

## HRL 2018 look & feel verification report for Grassland 2018 (GRA) in Finland

### I. Administrative part

HRL	<i>Grassland 10m</i>
Verified area, region	<i>Finland, overview map</i>
Institution carrying out the work	Finnish Environment Institute (SYKE)
Overall visual checking done by (name, position and e-mail)	Eetu Jutila, coordinator, eetu.jutila@syke.fi
Look & feel verification done by (name, position and e-mail)	Eetu Jutila, coordinator, eetu.jutila@syke.fi
Statistical verification done by (name, position and e-mail)	Eetu Jutila, coordinator, eetu.jutila@syke.fi Markus Törmä, research engineer, markus.torma@syke.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 colour/ infrared ortho photos Resolution: 0.25-0.5m Reference years: 2017-2019 (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: 2012-2015 &amp; 2018 (2016 not available, 2017 missing polygons)</i>
	<i>National high resolution Corine Land Cover 2012 and 2018 (HR CLC12-18) National Corine raster dataset Resolution 20x20m</i>
	<i>Topographic Database/The National Land Survey Compilations of object groups (meadows, parks, sports fields, wetlands and swamps) Vector data Reference years: 2012,2014, 2016, 2017 &amp; 2019</i>
	<i>Biotope data/Metsähallitus Biotope classes used: moors, heathlands, meadows Reference years 2005-2015</i>
	<i>IMAGE 2012 Sentinel-2 satellite image mosaics for spring and summer Reference year: 2012</i>
	<i>Sentinel-2 Global mosaics (S2GM)/Copernicus Programme Sentinel-2 natural and false colour satellite image mosaics for summer Reference year: 2018</i>
	<i>Esri World Imagery / Esri, DigitalGlobe, GeoEye, i-cubed, USDA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community High-resolution satellite and aerial images. Used in areas where NLS's ortho photos are not available.</i>

	<i>Reference years: 2016-2021</i>
Reporting done by (name, position and e-mail)	Eetu Jutila, coordinator, eetu.jutila@syke.fi
Date and place of writing the report	22.6.2021

## II. General overview of the verified data

The total area of the HRL Grassland feature layer (HRL GRA18) is 17131 km<sup>2</sup> covering about 5,6 % of the Finnish land area. The overall statistics of this layer can be seen in Table 1 and Figure 1 shows the corresponding grassland areas. There is no ready to use in-situ data available about grasslands. However, its extent was estimated by combining existing datasets including dominant plant species data from the Land Parcel Information System (LPIS) for years 2012-2015 & 2018, the National High Resolution Corine Land Cover 2018 & 2012 (HR CLC18, HR CLC12) data on pastures, abandoned arable land, natural grassland, golf courses, and meadows & parks from the Topographic Database. <sup>1</sup>

Table 2 shows the statistical comparison between the HRL GRA18 and the national reference layer. It also reveals that the grassland area (4875 km<sup>2</sup>) of the national reference layer is and is ca. 3,5 times smaller compared to the grassland area of the HRL GRA18 that is 17131 km<sup>2</sup>. One reason for this difference is the limitation of the national reference layer. However, it strongly indicates that the HRL Grassland feature layer is overestimating the grassland area. HRL GRA18 was also compared against HR CLC18. Table 3 displays the results of this comparison. The results align well with the experiences of the statistical and Look & Feel verification. Especially high commission errors seem to be found in Non-irrigated arable land, Peatbogs, Transitional woodland/shrub and Moors and heathland. The main commission difficulties of HRL GRA18 seem to be to distinguish between grassland and cropland and not to classify open bogs/swamps or fresh clear-cut areas as grassland. Major omission errors are in Pastures, Natural Pastures, Natural Grassland, Arable land outside farming subsidies and Agro-forestry areas. These errors are mainly related to the possible overestimation of tilled grasslands in the Ploughing Indicator (PLOUGH2018). Therefore, the area of the coinciding grassland between HRL-GRA18 and national reference layer is quite limited.

Table1. Overall statistics

HRL GRA18 Finland	Value	Km2	%
Non-grassland areas	0	329898,58	42,52 %
Grassland	1	17131,24	2,21 %
Unclassifiable	245	0,00	0,00 %
Outside area (no data)	255	428904,35	55,28 %
<b>SUM</b>		<b>775934,17</b>	<b>100,00 %</b>
<b>Total grassland surface</b>		<b>17131,24</b>	<b>2,21 %</b>

<sup>1</sup> 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 2012-2018 were considered as managed grasslands. Ploughing on these land parcels is yet allowed. Moreover, some elements, such as heathlands with high grass cover, sparsely vegetated grasslands and grasslands with scattered trees and shrubs covering a maximum 10 %, were excluded from the national in-situ data due inconsistent grass cover.



Figure 1. Overview map

Table 2. Comparison of relevant statistical values for HRL GRA18 with the national reference layer.

General overview of the verified data	km2	% of the area of Finland
HRL Grassland in Finland	17131,24	4,38 %
Grassland in national reference layer	4875,18	1,25 %
Coinciding grassland between HRL-GRA18 and national reference layer	2659,30	0,68 %
Grassland in HRL-GRA18 (not in national reference layer)	14471,94	3,70 %
Grassland in national reference layer (not in HRL-GRA18)	2215,89	0,57 %
Coinciding non-grassland in HRL-GRA18 and national reference layer	327688,17	83,83 %

 Table 3. HR CLC18 compared to impervious area in HRL GRA18. **Blue: Potential commission errors; Red: potential omission errors**

HR CLC18 code (Level 4)	HR CLC18 class name	Area in HR CLC18 (km2)	Share from country area (%)	Grassland area in HRL GRA18 (km2)	GRA content in HR CLC18 class (%)	Remaining area (not matching with HRL GRA18) in HR CLC18 (%)
1.1.1.1	Continuous urban fabric	170,9516	0,0 %	4,1	0,0 %	100,0 %
1.1.2.1	Discontinuous urban fabric	3176,1196	0,8 %	524,0	3,1 %	96,9 %
1.2.1.1	Commercial units	956,6492	0,2 %	114,0	0,7 %	99,3 %
1.2.1.2	Industrial units	623,8616	0,2 %	45,1	0,3 %	99,7 %
1.2.2.1	Road and rail networks and associated land	2378,1192	0,6 %	224,9	1,3 %	98,7 %
1.2.3.1	Port areas	39,8916	0,0 %	0,9	0,0 %	100,0 %
1.2.4.1	Airports	77,0064	0,0 %	12,6	0,1 %	99,9 %
1.3.1.1	Mineral extraction sites	421,9192	0,1 %	21,5	0,1 %	99,9 %
1.3.1.2	Open cast mines	28,9452	0,0 %	0,6	0,0 %	100,0 %
1.3.2.1	Dump sites	134,0024	0,0 %	16,0	0,1 %	99,9 %
1.3.3.1	Construction sites	27,4648	0,0 %	2,3	0,0 %	100,0 %
1.4.1.1	Green urban areas	33,8056	0,0 %	13,9	0,1 %	99,9 %
1.4.2.1	Summer cottages	1367,2284	0,3 %	83,9	0,5 %	99,5 %
1.4.2.2	Sport and leisure areas	134,89	0,0 %	25,7	0,2 %	99,8 %
1.4.2.3	Golf courses	87,0548	0,0 %	50,5	0,3 %	99,7 %
1.4.2.4	Racecourses	9,912	0,0 %	2,1	0,0 %	100,0 %
2.1.1.1	Non-irrigated arable land	21774,7388	5,6 %	5461,4	<b>31,9 %</b>	68,1 %
2.2.2.1	Fruit trees and berry plantations	62,6328	0,0 %	27,8	0,2 %	99,8 %
2.3.1.1	Pastures	39,6808	0,0 %	21,0	0,1 %	<b>99,9 %</b>
2.3.1.2	Natural pastures	93,9324	0,0 %	34,8	0,2 %	<b>99,8 %</b>
2.4.3.1	Arable land outside farming subsidies	2125,2284	0,5 %	859,8	5,0 %	<b>95,0 %</b>

<b>2.4.4.1</b>	Agro-forestry areas	35,3256	0,0 %	10,9	0,1 %	<b>99,9 %</b>
<b>3.1.1.1</b>	Broad-leaved forest on mineral soil	9795,4536	2,5 %	344,0	2,0 %	98,0 %
<b>3.1.1.2</b>	Broad-leaved forest on peatland	562,666	0,1 %	11,0	0,1 %	99,9 %
<b>3.1.2.1</b>	Coniferous forest on mineral soil	114100,1264	29,2 %	528,7	3,1 %	96,9 %
<b>3.1.2.2</b>	Coniferous forest on peatland	32643,194	8,4 %	116,1	0,7 %	99,3 %
<b>3.1.2.3</b>	Coniferous forest on rocky soil	3060,2108	0,8 %	4,0	0,0 %	100,0 %
<b>3.1.3.1</b>	Mixed forest on mineral soil	36411,738	9,3 %	677,9	4,0 %	96,0 %
<b>3.1.3.2</b>	Mixed forest on peatland	8723,216	2,2 %	73,0	0,4 %	99,6 %
<b>3.1.3.3</b>	Mixed forest on rocky soil	227,6856	0,1 %	2,2	0,0 %	100,0 %
<b>3.2.1.1</b>	Natural grassland	107,944	0,0 %	6,2	0,0 %	<b>100,0 %</b>
<b>3.2.2.1</b>	Moors and heathland	7381,9808	1,9 %	1388,2	<b>8,1 %</b>	91,9 %
<b>3.2.4.1</b>	Transitional woodland/shrub cc <10%	8477,8572	2,2 %	1523,4	<b>8,9 %</b>	91,1 %
<b>3.2.4.2</b>	Transitional woodland/shrub, cc 10-30%, on mineral soil	14768,9384	3,8 %	1914,9	<b>11,2 %</b>	88,8 %
<b>3.2.4.3</b>	Transitional woodland/shrub, cc 10-30%, on peatland	8623,5204	2,2 %	575,9	3,4 %	96,6 %
<b>3.2.4.4</b>	Transitional woodland/shrub, cc 10-30%, on rocky soil	1542,948	0,4 %	15,2	0,1 %	99,9 %
<b>3.2.4.6</b>	Transitional woodland/shrub under power lines	384,2592	0,1 %	83,4	0,5 %	99,5 %
<b>3.3.1.1</b>	Beaches, dunes, and sand plains	64,058	0,0 %	6,2	0,0 %	100,0 %
<b>3.3.2.1</b>	Bare rock	1779,8612	0,5 %	58,6	0,3 %	99,7 %
<b>3.3.3.1</b>	Sparsely vegetated areas	541,6728	0,1 %	33,4	0,2 %	99,8 %
<b>4.1.1.1</b>	Inland marshes, terrestrial	373,6288	0,1 %	42,2	0,2 %	99,8 %
<b>4.1.1.2</b>	Inland marshes, aquatic	1066,2452	0,3 %	35,5	0,2 %	99,8 %
<b>4.1.2.1</b>	Peatbogs	19087,9204	4,9 %	1888,9	<b>11,0 %</b>	89,0 %
<b>4.1.2.2</b>	Peat production sites	1026,1972	0,3 %	79,5	0,5 %	99,5 %
<b>4.2.1.1</b>	Salt marshes, terrestrial	300,594	0,1 %	63,7	0,4 %	99,6 %
<b>4.2.1.2</b>	Salt marshes, aquatic	292,2168	0,1 %	20,6	0,1 %	99,9 %
<b>5.1.1.1</b>	Water courses	1168,5224	0,3 %	18,8	0,1 %	99,9 %
<b>5.1.2.1</b>	Water bodies	32299,1008	8,3 %	27,0	0,2 %	99,8 %
<b>5.2.3.1</b>	Sea and ocean	52197,4096	13,4 %	4,6	0,0 %	100,0 %

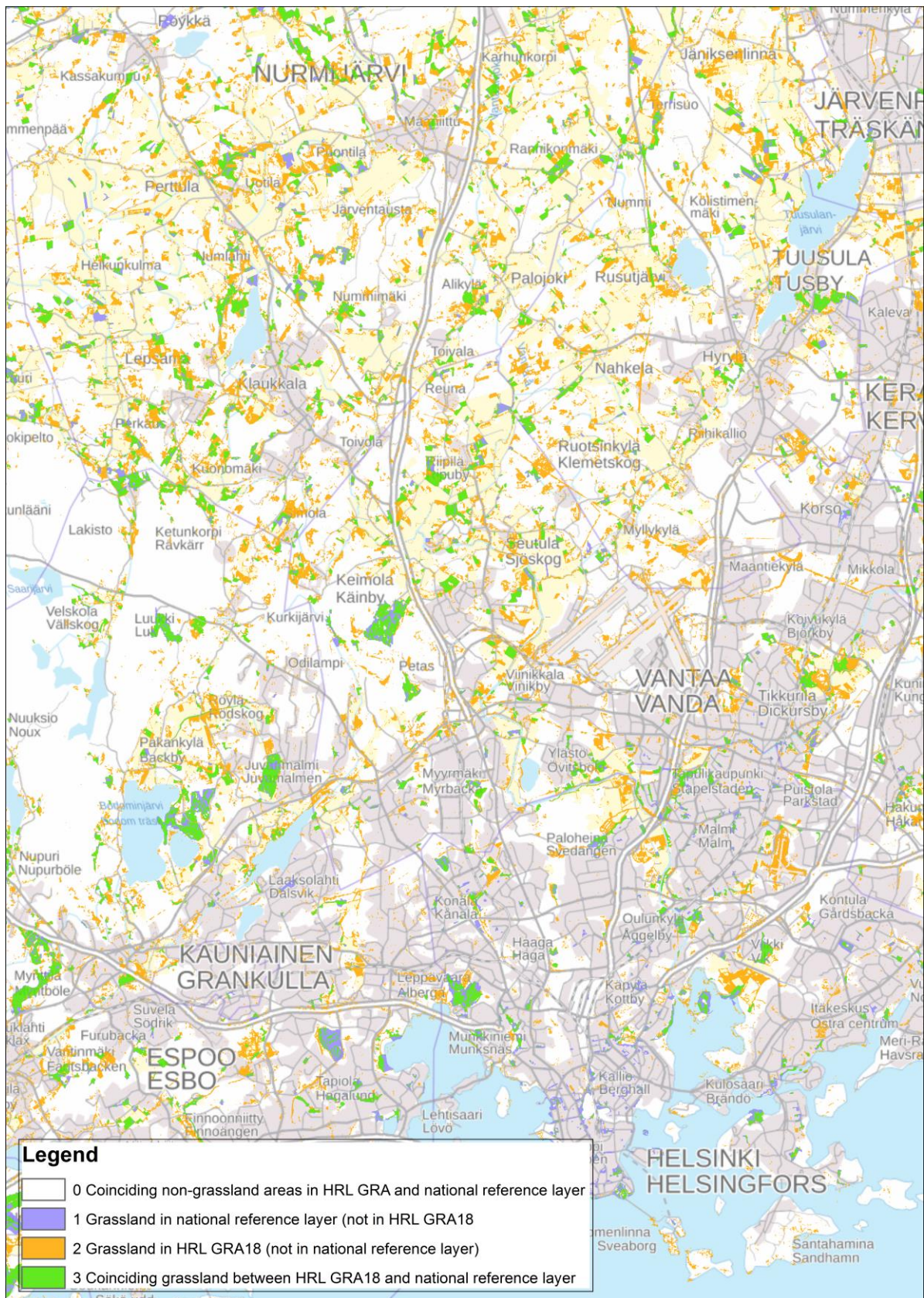


Figure 2. Comparison of HRL GRA18 and HR CLC18 in the Helsinki Metropolitan area. Purple areas indicate possible omission errors and orange areas commission errors.

### III. Overall visual checking

Positional accuracy			
Relative positional accuracy	Quick visual comparison of HRL data with available EO imagery (identifying large positional errors)	OK	<i>Large positional errors were not detected in the data</i>
Thematic accuracy			
Classification correctness	Simple look & feel thematic check (identifying basic thematic mistakes)	NOT OK	The quick visual comparison of the HRL GRA18 data with national orthophoto images indicate that the HRL GRA18 generally overestimates grassland in croplands, peatbogs and clear-cut areas. Moreover, potential omission errors due Ploughing Indicator exist.



#### IV. Look & feel verification results

##### Details of look & feel verification

##### 1. Included elements, possible OMISSIONS

Stratum	Name of the stratum (see proposed strata in Tables 5.2.2.x.b)	Number of samples verified	Results of the verification by strata (excellent, good, acceptable, insufficient, very poor): see chapter 5.2 of the guidelines
1	Natural, semi-natural, agricultural / managed grass-covered surfaces	10 (polygons)	Very poor (1)
2	Grasslands with scattered trees and shrubs covering a maximum 10 %	10 (points)	Very poor (1)
3	Heathland with high grass cover, maximum of 10 % non-grass cover	10 (points)	Acceptable (3)
4	Coastal grasslands, such as grey dunes and salt meadows located in intertidal flat areas with at least 30 % graminoid species of vegetation cover.	20 (polygons)	Insufficient (2)
5	Sparsely vegetated grasslands (> 30% vegetation cover)	12 (polygons)	Acceptable (3)
6	Grasslands in urban areas: parks, urban green spaces in residential and industrial areas	22 (polygons)	Acceptable (3)

7	Grass cover at airports	10 (points)	<i>Insufficient (2)</i>
8	Grass cover at sport and recreation areas	10 (polygons)	<i>Insufficient (2)</i>
9	Wet grasslands along rivers & lakes	16 (polygons)	<i>Insufficient (2)</i>
10	Agro-forestry areas including grass cover	13 (polygons)	<i>Acceptable (3)</i>
Overall evaluation			<i>Insufficient (2)</i>
Comments, overview of results			<p>The polygons and points verified are attached as separate shapefiles to this report (GRA2018_LookFeelSamples_polygons_FI.shp and GRA2018_LookFeelSamples_points_FI.shp).</p> <p>Two recommended strata were not evaluated:</p> <ul style="list-style-type: none"> <li>- Semi-arid steppes with scattered Artemisia scrub could not be located (not with in-situ data or visual scanning)</li> <li>- Natural grasslands on military sites were not evaluated due national security issues</li> </ul> <p>The classification accuracy varied significantly both between and within the checked strata. Sparsely vegetated grasslands, grasslands in urban areas and grassy Agro-forestry areas were detected with an acceptable accuracy. Yet, major omission errors were detected in stratum 1, 4 and 9 that are mainly due the PLOUGH18 that might be overestimating tilled grasslands. Grassland with scattered trees and shrubs consisted mainly of clear-cut areas older than 6 years that were mainly omitted.</p>

## 2. Excluded elements, possible COMMISSIONS

Stratum	Name of the stratum (see proposed strata in Tables 5.2.2.x.c)	Number of samples verified	Results of the verification by strata (excellent, good, acceptable, insufficient, very poor): see chapter 6.3 of the guidelines
1	Peat forming ecosystems dominated by sedges	50	Very poor (1)
2	Reed beds and helophytes dominated systems	20	Very poor (1)

3	Tall forbs, fern, shrub dominated vegetation	10	Insufficient (2)
4	Grasslands that have been observed as tilled (in the reference year or a certain period before, in that case they are considered as arable fields)	10	Very poor (1)
5	Vineyards, orchards, olive groves, (if more than 10 % shrubs or trees)	10	Very poor (1)
6	Tundras dominated by shrubs and lichens	10	Very poor (1)
7	Clear-cut areas in woods	10	Very poor (1)
8	Cropland	10	Very poor (1)
Overall evaluation			Very poor (1)
Comments, overview of results			<p>The polygons verified are attached as a separate shapefile to this report (GRA2018_LookFeelSamples_polygons_FI.shp).</p> <p>Rice fields were not evaluated since they do exist in Finland.</p> <p>Clear-cut areas were visually selected from areas where the forest had been chopped after between years 2012 and 2018.</p> <p>Overall, the accuracy between the checked strata was very poor. This confirms that the HRL GRA18 radically overestimates the grassland areas in Finland. Major problems exist with cropland, peat forming ecosystems and clear-cut areas in woods that are often classified as grassland.</p>

	<p>As a reference, the Finnish Food Authority produces a dataset about permanent grassland in Finland. The data is based on The Finnish Land Parcel Information System (FLPIS) and has an area of 1769 km<sup>2</sup> that is substantially smaller than the national reference layer used in this verification. The major differences are that it consists only of FLPIS data and has more strict classification criteria.</p>
--	---

## V. Documentation of errors and critical findings

### Omission errors

Examples and screenshots of omission errors of the strata checked in look & feel verification are presented in this chapter. The SAMPLE\_ID in the figure captions is referring to the associated GIS-data-files: GRA2018\_LookFeelSamples\_polygons\_FI and GRA2018\_LookFeelSamples\_points\_FI. The background image is an aerial ortho photo from reference years 2017-2019 (The National Land Survey)

**Natural, semi-natural, agricultural / managed grass-covered surfaces:** This stratum is poorly detected in the HRL GRA18 data and major omission errors occur. These are mainly related to the Ploughing Indicator that may be overestimating tilled grasslands. The ploughing information was not directly available for the national in-situ data. Thus, areas that remained in the in-situ data every year during the reference period (2012-2018) were considered as grassland. Yet, it does not guarantee that the grassland areas are not ploughed. During the look & feel and quantitative verification time series of ortho photos were used to assess whether the field was tilled or not. This was sometimes challenging as signs of tilling and mowing look quite similar from the ortho photos that were mainly captured in the summer.



Figure 3. Omission errors in an environmental subsidy area of permanent grassland. The turquoise polygon is the contradicting area derived from the GIS comparison of in-situ data and the HRL GRA18. GRA18 data is presented with light green and PLOUGH18 with a shades of red. SAMPLE\_ID 95, scale 1:10 000, coordinates (ETRS\_1989\_LAEA) E: 5046591 N: 4716681. Ortho photo reference year 2017.

### Commission errors

Examples and screenshots of commission errors of the strata checked in look & feel verification are presented in this chapter. The SAMPLE\_ID in the figure captions is referring to the associated GIS-data-files: GRA2018\_LookFeelSamples\_polygons\_FI and GRA2018\_LookFeelSamples\_points\_FI. The background image is an aerial ortho photo from reference years 2017-2019 (The National Land Survey).

**Peat forming ecosystems dominated by sedges:** This stratum is one of the most problematic strata of the look & feel verification as peat forming ecosystems are frequently classified as grassland. There exist various swamp types in Finland and all of them are not dominated by sedges. However, their biotope substantially differs from those associated with grasslands. Figure 4 shows an example of a large open bog that is almost completely misclassified.

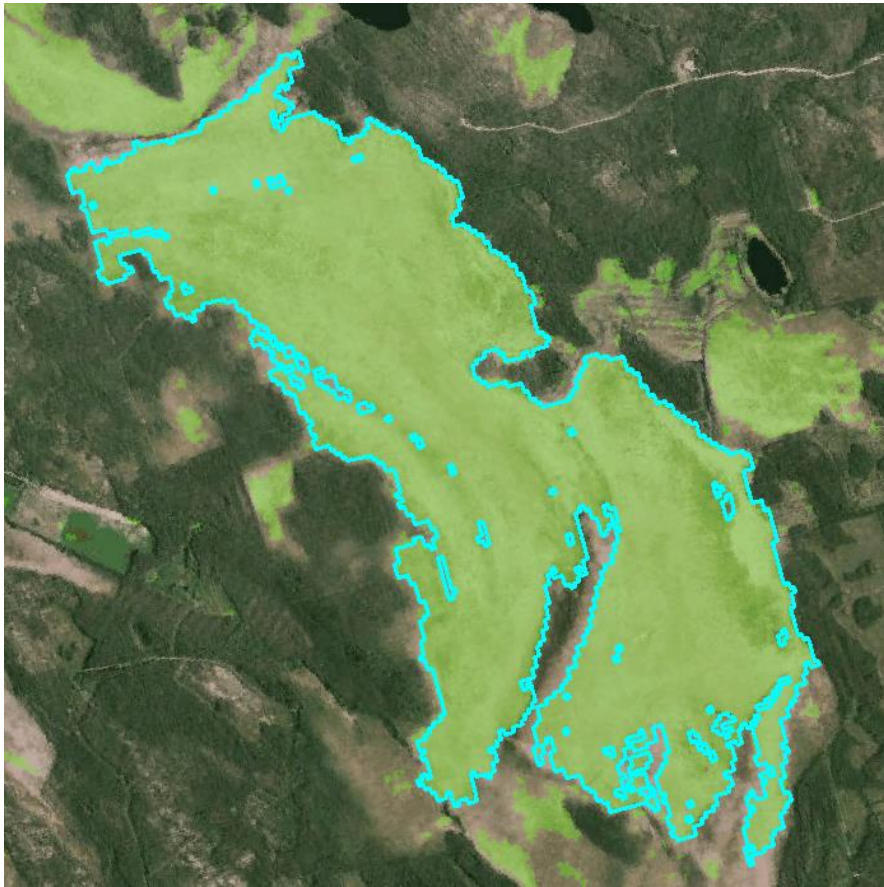


Figure 4. Commission errors in a swamp. The turquoise polygon is the contradicting area derived from the GIS comparison of in-situ data and the HRL GRA18. GRA18 data is presented with light green. SAMPLE\_ID 147, scale 1:20 000, coordinates (ETRS\_1989\_LAEA) E: 5019219 N: 4431245. Ortho photo reference year 2018.

**Reed beds and helophytes dominated systems:** The classification accuracy of this strata was also poor as many reed beds were frequently classified as grassland. Especially reed beds around small lakes or ponds were problematic areas. Figure 5 is an example of this kind of areas.

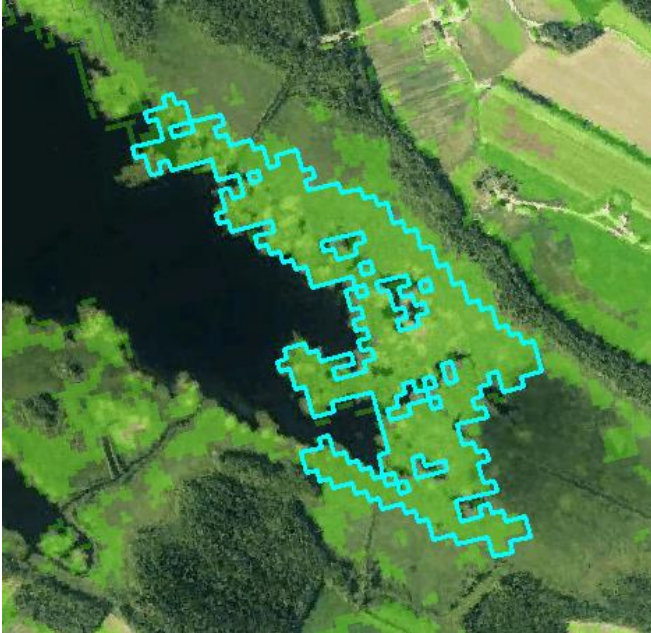


Figure 5. Commission errors in a reed bed around a small lake. The turquoise polygon is the contradicting area derived from the GIS comparison of in-situ data and the HRL GRA18. GRA18 data is presented with light green. SAMPLE\_ID 158, scale 1:10 000, coordinates (ETRS\_1989\_LAEA) E: 5145699 N: 4613173. Ortho photo reference year 2018.

**Tall forbs, fern, shrub dominated vegetation:** The sample locations for this stratum were selected from the samples of the quantitative verification that were observed to contain these elements. This may be the reason why the classification accuracy is slightly higher than in other commission error strata. Yet, the accuracy is insufficient. One example of the stratum can be seen from Figure 6.



Figure 6. Commission errors in a shrub dominated vegetation. The turquoise polygon is the selected sample pixel from the quantitative verification that was observed to contain shrubs. GRA18 data is presented with light green. SAMPLE\_ID 189, scale 1:2500, coordinates (ETRS\_1989\_LAEA) E: 5028417 N: 4411245. Ortho photo reference year 2018.



**Grasslands that have been observed as tilled (in the reference year or a certain period before, in that case they are considered as arable fields) :** This stratum is one of the most problematic strata of the look & feel verification as tilled grasslands are often classified as grassland. In the in-situ data land parcels were considered as tilled if the dominant plant species had changed from a class that is not considered as grassland (e.g. from previous cropland) to a class with assumed grass cover. These classes are listed in the stratification description of the quantitative verification. The conflicting area between the tilled grassland defined this way and the non-grassland areas of the HRL GRA18 is the largest (2333km<sup>2</sup>) in this section which makes it the most erroneous stratum. However, the definition used to find tilled grasslands is also limited as the dominant species can sometimes change without ploughing. Thus, the actual conflicting area might be smaller than the area described here. Figure 7 presents an example of the tilled grasslands.



Figure 7. Commission errors in a tilled grassland where a previous wheat field was transformed to grassland. The turquoise polygon is the contradicting area derived from the GIS comparison of in-situ data and the HRL GRA18. GRA18 data is presented with light green. SAMPLE\_ID 184, scale 1:1000, coordinates (ETRS\_1989\_LAEA) E: 5080185 N: 4644008. Ortho photo reference year 2020.

**Vineyards, orchards, olive groves, (if more than 10 % shrubs or trees):** In this stratum, only orchards were evaluated as the other elements are not relevant in Finland. Apple and blackcurrant plantations were quite often classified as grassland which seem to be the major commission issues of the stratum. One example can be seen in Figure 8 where a black currant plantation is classified as grassland. This is a good example about the fact that the orchards can be easily mixed with proper grassland, especially if the shrubs are low.



Figure 8. Commission errors in a blackcurrant plantation. The turquoise polygon is the contradicting area derived from the GIS comparison of in-situ data and the HRL GRA18. GRA18 data is presented with light green. SAMPLE\_ID 197, scale 1:5000, coordinates (ETRS\_1989\_LAEA) E: 5217437 N: 4292930. Ortho photo reference year 2018.

**Tundras dominated by shrubs and lichens:** Tundras dominated by shrubs and lichens were very poorly identified in the HRL GRA18 data (Figure 8). This can be partly explained by the fact that they can be quite easily confused with heathland. It was also quite challenging to separate these two strata during the verification. Typically, they were also blend with brushes. Luckily the Biotope Maps (Metsähallitus) could give some hints whether the ground cover consisted of lichens or heathland.

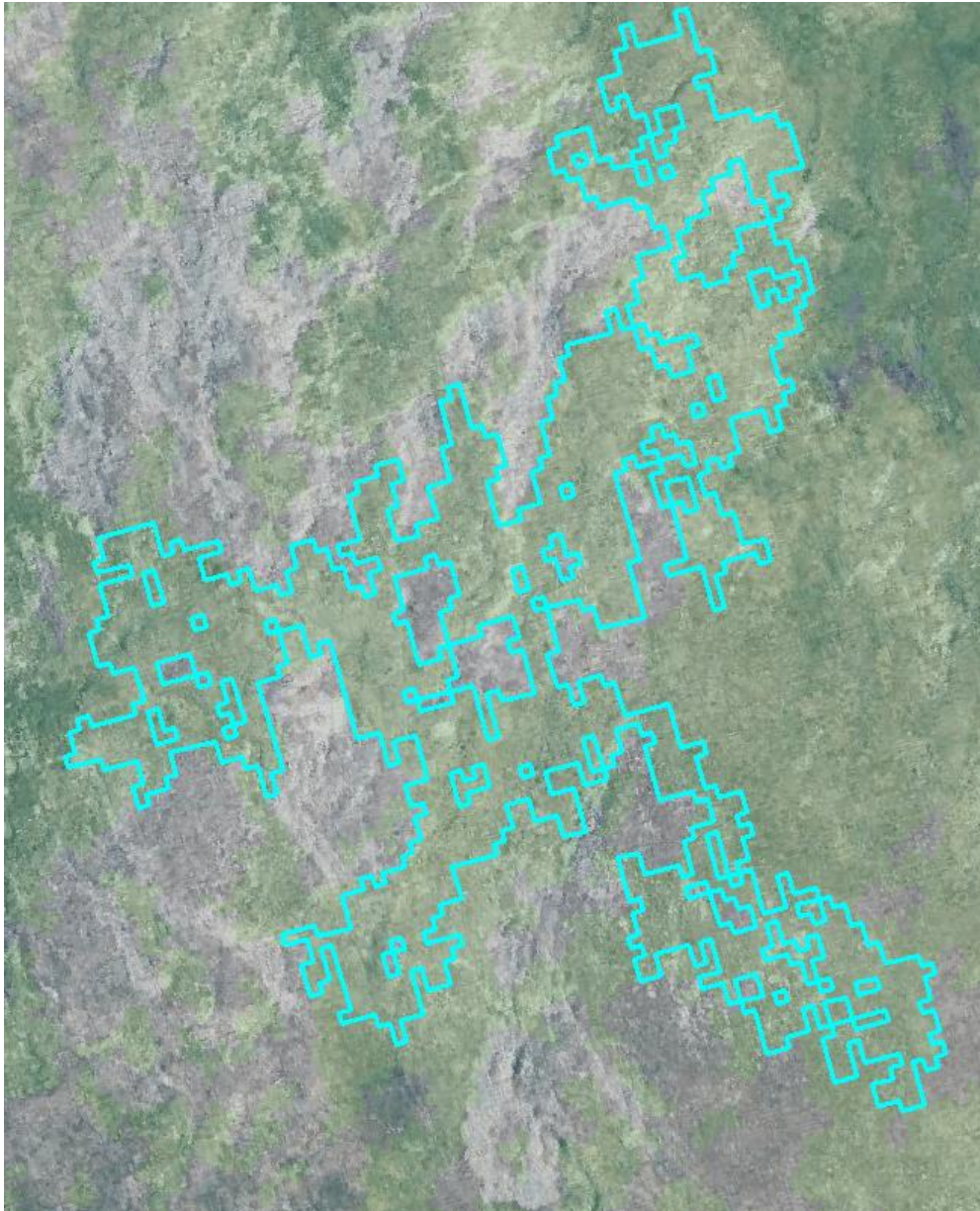


Figure 8. Commission errors in a tundra dominated by lichens. The turquoise polygon is the contradicting area derived from the GIS comparison of in-situ data and the HRL GRA18. GRA18 data is presented with light green. SAMPLE\_ID 208, scale 1:5000, coordinates (ETRS\_1989\_LAEA) E: 4800190 N: 5108828. Ortho photo reference year 2012.

**Clear-cut areas in the woods:** Clear-cut areas were very poorly identified in the HRL GRA18 data. The notes varied between 1 and 2. Clear-cut areas seem to be one of the major sources of commission errors as they are frequently misclassified as grassland. Figure 9 shows an example of a clear-cut area where the forest was chopped between years 2012 and 2016.

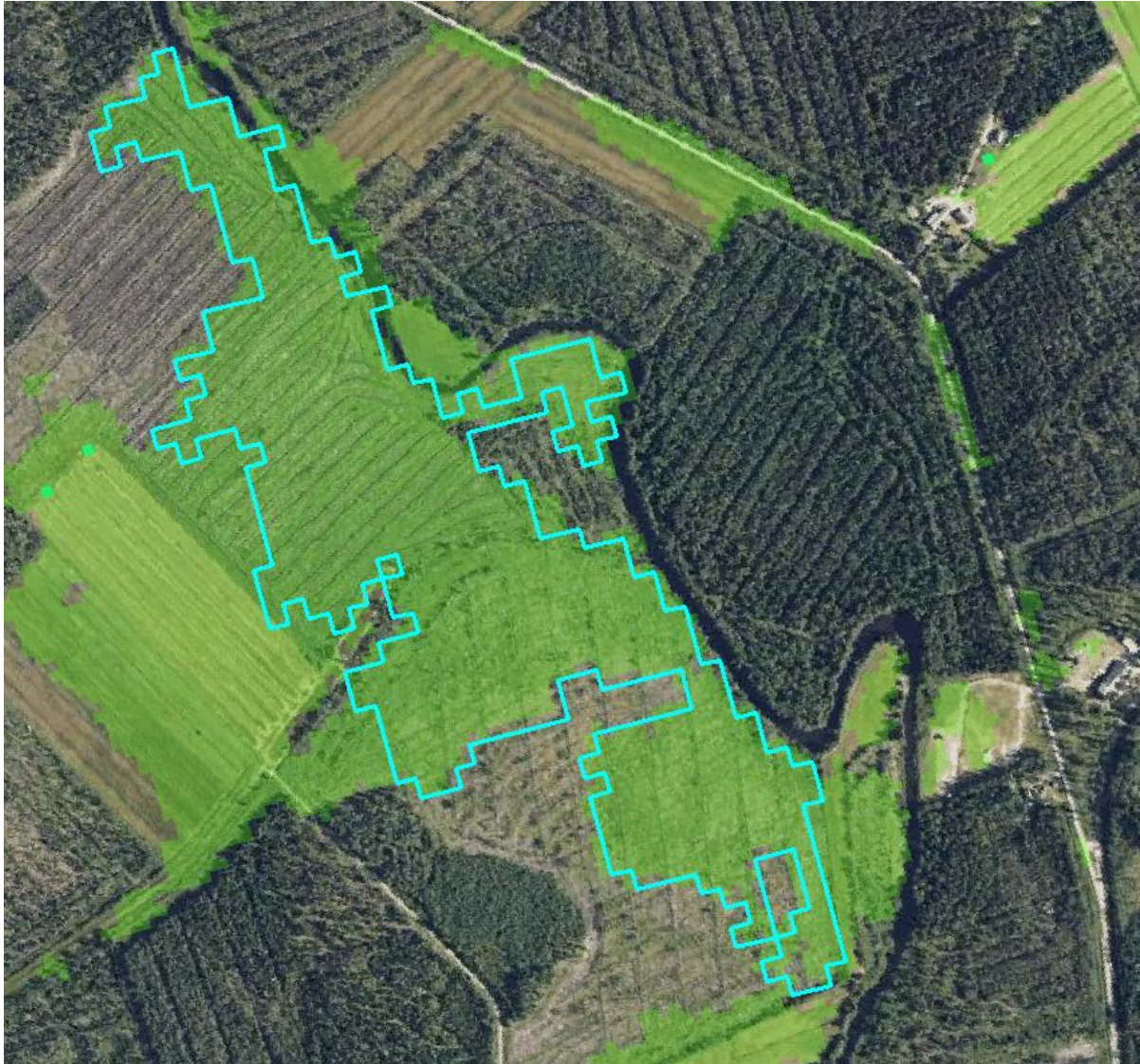


Figure 9. Commission errors in a clear-cut area where the logging was performed between years 2012 and 2016. The turquoise polygon is the contradicting area derived from the GIS comparison of in-situ data and the HRL GRA18. GRA18 data is presented with light green. SAMPLE\_ID 213, scale 1:5000, coordinates (ETRS\_1989\_LAEA) E: 4989585 N: 4865124. Ortho photo reference year 2016.

**Cropland:** Cropland was also very poorly interpreted in the HRL GRA18 data. Additionally, it is a major source of Commission errors as cropland was often misclassified as grassland. These areas were mostly previous grassland that was changed to cropland. Oat and barley fields were the most usual crop types that caused classification errors. Figure 10 displays an example of cropland that was, according to FLIPS, grassland before 2018.

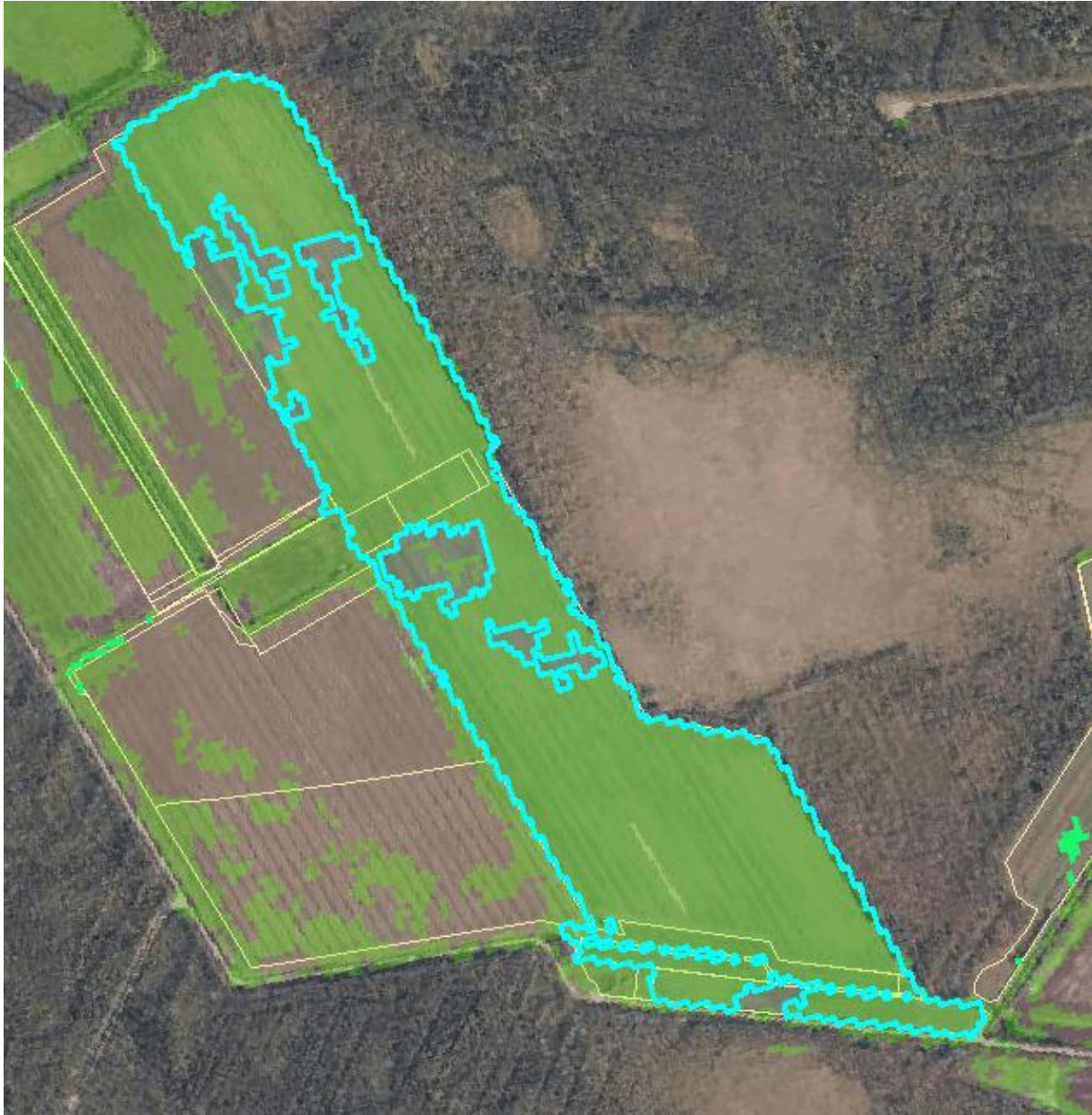


Figure 10. Commission error in cropland that was previously grassland. The turquoise polygon is the contradicting area derived from the GIS comparison of in-situ data and the HRL GRA18. GRA18 data is presented with light green. SAMPLE\_ID 228, scale 1:10000, coordinates (ETRS\_1989\_LAEA) E: 5121874 N: 4635013. Ortho photo reference year 2018.

VI. **Statistical verification** (optional)

<p>Description of methodology and software</p>	<p>Statistical verification was performed using GIS-software. Samples were selected with Matlab (rand-function) and they were validated against national in-situ datasets using ArcMap 10.8. Random samples were selected as following:</p> <ul style="list-style-type: none"> <li>- All non-grassland areas: 300 samples were selected from stratified non-grassland area of the HRL GRA18 data.</li> <li>- Grassland areas: 300 samples were selected from the grassland area of the HRL GRA18 data</li> </ul> <p>These sample pixels were checked against relevant in-situ datasets and assessed as correct/in-correct. In case of incorrect, an accurate class was given. For most of the pixels, comments on the land cover of the location was given. These pixels are attached to the report (GRA2018_QuantitativeSamples_FI.shp).</p> <p>The results of the statistical verification were used as an input to the Map Accuracy Tool using following steps:</p> <p>Step1: Create a matrix with the probabilities (number of samples in a raster cell / sum of row)</p> <p>Step 2: Convert probabilities to areas (area of stratum * probability)</p> <p>Step 3: Calculate Producer accuracies (area of agreement / sum area in the column)</p> <p>This way the areas of the different strata were used as weights in the tool to calculate Producer's accuracies. These areas were:</p> <ul style="list-style-type: none"> <li>- Non-grassland: 4875 km<sup>2</sup></li> <li>- Grassland: 17131 km<sup>2</sup></li> </ul> <p>These steps are also demonstrated in the attached Excel-file: GRA2018_MapAccuracyToolResults.xlsx</p>
<p>Stratification</p>	<p>Stratification was used to select non-grassland sample points. These points were selected from an area, that national in-situ data indicates to be grassland but is non-grassland according to the HRL GRA18 data. The in-situ data used was a combination of the following datasets:</p> <ul style="list-style-type: none"> <li>- National High Resolution Corine Land Cover Data for years 2012 and 2018 where pixels considered as grassland if they belonged during the both reference years to one of the following classes.</li> </ul>

	<p>2311=Pasture, 2312= Natural pasture, 2431=Arable land outside farming subsidies, 1421= Golf Courses,3211=Natural grassland.</p> <ul style="list-style-type: none"> <li>- The Topographic Database where polygons were considered as grassland if they belonged to the class 32900=Park in all the reference years 2012,2017 and 2019 or to the class 32800=Meadows in all the reference years 2012, 2014,2016 and 2017)</li> <li>- The Finnish Land Parcel Information System (FLPIS) where polygons were considered as grassland if their dominant plant species belonged (in all reference years 2012-2015 &amp; 2018) to one of the following classes: Special subsidy contract area, permanent pasture; pasturage (open); Natural treatment field, herbaceous plants; Natural pasture and meadow; Perennial environmental turf; Perennial dry hay, silage and fresh fodder grassland; Perennial pastures; Perennial cultivated grassland for seed; Permanent dry hay, silage or fodder field (min 5, max 10 years); Permanent pasture (min 5, max 10 years); Grassed buffer; Exclusion area; Grassed Exclusion area; green fallow; natural subsidy zone with permanent grassland.</li> </ul> <p>This area was in total 4875 km2.</p>
Comments	



## Confusion Matrix

		Reference Data			
		Non-grassland	Grassland	User Accuracy	UserAccuracyVariance
Classification Data	Non-grassland	1073	3803	22,01 %	0,01163
	Grassland	9650	7481	43,67 %	0,007427
	Weights	4875	17131		
	ProducerAccuracy	10,00 %	66,30 %		
	ProducerAccuracyVariance	0,004904	0,005054		
	PortmanteauAccuracy	38,87 %	38,87 %		
	PortmanteauAccuracyPartial	7,39 %	35,74 %		

<b>OverallAccuracy</b>	<b>0,388702</b>
<b>OverallAccuracyVariance</b>	<b>0,00633</b>
<b>AllocationDisagreement</b>	<b>0,345562</b>
<b>Shift</b>	<b>0</b>
<b>Exchange</b>	<b>0,345562</b>
<b>QuantityDisagreement</b>	<b>0,265736</b>
<b>AMI</b>	<b>0,061665</b>
<b>AMIAdjusted</b>	<b>-0,13464</b>
<b>AMIVariance</b>	<b>0,005302</b>
<b>Kappa</b>	<b>-0,24022</b>
<b>KappaVariance</b>	<b>0,020625</b>