*Enzo Zerega*

# Tree species identification in forest boundaries and tree volume estimation using an advanced prediction algorithm and satellite images*

In forestry, a crucial aspect is the characterization of forest stands. Here, it is important the identification of the tree species and to correctly define the stand boundaries. The most common procedure used for this is land cover classification using satellite images. However, when classifying forests, misclassification of stand boundaries can be substantial, yet classification studies often exclude these areas from the accuracy assessment. With the launch of the Sentinel-2 mission, which provides medium-high spatial resolution satellite images, global coverage and high revisit times, new methodologies for defining and estimating stand parameters are being developed. In combination with advanced prediction algorithms, such as the so-called machine learning algorithms, the use of Sentinel-2 data to predict forest variables has demonstrated to generate highly accurate results. Random Forest is one of these algorithms and has become increasingly popular in environmental studies during the last decade.

This study presents a methodology based on a combination of field data and Sentinel-2 images that were analyzed using the Random Forest algorithm in southern Sweden. The aim was to perform a land cover classification to identify forest patches of three tree species (Scots pine, Norway spruce and birch), testing different approaches to train the RF algorithm, with the focus on improving the results at the forest bordering areas. For this, a segmented accuracy analysis was proposed, where the accuracy was assessed for interior, intermediate and edge areas, as well as for entire forest patches. The RF algorithm was also used to estimate growing stock volume, a standardized parameter that quantifies tree volume. The results indicate that accuracies at bordering areas can be improved when using some of the proposed approaches to train the algorithm. The growing stock volume estimation yielded inferior results but was able to distinguish the satellite bands most correlated with tree volume. The present study contributes to a better characterization of forest stands and, consequently, to facilitate the generation of forest data required by environmental scientists and the forestry sector.

Keywords: Physical Geography, Ecosystem Analysis, Random Forest, remote sensing, accuracy assessment, land cover classification, training stage, multi-temporal, edge areas

Supervisor: **Lars Eklundh**
Co-Supervisor: **Abdulhakim M. Abdi**
Master degree project 30 credits in *Geomatics*, 2018
Department of Physical Geography and Ecosystem Science, Lund University. Student thesis series INES nr 458

*original title: Assessing edge pixel classification and growing stock volume estimation in forest stands using a machine learning algorithm and Sentinel-2 data