Supporting material 1

The details of estimating tree density and the total number of trees

Tree densities and total number of large trees by dbh classes

Our aim was to estimate tree densities (trees per hectare) and total numbers of living large and/or old trees for boreal subzones. To reach these goals we also need estimates of forest area. Its development is essential in itself, as the definition of land use classes has changed from the 1920s.

The results of the Finnish NFIs are usually calculated for regions, for which land areas are available from the statistics provided annually by the National Land Survey of Finland. From here on, we call these *calculation regions (c).* We aimed at calculating the results for *boreal subzones (b)*. The borders of subzones *(b)* do not match with the area statistics of National Land Survey of Finland and the exact land areas of subzones are thus unknown. Therefore, we calculated first the land area represented by a plot center (plot inventories NFI6-NFI11) or a meter of measured line (line survey NFI1) in calculation regions and derived the estimates of forest areas, tree densities and total numbers of trees for the intersections of regions c and boreal subzones b ($b \cap c$) using the estimators presented in Table S.1. The total number of trees in a boreal subzone b was derived as a sum of the estimates for intersections $b \cap c$. Tree density (stems per hectare) in each subzone was finally estimated as a ratio of the total number of stems and corresponding forest area.

In estimating tree densities (stems per hectare) from the line survey data of NFI1, we used the stand-level ocular assessment of growing stock volume (V, m^3ha^{-1}) as an auxiliary variable.

The ocular assessment of growing stock was available for a total of 62559 compartments intersected by the inventory lines with mean values of 74 and 48 m^3ha^{-1} for the southern and northern parts of Finland, respectively. The assessed growing stock volumes of stands were classified into four strata: $0 \leq V \leq 15$ (h =1), $15 < V \leq 35$ (h=2), $35 < V \leq 85$ (h=3) and $V > 85$ m^3ha^{-1} (h=4). The stratum of a measured plot was determined based on the ocular estimate of growing stock volume of the stand, where the plot was located. The aim of stratification was also to reduce the possible bias caused by shifting the sample plots from open forest stands to stands with higher volumes.

For NFI1, the sampling errors of forest areas were estimated using the method developed for line survey by Lindeberg (1924). The error estimation is based on differences between adjacent line transects (e.g., Tomppo & Heikkinen, 1999). Lindeberg (1924) suggested dividing the long inventory lines to shorter line transects to improve the error estimation. For the error estimation, we split Finland into slices with 50 km intervals in South-North direction and divided the survey lines (SW-NE direction) according to these slices.

In estimating tree densities from NFI1 data, we used growing stock volumes assessed on survey lines as auxiliary variables in stratifying the sample plot data. The sampling variances of estimated tree densities (Table S.1) consist of two parts, which are due to the two phases of sampling (e.g., Sukhatme et al. 1984, p. 139), i.e. survey lines and sample plots on lines. The increase of variance caused by the estimation of stratum sizes using the survey line data is, however, negligible here since the sample size was large. The sampling variances of estimated tree densities were thus obtained by

$$v(\hat{\hat{y}}_{F,bc,d}) = \sum_{h=1}^{4} w_{h,bc}^2 v(\hat{\hat{y}}_{F,bc,h,d}) \qquad \text{(S.1)},$$

(e.g., Cochran 1977, pp. 333-335). We used the method of Lindeberg (1924) to estimate the within strata variances $v(\hat{\hat{y}}_{h,bc,d})$ in (S.1).

The stratification of NFI1data had only minor effect on the estimated numbers of trees. For example, the estimated number of large trees (*dbh* ≥40 cm) was 16.61 million stems using the stratification and 16.45 million stems without stratification. The estimated variance for the number of stems was, however, 12 % smaller with stratification.

The sample plots of NFI11 were located using systematic cluster sampling. Clusters consisting of 9 to 14 plots were spread evenly over the inventory region (Korhonen *et al.,* 2017). Estimation of sampling error was based on the variability of cluster-level residuals. In order to acknowledge the greater efficiency of systematic over random sampling, estimates of sampling variances were based on local quadratic forms (Matérn, 1960) within groups of clusters closed to each other. For further details, see Tomppo *et al.* (2011, section 3.5).

Table S1. The estimators of forest area, tree density and total number of trees in the intersections of the boreal subzones (b) and NFI calculation regions (c), where land areas A_c are taken from area statistics (Land Survey Finland). For NFI1, land areas A_c within the present geographical regions of Finland used (Suomen tilastollinen …, 1943). For NFI6-NFI11, areas A_c of regions used in the calculation of NFI results were employed.

Variable	Inventory	
	NFI1 (1921-1924) Line survey with sample plots of fixed area	NFI6-NFI11 (1971-2013) Angle gauge sample plots
Forest area (km^2)	$\hat{A}_{F,bc} = al_c\, l_{F,bc}$, where $al_c = A_c/l_c$, A_c is land area (km^2) of region *c*, l_c is survey line length on land in region *c*, $l_{F,bc}$ is survey line length on forest land on combined productive forest land and forest land of poor growth in the intersection of region *c* and boreal subzone *b* ($b \cap c$)	$\hat{A}_{F,bc} = a_c\, n_{F,bc}$, where $a_c = A_c/n_c$ A_c is land area (km^2) of region *c*, n_c is number of sample plot centers on land in region *c*, $n_{F,bc}$ is number of sample plot centers on combined forest land and poorly productive forest land in the intersection of region *c* and boreal subzone *b* ($b \cap c$)
Tree density (stems per ha) in forest in *dbh* class *d*	$\hat{\hat{y}}_{F,bc,d} = \sum_{h=1}^{4} w_{bc,h}\, \hat{\hat{y}}_{F,bc,d,h}$, where $w_{bc,h} = l_{F,bc,h} / \sum_{h=1}^{4} l_{F,bc,h}$ is the relative size of stand volume stratum *h*, $\hat{\hat{y}}_{F,bc,d,h} = \sum_{k \in s_h} m_{d,k} \,/\, \sum_{k \in s_h} a_k$, $m_{d,k}$ is the number of sampled trees in *dbh* class *d* on plot *k*, a_k is the area (ha) of plot *k*, s_h is the set of sample plots on combined productive forest land and forest land of poor growth belonging to stand volume stratum *h* in region ($b \cap c$)	$\hat{\hat{y}}_{F,bc,d} = \frac{10000}{n_{F,bc}} \sum_{i \in S} \frac{1}{\pi r_i^2}$, where $i \in S$ indicates that tree *i* is growing on combined forest land and poorly productive forest land in region ($b \cap c$), its *dbh* equals *d* and it was included in the sample, $r_i = \min\{50 dbh_i / 100\sqrt{q}, r_{max}\}$, q is the basal area factor, r_{max} is the maximum plot radius (m). (Tomppo *et al.*, 2011)
Total number of trees (million stems) in *dbh* class *d*	$\hat{Y}_{F,bc,d} = \hat{A}_{F,bc} \times \hat{\hat{y}}_{F,bc,d} / 10000$	$\hat{Y}_{F,bc,d} = \hat{A}_{F,bc} \times \hat{\hat{y}}_{F,bc,d} / 10000$

Tree densities and total numbers of large trees by (tree age × dbh) classes

The estimates of tree age distributions were based on tree ages measured from a sub-sample of trees. Tree ages, which include the number of years to reach 1.3 m, were classified into the following classes: 1-49, 50-99, 100-149 and ≥150 years. The proportions of trees in the age classes within each *dbh* class were estimated from the sample trees for which tree age had been measured. We used data only from temporary plots in estimating tree age distributions. The estimated proportions of age classes were multiplied by the number of trees in the corresponding *dbh* class, where the estimates were based on all trees with a measured *dbh*, including both temporary and permanent sample plots.

Supporting material 2

The variation of tree age in *dbh* classes

We created boxplots using R software and ggplot2 package (Wickham 2009) to visualize the differences in tree age distributions in the 1970s and the 2010s. Tree age distributions within each *dbh* class and subzone have changed quite a bit between 1971-2013. The median ages of trees ≤ 25 cm were smaller in 2009-2013 than in the early 1970s. Due to the fast early development the age range of small trees is narrow today compared to 1970s, but today's age distributions of large trees are wider (Fig. S1).

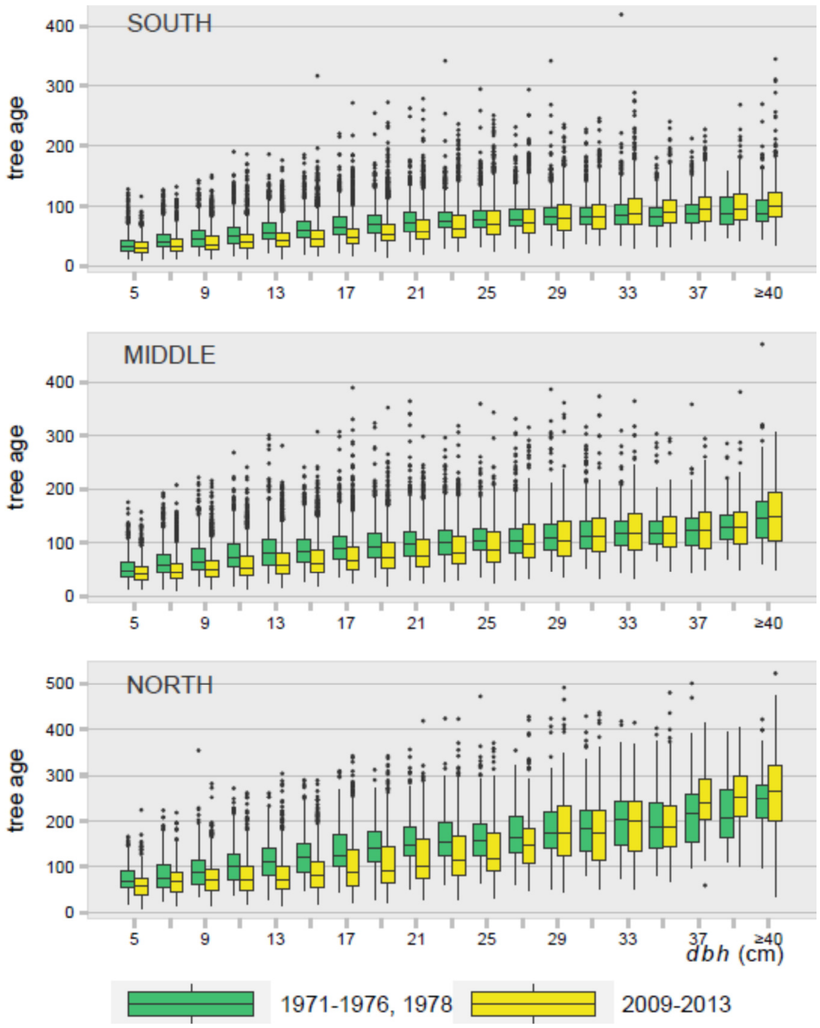

Figure S1. Boxplots for tree age in 2 cm *dbh* classes for the boreal subzones. SOUTH=southern boreal (incl. hemiboreal zone), MIDDLE=middle boreal, NORTH=northern boreal. In boxlots, the bottom and top of the box are the 25th (Q1) and 75th (Q3) percentiles, the band within the box is the median. The maximum length of the whiskers (thin lines) is 1.5×(Q3-Q1). The dots are outlier observations outside the whiskers.

Supporting material 3

Densities of trees ≥ 100 years for 2 cm *dbh* classes

Similarly to trees ≥ 150 years, most trees ≥ 100 years are quite small. Small trees in this age category have become less common and large trees more common than they were in the 1970s. The total number of trees ≥100 years has decreased by 12 % between 1971-2013. As this age category acts as an early warning signal for older trees – more useful for biodiversity – the reduction suggest that the quantities of old trees need to be monitored in future as well.

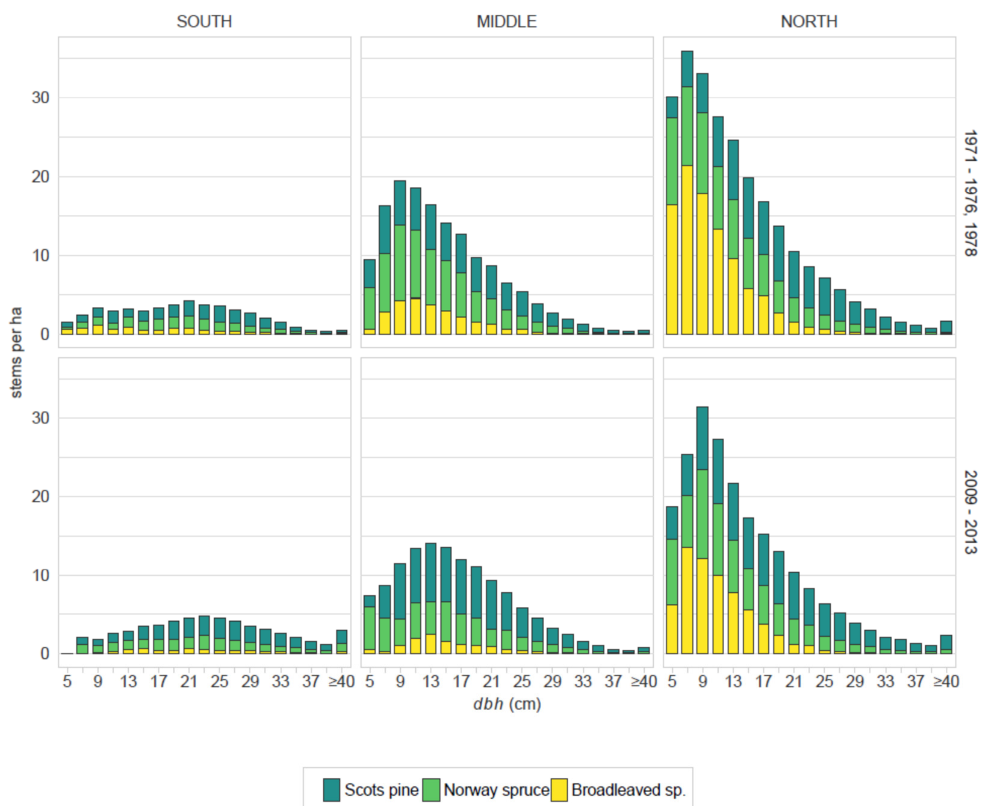

Figure S2. Tree densities (stems per ha) of trees ≥100 years) in the boreal subzones by *dbh* classes and tree species groups in 1971-1978 and 2009-2013. SOUTH=southern boreal (incl. hemiboreal zone), MIDDLE=middle boreal, NORTH=northern boreal.

Supporting material 4

Densities of old trees on peatlands

Peatlands are common in Finland (29% of the area of forest land and poorly productive forest land). The densities of old trees are higher on peatlands (41 trees ≥150 years $ha^{-1}a^{-1}$) than on mineral soils (33 trees $ha^{-1}a^{-1}$). The old trees on peatlands are on average much smaller than the old trees on mineral soils (Figure S3). The densities are especially high on peatlands of the northern boreal subzone.

The situation is similar for trees ≥100 years (Figure S3, lower panels): most of these trees on peatlands are small, in particular in the northern boreal subzone. A large share of the small old trees on mineral soils in the north consists of broadleaves. This not the case for peatlands, where the large majority of small old trees consists fairly evenly from pines and spruces.

Figure S3. Tree densities (stems per ha) of old trees (≥150 years, the two upper panels, and ≥100 years, the two lower panels) by *dbh* classes and tree species groups on mineral soils and peatlands in 2009-2013. SOUTH=southern boreal (incl. hemiboreal zone), MIDDLE=middle boreal, NORTH=northern boreal.

References

Cochran, W.G. (1977). *Sampling techniques*. 3rd Ed. John Wiley & Sons, Inc. 428 p.

Korhonen, K. T., Ihalainen, A., Ahola, A., Heikkinen, J., Henttonen, H. M. Hotanen, . . . Viiri, H. (2017). Suomen metsät 2009–2013 ja niiden kehitys 1921–2013. *Luonnonvara- ja biotalouden tutkimus*, 59/2017 , 86 p.

Lindeberg, J. W. (1924). Über die Berechnung des Mittelfehlers des Resultates einer Linientaxierung. *Acta Forestalia Fennica*, **25.5**, 1-22. doi: 10.14214/aff.7080.

Matérn, B. (1960). Spatial variation. *Meddelanden från statens skogsforskninginstitution*,**49(5)**, 1-144. Also published as *Lecture Notes in Statistics 36*. Springer-Verlag New York 1986. https://doi.org/10.1007/978-1-4615-7892-5.

Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S., & Asok, C. (1984). *Sampling Theory of Surveys with Applications*. *3rd Ed*. Iowa State University Press, Ames, Iowa (U.S.A.) and Indian Society of Agricultural Statistics, New Delhi. 526 p.

Suomen tilastollinen vuosikirja (1943). Uusi sarja **XLI,** 1-359. Tilastollisen päätoimiston julkaisema. Annuaire Statistique de Finlande. Nouvelle série 41e. Publication du Bureau Central de Statistique.
Available at
http://www.doria.fi/bitstream/handle/10024/69247/stv_1943.pdf?sequence=1&isAllowed=y

Tomppo, E., & Heikkinen, J. (1999). National Forest Inventory of Finland - past, present and future. In: Alho, J. (ed.). *Statistics, registries, and science*. *Experiences from Finland*. Statistics Finland, p. 89-108.

Tomppo, E., Heikkinen, J., Henttonen, H. M., Ihalainen, A., Katila, M., Mäkelä, H., Tuomainen, T., & Vainikainen, N. (2011). Designing and conducting a forest inventory - case: 9th National Forest Inventory of Finland. *Managing Forest Ecosystems* **21**. Springer, Heidelberg, Dordrecht, London, New York, 270 p. doi:10.1007/978-94-007-1651-3

Wickham, H. 2009. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. 213 p. doi: 10.1007/978-0-387-98141-3