Norway grants

The National Centre for Research and Development

Can we improve the accuracy of the land cover classification by pre-selection of the reference samples and applying DEM in the mountain area in Norway?

Adam Waśniewski[1], Agata Hościło[1], Linda Aune-Lundberg[2]

[1]Institute of Geodesy and Cartography, Centre of Applied Geomatics, Poland

[2]Norwegian Institute of Bioeconomy Research, Division of Survey and Statistics, Norway

Contact: adam.wasniewski@igik.edu.pl

IGiK NIBIO NORWEGIAN INSTITUTE OF BIOECONOMY RESEARCH IOŚ-PIB Lodz University of Technology EVERSIS

InCoNaDa Project

InCoNaDa Project– “Enhancing the user uptake of Land Cover / Land Use information derived from the integration of Copernicus services and national databases”

One of the goals:

- to developed the most accurate land cover map based on a time series of Sentinel-2 data using machine learning approach.

Consortium

- Institute of Geodesy and Cartography (IGiK),
- Norwegian Institute of Bioeconomy Research (NIBIO),
- Institute of Environmental Protection – National Research Institute (IEP-NRI),
- Łódź University of Technology (LUoT),
- Eversis Sp. Z o.o.

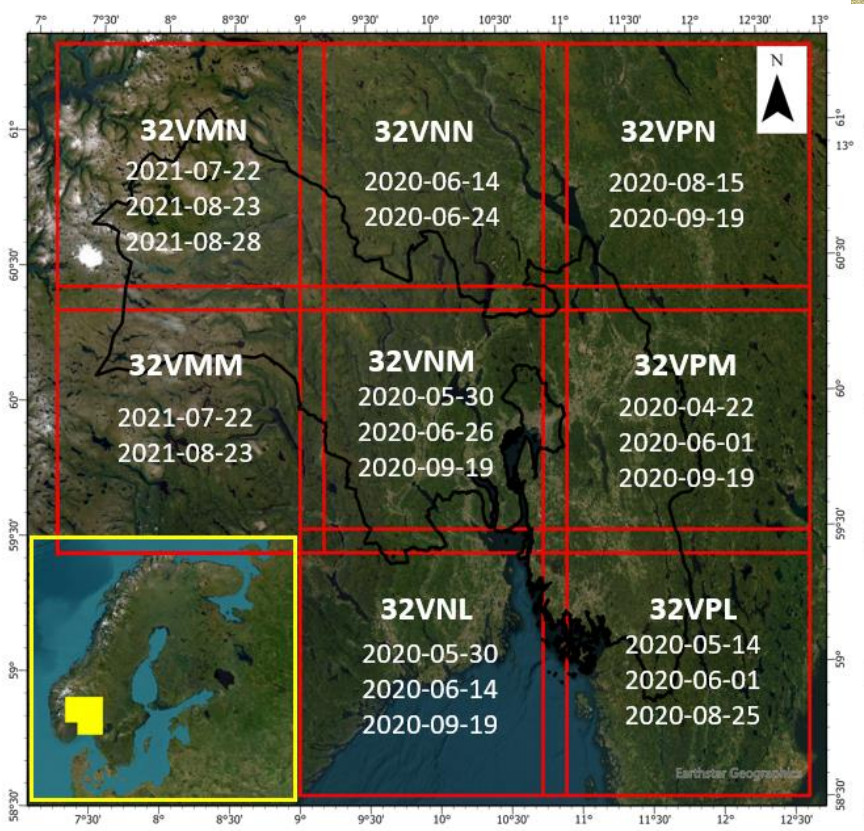
Study area and land cover classes

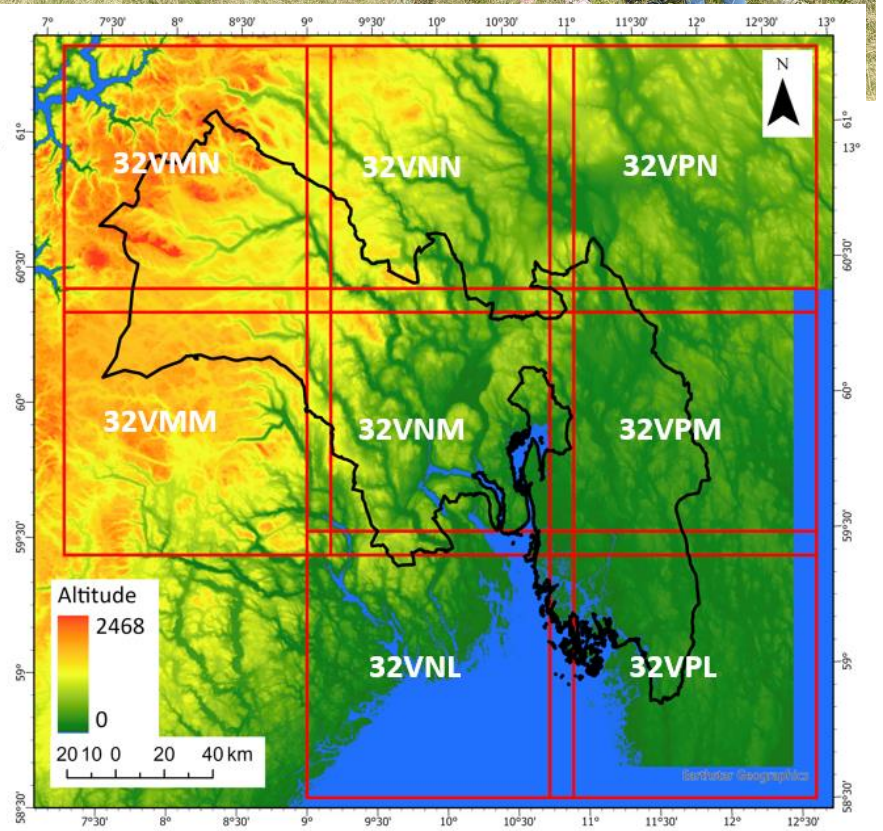
Objectives:

- to develop land cover map,
- to assess the impact of reference data on the classification result,
- to assess the impact of the Digital Elevation Model on the accuracy of classification.

The land cover classes:

- sealed surfaces,
- woodland coniferous,
- woodlands broadleaved,
- low vegetation,
- permanent herbaceous,
- periodically herbaceous,
- mosses,
- non- and sparse vegetation,
- water,
- snow and ice.





Reference dataset

AR50 is the Norwegian medium-resolution land resource dataset, which covers the whole of mainland Norway, scale of 1:50 000. Features in AR50 are polygons with attributes assigned according to the AR50 classification criteria. The primary classification is using nine land types representing a combination of land cover and land use. The primary classes are built-up areas, agriculture, forest, peat bogs, open areas, glaciers, fresh water, oceans and not mapped area. The minimum mapping unit for the primary classes in AR50 is 1.5 hectares.

AR5 is a more detailed version of the AR50 database and includes land cover information below the tree line.

Reference points:

- points created randomly,
- an internal buffer of 10 m and 20 m for water was applied,
- proportions: 2 points per km^2 ,
- minimum of 20 m distance between points,
- minimum of 200 points for one class on one granule,
- points divided into: 60% training and 40% verification.

ARTYPE/arealtype (land type)

Value	Land type
10	Built-up area
20	Agriculture
30	Forest
50	Open land
60	Peat bog
70	Glacier and permanent snow
81	Fresh water
82	Ocean
99	Not mapped

ARVEGET/vegetasjonsdekke (open land)

Value	Open land
51	Not vegetated
52	Sparse vegetation
53	Lichen
54	Intermediate vegetation
55	Vigorous vegetation
98	Not relevant
99	Not recorded

ARSKOGBON/skogbonitet (forest site class)

Value	Forest site class
18	High/particularly high productivity
13	Medium productivity
12	Low productivity
11	Non-productive
98	Not relevant
99	Not recorded

ARTRESLAG/treslag (tree type)

Value	Tree type
31	Coniferous forest
32	Deciduous forest
33	Mixed forest
39	Not forested
98	Not relevant
99	Not recorded

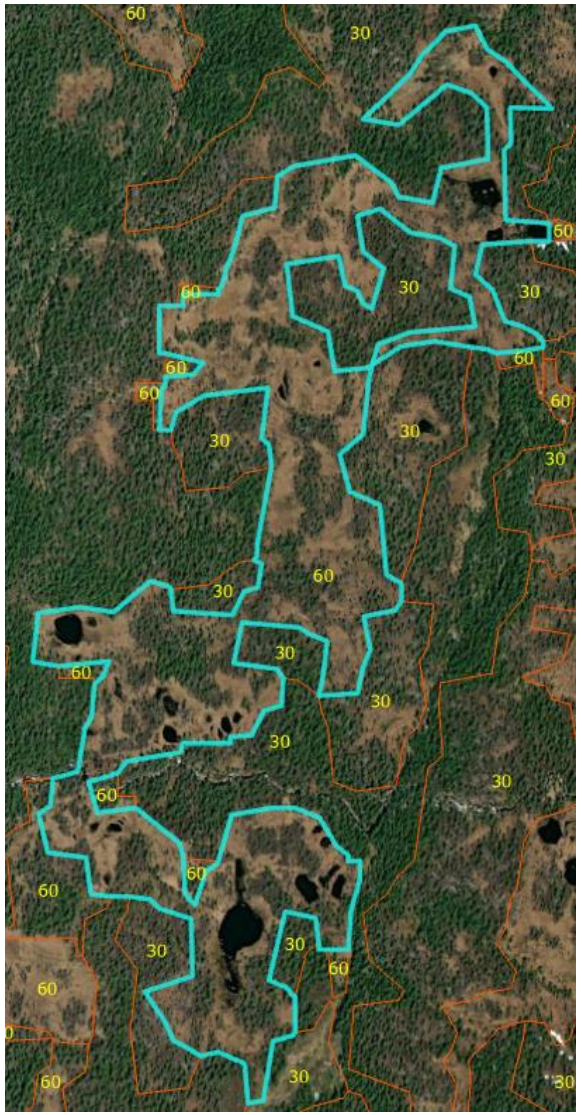
ARJORDBR/jordbruk (agricultural)

Value	Arable land
24	Fully and surface cultivated land
25	Pasture-land
98	Not relevant
99	Not recorded

ARDYRKING/dyrkbariord (arable land)

Value	Arable land
81	Non-cultivable soil
82	Cultivable soil
98	Not relevant
99	Not recorded

Problematic land cover classes



Mosses

Permanent herbaceous

Land cover classification

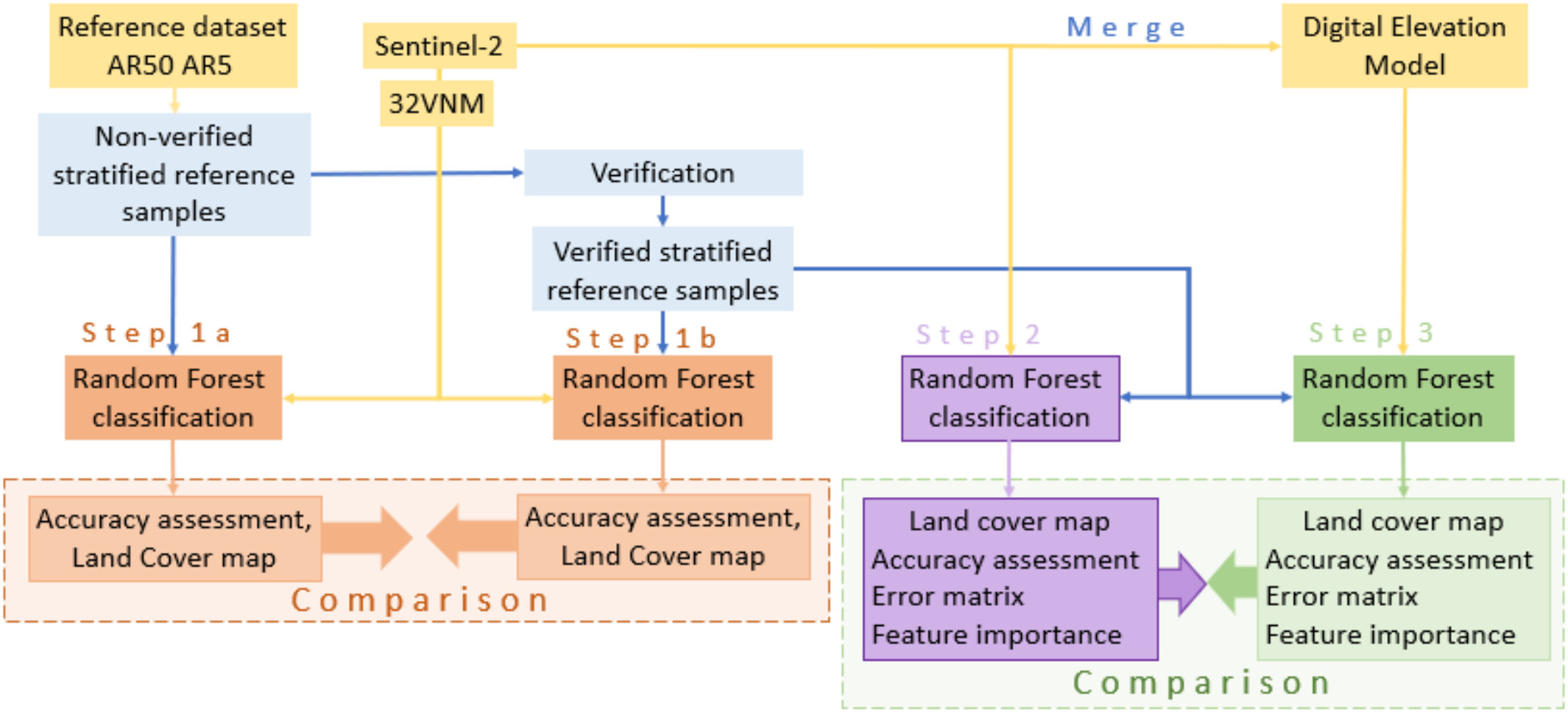
Land cover classification:

- Data – set of cloud free Sentinel-2A and -2B images captured between end of April and September 2018 and 2020.
- Clarification was carried out using the Random Forest machine learning algorithm.
- The analyses were carried out in the Python programming environment.
- The stability of the classification model was assessed using iterative accuracy assessment (iteration 100 times).
- Analyses were performed in cloud computing environment Amazon Web Services (AWS).
- After preliminary analyses, a Digital Elevation Model was included in the classification.

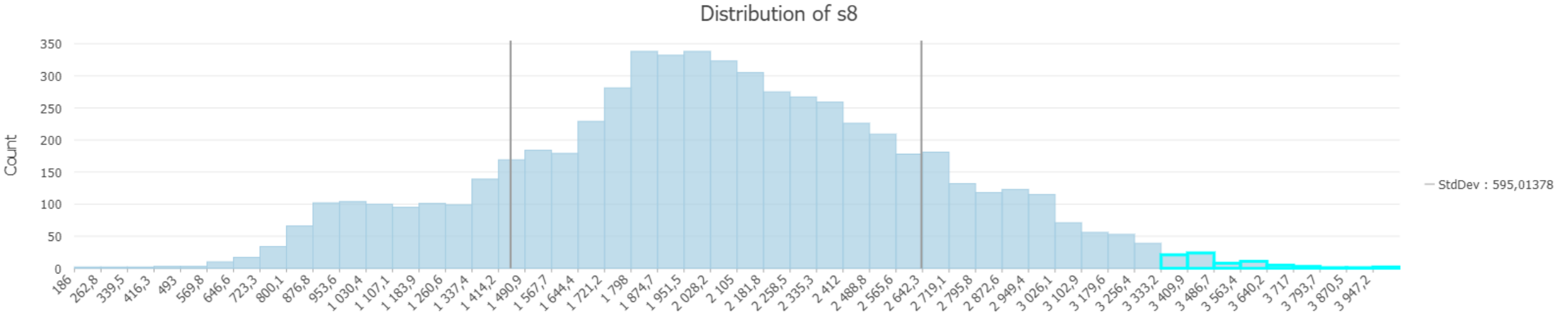
Digital Elevation Model:

- Shuttle Radar Topography Mission data (SRTM) is available from 56°S to 60°N.
- The freely available Digital Elevation Model was used for continental Norway with a spatial resolution of 10 meters, provided at geonorge.no.

Methodology scheme

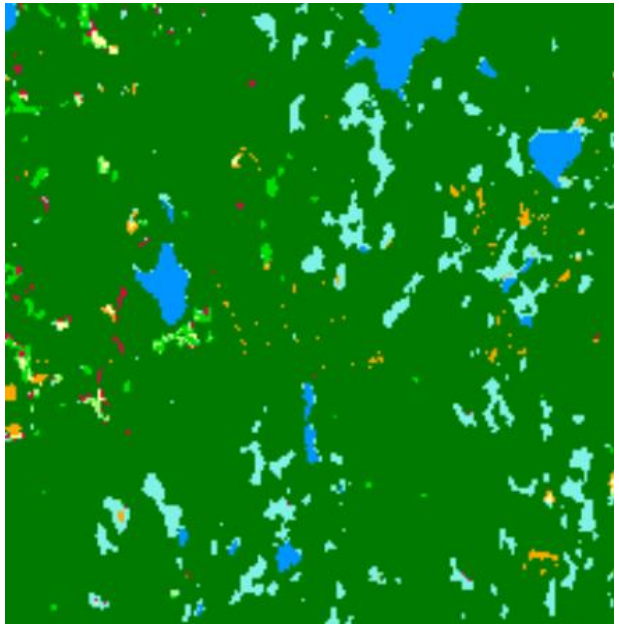


Histogram analysis – locating and removing points

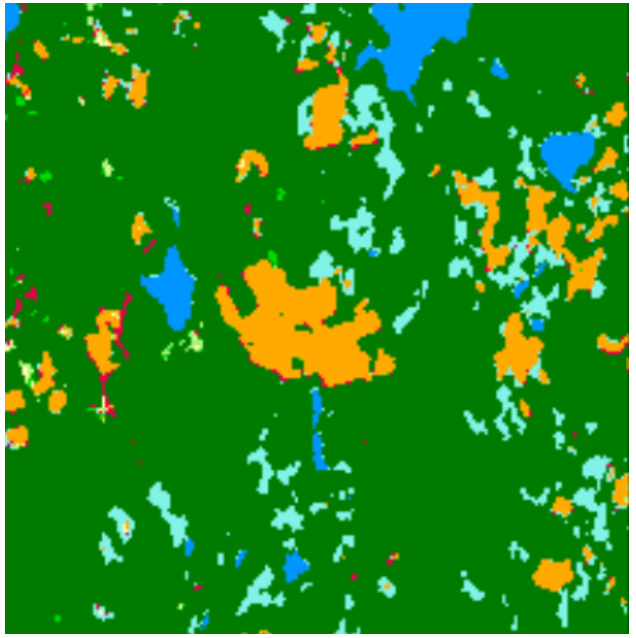


Locating points on clear-cuts

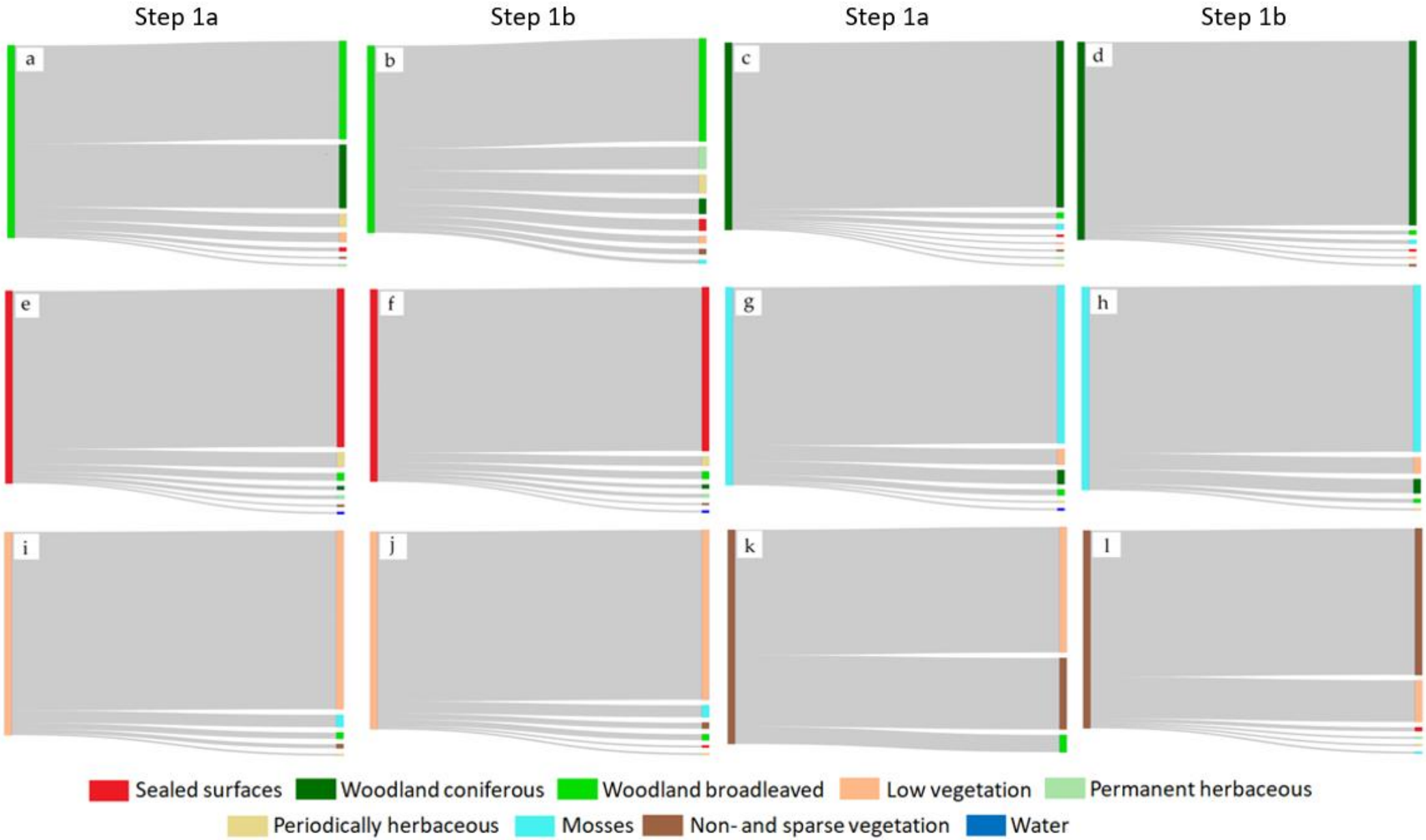
Classification result before removing points

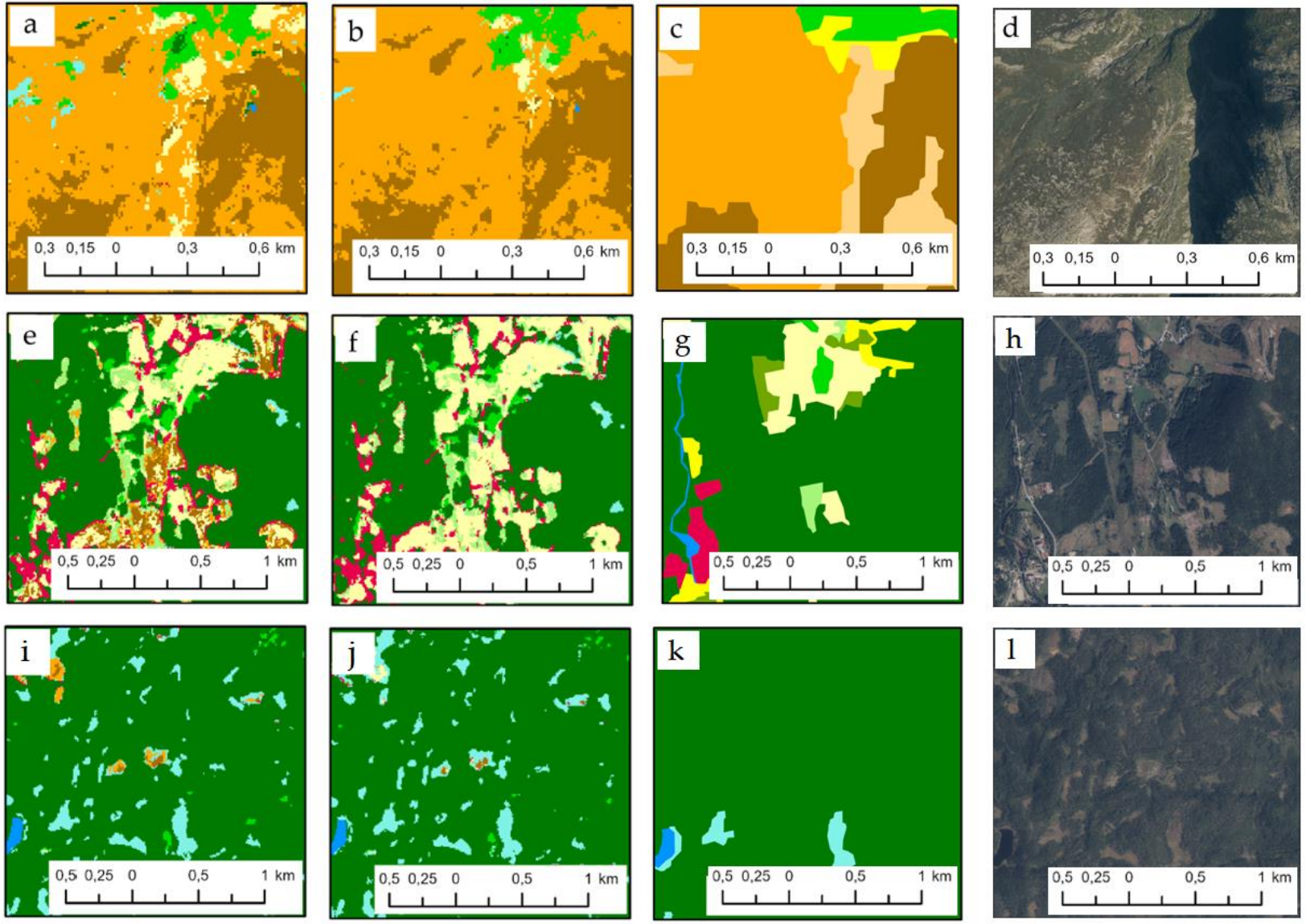


Classification results after removing points



Impact of reference samples





a, e, i – Non-verified stratified reference samples (step 1a)

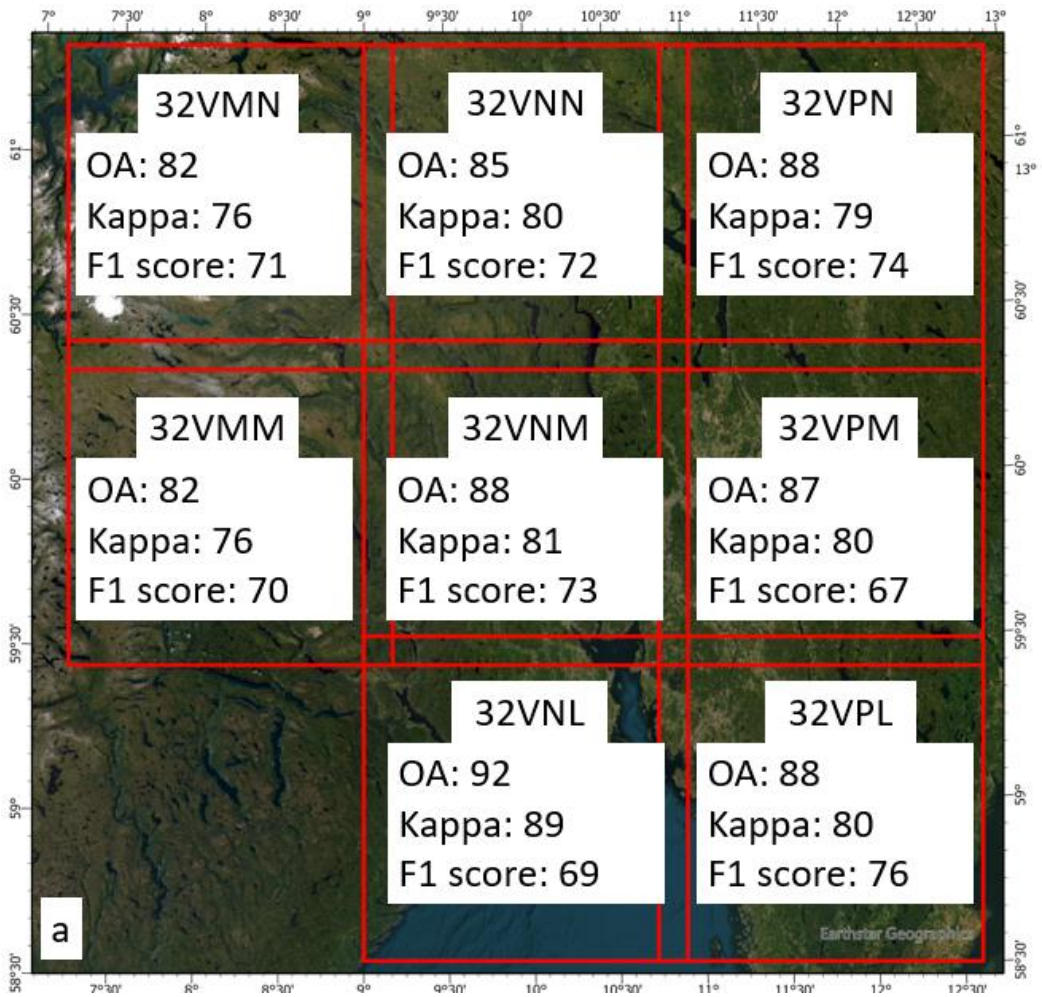
b, f, j – Verified stratified reference samples (step 1b)

c, g, k - AR50

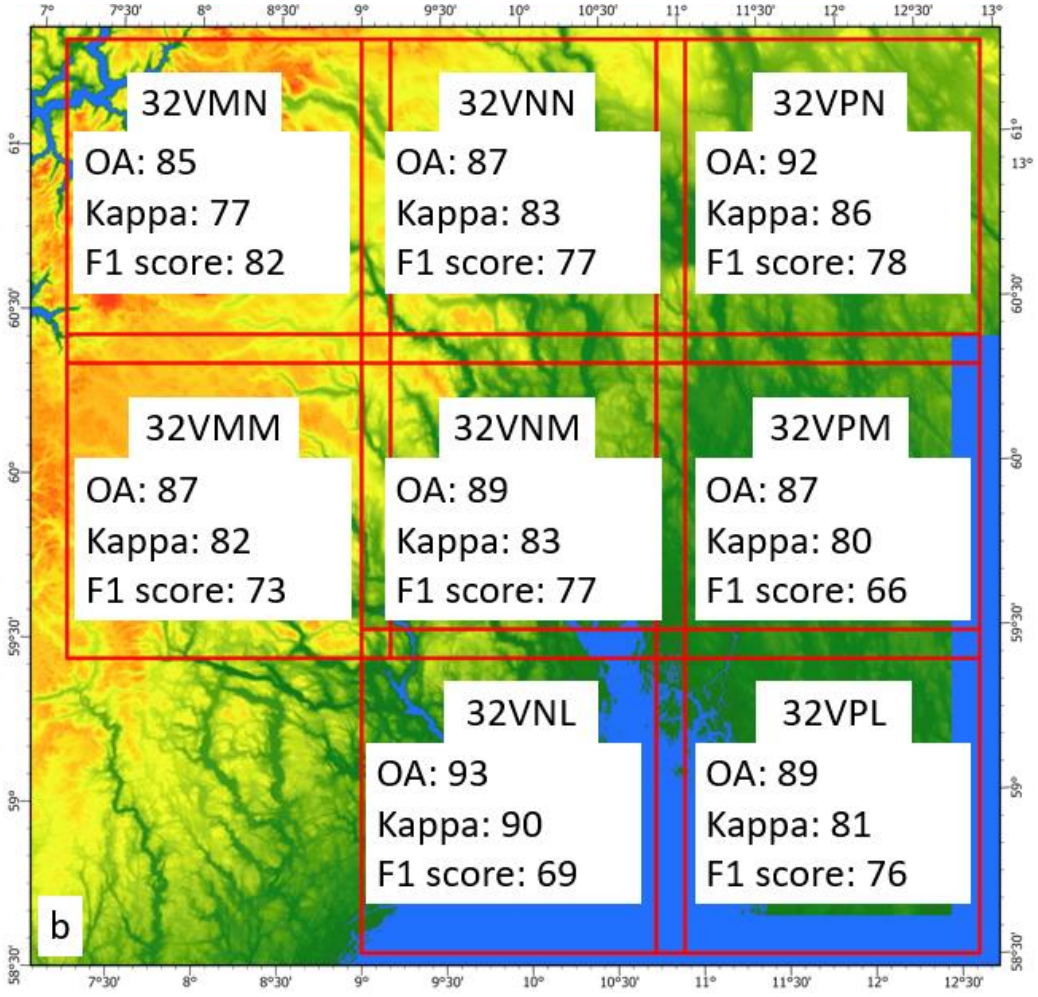
d, h, l - Ortophoto © Norge digitalt

Classification accuracy

Sentinel-2



Sentinel-2 + DEM

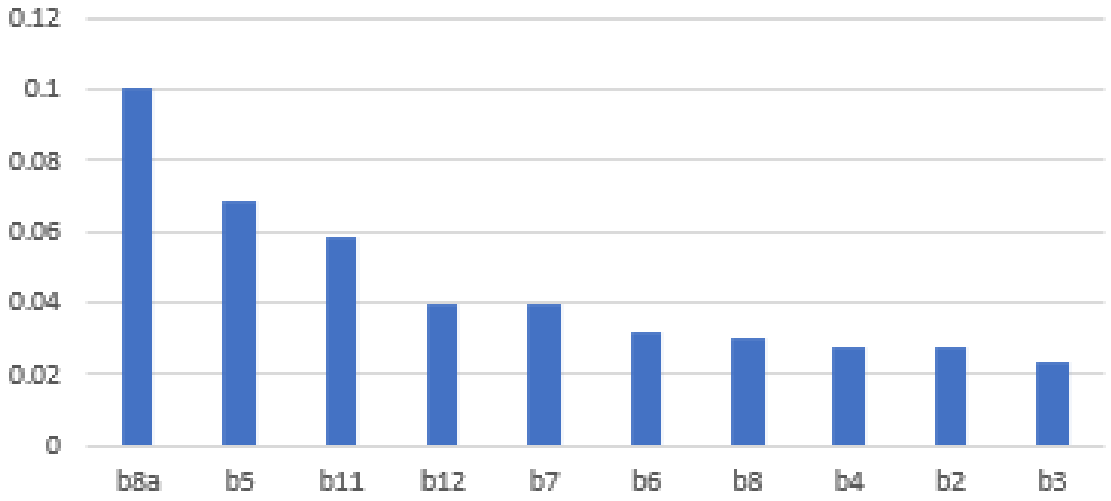


Variable importance

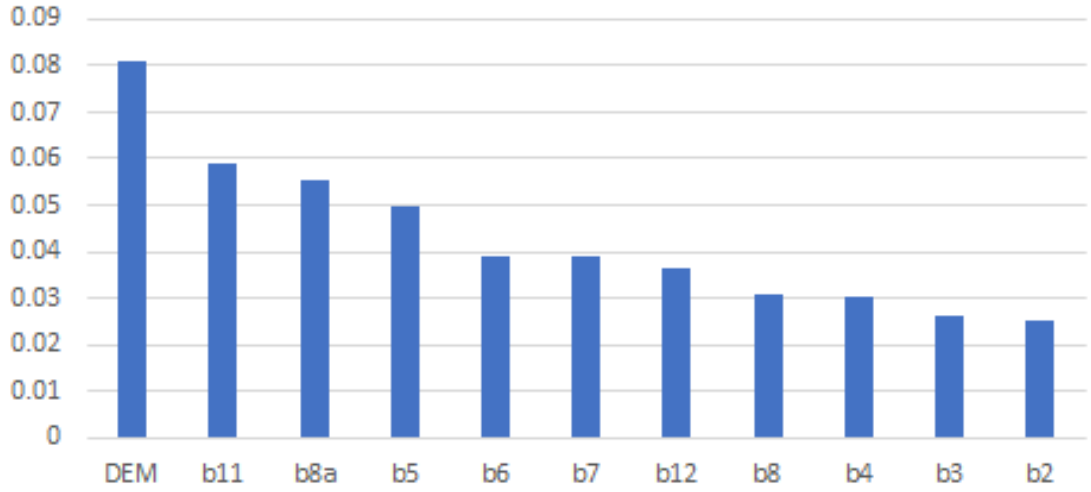
Stability of each classification model varied from 1 to 3 percentage points.

Mean variable importance from all classifications

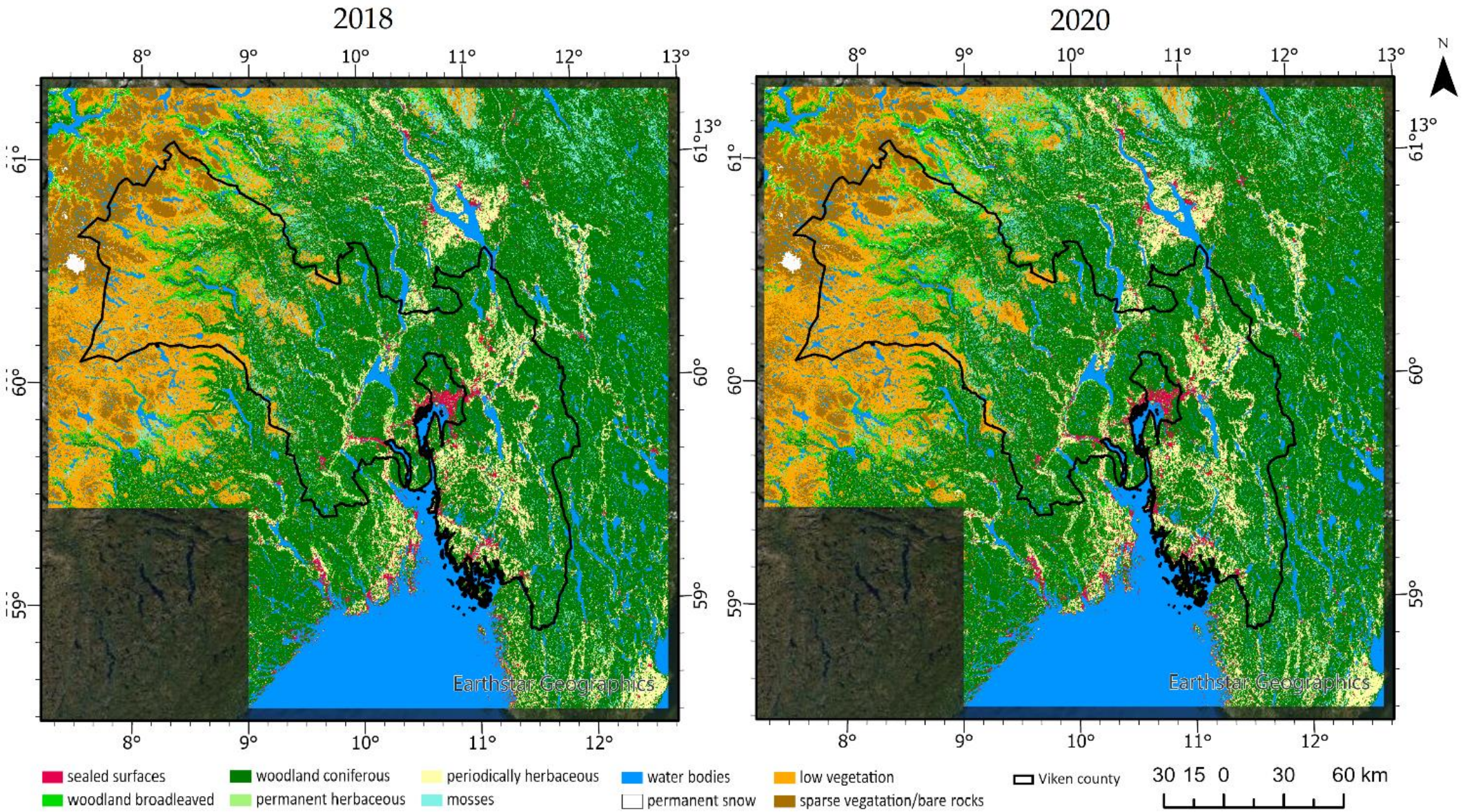
Sentinel-2



Sentinel-2 + DEM



Land cover classification Viken 2018 i 2020



Summary

- The smaller the scale of the reference data, the less accurately the reference points will be located.
- The rate of generalization of reference data affects the accuracy of classification.
- It is important to know the land characteristics and land cover class definitions in the reference databases.
- During Sentinel-2 data selection the phenology sholud be taken into consideration.
- Heterogeneous classes are more difficult to classify.
- Eliminating mislocated points increases the overall accuracy of the classification.
- DEM increases the accuracy of the classification and is the most informative variable used in classification.
- By adding DEM the accuracy for classes located at higher altitudes increased from 4 to 8 percentage points.

The impact of selection of reference samples and DEM on the accuracy of land cover classification based on Sentinel-2 data

Adam Waśniewski[1], Agata Hościło[1], Linda Aune-Lundberg[2]

[1] Institute of Geodesy and Cartography, Centre of Applied Geomatics, Poland
[2] Norwegian Institute of Bioeconomy Research, Division of Survey and Statistics, Norway

Corresponding author: adam.wasniewski@igik.edu.pl

Remote Sensing Applications: Society and Environment

Thank you

contact: adam.wasniewski@igik.edu.pl

The research leading to these results has received funding from the Norway Grants 2014-2021 via the Polish National Center for Research and Development [grant no: NOR/POLNOR/InCoNaDa/0050/2019-00]

More information: https://inconada.eu/

Enhancing the user uptake of Land Cover / Land Use information derived from the integration of Copernicus services and national databases (InCoNaDa)

NEWS ABOUT INCONADA TASKS CONSORTIUM PUBLICATIONS CONTACT