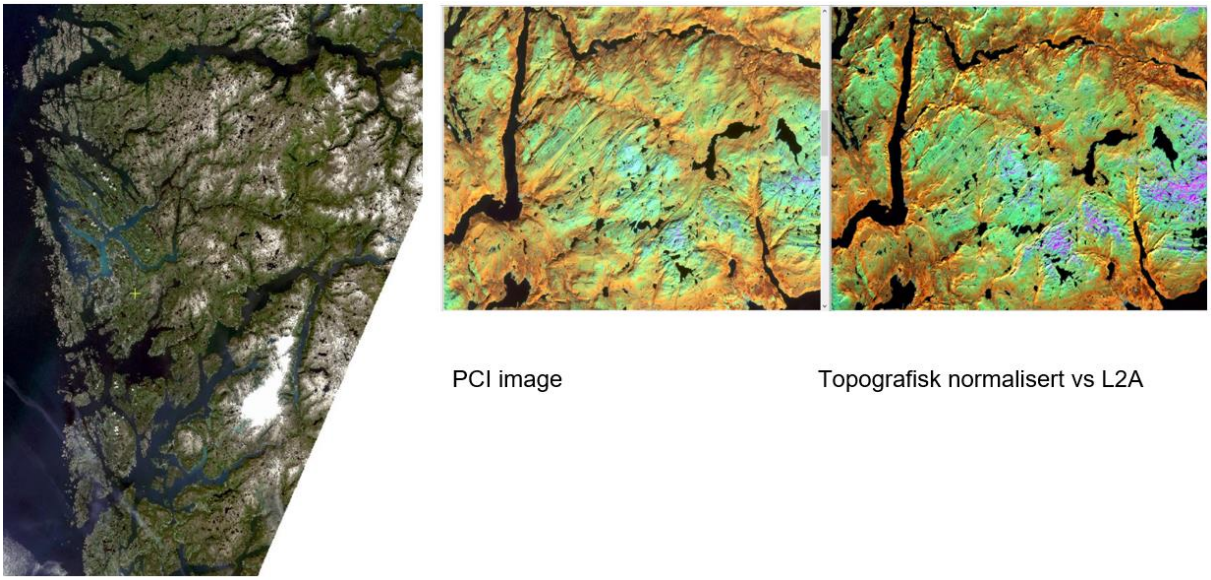
Threshold analysis combined with Deep Learning / Convolutional Neural Network DL/CNN

This document shows how to classify large areas in eCognition with a one button ruleset.

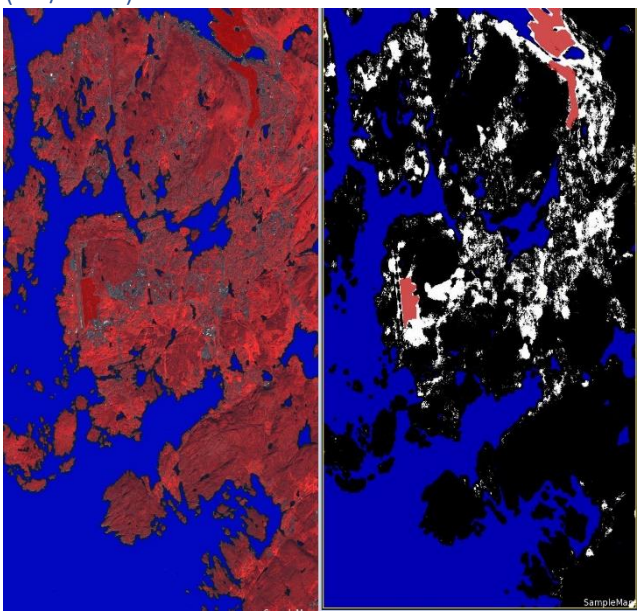


PCI image

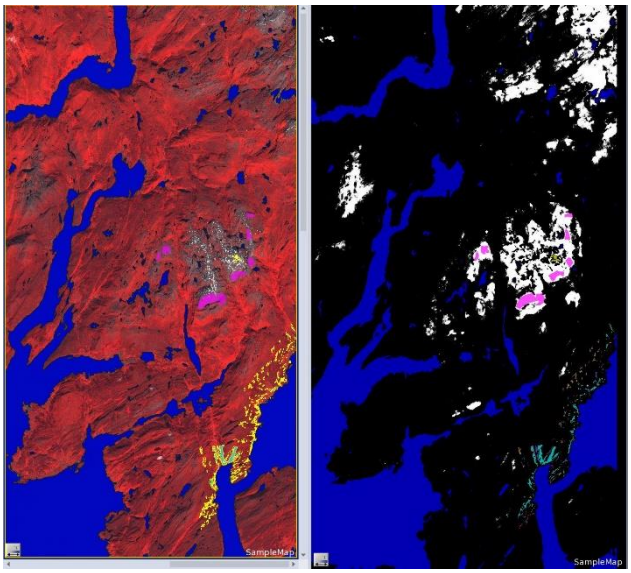
Topografisk normalisert vs L2A

In 2019 a consortium of University in Bergen, TerraNor and PCI Geomatics did a project for Norwegian Environmental Agency. The project called Falk, did an analysis of nature with use of satellite imagery, lidar terrain data and different other data sources. Trimble eCognition team helped with parts of the project.

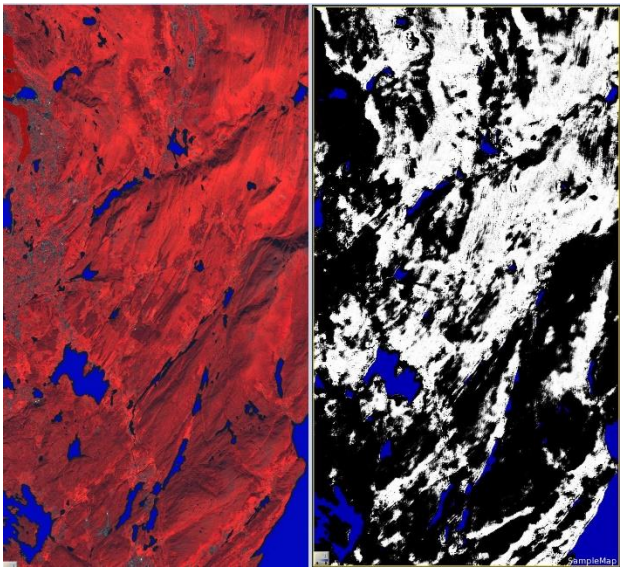
Threshold analysis combined with Deep Learning / Convolutional Neural Network (DL/CNN).



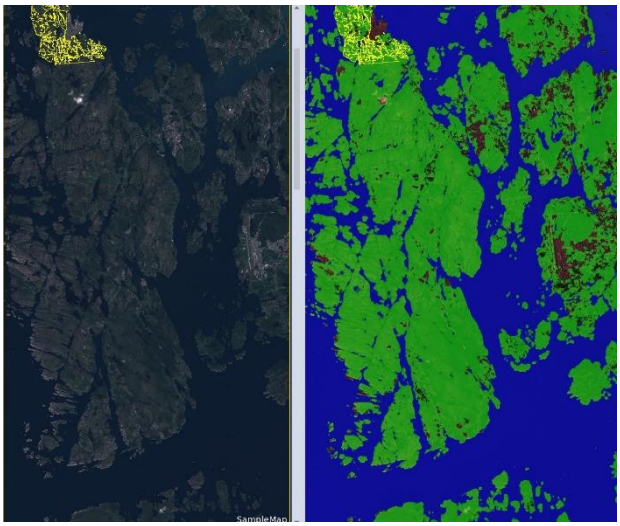
Urban areas:
We picked samples of urban areas from Sentinel 2 image. Image to the left shows Bergen with area around. The red colour on the left is due to Infrared colour in the red band on the computer screen. If you use Near infrared colour in the red band, the red colour in the screen becomes very strong in all areas with vegetation. All non-red areas are non-vegetation or water.
We used **DL/CNN to classify water and urban** areas. To the right urban are white and water is blue.
DL/CNN will accept trees, parks and other vegetation in urban areas as urban. This is a huge benefit with DL/CNN method in eCognition. The result was surprisingly good even with few training samples.



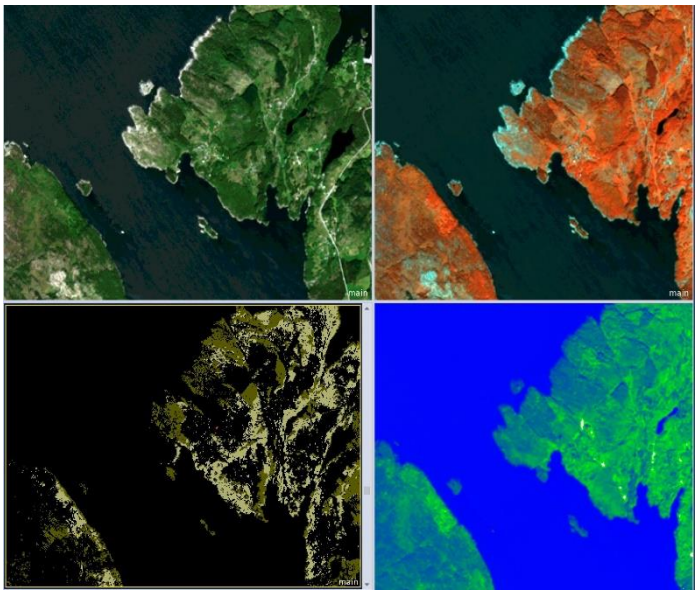
Bare mountain was found the same way as urban and water. The problem is that bare mountain and urban areas are similar. We used dtm height above sea level to distinguish the two. With more samples it is possible to get a higher score.



Samples of **calcareous elm, linden, hazel** was picked from NEA nature database and used to train the model. Due to few samples we ended up with **a map of deciduous trees**. NEA database has single points for each tree listed. The Sentinel 2 image with 10m resolution picked up several trees for the same spot. Sentinel 2 data is very good to find forest stand, but is not so good to find single trees.
For single trees we suggest to purchase Maxar data from TerraNor.



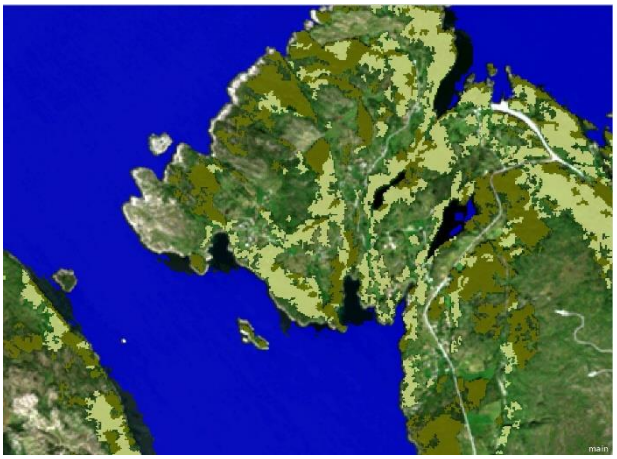
Coastal heath was not possible to capture. In NEA nature database, all areas that contained coastal heath or had been coastal heath, was classified as coastal heath. It turned out that we had polygons with forest, grass, bare rock and bogs classified as coastal heath. When we used these polygons to train eCognition, the classification showed that the whole of Hordaland was coastal heath.
This shows the importance of good training samples.



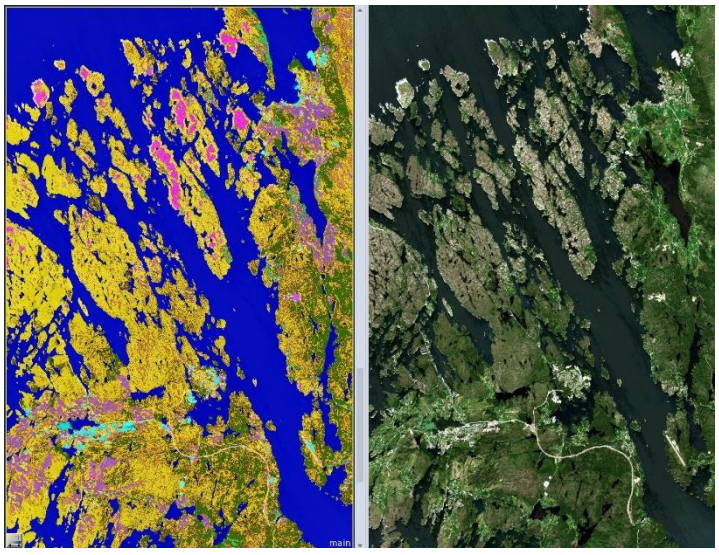
Threshold analysis is fast and can give excellent results in short time.
Sentinel 2 has 13 channels of high quality. We used PCI GXL MRA pan sharpening techniques to create 10 m resolution of all channels. This is described in a separate document.
First, we used water vapour channel and DL/CNN to map water. Then we used NDVI (NIR-RED)/(NIR+RED) to classify vegetation and non-vegetation. With use of lidar data we used vegetation height to distinguish forest from non-forest. In last step we used NIR2 to distinguish deciduous from coniferous.
This whole process is fully automatic in one eCognition ruleset.

Top left: original image.
Top right: use of near infrared in the computer red channel.
Bottom left: Deciduous light green, coniferous dark green
Bottom right: vegetation and water

Issues in classification



This image shows water, deciduous and coniferous classified with threshold values. The forests are very accurate. Water classification has several issues. We fixed it with DL/CNN, see example below.

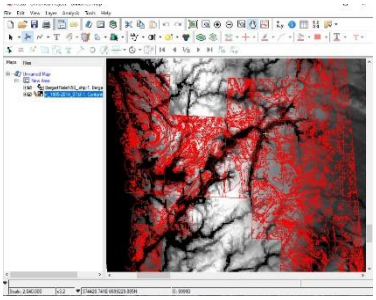


Here we used the first threshold classification of water to train DL/CNN and reclassify water class with DL/CNN. Result is far better than we managed with threshold values only.

Bare mountains, urban areas, forests and water have good results. More test surfaces will give even better results.

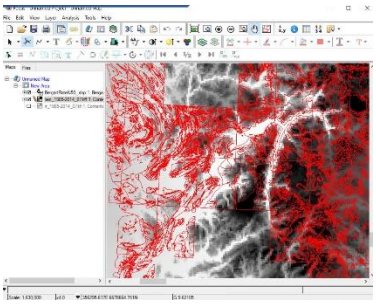
Geology, climate, dtm derivatives: derived data analyses

Most countries have good datasets for geology, weather and terrain. Met institutes have studied weather and climate for hundreds of years and have good digital datasets for wind, rainfall and temperature. Geology maps showing quaternary geology are important. Terrain model based on modern lidar data show both terrain and surface height.
We did not have time to use these datasets in the FALK project. Here we explain how they can be used.

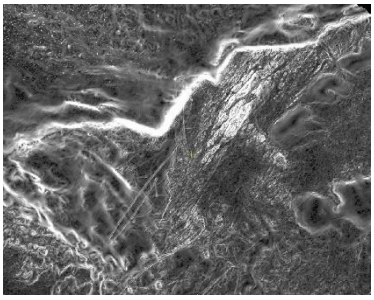


To the left we see rainfall and geology. Higher rainfall is brighter. You can see how the west coast and mountains have more rain than Eastern Norway.

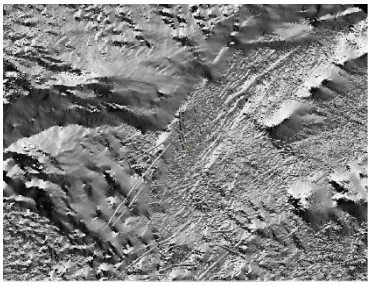
Quaternary geology maps combined with rainfall can tell you where to expect landslides.



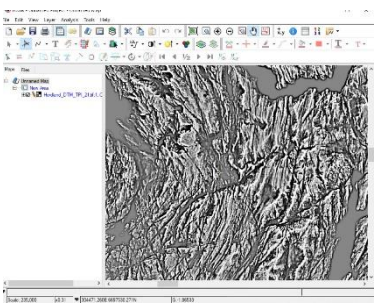
Temperature over the year with bright colour shows higher temperature. This information combined with rainfall and geology can give a very good indication of what species you can expect to find in different areas.



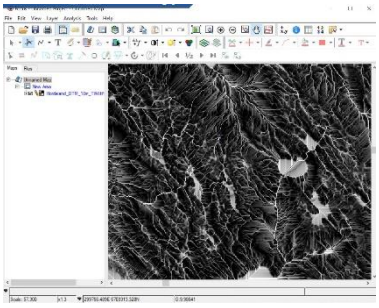
DTM gives us terrain height and surface height. This image shows slope.
Slope and quaternary geology maps combined with rainfall can tell you where to expect landslides.



Aspect shows how much sun light areas receive over the day and over the year.



TPI: topographic position index (check Wikipedia) is a way to describe terrain. Hill tops receive high value above zero while valleys receive low values below zero.



TPW: topographic wetness index shows how water is moving in the terrain.

All these derivatives can be combined to analyse vegetation. Any biologist can tell you how, or you can use machine learning to do the job for you. **If you have good samples of the classes you want to map, eCognition machine learning will help you use the features that distinguishes the classes**.
As you saw with coastal heath the samples must have good quality.