Data in Brief 30 (2020) 105573

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

Dataset of tree canopy structure and variation in understory composition in a boreal forest site

Titta Majasalmi[a,*], Miina Rautiainen[a,b]

[a] *School of Engineering, Department of Built Environment, Aalto University, P.O. Box 14100, Aalto 00076, Finland*
[b] *School of Electrical Engineering, Department of Electronics and Nanoengineering, Aalto University, P.O. Box 14100, Aalto 00076, Finland*

ARTICLE INFO

Article history:
Received 27 March 2020
Revised 3 April 2020
Accepted 7 April 2020
Available online 18 April 2020

Keywords:
Leaf area index
LAI
Canopy cover
Canopy openness
Site fertility
Forest canopy
Hyytiälä
Forest floor

ABSTRACT

A field data set from 301 forest plots was collected during peak-growing season (June 24 - July 17, 2013) around Hyytiälä forestry field station in Southern Finland (61° 50′ N, 24° 17′ E). For all plots, forest variables were collected following local forest inventory practice, and understory cover fractions were estimated using a traditional sampling quadrat. The understory layer in each plot was classified into four site fertility types: herb-rich, mesic, sub-xeric, and xeric. The upper understory layer fractional covers were estimated for: (1) dwarf shrubs, (2) pteridophytes and herbaceous species, and (3) graminoids, and the lower ground layer fractional covers for: (1) mosses, (2) lichens, and (3) litter (including all non-photosynthetic material). Canopy transmittance data were collected using two LAI-2000 device. The transmittance data were used to calculate effective leaf area index, true leaf area index, canopy openness and canopy cover for all plots. The data can be used to parameterize tree canopy and understory compositions in e.g., physically-based

* Corresponding author.
E-mail address: titta.majasalmi@aalto.fi (T. Majasalmi).

https://doi.org/10.1016/j.dib.2020.105573

reflectance models, land surface models, and regional carbon cycle models. Interpretations of the results are provided in the related article [1].

Specifications table

Subject	Environmental Science (General)
Specific subject area	Physically-based reflectance models and remote sensing of forests
Type of data	Table
How data were acquired	Data were collected using the LAI-2000 device [2] which measures blue light (320–490 nm) transmittance through plant canopies with a hemispherical lens. Tree canopy transmittance data was obtained by combining data from two device recording simultaneously below and above the tree canopy. Leaf area index, canopy cover and canopy openness were estimated from this data. Forest variables were collected following local forest inventory practice, and understory cover fractions were estimated using traditional sampling quadrats.
Data format	Raw Analyzed
Parameters for data collection	Field data was collected during peak-growing season (June 24 – July 17, 2013). Forest variables and understory cover fractions were measured during daylight hours, while tree canopy transmittance data were collected during diffuse sky conditions.
Description of data collection	The study area was divided into 16 subareas (1 km × 1 km), and within each 1 km^2 subarea, twenty plots were located using a systematic sampling scheme. From the 320 plots, 301 plots were located in a forest and contained understory vegetation. The sampling scheme used to collect the plot canopy transmittance data comprised eight measurement points. For each plot, we recorded information on soil type (i.e., mineral soil or peatland) and the presence/absence of ditches. The understory layer in each plot was classified into four site fertility types: herb-rich, mesic, sub-xeric, and xeric.
Data source location	Hyytiälä forestry field station Juupajoki Finland (61° 50′ N, 24° 17′ E)
Data accessibility	Repository name: Mendeley data Data identification number: 10.17632/dyt4nkp583.1 Direct URL to data: https://data.mendeley.com/datasets/dyt4nkp583/1
Related research article	Majasalmi T., & Rautiainen, M.A. The impact of tree canopy structure on understory variation in a boreal forest. Forest Ecology and Management 465. https://doi.org/10.1016/j.foreco.2020.118100. [1].

Value of the data

- Currently, there are no large open data sets on boreal forest leaf area index (LAI, m^2/m^2), nor do the current data sets explicitly link tree canopy LAI to understory cover fractions even though such data are urgently needed in many disciplines.
- The optically measured LAI is a key variable in radiative transfer equation for vegetation [3], and thus cannot be replaced with any other ecological variable in e.g., physically-based models in remote sensing of forests or in land surface modeling.
- As data needs differ between disciplines such as remote sensing, land surface modeling, and regional carbon cycle modeling, we provide the raw data to allow a range of analyses (e.g.,

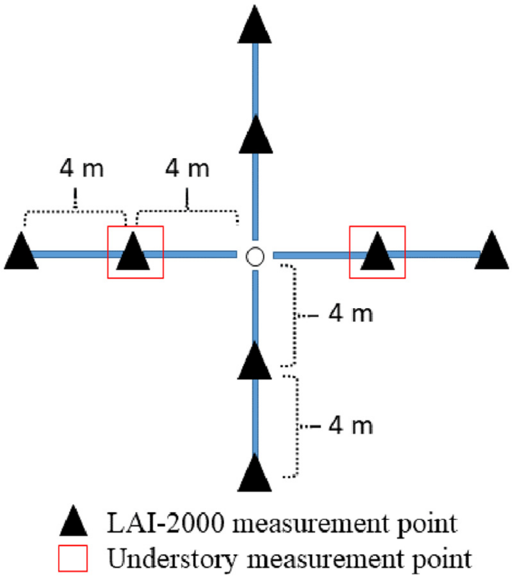

Fig. 1. Sampling scheme used to collect the plot data. The open circle is the plot center point, which was where the Bitterlich sampling, used to obtain traditional forest inventory variables, was conducted (i.e. Section 4.1), squares shows the locations of the two forest understory sub-plots (i.e. Section 3), and triangles the locations of the optical tree canopy measurements (i.e. Sections 4.1, 4.2).

forest reflectance modeling, further analysis of tree canopy and understory properties, and validation of global vegetation products based on satellite data) to suit different data needs.

- At present, for example, the development of land surface model components in climate models is hindered by the lack of data available to parameterize forest understories. The data presented here may be used to quantify the relationships between tree canopy LAI and understory compositions of different species groups, and thus foster development of more accurate regional land surface representations which may be expected to lead towards more accurate predictions of surface fluxes.

1. Data description

This article reports a dataset of boreal forest understory compositions and tree canopy properties collected from 301 forest plots in Finland. Understory composition, as described by cover fractions, was estimated with traditional sampling quadrats. Data on canopy structure (effective leaf area index, true leaf area index, canopy openness and canopy cover) were measured with the LAI-2000 device for all plots. In addition, forest variables were collected using common forest inventory practice. A cross-like sampling scheme was used to collect the plot data (Fig. 1) described in Tables 1 and 2.

The data file described in Table 1 was used as such in paper [1], which contains interpretations and conclusions related to the data. For each plot, the file contains (1) the mean understory composition, which is calculated assuming the upper understory to overlay equally the lower understory components, and (2) a shoot-level clumping corrected LAI of the canopy (as opposed to the effective LAI, which is the direct output of the LAI-2000 device). In addition, forest variables and tree species shares of total basal area, are included for each plot.

The data file described in Table 2 contains for each plot (1) vertical cover fractions of both upper and lower understory layers separately for all sub-plots as they were measured in the

Table 1
Data description including column names and variable definitions in "Processed.csv".

Region_ID:	Corresponds with a map shown in Fig. 1 in [4]
Soil_type:	1= mineral soil, 2= peatland
Ditched:	0= no, 1= yes
Fertility:	1= Herb-rich, 2= Mesic, 3= Sub-xeric, 4= Xeric
Litter:	Vertical fractional cover of litter (range 0–1)
Moss:	Vertical fractional cover of mosses (range 0–1)
Lichen:	Vertical fractional cover of lichens (range 0–1)
Shrub:	Vertical fractional cover of shrubs (range 0–1)
Herb:	Vertical fractional cover of herbaceous species (range 0–1)
Graminoid:	Vertical fractional cover of graminoid species (range 0–1)
Upper_story:	Vertical fractional cover of the upper understory (range 0–1). i.e., shrub+herb+graminoid
LAI:	Leaf area index of the tree canopy, m^2/m^2
CO:	Canopy openness, hemispherical quantity (range 0–1)
CC:	Canopy cover, vertical quantity (range 0–1)
dom_sp:	Dominating tree species, 1= pine, 2= spruce, 3= birch or deciduous
fDecid.:	Fraction of deciduous trees from stand total basal area (BA) (range 0–1)
fSpruce:	Fraction of spruce trees from stand total basal area (BA) (range 0–1)
fPine:	Fraction of pine trees from stand total basal area (BA) (range 0–1)
BA:	Stand total basal area, i.e. sum of cross-section of stems, m^2/ha
H:	Basal area weighted median tree height, m
CL:	Basal area weighted median crown length, m
DBH:	Median tree diameter at breast height, cm

Table 2
Data description including column names and variable definitions in "Raw.csv".

Region_ID:	Corresponds with a map shown in Fig. 1 in [4]
Soil_type:	1= mineral soil, 2= peatland
Ditched:	0= no, 1= yes
Fertility:	1= Herb-rich, 2= Mesic, 3= Sub-xeric, 4= Xeric
p1_Litter*:	Vertical fractional cover of litter in sub-plot p1 (range 0–1)
p1_Moss*:	Vertical fractional cover of mosses in sub-plot p1 (range 0–1)
p1_Lichen*:	Vertical fractional cover of lichens in sub-plot p1 (range 0–1)
p1_Shrub*:	Vertical fractional cover of shrubs in sub-plot p1 (range 0–1)
p1_Herb*:	Vertical fractional cover of herbaceous species in sub-plot p1 (range 0–1)
p1_Graminoid':	Vertical fractional cover of graminoid species in sub-plot p1 (range 0–1)
Upper_story:	Vertical fractional cover of the upper understory (range 0–1). i.e., shrub+herb+graminoid
LAI_eff:	Effective Leaf area index (LAI) of the tree canopy, m^2/m^2
CO:	Canopy openness, hemispherical quantity (range 0–1)
CC:	Canopy cover, vertical quantity (range 0–1)
Pine_Nha	Number of pine stems per hectare
Spruce_Nha	Number of spruce stems per hectare
Decid_Nha	Number of deciduous tree species stems per hectare
Pine_BA	Pine basal area (BA) i.e. sum of cross-section of stems, m^2/ha
Spruce_BA	Spruce basal area (BA)
Decid_BA	Basal area (BA) of deciduous tree species
BA:	Forest plot total basal area (BA)
H:	Basal area weighted median tree height, m
CL:	Basal area weighted median crown length, m
DBH:	Median tree diameter at breast height, cm

* Note, data contain two sub-plot data pairs (i.e. 'p1_' and 'p2_'). Examples provided only for 'p1_' to avoid repetition.

field conditions, and (2) effective canopy LAI (which is the direct output of the LAI-2000 device). In addition, forest variables for each plot are provided. The column names and definitions of variables of the preprocessed data file ("Processed.csv") are provided in Table 1 and that of the raw data file ("Raw.csv") in Table 2.

2. Experimental design, materials, and methods

2.1. Study site

Study site was located in Hyytiälä forestry field station, Juupajoki, in Southern Finland (61° 50′ N, 24° 17′E). Field measurements were conducted during peak-growing season (June 24 – July 17, 2013) in an area of ~16 km^2. The mean annual precipitation in the study site is 700 mm, and mean annual temperature is 3 °C. Hyytiälä area is under common forest management practices with periodical thinnings (e.g., rotation period varies from 60 to 120 years), and the forest land is owned by the state and private sector. Forests around Hyytiälä are dominated by Norway spruce (*Picea abies*), Scots pine (*Pinus sylvestris*) and birches (*Betula pubescens* and *Betula pendula*). Although monocultural birch stands are rare, birch is common in mixed species stands.

2.2. Sampling

Field plots were spatially distributed into 16 subareas (each 1-km × 1-km). In each subarea, a cluster of 20 plots was located using a systematic sampling scheme. The distance between the plots within each cluster was 100 m in the south-north direction, and 150 m in the east-west direction. If the individual plot location did not fall in a forest (e.g., it was on a road), it was moved in steps of 10 m (but not exceeding a total of 30 m) in either of the cardinal directions. The initial number of plots was 320, of which 307 were located in a forest, but only 301 had understory vegetation. Due to private land ownership, the plot coordinates are not provided (however Fig. 1. in [4] shows the spatial mapping of plot locations (i.e. 'Region_ID' -column) around Hyytiälä area with respect to in MODIS satellite data pixels).

3. Measurements of understory

The understory layer in each plot was classified into four site fertility types: herb-rich, mesic, sub-xeric, and xeric. Note that the same terminology has been previously used in e.g., [5,6] to study spectral properties of understory vegetation. For each plot, information on soil type (i.e. mineral soil or peatland) and thus also the presence/absence of ditches was recorded.

The cover fractions of understory were estimated from two 1 m × 1 m understory sub-plots, located 4 m west and east from the plot center, using traditional sampling quadrats (i.e. visually based estimates of the vertical cover fractions). The upper understory layer fractional covers (in the vertical direction) were estimated for: (1) dwarf shrubs, (2) pteridophytes and herbaceous species (later called as 'herbs'), and (3) graminoids, and the lower ground layer fractional covers for: (1) mosses, (2) lichens, and (3) litter (including all non-photosynthetic material). The same person estimated all cover fractions. Example species compositions in different site fertility types are provided in Table 1 in [1].

4. Measurements of tree canopy

4.1. Traditional forest inventory variables

The sample plot ($n = 301$) centers were located using a GPS, and the plot center points was marked using a fiber ribbon. From the plot center Bitterlich-sampling (i.e., sampling proportional to size of the breast-height-diameter (DBH, cm) of the trees) was carried out. Basal area (BA, m^3/ha) was measured from the plot center, whereas the DBH and tree height (H, m), and crown length (CL, m) were measured for the median basal area tree. For plots with small trees, the BA was estimated from the stem number and DBH (i.e. N*pi/4*DBH^2, where DBH is in meters, and

N is the number of stems per hectare). The plots were classified based on tree species with the largest BA into spruce (i.e. fSpruce), pine (i.e. fPine) or deciduous (fDecid.) dominated.

4.2. Optical tree canopy measurements

Optical data was measured using two simultaneously operating LAI-2000 devices [2], which measured diffuse sky radiation (wavelength region 320–490 nm) in five zenith angle bands (ranges: 0°–13°, 16°–28°, 32°–43°, 47°–58° and 61°–74°) with a hemispherical lens. The field-of-view (FOV) of the lens is ~150°. Measurements were conducted while the Sun elevation was lower than 16° above the horizon, or under fully overcast conditions, to avoid direct radiation reaching the sensors FOV. No view restrictors were used (i.e. to mask part of the hemisphere) during measurements. The two LAI-2000 units were used simultaneously to measure above and below the tree canopies, and they were inter-calibrated before the measurements. The reference LAI-2000 unit was located on top of the tower for radiation measurements in Hyytiälä in automatic mode to log readings every 15 s, while the other LAI-2000 unit was operated in manual mode under the tree canopies in forest plots. The below-canopy LAI-2000 unit measurement height was 1.6 m (i.e., the person making the measurements was able to stand while recording the readings of individual measurement points). LAI-2000 data was always collected after all other field measurements had been completed in order to avoid walking on the understory plots.

The sampling scheme used to collect the below canopy readings (Fig. 1) was a modified version that used by the Validation of Land European Remote Sensing Instruments (VALERI) cross-scheme. While VALERI contains twelve measurement points in each cardinal direction at two, six and ten meters from the plot center point) [7], the modified version used eight measurement points in each cardinal direction at four and eight meters distance from the plot center point. The modified version was used because it allows time saving in field measurements, without compromising the accuracy and bias of the original VALERI scheme [8].

For each plot, the mean canopy transmittance was obtained as the ratio of means of below and above canopy sensor readings from the two LAI-2000 units. The data from the two sensors was combined and processed by the software that comes with the LAI-2000 device (i.e., FV2000) [2]. Three outputs were obtained from the FV2000 software: the effective canopy leaf area index (LAI_{eff}), diffuse non-interceptance (i.e., canopy openness, CO) and zenith gap fraction (i.e., canopy cover, CC).

The FV2000 software calculates LAI_{eff} based on gap fraction data of five concentric rings centered at zenith angles θ based on the inversion of canopy transmittance $T(\theta)$ according to Beer's law equation [2,9] as:

$$LAI_{eff} = 2\sum_{0}^{\frac{\pi}{2}} -ln[T(\theta)]\cos(\theta)\sin(\theta)d\theta \tag{1}$$

For conifer canopies the LAI calculation implemented by the FV2000 software (Eq. (1)) underestimates the 'true' LAI due to needles clustering into shoots. Thus a standard correction (i.e. called as shoot-level clumping (SLC) correction) was applied for LAI_{eff} (see e.g., [10]). The true LAI was obtained by dividing the LAI_{eff} with a BA-weighted mean SLC factor. The SLC factors were 0.59 for pine [11] and 0.64 for spruce [12]. No SLC correction was applied for deciduous species.

The smallest zenith angle ring (of 0–13°) transmittance data was used to approximate CC as: 1-$T(\theta)$, whereas CO was calculated by the FV2000 software [2] as:

$$CO = \frac{\int_0^{\frac{\pi}{2}} \Gamma(\theta)T(\theta)\cos(\theta)\sin(\theta)\ d\theta}{\int_0^{\frac{\pi}{2}} \Gamma(\theta)\cos(\theta)\sin(\theta)\ d\theta} \tag{2}$$

where $\Gamma(\theta)$ is the intensity distribution of the sky radiation above the canopy.

Acknowledgments

We thank Anni-Sofia Hoppi and Johanna Karjalainen for field work. The data were collected in a project funded by the Academy of Finland in 2010–2015 (grant number: 135502). Completion of this article has received funding from Aalto ENG postdoctoral funds and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 771049). The text reflects only the author's view and the Agency is not responsible for any use that may be made of the information it contains.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

References

[1] T. Majasalmi, M.A. Rautiainen, The impact of tree canopy structure on understory variation in a boreal forest, For. Ecol. Manage. 465 (2020), doi:10.1016/j.foreco.2020.118100.
[2] LI-COR, LAI-2000 Plant Canopy Analyzer: Instruction Manual, LI-COR, Inc., Lincoln, Nebraska, 1992 https://licor.app.boxenterprise.net/s/q6hrj6s79psn7o8z2b2s.
[3] J. Ross, The Radiation Regime and Architecture of Plant Stands, Springer, The Hague, 1981, p. 391.
[4] T. Majasalmi, M. Rautiainen, P. Stenberg, T. Manninen, Validation of modis and GEOV1 fPAR products in a boreal forest site in Finland, Remote Sens (Basel) 7 (2) (2015) 1359–1379.
[5] M. Rautiainen, M. Mõttus, J. Heiskanen, A. Akujärvi, T. Majasalmi, P. Stenberg, Seasonal reflectance dynamics of common understory types in a northern European boreal forest, Remote Sens. Environ. 115 (2011) 3020–3028, doi:10.1016/j.jag.2014.11.012.
[6] M. Nikopensius, J. Pisek, K. Raabe, Spectral reflectance patterns and seasonal dynamics of common understory types in three mature hemi-boreal forests, Int. J. Appl. Earth Observat. Geoinformat. 43 (2015) 84–91, doi:10.1016/j.rse.2011.06.005.
[7] Validation, 2020. Validation of land European remote sensing instruments (VALERI) network. [Accessed on March 4th 2020]. Online at: http://www.avignon.inra.fr/valeri/.
[8] T. Majasalmi, M. Rautiainen, P. Stenberg, H. Rita, Optimizing the sampling scheme for LAI-2000 measurements in a boreal forest, Agric. For. Meteorol. 154 (2012) 38–43.
[9] M. Monsi, T. Saeki, The light factor in plant communities and its significance for dry matter production, Jap. J. Botany 14 (1953) 22–52.
[10] T. Majasalmi, M. Rautiainen, P. Stenberg, P. Lukeš, An assessment of ground reference methods for estimating LAI of boreal forests, For. Ecol. Manage. 292 (2013) 10–18, doi:10.1016/j.foreco.2012.12.017.
[11] H. Smolander, P. Stenberg, S. Linder, Dependence of light interception efficiency of scots pine shoots on structural parameters, Tree Physiol. 14 (1994) 971–980, doi:10.1093/treephys/14.7-8-9.971.
[12] P. Stenberg, S. Linder, H. Smolander, Variation in the ratio of shoot silhouette area to needle area in fertilized and unfertilized Norway spruce trees, Tree Physiol. 15 (1995) 705–712, doi:10.1093/treephys/15.11.705.