Detecting Alpine Treeline Ecotones – an Automated Remote Sensing Approach

Thesis Paper for the Project Seminar: Treelines of the World in SS 2020

University Lecturer: Dr. Maaike Bader, Physical Geography, Philipps-Universität Marburg

Student: Agnes Schneider

Matriculation number: 2689766

E-Mail: Schnei7v@students.uni-marburg.de; euboia@gmail.com

Date of Submission: 01.01.2022

Abstract

1 Introduction

Alpine Treeline Ecotones (ATEs) are transitional zones between subalpine Forest and Alpine (tundra) ecotones (Holtmeier & Broll (2005), Winings (2013)) also referred to as upper-treeline (Elliott (2017)) and occur globally (Singh et al. (2015), Bader et al. (2021)). They span between the actual Timberline/Economic Forest Line though the Upper/Physiognomic-Biologic Forest Line and the tree species line which is adjoining the actual Alpine zone (Chhetri & Thai (2019), 1543). The position of the treeline is influenced by multiple factors at local and regional level, but temperature has been identified as the global driving factor (Körner (1998), Körner & Paulsen (2004), Holtmeier & Broll (2005), Bader et al. (2007), Barredo Cano et al. (2020)). The global pattern can be described by spatial patterns in the x-y plane (discrete or diffuse, Figure 1) and by changes in tree stature (abrupt or gradual, Figure 2) in a multi-dimensional space (Harsch & Bader (2011), Bader et al. (2021)).

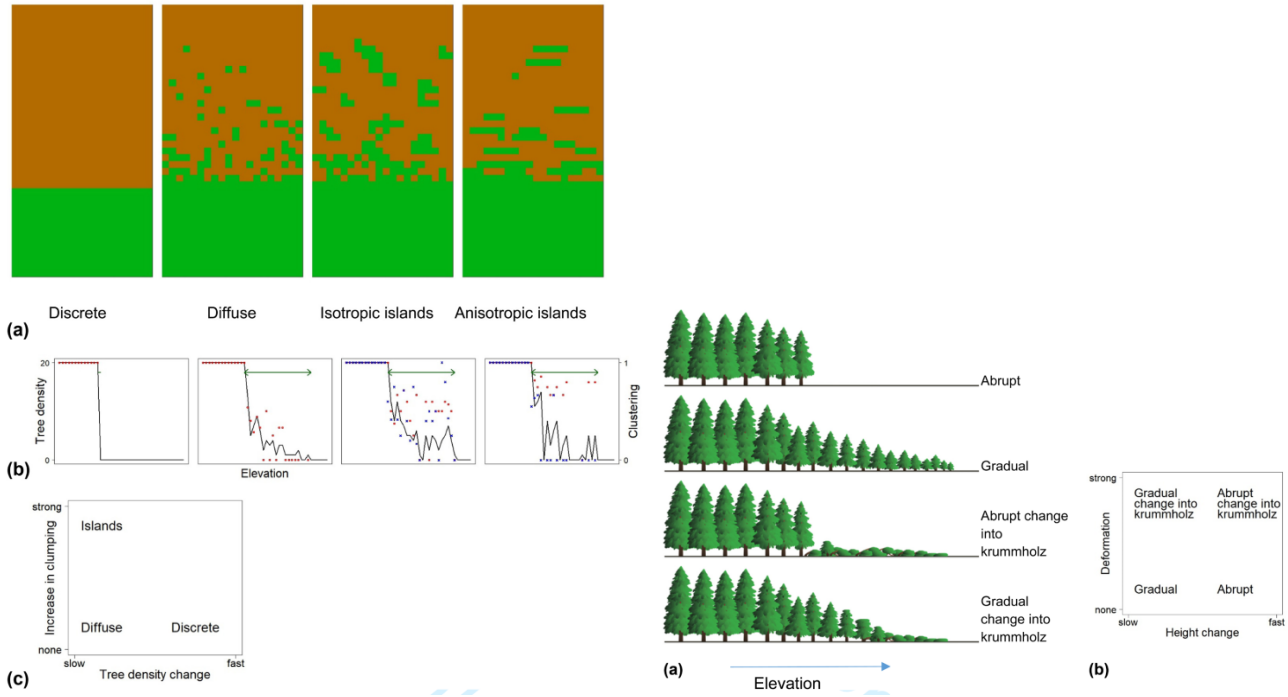

Figure 1: Left: Scheme of the spatial pattern of alpine treelines on the 2D x-y plane. a) Depicts the treeline as seen from above, while b) depicts the change of the treeline in the y direction (clustering of islands). c) Represents an abstraction of the pattern of treelines based on tree density change and the clustering of individual trees. Source: @baderGlobalFrameworkLinking2021, Figure 1. Right: Scheme of (discrete) tree stature/height change responding to change in elevation. a) Vertical cross section. b) Abstraction of tree stature change based on height change and deformation of tree shape. Source: @baderGlobalFrameworkLinking2021, Figure 2.

Although recognized, the distribution of ATE patterns have neither been mapped, nor been

described yet, let alone explained. Earlier studies have identified abrupt, diffuse, island and krummholz spatial patterns of ATEs (Harsch et al. (2009), Harsch & Bader (2011), Figure 2). As seen, treelines dispaly a high variability and differ in multiple dimensions. A comparison of multiple studies suggest, that the different spatial patterns of ATEs reflect fundamental ecological controlling processes and that different ATEs react differently to climate change (Harsch & Bader (2011), Figure 1) Figure 3.

	a) Rate of change	b) Type of change	Controlling processes (first-level)	Controlling processes (second- and third-level)
2-D pattern	**Discrete**		Mortality above but not in the ecotone	Any type of damage or stress. If gradual, growth limitation above the ecotone
	Ecotone	**Diffuse**	Mortality within the ecotone	Environmental heterogeneity or stochasticity. And/or ongoing seed-limited colonization
		Island	Mortality within the ecotone	Positive feedback due to radiation damage, freezing damage, wind damage, snow redistribution (snow moulds, growing-season length), clustered seed dispersal, clonal reproduction
Tree stature	**Abrupt**	Into alpine	Mortality	Overheating, desiccation, photoinhibition, freezing damage, soil limitations
		Into krummholz	Dieback	Winter desiccation, snow moulds, browsing, abrasion
	Gradual	No krummholz	Growth	Tissue formation (limited by a lack of heat)
		Into krummholz	Growth and dieback	Tissue formation + winter desiccation, snow moulds, browsing, abrasion

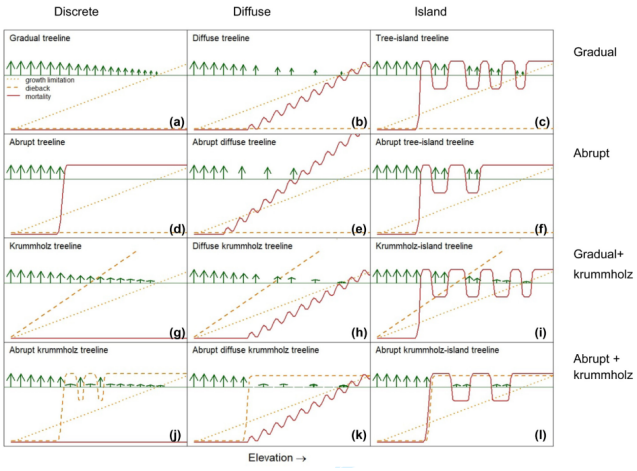

Figure 2: Left: Matrix of 2D spatial pattern, stature change and ecological processes which can contribute to the different types of ATEs. Source: @baderGlobalFrameworkLinking2021, Table 1. Right: Matrix of the multidimensional state space of treeline forms depicting extreme cases of the different dimensions. Columns represent the spatial patterns in the x-y plane and rows the change in tree stature (size and shape). The lines represents the hypothesized first-level ecological processes behind the patterns along an elevational gradient. The dotted line displays the growth limitation, the dashed line the dieback and the continuous line the mortality. Growth limitation always occurs while dieback only affects if krummholz is involved. Source: @baderGlobalFrameworkLinking2021, Figure 3.

To better understand and categorize ATEs a standardized description and terminology of spatial patterns has been proposed on hillslope and landscape scale by Bader et al. (2021), including hypotheses for the general mechanisms behind the patterns. The terminology and the multidimensional state-space can be most clearly understood from Figure 4.

After the definition and the characterization of the spatial patterns of ATEs, the following overview of ATE research shall serve as a basis for the workflow proposed in this methodological paper to globally detect ATEs.

2 Overview of Alpine Treeline Ecotone Research with focus on methodology

As already implied, Alpine Treeline Ecotone analysis has been carried out so far at local, regional and global scales with different focus. Chhetri & Thai (2019) investigate the use of GIS and remote sensing methods in ATE research between 1980 and 2017 and demonstrate that with the advances in sensor technology the access to higher resolution data and different data types the analysis methods diversify. Often it is hard to distinguish between the different methodological approaches because they all rely and depend on one-another and studies have usually applied multiple methods and methods have been applied for diverse scopes. The following overview of methods is crude and by no means comprehensive.

2.1 Statistical – analytical approach

Until the dispersal of openly available remote sensing products the approach was mainly concentrated on physiological/response based investigations with statistical analysis of local indicators based on in situ measurements. The main question was mainly: which environmental drivers control the location of ATE's and how complex are these (Körner 1998)? Drivers were differentiated on global, regional and local levels (Holtmeier – Broll 2005). As the main global driver temperature was defined (Körner (1998), Koerner (2012), Holtmeier & Broll (2007), 2, Bonanomi et al. (2018)), but at regional and local levels it looks more diverse: topography (Brown (1994), Virtanen et al. (2004) and Bader & Ruijten (2008)), geomorphologic processes, herbivory or anthropogenic disturbance (Chhetri & Thai (2019)). Analysis of topographic actors with logistic regression was used frequently (e.g Brown (1994), Virtanen et al. (2004) and Bader & Ruijten (2008)).

2.2 Remote sensing approach

The detection or localization of ATEs is facilitated by the availability of remote sensing data (satellite, airborne lately hyperspectral and UAV derived ortho/imagery). Specialized sensors operating in the R, G, B, IR, short-wave, hyperspectral and thermal regions of the electromagnetic spectrum give the possibility to approach and work with the specific spectral signatures (due to different absorption and reflection of radiation) of different vegetation entities. LiDAR

data enables to include the factor height thus moving the analysis of treelines in a 3D space. The main focus of this methodological paper is on the remote sensing approach (automated detection) and exemplary papers are accentuated in a short summary. The use of GIS and remote sensing has been constantly increasing in treeline studies since 2000, with a few preceding pioneers. Earlier studies concentrated on mapping treeline positions and lately the interest shifted towards factors that control treeline variation (Chhetri & Thai (2019), 1543). It can also be seen that with the development of the respective sensors the interest and use moved to data with higher spatial resolution (LiDAR data, Hyperspectral data), which on the other hand attracts more cost and thus often thins out the studies due to the lack of monetary resources and also the use of proprietary software, which shows that there is a lot to do in means of open-source and reproducible best-practice applications and code. Usually there is little information on the software used. The use of remotely sensed data is usually combined with classical approaches like statistical analysis in complex research questions, from which the following main directions emerge:

2.2.1 *Mapping of ATEs*

To map ATEs, studies use aerial photographs and Landsat imagery to identify and quantify treelines (Brown (1994), W. L. Baker et al. (1995), Allen & Walsh (1996), Kimball & Weihrauch (2000), Virtanen et al. (2004), and Resler et al. (2004)) but often also vegetation indices are used Myneni et al. (1997), Singh et al. (2015), and Mohapatra et al. (2019)) to analyse treeline elevation (Allen & Walsh (1996) and Kimball & Weihrauch (2000)) and topographic variables/geomorphological parameters (slope, angle, curvature, relief) to explain treeline structure (patch-metrics) (Kimball & Weihrauch (2000)). Tree population parameters are derived via PCA (e.g W. Baker & Weisberg (1997)) and also species distribution modelling is a usual application (Chhetri et al. (2017)).

2.2.2 *Montitoring ATEs/Change detection*

ATEs are space and time related phenomenons and they respond to changing environmental conditions, that is they can be sensitive to climate change (Singh et al. (2015), Holtmeier & Broll (2005), Holtmeier & Broll (2007), Harsch & Bader (2011), and Bader et al. (2021)). The rise in global average temperatures seems to lead to the geographically varying shifting

of ecotones: on regional level to upward shift ((Mohapatra et al., 2019)) but also stable or retracting ATEs can be determined (Winings (2013)). It still has to be understood if the results are due data quality. The identification and quantification of change in the ATEs can be carried out with regional and global monitoring of ATEs (Chhetri & Thai (2019)).

2.2.3 *Automated detection and mapping of ATEs*

Recently several research projects, Master theses and PhDs have investigated (semi-)automated detection and mapping methods of ATEs (see a list until 2013 in Winings (2013) and the also the recent literature). The availability of high resolution data facilitate the use of more and more sophisticated methods. In this methodological paper we are concentrating specifically on automated analysis.

Automated methods imply the use of specific algorithms to extract information from remote sensing data, either pixel- or object based or recently also via Deep Learning (mainly CNNs). Some studies compare object- and pixel-based or different object-based segmentation methods (Immitzer et al. (2012), Winings (2013), and Kupková et al. (2017)). Also it has to be emphasized, that the automated detection of ATEs relies heavily on tree detection. Parallel to elaborated workflows for the automated detection of ATEs improvements are made continuously on tree detection methods and tree cover estimation (Whiteside et al. (2020)) closely connected to the development of sensors and the new data processing methods (Qiu et al. (2020), Weinstein et al. (2019) and Weinstein et al. (2020)) and form an important basis for automated analysis. In the following the most prominent automated methods are presented shortly including a few case studies.

Pixel-based Image Analysis is working with the information encoded in pixels – it assigns each pixel to a specific class on the basis of the respective values of the spectral bands or index or morphometric information (slope, aspect, etc.). One drawback is, that the context of the pixels and it's neighborhood gets neglected and the pixel values can be affected by circumstantial effects, like reflectance differences (Stueve et al. (2011)), shadow or clouds (Allen & Walsh (1996)). Also it doesn't deal, with textures per se, and for this a textural analysis has to be done by using different filters (mean, sobel, focal, etc.).

Resler et al. (2004) use panchromatic aerial imagery where they incorporate spectral (brightness values) and spatial (textural) information to classify 4 classes (tundra/bare, alpine meadow,

open forest/krummholz and closed canopy) representing the ATE using the maximum likelihood algorithm. A classification with and without textural information was done to assess the meaningfulness of textural information. The ERDAS modeler was used to extract textures.

Král (2009) uses CIR orthophotos to do a "classical" landcover classification using the maximum likelihood classifier, which then is reclassified into 2 classes (spruce canopies and other). Subsequently a focal filter is applied to the spruce canopy closure class for texture analysis, which is then reclassified into 5 classes (no trees, emergent trees, groups of trees, open-canopy forest, closed canopy forest). Class 3, that is groups of trees (26-50% spruce), neighbouring to alpine grassland and open-canopy forest was defined as ATE.

Immitzer et al. (2012) used WW-2 satellite data (8 + 4 bands) for the identification of 10 tree species by means of Random forest classification (object-based vs. pixel-based) using spectra of a) manually delineated tree crowns b) derived tree crown polygons and reference samples for tree species.

Winings (2013) used high resolution aerial imagery and LiDAR data in her Master's thesis to map the alpine treeline. She compared pixel- and object based classification. She used four different data input for both classification methods: NDVI, NDVI + multispectral aerial imagery, NDVI + tree height or NDVI + multispectral aerial imagery + tree height. In the case of the pixel-based classification the maximum likelihood and the unsupervised ISODATA (Iterative Self-Organizing Data Analysis Technique) clustering algorithms were compared. For the object-based image analysis multi-resolution segmentation was conducted, using colour and shape homogeneity. After the segmentation, the classes (tree vs. non-tree) were assigned based on object feature threshold. The accuracy for the pixel-based classifications was between 85.3 and 88.4 and for the object-based classification between 81.5 and 92.9 %, resulting in the best classification on the dataset with NDVI + multispectral aerial imagery + tree height. For the pixel-based image analysis ENVI and ERDAS and for the object-based analysis eCognition was used.

Kupková et al. (2017) used airborne hypersepctral (APEX and AISA DUAL) and Sentinel-A data for the classification of tundra vegetation by comparing pixel-based and object-based image analysis. Reference data was collected corresponding to 8 vegetation classes (anthropogenic areas, picea abies, pinus mugo dense, pinus mugo sparse, closed alpine grassland, grasses, alpine

heathlands, wetlands and peat bogs; with a detailed and a simplified legend). Based on the difference in resolution the hyperspectral data and the Sattelite imagery was classified separately. Latter was only classified pixel-based and with SVM (Support Vector Machines with radial basis function), NN (Neural Net) and MLC (Maximum Likelihood Classification) algorithms. The hyperspectral data, having a higher spatial resolution was classified pixel- and object-based. For the pixel-based classification the SVM, NN and MLC algoriths were used. For the objcet-based classification Edge-based segmentation was used on the hyperspectral datasets. The hyperspectral data yielded better classification result thean the Satellite data, with SVM pixel-based classification. ENVI was used for the study.

Object-based Image Analysis (GeOBIA) on the other hand is dealing with the grouping of pixels in homogeneous groups, that is segments which bear similar spectral, spatial and textural information. From each segment additional information can be extracted (statistical information, size, shape and context). Different segmentation algorithms exist, which treat the image and the segments different.

Middleton et al. (2008) used the Feature Extraction Module (Fx) implemented in ENVI to extract tree crowns from two aerial photographs (one from 1947 and one from 2003) via segmentation and feature classification with SVMwith textural, spatial and spectral information. The results were compared to forest inventory information and an upward shift was recorded on Lommoltunturi fell.

Ranson et al. (2011) used MODIS VCF (Vegetation Continuous Fields) tree cover data and segmentation to delineate the circumpolar taiga-tundra ecotone (TTE). The multi-annual VCF was adjusted usingf linear regressions and a vector layer was applied with previously delineated taiga and tundra biomes. Also the water bodies were masked out. Subsequently multi-resolution segmentation was carried out with eCongnition based on the homogeneity criterion. The resulting polygons were then classified on a specific range of adjusted VCF values which represent the TTE.

Mishra et al. (2018) used a UAV equipped with a Parrot Sequoia multispectral (Red, Green, Blue, Red Edge, Near Infra-Red) camera to acquire high resolution Imagery. Subsequently an SfM Orthoimage was calculated and then multi-resolution (based on the homogeneity criterium of scale, shape/colour and compactness/smoothness) and spectral difference segmenta-

tion (merging neighbouring objects based on a spectral threshold) was combined in eCognition to generate optimal feature space variables for the classes. Then the Random Forest Classifier was used for classification with 3 sets of features (spectral features; spectral features + geometric/shape features; spectral features + geometric/shape features + textural features) for species-level mapping of vegetation in the Himalayas.

Whiteside et al. (2020) used derivatives of aerial imagery and WW2 satellite data (TGI, NDVI) resampled to 1 m filtered by a low-pass filter. Then a threshold-based multi-resolution segmentation was conducted with eCognition to assess the tree cover (in percentage) for each date (1964, 1976, 1981, 2010). The results were compared by date to assess the tree cover reduction (4%) during the 36 years.

Geping Luo & Li Dai (2013) used aerial imagery from 1962 and 1981, QuickBird Satellite image from 2006 was used as data input to map vegetation distribution after orthorectification, and generating a DEM. The land-cover types were delineated: Schrenkiana, Sabina and other. Multi-resolution (?) segmentation was conducted in eCognition subsequently combined with a k-nearest neighbour classification. The result was compared with fieldwork data collection (2010, 2011) of the two species. With the post-classification approach the land-cover change was examined between 1962, 1981 and 2006.

Qiu et al. (2020) proposed a new spectral multi-scale (SMS) individual tree crown (ITC) delineation method using both brightness and spectra of high-resolution multrispectral imagery to be able to better delineate tree crowns in deciduous or mixed forests, where adjacent tree crowns are very close to each other. As the first step a morphological gradient map is calculated of multispectral images, then as a second step an inverse gradient image. Then initial treetops were extracted by multi-scale filtering and morphological operations with regard to tree crown shape which then were refined with the spectral reference of the neighbouring treecrowns (tree trops map). Subsequently the morphological gradient map is segmented by marker-controlled watershed segmentation which is then refined by the tree tops map, to receive an individual tree crown delineation map.

Deep Learning – contrary to pixel-based and GeOBIA – works on scene level and enables thus to deal better with the complex semantic structure of the increasing resolution of remote sensing images. A multitude of different Deep Learning models exist with different structures

to fulfill different aims (e. g. segmentation, classification). The most common Deep Learning model structure are CNNs – Convolutzional Neural Networks, which are multi-layer networks with learning ability that consists of convolutional layers, pooling layers, and fully connected layers.

Fricker et al. (2019) used airborne hyper-spectral imagery, LiDAR data and a CNN (Convolutional Neural Network) to automate tree species classification. 7 dominant tree species and a dead tree class were identified to serve as reference data for the CNN. A LiDAR derived CHM was used to digitize the individual tree canopies to prepare their pixels for the species labelling for the CNN. The classification was executed separately on the RGB and the hyper-spectral data. The classification with the hyper-spectral data (0.73 – 0.90) yielded better classification results than the RGB classification (0.41 – 0.88). All code and data to ensure reproducibility can be found online.

Weinstein et al. (2019) proposed a semi-supervised CNN workflow based on the comparision of 3 unsupervised tree-crown segmentation algorithms. The result of the chosen tree crown segmentation (clustering of a CHM by tree height and crown width) of the LiDAR data was extracted as a bounding box from the RGB image, which dataset is then labeled self-supervised pretrained by a retinanet CNN. Then the CNN was retrained with a small hand-annotated dataset (supervised classificiation), to correct errors from the intitial un-supervised segmentation, which indeed improved the results of the prediction.

Weinstein et al. (2020) build on the results from Weinstein – Marconi – Bohlman – Zare – White 2019 and tested if training datasets can be generalized and be used on completely different forested areas. Generally the performance of the model performance decreased, but when they were applied to spatially and spectrally similar forested areas the performance increased. Best was again, when the CNN was retrained by a handful of hand-annotated data from the same area.

3 Methodology

3.1 Description of the research area and the data sets used in the case study

The workflow is presented on a case study area in Vorarlberg, Alps…

3.2 The reproducible workflow

The reproducible workflow is contained in the files: **00_library_n_prep.R**, **0_set_up_working_envrnmt.R**, **1_data_prep_general.R**, **2_segmentation_area_1.R**, **3_segmentation_validation_ROI.R**, **4_data_prep_classification_area_1.R**, **5_classification_validation_area_1.R** and **6_data_prep_classification_ROI.R**.

These R files contain the whole selection process of the best suitable variables. This .Rmd only present the best suitable variables, with references to the respective .R files.

To reproduce the results of the whole project you can do that by running the code in the .R files, numbered sequentially. 00_ is the library file, you do not have to run it. **0_set_up_working_envrnmnt.R** is going to set up the project folder structure on your computer/laptop by adapting the path of the project directory to your computer/laptop. Then you have to download the data from Google Drive in the respective folders (because a)**github** limits the size of data to be uploaded and b) the data used is not openly available, that is proprietary) and put the data into the respective folders of the folder structure you just created on your computer: **dsm/**, **dem/**, **RGB_IR/**, **RGB_IR/**, **treepos/**, **train/**.

The workflow consist of to main workflows: the pixel-based image classification and the Object Based Image Analysis which will be combined at the end. First the values of the variables to be used were elaborated/tested on test area 1 and then applied to the ROI. Thus you will first work with test area 1 and then apply the result on the ROI. Also only the successful result will be used in this .Rmd. You can learn about the testing sequence by checking it in the respective **.R** files.

After the folder structure on the computer/laptop was set up by executing the code in **0_set_up_working_envrnmnt.R**, the raw data will be prepared and the CHM for the

ROI is generated by subtracting the DEM from the DSM! Following that the position and the properties of the CHM just created are checked and the layer is renamed. Then the CHM is read in again and plotted to check it's crs.

Flowchart of the workflow

The .R files are commented in a manner, that they can be used as a manual or tutorial. In the next chapter only the ideas behind the workflow and the results will be discussed for shortness.

3.3 Data preparation (.R file: 1_data_prep_general.R)

Originally masks for each 4 test area were created in QGIS and based on them, the CHM and the RGB and IR imagery was cropped to the size of the test area masks. When it was found that the minimal computable raster size (because of the restrictions of the 'raster' package) was a lot smaller than the maximum size of the actual ROI originally decided, the size of the ROI was reduced and the sizes of the test areas were defined by using the **'Clip raster by Extent'** tool, with ***'use map canvas'*** setting in QGIS, to 4 approximately 32 x 32 m test areas. Then the respective CHM test areas were used as masks/extents to clip the respective RGB and IR test areas in QGIS.

3.4 Segmentation of test area 1 (.R file: 2_segmentation_area_1.R)

The aim of this step is to find an accurate segmentation for test area 1 which can be tested on the other areas and then applied to the ROI.

The first task is to set verification points, called tree positions for CHM_1. First of all the minimum height of a tree has to be decided (which of course is region and area dependent). In the Alpine region we also have shrubs and Krummholz, which has to be distinguished from trees. If one also would like to consider the young trees, then the highest point of the shrubs has to be determined and the height for the trees has to be set above the highest shrubs. If one would be concentrating on grown trees, still the CHM in QGIS has to be checked and the minimum height of a tree has to be decided on that basis and this cannot be modified. The best is if one classifies the CHM in QGIS based on the tree heights and check with the RGB/IR

imagery where the shrubs 'finish' and the trees 'start.'

Fist load CHM_1 and the corresponding tree positions and plot them together:

Based on the classification of the CHM in QGIS it was decided for CHM_1, that h = 4 is a good height to get the lowest young trees but not to get involved with the shrubs and Krummholz.

With the function **TreeSeg** it is possible to find the minimal and maximal values for *a*, *b* (horizontal and vertical values (x, y) in meters) and *h* (height in meters) based on the CHM (and vp - verification points/tree positions). This means it is possible to test various settings beforehand and then inspect the results in QGIS. Feeding long variable sequences to **BestSegVal** can make the calculation time long and if one wants to compare what the different settings of the different variables do, it is easier to test it without sequences.
It can be tested if there is a need to use a *sum* filtered CHM (see the difference between **test_0** and **test_1**) or how big or small the variable *a*, *b* and *MIN* have to be set. Based on the fact that there are 15 trees in test area 1 we can estimate how accurate the segments might be. Taking a look at test 1, 2 & 3 we can see how the reduction of *a* and *b* elevated the number of trees.

Thus we can move on to function of **BestSegVal**, which works with value sequences for all variables and takes also the tree positions (vp) in account. In the first run we test the following values: **a=seq(0.05, 0.1, 0.01), b=seq(0.05, 0.1, 0.01), h=seq(3.5, 6, 0.5), MIN=seq(0.1, 0.5, 0.1)**, to go a bit lover and above the actual height of the trees and to determine the best *a*, *b* and *MIN* values. *CHMfilter* is set at 3, because we want to do only minimal filtering to close the gaps. In the first run we got 1080 results and 44 maximal segments. Because there are 15 trees in test area 1, the tibble was filtered to 20 segments (a bit above our desired count) and arranged according to the best hitrate (which was 0.733333) and still gave 668 results. These are still too much to decide which combination of settings should be used on the other areas. Thus the results were filtered to a **hitrate >= 0.7** and **height <= 4.0.** This resulted in 40 observations, from which the first 4 were actually calculated with *TreeSeg* and compared in QGIS. There was only minimal difference between **test_vp0** and **test_vp1**, based on the difference of 0.5 m height. It became clear that MIN (the size of the smallest segment) has to be set to 0.1.

Thus in the second run the following variables were set to a fix value or respective to the same

sequence: **MIN=0.1, h=seq(3.5, 6, 0.5)** and *a* and *b* were tested for a sequence between *0.01 and 0.1 with 0.01 steps*, to find better segment sizes. The second run (600 results) was filtered the same way (filtering to 20 segments lead to 204 results - the best hitrate being still 0.733333; the filtering to **hitrate >= 0.7** and **height <= 4.0**. lead to 8 observations) resulted in the same six results of **head(best_seg_vp_1_filt)** and two additional, so these 8 results will be inspected in the next step and only the second run will be displayed here. You can trace all steps in the respective .R files.

3.5 Cross-validation of test areas and Segmentation of ROI (.R file: 3_segmentation_validation_area_1_ROI.R)

The segmentation of test area 1 resulted in 8 possible segmentations which are going to be tested if they fit the other test areas. Up to some degree it is a kind of guess work to test the possibly best segmentation which fits the other test areas and also fits the ROI. That is why all 8 results from **best_seg_vp_1_filt** were processed. The 8 results contain also results with h=3.5 m, so it can be checked which height can really represent the whole ROI. After the visual check in QGIS it became clear, that h=3.5 m is too low because it may result in segmenting some of the shrub too. Of course this depends on the definition of tree height and if one wants to segment also the seedlings. Because of the similar height of the seedlings and shrubs it is practically impossible to distinguish shrubs and seedlings based purely on height.

While inspecting **segm_ROI -> segm_ROI_8** (demonstrating almost the same overall performance) it became clear that **segm_ROI_8** delivers the best result. But what happens with heights a bit above 4 m? 4.5, 5 and 5.1 m were tested to test if the segmentation can be even better than **segm_ROI_8**. We can see - as expected - that with the increase of the tree height less seedling and tree crown area is included. Thus the optimal result is **segm_ROI_8**. You can see the results of **segm_ROI_9** to **segm_ROI_12** in the respective script.

3.6 Data preparation for the Classification (.R file: 4_data_prep_classification:area_1.R)

For the preparation for the spectral classification first we will compute RGB and then multispectral (MS) indices, which we will filter and then test on correlation. Then 10 predictor sets

are built which will be tested in the next step. The indices are used to enhance the spectral differences between the different object classes which are to be detected. Currently 11 RGB and 4 MS indices are implemented. Additionally the filter enhance special properties of indices and the overall homogeneity of spectral areas and enhance the edges and between the different objects. During the testing it was understood, that there are certain indices which contain too homogeneous areas from the beginning thus filtering them would distort the values and they would not be kicked out by the correlation and would deform the prediction. Thus three different index stacks were complied: *RGB_1_ind (VVI, VARI, NDTI, RI, CI, BI,SI,HI, TGI, GLI, NGRDI) RGB_1_ind2 (VVI, NDTI, CI, BI, SI, TGI, GLI, NGRDI)* Which were then stack together to form the following data stacks: *ALL_1_stack (RGB_1_ind, RGBNIR_1_ind)* 15 layer *ALL_1_stack2 (RGB_1_ind2, RGBNIR_1_ind)* 12 layer *ALL_1_stack3 (RGB_1_ind, RGBNIR_1_ind, RGBNIR_1)* had been corrected by subtracting HI: 18 layer Different data stacks were tested to see if more data would filter out/overlay the shortcomings of certain indices. The homogeneity test filtered out the homogeneous indices, that is also VARI and RI, but HI had to be extracted separately.

4 Results

5 References

Allen, T. R., & Walsh, S. J. (1996). Spatial and compositional pattern of alpine treeline, Glacier National Park, Montana. *Photogrammetric Engineering and Remote Sensing (USA)*. https://scholar.google.com/scholar_lookup?title=Spatial+and+compositional+pattern+of+alpine+treeline%2C+Glacier+National+Park%2C+Montana&author=Allen%2C+T.R.+%28University+of+Vermont%2C+Burlington%2C+VT.%29&publication_year=1996

Bader, M. Y., Llambí, L. D., Case, B. S., Buckley, H. L., Toivonen, J. M., Camarero, J. J., Cairns, D. M., Brown, C. D., Wiegand, T., & Resler, L. M. (2021). A global framework for linking alpine-treeline ecotone patterns to underlying processes. *Ecography*, *44*(2), 265–292. https://doi.org/10.1111/ecog.05285

Bader, M. Y., Rietkerk, M., & Bregt, A. K. (2007). Vegetation Structure and Temperature Regimes of Tropical Alpine Treelines. *Arctic, Antarctic, and Alpine Research*, *39*(3), 353–364. https://doi.org/10.1657/1523-0430(06-055)%5BBADER%5D2.0.CO;2

Bader, M. Y., & Ruijten, J. J. A. (2008). A topography-based model of forest cover at the alpine tree line in the tropical Andes. *Journal of Biogeography*, *35*(4), 711–723. https://doi.org/10.1111/j.1365-2699.2007.01818.x

Baker, W. L., Honaker, J. J., & Wiesberg, P. J. (1995). *Using aerial photography and GIS to map the forest-tundra ecotone in Rocky Mountain National Park, Colorado, for global change research.* https://scholar.google.com/citations?view_op=view_citation&hl=en&user=kgaHKiwAAAAJ&alert_preview_top_rm=2&citation_for_view=kgaHKiwAAAAJ:qjMakFHDy7sC

Baker, W., & Weisberg, P. (1997). Using GIS to model tree population parameters in the Rocky Mountain National Park forest–tundra ecotone. *Journal of Biogeography*, *24*(4), 513–526. https://doi.org/10.1111/j.1365-2699.1997.00130.x

Barredo Cano, J., Mauri, A., & Caudullo, G. (2020). *Impacts of climate change in European mountains: Alpine tundra habitat loss and treeline shifts under future global warming : JRC PESETA IV project : Task 8.* (No. JRC115186). Publications Office of the European Union. https://data.europa.eu/doi/10.2760/653658

Bonanomi, G., Rita, A., Allevato, E., Cesarano, G., Saulino, L., Di Pasquale, G., Allegrezza, M., Pesaresi, S., Borghetti, M., Rossi, S., & Saracino, A. (2018). Anthropogenic and environmental factors affect the tree line position of Fagus sylvatica along the Apennines (Italy). *Journal of Biogeography*, *45*(11), 2595–2608. https://doi.org/10.1111/jbi.13408

Brown, D. G. (1994). Predicting Vegetation Types at Treeline Using Topography and Biophysical Disturbance Variables. *Journal of Vegetation Science*, *5*(5), 641–656. https://doi.org/10.2307/3235880

Chhetri, P. K., Shrestha, K. B., & Cairns, D. M. (2017). Topography and human disturbances are major controlling factors in treeline pattern at Barun and Manang area in the Nepal Himalaya. *Journal of Mountain Science*, *14*(1), 119–127. https://doi.org/10.1007/s11629-016-4198-6

Chhetri, P. K., & Thai, E. (2019). Remote sensing and geographic information systems techniques in studies on treeline ecotone dynamics. *Journal of Forestry Research*, *30*(5), 1543–1553. https://doi.org/10.1007/s11676-019-00897-x

Elliott, G. P. (2017). Treeline Ecotones. In D. Richardson, A. Castree, M. F. Goodchild, A. Kobayashi, W. Liu, & R. A. Marston (Eds.), *The International Encyclopedia of Geography* (pp. 1–10.). Wiley & Sons. https://doi.org/10.1002/9781118786352.wbieg0539

Fricker, G. A., Ventura, J. D., Wolf, J. A., North, M. P., Davis, F. W., & Franklin, J. (2019). A Convolutional Neural Network Classifier Identifies Tree Species in Mixed-Conifer Forest from Hyperspectral Imagery. *Remote Sensing*, *11*(19), 2326. https://doi.org/10.3390/rs11192326

Geping Luo, & Li Dai. (2013). Detection of alpine tree line change with high spatial resolution remotely sensed data. *Journal of Applied Remote Sensing*, *7*(1), 1–14. https://doi.org/10.1117/1.JRS.7.073520

Harsch, M. A., & Bader, M. Y. (2011). Treeline form – a potential key to understanding treeline dynamics. *Global Ecology and Biogeography*, *20*(4), 582–596. https://doi.org/10.1111/j.1466-8238.2010.00622.x

Harsch, M. A., Hulme, P. E., McGlone, M. S., & Duncan, R. P. (2009). Are treelines advancing? A global meta-analysis of treeline response to climate warming. *Ecology Letters*, *12*(10), 1040–1049. https://doi.org/10.1111/j.1461-0248.2009.01355.x

Holtmeier, F.-K., & Broll, G. (2005). Sensitivity and response of northern hemisphere altitudinal and polar treelines to environmental change at landscape and local scales. *Global Ecology and Biogeography*, *14*(5), 395–410. https://doi.org/10.1111/j.1466-822X.2005.00168.x

Holtmeier, F. K., & Broll, G. E. (2007). Treeline advance - driving processes and adverse factors. *Landscape Online*, *1*, 1–33. https://doi.org/10.3097/LO.200701

Immitzer, M., Atzberger, C., & Koukal, T. (2012). Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. *Remote Sensing*, *4*(9), 2661–2693. https://doi.org/10.3390/rs4092661

Kimball, K. D., & Weihrauch, D. M. (2000). Alpine vegetation communities and the alpine-treeline ecotone boundary in New England as biomonitors for climate change. *In: McCool, Stephen F.; Cole, David N.; Borrie, William T.; O'Loughlin, Jennifer, Comps. 2000. Wilderness Science in a Time of Change Conference—Volume 3: Wilderness as a Place for Scientific Inquiry; 1999 May 23–27; Missoula, MT. Proceedings RMRS-P-15-VOL-3. Ogden, UT: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. P. 93-101*, *015*. http://www.fs.usda.gov/treesearch/pubs/21974

Koerner, C. (2012). Was steuert das Pflanzenwachstum? *Biologie in Unserer Zeit*, *42*(4), 238–243. https://doi.org/10.1002/biuz.201210484

Körner, C. (1998). A re-assessment of high elevation treeline positions and their explanation. *Oecologia*, *115*(4), 445–459. https://doi.org/10.1007/s004420050540

Körner, C., & Paulsen, J. (2004). A world-wide study of high altitude treeline temperatures. *Journal of Biogeography*, *31*(5), 713–732. https://doi.org/10.1111/j.1365-2699.2003.01043.x

Král, K. (2009). Classification of Current Vegetation Cover and Alpine Treeline Ecotone in the Praděd Reserve (Czech Republic), Using Remote Sensing. *Mountain Research and Development*, *29*(2), 177–183. https://doi.org/10.1659/mrd.1077

Kupková, L., Červená, L., Suchá, R., Jakešová, L., Zagajewski, B., Březina, S., & Albrechtová, J. (2017). Classification of Tundra Vegetation in the Krkonoše Mts. National Park Using APEX, AISA Dual and Sentinel-2A Data. *European Journal of Remote Sensing*, *50*(1), 29–46. https://doi.org/10.1080/22797254.2017.1274573

Middleton, M., Närhi, P., Sutinen, M., & Sutinen, R. (2008). *Object based change detection of historical aerial photographs reveals altitudinal forest expansion.* https://www.semanticscholar.org/paper/OBJECT-BASED-CHANGE-DETECTION-OF-HISTORICAL-AERIAL-Middleton-N%C3%A4rhi/373b86df6f5979767ef6adbcfcc0b987fa37693f

Mishra, N. B., Mainali, K. P., Shrestha, B. B., Radenz, J., & Karki, D. (2018). Species-Level Vegetation Mapping in a Himalayan Treeline Ecotone Using Unmanned Aerial System (UAS) Imagery. *ISPRS International Journal of Geo-Information*, *7*(11), 445. https://doi.org/10.3390/ijgi7110445

Mohapatra, J., Singh, C. P., Tripathi, O. P., & Pandya, H. A. (2019). Remote sensing of alpine treeline ecotone dynamics and phenology in Arunachal Pradesh Himalaya. *International Journal of Remote Sensing*, *40*(20), 7986–8009. https://doi.org/10.1080/01431161.2019.1608383

Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G., & Nemani, R. R. (1997). Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature*, *386*(6626), 698–702. https://doi.org/10.1038/386698a0

Qiu, L., Jing, L., Hu, B., Li, H., & Tang, Y. (2020). A New Individual Tree Crown Delineation Method for High Resolution Multispectral Imagery. *Remote Sensing*, *12*(3), 585. https://doi.org/10.3390/rs12030585

Ranson, K. J., Montesano, P. M., & Nelson, R. (2011). Object-based mapping of the circumpolar taiga–tundra ecotone with MODIS tree cover. *Remote Sensing of Environment*, *115*(12), 3670–3680. https://doi.org/10.1016/j.rse.2011.09.006

Resler, L. M., Fonstad, M. A., & Butler, D. R. (2004). Mapping the Alpine Treeline Ecotone with Digital Aerial Photography and Textural Analysis. *Geocarto International*, *19*(1), 37–44. https://doi.org/10.1080/10106040408542297

Singh, C. P., Mohapatra, J., & Dharaiya, N. A. (2015). Remote Sensing of alpine Treeline Dynamics. *ISG Newsletter*, *21*(4), 3–8. https://www.researchgate.net/publication/283837476_Remote_Sensing_of_Alpine_Treeline_Dynamics

Stueve, K. M., Isaacs, R. E., Tyrrell, L. E., & Densmore, R. V. (2011). Spatial variability of biotic and abiotic tree establishment constraints across a treeline ecotone in the Alaska

Range. *Ecology*, *92*(2), 496–506. https://doi.org/10.1890/09-1725.1

Virtanen, T., Mikkola, K., Nikula, A., Christensen, J. H., Mazhitova, G. G., Oberman, N. G., & Kuhry, P. (2004). Modeling the Location of the Forest Line in Northeast European Russia with Remotely Sensed Vegetation and GIS-Based Climate and Terrain Data. *Arctic, Antarctic, and Alpine Research*, *36*(3), 314–322. http://www.jstor.org/stable/1552640

Weinstein, B. G., Marconi, S., Bohlman, S. A., Zare, A., & White, E. P. (2020). Cross-site learning in deep learning RGB tree crown detection. *Ecological Informatics*, *56*, 101061. https://doi.org/10.1016/j.ecoinf.2020.101061

Weinstein, B. G., Marconi, S., Bohlman, S., Zare, A., & White, E. (2019). Individual Tree-Crown Detection in RGB Imagery Using Semi-Supervised Deep Learning Neural Networks. *Remote Sensing*, *11*(11), 1309. https://doi.org/10.3390/rs11111309

Whiteside, T. G., Esparon, A. J., & Bartolo, R. E. (2020). A semi-automated approach for quantitative mapping of woody cover from historical time series aerial photography and satellite imagery. *Ecological Informatics*, *55*, 101012. https://doi.org/10.1016/j.ecoinf.2019.101012

Winings, C. J. (2013). Mapping alpine treeline with high resolution imagery and LiDAR data in North Cascades National Park, Washington. *WWU Graduate School Collection*. https://doi.org/https://doi.org/10.25710/9zkh-h696