HRL verification report template for dominant leaf type in Finland

I. Administrative part

HRL	Dominant Leaf Type
Country (and region, if regions are verified separately)	Finland
Institution carrying out the work	Natural Resources Institute Finland (Luke)
General overview of data quality done by (name, position and e-mail)	Hanna Huitu, research scientist, hanna.huitu@luke.fi Matti Katila, research scientist, matti.katila@luke.fi Sakari Tuominen, senior scientist, sakari.tuominen@luke.fi
Look-and-feel analysis done by (name, position and e-mail)	Hanna Huitu
Statistical verification done by (name, position and e-mail)	Hanna Huitu Matti Katila
In situ data used. *Replace Data-x with the full name of the dataset. Mention quality issues if relevant.*	National forest inventory (NFI) field plots, from systematic cluster sampling 2013, except for northern Lapland (see Fig. 3, sampling regions) 2012 (n=9766). All land cover types (except sea) are represented in the data.
	Finnish multisource-NFI thematic map of canopy cover and canopy cover of broadleaved trees 2015.
	False Colour Aerial Photographs from National Land Survey, year 2015 (used as a WMS layer)
	Corine CLC2012 land cover map
	Topographic database of National Land Survey
Internal quality control done by (name, position and e-mail)	
Date and place of writing the report	8.2.2019

II. General overview of data quality

Results of the general overview of data quality (obligatory)	*Object to the verification was the High Resolution Layer for Dominant Leaf Type (2015). The verified data set was a raster at 20 m * 20 m resolution, with pixel values denoting whether the tree cover was dominated by coniferous or broadleaved tree species, or whether there was no tree cover at all (tree cover density less than 1 %). No minimum mapping unit was used.* ***Classification error matrices*** *are placed to the end of part IV Statistical verification (Table 1a. and Table 1b.).* ***Geometric accuracy:*** *Level of geometric accuracy was good. Based on overlaying this product with topographic data layers, the product was not found to contain shifts or other major problems of geometric accuracy.* ***Thematic accuracy:*** *In general, leaf type dominance was detected at a reasonably good level, and areal patterns of coniferous- and broadleaved dominance were in line with the reference data used. Detection of the presence or absence of tree cover (either coniferous or broad-leaved) was sometimes a problem. For these errors see the evaluation report for HRL TCD product for additional information.* ***Errors of commission and omission*** *are assessed in detail in parts III and IV.* ***Issues found in this verification:*** *i. Especially in young mixed forests, forests dominated by coniferous species were often erroneously classified as dominated by broadleaved species, and vice versa. Classification accuracy improved as trees matured (See Table 2)* *ii. Areas with young forest or low tree cover were often erroneously classified as not having any trees, and as a consequence, the HRL layer did not contain information of their leaf type dominance.* **Observations from the visual comparison** Visual comparison was carried out between the HRL DLT product and dominant leaf type of the MS-NFI derived from the canopy cover theme and the canopy cover of the broadleaved trees. Comparing the two products there are more non-tree covered areas in the HRL DLT product. This is partly due to the more even distribution of the canopy cover in the NFI field data used as training data for the MS-NFI product and partly due to the k-NN estimation method employed. Therefore a minimum canopy cover of 2 % for tree species to be defined was used with the MS-NFI. There are more broadleaved dominated areas in the HRL DLT product in the south of Finland. Otherwise the forest stands in the broadleaved class are distributed in a similar way over the country; see section V Fig. 4 HRL TCD and Fig. 5 MS-NFI-2015 canopy cover.

i Mixed forests were challenging to classify

About 13 % of the (true) coniferous-dominated plots were erroneously classified as being broad-leaved dominated (See Fig. 1 map on the right). Likelihood of this classification error increases, as the true share of broad-leaved canopy approaches 50 %. Share of true broadleaved-dominated plots erroneously classified as coniferous-dominated was roughly at the same level (11 %). Note that in Fig.1, the higher number of errors on the map on the right compared to map on the left reflects higher number of observations in coniferous-dominated ground truth plots.

Fig. 1 Errors of mixing broadleaved- and coniferous dominance. On the left, cases where ground truth information is broad-leaved dominated, but HRL classification finds coniferous-dominated tree cover. On the right, cases where ground truth information is coniferous-dominated tree cover, but HRL classifies the plots broadleaved-dominated.

ii) For small trees or low tree cover densities, HRL did not find any (either broad-leaved or coniferous) tree cover

On 14 % of the examined 9766 plots, field inventory had recorded trees on the plots, but no tree cover was present in the HRL layer (see the right map on Fig. 2). As the DLT layer and tree cover density layer are connected, this error seems to reflect errors of tree cover density layer.
Young forest stands were difficult to classify correctly but mature forest was easier. The share of correctly classified plots increased in a consistent way for both leaf types, as forest matured through the 5 consecutive forest development stages from seedlings to mature forest (see Table 2).

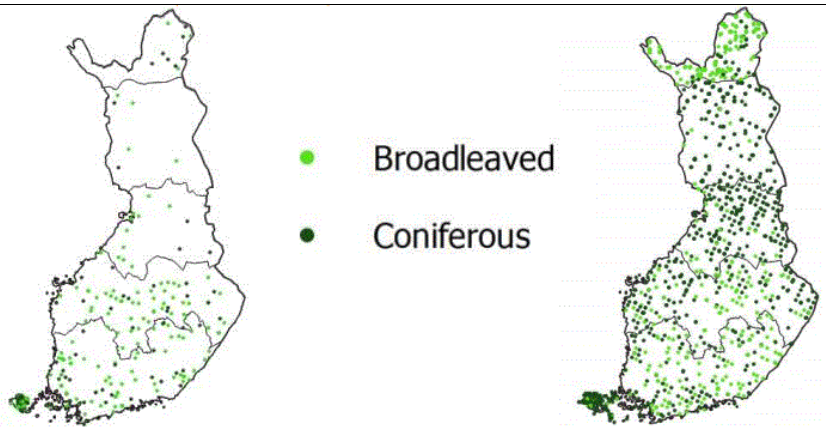

Fig. 2 Errors on detecting presence of tree cover. On the left, cases where ground truth information is no tree cover, but HRL classification finds either coniferous- or broadleaved dominated tree cover. On the right, cases where ground truth information is either broadleaved or coniferous-dominated tree cover, but HRL classification finds no tree cover.

III. Look-and-feel (obligatory)

Stratum	Name of the stratum (see proposed strata in Tables 17-21)	Number of samples verified	
1	Urban vegetation (Trees in parks, cemeteries, etc.)	10	**Good.** Open green areas (with no tree cover) were often classified correctly. Some cases found where open grass fields falsely classified as broad-leaved trees. Raster cell size (20 m) is large in context of small-scale urban landscape, posing challenge to the classification.
2	Trees in sport and recreation areas	10	**Acceptable**. Both open green areas and groups of trees were often classified correctly. Some cases where open fields had been erroneously classified as broad-leaved tree cover.
3	Orchards, fruit trees	5	**Acceptable**. If tree crowns were found, they were correctly classified to respective classes.
.4	Forest along rivers & lakes	5	**Acceptable.** Separation of broad-leaved and coniferous forests was challenging in cases of mixed tree cover.
5	Coastal forests	5	**Acceptable**. Separation of broad-leaved and coniferous forests was challenging in cases of mixed tree cover.
6	Agricultural areas with scattered small forest patches (if ≥ 0.5 ha)	10	**Acceptable**. Small forest patches on agricultural areas are often broadleaved-dominated. This was captured correctly. Separation of broad-leaved and coniferous forests was challenging in cases of mixed tree cover.
7	Non-tree woody vegetation (Transitional woodland-shrub, moors and heathland, sclerophyllous vegetation)	10	**Insufficient.** Checked based on in situ information on the existing tree cover. Several commission errors in cases with zero tree canopy ground truth, many of them in the northern vegetation zones(Lappland inventory area).See V for details.
8	Wetland	10	**Insufficient.** Patchy, detailed textures of vegetation and water on wetlands may resemble tree canopies, and most error-prone locations were visually looked up and checked for commission errors. Some commission errors were found. See V for details.
A1	Peatland areas with scattered small forest patches (if ≥ 0.5 ha)	5	**Insufficient**. While most of the inspected ≥ 0.5 ha forest patches were found, the classification was not able to sufficiently distinguish them from the surrounding peatland area of no tree cover, but commissions were occurring and thus also dominant leaf type was not correctly classified. Use of a peatland mask is recommended to support future production efforts
A2	Peat production areas	5	**Acceptable.** Peat production areas in use were correctly classified as not having tree cover, and forest re-growth on areas no longer in use was often detected correctly
A3	Seedling stands	10	**Insufficient**. The HRL classification of dominant leaf type commonly showed disagreements with measured ground truth data. However, detection of dominant leaf type in young development classes in general is very difficult, due to small size of trees and natural mixing of broad-leaved and coniferous trees.
Overall evaluation			*acceptable*
Comments			Overall, dominance of the leaf types (broad-leaved / coniferous) for majority of area is well predicted. Groups of trees in urban area were well classified to their leaf types, and broadleaved- and coniferous-dominated forests were visually located in similar way in forestry areas. Classification errors in this layer were often connected to omission errors in tree cover density layer (low tree cover values were given zero value, and thus no leaf type was present)

	Mixed forests are common, yet based on remote sensing materials and without accurate in situ data it is very difficult to distinguish which of the leaf types (broad-leaved or coniferous) dominates the tree cover. Additional complexity is caused by regularly occurring silvicultural practices such as clearing or thinning, which often steer the forest stand from early broadleaved-dominated development stages towards more coniferous-dominated later stages.

VERY IMPORTANT:* In case of critical findings and to allow traceability, *please, document errors, together with* justifications/explanations/meaningful examples & screenshots, *in section V of this document (see instructions in Ch. 6.3. in Guidelines)

IV. Statistical verification1

For statistical verification of the HRL forest layers, there is an extensive field sample available based on systematic cluster sampling. A set of plots from the 11th National Forest Inventory from year 2013 (NFI11) covering all the land use classes forest land, built-up, arable land, roads and power lines and inland waters (https://www.luke.fi/en/natural-resources/forest/forest-resources-and-forest-planning/forest-resources/). The northernmost Lapland was an exception, the field sample was selected based on double sampling with stratification and originated from the year 2012 (NFI11). There were 9766 field plots selected for quantitative verification. All the field plots on land and inland water were included. In order to follow verification guidelines regarding minimum sapled patch size (section 5.3), it was required that minimum distance to the nearest stand boundary was 20 m on national forestry land and 12.5 m on non-forest land. The radius of the of the NFI 11 field plot is 12.52 m or 12.45 m in South Finland and North Finland, correspondingly. Field plots where a drastic change of land cover or a clearcut of forest had occurred between the field measurement date and image acquisition date (30.6.2015 was assumed for the HRL product) were removed using MS-NFI2015 satellite images and land use change monitoring data from Greenhouse gas reporting project.

The canopy cover (cc) percentage and the cc of broadleaved trees were readily modeled for the field plots on the forest, poorly productive forest land and unproductive land (national land classes) plots (.Mäkisara, K., Katila, M., Peräsaari, J. & Tomppo, E. 2016. The Multi-Source National Forest Inventory of Finland -methods and results 2013. Natural resources and bioeconomy studies 10/2016, Natural Resources Institute Finland. 215 p. http://urn.fi/URN:ISBN:978-952-326-186-0). For more details about estimating the canopy cover for the NFI field plots see the Tree Cover Density verification report section IV. The broadleaved-coniferous dominance was derived from the proportion between cc of broadleaved trees and total cc. In NFI11 trees were also tallied on field plots outside forest (i.e. non-forestry land) and the canopy cover was predicted using statistical models estimated using NFI10 field plots on forest land mineral soils. On the plots outside forestry land, the tree species dominance was defined based on basal area of the tree species tallied.

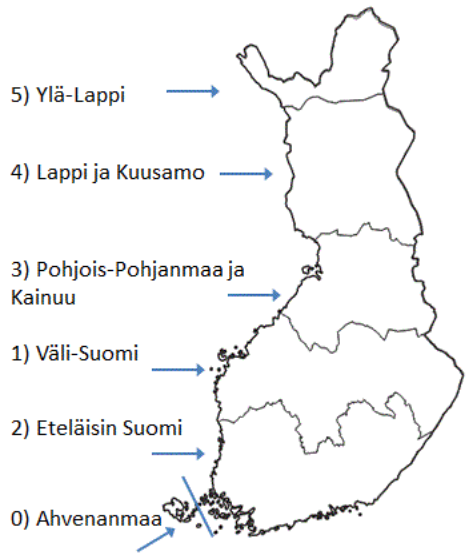

Fig. 3 Sampling regions for the Finnish National Forest Inventory

[1] not relevant for Grassland product, and also not relevant for permanent/temporary wet, and temporary water classes of WAW product

Quantitative error estimates were reported according to the Copernicus Gioland verification guidelines version 1.4 , possibly broken down to geographical regions of different sampling regions in the Finnish NFI (Fig. 3). As the NFI11 data set covered the whole country and was based on systematic sampling, it was found reasonable to calculate the omission and the commission errors solely based on the confusion matrices from NFI field plot points between NFI dominant leaf type vrs. DLT class from the HRL layer.

Stratification	no stratification
Comment on stratification	Field measurements from the national forest inventory (NFI) were used as ground truth data in this verification. NFI is based on systematic cluster sampling over all land use classes and ownership types. Number of field plots per area decreases towards north. The country is divided into six inventory areas (Fig.3.), and results are presented also for these sub-regions. In Finland, over 78 % of the land area is covered by forestry land, and tree cover is found also on other land use classes. Due to sampling methodology and high prevalence of the class to be inspected, no stratification was used.
Number of random samples for finding omission errors	*9766*
Number of valid (applicable) samples for finding omission errors	*Class 0 = No tree cover* *Samples for finding omission error: 3573* samples by inventory regions: Ahvenanmaa 278 Väli-Suomi 1152 Eteläisin Suomi 1518 Pohjois-Pohjanmaa ja Kainuu. 376 Lappi ja Kuusamo 103 Ylä-Lappi 146 *Class 1 = Broadleaved dominance* *Samples for finding omission error: 1299* samples by inventory regions: Ahvenanmaa 51 Väli-Suomi 283 Eteläisin Suomi 337 Pohjois-Pohjanmaa ja Kainuu. 95 Lappi ja Kuusamo 52 Ylä-Lappi 481 *Class 2 = Coniferous dominance* *Samples for finding omission error: 4894* samples by inventory regions: Ahvenanmaa 277 Väli-Suomi 1324 Eteläisin Suomi 1093 Pohjois-Pohjanmaa ja Kainuu. 863 Lappi ja Kuusamo 831 Ylä-Lappi 506
Omission error (%)[2] with uncertainty	*Class 0 = No tree cover*

[2] Producer's accuracy (%) = 1 – omission error (%)

	Omission error: 6.69 % uncertainty 0.82 % Omission error by inventory regions: Ahvenanmaa 15.8 % 4.29 % Väli-Suomi 6.08 % 1.38 % Eteläisin Suomi 6.19 % 1.21 % Pohjois-Pohjanmaa ja Kainuu. 4.26 % 2.04 % Lappi ja Kuusamo 5.83 % 4.52 % Ylä-Lappi 6.16 % 3.90 % *Class 1 = Broadleaved dominance* *Omission error: 49.5 % uncertainty 2.72 %* Omission errors by inventory regions: Ahvenanmaa 45.1 % . 13.66 % Väli-Suomi 37.8 % 5.65 % Eteläisin Suomi 41.8 % 5.27 % Pohjois-Pohjanmaa ja Kainuu. 47.4 % 10.04 % Lappi ja Kuusamo 71.2 % 12.31 % Ylä-Lappi 60.3 % 4.37 % *Class 2 = Coniferous dominance* *Omission error: 30.8 % uncertainty 1.29 %* Omission errors by inventory regions: Ahvenanmaa 49.5 % 5.89 % Väli-Suomi 29.2 % 2.45 % Eteläisin Suomi 28.4 % 2.67 % Pohjois-Pohjanmaa ja Kainuu. 36.6 % 3.21 % Lappi ja Kuusamo 30.1 % 3.12 % Ylä-Lappi 20.9 % 3.55 %
Comment on omissions	Not OK The accuracy of the broadleaved dominant forest type classification is poor. Broadleaved tree species are often in mixed forest stands with coniferous species, making correct classification a difficult task. Accuracy of the coniferous dominance is also weak. Large portion of omissions was detected to recently planted forests. Often omission happened as a result of HRL classification not finding any tree cover, on plots that had small trees or a low tree cover in reality.
Number of random samples for finding commission error	*9766*
Number of valid (applicable) samples for finding commission error	*Class 0 = No tree cover* *Total 4691* samples by inventory regions: Ahvenanmaa 359 Väli-Suomi 1325 Eteläisin Suomi 1648 Pohjois-Pohjanmaa ja Kainuu. 614 Lappi ja Kuusamo 315 Ylä-Lappi 430 *Class 1 = Broadleaved dominance* *Total 1464* samples by inventory regions: Ahvenanmaa 70

	Väli-Suomi 449 Eteläisin Suomi 468 Pohjois-Pohjanmaa ja Kainuu. 155 Lappi ja Kuusamo 63 Ylä-Lappi 259 *Class 2 = Coniferous dominance* *Total 3611* samples by inventory regions: Ahvenanmaa 177 Väli-Suomi 985 Eteläisin Suomi 832 Pohjois-Pohjanmaa ja Kainuu. 565 Lappi ja Kuusamo 608 Ylä-Lappi 444
Commission error (%)[3] with uncertainty	*Class 0 = No tree cover* *Commission error: 28.9 % uncertainty 1.30 %* Commission error by inventory regions: Ahvenanmaa 34.8 % 4.93 % Väli-Suomi 18.3 % 2.08 % Eteläisin Suomi 13.6 % 1.65 % Pohjois-Pohjanmaa ja Kainuu. 41.4 % 3.90 % Lappi ja Kuusamo 69.2 % 5.10 % Ylä-Lappi 68.1 % 4.40 % *Class 1 = Broadleaved dominance* *Commission error: 55.2 % uncertainty 2.55 %* Commission errors by inventory regions: Ahvenanmaa 60.0 % 11.48 % Väli-Suomi 60.8 % 4.52 % Eteläisin Suomi 58.1 % 4.47 % Pohjois-Pohjanmaa ja Kainuu. 67.7 % 7.36 % Lappi ja Kuusamo 76.2 % 10.52 % Ylä-Lappi 26.3 % 5.36 % *Class 2 = Coniferous dominance* *Commission error: 6.18 % uncertainty 0.79 %* Commission errors by inventory regions: Ahvenanmaa 20.9 % 5.99 % Väli-Suomi 4.87 % 1.34 % Eteläisin Suomi 5.89 % 1.60 % Pohjois-Pohjanmaa ja Kainuu. 3.19 % 1.45 % Lappi ja Kuusamo 4.44 % 1.64 % Ylä-Lappi 9.91 % 2.78 %
Comment on commissions	Not OK. Errors of commission were large on broadleaved dominated class, typically as a classification error of a mixed forest stand where the ratio of broadleaved and coniferous trees is difficult to estimate. Coniferous-dominated class experienced this same error of commission for mixed stand, but its magnitude is smaller due to larger preva-

[3] User's accuracy (%) = 1 – commission error (%)

	lence of the coniferous class. Errors of commission for non-tree covered class were often young forests.
Overall evaluation	Looking at the confusion matrix (Table 1a+b) there is overestimation of 'no tree cover' class and underestimation of 'coniferous dominance' class. The layer predicted the leaf type dominance reasonably well in Finland , and types of errors were not surprising but linked with the intensive forest management with a large share of young forests(see effect of forest development stage to classification accuracy in Table 2.), and detailed and high areal variation in the mixture of broadleaved and coniferous species in Finland.

Table 1a Classification error matrix: number of plots examined

TRUE / ESTIMATED (HRL)	No tree cover	Broadleaved	Coniferous	TOTAL
No tree cover	**3 334**	**500**	**857**	4 691
Broadleaved	**159**	**656**	**649**	1 464
Coniferous	**80**	**143**	**3 388**	3 611
TOTAL	3 573	1 299	4 894	**9 766**

Table 1b Classification error matrix: percentage of plots examined

TRUE / ESTIMATED (HRL)	No tree cover	Broadleaved	Coniferous	TOTAL
No tree cover	**34.1 %**	**5.1 %**	**8.8 %**	48.0 %
Broadleaved	**1.6 %**	**6.7 %**	**6.6 %**	15.0 %
Coniferous	**0.8 %**	**1.5 %**	**34.7 %**	37.0 %
TOTAL	36.6 %	13.3 %	50.1 %	**100 %**

Table 2 Correctly classified broadleaved- and coniferous dominated reference sites by forest development stages

Development stage	Broadleaved	Coniferous
2: Young seedling stand	54 %	58 %
3: Advanced seedling stand	57 %	58 %
4: Young thinning stand	84 %	82 %
5: Advanced thinning stand	92 %	91 %
6: Mature stand	95 %	93 %

V. Documentation of errors and critical findings.

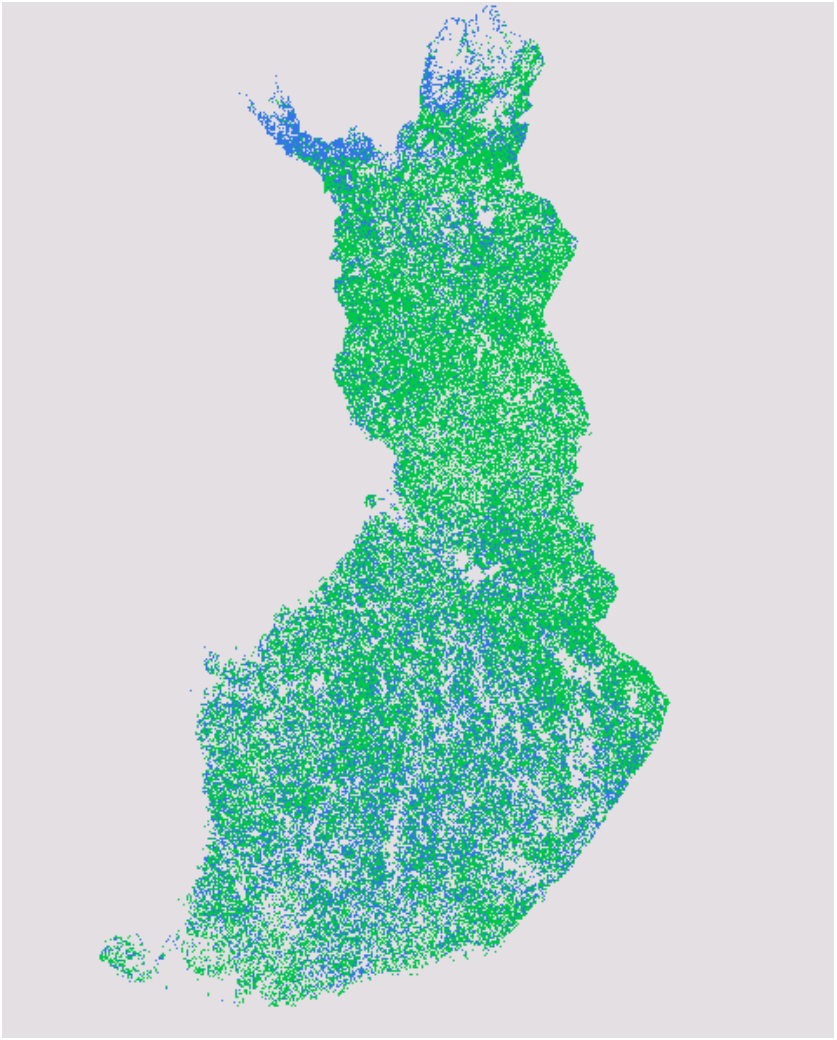

Fig. 4 HRL dominant leaf type; broadleaved forest –blue and coniferous green

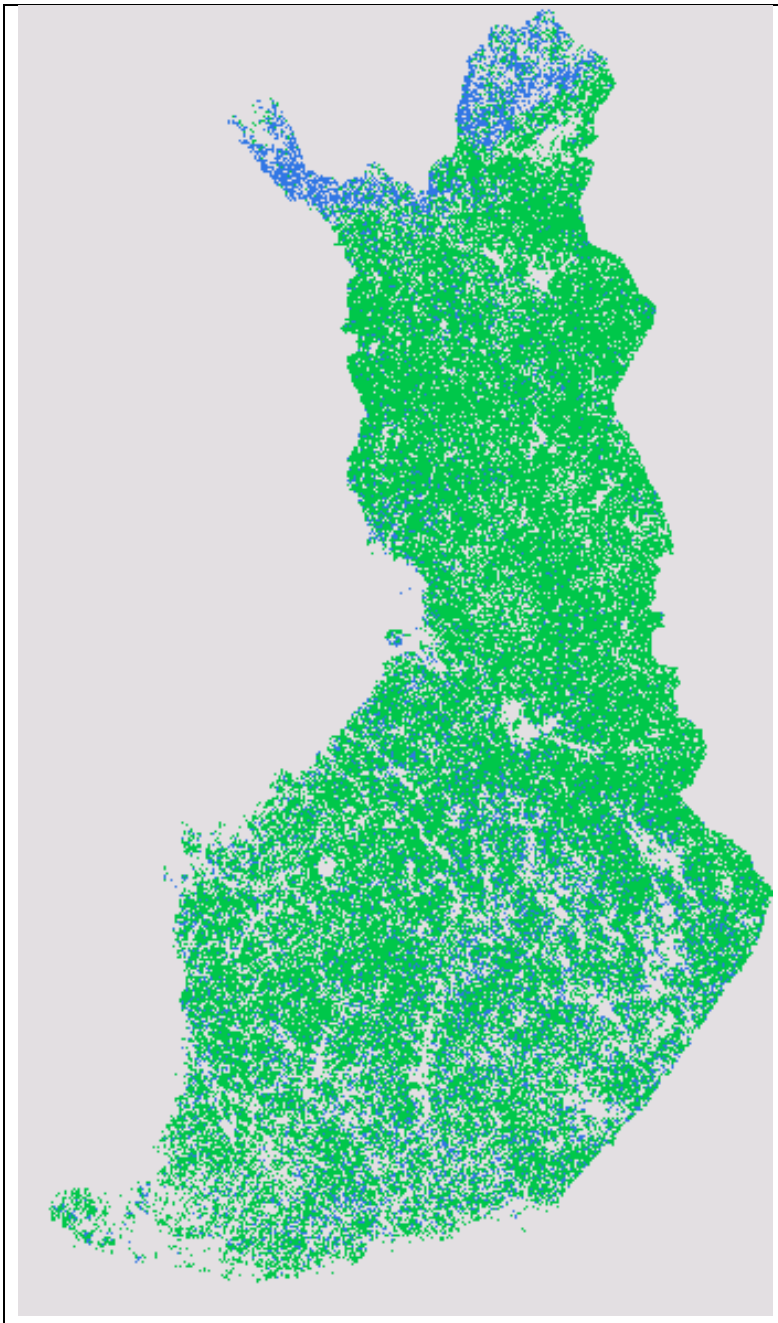

Fig. 5 Multisource-NFI 2015 derived dominant leaf type; broadleaved forest –blue and coniferous green. Minimum canopy cover 2 % for tree species definition.

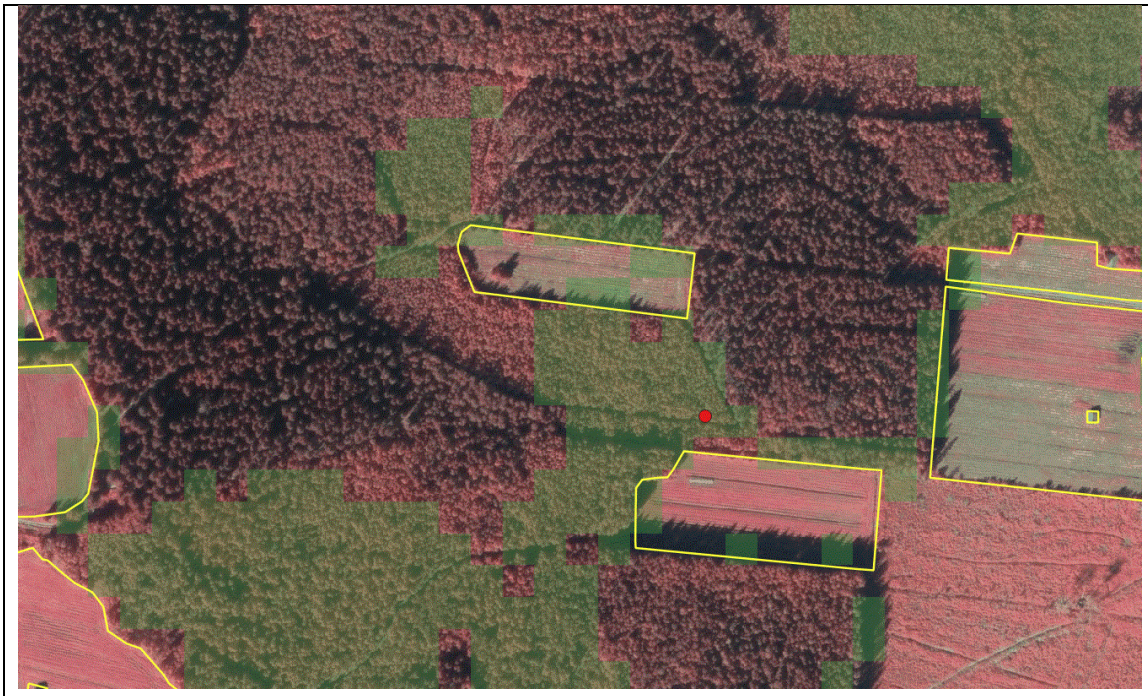

Fig. 6 Commission error on broadleaved dominated class. HRLclassification showing broadleaved dominance (green raster), overlaid over aerial imagery. Ground truth data measured on the plot (red dot) indicates coniferous dominance. Yellow delineation indicates agricultural fields (7040053, 505030) .

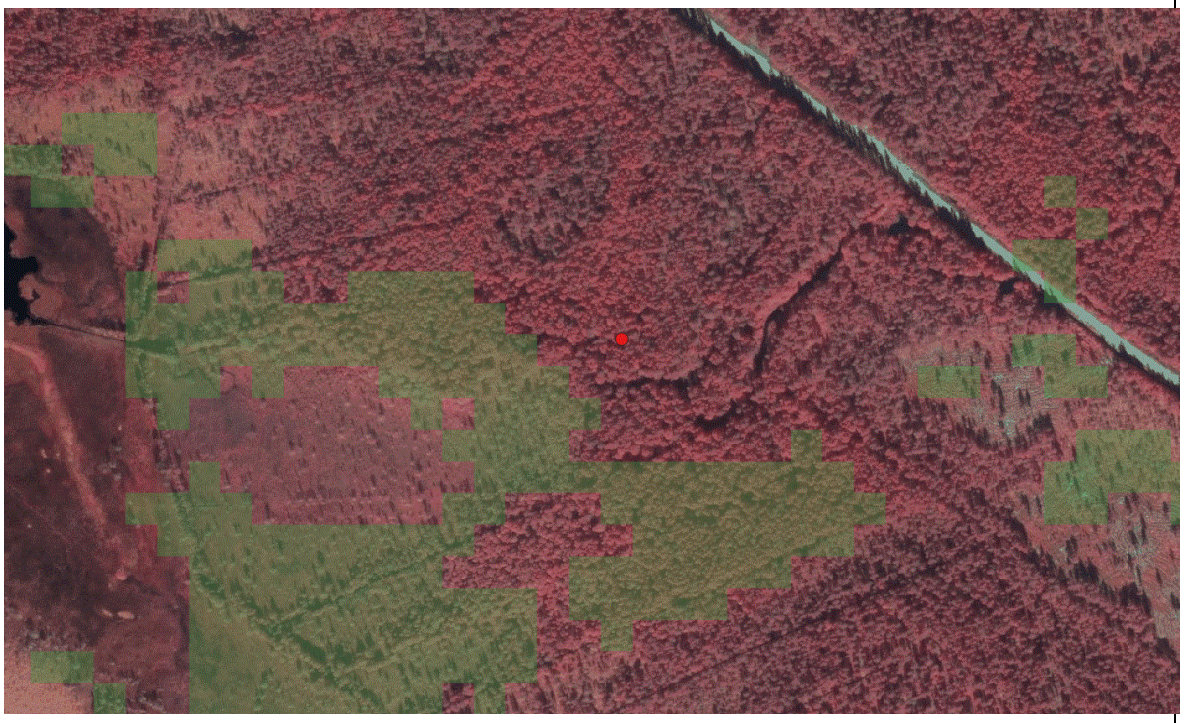

Fig. 7 Omission error on broadleaved dominated class. HRLclassification showing broadleaved dominance (green raster), overlaid over aerial imagery. Ground truth data measured on the plot (red dot) indicates broadleaved dominance, while HRL layer classifies the area dominated by coniferous trees (7212141, 492537)

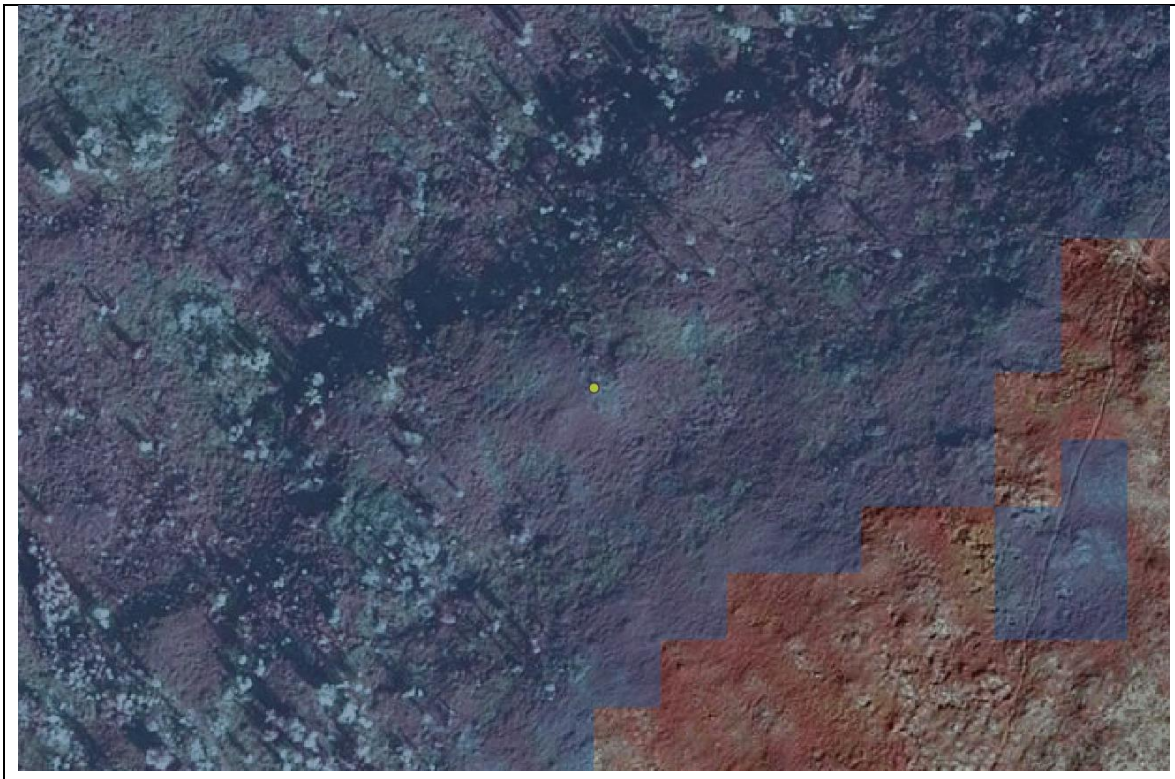

Fig. 8 Commission error (for broadleaved-dominated tree cover class) in bushy, non-tree vegetation in Northern Finland. Large area around scattered individual tree tops was erroneously classified as tree cover (7774815, 533900).

Fig. 9 Omission error for young forest stand. Ground truth: age 12 yrs, broad-leaved dominance, crown cover = 24 % (6817144, 374131).

VI Documentation of software used for verification

Detailed information on the software type and exact version of software used for the validation.
R version 3.4.4 + RStudio Version 1.1.442. (Base R + packages:dplyr,ggplot2,gridExtra,psych) SAS 9.4. QGis 3.0, ArcMap 10.3.1.