UNIVERSITY OF TARTU
Institute of Computer Science
Computer Science Curriculum

Kerstin Äkke

Snow cover detection in Estonia from SAR images using machine learning methods

Master's Thesis (30 ECTS)

Supervisor: Viacheslav Komisarenko, MSc
Supervisor: Anti Gruno, MSc

Tartu 2020

Snow cover detection in Estonia from SAR images using machine learning methods

Abstract:

Usability of optical satellite data for monitoring snow cover can be limited in regions with frequent high cloud coverage. Synthetic aperture radar (SAR) could theoretically be used to monitor snow regardless of clouds or lack of illumination. There are several factors that complicate the task in Estonia such as dense vegetation and quickly changing snow conditions. So far most studies on using SAR for snow detection have been done in mountainous regions and over short time periods. The aim of this study was to test applicability of a method that combines most common features for snow detection extracted from SAR images in a machine learning model. This method had shown good transferability in mountain regions, however the modelling results on Estonian data were unsatisfactory. Analysis of features derived from SAR images revealed poor separability of snow free and snow covered classes. This suggest the main issue is with the feature extraction methods rather than machine learning. Possibly the processing chain could be optimized for Estonia and other regions with flat topography and predominantly dense vegetation. This thesis did not result in a usable model, but could serve as a basis for further studies.

Keywords:
synthetic aperture radar, backscattering, InSAR, PolSAR, snow classification, random forest

CERCS:P170 - Computer science, numerical analysis, systems, control

Lume tuvastamine Eestis tehisavardari piltidelt masinõppe meetoditega

Lühikokkuvõte:

Optiliste stelliitide kasutamine lumikatte jälgimiseks võib olla piiratud sagedase pilvkattega aladel. Tehisavaradar (i.k. *Synthetic aperture radar*) sobib teoorias tuvastamiseks sõltumata pilvkattest või päikesevalguse puudumisest. Eestis teeb kaugseirega lume tuvastamist keerulisemaks taimkate ja kiirelt muutuvad lumeolud. Praeguse seisuga on suurem osa uuringutest, mis kasutavad tehisavaradarit, tehtud mägistes piirkondades. Käesoleva töö eesmärk oli testida meetodit, mis ühendas kõige tavalisemad tehisavaradarist tuletatud tunnused masinõppemudeliga, Eesti oludes. See meetod on näidanud head ülekantavust mägistes regioonides, kuid sellega lume modelleerimise tulemused Eesti kontekstis olid mitterahuldavad. Tunnuste analüüs näitas et lumikatte ja lume vabad klass ei olnud tehisavaradarist tuletatud tunnuste põhjal hästi eraldatavad. Sellega näidati, et probleem on pigem tunnuste eraldamise kui masinõppe osas. Võimalik, tehisavaradari piltide töötlemist saab optimiseerda Eesti ja teiste lameda topograafia ing tihke taimestikuga alade jaoks. Käesoleva uurimistöö tulemusena ei saadud usaldusväärset mudelit, kuid töö sobib aluseks edasistele uuringutele.

Võtmesõnad:
tehisavaradar, tagasihajumine, InSAR, PolSAR, lume klassifikatsioon, juhuslik mets

CERCS:P170 - Arvutiteadus, arvutusmeetodid, süsteemid, juhtimine (automaatjuhtimisteooria)

Contents

1 Synthetic Aperture Radar **8**

1.1 General SAR properties 8

1.2 Sentinel-1 9

2 Study Area and Duration **11**

3 Data Sources **13**

3.1 Digital elevation model 13

3.2 Global SnowPack Snow Cover 14

3.3 Optical snow cover 15

3.4 Land Cover 15

3.5 Weather Data 17

4 Processing SAR images **19**

4.1 Backscatter ratio 20

4.1.1 Processing steps 21

4.1.2 Processing Output 22

4.2 Interferometric SAR 23

4.2.1 Processing steps 24

4.2.2 Processing output 25

4.3 Polarimetric SAR 26

4.3.1 Processing steps 27

4.3.2 Processing output 28

5 Experiments **29**

5.1 Unbalanced Classes 30

5.2 Balanced Classes 31

5.3 Mixed Seasons and Balanced Classes 32

6 External validation **36**

6.1 Sentinel-2 and Fmask 36

6.2 Weather station measurements 37

7 Conclusion **40**

References **43**

Appendix **44**

II. Polarimetric decomposition results 44

II. Licence 46

Introduction

Snow cover has an effect on climate change, ecosystems as well on human activities such as transport, tourism and infrastructure maintenance [1]. There have been several studies on the possibility of using Synthetic aperture radar (SAR) images to detect snow ([2], [3], [4], [5]).

There are three main approaches to obtain information about snow cover from SAR images: backscatter, interferometry (InSAR) and polarimetry (PolsSAR) based approach [2]. A recent study by Tsai et al. [3] that combines all three methods as well as land cover and topographical information was chosen as a basis for the current thesis. The main reason for choosing this specific study was its generalized approach.

The main advantages of using SAR images are that unlike optical images they can be used regardless of cloud cover or darkness. This study focuses on detecting snow in Estonia where the average annual total cloud cover is 7 points with the most coverage during the winter season [6].

Most previous studies about using SAR for snow detection, including the article by Tsai et al. focused on mountain regions [2]. The main goal of this thesis is to test the effectiveness of the method from the base article in Estonian circumstances where topography is relatively flat. Furthermore, land cover in Estonia is mostly dense vegetation on which most SAR snow detection methods are less effective [2].

The first chapter introduces the working principles of SAR Systems. It explains how SAR images are created and what are their general properties. The chapter also describes specific radar properties of Sentinel-1 mission that was the source of SAR images for this thesis. Finally the chapter explains the structure and attributes of the Sentinel-1 products that were used in this study.

In the second chapter the study area and duration are established. The chapter explains

limitations in choosing dates caused by the combination of data unavailability and chosen methods.

The third chapter lists all data sources used for constructing, training and validating the models. For each data source except SAR images there is a subsection detailing the origin of the data set. The subsections also contained details of used tools and methods for any dataset that needed further processing. Data from weather station and land cover was described in respective chapters to understand characteristic of chosen study area period.

The fourth chapter is about processing SAR images. The chapter describes the tools used for processing SAR images and processing steps that all backscatter, InSAR, and PolSAR methods have in common. The chapter also contains a subsection for each approach. Subsections describe the theory behind each method followed by implementation of processing pipeline. Finally the extracted features are analyzed to estimate snow free and snow covered ground separability based on single features.

The fifth chapter focuses on the machine learning experiments. First general approach to model construction, model parameters as well as testing results are described. There are subsections about individual models explaining specifics of each experiment set up. The subsections also contain detailed comparisons between training and testing metrics per land class. The experiment results are analyzed to find ways for improving model performance.

The final chapter validates machine learning experiment results on external data sets. There are two sources of validation data: snow cover derived from optical satellite and weather station snow depth measurements. Finally the models are compared to each other based on validation results.

1 Synthetic Aperture Radar

Synthetic Aperture Radar is an active sensing method for acquiring high resolution images with a Side-Looking Airborne Radar (SLAR) system. SLAR systems consist of radar mounted on a moving airborne platform such as aircraft or satellite. Observation geometrt of SLAR with relevant terms is shown in Figure 1. SLAR collects observations by transmitting pulses of microwave frequencies toward ground away from nadir at a look angle Θ_l (/Figure 1) and recording reflections (backscatter) of transmitted signal. [7]

1.1 General SAR properties

The spatial resolution of SLAR along track (azimuth) direction is mainly determined by antenna length of the radar and carrier altitude. Resolution deteriorates at higher altitudes and increases with antenna length. For spaceborne SLAR the length of antenna needed to get images with useful azimuth resolution can be hundreds of meters. Deploying hardware with such specifications is unfeasible. The SAR method proposed by Carl Wiley [8] solves this by simulating a virtual antenna (synthetic aperture). Modern SAR systems usually have azimuth resolution between 0.5-20m. [7]

Ground range resolution depends on range resolution and Local Incidence Angle (LIA) Θ_i (Figure 1). Dependence on LIA means the ability to distinguish between ground objects is different throughout the swath. The sensing geometry of SAR can make direct visual comparison of SAR images with other data sources difficult. Due to this SAR images are often converted from slant-range geometry to a cartographic projection with terrain-correction or geocoding as the last step of processing. [7]

SAR sensors record both amplitude and phase of returned signal for each resolution cell. There are usually multiple scatterers within the resolution cell and interaction between

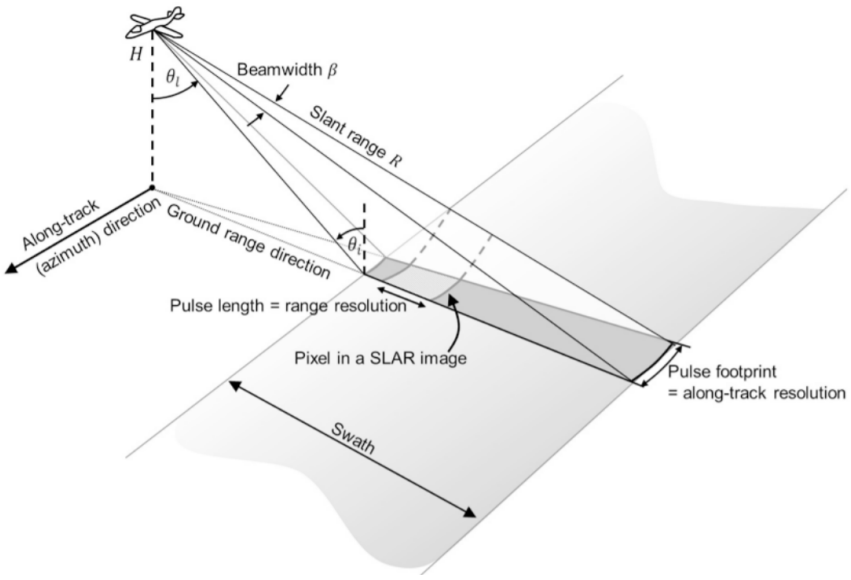

Figure 1. Side-Looking Airborne Radar [7]

those scatterers can cause measurements between resolution cells to vary greatly even on the same surface (speckle). Speckle interferes with determining actual surface properties and several methods have been developed to reduce speckle. The sensed object properties with most effect on reflected signal are surface roughness and dielectric properties. The exact effect of these depends on the sensor wavelength as well as sensor polarization. [7]

Active sensing methods can control the polarization of emitted and received signals. The signal polarization is usually either vertical (V) or horizontal (H). Comparing the differences between images sensed with each combination can give information about land cover. [7] One example of that is the PolSAR method described in Chapter 4.3.

1.2 Sentinel-1

All SAR images used in this thesis were acquired by the European Space Agency (ESA) Sentinel-1 mission. The mission constellation consists of two satellites: Sentinel-1A

and Sentinel-1B. The satellites share the same orbital plane. A single satellite makes a repeat pass every 12 days. The mission began in 2014 with the launch of Sentinel-1A. Sentinel-1B became operational in 2016 reducing the exact repeat cycle to 6 days. [9]

The main acquisition mode of Sentinel-1 is Interferometric Wide swath (IW). It acquires data on C-band (5cm) wavelength by 250-km swaths. Sentinel-1 is offers dual polarized data. The offered polarisation combinations are vertical transmit and vertical receive (VV) or vertical transmit and horizontal receive (VH)[10].

The IW products of Sentinel-1 are sensed using Terrain Observation with Progressive Scanning SAR (TOPSAR) sensing technique. The aim of TOPSAR is to get homogenous image quality throughout the swath. It is achieved by sensing in bursts along azimuth direction and switching between sub-swaths in several bursts. [7] The bursts are synchronized between passes to guarantee alignment of sensing pairs [10]. This is necessary for the inSAR method described in Chapter 4.2.

The Sentinel-1 IW Single Look Complex (SLC) products contain complex (amplitude and phase) pixel values in slant-range observation geometry for each sub-swath/polarization pair [7]. For Sentinel-1 there are three sub-swaths and two polarization modes, thus a total of six images. Each sub-swath image contains a number of bursts separated by black-fill demarcation. During post-processing the bursts may be joined into a continuous image (deburst). [10] In the current paper SLC products are used for InSAR (Chapter 4.2) and PolSAR (Chapter 4.3) processing.

The IW SLC product may be processed further into a Ground Range Detected (GRD) product. To produce GRD from SLC the product is detected, multi-looked and projected to ground range. Furthermore, the bursts and swaths of each polarization are merged into a single continuous image. GRD products no longer have phase information. As the result of multilooking, resolution is smaller compared to SLC, but speckle is also reduced. [9] The GRD images are used for calculating backscatter ratio in Chapter 4.1.

2 Study Area and Duration

The scope of this study was limited to reduce preprocessing and model construction complexity. The area of interest (AOI) shown in Figure 2 has total ground surface of approximately $10000km^2$. Detecting snow on water bodies was not of included in the scope of current thesis. The location of AOI was chosen such that the preprocessing workflow could be extended over the whole country by processing sections of similar size. The usual Estonian snow period lasts from December to the beginning of April [6]. Thus a season in the context of this thesis is considered to last from 1 November to 30 April. Seasons 2016-2017 and 2015-2016 were chosen for this study because these are the most recent seasons with both SLC and GRD images available on ESTHub.

Figure 2. The study area

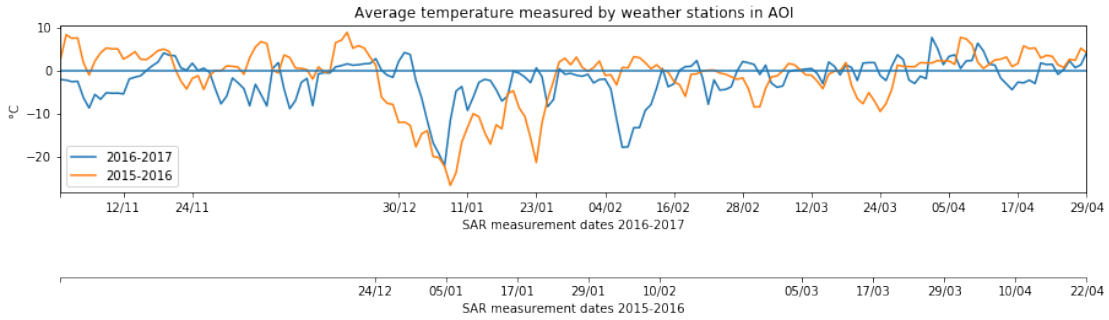

Figure 3. SAR acquisition dates and daily average temperature measurements

The dates of SAR measurements included in the study are displayed on the x-axis of Figure 3 with corresponding average temperatures on the y-axis. There were some dates in both chosen seasons where SAR images were missing or incomplete. Temporal requirements of Insar caused images from consequent pass to also be excluded. Most

notably there were no images included from before the second half of December in season 2015-2016 (Figure 3). Furthermore, the SAR image used as reference for backscatter (2016.11.12) could not be included in training.

The Backscattering method requires a pair of images sensed with the same geometry [2]. That means both images must be taken from the same relative orbit. Images taken by Sentinel-1A on relative orbit 7 were used. Images acquired by Sentinel-1B were not used because it became operational in April 2016, meaning it could only be used for one of the chosen seasons. Including 1B data in parallel to 1A without reducing time between InSAR pairs could be considered in the future.

3 Data Sources

There were several data sources used in thesis thesis. Sentinel-2 optical satellite data and weather station measurements were used for validation and seasonal analysis. Four data sources were used to train the models:

- SAR images
- digital elevation model
- land cover
- snow cover

DEM and SAR are used to derive the model features. Land cover classes are used to train a separate Random Forest (RF) for each land class. Random Forest training belongs to the family of supervised learning models so an existing snow cover data set Global SnowPack is used as ground truth.

All of the data processing, other than the processing of SAR images described in chapter 4, was done with Python 3.6 and Jupyter notebook. Land cover and DEM datasets were geographically aligned and resampled to target resolution of 100m. The target resolution was chosen to be comparable with the study by Tsai et al. [3] who chose this value because it is necessary for snow cover mapping in mountain regions. This combined product of DEM and land cover was used as a reference to which to align other data sources.

3.1 Digital elevation model

While Estonia does not have considerable variations in height unlike mountain regions, the effect of surface geometry on snow gathering should still be taken into account. There

are four model parameters that are calculated from the DEM: elevation, aspect, slope and curvature. While elevation can be read directly, the other parameters were calculated with RichDEM python library [11].

National Aeronautics and Space Administration (NASA)'s Shuttle Radar Topographic Mission (SRTM) provides freely accessible digital elevation models with 80% global coverage. For this thesis, the processed SRTM data version 4.1 with 90m spatial resolution and less than 16m vertical error was used. The processed version was preferred over the original SRTM data because regions of no data have been filled using interpolation methods. [12]

3.2 Global SnowPack Snow Cover

Global SnowPack (GSP) [13] daily snow cover extent (SCE) product by The German Aerospace Center (DLR) was used as ground truth for training the model. The GSP dataset is derived from Moderate Resolution Imaging Spectroradiometer (MODIS) Aqua and Terra daily products.

GSP has filled the gaps in combined MODIS data caused by cloud coverage with three interpolation steps:

1. DEM referenced snowline interpolation
2. three-day temporal interpolation
3. seasonal interpolation over the whole hydrological year

The accuracy of GSP data sets $(77\% - 84\%)$ is lower than MODIS daily products $(\ 93\%)$ [13]. However, the GSP dataset preserves information about which method was used to derive values of each pixel so it is possible to exclude pixels from any interpolation step to reduce noise.

3.3 Optical snow cover

One of the validation sources used by Tsai et al. [3] was optical-SCE from Landsat and Sentinel-2 data. Landsat was omitted from the current thesis as it is not available on ESTHub. Because the testing dataset consisted primarily of 2015-2016 data for all experiments, Sentinel-2 images were also chosen from that season.

Several images were discarded due to severe cloud contamination. All Sentinel-2 images overlapping the AOI were processed with python-fmask library version 0.4.5 [14]. The library implements Fmask4, an algorithm for differentiating between clouds, cloud shadows, snow and water bodies [15].

Sentinel-2 acquires images in large tiles consisting of multiple sub-images. The processed images taken on the same date were combined using SNAP and subset to the extent of the study region. Even though this resulted in only partial overlap with the study area, six dates with enough cloud free pixels were selected and paired with the closest SAR observation date. The results of comparing FMASK and predictions of model trained in this paper are described in Chapter 6.1.

3.4 Land Cover

The European Space Agency releases annual land cover (LC) maps with 300m resolution. For years 1992-2015 these maps have been released under ESA Climate Change Initiative (CCI). The Global Land Cover products for 2016, 2017 and 2018 have been released under Copernicus Climate Change Service (C3S) and are consistent with the CCI 1992 – 2015 map series. The products have 22 land cover classes, some of which have been further divided by vegetation density. The classes were remapped to more general land classes the same way as Tsai et al.[3] did to reduce complexity.

All land classes represented in the AOI are listed in Figure 4. The proportion of each

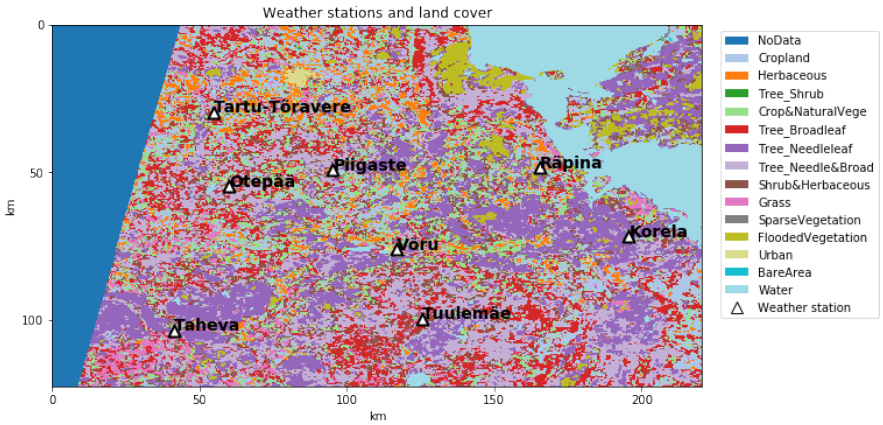

Figure 4. Map of the study area with land classes and Weather stations. NoData also includes area outside of the AOI

land class in Estonia and AOI is shown in Table 1. The percentages of classes were calculated over bounding box containing Estonia not within actual country borders. The *Water* class was excluded as snow detection was limited to ground on this thesis.

The classes are divided into dense and sparse vegetation. Dense vegetation is further divided into forests as forest classes had significantly lower accuracy in baseline model by Tsai et al. [3]. As seen in Table 1 majority of Estonia is covered by forests. Sparse vegetation accounts for only about 17%.

In this thesis the land cover for year 2016 was used. It is temporally close enough to both training and testing period that no major changes in land cover are expected. Land classes with less than 0.2% representation in AOI were excluded from all experiments due to too few samples. Land class with the lowest representation included in an experiment was *Urban* 6.2.

Table 1. Comparison of land cover class percentage in Estonia, representation in AOI and whther vegetation type is dense (D), sparse (S) or forets (F)

LCtype	%in EST	%in AOI	%Type
$Cropland$	12.43%	11.38%	S
$Herbaceous$	3.93%	5.49%	D
$Tree\_Shrub$	0.01%	0.01%	F
$Crop\&NaturalVege$	8.40%	10.73%	D
$Tree\_Broadleaf$	17.65%	13.73%	F
$Tree\_Needleleaf$	20.00%	19.84%	F
$Tree\_Needle\&Broad$	22.42%	21.01%	F
$Shrub\&Herbaceous$	6.33%	9.20%	D
$Grass$	4.74%	4.59%	S
$SparseVegetation$	0.09%	0.08%	S
$FloodedVegetation$	3.46%	3.60%	D
$Urban$	0.49%	0.34%	S
$BareArea$	0.04%	0.01%	S

3.5 Weather Data

The weather data was requested from Estonian Weather Service. There are 8 weather stations in the chosen AOI (Figure 4). The weather parameters included in this study were temperature and snow depth.

Temperature data was used for selecting the reference image for backscatter method 4.1. The snow depth data was used to validate the model. The validation results are under Chapter 5.

As seen on Figure 5 the snow period started earlier than usual in 2016. Snow depth measured by most stations in 2016 is similar. There are multiple melting periods throughout the 2016-2017 season and longest snow cover duration is less about a month. Exception to this is Tuulemäe which had lasting snow cover between January and March (Figure 5)

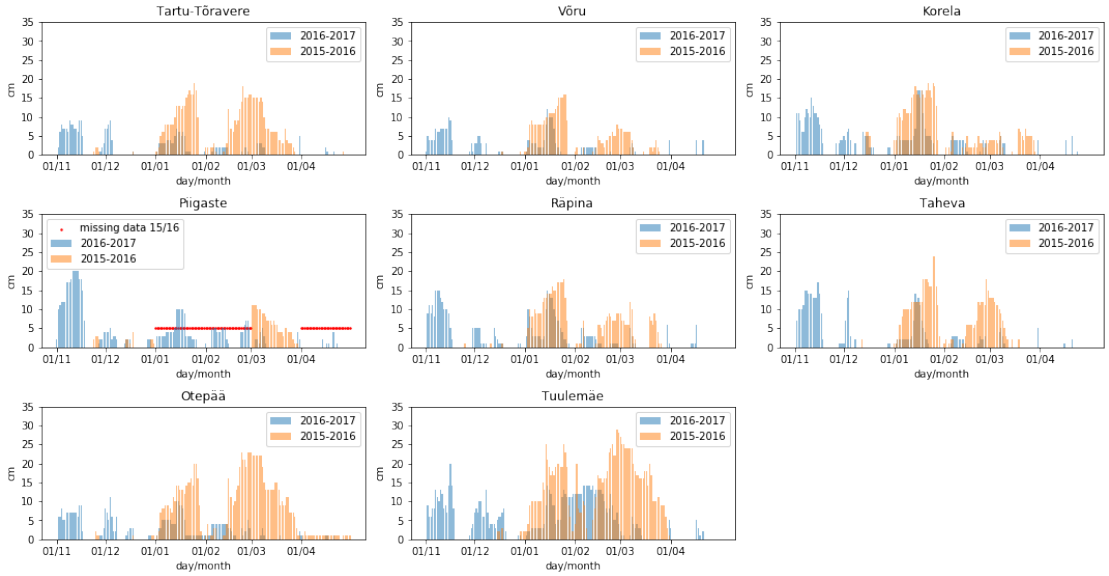

Figure 5. Seasonal daily snow depth comparison for each weather station inside AOI

There are some significant differences between the two observed seasons. The snow cover during the 2015-2016 season lasted from January until the end of March with one melting period in between. The average snow depth during that period was also higher than the average of most snow periods in 2016-2017 in each respective station (Figure 5). There were also differences in temperatures throughout the season with temperatures in 2016-2017 fluctuating around 0 degrees in most months.

Grain size, water content, and snow density influence how SAR beam interacts with snow [2]. These parameters are affected by both temperature and snow cover duration. Differences observed between seasons might make predicting one season based on another more difficult. For this reason, in Chapter 5.3 a model was trained with data from both seasons combined.

4 Processing SAR images

Working with Sentinel-1 images can be difficult on personal computers due to hardware requirements. SLC files are generally over 4GB. The memory required for working with multiple SAR images while also storing intermediate processing steps can easily reach several hundred GB.

National satellite data center (ESTHub) [16] offers an environment for satellite data processing. It stores satellite images covering Estonia and other regions of interest to provide faster download to users as well as the opportunity to process images within the server. ESTHub provides two access modes for their users: virtual machine or graphical user interface. The first option was chosen in this work due to its flexibility.

All the satellite image processing in this thesis was done with sentinel application platform (SNAP) [17]. SNAP is a free open-source software developed by the European Space Agency (ESA). This software offers the possibility for users to define processing graphs. This makes it easy to apply multiple methods in succession for bulk processing.

To share resources between users in ESTHub, processing jobs are scheduled with Calvalus Hadoop Tool (CHT). The specifications for a job using SNAP are given in JSON file with parameters of interest:

- inputPath - an absolute path to file or a list of file paths
- processorName - name of a SNAP operator or user defined processing graph
- processorParameters - a list of graph or operator processing parameters
- regionGeometry - area of interest (optional)

ESTHub processing chain will not process any images listed in the inputPath parameter that do not overlap with given regionGeometry. If the processor requires more than one

input file such as for InSAR, then only one of the files can be specified through inputPath parameter. All other input files must be given as processor parameters.

The chosen study region was covered by two SAR images. All methods include merging the resulting images from the same date into a single file while re-sampling to target resolution with SAR-Mosaic operator. Finally the images were also collocated with the DEM and LC combined raster (3) to make sure they align with each other.

4.1 Backscatter ratio

The water in melting snow significantly reduces the backscatter coefficient. Nagler's method is still the most common method for detecting wet snow from SAR images. [2] It is by Equation 1 [18].

$$\frac{\sigma_{snow}}{\sigma_{reference}} < -3dB \tag{1}$$

The method calculates the ratio between backscatter coefficients of snow-free of dry snow reference image ($\sigma_{reference}$) and target image (σ_{snow}) with the formula shown in Equation 1. Then it applies a fixed threshold to classify pixels as snow covered or snow free. Threshold of -3dB is common for most applications.[2]

The article by Tsai et al. [3] also calculates the backscatter ratio between reference image and snow cover based on Nagler's method. However, instead of applying threshold the ratio was used as one of the inputs to training a random forest. Furthermore, they used terrain flattened backscatter coefficients. Terrain flattening removes radiometric distortions caused by surface topography [7].

The article does not specify whether VV of VH polarization was used. [3] In this thesis both polarizations were included in the model because combining backscattering ratio of different polarizations has been used to improve separability of classes [19].

It is recommended to choose an image with frozen ground covered with dry snow as

reference because soil moisture may affect the result [20]. The image taken on 2016.11.12 was deemed suitable. The ground on that day was estimated to be frozen as the average temperatures were constantly below zero at least since the beginning of November (Figure 3).

4.1.1 Processing steps

The processing chain was divided into two SNAP graphs. The first processing segment shown in Figure 6a takes a single GRD product as input and returns backscatter coefficient. The input image is first calibrated to get the genuine brightness values. Then thermal noise removal, Lee-Sigma speckle filter and Terrain flattening are applied with default parameters. After that terrain correction is done based on SRTM 3Sec digital elevation model. Finally the result is cropped to the extent of the study area. The same processing

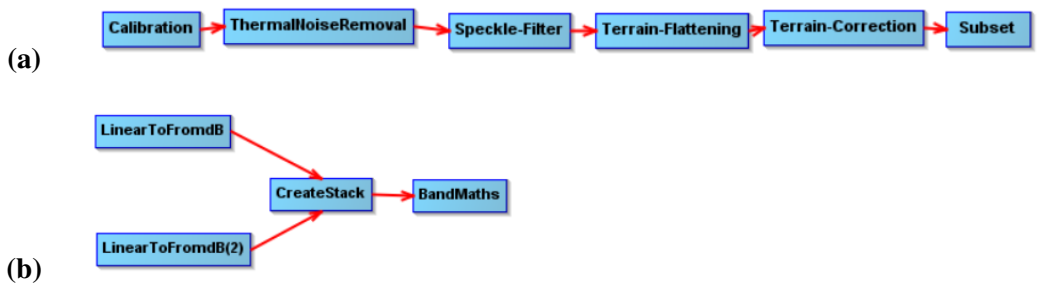

Figure 6. Backscatter coefficient extraction (a) and ratio calculation (b) processing chains

chain is also applied on the reference image. The area chosen for this study was covered by two SAR images resulting in two separate reference images. These references were merged into a single image with SAR-Mosaic operator before the next processing segment for convenience.

Figure 6b shows the processing steps used to calculate the ratio between backscatter

coefficients. Firstly either VV or VH polarization backscatter coefficient band is selected and converted to decibels for both the reference and snow image. The bands are then collocated with the CreateStack operator. Finally in the BandMaths operator the backscatter ratio is calculated with the Equation 1 from Nagler's method [18].

Initially this segment was done with SNAP, however, applying the final SAR processing steps described in Chapter 4 directly after the first segments (Figure 6a) and calculating the ratio with Python proved more convenient. In this case the create stack step is unnecessary as the images are already aligned with each other after collocation with DEM and LC raster. The main benefit of using Python for ratio calculation is that it enables easier switching between different reference images.

4.1.2 Processing Output

As demonstrated in Figure 7, snow-covered and snow-free pixels are not clearly separable by backscatter ratio for any land class. This could be explained by the size of the area and range of analysed dates. The snow-covered pixels Figure 7 include include both wet and dry snow. While wet snow should have considerably lower backscatter than snow-free ground, dry snow is nearly indistinguishable [21].

Refreezing of the surface layer during the melting period can also increase backscatter coefficient [19]. In Estonia, during the melting period, the temperatures dropping overnight near ground is a common occurrence [6]. Koskinen et al. demonstrated that dense vegetation can further reduce contrast between snow-free and snow-covered measurements [21].

A possible solution for increasing the separability of classes would be adding temperature data to model features. A shortcoming of this would be that the temperatures measured by stations can be different from ground temperature [6]. Additionally, interpolation between sparse data from weather stations may be inaccurate.

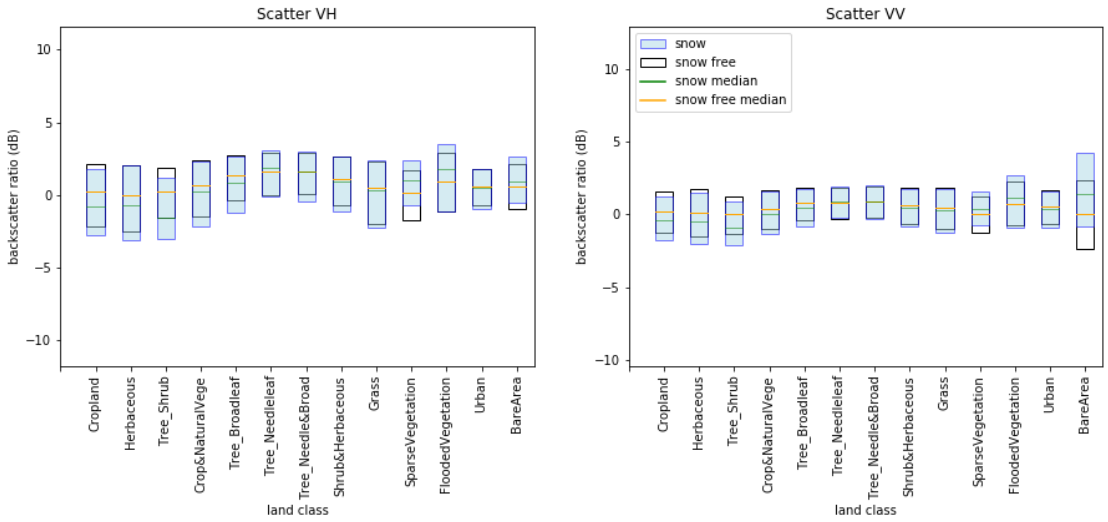

Figure 7. Boxplots showing backscatter ratios (median, first and third quartile) of snow free and snow covered pixels for different polarizations and land classes between 2016.11 - 2017.04

4.2 Interferometric SAR

Interferometry can be used to detect similarity between images acquired on different dates [22] Both wet and dry snow cause differences in backscatter mechanism which lowers the coherence between images. Using InSAR to detect snow requires images taken from the same orbit at consequent passes. This limits temporal variability and supports differentiating changes in coherence caused by snow from those caused by other factors. [4]

If there have been no changes in snow cover between passes this method will not work. [4] However this should not be an issue in Estonian context as Figure 3.5 demonstrates constant changes in snow depth. On the other hand, in densely vegetated or forested areas where coherence is constantly low, changes caused by snow will not cause a significant drop in coherence [2].

4.2.1 Processing steps

The InSAR method uses SLC images that contain both amplitude and phase information. The study area of the current thesis was covered by two sub-swaths, however for future scalability steps for processing all swaths were included in the processing graph.

As a first processing step ApplyOrbitFile operator was called. It is a recommended step for InSAR processing, because the orbit information within the SAR product is inaccurate. This may be done for all swaths simultaneously. [22] Separating this step from the processing graph in Figure 8 was done to filter out images in the selected time period that overlapped the study area.

The slave image's path was given in the processorParameter field that ESTHub does not compare with the area specified in regionGeometry parameter automatically. If the slave does not overlap the master, then the result is an image filled with NaN values. Any such images were removed during the SAR-Mosaic step. Narrowing the amount of potentially suitable slaves beforehand reduced the amount of empty images which made the process faster and reduced the storage space required.

The images collected in the previous step were processed with the graph displayed in Figure 8. Coregistration (Back-Geocoding) and interferogram formation must be done for each swath separately [22]. Before interferogram formation Enhanced-Spectral-Diversity operator was applied. It is unnecessary if working with a single burst, however, with multiple bursts it should increase the quality of the coregistration [22].

After the interferogram formation the swaths were deburst and merged. Finally the coherence band calculated by the Interferogram operator was terrain corrected, cropped to study area. This process was done separately for VH and VV polarization.

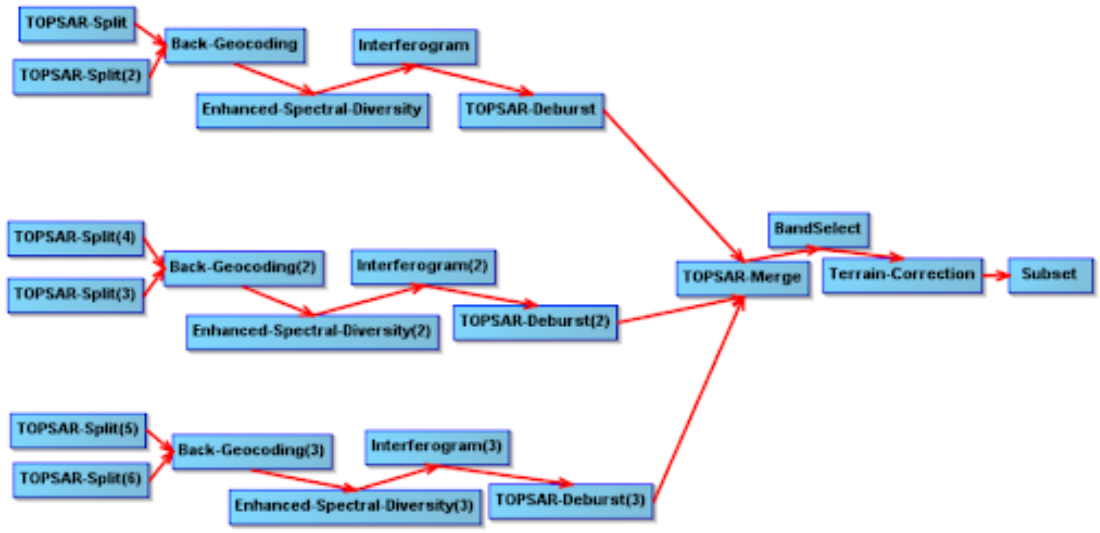

Figure 8. InSAR processing pipeline

4.2.2 Processing output

One of the methods for detecting snow is applying threshold to coherence [2]. The InSAR processing outputs of 2016-2017 season were analyzed to estimate class separability over the entire season. The value distribution over positive and negative class coherence values are compared in Figure 9. The data is separated by polarisation and land classes.

There was greater difference between median coherence values acquired from VV polarized data which is consistent with observations made by He et al. [23]. There is very little difference in coherence distributions for forest land classes. This was expected as dense vegetation has naturally low coherence regardless of snow [23].

Coherence is also correlated with local incidence angle [23]. On current data the effect should be negligible as LIA remained between 20-60 degrees which is Combining data from Sentinel-1 A and B satellites might help as it reduces the temporal variability.

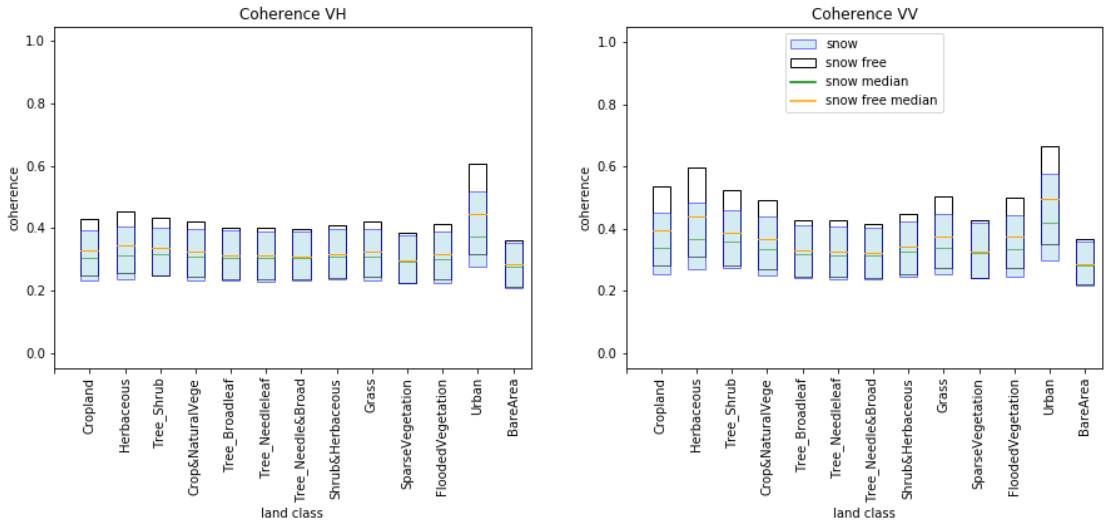

Figure 9. Boxplots comparing coherence distribution of snow free and snow covered pixels for different polarizations and land classes between 2016.11 - 2017.04

4.3 Polarimetric SAR

Polarimetric decomposition is used to determine geometrical scattering characteristics of ground features. Polarimetric methods have only recently become available for dual polarised sensors such as Sentinel-1. PolSAR can be used to detect both wet and dry snow. [2]

There are multiple polarimetric decomposition methods, however the one used in this thesis Entropy, Alpha and Anisotropy (H/A/α) polarimetric decomposition which was also used in the base article by Tsai et al. [3]. The method relies on eigenvalue analysis of the coherency matrix to estimate randomness (Entropy), pirmary (alpha) and secondary (anisotropy) of underlying scattering mechanisms [5].

4.3.1 Processing steps

Polarimetric decomposition requires complex values and thus is only applicable to SLC products. The chain of operations for polarimetric decomposition was separated into two processing graphs both of which are displayed on Figure 10. The first processing segment (Figure 10a applies orbit file, calibrates the data and finally debursts and merges all swaths creating a single continuous image.

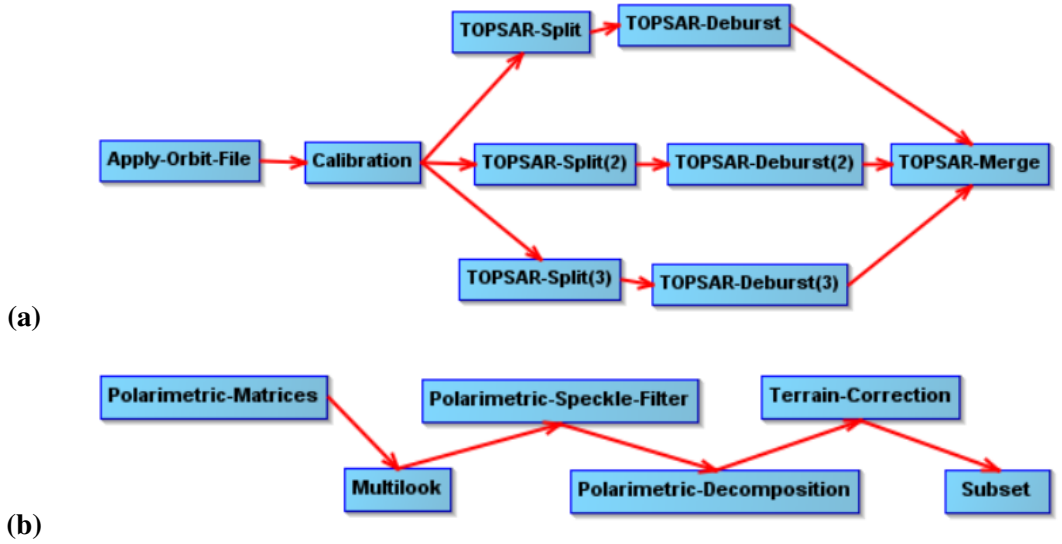

Figure 10. Continuous image creation (a) and polarimetric decomposition (b) processing chains

In the second processing segment (Figure 10b) polarimetric decomposition is performed. First a 2x2 covariance matrix [C2] is generated. The pixels are then multilooked and filtered with The refined Lee filter to reduce SAR inherent speckle noise. The next step applies H/A/α polarimetric decomposition. Finally the result was Terrain corrected and output to the extent on the study area.

4.3.2 Processing output

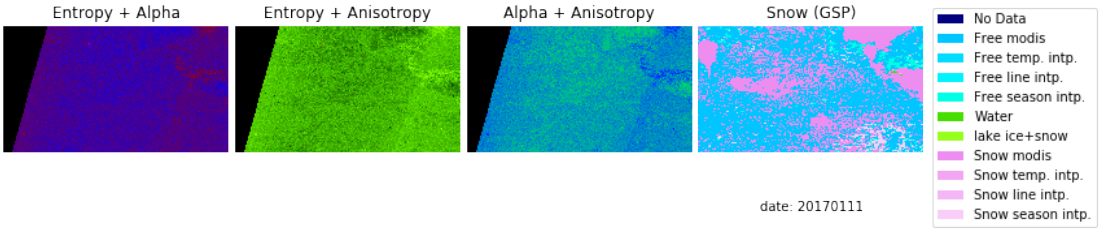

Figure 11. Color composites of H/a/α Polarimetric decomposition parameters and snow cover from Global Snowpack (GSP) for 2017.01.11

On visual comparison of decomposition parameter color composites (red=H, green=A, blue=α) and snow cover Figure 11 there is no obvious connection between decomposition parameter pairs and snow cover. Figures for observation dates between 2016 November and 2017 february are in Appendix 7.

It can be seen in the example that MODIS dataset and therefore GSP (Chapter 3.2) have misclassified water bodies. This happened on multiple dates (Appendix 7, however it is not an issue in current thesis as water bodies are excluded during separation by land class.

There is a visible artifact where subswaths are joined (about quarter width from the right side of image) on color composites of several observation dates (Appendix 7). The artifact looks like thermal noise [24], however if that is the case, it can not be removed with currently available methods. Removal of thermal requires detected values [25]. This is incompatible with PolSAR that requires complex input. Possible solution might be limiting the AOI to discard edges of sub-swath where the noise is the most severe.

5 Experiments

The baseline for experiments was a set of models by Tsai et al. [3]. The focus in the current study was to optimize F1 score over combined land classes, as there was more interest in correctly predicting the positive class than getting high accuracy. All experiments were run on Windows 10 operating system with 16GB or RAM and Intel(R) Core(TM) i5-6200U Processor.

The baseline model by Tsai et al. [3] was a random forest classifier with following parameters:

- 600 trees
- 'sqrt' features
- 0.63 of total training samples per tree

The train-validation split was $70\% - 30\%$ [3]. The sample sizes before split were set to minimum 15000 and maximum 600000 samples per land class for current study. The upper limit for samples was set to keep the training time of each individual model under 20 minutes.

Table 2. Model scores over all land classes. The evaluated models are Unbalanced (UB) Unbalaced with prediction threshold (UB-T), Under-sampled (US), Mixed seasons (MX)

	UB	UB-T	US	MX
Accuracy	0.69	0.68	0.63	0.61
Precision	0.54	0.44	0.44	0.37
Recall	0.03	0.15	0.53	0.47
F1	0.05	0.22	0.48	0.41

There were total of three models trained. First two models were trained only on 2016-2017 season data and tested on 2015-2016 season. The third model was trained and tested on mixed datasets. In addition, the first model was also evaluated. The scores of all models on respective test sets over combined land classes are shown in Table2.

The four models all showed generally weak performance (Table 2). This was in comparison to the baseline models that had minimum accuracy of 0.75 and F1 0.7 [3]. The exact experiment setups and more detailed results are described in the next three chapters.

5.1 Unbalanced Classes

The first model (UB) was trained on SAR images taken between 2016.11.24-2017.03.24. When including April of 2017, the classes were severely unbalanced in favour of negative class. There was about 10% of snow covered pixels in training for any land class. Excluding April that had little snow raised the proportion of positive class samples to around $14\% - 16\%$ for most land classes except *FloodedVegetation* and *Urban* which remained under with 8.1% and 2.3% respectively (Table 3 2.col).

The unbalanced classes expectedly lead to low recall and thus low F1 score on positive class as shown in Figure 12. The F1 score over all classes combined was below 0.1, which is significantly lower than F1 scores over 0.7 on all test sites reported by Tsai et al. [3].

There was an attempt to reduce the model bias by changing prediction threshold of the UB model thus creating model UB-T. It was determined that threshold 0.3 gave the maximum increase in F1 for all models without significant (< 0.05) loss of accuracy. However both recall remained below 0.4 for all land classes. For most classes, except *Tree_Broadleaf* and *Shrub&Herbaceou*s, recall was under 0.2. (Figure 12)

While the overall accuracy of both UB and UB-T over the entire 2015-2016 season was

Table 3. Ratio of positive samples in training set before balancing

Landcover	% snow UB/US	% snow MX
$Cropland$	16.3%	19.9%
$Herbaceous$	15.0%	20.0%
$Crop\&NaturalVege$	14.0%	18.5%
$Tree\_Broadleaf$	23.8%	30.4%
$Tree\_Needleleaf$	16.3%	24.1%
$Tree\_Needle\&Broad$	19.4%	26.4%
$Shrub\&Herbaceous$	14.3%	18.9%
$Grass$	14.8%	16.9%
$FloodedVegetation$	8.1%	12.4%
$Urban$	2.3%	3.3%

about 0.7 (Table 2), scores on individual dates depended mostly on the ratio of snow free to snow covered samples. F1 score was mostly reduced by low recall on positive class. The highest recall was on classes *Tree_Broadleaf* and *Tree_Needle&Broad* that had the highest proportion of snow covered pixels in training data. Thus the following models were trained on balanced classes.

5.2 Balanced Classes

Second model (US) setup was similar to previous one except the classes were balanced by under-sampling the negative class. This caused *Urban* land cover class to have too few training samples (less than 10000) to be included in the model. The training and test score comparisons of other land classes are shown in Figure 13.

There was a significant difference between training and testing scores. However, the goal

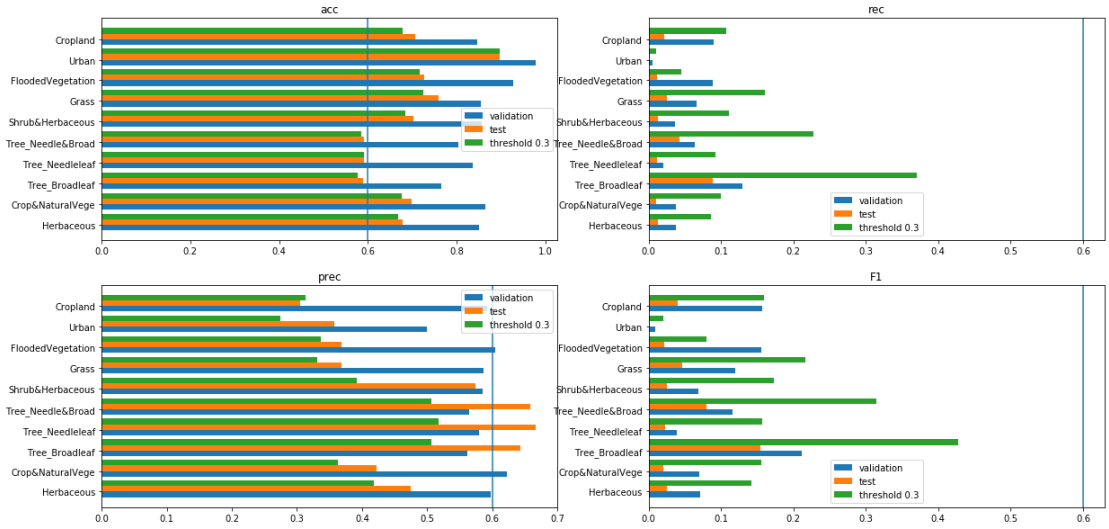

Figure 12. Training and test scores comparison of Unbalanced model with (UB-T) and without thresholding (UB)

of raising F1 score was achieved for all land classes. The overall accuracy of the model was 0.58 which is better than random, but too low for practical use. Recall and precision were better compared to the UB and UB-T models (Table 2).

The difference between training and testing scores may come partially from overestimating positive class on the unbalanced testing data. Another reason might be that there are significant differences in snow characteristics between seasons (Chapter 3.5).

5.3 Mixed Seasons and Balanced Classes

Mixing multiple seasons should include more variations in snow cover and thus make the model more robust. For the third model (MX) training set was still mostly 2016-2017 season. Four dates were from season 2015-2016 were added based on snow depth (Figure 5) and measured temperature (Figure 3) with the aim to include snow conditions not previously in the set. For balance, dates from March 2017 were moved to testing set.

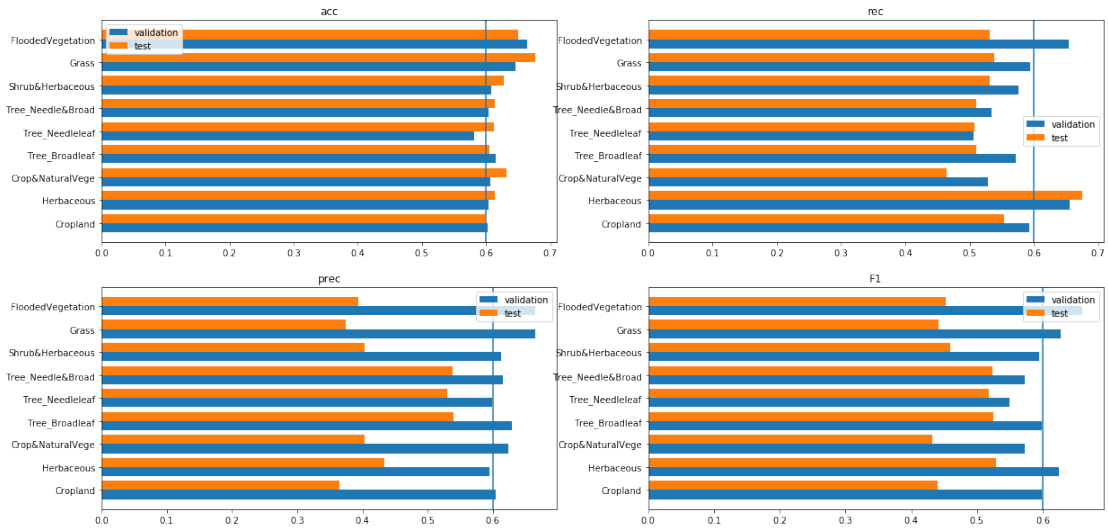

Figure 13. Training and test scores comparison of Under-sampled (US) model

April images of both seasons were excluded due to small amount of snow. The training and testing dates are listed in Table 4

Table 4. Table separation to training and test set

	2015-2016 season				2016-2017 season			
Train	2015.12.24	2016.01.05	2016.03.17	2016.03.29	2016.11.24	2016.12.30	2017.01.11	2017.01.23
					2017.02.04	2017.02.16	2017.02.28	
Test	2016.01.17	2016.01.29	2016.02.10	2016.03.05	2017.03.12	2017.03.24		

Similarly to US model, the training set was balanced by under sampling negative class. Despite the overall increase in samples *Urban* land class was excluded from this model as well due to lack of samples.

The evaluation scores over all land classes are lower than the US model (Figure2). Most other classes maintained accuracy similar to UB model however precision and recall dropped for all LC classes. This is most likely caused by composition of the testing sets. March 2017 had small ratio of snow covered pixels which skewed the test set in favour

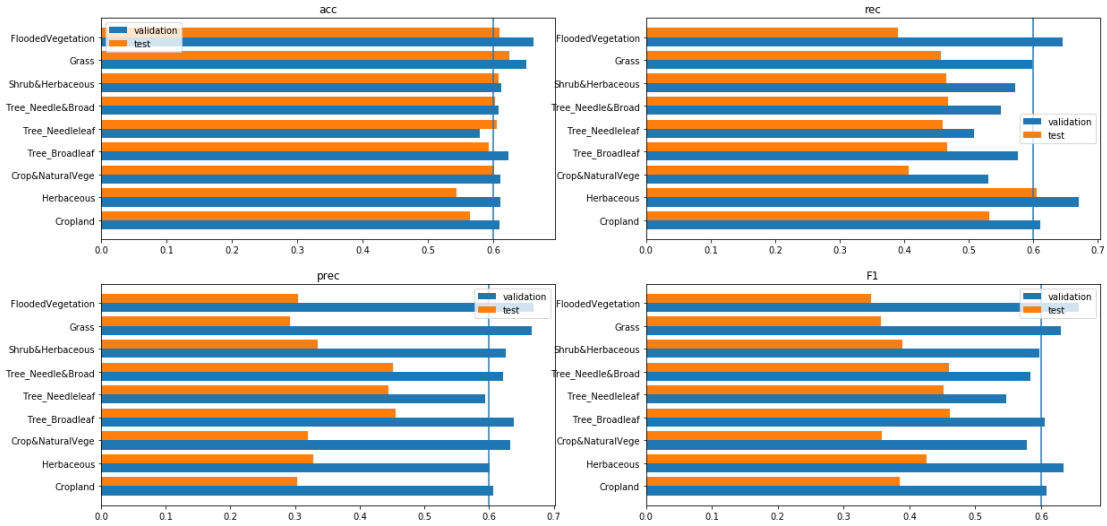

Figure 14. Training and test scores comparison of mixed seasons (MX) model

of snow free pixels which is why accuracy could be maintained despite misclassifying more positive samples.

The negative effect of mixing seasons suggests that adding seasonal variations increases confusion between classes. There are some options that could be explored to reduce the confusion:

- Tuning SAR processing parameters - value distributions of features detected from SAR images show large overlap for snow free and snow covered ground (Chapters 4.1.2, 4.2.2 & 4.3.2)
- Trying different backscatter ratio and PolSAR methods - the methods used in creating the baseline model by Tsai et al. [3] are not necessarily the most effective for Estonian conditions. There exist several alternatives [2]
- Including sentinel-1B data - reducing time between InSAR pairs should reduce noise caused by temporal variability [4]. This is not applicable for data before 2016-2017 season but could be applied for future seasons.

- Tuning model parameters - the number of trees may be insufficient to model variability in snow conditions

Dense vegetation has negative effect on separating classes based on features derived from SAR images [2]. Based on the composition of land classes in Estonia, deriving good features for accurate predictions may not be possible with existing methods.

6 External validation

The model predictions were compared with Results from optical snow detection algorithm Fmask (Chapter 3.3) and weather station measurements (Chapter 3.5). The comparisons were done on 2015-2016 season data for all models.

6.1 Sentinel-2 and Fmask

To compare Fmask and trained model results 10000 points from combined land classes were randomly sampled from every chosen Sentinel-2 image. Both classes were equally represented by 5000 points.

There is no exact match between Sentinel-1 and sentinel-2 acquisition dates. The Fmask results were compared with predictions for the closest sentinel-1 acquisition date. The pairs Sentinel-2 and Sentinel-1 observation dates chosen are in Table 5.

Table 5. Pairs of relatively cloud free Sentinel-2 and closest Sentinel-1 acquisition dates separated by acquisition month

	January		**February**	**March**		
Sentinel-2	20160101	20160131	20160217	20160311	20160318	20160328
Sentinel-1	20160105	20160129	20160210	20160305	20160317	20160329

Figure 15 has monthly evaluation scores for January February and March. January corresponds to first snow spike, February to mid season low point and March to end of season melting period of 2015-2016 season (Figures 3 & 5). The months were represented by two, one and three Sentinel-2 images respectively.

UB-T

(%) True	Predicted **Snow**	Predicted **Bare**
Snow	3	47
Bare	2	48

January

(%) True	Predicted **Snow**	Predicted **Bare**
Snow	3	47
Bare	2	48

February

(%) True	Predicted **Snow**	Predicted **Bare**
Snow	3	47
Bare	2	48

March

US

(%) True	Predicted **Snow**	Predicted **Bare**
Snow	17	33
Bare	12	38

(%) True	Predicted **Snow**	Predicted **Bare**
Snow	15	35
Bare	6	44

(%) True	Predicted **Snow**	Predicted **Bare**
Snow	17	33
Bare	11	39

MX

(%) True	Predicted **Snow**	Predicted **Bare**
Snow	18	32
Bare	13	37

(%) True	Predicted **Snow**	Predicted **Bare**
Snow	16	34
Bare	6	44

(%) True	Predicted **Snow**	Predicted **Bare**
Snow	17	33
Bare	11	39

Figure 15. Monthly confusion matrix over combined land classes

The UB model predicted positive class on less than 0.05% of samples for all months and was not included in the Figure 15. The model UB-T overwhelmingly predicted the negative class. The under-sampled models (US and MX) also somewhat favoured the negative class, but to a lesser extent. There was little difference between monthly scores, which suggests the snow conditions did not have strong effect on predictions.

6.2 Weather station measurements

Weaher station measurements are the most accurate. Unfortunately they are sparse. There were disagreements between GSP and weather station data (Figure 16). The data set marking snow covered ground as ground free happened more frequently when measured snow depth was under 15mm and GSP pixel had been interpolated.

Snow depth for each weather station in Estonia is measured from three places over area

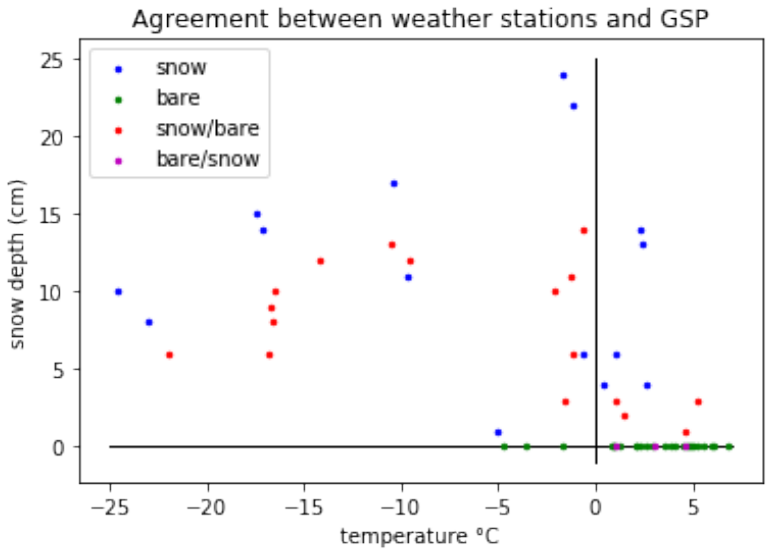

Figure 16. Comparison between snow depth measurements and GSP classification

of no less than 400_m [6]. The sample nearest to a station does not necessarily exactly overlap the snow measurement area meaning in some cases the disagreement could be caused by misalignment.

All models were validated their respective test sets. The validation results are listed in Figure 17. The ratio of agreement with weather station measurements was greater than with Fmask algorithm (Figures 17 & 10b). This could be attributed to limited selection of land classes in weather station locations.

Four out of eight stations were in sparsely vegetated areas: two in *Herbaceous* and one in *Grass* and *Urban* land cover each. The other stations were in densely vegetated areas however only one was in forest region and it was missing measurements on half of the observation dates. This gives severely inaccurate representation of actual land classes (Figure 1).

External validation results are consistent with scores on GSP data. The unbalanced models show strong bias toward the negative class. Setting a lower prediction threshold

(%)	Predicted	
True	**Snow**	**Bare**
Snow	2	48
Bare	0	50

UB

(%)	Predicted	
True	**Snow**	**Bare**
Snow	17	33
Bare	7	43

UB-T

(%)	Predicted	
True	**Snow**	**Bare**
Snow	29	23
Bare	12	37

US

(%)	Predicted	
True	**Snow**	**Bare**
Snow	23	21
Bare	21	36

MX

Figure 17. "Confusion matrix of each model on weather station measurements"

for positive class had a weak effect on raising the rate positive predictions. The approach of under-sampling the majority class The MX model has higher confusion rate compared to the US model. The US model is the best at recognising snow out of the four models compared, but has accuracy lower than 0.7 and F1 score around 0.5.

7 Conclusion

A method previously tested in mountain regions was applied for detecting seasonal snow cover in Estonia. Backscatter, InSAR and PolSAR based features were extracted from SAR images and combined in a rndom forest approach. Separability of land classes was analyzed based on the extracted features. Lastly several models were trained and evaluated both with the chosen ground truth data set as well as external data sources. It was discovered that the snow covered and snow free classes largely overlap for individual features. There was further analysis on possible causes for this. The overlap in classes was also reflected in all trained models having accuracy between 0.6-0.7. The F1 scores on positive class ranged from 0.03-0.53 which is too low to be applied to real situations. Future work on the subject should mainly focus on improving class separability. This study indicates that finding optimal parameters for feature extraction from Sentinel-1 SAR images could be effective. For backscatter this could include changing the method of ratio calculation. For InSAR based coherence one option might be including Sentinel-1B data to reduce temporal distance between interferometric pairs. Furthermore, including additional parameters such as temperature and increasing the specialisation of sub models by more detailed land class separation could also be explored.

References

[1] European Environmental Agency. Snow cover. https://www.eea.europa.eu/data-and-maps/indicators/snow-cover-3/assessment, 2016.

[2] Ya-Lun Tsai, Dietz Andreas, Natascha Oppelt, and Claudia Kuenzer. Remote sensing of snow cover using spaceborne sar: A review. *Remote Sensing*, 11:1456, 06 2019.

[3] Ya-Lun Tsai, Dietz Andreas, Natascha Oppelt, and Claudia Kuenzer. Wet and Dry Snow Detection Using Sentinel-1 SAR Data for Mountainous Areas with a Machine Learning Technique. *Remote Sensing*, 11:895, 04 2019.

[4] Snehmani, Mritunjay Kumar Singh, R.D. Gupta, Anshuman Bhardwaj, and Pawan Kumar Joshi. Remote sensing of mountain snow using active microwave sensors: a review. *Geocarto International*, 30(1):1–27, 2015.

[5] Shane R. Cloude and Eric Pottier. An entropy based classification scheme for land applications of polarimetric sar. *IEEE Trans. Geosci. Remote. Sens.*, 35:68–78, 1997.

[6] Estonian Weather Service. Ilmatarkus. https://www.ilmateenistus.ee/ilmatarkus/.

[7] Africa Ixmucane Flores-Anderson, Kelsey E. Herndon, Rajesh Bahadur Thapa, and Emil Cherrington. The Synthetic Aperture Radar (SAR) Handbook: Comprehensive Methodologies for Forest Monitoring and Biomass Estimation. 04 2019.

[8] C. A. Wiley. Synthetic aperture radars-a paradigm for technology evolution. *IEEE Transactions on Aerospace and Electronic Systems*, AES-21(3):440–443, 1985.

[9] European Space Agency. Sentinel-1 User Guide. https://sentinel.esa.int/web/sentinel/user-guides/sentinel-1-sar.

[10] European Space Agency. Sentinel-1 SAR Technical Guide. https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-1-sar/.

[11] Richard Barnes. RichDEM: Terrain Analysis Software. https://github.com/r-barnes/richdem, 2016.

[12] A. Nelson E. Guevara Jarvis A., H.I. Reuter. Hole-filled seamless SRTM. http://srtm.csi.cgiar.org, 2008.

[13] Andreas J. Dietz, Claudia Kuenzer, and Stefan Dech. Global snowpack: a new set of snow cover parameters for studying status and dynamics of the planetary snow cover extent. *Remote Sensing Letters*, 6(11):844–853, 2015.

[14] Neil Flood. Python Fmask. http://www.pythonfmask.org/en/latest/.

[15] Shi Qiu, Zhe Zhu, and Binbin He. Fmask 4.0: Improved cloud and cloud shadow detection in landsats 4-8 and sentinel-2 imagery. *Remote Sensing of Environment*, 231:111205, 09 2019.

[16] Republic of Estonia Land Board. National mirror site (ESTHub). https://www.maaamet.ee/en/objectives-activities/geoinformatics/national-mirror-site-esthub.

[17] European Space Agency. Sentinel Application Platform. https://step.esa.int/main/toolboxes/snap/, 2016.

[18] T. Nagler and H. Rott. Retrieval of wet snow by means of multitemporal sar data. *IEEE Transactions on Geoscience and Remote Sensing*, 38(2):754–765, 2000.

[19] Thomas Nagler, Helmut Rott, Elisabeth Ripper, Gabriele Bippus, and Markus Hetzenecker. Advancements for snowmelt monitoring by means of sentinel-1 sar. *Remote Sensing*, 8(4):348, Apr 2016.

[20] N. Baghdadi, J.-P. Fortin, and M. Bernier. Accuracy of wet snow mapping using simulated radarsat backscattering coefficients from observed snow cover characteristics. *International Journal of Remote Sensing*, 20(10):2049–2068, 1999.

[21] Jarkko Koskinen, Jouni Pulliainen, and Martti Hallikainen. The use of ers sar data for snow melt monitoring. *IEEE Transactions on Geoscience and Remote Sensing*, 35:601 – 610, 06 1997.

[22] Andreas Braun and Luis Veci. TOPS Interferometry Tutorial. `http://step.esa.int/docs/tutorials/S1TBX%20TOPSAR%20Interferometry%20with%20Sentinel-1%20Tutorial_v2.pdf`, 2020.

[23] Guangjun He, Pengfeng Xiao, Xuezhi Feng, Xueliang Zhang, Zuo Wang, and Ni Chen. Extracting snow cover in mountain areas based on sar and optical data. *IEEE Geoscience and Remote Sensing Letters*, 12:1–5, 05 2015.

[24] Jeong-Won Park, Anton Korosov, Mohamed Babiker, Stein Sandven, and Joong-Sun Won. Efficient thermal noise removal for sentinel-1 topsar cross-polarization channel. *IEEE Transactions on Geoscience and Remote Sensing*, PP:1–11, 12 2017.

[25] S-1 Mission Performance Centre. Thermal Denoising of Products Generated by the S-1 IPF . `https://sentinel.esa.int/documents/247904/2142675/Thermal-Denoising-of-Products-Generated-by-Sentinel-1-IPF`, 11 2017.

Appendix

I. Polarimetric decomposition results

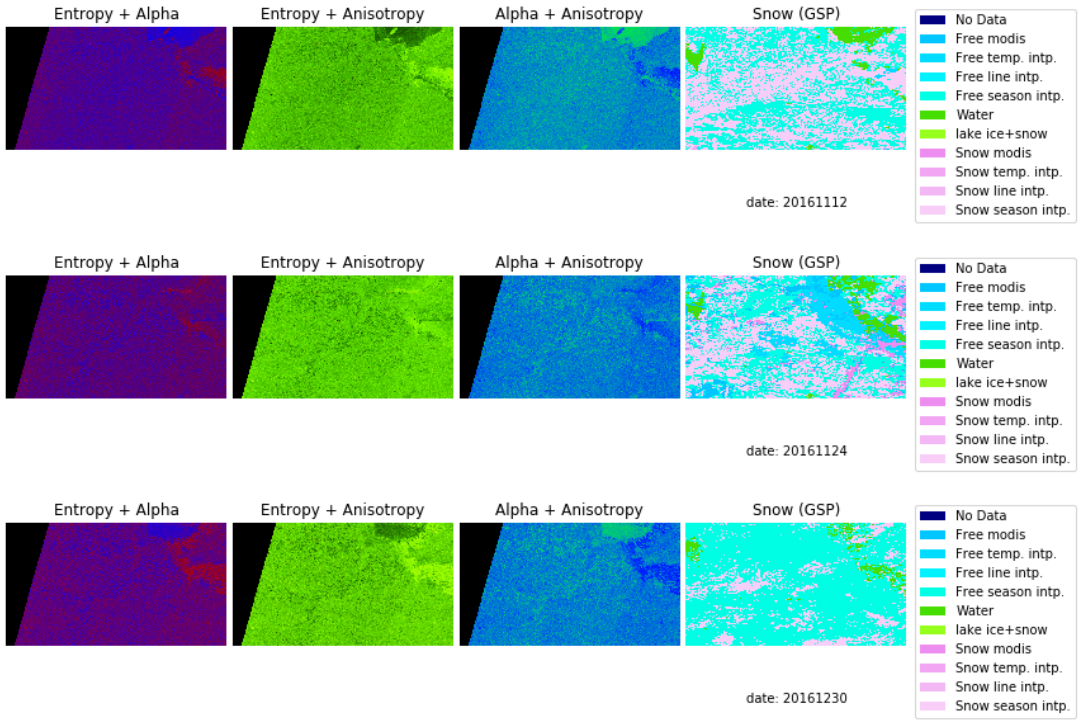

Figure 18. H/a/α Polarimetric decomposition results and snow cover from Global Snowpack (GSP)

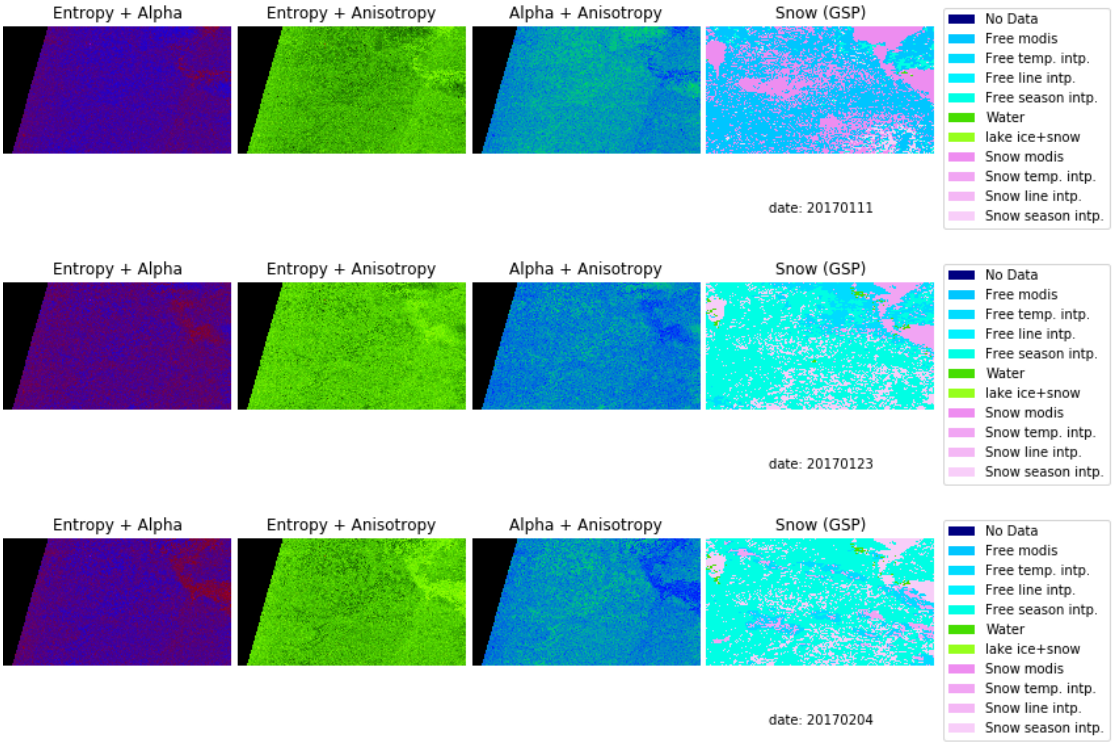

Figure 19. H/a/α Polarimetric decomposition results and snow cover from Global Snowpack (GSP)

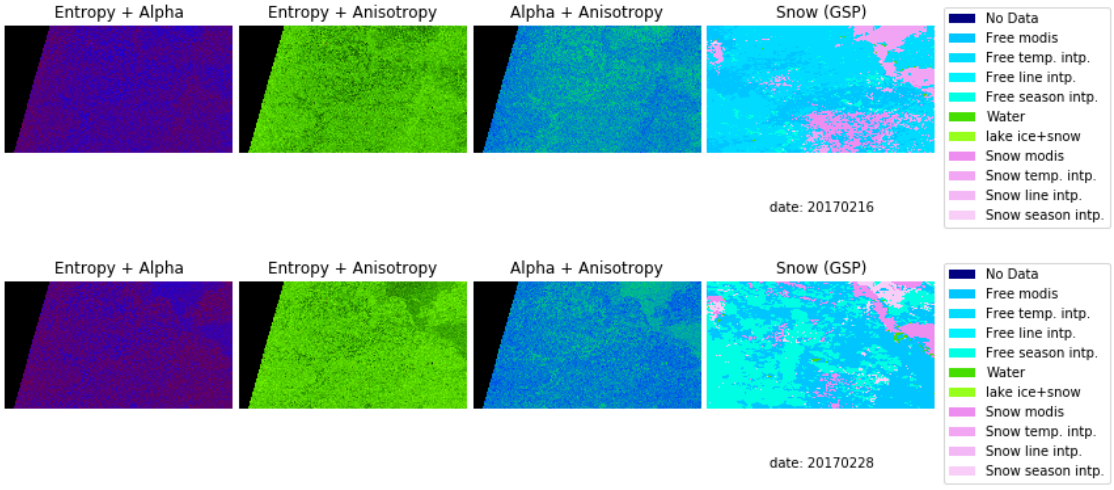

Figure 20. H/a/α Polarimetric decomposition results and snow cover from Global Snowpack (GSP)

Licence

Non-exclusive licence to reproduce thesis and make thesis public

I, **Kerstin Äkke**,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to

   reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

   Snow cover detection in Estonia from SAR images using machine learning methods,

   supervised by Viacheslav Komisarenko and Anti Gruno

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Kerstin Äkke
15/05/2020