

HRL verification report of HRL Grassland in Finland

I. Administrative part

HRL	<i>type the name of the verified layer</i>
Country (and region, if regions are verified separately)	Finland
Institution carrying out the work	Finnish Environment Institute (SYKE)
General overview of data quality done by (name, position and e-mail)	lida Autio, coordinator, iida.autio@ymparisto.fi
Look-and-feel analysis done by (name, position and e-mail)	lida Autio, coordinator, iida.autio@ymparisto.fi
Statistical verification done by (name, position and e-mail)	lida Autio, coordinator, iida.autio@ymparisto.fi Markus Törmä, research engineer, markus.torma@ymparisto.fi
In situ data used. <i>Replace Data-x with the full name of the dataset. Mention quality issues if relevant.</i>	<i>National Ortho photo database/The National Land Survey Natural color/black and white ortho photos Resolution: 0.25-0.5m Reference years: 2009-2016 (partial coverages)</i>
	<i>The Finnish Land Parcel Information System (FLPIS) Based on farming subsidy reports Information of the dominant plant species of the field plots Vector data Reference years: 2009-2015</i>
	<i>National high resolution Corine Land Cover 2006, 2012 and 2018 (HR CLC06-18) National Corine raster dataset Resolution 20x20m (25x25m, year 2006)</i>
	<i>Corine Land Cover change layers 2006-2012 and 2012-2018 Raster data Resolution 0.5ha</i>
	<i>Topographic Database/The National Land Survey Compilations of object groups (meadows, sports fields) Vector data Reference year: 2011 and 2015</i>
	<i>IMAGE 2017 Sentinel-2 satellite image mosaics for spring and summer 2017</i>
	<i>Biotope Maps/Metsähallitus Vector data Reference years (only one year/location): 1996-1998</i>
Internal quality control done by (name, position and e-mail)	Pekka Härmä, project manager, pekka.harma@ymparisto.fi; Minna Kallio, coordinator, minna.kallio@ymparisto.fi; Markus Törmä, research engineer, markus.torma@ymparisto.fi
Date and place of writing the report	20.2.2019 Helsinki

II. General overview of data quality

The total area of the HRL Grassland feature layer is 8402 km² covering about 2,5 % of the Finnish land area. There is no ready to use data available on grasslands in Finland, but an estimation of its extent can be made by combining existing datasets. Appropriate datasets available are the Land Parcel Information System (LPIS) dataset on dominant plant species for years 2009-2015, the national HR Corine Land Cover data on pastures, abandoned arable land and golf courses in years 2006, 2012 and 2018 and meadows in the Topographic database of 2011 and 2015¹. They indicate that there would be ca. 4000 km² of grasslands in the country. This indicates that the Grassland feature layer may be overestimating the grassland area.

Pan-European LUCAS survey of EUROSTAT has information on grasslands and its latest version has collected year 2015. In Finland, there are 13378 LUCAS in-situ points and of these 671 were classified as grassland. Based on these, EUROSTAT has estimated that the area of grasslands in Finland is 14852 km². The large difference in area compared to the HRL class area or national reference data may be explained by the more inclusive class definition of grassland in the LUCAS survey², for example clear cut areas with grassy ground vegetation in forests may be classified as grassland in LUCAS field survey.

In the general overview HRL Grassland and the Finnish national HR CLC12 were combined covering the whole country. The content of HRL Grassland was analysed by calculating the distribution of land cover within HRL Grassland as mapped in HR CLC12 (Table 1).

Table 1. The distribution of the most common land cover classes (using national HR CLC12) in areas which are mapped as grasslands in HRL Grassland feature layer.

Corine Land cover class	Distribution of HRL Grassland according to CLC12.
Non-irrigated arable land (2.1.1)	60,0 %
Arable land no longer in use (2.4.3.1)	3,5 %
Coniferous forest (3.1.2)	4,5 %
Moors and heathland (3.2.2)	3,7 %
Transitional woodland and scrub (3.2.4)	13,4 %

The results show that majority (60,0 %) of the HRL grassland layer is mapped as arable land in HR CLC12. Besides cropland, CLC arable land class (2.1.1) includes also managed grasslands and thus this figure overestimates the share of agricultural crop land on HRL Grassland layer. A rather big share of HRL Grassland (13,4 %) is classified as transitional woodland in CLC (3.2.4.). This CLC class is mainly clear-cuts and young forests incorrectly classified as HRL Grassland.

An overlay analysis was performed combining HRL Grassland feature layer with national data on grasslands (as described above). Only 17,0 % of the HRL Grassland was located in grassland areas according to the national data. This is illustrated in Figure 1.

¹ The national in situ data used for the verification is not optimal, as information on tilling is not directly available. E.g. in the Land Parcel Information System (LPIS) parcels that had been growing perennial grass species during the period of 2009-2015 were considered as managed grasslands. Ploughing on these land parcels is yet allowed.

² LUCAS 2015 - Technical reference document C3 - Classification, Eurostat Technical Documents 2015, E4.LUCAS (ESTAT)

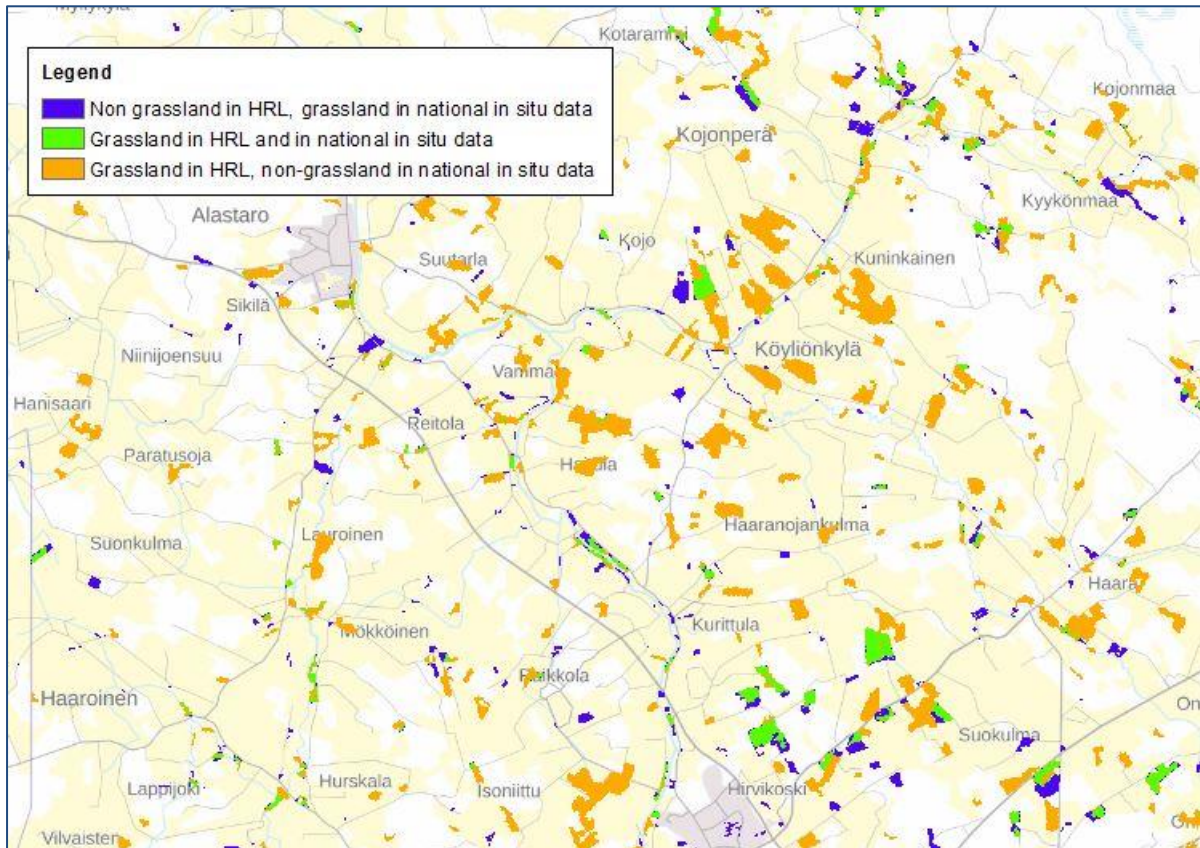


Figure 1. Overlay analysis showing potential commission (orange) and omission (blue) errors. Over-estimation of grassland area in HRL Grassland is clearly visible. Green color indicates areas where HRL grassland is in accordance with national data.

A field trip was organized in order to verify the accuracy of HRL-products in the field. Collected field observations (n=212) didn't allow statistical analysis, but general remarks were made. Most of the visited locations were on HRL Grassland -off areas in coastal meadows and pastures; and omission errors were detected especially at the water front of these habitats. Demonstrative images are included in section V (Figures 7 and 8).

III. Look-and-feel

Stratum	Name of the stratum	Number of samples verified	Results of the verification by strata (<i>excellent, good, acceptable, insufficient, very poor</i>)
1	Cropland	(*)	Very poor - Arable land is often included in the HRL Grassland layer
2	Clear-cut areas	(*)	Insufficient - Clear cut areas are classified as grassland in HRL Grassland (13,4 % of Transitional woodland and scrub)
3	Forest	(*)	Insufficient - Coniferous forest is classified as grassland in HRL Grassland layer (4,5 %).
4	Pastures	(*)	Insufficient - Pastures are often omitted from the HRL Grassland data
5	Abandoned fields	(*)	Insufficient - Abandoned fields are often omitted from HRL Grassland data
Overall evaluation			<i>insufficient (excellent, good, acceptable, insufficient, very poor)</i>
Comments			

(*) In the statistical verification totally 560 locations were interpreted and checked, which gave also detailed and statistically unbiased look-and-feel impression of HRL data including critical strata.

IV. Statistical verification³

Stratification	<i>Sample plots for determining omission errors were concentrated in areas of potential errors. These areas included CLC classes 142, 231, 243 and 321 from Finnish HR CLC2012 (20 m raster). These HRL-off areas were buffered by one pixel to increase the total area for sample selection. Border pixels of HRL-on areas were removed. These operations were performed in order to reduce the influence of possible positional errors and shifts in different data sets. A systematic network (200 meter interval) of potential sample plots was determined, from which random samples of 280 HRL-on and 280 HRL-off points were selected. Results are illustrated in Figure 2.</i>
Comment on stratification	
Number of random samples for finding omission errors	280
Number of valid (applicable) samples for finding omission errors	280
Omission error (%) ⁴ with uncertainty (calculated for the stratified HRL-off area)	30,7 %; uncertainty 300,8 % ⁶ (99,1 % ⁷)
Comment on omissions	
Number of random samples for finding commission error	280
Number of valid (applicable) samples for finding commission error	280
Commission error (%) ⁵ with uncertainty	76,8 %; uncertainty 2,5 % ⁸
Comment on commissions	
Overall evaluation	<i>General overview, look-and-feel as well as statistical verification indicate that the HRL Grassland layer does not succeed in identifying the grassland areas in Finland and unacceptable amount of commission and omission errors occur throughout the data. Classification of HRL Grassland needs radical improvements especially in separating arable land as well as clear cut areas from grassland.</i>

³ Performed, even though not obligatory for HRL Grassland product

⁴ Producer's accuracy (%) = 1 – omission error (%)

⁵ User's accuracy (%) = 1 – commission error (%)

⁶ Calculated as instructed in the Annex1 of the verification guide. The term "Area_{HRLclass}" in the formula is corrected for omission and commission errors (Area_{RealHRLclass}).

⁷ Calculated as instructed in the Annex1 of the verification guide. The term "Area_{HRLclass}" in the formula is NOT corrected for omission and commission errors.

⁸ Calculated to correspond to a significance level of appr. 68,3 % as instructed in Annex1 of the verification guidelines.

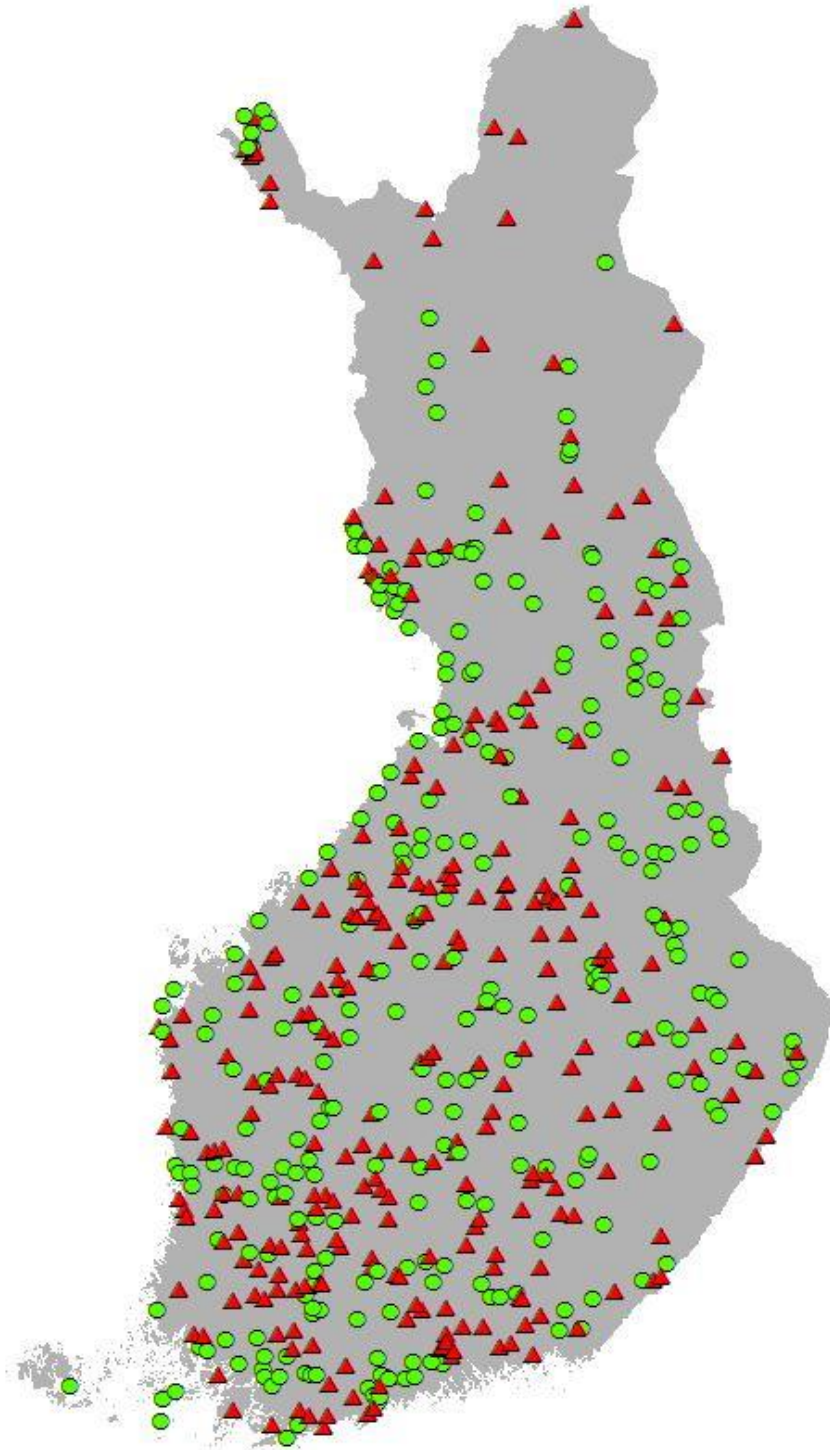


Figure 2. Areal distribution of sample plots in statistical verification. Green sample plots are correctly and red plots incorrectly interpreted as grassland in the HR data.

V. Documentation of errors and critical findings

Screenshots of typical mistakes in HRL Imperviousness data are displayed on top of true color ortophotos in scale 1:3000 - 1:4000. HRL Grassland is displayed as transparent brown. CLC data is displayed in light transparent colors specified in the captions.

Commission errors

The strata with the major commission errors are arable land and transitional forests. This was evident in all steps of the verification. Out of the total of 215 erroneously classified sample points on HRL-on areas, 146 are located on agricultural croplands. The production method doesn't seem to recognize arable land from grassland and this is a major problem in the dataset. No clear pattern can be found on what kind of croplands are misclassified. Many of them have been growing perennial grasses at some point during the reference period (2015 +/- 1 year and 6 preceding years) but not the whole time. Transitional woodlands that are misclassified as HRL Grassland are mostly clear cut areas. Commission errors are demonstrated in figures 3-5.



Figure 3. Arable land interpreted as HRL Grassland (light brown). Scale 1:8000, coordinates E:401150, N: 6751350.

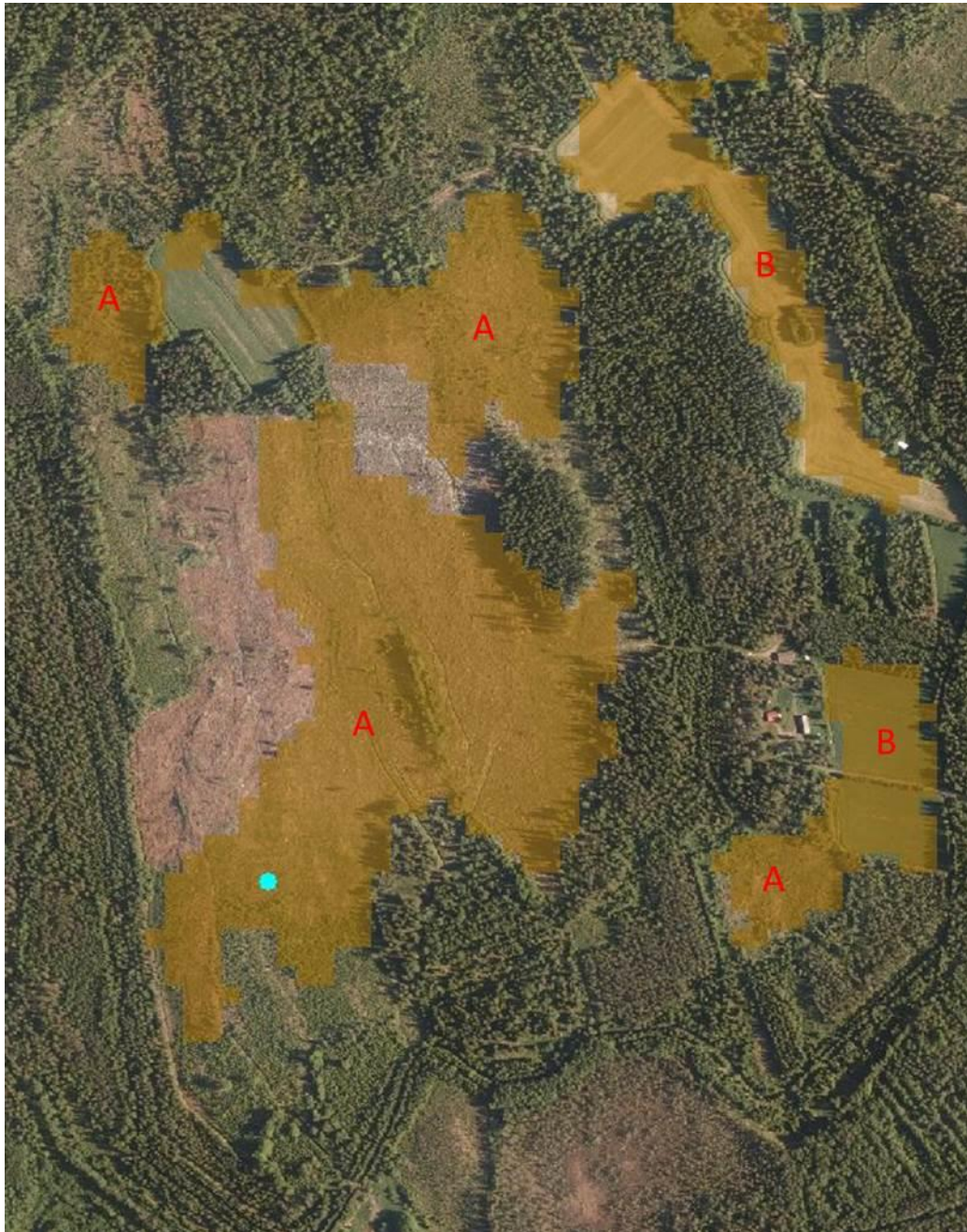


Figure 4. HRL Grassland (light brown) in forest clear-cut (A) and arable land (B). Scale 1:6000, coordinates E:435210, N: 7214100.

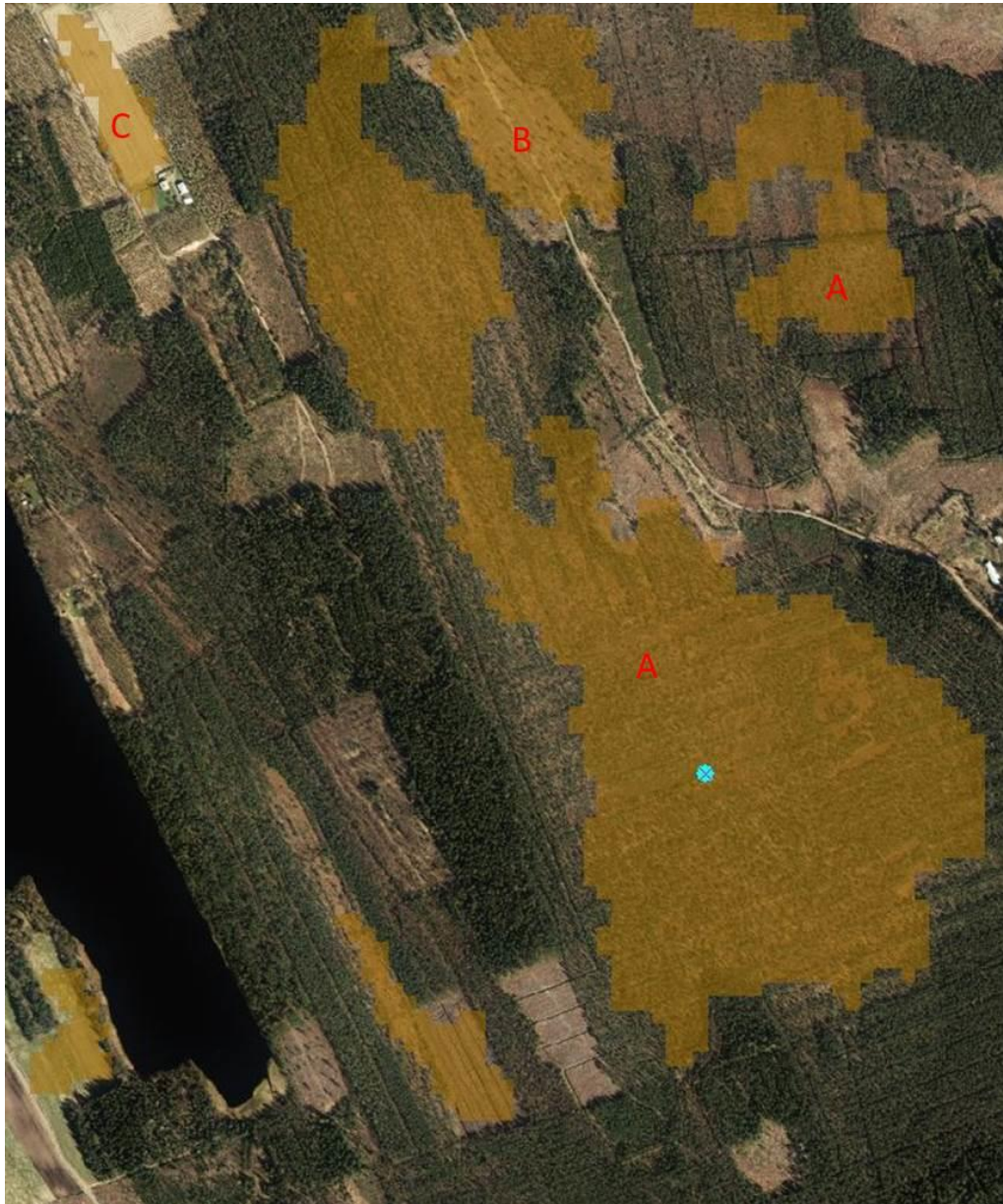


Figure 5. HRL Grassland (light brown) in forest (A), clear cut area (B) and arable land (C). Scale 1:8000, coordinates E:516177, N: 6867583.

Omission errors

Over half of the omission errors in the statistical verification were found on abandoned arable land. These areas have not been cultivated for several years and are currently grown over with grass and should thus be classified as grassland. Omission errors are also found in pastures. Omission errors are demonstrated in figures 6-7.

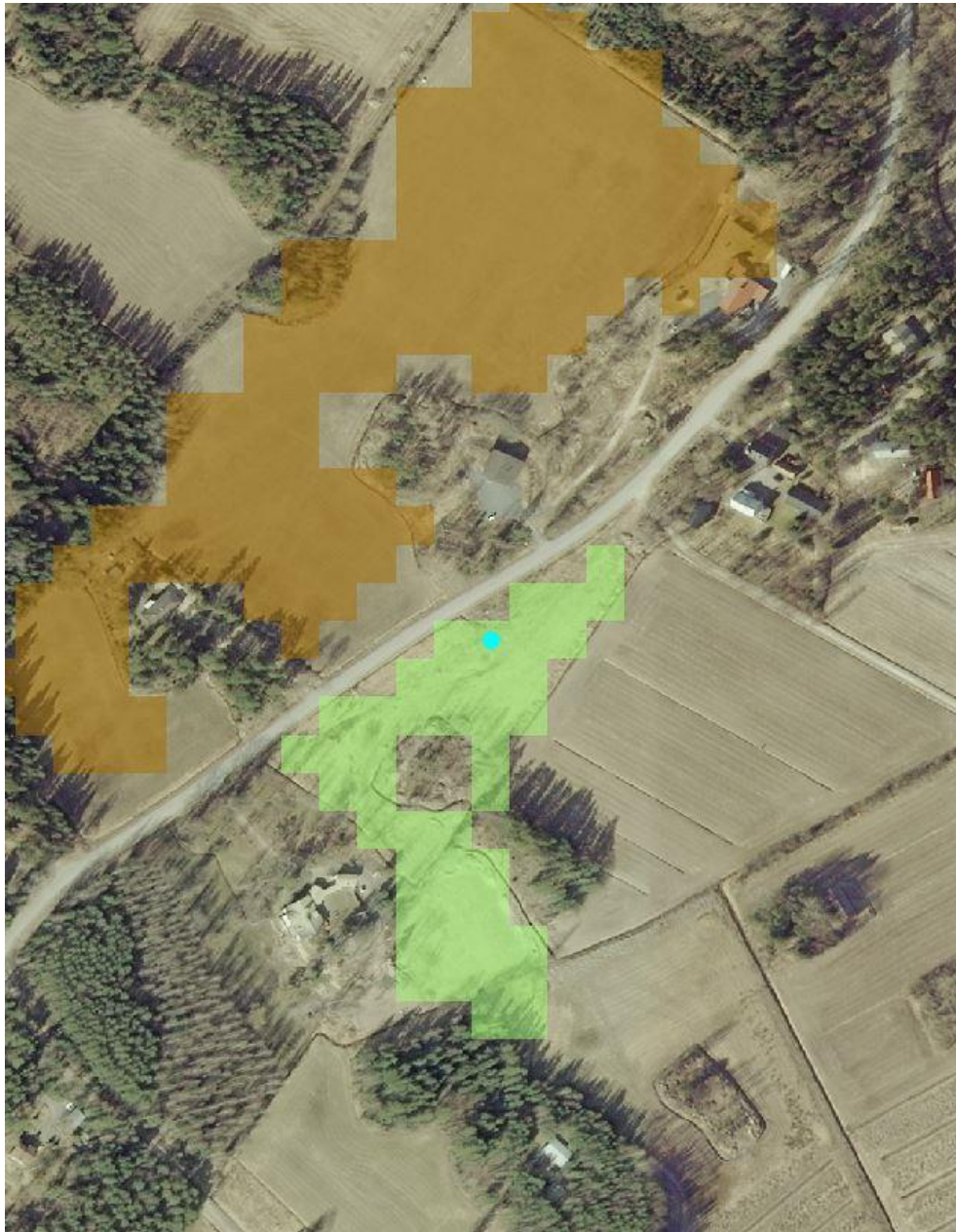


Figure 6. HR CLC12 Abandoned arable land (light green) not classified as HRL Grassland (omission). HRL Grassland (light brown) in arable land (commission). Scale 1:3000, coordinates E:217350, N: 6825750.

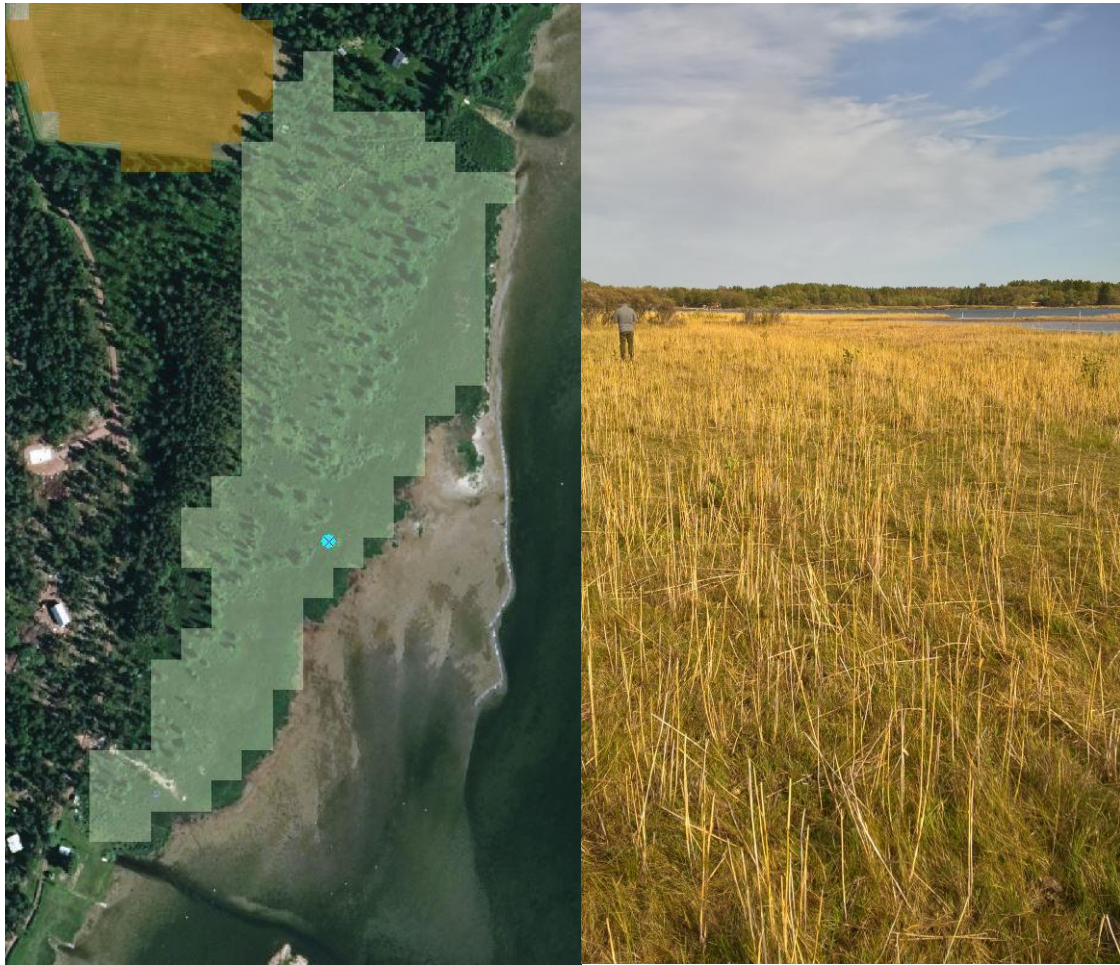


Figure 7. The coastal edge of a HR CLC12 pasture (light color) not mapped as grassland on HRL Grassland layer (light brown). Scale 1: 3000, E: 99320, N: 6684897 (picture on the left). The site photographed on a field visit in autumn 2018 (picture on the right).

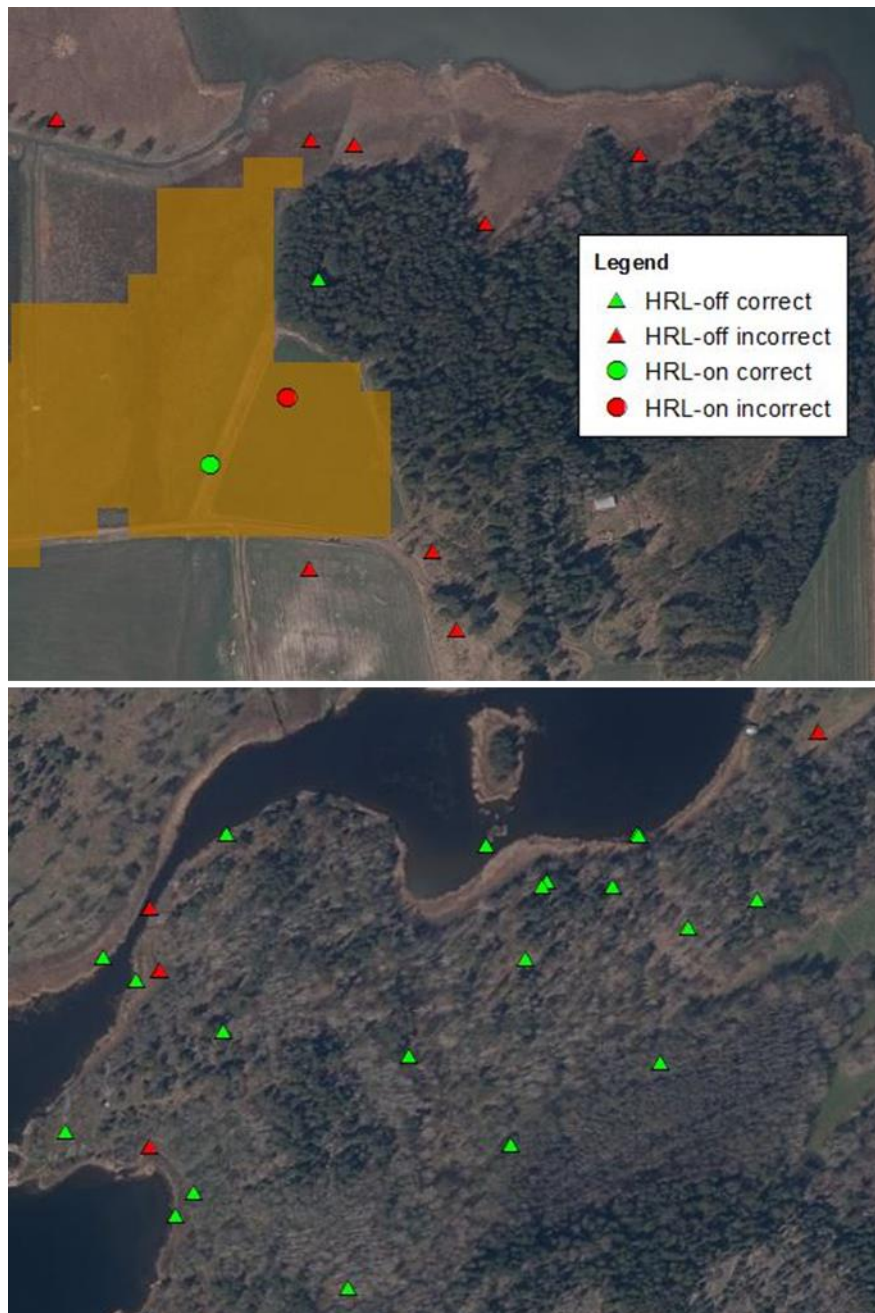


Figure 8. Examples of HRL grassland sample points visited and validated during a field trip in autumn 2018. HRL Grassland is shown in light brown. Misclassified sample points are concentrated in the beachfront grasslands. Scale 1:3000, coordinates E: 115431, N: 6705637 (image above), E: 108712, N: 6712576 (image below).

VI. Documentation of software used for verification

The software type and exact version of software used for the validation:

General overview & Look-and-feel:

- ArcGIS 10.5.1 desktop
- Excel 2010

Statistical verification:

- ArcGIS 10.5.1 desktop
- ERDAS IMAGINE 2016
- Matlab R2016b
- Excel 2010