HRL verification report template for HRL Imperviousness in Finland

I. Administrative part

HRL	*type the name of the verified layer*
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)	Iida Autio, coordinator, iida.autio@ymparisto.fi
Look-and-feel analysis done by (name, position and e-mail)	Iida Autio, coordinator, iida.autio@ymparisto.fi
Statistical verification done by (name, position and e-mail)	Iida Autio, coordinator, iida.autio@ymparisto.fi Markus Törmä, research engineer, markus.torma@ymparisto.fi
In situ data used. *Replace Data-x with the full name of the dataset. Mention quality issues if relevant.*	*National Ortho photo database/The National Land Survey* *Natural color/black and white ortho photos* *Resolution: 0.25-0.5m* *Reference years: 2014-2016 (partial coverages)*
	National high resolution Corine Land Cover 2012 (HR CLC2012) *National Corine raster dataset* *Resolution 20x20m*
	Topographic Database/The National Land Survey *Compilations of object groups (buildings)* *Vector data* *Reference year: 2015*
	The National Road and Street Database, Digiroad *Vector dataset* *Reference year: 2015*
	Google street view photos
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 HRL built-up areas (in the built-up map derived from the impervious HRL-layer[1]) is 1930 km^2. The built-up area according to the National high resolution Corine Land Cover 2012 data (20x20m) is 7162 km^2. This includes CLC12 classes 1.1.1, 1.1.2, 1.2.1, 1.2.2, 1.2.3 and 1.2.4 as recommended by the guidelines of verification. This indicates that HRL imperviousness underestimate impervious surfaces in Finland. This is partly due to the fact that the built-up map includes only the areas with ≥30% imperviousness. If areas with 1-29 % imperviousness are added, the total area of build-up surfaces increase into 4589 km2. On the other hand, these two datasets are not fully comparable, since the HRL imperviousness represents pure land cover, while HR CLC12 is a mixture of land cover and land use. Thus discontinuous urban fabric class 1.1.2 includes significant amount of green areas around houses.

The HRL built-up map (IMD > 30%) was combined with the national HR CLC12 and the content of the built-up map was analyzed by calculating distribution of land cover classes within HRL build-up map as mapped in the HR CLC12 data. The largest shares are presented in Table 1.

Table 1. Shares of national HR CLC12 data within HRL built-up map

Corine Land cover class	**Distribution of HRL built-up according to HR CLC12.**
Continuous urban fabric (1.1.1.1)	5,5 %
Discontinuous urban fabric (1.1.2.1)	25,1 %
Industrial or commercial units (1.2.1.1)	14,4 %
Industrial or commercial units (1.2.1.2)	12,5 %
Road and rail networks and associated land (1.2.2.1)	17,8 %
Coniferous forest (3.1.2.1)	4,9 %
Transitional woodland/shrub (3.2.4.1)	6,3 %

The HRL built-up map includes mostly appropriate HR CLC12 classes but also forest areas (4,9 %) and transitional forest and shrub (6,3 %).

The overlay analysis also revealed areas which are built-up in HRL-data and non-built-up in the national HR CLC12 dataset (commission errors). Table 2.indicates the shares of national HR CLC12 classes that are misclassified as impervious in the HRL-feature layer.

[1]Built-up map (derived from HRL imperviousness) decision rule: a 20m x 20m area is considered built-up, if imperviousness ≥ 30% .

[2]Rakennetun alueen pinta-alan ennakointi – paikkatietoaineistojen ja -menetelmien hyödyntäminen rakennetun alueen muutosten laskennassa, Suomen ympäristökeskuksen raportteja 28/2015.

Table 2.Shares of national HR CLC12 classes misclassified as built-up in HRL dataset

Corine Land cover class	**Share in HRL built-up/CLC12 non-built-up**
Sport and leisure facilities (1.4.2.2)	7,0 %
Non-irrigated arable land (2.1.1.1)	7,3 %
Coniferous forest (3.1.2.1)	21,9 %
Transitional woodland/shrub, cc<10 % (3.2.4.1)	28,5 %
Transitional woodland/shrub, cc 10-30 % (3.2.4.2)	10,2 %

Sport and leisure facilities (CLC 1.4.2.2) seem to be often included in the HRL built-up map even though their land cover is mostly natural grass or sand fields. One fifth (21 %) of the misclassified built-up area is coniferous forest. The misclassified areas are mostly small (1-2 pixels). This indicates that the high percentage is due to a large amount of small single misinterpreted pixels around and amongst built-up areas and roads. Same applies to the commission errors on arable land (7,3 %).

Omissions were found in areas belonging to the selected artificial surface classes of the HR CLC12 (classes 1.1.1, 1.1.2, 1.2.1, 1.2.2, 1.2.3 and 1.2.4). Largest proportion of these areas (46%) belongs to CLC class 1.1.2. This is mainly due to differences in class definitions as mentioned above. The CLC class with majority of the omission errors was Road and rail networks and associated land 1.2.2 (35%). Discontinuity of the roads is clearly visible throughout the IMD data. Some roads are left out of the HRL built-up map due to resolution of input satellite data. Many of the roads in the HR CLC12 data are not paved and thus appropriately left out of the HRL built-up map.

The differences between the overlaid datasets are presented in Figure 1.

By visual scanning of the HRL built-up/CLC12 non-built-up areas, a slight (<1 pixel) geometric shift can be detected on roads. There is no consistent direction to this shift. There seems to be a slight misinterpretation of the impervious surfaces in the south western coastal zone of the Baltic Sea, where the built-up map includes areas of water right at the coastline. This mistake is restricted to a fairly small area and cannot be considered as geometric shift as it applies to all directions (see figure 6.).

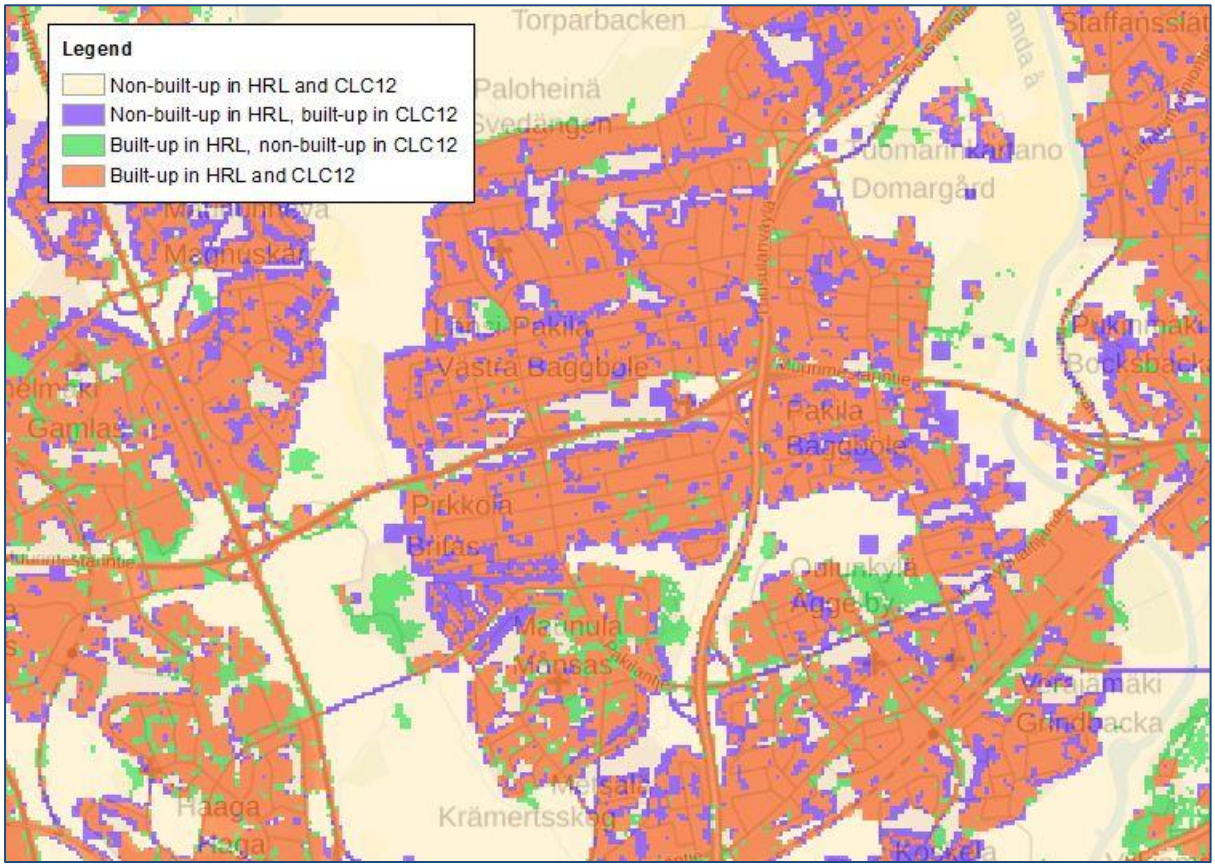

Figure 1. Overlay analysis showing potential commission (green) and omission (purple) errors. Built-up areas in HR CLC12 include classes 1.1.1, 1.1.2, 1.2.1, 1.2.2, 1.2.3 and 1.2.4.

III. Look-and-feel

Stratum	Name of the stratum	Number of samples verified	*Results of the verification by strata (excellent, good, acceptable, insufficient, very poor)*
1	Agriculture areas around built-up	(*)	Good - large areas of agricultural land was not found in HRL built-up.
2	Major cities	(*)	Good - small green areas amongst buildings are mis-classified as HRL built-up but in general, stratum is well recognized.
3	Continuity of major highways	(*)	Insufficient - Discontinuity even on major roads.
4	Sport and recre-ation areas	(*)	Acceptable - Sports fields covered by sand or grass are often included in HRL built-up.
.5	Coniferous for-est/transitional forest	(*)	Acceptable - total area of misclassified forests amongst urban fabric and roads is large but single areas are small.
Overall evaluation			*good/acceptable (excellent, good, acceptable, insuffi-cient, very poor)*
Comments			

(*) In the statistical verification totally 559 locations were interpreted and checked, which gave also detailed and statistically unbiased look-and-feel impression of HRL data including critical strata.

IV. Statistical verification

Stratification	*Sample plots for determining omission errors were concentrated in areas of potential errors. These areas included CLC classes 111, 112, 121, 122, 123, 124 and 142 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.*
Comment on stratification	
Number of random samples for finding omission errors	*280*
Number of valid (applicable) samples for finding omission errors	*280*
Omission error (%)[3] with uncertainty (calculated for the stratified HRL-off area)	43,4 %; uncertainty 195,6 %[5] (262,0 %[6])
Comment on omissions	
Number of random samples for finding commission error	*280*
Number of valid (applicable) samples for finding commission error	*279: one sampling point was selected twice*
Commission error (%)[4] with uncertainty	25,1%; uncertainty 2,6 %[7]
Comment on commissions	
Overall evaluation	*General overview, look-and-feel as well as statistical verification indicate that the HRL Imperviousness layer has succeeded fairly well in mapping the sealed areas in Finland. There are some systematic errors such as discontinuity of major roads and misinterpretation of forest as built-up in and around discontinuous urban fabric.*

[3] Producer's accuracy (%) = 1 – omission error (%)
[4] User's accuracy (%) = 1 – commission error (%)
[5] Uncertainty calculated as instructed in the Annex1 of the verification guide. The term "$Area_{HRLclass}$" in the formula is corrected for omission and commission errors ($Area_{RealHRLclass}$).
[6] Uncertainty 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.
[7] Calculated to correspond to a significance level of appr. 68,3 % as instructed in Annex1 of the verification guidelines.

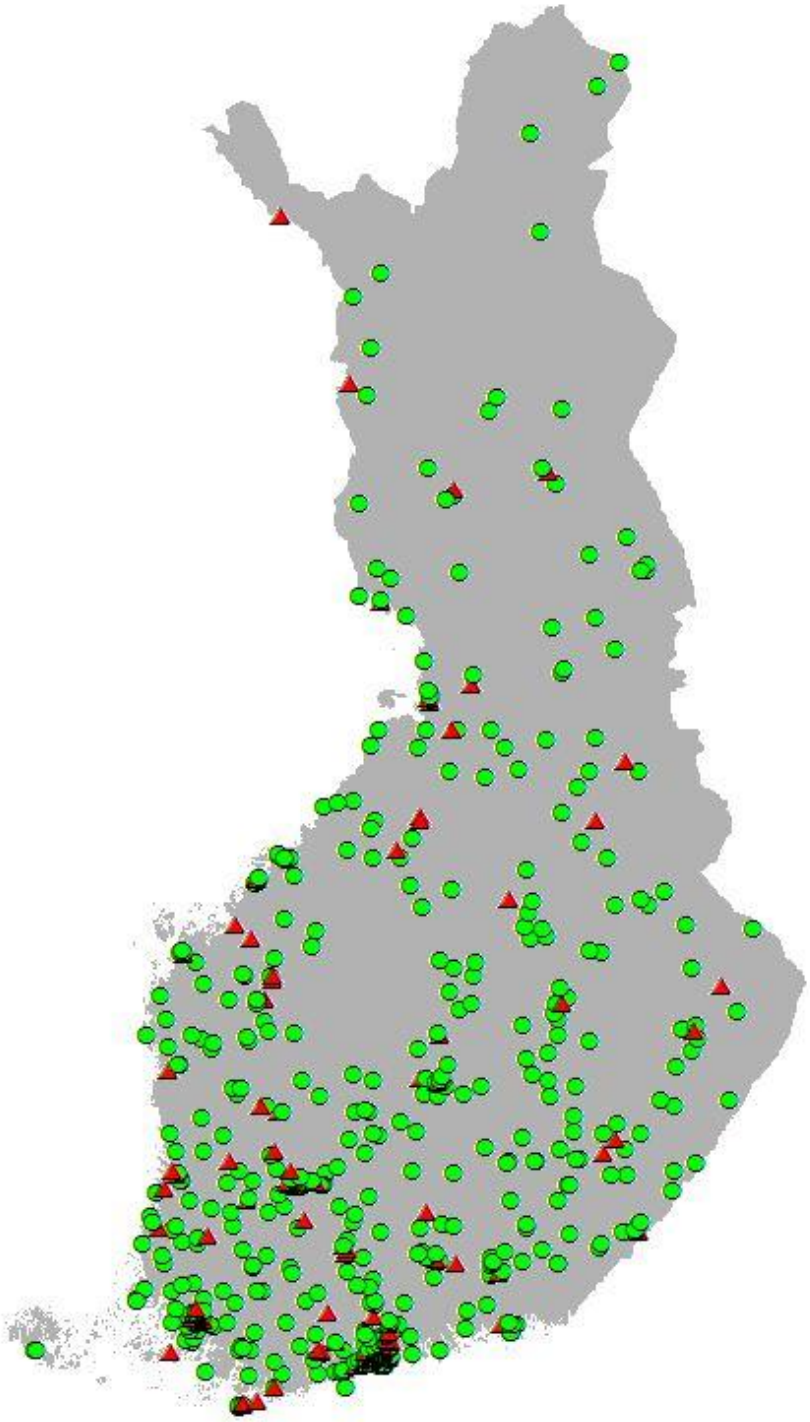

Figure 2. Areal distribution of sample plots in statistical verification. Green sample plots are correctly (both commission and omission) and red plots incorrectly interpreted as built-up in the HR data.

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 1:2000 - 1:4000. HRL built-up map is displayed as transparent purple. In the first image also imperviousness densities of 1-30% are displayed in a scale from yellow to red.

Figure 3. Example of the discontinuity of roads. HRL built-up (purple) does not cover the whole road area and even HRL-on area in densities of 1-30% (a scale from yellow to red) is discontinuous. Scale 1:3000, coordinates E: 524550, N: 7397750.

Figure 4. Green areas (forest) >MMU amongst and in the edges of buildings are incorrectly included in the HRL built-up (purple). Scale 1:2000, coordinates E: 377150, N: 6678950.

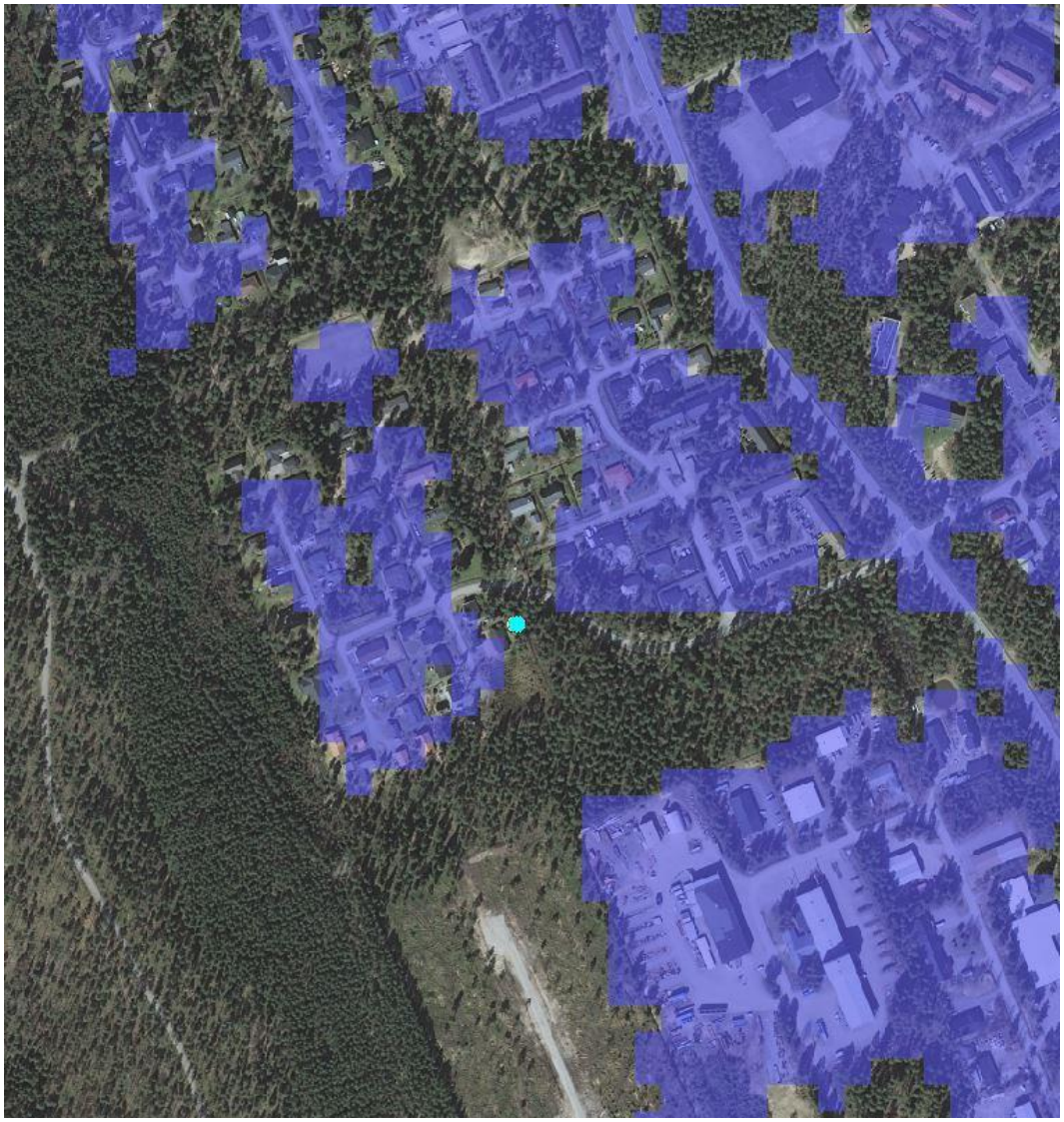

Figure 5. Forest area is incorrectly included in the HRL built-up (purple) amongst and in the edges of urban fabric. Scale 1:4000, coordinates E: 263950, N: 6707550.

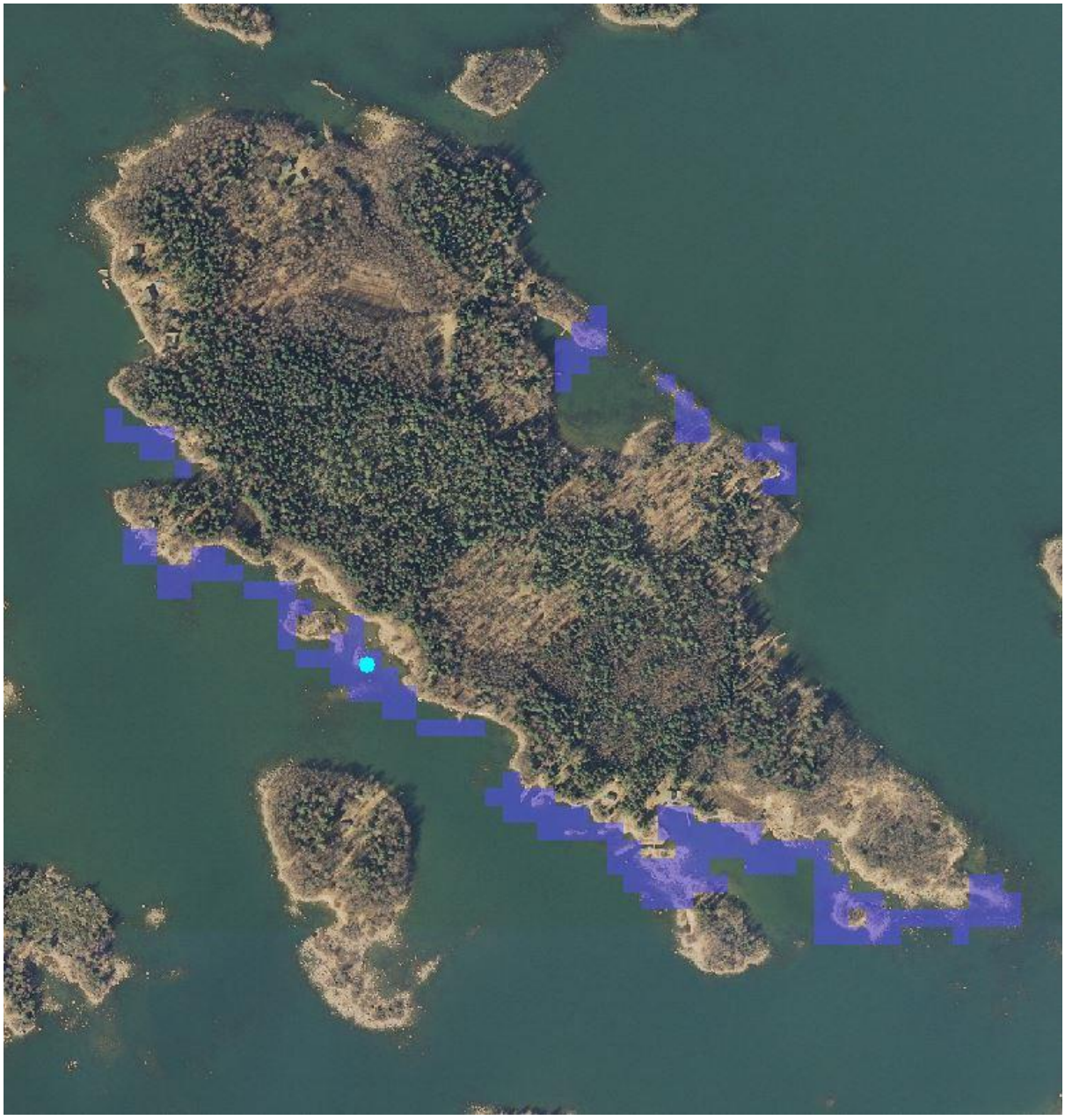

Figure 6. HRL built-up (purple) is misclassified with water at the south western Baltic Sea coast-line. Scale 1:6000, coordinates E: 201163, N: 6819783.

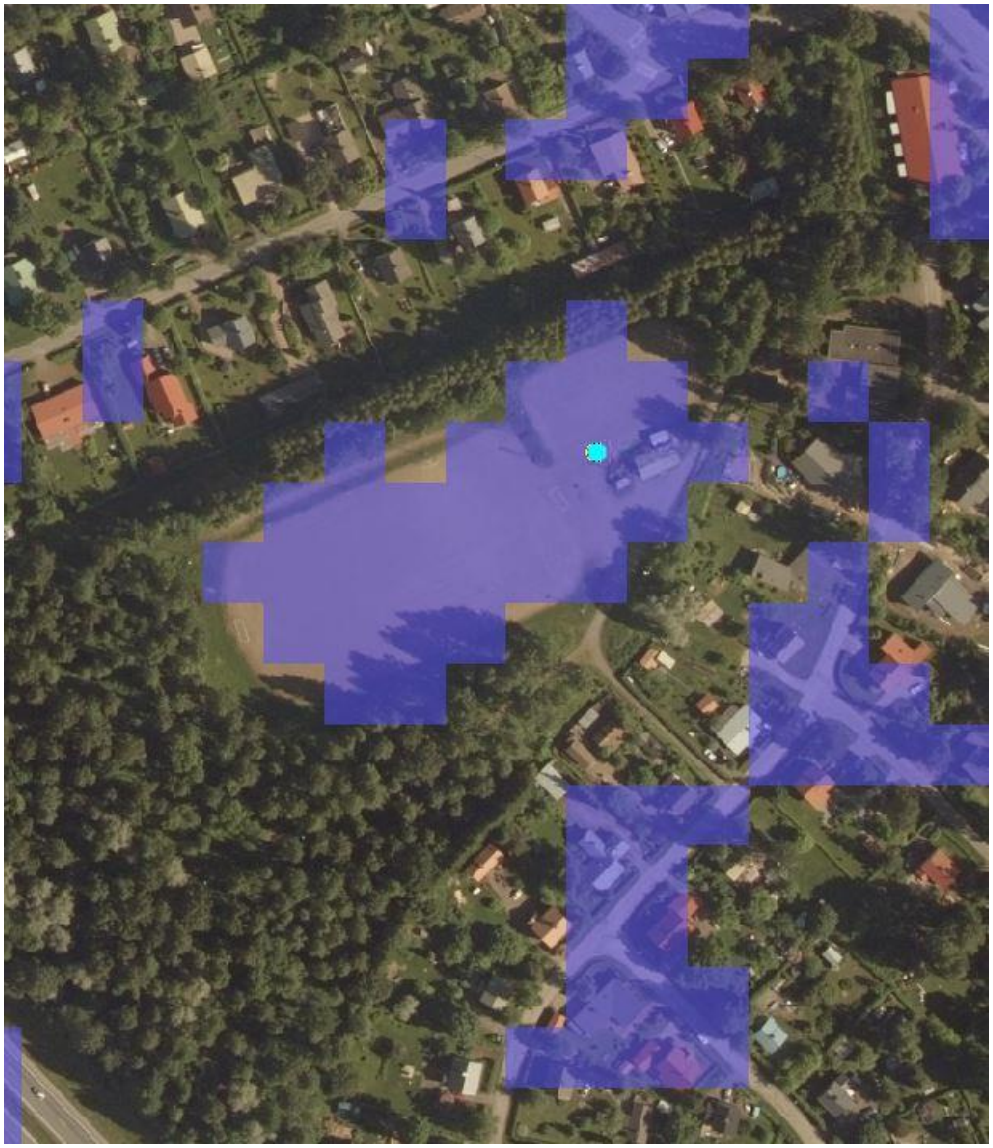

Figure 7. A sandy sports field incorrectly included in the HRL built-up (purple). Scale 1:2000, co-ordinates E: 361350, N: 6766550.

VI. Documentation of software used for verification

The software type and 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