Supplement of Biogeosciences, 15, 399–412, 2018
https://doi.org/10.5194/bg-15-399-2018-supplement

Biogeosciences
Open Access

Supplement of

An enhanced forest classification scheme for modeling vegetation–climate interactions based on national forest inventory data

Titta Majasalmi et al.

Correspondence to: Titta Majasalmi (titta.majasalmi@nibio.no)

S1. Preprosessing of Norwegian MS-NFI data

As the northernmost forest area (i.e. Finnmark county) is not covered by SAT-SKOG (Gjertsen, 2009) a forest resource map called 'AR5' (Ahlstrøm et al., 2014) and NFI field plot data were used to model the forest cover information for Finnmark. The AR5 forest mask was used to define the forest extent in Finnmark. A moving window interpolation was used to fill in data gaps between the NFI field plots located inside the forest maks. As SAT-SKOG does not contain tree height information, it was modeled based on available data of tree species, tree age at breast-height (i.e. 1.3 m above ground surface), and site index using equations by Tveite and Braastad (1981). Separate equations were used for pine, spruce, and birch (birch model was applied to all deciduous species). To calibrate the mean height to correspond with the Lorey's height (H), a separate model was developed based on Norwegian NFI data to scale the plot mean height into H. The H model was developed based on Norwegian NFI field plot data to predict the H based on plot median tree height (H_m) and plot total stem volume (V) (**Table S1.**). Median tree was used instead of mean tree while developing the model, because median tree exists in the data and is less affected by outliers than the mean. Coefficient of determination (variance) explained by fixed effects was 0.82, and both fixed and random effects was 0.98. Model RMSE was 1.40. The model was not applied if either pixel H_m or V was zero.

Table S1. Description of the Lorey's height model (H ~ H_m+V + ε.). Abbreviations: H=Lorey's height (m), H_m=median tree height (m), V=plot total stem volume (m^3/ha), sd=standard error and ε =residual.

			Fixed effects:	Random effects:	
Area	Species		value	sd (plot)	ε
Norway	All	Intercept	3.015	1.841	0.692
		H_m	0.756		
		V	0.013		

References

Ahlstrøm, A. P., Bjørkelo, K. and Frydenlund, J. AR5 KLASSIFIKASJONSSYSTEM: Klassifikasjon av arealressurser. Rapport fra Skog og landskap 06/14: III, 38s. http://www.skogoglandskap.no/filearchive/rapport_06-2014.pdf , 2014.

Gjertsen, A.K. SAT-SKOG kom på Internett i 2009. Årsmelding fra Skog og landskap 2009: 27. http://www.skogoglandskap.no/filearchive/sat_skog_kom_pa_internett_i_2009.pdf , 2009.

Tveite, B. and Braastad H. Bonitering for gran, furu og bjørk. Norsk Skogbruk 27, 17-22, 1981.

S2: Leaf area index calculation

Allometric tree-wise foliage mass models were first used to estimate total foliage mass of trees within a sample plot, and LAI was calculated by dividing plot total foliage mass with plot area and multiplying with Specific Leaf Area (SLA) value (i.e. conversion from mass to area). Foliage masses for pine and spruce trees were calculated using Marklund's (1988) biomass models. Foliage mass of deciduous trees was calculated using the biomass model of Smith et al., (2014), because Marklund does not have separate model for deciduous trees. All three models were developed under Fennoscandic boreal forest conditions. NFI plot data contained measurements of diameter-at-breast-height (dbh) and tree height (h) from all trees within a circular 8.92 m radius sample plot (Note, in Swedish NFI the plot radius was 10 m). In Norwegian NFI plot data, small trees were counted based on either h or dbh (i.e. three classes: h<1.3 m, dbh=0-2.5 cm and dbh=2.5-5 cm), whereas in Swedish NFI data, small trees were counted only based on h (i.e. two classes: h<0.5 m and h<1.3 m).

Marklund's tree dbh- and h -dependent models were used to calculate foliage mass of pine and spruce trees (T-18 and G-16, respectively), whereas for small trees the foliage mass was calculated using Marklund's dbh-dependent models (T-17 and G-15). For Norwegian small trees, foliage mass was obtained based on class (i.e. dbh= 0-2.5 cm and dbh= 2.5-5 cm) mean diameter values (i.e. 1.25 cm and 3.75 cm, respectively) and number of trees. For the smallest height class (i.e. h<1.3 m), the dbh-constant was set to 0.3 cm, because dbhs for trees shorter than 1.3m were not measured. For the two Swedish height classes (i.e. h<0.5 m and h<1.3 m), the dbh-constants were set to 0.1 cm and 0.5 cm. LAI was calculated by dividing plot total foliage mass with plot area and multiplying with SLA value. The SLA values were: 6.2 m^2/kg for pines (Palmroth and Hari, 2001), 4.95 m^2/kg for spruces (Stenberg et al., 1999), and 13.55 m^2/kg for deciduous trees (Lintunen et al., 2011). While the clustering analysis was performed using Swedish and Norwegian NFI data, the LUT may be assumed applicable in Finland because the biomass functions by Marklund (1988) are applicable in Finland (Kärkkäinen, 2005) and given the similarities in commercial species and forest management practices in Fennoscandia.

The highest mean canopy LAI values (LAI_{canopy}) were found from spruce dominated plots and the smallest in plots dominated by either pine or deciduous species (**Table S2.**). For conifer dominated plots the mean LAI_{canopy} values were higher in Norway than in Sweden. The influence of small trees (LAI_u) on plot mean LAI were systematically larger in Norway than in Sweden, due to differences in definition of small trees in NFI measurements (i.e. in Norwegian NFI trees with dbh less than five centimeters are recoded as small trees, whereas in Sweden the breast height (1.3 m) threshold is used to count small trees). In addition, as Norway has alpine birches growing in mountainous areas the influence of LAI_u on plot total mean LAI was larger in Norway than in Sweden.

Our results showed that the LAI_u accounted for on average 2-8 % of total LAI, and thus should be taken into account while estimating the estimating the forest LAI. The contribution of LAI_u on total LAI was the largest in deciduous plots, which are in the center of climate change mitigation and adaption by forest management. To our knowledge, this is the first paper to report how the information of small trees in the NFI data may be used to estimate LAI of understory trees to approximate forest total canopy LAI. Error propagation of LAI estimates is not possible, because the ‘true’ LAI is not known (for details see Majasalmi et. al., 2013). Note, in our paper the sum of LAI_{canopy} and LAI_u is called LAI_{max}, and it may be interpreted as the maximum growing season LAI of all forest canopy layers.

Table. S2. The mean Leaf Area Index (LAI) of plots dominated by different species groups: LAI_{canopy} is the forest canopy LAI, LAI_u refers to LAI of small trees i.e. understory trees (Note, the different counting method of small trees in Norway and in Sweden), and $LAI_{u\%}$ is the percentage of LAI_u for the total mean LAI (LAI_{canopy}+LAI_u).

	Norway			Sweden		
	LAI_{canopy}	LAI_u	$LAI_{u\%}$	LAI_{canopy}	LAI_u	$LAI_{u\%}$
Spruce	5.59	0.24	4.05	4.90	0.10	1.91
Pine	2.58	0.12	4.34	2.31	0.07	2.83
Deciduous	2.33	0.21	8.31	2.76	0.15	5.17

References

Kärkkäinen, L. Evaluation of performance of tree-level biomass models for forestry modeling and analyses. Finnish Forest Research Institute, Research papers, 940: p123, 2005.

Lintunen, A., Sievänen, R., Kaitaniemi, P., and Perttunen, J. Models of 3D crown structure for Scots pine (Pinus sylvestris) and silver birch (Betula pendula) grown in mixed forest. Can. J. For. Res., 41, 1779-1794, 2011.

Majasalmi, T., Rautiainen, M., Stenberg, P., and Lukeš, P. An assessment of ground reference methods for estimating LAI of boreal forests. For. Ecol. Manage., 292, 10-18, 2013.

Marklund, L. G. Biomassafunktioner för tall, gran och björk i Sverige. Sveriges lantbruksuniversitet, Institutionen för skogstaxering, 1988.

Palmroth, S. and Hari, P. Evaluation of the importance of acclimation of needle structure, photosynthesis, and respiration to available photosynthetically active radiation in a Scots pine canopy. Can. J. For. Res., 31, 1235-1243, 2001.

Smith, A., Granhus, A., Astrup, R., Bollandsås, O. M. and Petersson, H. Functions for estimating aboveground biomass of birch in Norway. Scand. J.Forest Res., 29, 565-578, 2014.

Stenberg, P., Kangas, T., Smolander, H., and Linder, S. Shoot structure, canopy openness, and light interception in Norway spruce. Plant Cell Environ., 22, 1133-1142, 1999.

S3: Crown ratio models

Crown Length (CL) can be modeled directly or alternatively via Crown Ratio (CR), which is defined as CL divided by h (i.e. CL = h*CR). We used a mixed effects modeling approach with maximum likelihood estimator to develop the CR models for the three main species (i.e. spruce, pine, and birch) in Norway and in Sweden. Separate CR models were developed for Norway and Sweden to account for possible differences in forest management (i.e. harvesting) and environmental conditions. The number of trees used to develop the CR models was nearly 24,000 (i.e. spruce: 10,403, pine: 7,830 and deciduous: 5,828) in Norwegian NFI data. In Swedish NFI data the number of trees used to compile the models was almost 29,500 (i.e. spruce: 12,228, pine: 13,171 and deciduous: 4,081). The model was created based on forest variables which are influenced by forest management operations and are widely available from forest inventory databases (i.e. species, h, and V). The number of predictor variables was selected based on the Akaike Information Criterion (AIC), which quantifies the trade-off between model complexity and fit. To account for non-independencies between observations (i.e. several model trees from one plot) model intercept was allowed to vary between individual trees and plots. CR models were developed based on NFI trees for which both h and CL were measured (**Fig. S1**).

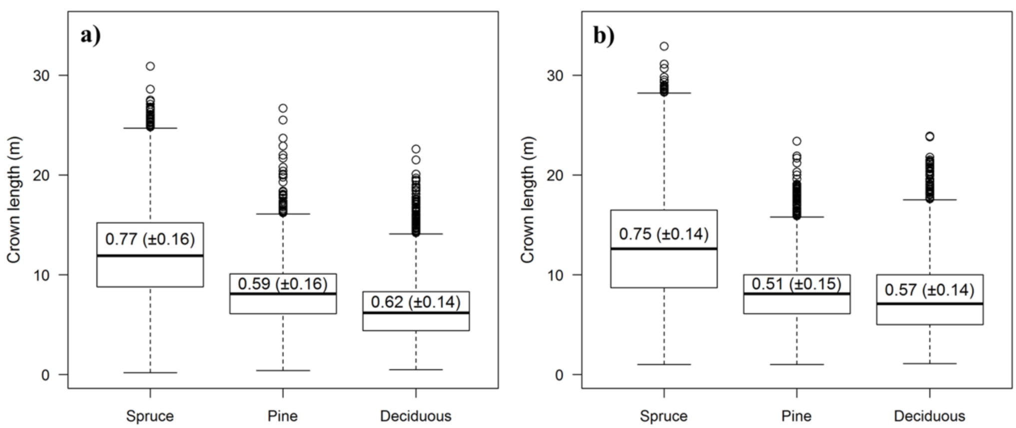

Figure S1. Individual tree data for developing the Crown Ratio (CR) models based on **a)** Norwegian National Forest Inventory (NFI) trees, and **b)** Swedish NFI trees. Mean and standard deviation (±sd) of measured CR are plotted inside the boxes. Box contains 50% of the values and dark horizontal line denotes the median value.

The fixed effect of the CR was modeled based on h and V (**Table S3.**). Both h and V were found to be significant predictors for CR (p=0.00) for each species (i.e spruce, pine, and birch). The variance explained by fixed effects of the CR models varied between 0.11-0.37 (**Table S4.**). In Norway, the largest portion of variance explained by fixed effects was noted for spruce CR model, whereas in Sweden the pine CR model explained the highest portion of variance of fixed effects. For both countries the variance explained by fixed effect remained the lowest for birch CR models. Including both fixed and random effects increased the portion of explained variance to range between 0.45 and 0.63.

The relatively low explanatory power of the CR models is due to large scatter of CR estimates. Higher prediction power could have been obtained by developing models for CL, which has larger dynamic range compared CR. However, by modeling CR instead of CL, the problem of negative CLs was avoided, and as a tradeoff we had to accept the lower prediction power of our CR models. CL reflects the vigor of trees and is influenced by both stand density via competition over time and environmental factors (e.g. water availability and temperature). Increasing stand density reduces CL due to decreasing light conditions near the crown base (i.e. foliage respiration exceeds the net photosynthesis) and mechanical damage as tree canopies collide with each other. Thus, thinnings have important role in controlling CL development and hence to the aerodynamic properties of forests.

Table S3. Description of the Crown Ratio (CR) models (CR ~ h + V + ε.). Abbreviations: h=tree height (dm),V=plot total stem volume (m^3/ha), sd=standard error and ε=residual.

		Fixed effects:		Random effects:		
Area	Species		value	sd (plot)	sd (tree)	ε
Norway	Spruce	Intercept	0.821	0.055	0.055	0.100
		h	0.008			
		V	-0.001			
	Pine	Intercept	0.659	0.067	0.067	0.102
		h	0.005			
		V	-0.001			
	Birch	Intercept	0.612	0.058	0.058	0.105
		h	0.008			
		V	-0.001			
Sweden	Spruce	Intercept	0.802	0.058	0.056	0.101
		h	0.006			
		V	-0.001			
	Pine	Intercept	0.618	0.065	0.066	0.091
		h	0.001			
		V	-0.001			
	Birch	Intercept	0.606	0.059	0.059	0.105
		h	0.005			
		V	-0.001			

Table S4. Variance explained by fixed effects and both fixed and random effects of the Crown Ratio (CR) models. RMSE is the Root Mean Square Error.

Area	Species	RMSE	Fixed effects	Fixed + Random effects
Norway	Spruce	0.11	0.37	0.61
	Pine	0.11	0.26	0.60
	Birch	0.11	0.11	0.45
Sweden	Spruce	0.11	0.20	0.51
	Pine	0.10	0.24	0.63
	Birch	0.10	0.14	0.47

S4. Background information concerning MS-NFI data

The forest masks used to apply the kNN are developed by forest authorities, and thus mask definitions may slightly vary between countries. In addition, statistical calibration is applied at municipality level to correct errors in the map data. The calibration is based on a confusion matrix between land use classes of the sample plots recorded in the field and extracted from the raster map. The forest area on the NFI maps is not corrected, and thus it is commonly overrepresented on the map (i.e. a pixel inside forest mask may have V=0 or H=0. Note, however, that in our classification such pixel would not be classified as forest). Due to local calibration, the values from the NFI maps do not directly correspond with the official forestry statistics. If forest authorities would provide these calibrated MS-NFI maps instead of the uncalibrated maps the MS-NFI estimates and forest statistics would agree. The underlying assumptions of our classification scheme are that NFI data represents different forest types within a country (i.e. samples the whole population) and that MS-NFI data represent the extent and structural variation of forests spatially. MS-NFI data are intended to represent large forest areas, and while the errors at pixel level are relatively high at fine spatial resolution, the error decreases as the size of the estimation area increases. For example, in Finnish MS-NFI the average error of the V estimates at pixel level is 57.8 m^3/ha and H is 4.6 m (errors calculated as an average of mineral soil and peatland estimates for Finnish MS-NFI 2009 products and are reported in MS-NFI 2013 metadata). According to Tomppo et al. (2014): “For a sufficiently large area consisting of a group of pixels, e.g., for areas of 200 000–300 000 ha, the MS-NFI estimates are compared to the estimates and error estimates based solely on field data”. As the land area of e.g. Finland is 33,842,400 ha, and majority of the land area is covered by forest, the estimation errors of the MS-NFI data may be assumed small.

Tomppo, E., Katila, M., Mäkisara, K. and Peräsaari, J. The Multi-source National Forest Inventory of Finland - methods and results 2011. Metlan työraportteja / Working Papers of the Finnish Forest Research Institute 319. 224 p. http://www.metla.fi/julkaisut/workingpapers/2014/mwp319.htm, 2014.

S5. Enhanced LC-product and percentage layers

In the final product, the forest pixels are recoded as follows: The twelve forest classes were recoded by adding 300 (e.g. "Deciduous forest subgroup 4" would be coded as 312 - see **Table 2.**). In addition, two digits were added after recoding the forest class number to indicate the presence of forest cover within the LC-pixel and whether the pixel is a 'true' forest pixel based on MS-NFI data data or gapfilled. The fourth digit is used to indicate the fraction of forested pixels within an LC-pixel: Value '1' indicates that the fraction of forested pixels within an LC-pixel is >40%, and value '2' denotes that the fraction of forested pixels within an LC-pixel is between 15-40%. For 'true' forested pixels the last digit is '0', whereas for gapfilled pixels the last digit is '1'. For example, a pixel with value "30210" can be interpreted as "Spruce, subgroup 2, >40% forest cover fraction, "true-forest" (i.e. non-gapfilled).

For land models which allow more than one land cover type, the percentages of each subgroup for each LC-pixel are provided as separate layers. Note, these layers do not sum up to 100, unless each MS-NFI pixel within the LC-pixel is classified as forest. LC- pixels where the portion of forested pixels in MS-NFI data was less than 15% were removed from these layers to correspond with the enhanced LC-product. In addition, gapfilled pixels are provided as separate layers which allow users to either include or exclude these from computations. To construct a complete surface representation, the other land cover types (i.e. non-forested pixels) from the ESA LC-product are provided as a separate layer which can be summed with the percentage layers.