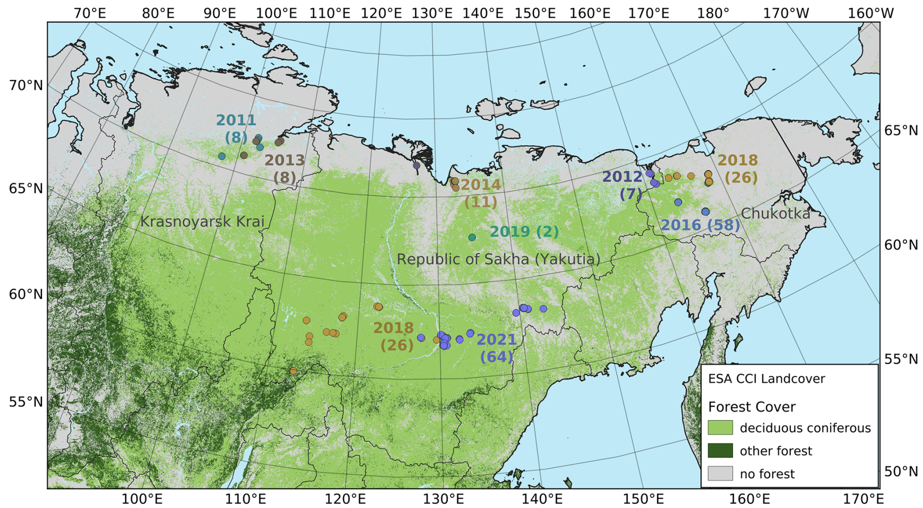

Figure 1. The vegetation in the larch-dominated forests of northeastern Russia. Numbers indicate the year and the number of vegetation plots on each expedition.

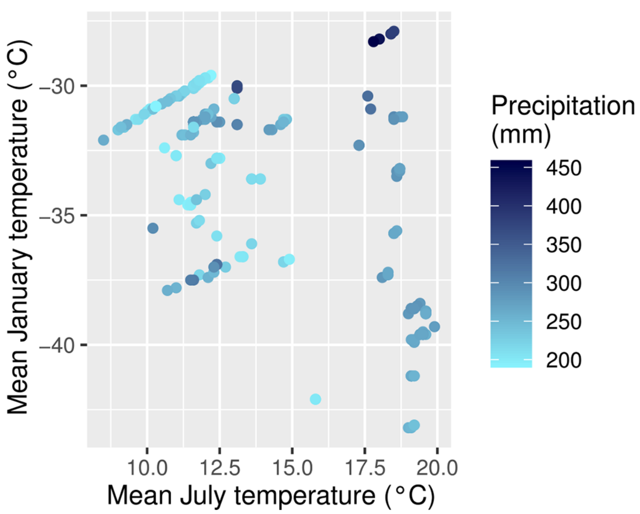

Figure 2. The climate on the plots according to CHELSA (Climatologies at high resolution for the Earth's land surface areas) data. Mean January temperature is shown on the y axis, mean July temperature is shown on the x axis, and mean annual precipitation as coloured dots.

data. The goal was to cover a wide range of conditions such as tree cover percentage and reflectance values in the region of each expedition. The exact positioning of the survey plot was finalized on site, with the aim to have each plot representing a homogeneous vegetation type. Not all vegetation survey plots contain forest or even single trees; some were used to record ground vegetation and tree recruitment, while taller trees were absent.

The geographic coordinates of the plot centre were recorded with a GPS device, using the datum WGS 84 (World Geodetic System). Plots were either rectangular or circular. Rectangular plots were more commonly used in the tundra–taiga ecotone. They would typically be squares of 20 m × 20 m, but their size was sometimes increased in areas with very few trees per hectare or decreased in size if the vegetation or topography demanded it. A grid of 2 m × 2 m was laid out over the plot in order to locate trees precisely inside of it. In a rectangular plot, every tree was recorded in detail and the following variables were noted: species, height, vitality estimate (on a discrete scale, from very vital, vital, mediocre, low, and very low to dead), growth form, basal diameter, diameter at breast height (DBH), maximum crown diameter, and the smaller crown diameter, which was measured perpendicular to the maximum.

Circular plots had a diameter of 15 m, except for occasions in which the forest was too dense to record all trees in this range; in these cases, the diameter was reduced to 10 m. Of the trees in the circular plot area, a minimum of 10 trees was chosen for the detailed inventory as above. The goal was to choose 10 trees per species so that they covered the entire range of height and diameter variation present on the plot. If there were more than two species, then the number of chosen trees per species was reduced due to time constraints, with the focus on coniferous trees. After making the detailed inventory of the chosen trees, all trees on the plot were recorded, noting only the species, estimated height, and general remarks, for example, whether the tree had low vitality, was dead, inclined, or not of upright growth form. Based on this, the variable growth type can take the values of tree (T), shrub (S), tree lying (TL for lying deadwood), multistem (M, if several shoots emerged from the same base), and krumholz (K). *Larix* can occur both in the tree form and in the krumholz form. The criterion for the latter is the lack of a straight, upright stem (Kruse et al., 2020b). The variable survey protocol tells us if the tree was recorded on a rectangular plot (PLOT), outside of a plot (EXTRA), or on a circular plot. In the latter case, the variable takes the value PLOTHEIGHT, if only height was measured, and CIRCLEPLOT, if it is the detailed inventory. Those trees were recorded twice, with different IDs, i.e. once as PLOTHEIGHT and once as CIRCLEPLOT. In order to avoid duplicate entries, the data set was later split up into two tables, so that, within each table, any tree would only be mentioned once.

Tree height was measured with a clinometer for some trees and, for others, visually estimated by making a comparison with the measured trees or objects of known height. According to experience, the error in this method was below 10 % for smaller trees or below 1 m for larger ones. Generally, all trees at least 40 cm in height were measured. Additionally, for many plots along the treeline, where recruitment was the focus of the research, smaller individuals were recorded on subplots. Stem diameters were measured either with a measuring tape (as circumference) or a calliper, recording the basal diameter just above the root collar and DBH at 1.3 m above the ground. Crown diameters were estimated from below, with the help of ground measurements using a measuring tape.

Parts of the data set presented here has already been published in other data publications and are available individually:

- Wieczorek et al. (2017) (including data of the 2011 and 2013 expeditions) at https://doi.org/10.1594/PANGAEA.874615,
- Kruse et al. (2020a) at https://doi.org/10.1594/PANGAEA.923638, and
- van Geffen et al. (2021) at https://doi.org/10.1594/PANGAEA.932821.

2.3 Processing of the data

In the two tables of the tree database, i.e. tree heights and tree measurements, every entry contains information about one tree. Some processing was done prior to the analysis to derive variables that were not present in the original data set. The entire list, displaying which variables were recorded directly on site and which were derived from other measurements, can be found in Appendix B.

The species of each tree was recorded differently, depending on the surveyor. This led to differences in the naming convention, for example, *Betula pendula* on some plots and *Betula spec.* on others. Therefore, the 23 taxa entries were harmonized into 10 species groups, as identified by the genus name. The species *Larix gmelinii* and *Larix cajanderi* were grouped together in the species group *Larix*. An exception is the genus *Pinus*, where *Pinus pumila* (Pall.) Regel was excluded from the *Pinus* group due to its shrub-like growth form.

Height was recorded for all trees, but diameters only for selected ones, and the existing diameters were used to calculate allometries from which the diameters were then reconstructed from the height for those trees where they were not measured. For each species group, a power function of the form

$$D_{\mathrm{BS}} = a_1 \cdot H^{a_2}$$

was fitted with the least squares method (where D_{BS} is the diameter at base, H is the height, and a_1 and a_2 are the optimization coefficients). For diameter at breast height (D_{BH}), the function is as follows:

$$D_{\mathrm{BH}} = a_1 \cdot (H - 1.3)^{a_2}.$$

Initial analyses with this function revealed that the diameter estimations were biased on some plots. On densely forested plots, trees tended to have smaller diameters at the same height compared to sparsely forested plots, especially in the lower half of the height range. As the power functions computed for the different stand density groups (measured in trees per hectare) differed both in exponent and in factor, we used the adjusted power function,

$$D_{\mathrm{BS}} = (a_1 + a_3 \cdot S) \cdot H^{(a_2 + a_4 \cdot S)},$$

where the stand density S was computed from T_{ha}, the number of trees per hectare, as follows:

$$S = \max(\log_{10}(T_{\mathrm{ha}}), 2)).$$

The formula for D_{BH} was analogous, replacing H with $(H - 1.3)$. The latter formulas were only applied to the species group *Larix*, as all other species were not present on enough different plots to prevent overfitting. For all other species groups, the former, simpler formulas were used.

Having thus obtained the variables for the predicted diameter at base (DBS) and predicted DBH for all trees, it was possible to calculate further metrics, including the basal area (BA),

$$\mathrm{BA} = \frac{\pi}{4} {D_{\mathrm{BH}}}^2,$$

and stem volume (V), which was obtained using the Smalian volume formula (Cailliez and Alder, 1980) for trees taller than breast height,

$$V = \frac{{D_{\mathrm{BS}}}^2 + {D_{\mathrm{BH}}}^2}{2} \cdot \frac{\pi}{4} \cdot 1.3 + \frac{{D_{\mathrm{BH}}}^2}{2} \cdot \frac{\pi}{4} \cdot (H - 1.3),$$

and for trees smaller than breast height, respectively, and

$$V = \frac{{D_{\mathrm{BS}}}^2}{2} \cdot \frac{\pi}{4} \cdot H.$$

After calculating these variables for the individual trees, the database was split up into two tables to avoid a situation in which individual trees would appear twice in the same table under two different survey protocols. The table of tree heights includes all standing trees inside of the plots with heights from 0.4 m upwards. From the circular plots, the table includes only the PLOTHEIGHT measurements. The table of the tree measurements includes all other entries, including all CIRCLEPLOT and PLOT measurements, even though the PLOT measurements are also included in the table of tree heights. This was done so that all entries with diameter measurements would be included in the tree measurements table. Of the information in the tree heights table, several variables were aggregated at the plot level by calculating mean and selected quantiles of height in addition to the sum of the basal area and stem volume. The latter variables were then divided by the plot area to obtain the respective values per hectare.

Another measure we calculated for the height distributions of each plot is the Gini coefficient (Gini, 1912). It ranges between 0 and 1, assuming a value of 0 if all trees have the same height and approaching 1 if there are a few very big trees alongside many very small ones. Let h_i be a collection of height measurements in ascending order and i in $\{1,\ldots,n\}$. Then, the Gini coefficient is defined as follows:

$$1-2\frac{\sum_{i=1}^{n}(h_i\cdot(n-i+0.5))}{\sum_{i=1}^{n}(h_i\cdot n)}.$$

2.4 Data products for comparison

In the study, we used several gridded, mostly remote-sensing-derived data products on climate, biomass, height, forest cover loss, and stand age to compare with and relate to the forest inventory.

2.4.1 Climate

CHELSA (Climatologies at high resolution for the Earth's land surface areas; Karger et al., 2017, 2021) is a global raster data set containing many different variables, like monthly mean temperatures and precipitation sums, and several different bioclimatic variables. This study uses the temperature means for the months of January and July, the sum of the monthly precipitation, and the growing degree days above 0 °C (GDD0).

All values are means for the period 1981–2010, with a spatial resolution of 30 degree seconds, which is less than 1 km.

2.4.2 Forest biomass

The GlobBiomass data set (Santoro et al., 2018, 2021) covers the Earth's land surface with a pixel size of 1 ha. It provides values for the aboveground biomass (AGB) and growing stock volume (GSV) for the year 2010, in addition to the standard errors derived from satellite-based synthetic aperture radar, and an extensive set of ground measurements. The authors note that their data set is not precise at the pixel level but is better when aggregated over larger areas.

2.4.3 Forest height

The forest canopy height product (Simard et al., 2011) is a raster data set with a resolution of 1 km^2. It estimates the maximum canopy height in each pixel from the Geoscience Laser Altimeter System (GLAS) satellite-borne lidar, using additional data about climate, elevation, and canopy cover. All values are for the year 2005.

2.4.4 Tree cover loss

We used the tree cover loss product from the Global Forest Watch project (Hansen et al., 2013), which is based on yearly observations of Landsat images (30 m resolution). The project publishes various related data sets, e.g. a product about forest cover gain, and most products are updated regularly. The tree cover loss product detects, for each pixel, if it has been converted from containing tree cover (yes/no) to not containing tree cover in the time from 2000 to 2019. It assigns the year of the loss to a given pixel or 0 if no loss has taken place since the year 2000.

2.4.5 Siberian larch stand age

The Distribution of Estimated Stand Age Across Siberian Larch Forests (Chen et al., 2017) is related to the former data set and is also mainly based on Landsat images with 30 m resolution. It incorporates some more analysis to detect stand-replacing forest fires, but it only covers a part of eastern Siberia, including 54 of our vegetation survey plots, and spans the years 1989–2012. For every pixel, it gives the age of the forest stand as if it has experienced a stand replacing fire since 1989, a value of 100 if there has been no fire between 1989–2012, or no data if the pixel does not contain larch forest.

2.5 Analysis methods

The remote sensing products that were used all consisted of raster data. The values at the locations of the plot centres were extracted using QGIS 3.16 (QGIS, 2021).

From the CHELSA climate data set, four variables were chosen for further analyses, namely annual precipitation sum (Prec.), January mean temperature (T01), July mean temperature (T07), and growing degree days above 0 °C (GDD0). Univariate linear regressions were calculated between every single variable and four forest inventory variables.

To compare the GlobBiomass product and the forest height product with our data, linear regressions were calculated between the remote-sensing-derived variables and suitable variables of our forest plot data, like stem volume.

We compared the quotient of the living basal area over the total basal area for plots with recent tree cover loss and plots without recent tree cover loss as assessed by a two-sided t test.

All analysis was performed in R 4.1.0 (R Core Team, 2021).

3 Results

3.1 Description of the data

3.1.1 Descriptive statistics

The tree database comprises 42 675 entries, describing 40 289 trees. This is due to the fact that, on circular plots, the trees that were subject to detailed inventory are also recorded again in the height-only inventory. Of these, 33 513 individuals were used for aggregation at the plot level. The rest were excluded for being smaller than 40 cm, because such trees were not recorded on every plot, or for being located outside of the vegetation plots listed in the plot database.

The plot database includes 226 vegetation plots, of which 162 contain trees taller or equal to 40 cm. Of the 40 289 trees, 4660 (11.6 %) were dead and 35 629 (88.4 %) living at the time of recording. All entries in the tree database have a recorded height, which ranges up to 28.5 m. The species is recorded for all but 31 entries. The most frequent species are *Larix cajanderi* (44.4 % of database) and *Larix gmelinii* (25.7 %). The two *Larix* species never occur together on the same plot. Other frequent taxa are *Betula pendula* Roth (13.9 %), *Picea obovata* Ledeb. (5.8 %), *Pinus sylvestris* L. (5.0%), and the genus *Salix spec.* (3.2 %). Among the less frequent taxa are *Populus tremula* L., *Alnus spec.*, *Pinus pumila* Regel, *Pinus sibirica* Du Tour, and *Abies sibirica* Ledeb.

Values for basal diameter are present for 2583 entries. They range from 0 to 97.7 cm, with the median at 6.99 cm and mean at 11.08 cm. For diameter at breast height (DBH), there are 2095 values in the data set, almost all of which are trees for which the basal diameter is also given. DBH is almost always lower than basal diameter, on average by the factor 0.628. DBH ranges up to 71.6 cm, with the median at 6.4 cm and mean at 9.02 cm. Maximum crown diameter and smaller crown diameter (measured perpendicular to maximum) are given for 2079 entries and range from 0 to 16 m. The quotient of the two diameters is, on average, 0.81. Tree crown area, which is the product of the two values and the factor $\frac{\pi}{4} \cdot \frac{1\,\mathrm{m}^2}{10\,000\,\mathrm{cm}^2}$ is, on average, 4.77 m^2, with a median of 1.43 m^2.

3.1.2 Diameter–height allometry

The power function allometries for the different species differ notably, as can be seen in Fig. 3. The basal diameter of birches (*Betula*), for example, is obtained from height with an exponent of $a_1 = 1.15$ and a factor of $a_2 = 0.91$, while for *Abies*, the exponent is $a_1 = 0.66$ and the factor is $a_2 = 2.69$. The genus *Populus* differs strongly from the other species groups, with an exponent of $a_1 = 2.29$ and a factor of $a_2 = 0.06$. In the DBH-model *Populus* differs remarkably from the others, too, even if not that strongly. All factors and exponents are displayed in Appendix C.

The graphs in Fig. 3k and u show the diameter–height allometries for the genus *Larix* when taking into account the number of trees per hectare. When tree measurements are grouped by stand density, the resulting power functions differ by more than the respective standard errors for the coefficients, especially for heights between 4 and 12 m, where a higher number of trees on the plot have smaller diameters.

3.1.3 Height distributions

Tree heights show a nearly exponential distribution, with the exception that values from approximately 15 m upward occur slightly more frequently than expected under an exponential distribution (Fig. 4). However, at the level of individual plots, the distribution patterns vary widely. This can be seen in Fig. 5. Although the tree heights on plot EN21-260 are close to an exponential distribution, suggesting a continuous recruitment rate, in EN21-253 the larger trees are overrepresented. Plot EN21-230 is missing the smallest cohort, and plot EN21-246 is an example of dense regrowth after a stand-replacing fire, where older trees taller than 7 m are absent. Plot EN21-226 is dominated by a cohort of middle-sized trees, lacking both small and very large ones. In EN21-219, some large and many small individuals are present, while medium-sized ones are missing.

The Gini coefficient is normally distributed with a mean of 0.363 and standard deviation of 0.123. Plot EN21-258 is an example of a plot with a high Gini value (0.679), and plot EN21-226 is at the lower end, with a Gini coefficient of 0.166. The Gini coefficients are negatively correlated with the geographic latitude of the plot (Fig. 5). The linear regression has a p value of 0.021, and $R^2 = 0.33$.

3.1.4 Species distribution

In accordance with the known ranges of the different species, we observe that species diversity tends to be higher on the plots in central and western Yakutia, which experience warmer summers and longer growing seasons than the plots near the northern treeline. All plots north of 70° N have only one species (larch), while, for the plots south of 65° N, there are, on average, 3.11 species, with a maximum of 9 tree species from 7 species groups.

The species *Pinus sylvestris, Picea obovata, Abies sibirica, Ulmus spec.*, and *Populus tremula* only occur on the plots south of 65° N, with a July temperature of at least 17 °C. More predominant among the southerly plots are *Betula pendula* and *Alnus spec.*, but they are also found at one and three plots, respectively, in Chukotka. The taxa *Pinus pumila* and *Salix spec.* occur frequently between 65 and 70° N. Of the plots with trees, all but one have *Larix* individuals. For the

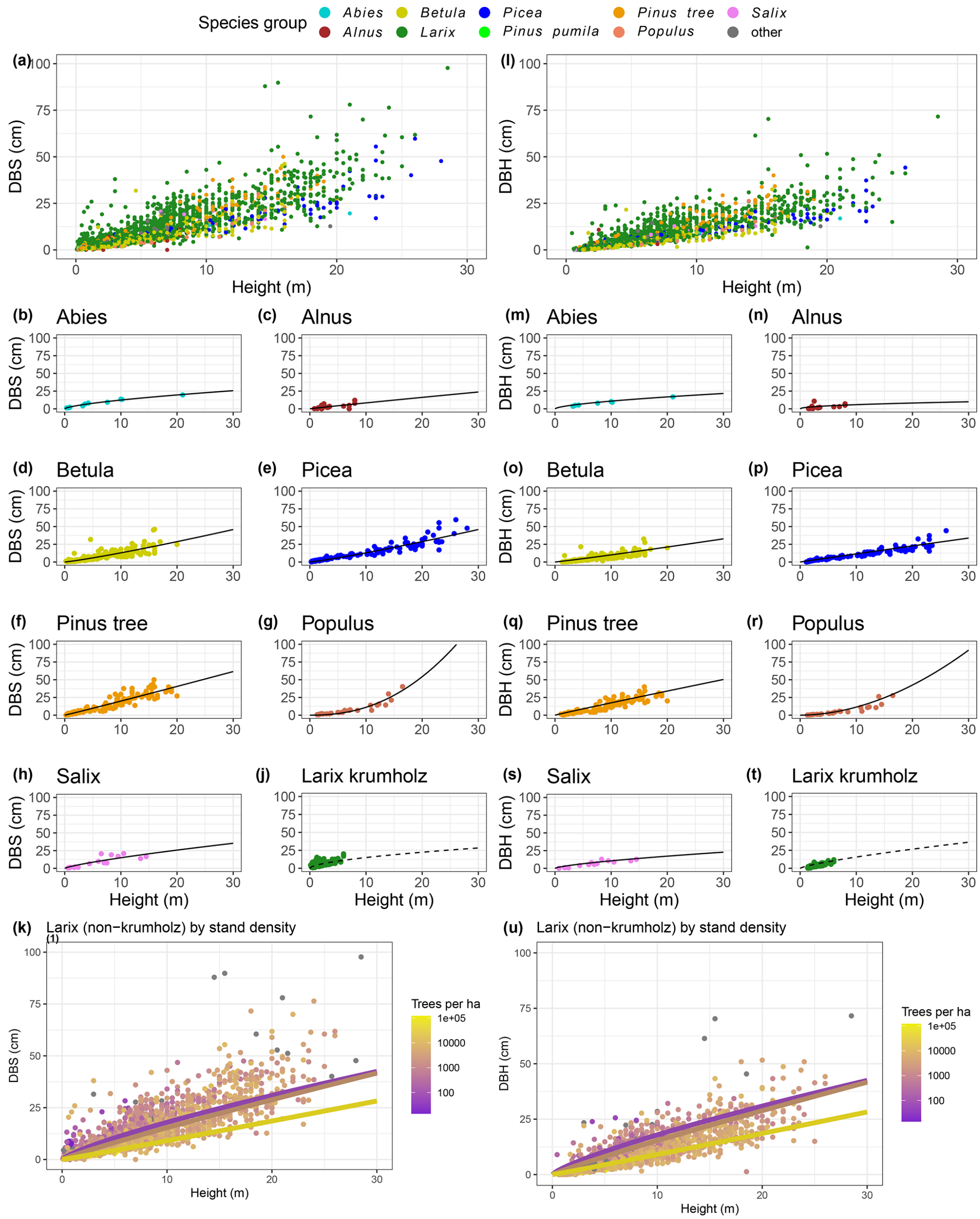

Figure 3. Diameter at base (DBS; left) and diameter at breast height (DBH; right) against the height per species. Power function allometries per species are shown. Panels **(k)** and **(u)** show *Larix* only, coloured by trees per hectare. The regression lines illustrate the allometry for three different stand densities (300, 3000, and 30 000 trees per hectare), while, in the actual allometric formula, the stand density is a continuous variable.

plots to the west of 130° E, it is *L. gmelinii*, and for the plots east thereof, it is *L. cajanderi*.

3.2 Remote sensing products as predictors

3.2.1 CHELSA Climate

The climate on the plots is strongly continental (see also Fig. 2), with mild to warm summers and extremely cold winters. The length of the growing season is between 63 and 132 d, and the GDD0 ranges from 565 to 1974.

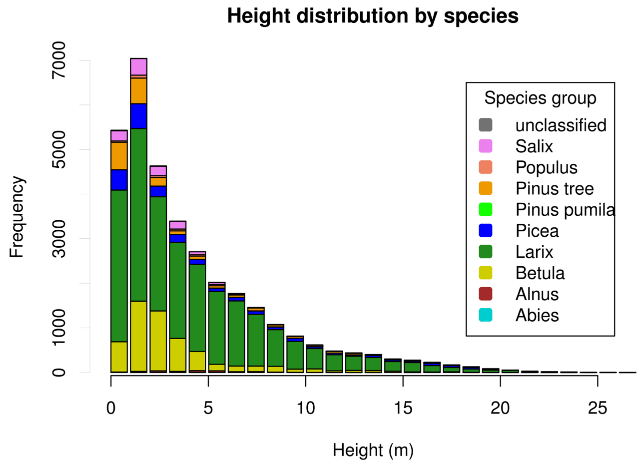

Figure 4. Height distribution of all trees on plots.

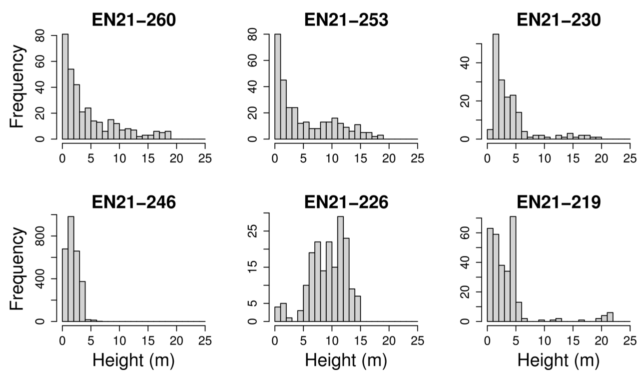

Figure 5. The height classes among all species for six different plots of the Yakutia 2021 expedition, which were chosen as examples for differing height distributions.

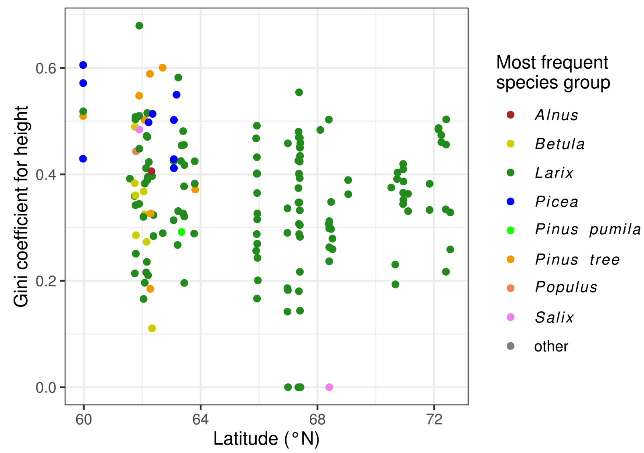

Figure 6. Gini coefficients for height against latitude, coloured by most frequent species on plot. (Plots with more diverse height distributions have higher Gini coefficients.)

Weak correlations between four climate parameters (precipitation, January temperature, July temperature, and GDD0) and four forest structure parameters (mean height, $\log_{10}$ (number of trees per hectare), basal area per hectare, and stem volume per hectare) are found (Fig. 7). The climate variables mean that the January temperature (T01) and precipitation have very low correlation coefficients with all forest metrics. The correlations between T01 and the forest metrics are even negative, although R^2 values are close to 0. Mean July temperature (T07) and GDD0 are more strongly correlated with several forest structure parameters, but the strength of the correlation is only intermediate and does not exceed $R^2 = 0.321$ in any combination. The predictor variables are also correlated among each other, especially T07 and GDD0 ($R^2=0.993$; all correlations are shown in Appendix D).

3.2.2 GlobBiomass

The two leading variables from the GlobBiomass data set – aboveground biomass (AGB) and growing stock volume (GSV) – are themselves strongly correlated ($R^2=0.989$ over all plots). Therefore, we focus on just one of them – GSV – which can be derived from our data with more confidence, since we did not measure the wood density and biomass expansion factors.

Remote-sensing-derived GSV and inventory-derived GSV follow the same tendency (correlation with $R^2=0.49$ and residual standard error 79.9; Fig. 8). But, for some plots, the two values differ by more than an order of magnitude.

3.2.3 Forest height

The values of the Simard et al. (2011) data are 0 (no forest) or have integers between 11 and 27 for the forest height in metres. On 125 of the plots, they record a value of 0, while we actually encountered trees on 60 of these plots in our inventory. A linear correlation between the Simard et al. (2011) canopy height and the maximum tree height on the plot (Fig. 9) has an intercept of 8.55, a slope of 0.298, and $R^2=0.20$. Other metrics, such as the 98th, 90th, or 75th percentiles of the observed tree height, have even less correlation (see Appendix E).

3.2.4 Forest loss

The data set for the Stand Age of Siberian Larch Forests by Chen et al. (2017) has data for 54 of our vegetation plots and finds that 6 plots have experienced stand-replacing events between 1989 and 2012. The Hansen et al. (2013) data set covers a wider area and different time range. However, there are five plots for which they detect forest loss in times and places where Chen et al. (2017) find that the stand age is at maximum. We encountered clear signs of recent disturbance in the vegetation at only 50 % of the plots where either of the data products detected forest loss.

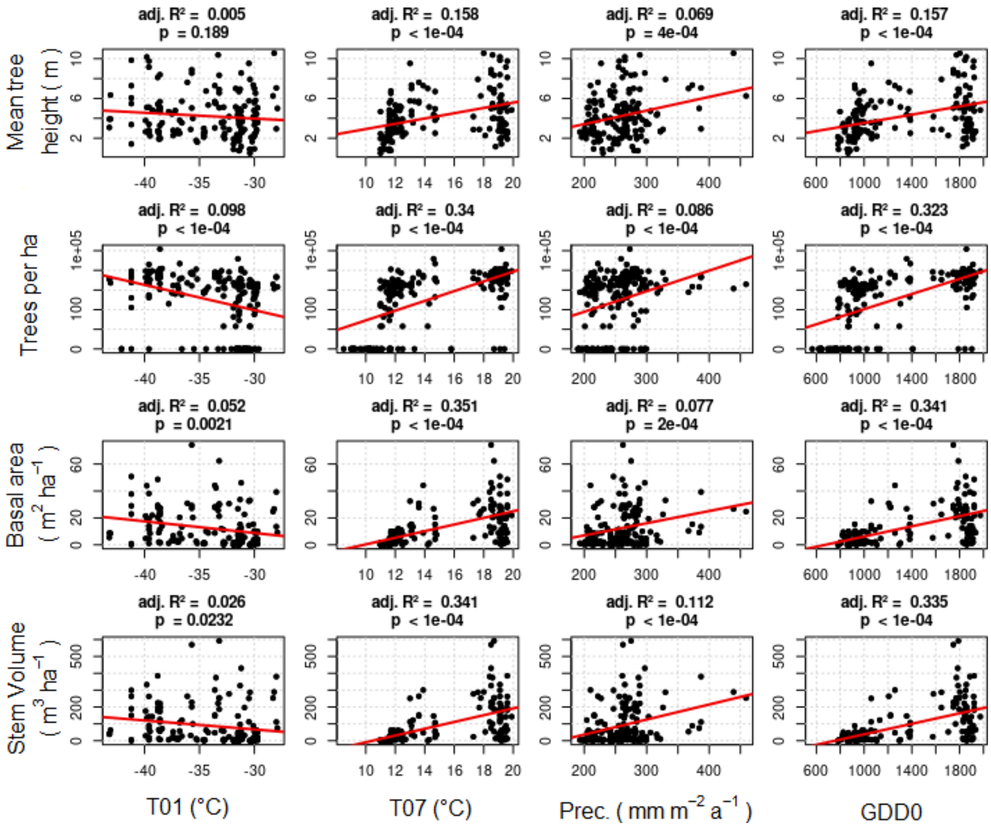

Figure 7. Comparison of forest inventory variables with climate variables. Linear regression lines are given in red.

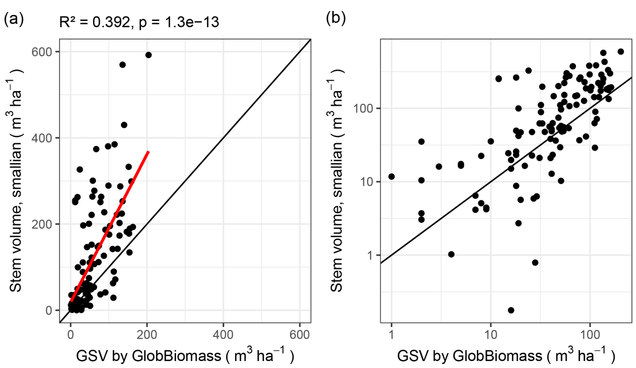

Figure 8. Stem volume calculations plotted against growing stock volume (GSV) from the GlobBiomass data set. **(a)** Linear scale. **(b)** Logarithmic scale, with zeroes removed.

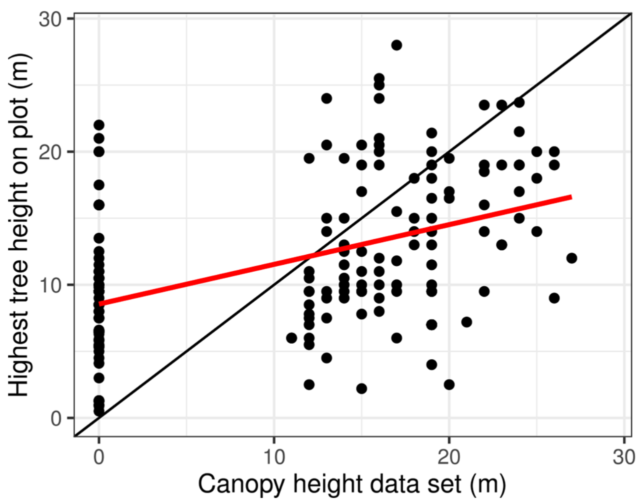

Figure 9. Highest tree of the inventory plots plotted against canopy height, according to the Simard et al. (2011) data set. Linear regression line in red.

The average quotient of the basal area of living trees to overall basal area is higher for the plots without disturbance than for the plots with forest loss, according to the Hansen et al. (2013) data set, which shows that there is more standing deadwood on plots with forest loss (Fig. 10). Although a t test finds that the two groups differ very significantly ($p = 4 \times 10^{-6}$), we see that there are also individual disturbance plots in which dead trees do not constitute a relevant number for the basal area. On most of these, field observations did not find signs of recent disturbance, except for one plot where the natural succession was at a pioneer stage.

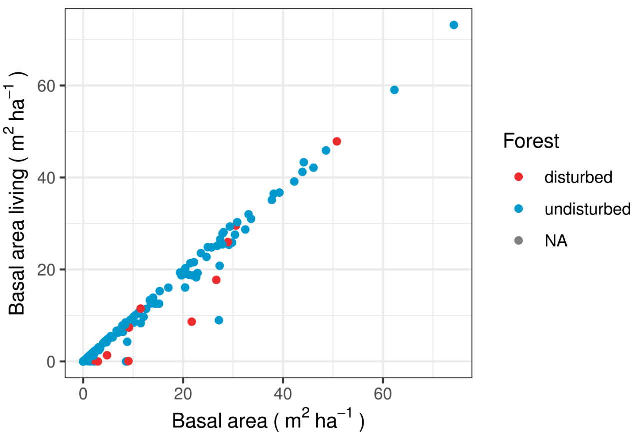

Figure 10. Living wood volume compared to overall wood volume. Plots with recent forest loss are marked red.

4 Discussion

4.1 Relevance of the data set

The data we present in this study are unique in their extent for the regions they cover. Schepaschenko et al. (2017) have compiled a vast number of forest inventories in Eurasia, but their coverage of our study region is sparse. For example, they include no data from Chukotka and the Kolyma area, where our data set has 91 plots. The same is true for the validation data set used by Yang and Kondoh (2020), who have only one location within our area of interest from the more than 400 literature sources they reviewed. This shows the lack of forest inventories from northeastern Siberia, which our data set aims to mend.

4.2 Validity of methods

The fieldwork was carried out according to scientific standards. Tree height was chosen as the leading variable because it is easy to have an overview of sparse stands, and it generally correlates well with other variables (stem diameter and biomass). Diameter at breast height (DBH), even though it is more commonly used as a predictor, is more laborious to determine for trees in sparse stands with low crowns. With frequent clinometer measurements, we assured precise height estimations, and the remaining errors can be expected to average out over the high number of observations, which were easily obtained due to the efficiency of the method. Drawbacks coming with this method are explained in the following. Since the diameter is only predicted from height, errors from this prediction propagate into derived variables like basal area and stem volume. And the initial measurement error, even if small, propagates along the same way. This error was not quantified systematically.

The correlations of the forest metrics with climate variables (Sect. 3.2.1) cannot be generalized because the distribution of the plots is not representative of the area. Even though the survey plots in each region cover the entire range of vegetation in any given zone, they are not weighed according to the occurrence of the vegetation type they represent. However, the relationships can still give us some idea of the general behaviour of the variables.

4.3 Tree species and heights distribution

We observe a higher species diversity in the more southerly stands which experience longer, warmer growing seasons. This is in accordance with the expectations and the known ranges of the observed tree species (Kuznetsova et al., 2010).

It is uncommon in the literature to record height distributions, but methodological analogues are age class or diameter distributions, which can be used to show recruitment patterns, e.g. Lin et al. (2005). While the close-to-exponential distribution of tree height suggests a continuous recruitment rate and continuous mortality throughout the age classes, a closer look at individual tree stands shows that they differ strongly from each other. This suggests that recruitment patterns are only continuous at the landscape scale, but discontinuous at the local scale, which is consistent with the well-known fact that stand-replacing fires regularly rejuvenate forests in the permafrost ecosystems of our research area (Kharuk et al., 2011).

4.4 Allometries

We see that the tree species have very different allometries. This may be partially due to the fact that they are actually different and partially due to the random effects of the plots and the small sample sizes for some species groups, like *Abies* (10 measurements for DBS) and *Populus* (27). The species groups with more than 100 measurements (*Betula, Larix, Picea,* and *Pinus*) have smaller differences among each other in the allometry coefficients. There is little literature with which to compare our results because the diameter, and not height, is commonly used as a predictor variable, as in Alexander et al. (2012) and Delcourt and Veraverbeke (2022), who both model biomass from diameter. We still chose to use height as the principal variable, as it is very easy to estimate in sparse forest stands. Nevertheless, using height as a predictor, Kajimoto et al. (1999) find a similar exponent for the *Larix gmelinii* stem weight as we found for volume.

4.5 Comparison of inventory and remote sensing

We find, for the examined remote sensing products, that predicting forest statistics on the plot base results in large errors. There are various factors that can lead to such a mismatch, as discussed by Houghton et al. (2007). Imprecision in the field measurements or the data processing may play a role (Picard et al., 2015). But likely another relevant factor is the coarse resolution of the remote sensing data, alongside the heterogeneity of the landscape on the scale between plot size and pixel size. The Simard et al. (2011) canopy height product, for example, has a resolution of 1 km^2, which is more than 1000 times our average vegetation plot size. Therefore, it cannot capture differences in canopy height below the kilometre scale, even though many landscape elements are smaller than this. This mismatch in resolution becomes especially relevant in the forest tundra, where the sparsity of the stands makes them difficult to detect in satellite images (Ranson et al., 2004; Montesano et al., 2016).

Another issue may be the lack of calibration of the remote sensing data sets, especially in the poorly researched area of northeastern Siberia. Zhang et al. (2019), who investigated numerous remote-sensing-based forest data sets, suggest that most of them suffer from a lack of validation and ground-truthing. Furthermore, Yang and Kondoh (2020) investigated the Simard et al. (2011) data set, and they find that it generally overestimates small canopy heights and underestimates large ones. When assessing the reliability of their biomass data product, Santoro et al. (2018, 2021) note that the relative AGB standard error in eastern Siberia is among the highest in the world, indicating a large uncertainty for this region.

A different source of error is the temporal mismatch between the acquisition of the inventory data and the remote sensing images. This varies throughout our data set, as the expeditions span a time range of 10 years, which is not accounted for in the comparisons, except for the comparison with the forest loss data sets. However, in the time ranges considered here, we can assume that the differences in variables such as stand height and growing stock volume are small, due to the very low growth rates of the forests in the region (Kajimoto et al., 2010). Only disturbances, such as wildfires and insect pests, could create large changes in the growing stock in a relatively short time.

We expect that all forest loss in our area is due to fire, as we did not find any signs of deforestation due to human activities on any of the surveyed plots. While the analysis of the forest loss data set led to the expected result that the plots with recent forest loss tend to have lower fractions of living basal area, it is still surprising that we saw some plots that were supposedly affected by forest loss, and thus by fire, with a large part of the stand alive, both in absolute and relative terms. This may be because many forest fires in Siberia are low-intensity fires (Ponomarev et al., 2022) which are detected as burned forest in 1 year, even though a large part of the trees recovers by the following year. Revisiting some of our survey plots in the future may help to improve the understanding of this topic.

4.6 The influence of climate on forest metrics

We find that the climate explains many of the quantitative forest metrics, albeit to a limited extent. Forest metrics such as basal area and stem volume are positively correlated with summer temperatures and growing degree days. However, the observed correlations are quite weak, and the range of the forest metrics is large. This suggests that the forest we observed is spatially heterogeneous and depends on properties which vary on smaller spatial scales than the climate.

It is counterintuitive that the investigated forest metrics are negatively correlated with January temperature in our data set, but it can be explained by the January temperature being negatively correlated with the July temperature (see Appendix D) and length of the growing season ($R^2 = 0.31$; slope is -0.107), which is another bioclimatic variable from the CHELSA data set. The plots near the Arctic Ocean have a less continental climate, meaning that they tend to have both milder winters and cooler summers than the more southerly ones. Thus, we cannot conclude that colder winters are favourable for forest growth.

There is scarcely any correlation between our observed forest metrics and precipitation, which suggests that water availability is not a limiting factor for forest growth in northeastern Siberia. Sugimoto et al. (2002) support this hypothesis by pointing out that larch forests in these regions have a good supply of water from snowmelt, rain, or thawing permafrost, depending on the weather in any given year. Opposed to this, Kharuk et al. (2019), who investigated a larch forest on the central Siberian Plateau, report that, since the 1990s, growth has been diminished by drought stress and extreme events, which are increasing under climate warming, like the 2020 Siberian heat wave (Collow et al., 2022). Kropp et al. (2017) and Walker et al. (2021) support findings that water availability is a limiting factor for *Larix cajanderi*.

4.7 Outlook

The analyses performed in this study do not exhaust the possibilities offered by this data set and serve purely to present the data. The fact that individual trees were measured, and related to the inventory plots, make it a very versatile data set. Some variables that were taken in the inventory can be analysed further. In particular, the crown diameters and crown base have not been assessed as yet. The forest inventory could be related to other, still unpublished data collections from the same expeditions, such as the projective crown cover estimations, ground vegetation surveys, soil profiles, genetic samples, stem increment cores, and stem discs. These additional samples were not collected for all individuals, but they could at least be related to a portion of the forest inventory data. Also, for some of the more recent expeditions, drone-based photogrammetric and lidar point clouds exist (e.g. SiDroForest) and could provide insight into the heterogeneity of the landscape and bridge the gap between survey plot size and pixel size of satellite-derived data. Furthermore, these centimetre-resolution point clouds are capable of capturing single tree measurements and bringing them to the landscape level. A different way to fill this gap and improve the predictions of the state of remote forests is with remote sensing products at a higher resolution, such as the boreal forest canopy height data set in connection with Potapov et al. (2020). They published a global canopy height data set with 30 m resolution for the tropical and temperate zones of the world, and the data for the boreal regions are expected to be released soon.

Our data set can also be used to calibrate and improve current and future remote sensing products. For this purpose, researchers can rely on the individual tree measurements, such as height, and on metrics aggregated at the plot level. The data set can serve to calculate or improve allometries for the investigated taxa, especially the two eastern Siberian larch species of *Larix cajanderi* and *Larix gmelinii*.

5 Data availability

The data are available at https://doi.org/10.1594/PANGAEA.943547 (Miesner et al., 2022).

6 Conclusions

We presented and analysed a data set resulting from forest inventories in various regions of northeast Siberia. A subset of the entries includes diameter measurements and height measurements, whereas the majority only includes height. Therefore, we computed diameter–height allometries, which are reasonably accurate overall but show a bias for some plots. It proved difficult to predict forest metrics at the plot level, for example, stem volume and basal area, from a selection of remote sensing products, as these were not strongly correlated. Among the climatic variables taken from the CHELSA data set, the mean July temperature is one of the best predictors, along with GDD0 and length of growing season, while the mean January temperature and precipitation proved almost insignificant. The GlobBiomass data set and the Simard et al. (2011) forest height product are correlated with the volume and height measurements on the survey plots but unsuitable for predicting the latter on a small scale. The data sets used for forest age and disturbance often differ from both each other and the observations made in the field. This leads us to conclude that, even in our time of widely available global remote sensing data sets, field measurements like the ones presented here are still vital for the understanding of remote ecosystems such as the larch-dominated forests of northeast Siberia.

Appendix A: Overview over all vegetation plots

Table A1. Overview over all vegetation plots. NA: not available. n/a: not applicable.

Site	Expedition	Latitude (° N)	Longitude (° E)	Area (m^2)	No. of trees	Most frequent species group
11-CH-02II	2011_Khatanga	71.83993	102.88387	400	88	*Larix*
11-CH-02III	2011_Khatanga	71.84179	102.87589	400	93	*Larix*
11-CH-06I	2011_Khatanga	70.66915	97.7121	400	31	*Larix*
11-CH-06III	2011_Khatanga	70.66498	97.7064	400	59	*Larix*
11-CH-12I	2011_Khatanga	72.3938	102.30144	2800	99	*Larix*
11-CH-12II	2011_Khatanga	72.40009	102.28725	9900	300	*Larix*
11-CH-17I	2011_Khatanga	72.24235	102.24565	480	101	*Larix*
11-CH-17II	2011_Khatanga	72.24144	102.22661	400	67	*Larix*
12-KO-02/I	2012_Kytalyk_Kolyma	68.38916	161.466171	400	219	*Larix*
12-KO-02/II	2012_Kytalyk_Kolyma	68.389936	161.448985	280	122	*Larix*
12-KO-03/I	2012_Kytalyk_Kolyma	68.516169	161.18194	320	258	*Larix*
12-KO-03/II	2012_Kytalyk_Kolyma	68.513173	161.195505	256	174	*Larix*
12-KO-04/I	2012_Kytalyk_Kolyma	69.051323	161.206493	400	118	*Larix*
12-KO-04/II	2012_Kytalyk_Kolyma	69.05362	161.205179	520	62	*Larix*
12KO05	2012_Kytalyk_Kolyma	69.11836	161.02342	NA	0	n/a
13-TY-02-VI	2013_Taymyr	72.54772	105.7316	33023.36	141	*Larix*
13-TY-02-VII	2013_Taymyr	72.54884	105.74576	7156.53	88	*Larix*
13-TY-04VI	2013_Taymyr	72.40887	105.44804	400	66	*Larix*
13-TY-04VII	2013_Taymyr	72.40401	105.45187	400	92	*Larix*
13-TY-07VI	2013_Taymyr	71.10012	100.81295	576	106	*Larix*
13-TY-07VII	2013_Taymyr	71.10598	100.8463	400	91	*Larix*
13-TY-09VI	2013_Taymyr	72.15067	102.09771	576	173	*Larix*
13-TY-09VII	2013_Taymyr	72.14365	102.06259	576	183	*Larix*
14-OM-02-V1	2014_Omoloy	70.74418	132.698523	400	450	*Larix*
14-OM-02-V2	2014_Omoloy	70.72644	132.658169	400	143	*Larix*
14-OM-11-V3	2014_Omoloy	70.957883	132.570074	400	0	n/a
14-OM-20-V4	2014_Omoloy	70.526707	132.914259	400	292	*Larix*
14-OM-TRANS1	2014_Omoloy	70.943542	132.777408	314.16	24	*Larix*
14-OM-TRANS2	2014_Omoloy	70.939004	132.790487	314.16	25	*Larix*
14-OM-TRANS3	2014_Omoloy	70.935714	132.820357	314.16	25	*Larix*
14-OM-TRANS4	2014_Omoloy	70.93332	132.854538	314.16	23	*Larix*
14-OM-TRANS5	2014_Omoloy	70.935817	132.868951	314.16	24	*Larix*
14-OM-TRANS6	2014_Omoloy	70.944295	132.8777	314.16	22	*Larix*
14-OM-TRANS6-7	2014_Omoloy	70.948754	132.884332	NA	0	*Larix*
16-KP-V01	2016_Keperveem	67.3618	168.2542	706.86	37	*Larix*
16-KP-V02	2016_Keperveem	67.366	168.2366	706.86	7	*Larix*
16-KP-V03	2016_Keperveem	67.3664	168.2948	624	128	*Larix*
16-KP-V04	2016_Keperveem	67.3736	168.31	706.86	13	*Larix*
16-KP-V05	2016_Keperveem	67.3769	168.3122	706.86	107	*Larix*
16-KP-V06	2016_Keperveem	67.35	168.1885	706.86	107	*Larix*
16-KP-V07	2016_Keperveem	67.3456	168.1842	706.86	0	*Larix*
16-KP-V08	2016_Keperveem	67.3449	168.1802	706.86	1	*Larix*
16-KP-V09	2016_Keperveem	67.3538	168.2157	706.86	0	*Larix*
16-KP-V10	2016_Keperveem	67.3452	168.2013	706.86	24	*Larix*
16-KP-V11	2016_Keperveem	67.35	168.2009	706.86	85	*Larix*
16-KP-V12	2016_Keperveem	67.3531	168.2264	706.86	68	*Larix*
16-KP-V13	2016_Keperveem	66.9731	163.4177	706.86	187	*Larix*
16-KP-V14	2016_Keperveem	66.9874	163.3981	706.86	14	*Larix*
16-KP-V15	2016_Keperveem	66.9914	163.3843	706.86	1	*Larix*
16-KP-V16	2016_Keperveem	66.9715	163.4021	706.86	31	*Larix*
16-KP-V17	2016_Keperveem	66.9869	163.455	480	190	*Larix*
16-KP-V18	2016_Keperveem	66.9699	163.3845	50	192	*Larix*
16-KP-V19	2016_Keperveem	66.9706	163.3948	706.86	238	*Larix*

Table A1. Continued.

Site	Expedition	Latitude (° N)	Longitude (° E)	Area (m^2)	No. of trees	Most frequent species group
16-KP-V20	2016_Keperveem	65.9249	166.3609	706.86	107	*Larix*
16-KP-V21	2016_Keperveem	65.926	166.3609	706.86	48	*Larix*
16-KP-V22	2016_Keperveem	65.9352	166.3905	706.86	6	*Larix*
16-KP-V23	2016_Keperveem	65.9352	166.3933	706.86	0	*Larix*
16-KP-V24	2016_Keperveem	65.9365	166.389	706.86	0	*Larix*
16-KP-V25	2016_Keperveem	65.9372	166.3906	706.86	0	*Larix*
16-KP-V26	2016_Keperveem	65.9369	166.3861	706.86	76	*Larix*
16-KP-V27	2016_Keperveem	65.9369	166.385	706.86	114	*Larix*
16-KP-V28	2016_Keperveem	65.9231	166.3683	1296	96	*Larix*
16-KP-V29	2016_Keperveem	65.9252	166.3882	706.86	49	*Larix*
16-KP-V30	2016_Keperveem	65.9579	166.3333	706.86	4	*Larix*
16-KP-V31	2016_Keperveem	65.9585	166.3368	706.86	0	*Larix*
16-KP-V32	2016_Keperveem	65.9468	166.3561	706.86	6	*Larix*
16-KP-V33	2016_Keperveem	65.9459	166.3577	706.86	0	*Larix*
16-KP-V34	2016_Keperveem	65.9415	166.3486	706.86	140	*Larix*
16-KP-V35	2016_Keperveem	65.9329	166.2618	706.86	125	*Larix*
16-KP-V36	2016_Keperveem	65.9294	166.291	706.86	2	*Larix*
16-KP-V37	2016_Keperveem	65.9002	166.419	576	90	*Larix*
16-KP-V38	2016_Keperveem	65.9003	166.4168	706.86	135	*Larix*
16-KP-V39	2016_Keperveem	65.9217	166.3139	706.86	205	*Larix*
16-KP-V40	2016_Keperveem	67.7969	168.7096	706.86	0	n/a
16-KP-V41	2016_Keperveem	67.8171	168.6865	706.86	0	n/a
16-KP-V42	2016_Keperveem	67.8171	168.6885	706.86	0	n/a
16-KP-V43	2016_Keperveem	67.8195	168.6976	706.86	0	n/a
16-KP-V44	2016_Keperveem	67.8196	168.6963	706.86	0	n/a
16-KP-V45	2016_Keperveem	67.82	168.714	706.86	0	n/a
16-KP-V46	2016_Keperveem	67.8199	168.7115	706.86	0	n/a
16-KP-V47	2016_Keperveem	67.8048	168.7037	706.86	0	n/a
16-KP-V48	2016_Keperveem	67.8002	168.6379	706.86	0	n/a
16-KP-V49	2016_Keperveem	67.8026	168.6359	706.86	0	n/a
16-KP-V50	2016_Keperveem	67.8051	168.6297	706.86	0	n/a
16-KP-V51	2016_Keperveem	67.8055	168.6327	706.86	0	n/a
16-KP-V52	2016_Keperveem	67.8069	168.6311	706.86	0	n/a
16-KP-V53	2016_Keperveem	67.8079	168.6323	706.86	0	n/a
16-KP-V54	2016_Keperveem	67.8096	168.6299	706.86	0	n/a
16-KP-V55	2016_Keperveem	67.8091	168.6336	706.86	0	n/a
16-KP-V56	2016_Keperveem	67.8082	168.6355	706.86	0	n/a
16-KP-V57	2016_Keperveem	67.8076	168.645	706.86	0	n/a
16-KP-V58	2016_Keperveem	67.8086	168.645	706.86	0	n/a
18-LD-VP012-Tit-Ary	2018_Lena	71.967274	127.092825	900	0	*Larix*
B19-T1	2019_Batagay	67.58117	134.785314	706.86	0	n/a
B19-T2	2019_Batagay	67.580618	134.78351	706.86	0	n/a
EN18000	2018_Chukotka	68.097147	166.375447	706.86	111	*Larix*
EN18001	2018_Chukotka	67.39273	168.34662	706.86	50	*Larix*
EN18002	2018_Chukotka	67.386775	168.336731	706.86	0	n/a
EN18003	2018_Chukotka	67.39691	168.34702	706.86	37	*Larix*
EN18004	2018_Chukotka	67.397489	168.351225	706.86	6	*Larix*
EN18005	2018_Chukotka	67.419652	168.387511	706.86	1	*Larix*
EN18006	2018_Chukotka	67.414969	168.402874	706.86	141	*Larix*
EN18007	2018_Chukotka	67.403274	168.371965	706.86	181	*Larix*
EN18008	2018_Chukotka	67.402135	168.375284	706.86	0	*Larix*
EN18009	2018_Chukotka	67.400725	168.379683	706.86	4	*Larix*
EN18010	2018_Chukotka	67.402371	168.3662	706.86	11	*Larix*
EN18011	2018_Chukotka	67.404042	168.364252	706.86	0	*Salix*
EN18012	2018_Chukotka	67.402142	168.378078	706.86	80	*Larix*
EN18013	2018_Chukotka	67.405174	168.355304	706.86	0	*Salix*
EN18014	2018_Chukotka	67.395309	168.349106	1600	59	*Larix*

Table A1. Continued.

Site	Expedition	Latitude (° N)	Longitude (° E)	Area (m^2)	No. of trees	Most frequent species group
EN18015	2018_Chukotka	67.420379	168.33061	706.86	0	*Salix*
EN18016	2018_Chukotka	67.426726	168.390047	706.86	0	*Larix*
EN18017	2018_Chukotka	67.43229	168.383376	706.86	0	*Salix*
EN18018	2018_Chukotka	67.456295	168.405961	706.86	0	n/a
EN18019	2018_Chukotka	67.457073	168.408963	706.86	0	n/a
EN18020	2018_Chukotka	67.459159	168.411934	706.86	0	n/a
EN18021	2018_Chukotka	67.392129	168.328815	706.86	116	*Larix*
EN18022	2018_Chukotka	67.401024	168.348006	706.86	0	*Larix*
EN18023	2018_Chukotka	67.399236	168.351285	706.86	0	*Pinus pumila*
EN18024	2018_Chukotka	67.370964	168.426362	706.86	120	*Larix*
EN18025	2018_Chukotka	67.367027	168.42381	706.86	97	*Larix*
EN18026	2018_Chukotka	67.396089	168.354297	706.86	77	*Larix*
EN18027	2018_Chukotka	67.393408	168.35905	706.86	54	*Larix*
EN18028	2018_Chukotka	68.46781	163.357622	706.86	97	*Larix*
EN18029	2018_Chukotka	68.465606	163.352262	706.86	71	*Larix*
EN18030	2018_Chukotka	68.405539	164.532731	706.86	669	*Larix*
EN18031	2018_Chukotka	68.404918	164.545351	706.86	100	*Larix*
EN18032	2018_Chukotka	68.404868	164.551181	706.86	1	*Salix*
EN18033	2018_Chukotka	68.403212	164.551805	706.86	0	*Salix*
EN18034	2018_Chukotka	68.403486	164.548043	706.86	35	*Larix*
EN18035	2018_Chukotka	68.403166	164.590932	706.86	168	*Larix*
EN18051	2018_Chukotka	67.80261	168.7047	706.86	0	n/a
EN18052	2018_Chukotka	67.79941	168.7083	706.86	0	n/a
EN18053	2018_Chukotka	67.79729	168.7107	706.86	0	n/a
EN18054	2018_Chukotka	67.79766	168.6904	706.86	0	n/a
EN18055	2018_Chukotka	67.79103	168.6825	706.86	0	n/a
EN18061	2018_Yakutia	62.076376	129.618586	706.86	611	*Pinus* tree
EN18062	2018_Yakutia	62.179065	127.805796	706.86	418	*Larix*
EN18063	2018_Yakutia	63.776636	122.501003	706.86	459	*Larix*
EN18064	2018_Yakutia	63.814594	122.209683	706.86	435	*Pinus* tree
EN18065	2018_Yakutia	63.795223	122.443715	304	242	*Larix*
EN18066	2018_Yakutia	63.797119	122.438071	706.86	115	*Larix*
EN18067	2018_Yakutia	63.076368	117.975342	706.86	339	*Larix*
EN18068	2018_Yakutia	63.074232	117.98207	706.86	74	*Larix*
EN18069	2018_Yakutia	63.173288	118.132507	706.86	543	*Picea*
EN18070_centre	2018_Yakutia	63.082476	117.985333	300	81	*Picea*
EN18070_edge	2018_Yakutia	63.082983	117.984938	300	224	*Picea*
EN18070_end	2018_Yakutia	63.08341	117.984574	200	0	n/a
EN18070_transition	2018_Yakutia	63.082733	117.985156	300	142	*Picea*
EN18071	2018_Yakutia	62.225093	116.275603	706.86	236	*Larix*
EN18072	2018_Yakutia	62.199571	117.379125	706.86	688	*Larix*
EN18073	2018_Yakutia	62.188712	117.409917	706.86	837	*Larix*
EN18074	2018_Yakutia	62.215192	117.021599	706.86	275	*Picea*
EN18075	2018_Yakutia	62.696991	113.676535	706.86	274	*Pinus* tree
EN18076	2018_Yakutia	62.70089	113.67341	706.86	582	*Larix*
EN18077	2018_Yakutia	61.892568	114.288623	706.86	546	*Pinus* tree
EN18078	2018_Yakutia	61.575058	114.29995	706.86	236	*Larix*
EN18079	2018_Yakutia	59.974919	112.958985	706.86	305	*Pinus* tree
EN18080	2018_Yakutia	59.977106	112.961379	706.86	339	*Picea*
EN18081	2018_Yakutia	59.970583	112.987096	706.86	83	*Picea*
EN18082	2018_Yakutia	59.97764	112.98218	706.86	101	*Larix*
EN18083	2018_Yakutia	59.974714	113.002874	706.86	138	*Picea*
EN21-201	2021_Yakutia	63.217776	139.543709	NA	0	*Larix*
EN21-202	2021_Yakutia	63.32516	141.07455	706.86	160	*Larix*
EN21-203	2021_Yakutia	63.430107	140.412509	706.86	126	*Larix*
EN21-204	2021_Yakutia	63.44253	140.40282	706.86	118	*Larix*

Table A1. Continued.

Site	Expedition	Latitude (° N)	Longitude (° E)	Area (m^2)	No. of trees	Most frequent species group
EN21-205	2021_Yakutia	63.43858	140.40688	706.86	44	*Larix*
EN21-206	2021_Yakutia	63.34379	141.07071	706.86	81	*Larix*
EN21-207	2021_Yakutia	63.344383	141.069788	NA	3	*Pinus pumila*
EN21-208	2021_Yakutia	63.34528	141.06827	NA	0	n/a
EN21-209	2021_Yakutia	63.39854	140.55406	706.86	50	*Larix*
EN21-210	2021_Yakutia	63.397717	140.55925	NA	0	n/a
EN21-211	2021_Yakutia	63.40056	140.55357	706.86	109	*Larix*
EN21-212	2021_Yakutia	63.232626	142.962381	706.86	251	*Larix*
EN21-213	2021_Yakutia	63.230378	142.963774	100	219	*Larix*
EN21-214	2021_Yakutia	63.23257	142.9577	NA	0	n/a
EN21-215	2021_Yakutia	63.210719	139.540937	706.86	14	*Larix*
EN21-216	2021_Yakutia	63.212267	139.541692	NA	0	n/a
EN21-217	2021_Yakutia	63.438697	140.597609	706.86	41	*Larix*
EN21-218	2021_Yakutia	63.428277	140.579547	706.86	0	n/a
EN21-219	2021_Yakutia	63.425647	140.588331	706.86	284	*Larix*
EN21-220	2021_Yakutia	62.07984	132.3668	NA	0	n/a
EN21-221	2021_Yakutia	62.083241	132.372643	706.86	28	*Betula*
EN21-222	2021_Yakutia	62.08595	132.370772	706.86	640	*Larix*
EN21-223	2021_Yakutia	62.087193	132.370561	706.86	306	*Larix*
EN21-224	2021_Yakutia	62.042778	132.388521	706.86	0	n/a
EN21-225	2021_Yakutia	62.044236	132.391202	706.86	452	*Betula*
EN21-226	2021_Yakutia	62.045558	132.389098	706.86	168	*Larix*
EN21-227	2021_Yakutia	62.040546	132.396302	314.16	4	*Larix*
EN21-228	2021_Yakutia	62.384988	133.748979	706.86	268	*Larix*
EN21-229	2021_Yakutia	62.384468	133.750727	314.16	109	*Larix*
EN21-230	2021_Yakutia	62.334507	133.688018	706.86	163	*Larix*
EN21-231	2021_Yakutia	62.334694	133.68405	NA	22	*Betula*
EN21-232	2021_Yakutia	62.172203	130.911195	706.86	652	*Larix*
EN21-233	2021_Yakutia	62.169607	130.903851	706.86	308	*Larix*
EN21-234	2021_Yakutia	62.287013	130.377589	706.86	39	*Pinus* tree
EN21-235	2021_Yakutia	62.275634	130.37659	706.86	141	*Pinus* tree
EN21-236	2021_Yakutia	62.262231	130.327876	706.86	234	*Pinus* tree
EN21-237	2021_Yakutia	62.13009	130.874837	706.86	288	*Larix*
EN21-238	2021_Yakutia	62.133528	130.873521	706.86	176	*Larix*
EN21-239	2021_Yakutia	62.316127	130.116028	314.16	290	*Alnus*
EN21-240	2021_Yakutia	62.353399	130.151416	706.86	645	*Picea*
EN21-241	2021_Yakutia	62.148377	130.65177	706.86	29	*Larix*
EN21-242	2021_Yakutia	62.148415	130.653568	706.86	445	*Betula*
EN21-243	2021_Yakutia	62.149423	130.654024	706.86	0	n/a
EN21-244	2021_Yakutia	62.156934	130.659589	314.16	628	*Larix*
EN21-245	2021_Yakutia	61.78444	130.48492	706.86	299	*Populus*
EN21-246	2021_Yakutia	61.78305	130.49245	225	2713	*Betula*
EN21-247	2021_Yakutia	61.77975	130.49998	706.86	76	*Larix*
EN21-248	2021_Yakutia	61.747877	130.530323	706.86	405	*Betula*
EN21-249	2021_Yakutia	61.745655	130.530715	706.86	835	*Larix*
EN21-250	2021_Yakutia	61.745696	130.532625	706.86	539	*Betula*
EN21-251	2021_Yakutia	61.740083	130.528577	706.86	149	*Larix*
EN21-252	2021_Yakutia	61.897154	130.482395	706.86	352	*Salix*
EN21-253	2021_Yakutia	61.89501	130.4848	706.86	290	*Larix*
EN21-254	2021_Yakutia	61.894779	130.488766	706.86	291	*Larix*
EN21-255	2021_Yakutia	61.769113	130.386747	706.86	871	*Larix*
EN21-256	2021_Yakutia	61.76639	130.83875	706.86	596	*Betula*
EN21-257	2021_Yakutia	61.770502	130.391538	NA	0	n/a
EN21-258	2021_Yakutia	61.899226	130.423401	706.86	506	*Larix*
EN21-259	2021_Yakutia	61.901329	130.500516	706.86	492	*Larix*
EN21-260	2021_Yakutia	61.76387	130.47968	706.86	309	*Larix*
EN21-261	2021_Yakutia	61.766817	130.457716	706.86	329	*Larix*
EN21-262	2021_Yakutia	61.76123	130.47043	NA	0	n/a
EN21-263	2021_Yakutia	62.209135	127.691498	NA	0	n/a
EN21-264	2021_Yakutia	62.216896	127.717821	NA	0	n/a

Appendix B: Overview over all variables

Table B1. List of all variables of the tree database. The spelling, capitalization, grammar and units reflect the original database.

Tree database variables (tables "Tree heights" and "Tree measurements" have the same column names)		
Column name in *.tab file	Original measurement or derived variable?	Unit/values
Tree ID	original	
Event	original	
Campaign	original	
PI	original	
Date/Time	original	Date (YYYY-MM-DD)
Lat C (Plot latitude)	derived (from PlotDataBase)	° N
Long C (Plot longitude)	derived (from PlotDataBase)	° E
Latitude	original	° N
Longitude	original	° E
Tree, survey protocol	original	Category with levels PLOT, PLOTHEIGHT, CIRCLEPLOT, EXTRA
Subsample ID	original	
Species	original	
Genus (Species group)	derived (from Species)	
Growth form (T = Tree, K = Krumholz, S = S ...)	original	Category with levels T = Tree, K = Krumholz, S = Shrub, M = Multistem, TL = Tree lying
Tree height [m]	original	m
Crown diam [m] (Maximum)	original	m
Crown diam [m] (Smaller, diameter measured pe ...)	original	m
Vitality (++ = very high vitality, + = ...)	original	Category with levels "++", "+", "0", "-", "–", and "dead"
Comment (Vitality estimate comment)	original	
Tree D base [cm]	original	cm
DBH [cm]	original	cm
Tree crown base [m]	original	m
Tree D base [cm] (Predicted)	derived (from Tree height)	cm
DBH [cm] (Predicted)	derived (from Tree height)	cm
Tree BA base [m**2]	derived (from Tree height via predicted diameters)	m^2
Tree BA breast [m**2]	derived (from Tree height via predicted diameters)	m^2
Tree vol conical [m**3]	derived (from Tree height)	m^3
Tree vol smallian [m**2]	derived (from Tree height)	m^3

Table B2. List of all variables of the plot database. The spelling, capitalization, grammar and units reflect the original database.

Plot database variables		
Column name in *.tab file	Original measurement or derived variable?	Unit/values
Event	derived (from Site and existing PANGAEA data sets)	
Site	original	
Campaign	original	
Area (Federation Subject)	original	
Area (District)	original	
Elevation [m a.s.l.]	derived (from Latitude and Longitude)	m
PI	original	
Reference	original	
Area (Camp Location)	original	
Latitude	original	° N
Longitude	original	° E
Date/Time	original	
Comment (Area comment)	original	
Area [m**2]	derived (from plot area shape)	m^2
Plot (Area shape in m)	original	
Plot (Area seedlings in m)	Original	
Forest type	original	
Trees [#]	derived (from Tree Heihts data base)	
Trees [#/ha]	derived (from Tree Heihts data base)	ha^{-1}
Tree height [m] (Mean values)	Derived (from Tree Heihts data base)	m
Tree height [m] (Living, Mean values)	Derived (from Tree Heihts data base)	m
Tree height [m] (Median values)	derived (from Tree Heihts data base)	m
Tree height [m] (Living, Median values)	derived (from Tree Heihts data base)	m
Height quantile [m] (Quantile (25th))	derived (from Tree Heihts data base)	m
Height quantile [m] (Quantile (75th))	derived (from Tree Heihts data base)	m
Height quantile [m] (Quantile (90th))	derived (from Tree Heihts data base)	m
Height quantile [m] (Quantile (98th))	derived (from Tree Heihts data base)	m
Height max [m]	derived (from Tree Heihts data base)	m
Trees [#] (Living)	derived (from Tree Heihts data base)	
Tree BA breast [m**2]	derived (from Tree Heihts data base)	m^2
Tree BA base [m**2]	derived (from Tree Heihts data base)	m^2
Tree vol conical [m**3]	derived (from Tree Heihts data base)	m^3
Tree vol smallian [m**3]	derived (from Tree Heihts data base)	m^3
Tree vol conical [m**3] (Living)	derived (from Tree Heihts data base)	m^3
Tree vol smallian [m**3] (Living)	derived (from Tree Heihts data base)	m^3
Tree BA breast [m**2] (Living)	Derived (from Tree Heihts data base)	m^2
Tree BA [m**2/ha]	derived (from Tree Heihts data base)	m^2/ha
Tree vol conical [m**3/ha]	derived (from Tree Heihts data base)	m^3/ha
Tree vol smallian [m**3/ha]	derived (from Tree Heihts data base)	m^3/ha
Tree BA [m**2/ha] (Living)	derived (from Tree Heihts data base)	m^2/ha
Tree vol conical [m**3/ha] (Living)	Derived (from Tree Heihts data base)	m^3/ha
Tree vol smallian [m**3/ha] (Living)	derived (from Tree Heihts data base)	m^3/ha
Gini coeff. (Height)	derived (from Tree Heihts data base)	
Genus (Most frequent species group)	derived (from Tree Heihts data base)	
H’	derived (from Tree Heihts data base)	
Spec No. (#)	derived (from Tree Heihts data base)	

Appendix C: Coefficients of diameter–height allometries

Allometries were calculated to obtain the diameter from the height of the tree with the following formula:

$$D = (a_1 + a_3 \cdot S) \cdot H^{(a_2 + a_4 \cdot S)},$$

where D is the diameter in centimetres. H is the height, and S is the stand density, which is obtained from the number of trees per hectare (T_{ha}) as follows:

$$S = \max(\log_{10}(T_{\mathrm{ha}}), 2)).$$

The coefficients a_1, a_2, a_3, and a_4 resulting from fitting with the least squares method are shown in Tables B1 and B2.

Table C1. Coefficients for diameter at base allometries.

Diameter at base					
Species group	a1	a2	a3	a4	Standard error
Larix	4.4264	0.7696	−0.7768	0.07843	4.9087
Salix	2.4481	0.7853	0	0	4.2406
Betula	0.9125	1.1514	0	0	4.3228
Alnus	0.9429	0.9483	0	0	2.3793
Pinus tree	1.8644	1.0282	0	0	4.5821
Picea	0.9178	1.1499	0	0	4.0817
Unclassified	3.4548	0.5980	0	0	3.1845
Abies	2.6910	0.6625	0	0	1.0015
Populus	0.05764	2.2859	0	0	2.9014
Larix krumholz	3.9278	0.5807	0	0	2.1763

Table C2. Coefficients for diameter at breast height allometries.

Diameter at breast height					
Species group	a1	a2	a3	a4	Standard error
Larix	5.5512	0.4757	−1.0528	0.1248	3.9600
Salix	2.1031	0.7011	0	0	1.7292
Betula	0.8402	1.0764	0	0	3.0624
Alnus	1.9435	0.4799	0	0	2.5714
Pinus tree	1.8775	0.9672	0	0	4.2074
Picea	1.1947	0.9826	0	0	3.0611
Unclassified	3.0454	0.5854	0	0	2.7661
Abies	2.5659	0.6275	0	0	0.5474
Populus	0.1440	1.8987	0	0	2.5191
Larix krumholz	2.6825	0.7662	0	0	1.1286

Appendix D: Correlation matrix of climate variables

Table D1. Correlations between the four climate variables, namely the January temperature (T01), July temperature (T07), annual precipitation (Prec.), and growing degree days above 0 °C (GDD0), from the CHELSA data set at the locations of our plots, calculated using the R function cor().

	T01	T07	Prec.	GDD0
T01	1	−0.645	−0.045	−0.640
T07	−0.645	1	0.443	0.997
Prec.	−0.045	0.443	1	0.429
GDD0	−0.640	0.997	0.429	1

Appendix E: Correlation coefficients for forest canopy height

In Sect. 3.2.3, linear correlations were calculated between the Simard et al. (2011) forest height product and different forest metrics (heights in metres), with the results shown in Table E1.

Table E1. Correlation coefficients for forest canopy height. Adj. R^2: adjusted R^2.

Variable	Adj. R^2	Standard error
Maximum height	0.190	5.42
Height 98th percentile	0.152	4.638
Height 90th percentile	0.0961	4.063
Height 75th percentile	0.0522	3.359
Height 25th percentile	0.0115	1.579
Mean height	0.0688	2.143

Author contributions. TM and UH conceptualized the paper and did the analysis. TM drafted the paper and performed the analysis. SK revised the data. TM, UH, LAP, MW, ESZ, AIK, PVD and SK did fieldwork and revised the paper.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. Many people of the logistic and scientific staff at the Alfred Wegener Institute (AWI) and North-Eastern Federal University Yakutsk (NEFU) contributed to the success of the expeditions. We thank Cathy Jenks, for the English language proofreading. And, finally, we are grateful to Elizabeth Webb and two anonymous reviewers, whose comments helped to improve the paper.

Financial support. The study has been supported by the ERC consolidator (grant no. 772852). Parts of the fieldwork have been funded by the Russian Foundation for Basic Research (grant no. 20-35-90081) and the Ministry of Science and Higher Education of Russia (grant no. FSRG-2020-0019). Publisher's note: Copernicus Publications has not received any payments from Russian or Belarusian institutions for this paper.

Review statement. This paper was edited by Hanqin Tian and reviewed by Elizabeth Webb and two anonymous referees.

References

Abaimov, A. P.: Geographical Distribution and Genetics of Siberian Larch Species, in: Permafrost Ecosystems: Siberian Larch Forests, edited by: Osawa, A., Zyranova, O. A., Matsuura, Y., Kajimoto, T., and Wein, R. W., Springer, 41–58, https://doi.org/10.1007/978-1-4020-9693-8_2, 2010.

Alexander, H. D., Mack, M. C., Goetz, S., Loranty, M. M., Beck, P. S. A., Earl, K., Zimov, S., Davydov, S., and Thompson, C. C.: Carbon Accumulation Patterns During Post-Fire Succession in Cajander Larch (*Larix cajanderi*) Forests of Siberia, Ecosystems 15, 1065–1082, https://doi.org/10.1007/s10021-012-9567-6, 2012.

Bonan, G. B.: Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests, Science, 320, 1444–1449, https://doi.org/10.1126/science.1155121, 2008.

Cailliez, F. and Alder, D.: Forest volume estimation and yield prediction (Vol. 1), Food and agriculture Organization of the United Nations, Rome, ISBN 92-5-100923-6, https://www.fao.org/3/ap354e/ap354e00.pdf (last access: 29 November 2022), 1980.

Chen, D., Loboda, T. V., Krylov, A., and Potapov, P.: Distribution of Estimated Stand Age Across Siberian Larch Forests, 1989–2012, ORNL DAAC [data set], https://doi.org/10.3334/ORNLDAAC/1364, 2017.

Collow, A. B. M., Thomas, N. P., Bosilovich, M. G., Lim, Y.-K., Schubert, S. D., and Koster, R. D.: Seasonal Variability in the Mechanisms Behind the 2020 Siberian Heatwaves, J. Climate, 35, 3075–3090, https://doi.org/10.1175/JCLI-D-21-0432.1, 2022.

Delcourt, C. J. F. and Veraverbeke, S.: Allometric equations and wood density parameters for estimating aboveground and woody debris biomass in Cajander larch (*Larix cajanderi*) forests of northeast Siberia, Biogeosciences, 19, 4499–4520, https://doi.org/10.5194/bg-19-4499-2022, 2022.

Dolman, A. J., Maximov, T. C., Moors, E. J., Maximov, A. P., Elbers, J. A., Kononov, A. V., Waterloo, M. J., and van der Molen, M. K.: Net ecosystem exchange of carbon dioxide and water of far eastern Siberian Larch (*Larix cajanderii*) on permafrost, Biogeosciences, 1, 133–146, https://doi.org/10.5194/bg-1-133-2004, 2004.

FAO: On definitions of forest and forest change, FRA Working Paper No. 33, Rome, https://www.fao.org/3/ad665e/ad665e03.htm#P199_9473 (last access: 19 December 2022), 2000.

FAO: Global Forest Resources Assessment 2020 – Key findings, Rome, https://doi.org/10.4060/ca8753en, 2020.

Gini, C.: Variabilità e mutabilità – contributo allo studio delle distribuzioni e delle relazioni statistiche, Bologna, 1912.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., and Townshend, J. R. G.: High-Resolution Global Maps of 21st-Century Forest Cover Change, Science, 342, 850–53, https://doi.org/10.1126/science.1244693, 2013.

Houghton, R. A., Butman, D., Bunn, A. G., Krankina, O. N., Schlesinger, P., and Stone, T. A.: Mapping Russian forest biomass with data from satellites and forest inventories, Environ. Res. Lett. 2, 045032, https://doi.org/10.1088/1748-9326/2/4/045032, 2007.

Jia, B. and Zhou, G.: Growth characteristics of natural and planted Dahurian larch in northeast China, Earth Syst. Sci. Data, 10, 893–898, https://doi.org/10.5194/essd-10-893-2018, 2018.

Kajimoto, T., Matsuura, Y., Sofronov, M. A., Voloktina, A. V., Mori, S., Osawa, A., and Abaimov, A. P.: Above- and below-ground biomass and net primary productivity of a *Larix gmelinii* stand near Tura, central Siberia, Tree Physiol., 19, 815–822, https://doi.org/10.1093/treephys/19.12.815, 1999.

Kajimoto, T., Osawa, A., Usoltsev, V. A., and Abaimov, A. P.: Biomass and Productivity of Siberian Larch Forest Ecosystems, in: Permafrost Ecosystems: Siberian Larch Forests, edited by: Osawa, A., Zyranova, O. A., Matsuura, Y., Kajimoto, T., and Wein, R. W., Springer, 41–58, https://doi.org/10.1007/978-1-4020-9693-8_6, 2010.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, P., and Kessler, M.: Climatologies at high resolution for the Earth land surface areas, Scientific Data, 4, 170122, https://doi.org/10.1038/sdata.2017.122, 2017.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., and Kessler, M.: Climatologies at high resolution for the earth's land surface areas, EnviDat [data set], https://doi.org/10.16904/envidat.228.v2.1, 2021.

Kharuk, V. I., Ranson, K. J., Dvinskaya, M. L., and Im, S. T.: Wildfires in northern Siberian larch dominated communities, Environ. Res. Lett., 6, 045208, https://doi.org/10.1088/1748-9326/6/4/045208, 2011.

Kharuk, V. I., Ranson, K. J., Petrov, I. A., Dvinskaya, M. L., Im, S. T., and Golyukov, A. S.: Larch (*Larix dahurica* Turcz) growth response to climate change in the Siberian permafrost zone, Reg. Environ. Change, 19, 233–243 https://doi.org/10.1007/s10113-018-1401-z, 2019.

Kropp, H., Loranty, M., Alexander, H. D., Berner, L. T., Natali, S. M., and Spawn, S. A.: Environmental constraints on transpiration and stomatal conductance in a Siberian Arctic boreal forest, J. Geophys. Res-Biogeo., 122, 487–497, https://doi.org/10.1002/2016JG003709, 2017.

Kruse, S., Bolshiyanov, D., Grigoriev, M., Morgenstern, A., Pestryakova, L., Tsibizov, L., and Udke, A.: Russian-German Cooperation: Expeditions to Siberia in 2018, Berichte zur Polar- und Meeresforschung, 734, 136–153, https://doi.org/10.2312/BzPM_0734_2019, 2019.

Kruse, S., Herzschuh, U., Schulte, L., Stuenzi, S. M., Brieger, F., Zakharov, E. S., and Pestryakova, L. A.: Forest inventories on circular plots on the expedition Chukotka 2018, NE Russia, PANGAEA [data set], https://doi.org/10.1594/PANGAEA.923638, 2020a.

Kruse, S., Kolmogorov, A. I., Pestryakova, L. A., and Herzschuh, U.: Long-lived larch clones may conserve adaptations that could restrict treeline migration in northern Siberia, Ecol. Evol., 10, 10017–10030, https://doi.org/10.1002/ece3.6660, 2020b.

Kuznetsova, L. V., Zakharova, V. I., Sosina,, N. K., Nikolin, E. G., Ivanova, E. I., Sofronova, E. V., Poryadina, L. N., Mikhailyova, L. G., Vasilyeva, I. I., Remigailo, P. A., Gabyshev, A. P., Ivanova, A. P., and Kopyrina, L. I.: Flora of Yakutia: Composition and Ecological Structure, in: The far North, edited by: Kuznetsova, L. V., Zakharova, V. I., Sosina, N. K., Nikolin, E. G., Ivanova, E. I., Sofronova, E. V., Poryadina, L. N., Mikhalyova, L. G., Vasilyeva, I. I., Remigalio, P. A., Gabyshev, V. A., Ivanova, A. P., and Kopyrina, L. I., Plant and Vegetation, 3, Springer, Dordrecht, 24–140, https://doi.org/10.1007/978-90-481-3774-9_2, 2010.

Lin, H., Yang, K., Hiseh, T., and Hiseh, C.: Species Composition and Structure of a Montane Rainforest of Mt. Lopei in Northern Taiwan, Taiwania, 50, 234–249, 2005.

Miesner, T., Herzschuh, U., Pestryakova, L. A., Wieczorek, M., Kolmogorov, A., Heim, B., Zakharov, E. S., Shevtsova, I., Epp, L. S., Niemeyer, B., Jacobsen, I., Schröder, J., Trense., D., Schnabel, E., Schreiber, X., Bernhardt, N., Stuenzi, S. M., Brieger, F., Schulte, L., Smirnikov, V., Gloy, J., von Hippel, B., Jackisch, R., and Kruse, S.: Tree data set from forest inventories in north-eastern Siberia, PANGAEA [data set], https://doi.org/10.1594/PANGAEA.943547, 2022.

Montesano, P. M., Sun, G., Dubayah, R. O., and Ranson, K. J.: Spaceborne potential for examining taiga–tundra ecotone form and vulnerability, Biogeosciences, 13, 3847–3861, https://doi.org/10.5194/bg-13-3847-2016, 2016.

Osawa, A. and Zyryanova, O. A.: Introduction, in: Permafrost Ecosystems: Siberian Larch Forests, edited by: Osawa, A., Zyranova, O. A., Matsuura, Y., Kajimoto, T., Wein, R. W., Springer, 3–15, https://doi.org/10.1007/978-1-4020-9693-8_1, 2010.

Overduin, P. P., Blender, F., Bolshiyanov, D. Y., Grigoriev, M. N., Morgenstern, A., and Meyer, H.: Russian-German Cooperation: Expeditions to Siberia in 2016, Berichte zur Polar- und Meeresforschung, 709, 130–137, https://doi.org/10.2312/BzPM_0709_2017, 2017.

Payette, S.: Fire as a controlling process in the North American boreal forest, in: A Systems Analysis of the Global Boreal Forest, edited by: Bonan, G. B., Shugart, H. H., and Leemans, R., Cambridge University Press, Cambridge, 144–169, https://doi.org/10.1017/CBO9780511565489.006, 1992.

Picard, N., Boyemba Bosela, F., and Rossi, V.: Reducing the error in biomass estimates strongly depends on model selection, Ann. For. Sci., 72, 811–823, https://doi.org/10.1007/s13595-014-0434-9, 2015.

Ponomarev, E., Zabrodin, A., and Ponomareva, T.: Classification of Fire Damage to Boreal Forests of Siberia in 2021 Based on the dNBR Index, Fire, 5, 19, https://doi.org/10.3390/fire5010019, 2022.

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C. E., Armston, J., Dubayah, R., Blair J. B., and Hofton, M.: Mapping and monitoring global forest canopy height through integration of GEDI and Landsat data, Remote Sens. Environ., 253, 112165, https://doi.org/10.1016/j.rse.2020.112165, 2020.

QGIS Geographic Information System, QGIS Association [software], http://www.qgis.org, last access: 10 October 2021.

R Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing [software], Vienna, Austria, https://www.R-project.org/, last access: 16 June 2021.

Ranson, K. J., Sun, G., Kharuk, V. I., and Kovacs, K.: Assesing tundra-taiga boundary with multi-sensor satellite data, Remote Sens. Environ., 93, 283–295, https://doi.org/10.1016/j.rse.2004.06.019, 2004.

Santoro, M., Cartus, O., Mermoz, S., Bouvet, A., Le Toan, T., Carvalhais, N., Rozendaal, D., Herold, M., Avitabile, V., Quegan, S., Carreiras, J., Rauste, Y., Balzter, H., Schmullius, C., and Seifert, F. M.: GlobBiomass global above-ground biomass and growing stock volume datasets, GlobBiomass [data set], http://globbiomass.org/products/global-mapping (last access: 11 November 2021), 2018.

Santoro, M., Cartus, O., Carvalhais, N., Rozendaal, D. M. A., Avitabile, V., Araza, A., de Bruin, S., Herold, M., Quegan, S., Rodríguez-Veiga, P., Balzter, H., Carreiras, J., Schepaschenko, D., Korets, M., Shimada, M., Itoh, T., Moreno Martínez, Á., Cavlovic, J., Cazzolla Gatti, R., da Conceição Bispo, P., Dewnath, N., Labrière, N., Liang, J., Lindsell, J., Mitchard, E. T. A., Morel, A., Pacheco Pascagaza, A. M., Ryan, C. M., Slik, F., Vaglio Laurin, G., Verbeeck, H., Wijaya, A., and Willcock, S.: The global forest above-ground biomass pool for 2010 estimated from high-resolution satellite observations, Earth Syst. Sci. Data, 13, 3927–3950, https://doi.org/10.5194/essd-13-3927-2021, 2021.

Scheffer, M., Hirota M., Holmgren, M., Van Nes, E. H., and Chapin, F. S.: Thresholds for boreal biome transitions, P. Natl. Acad. Sci. USA, 109, 21384–21389, https://doi.org/10.1073/pnas.1219844110, 2012.

Schepaschenko, D., Shvidenko, A., Usoltsev, V., Lakyda, P., Yunjian L., Vasylyshyn, R., Lakyda, I., Myklush, Y., See, L., McCallum, I., Fritz, S., Kraxner, F., and Obersteiner, M.: A dataset of forest biomass structure for Eurasia, Sci. Data, 4, 170070, https://doi.org/10.1038/sdata.2017.70, 2017.

Schepaschenko, D., Moltchanova, E., Fedorov, S., Karminov, V., Ontikov, P., Santoro, M., See, L., Kositsyn, V., Shivdenko, A., Romanovskaya, A., Korotkov, V., Lesiv, M., Bartalev, S., Fritz, S., Shchepashchenko, M., and Kraxner, F.: Russian forest sequesters substantially more carbon than previously reported, Sci. Rep., 11, 12825, https://doi.org/10.1038/s41598-021-92152-9, 2021.

Schuur, E., McGuire, A., Schädel, C., Grosse, G., Harden, J., Hayes, D., Hugelius, G., Koven, C., Kuhry, P., Lawrence, D., Natali, S., Olefeldt, D., Romanovsky, V. E., Schaefer, K., Turetsky, M., Treat, C., and Vonk, J.: Climate change and the permafrost carbon feedback, Nature, 520, 171–179, https://doi.org/10.1038/nature14338, 2015.

Simard, M., Pinto, N., Fisher, J. B., and Baccini, A.: Mapping forest canopy height globally with spaceborne lidar, J. Geophys. Res., 116, G04021, https://doi.org/10.1029/2011JG001708, 2011.

Sugimoto, A., Yanagisawa, N., Naito, D., Fujita, N., and Maximov, T. C.: Importance of permafrost as a source of water for plants in east Siberian taiga, Ecol. Res., 17, 493–503, https://doi.org/10.1046/j.1440-1703.2002.00506.x, 2002.

van Geffen, F., Schulte, L., Geng, R., Heim, B., Pestryakova, L. A., Herzschuh, U., and Kruse, S.: SiDroForest: Individual-labelled trees acquired during the fieldwork expeditions that took place in 2018 in Central Yakutia and Chukotka, Siberia, PANGAEA [data set], https://doi.org/10.1594/PANGAEA.932821, 2021.

Walker, X. J., Baltzer, J. L., Cumming, S. G., Day, N. J., Ebert, C., Goetz, S., Johnstone, J. F., Potter, S., Rogers, B. M., Schuur, E. A., Turetsky, M. R., and Mack, M. C.: Increasing wildfires threaten historic carbon sink of boreal forest soils, Nature, 572, 520–523, https://doi.org/10.1038/s41586-019-1474-y, 2019.

Walker, X., Alexander, H. D., Berner, L., Boyd, M. A., Loranty, M. M., Natali, S., and Mack, M. C.: Positive response of tree productivity to warming is reversed by increased tree density at the Arctic tundra-taiga ecotone, Can. J. For. Res. 51, 1323–1338, https://doi.org/10.1139/cjfr-2020-0466, 2021.

Wang, W., Zu, Y., Wang, H., Matsuuta, Y., Sasa, K., and Koike, T.: Plant Biomass and Productivity of *Larix gmelinii* Forest Ecosystems in Northeast China: Intra- and Inter- species Comparison, Eurasian J. For. Res., 8, 21–41, 2005.

Widagdo, F. R. A., Xie, L., Dong, L., and Li, F.: Origin-based biomass allometric equations, biomass partitioning, and carbon concentration variations of planted and natural *Larix gmelinii* in northeast China, Global Ecology and Conservation, 23, e01111, https://doi.org/10.1016/j.gecco.2020.e01111, 2020.

Wieczorek, M., Kruse, S., Epp, L. S., Kolmogorov, A., Nikolaev, A. N., Heinrich, I., Jeltsch, F., Pestryakova, L. A., Zibulski, R., and Herzschuh, U.: Field and simulation data for larches growing in the Taimyr treeline ecotone, PANGAEA [data set], https://doi.org/10.1594/PANGAEA.874615, 2017.

Xiao, R., Man, X., and Duan, B.: Carbon and Nitrogen Stocks in Three Types of *Larix gmelinii* Forests in Daxing'an Mountains, Northeast China. Forests, 11, 305, https://doi.org/10.3390/f11030305, 2020.

Yang, W. and Kondoh, A.: Evaluation of the Simard et al. 2011 Global Canopy Height Map in Boreal Forests, Remote Sens., 12, 1114, https://doi.org/10.3390/rs12071114, 2020.

Zhang, Y., Liang, S., and Yang, L.: A Review of Regional and Global Gridded Forest Biomass Datasets, Remote Sens., 11, 2744, https://doi.org/10.3390/rs11232744, 2019.