Petra Miletich Bakk.rer.nat.

European large area automatic forest mapping based on random sampling points using Sentinel-2A

MASTER'S THESIS

to achieve the university degree of

Master of Science

Master's degree programme: Geospatial Technologies

submitted to

Graz University of Technology

Supervisor

Univ.-Prof. Dr.rer.nat. Dipl.-Forstwirt, Mathias Schardts

Institute for Geodesy

Graz, March 2017

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis.

Date

Signature

Abstract

The major motivation for small scale forest spectral discrimination and mapping covering large areas is to provide fundamentals for reporting and policy decision-making. The challenge, therefore, is to produce accurate, complete and reliable forest information on a variety of spatial scales over large extents. The aim of this master's thesis is to develop methods for the classification of the forest attributes tree cover density and forest type using Sentinel-2A image data and randomly sampled plots as trainings data. The new launched Sentinel-2 Mission provides users with high spatial resolution multispectral image data with an acquisition rate of 5 to 10 days. Additionally, four new red edge bands are introduced to improve vegetation status and types differentiations. The continuous tree cover density prediction is performed with regression models, whereas the discrete forest types are classified into the specific classes 'broadleaved', 'coniferous' and 'mixed'. Regarding both classifications automated machine learning methods like Random Forest and K-Nearest Neighbours are applied and compared. To allow cost effective large area applicability visually interpreted sample plots are used, which are already available at European level. Independent validation is performed to access the accuracy. The tree cover density prediction shows moderate better results than the forest type differentiation, indicating that the sampled plots used as training data are not perfectly suited for forest type classification.

Kurzfassung

Die Erfassung von forstlichen Parametern dient als Grundlagen für politische Entscheidungsprozesse. Die Herausforderung liegt darin, genaue Informationen mit einem hohen Automatisierungsgrad erzeugen zu können. Ziel dieser Arbeit ist es die Möglichkeiten der Klassifikation von forstlichen Parametern, wie Waldtyp und Walddichte mithilfe von Sentinel-2A Szenen und zufallsgenerierten Stichprobenpunkten zu testen. Die vor kurzem gestartete Sentinel-2 Erdbeobachtungsmission bietet den Nutzern hochauflösende multispektrale Bilder in einem Akquirierungszeitraum von 5 bis 10 Tagen, mit zusätzlichen Red Edge Bändern, welche vor allem im Bereich der Vegetationsanalyse Verbesserungen bringen sollen. Für die Klassifikation der Baumdichte werden Vorhersagemodelle über Regressionsverfahren trainiert. Die Klassifikation der Waldtypen erfolgt über das Training von Klassifikationsmodellen. Bei beiden Klassifikationsmethoden werden europaweit zufällig generierte Stichprobenpunkten als Trainingsdaten verwendet. Innerhalb der Klassifikationen werden automatisierte *machine-learning* Methoden, wie Random Forest und k-Nearest Neighbour getestet und verglichen. Für die Klassifikation der Baumdichte können moderat bessere Ergebnisse als bei den Waldtypen erzeugt werden. Die Klassifikation der Waldtypen weist noch Verbesserungspotential auf, da sich die heterogenen Trainingsdaten nicht gut für die Differenzierung der Waldtypen eignen. Die Reduktion der Trainingsdatensätze auf homogene Gebiete erzielt bereits bessere Klassifikationsergebnisse.

Contents

List of Figures

List of Figures

List of Tables

List of Tables

List of Acronyms

EEA ..... European Environment Agency
LAEA ..... Lambert Azimuthal Equal Area
FAO ..... Food and Agriculture Organization of the United Nations
LUCAS ..... Land use and land cover survey
EEAVAL ..... Validation Services for the geospatial products of the Copernicus
..... land Continental and local components including in-situ data
HR ..... High Resolution
VHR ..... Very High Resolution
HRL ..... High Resolution Layer
TCD ..... Tree Cover Density
FTY ..... Forest Type
BOA ..... Bottom of Atmosphere
TOA ..... Top of Atmosphere
BRDF ..... Bidirectional reflectance distribution function
AOT ..... Aerosol Optical Thickness
WV ..... Water Vapour
APDA ..... Atmospheric Pre-corrected Differential Absorption
NDVI ..... Normalized Difference Vegetation Index
NDSI ..... Normalized-Difference Snow Index
SSA ..... Single scattering albedo
LUT ..... Look up table
SAFE ..... Standard Archive Format for Europe
OLI ..... Operational Land Imager

List of Tables

MSI ..... Multi Spectral Instrument

MODIS ..... Moderate-resolution Imaging Spectroradiometer

DU ..... Dobson Unit

GR ..... Ground resolution

VRE ..... Vegetation Red Edge

MMU ..... Minimum Mapping Unit

MMW ..... Minimum Mapping Width

BC ..... Band combination

CL ..... Classifier

TS ..... Training Samples

RF ..... Random Forest

kNN ..... k Nearest Neighbours

m10 ..... Mosaic with 10 m ground resolution

m20 ..... Mosaic with 20 m ground resolution

t ..... tree parameter

n ..... node parameter

s ..... sample parameter

d ..... depth parameter

N ..... number of training samples

NP ..... non-parametric

NM ..... non-metric

PA ..... parametric

P ..... Producer Accuracy

U ..... User Accuracy

B ..... Broadleaved Forest

C ..... Coniferous Forest

M ..... Mixed Forest

1. Introduction to European Forest Mapping

Information on spatial forest distribution provided by forest mapping is needed as fundamentals for climate change related agreements, reporting duties, forest policy-making as well as forest ecosystem management and other forest related applications (Laurin et al., 2016; Baffetta et al., 2012). "Forest policy and management in Europe are under the direct or indirect influence of a significant number of processes and organisations. Within the EU, forestry matters are addressed, for instance, within the council Working Party on the Forestry and the Standing Forestry Committee of the European Commission"(Michalak, 2014). Those organisations benefit from the most up-to-date forestry information as fundamental for their recommendation. Furthermore most of the international forest-related agreements are depended upon the reporting. The first forest international and European forest agreements have laid focus on the relationship between forest and economy. With time, various information on other forest related function and variables (e. g. forest health, biodiversity) are gathered to support the forest management process (Michalak, 2014).

In the same way K. Jia et al. (2014), X. Zhu and Liu (2014), Moreno et al. (2016), and Masek et al. (2015) highlight the importance of better understanding of forest ecosystem processes for economy, environment and society. After all forests play a major role in global ecosystems, affecting biodiversity, air pollution, water availability as well as quality, energy balance, carbon dynamics and climate regulations. Besides forest have suffered from large scale disturbances and mortality due to climate change, forest pest, forest fires and anthropogenic influences like clear-cuts. Therefore forestry aims for a sustainable management of forest resources. In order to be able to optimally exploit forest resources, accurate information regarding extent, condition and productivity over large areas is required constantly by forest industry and policy (Baffetta et al., 2012; Hyyppä et al., 2000).

There are several efforts of forests mapping at different scales, differentiating between monitoring forest resources at small scale, covering large areas, to estimate spatial and temporal variations and economic planning for smaller homogeneous forest areas or forest ecosystems to search for cause-effect relationships

(Hyyppä et al., 2000; Michalak, 2014). Since regional and small scale mapping products vary in spatial resolution, input data and forest definition, the importance of forest observation and information providing at European scale has been pointed out by Probeck et al. (2014):

> *"Consistent, accurate, reliable and up-to-date information on the state of forests in Europe is required by European countries for reporting and policy making in the frame of several European countries and international forest- and environment-relate policies, action plans and international agreements in the fields of environmental protection, protection of biodiversity and ecosystems, conservation planning, sustainable use of natural resources, climate change mitigation actions and environmental modelling [...]" (Probeck et al., 2014).*

1.1. European Forest Products

Developments in recent years provide several forest-related information on global and European scale, varying in detail, source of information and target groups. One of the major products is CORINE Land Cover from 1990, 2000, 2006 or 2012 representing, among other land cover types, forest cover information. Within the forest class following forest types are differentiated: Coniferous, broadleaved and mixed forests along with agroforestry and transitional woodland. Forest areas smaller than 25 ha are not included (Probeck et al., 2014; Moreno et al., 2016).

A global Remote Sensing Digital Map derived from AVHRR images has been produced in 1992. Images with 1 km spatial resolution are classified into 17 classes, including the five forest classes, evergreen coniferous forests, evergreen broadleaved forest, deciduous coniferous forests, deciduous broadleaved forests and mixed forests. The map has been updated for the years 1997 / 1998 (Probeck et al., 2014; Schuck et al., 2002).

Moreover, the manually derived European Community Forest Map of 1987 has been updated using Corine Land Cover and a Non-Forest / Forest map of 1992 derived from the ESA AVHRR forest map, which has been completed in 1997 (Probeck et al., 2014; Schuck et al., 2002).

Furthermore, the ESA has produced a global land cover classification within the Glob Cover project for two periods: December 2004 - June 2006 and January - December 2009. The product is based on 300 m MERIS data representing 20 land cover classes independent of national borders. Within those land cover classes broadleaved and coniferous forest information is spread over 12 classes (Probeck et al., 2014; ESA, 2016).

Additionally, the Forest Cover Map 2000, 2006 and 20011/12 and the Forest Type Map 2006 are provided by the forest group of the Joint Research Centre. The forest/non-forest information is derived from LANDSAT ETM+ imagery and CORINE Land Cover 2000, provided with 25 m raster resolution and validated with independent survey points. The first map from 2000 provides the user with forest / non-forest informations covering EU25 countries. Updated versions differentiate between broadleaved and coniferous forest types. Automatic classification techniques based on clustering algorithms are used as mapping approach (Probeck et al., 2014; JRC, 2015).

Finally, the most recent product published from the ESA is the GIO high resolution (HR) Forest Layer from 2012, created from automated classification methods. The objective of the HR forest layer product is to map forest cover characteristics such as tree cover density and forest type, for reporting. Both mapping characteristics are provided with 20 x 20 m pixel resolution and aggregated to 100 x 100 m product, with a minimum mapping unit from 0.5 or 1 ha (Probeck et al., 2014; Langanke et al., 2015).

1.2. Forest Mapping Approaches based on sampling

Many studies focus on the integration of remotely sensed data and forest inventory plots for mapping forests over large areas due to the fact that the most frequent forest information providers are the national inventory organisations. The forest inventories provide details like tree type, diameter at breast height, height, age, and damage, gathered from field measurements. Through detailed forest inventories, information about the state and dynamics of forests and for forest ecosystem management can be derived (Gjertsen, 2007; Thessler et al., 2008; Pippuri et al., 2016; Laurin et al., 2016; Baffetta et al., 2012; Ohmann et al., 2014).

Accordingly, forest inventories can be exploited as ground truth data for the classification process, representing trainings or validation samples. Nevertheless, Moreno et al. (2016) state that consistent inventory plots at the European scale are not openly available as by now. Furthermore, national forest inventory plots are suboptimal regarding large area forest mapping approaches, as they are thematically, spatially and temporally inconsistent (Probeck et al., 2014; K. Jia et al., 2014; Michalak, 2014). Researchers have to obtain the plots from each country individually and subsequently aggregate them to the required spatial and thematic level. To meet the demand for coherent forest information, the production of harmonised data from mapping approaches beyond national borders is necessary (Michalak, 2014). Even, Moreno et al. (2016) argue that by combining remotely sensed image data with inventory plots, factors like spatial variability, sampling

techniques and time, need to be considered. Those factors affect the confidence in the plot data, thus leads to reduce the reliability of the produced information.

As an alternative to forest inventory data independent survey points could be considered for forest mapping applications. The European Union Land Use/Cover Area frame Survey (LUCAS) provides statistics on land cover across EU-28 countries based on ground observations. The sample data collected between March and September 2012 provides harmonised information for many applications, including forestry. The three forest classes differentiated are broadleaved, coniferous and mixed forest (EUROSTAT, 2015). LUCAS points do not provide as detailed information on forests as national forest inventory plots, but serve land use and land cover information purposes. Aside from field measurements independent survey points can be gathered by visual interpretation, providing an opportunity for forest cover characteristics classification. Moreno et al. (2016) mention that with exactly located plots high resolution satellite data, such as LANDSAT, SENTINEL, SPOT or LIDAR information, can be considered within the training process as ground truth, in order to create products with higher spatial resolution.

1.3. Forest Mapping Expectations with SENTINEL

The existing pan-European mapping approaches are based on moderate or coarse resolution data representing the major data source for mapping forests at global or regional scale. Those products do not fulfil the requirements for local and inter-regional forest applications (Masek et al., 2015). However, Laurin et al. (2016) point out the need of more detailed information with high geometric resolutions, concerning the forest characteristics, forest type, tree cover density, productivity and health conditions. Recent studies focus on receiving a more detailed characterization of forests. With high spatial resolution satellite data covering large areas, forest information with a high thematic accuracy can be produced, suitable for repeated monitoring over large areas. Furthermore, remotely sensed data can be used to derive information on forest productivity, cover type and deforestation (Probeck et al., 2014; Bruzzone and Demir, 2014; Moreno et al., 2016).

In respect of the demand to generate reproducible, accurate, complete and reliable forest information at large scale, high resolution satellite images covering large areas are necessary (Musaoglu and Örmeci, 2000; Probeck et al., 2014). At this point it needs to be mentioned that the use of high spatial resolution images decreases the mixed pixel problem, while the class intern variability increases, resulting in a lower spectral variability between different classes. Accordingly, the classes are statistically worse separable in the spectral feature space, which may results in lower classification accuracies (Bruzzone and Demir, 2014). On

the other hand, using satellite images with high geometric resolution to identify forest attributes, allows the usage of textural information to increase the mapping accuracy (X. Zhu and Liu, 2014).

The launch of the Multispectral Image Sensor (MIS) on-board SENTINEL-2 with 10 m spatial resolution and three new bands in the red edge region, provided by the Copernicus initiative represents new opportunities to overcome limitations of coarse resolution data. (Hirschmugl et al., 2017; Laurin et al., 2016). Better compensation of atmospheric influences and cloud masking could be accomplished by the 13 bands available (Probeck et al., 2014). "The three S2 red edge bands are especially promising for their ability to detect fine differences in chlorophyll pigments; higher chlorophyll content can indicate higher canopy density [...]" (Laurin et al., 2016). In addition, SENTINEL-1 supplements forest management with the ability for "clear-cut and partial-cut detection, forest type classification, biomass estimation and disturbance detection" (ESA, 2015).

1.4. Study Objectives

The aim of this thesis is to introduce a methodological framework regarding forest type and tree cover density classification methods using SENTINEL-2A data and already available sampling plots as trainings data. The random sampling plots, provided by Joanneum Research, are collected through visual image interpretation carried out in the frame of the EEAVAL (Validation Services for the geospatial products of the Copernicus land Continental and local components including in-situ data) project. The HR Forest Layers provided by Copernicus are produced "through a combination of automatic processing and interactive rule based classification" (Sannier et al., 2015) for the reference year 2012 mainly based on 2011/2012 with 20 m resolution satellite images (Langanke et al., 2015). The validation process is based on a stratified random sampling process and a visual interpretation of 100 by 100 m plots with high resolution images (Sannier et al., 2015).

The major aim of this thesis is to investigate to what extent the provided plots reused as training areas are suitable for a pan-European classification approach based on SENTINEL-2. In this case the major advantage will be that the proposed method avoids time-consuming efforts to collect training areas within each satellite image and thus reaching a higher automation degree. Concerning the study objective following aspects will be considered and analysed. Regarding the SENTINEL-2 image data preprocessing, the recommended Sen2Cor processor presents a combined approach for atmospheric and topographic corrections. It will be investigated to what extent the Sen2Cor processor will provide satisfactory results regarding the SENTINEL-2 preprocessing. The resulting Level-2A

products will be compared to two sensors with similar spectral resolution to check the consistency. Furthermore, the potential of the new introduced red-edge bands will be examined regarding the forest type discrimination and the tree cover density estimation using signature and separability analysis. Additionally, the effects of several different band combinations on the classification accuracy will be investigated within the classification process to find the most suitable band combination. Moreover, the role of the two automated machine learning classifiers Random Forest and K-Nearest Neighbours will be examined, comparing the classification accuracies. Therefore, the continuous tree cover density prediction will be performed with regression models. Whereas the discrete forest types will be classified into three specific classes 'broadleaved', 'coniferous' and 'mixed'. With regards to both classifiers the impact of the parameters used within the training process on the classification accuracy will be assessed.

2. Automated Image Classification Methods

Image classification provides the basis for image analysis and pattern recognition considering various remote sensing applications including forest mapping (Lillesand et al., 2014; Campbell, 2002; Al-Doski et al., 2013; Rees, 2001). A wide range of automatic image classification algorithms has been developed for many different purposes most of them, suited for remote sensing applications. These algorithms for automatic image classification procedures are called classifiers and can be applied to single band images or to multispectral images (Mountrakis et al., 2011; Rees, 2001). They range from supervised to unsupervised, parametric to non-parametric, hard to soft, pixel-based to subpixel-based or object-based approaches as shown in table 2.1. The appropriate classifier must be selected for each specific task, considering image characteristics and the analysis purpose (Liaw and Wiener, 2002; Campbell, 2002; Al-Doski et al., 2013).

2.1. Supervised, Unsupervised and Hybrid Classification Logic

As mentioned in the previous chapter, classification methods can be distinguished by several characteristics. Generally they are based on either supervised or unsupervised classification logic (Rees, 2001; Kasischke et al., 2004).

In case of **unsupervised classification** methods prior knowledge of the region is not required. Pixels sharing spectral reflectance characteristics are associated with spectral clusters. The basic premise is that values of the same class are closer together in the feature space. Accordingly, they share similar reflectance values (Lillesand et al., 2014; Jensen et al., 2009; Al-Doski et al., 2013; Kasischke et al., 2004; Mather and Magaly, 1999). Rees (2001) and Campbell (2002) state that the major advantage of unsupervised methods is the minimization of human error through full automatic procedures, thus the user only has to specify the number of classes. The limited user control can be seen as disadvantage, due to homogeneous spectral classes within the data, which do not correspond to

Table 2.1.: Satellite image classification methods

	Pixel-based / Object-based classifiers		Sub-pixel-based classifiers		Segmentation algorithms
Supervised	Maximum Likelihood	PA	fuzzy classification	NP	
	Minimum Distance-to-Means	PA	Spectral Mixture Analysis	NP	
	Mahalanobis Distance	PA	Neural Networks	NP	
	Paralellepiped	PA	Regression Modeling	NP	
	k-Nearest Neighbour	NP	Regression Tree Analysis	NP	
	Artificial Neural Networks	NP	Fuzzy-spectral Mixture Analysis	NP	
	Decision Tree	NM			
	Support Vector Machine	NP			
	Random Forest	NM			
	Linear Discriminant Analysis	PA			
Unsupervised	k-means				Region-Growing
	Iterative Self-Organizing Data Analysis (ISODATA)				Markovian Methods
	Self-Organizing Maps (SOM)				Watershed Methods
					Hierarchical Algorithms
	NP = non-parametric		PA = parametric		NM = non - metric

thematic classes (Rees, 2001; Campbell, 2002). The clustering process results only in spectral clusters without thematic labels. Consequently, the analyst needs to label and combine the clusters into thematic classes. Therefore, the human error is still an issue to be considered (Lillesand et al., 2014).

Unlike the unsupervised classification methods the **supervised techniques** require prior information about the thematic classes. These samples of known identity are defined by the analyst and should be homogeneous and representative for each feature type of interest. These representative training sites may be derived from fieldwork, visual interpretation of aerial photos / VHR satellite images or relevant maps (Mather and Magaly, 1999; Rees, 2001; Lillesand et al., 2014). The trainings data preparation can be time-consuming and expensive. Nevertheless, the advantage of supervised classification processes are the resulting thematic classes. The method is often used for quantitative analysis and a variety of algorithms are available (Campbell, 2002).

Hybrid classifiers share characteristics from both unsupervised and supervised classification methods. The advantage of combining the methods is that spectral clusters can be efficiently determined with unsupervised classification methods and with available ground truth data these spectral classes can be compared to the training data. After the clustering supervised methods perform the classification (Richards and X. Jia, 1999; Rees, 2001; Jensen et al., 2009; Al-Doski et al., 2013).

Considering the independent survey points as trainings data, which are already available at European level, supervised classification methods are taken into account regarding the forest mapping approach in this thesis.

2.2. Pixel-based and Object-based Methods

Another kind of distinction in image classification differentiates between object-based and pixel-based classification methods. Generally, classification methods can make use of spectral or spatial pattern recognition approaches. The basis for the pixel-based spectral pattern recognition is the spectral information of the pixel itself. The spectral information of surrounding pixels are not taken into consideration. **Pixel-based classifiers** automatically associate each pixel with a thematic class using the spectral pattern recognition (Lillesand et al., 2014; Campbell, 2002; Mather and Magaly, 1999).

Spatial pattern recognition approaches categorize pixels or groups of pixels on basis of their spatial relationship. Parameters like texture, shape, size and context are taken into account. **Object-based classifiers** combine both spectral and spatial reflectance pattern recognition methods, assuming that the image contains homogeneous segments, which are larger than the pixel size (Lillesand et al., 2014; Jensen et al., 2009). "With object-based models, geographical objects, instead of individual pixels, are considered as the basic unit" (M. Li et al., 2014). First, image segmentation is used to create homogeneous objects using the spatial, spectral, textural and contextual information from the input data. Second, these objects are classified into thematic classes. Object-based image classification techniques are embedded in various algorithms like Region Growing, Markovian methods, Watershed methods and Hierarchical methods. The majority of researcher consider these methods more appropriate with high-spatial resolution image data compared to pixel-based classifiers (M. Li et al., 2014; Lillesand et al., 2014; Jensen et al., 2009; Campbell, 2002). Considering the spatial resolution of Sentinel-2a image data with 10 or 20 m and the mapping subjects, represented by tree patches which are larger than pixel size, object-based methods are suboptimal regarding large area forest mapping applications.

2.3. Hard and Soft Classifiers

Pixel-based classifiers are further classified into two categories: Hard- and soft classifiers. **Hard classifiers** assume that only one thematic class exists within each pixel and each individual pixel is given a single label. This assumption

is often wrong for data with mixed pixels. Such pixels occur when the feature target scale is smaller than the pixel size (Lillesand et al., 2014; Al-Doski et al., 2013; Mather and Magaly, 1999). M. Li et al. (2014) and Al-Doski et al. (2013) pointed out that the most commonly used unsupervised pixel-wise classifiers are k-means, Self-Organizing Data Analysis technique and the Self-Organizing Maps method. Furthermore, large variations of supervised pixel-based classification methods have been developed including Gaussian Maximum Likelihood, Minimum Distance-to-Means, Mahalanobis Distance, Parallelepiped, K-Nearest Neighbours, Artificial Neural Network, Decision Tree, Random Forests, Support Vector Machine, and Genetic Algorithms (Mather and Magaly, 1999; Abburu and Golla, 2015; Al-Doski et al., 2013; Grinand et al., 2013; Lillesand et al., 2014; Hassanat et al., 2014; Liaw and Wiener, 2002).

As for the classification process mixed pixels represent a significant problem, **soft classifiers** are developed. These classifiers assign each pixel with a measure of probability for each thematic class and work on sub-pixel scale. Consequently, it leaves the decision for the final class to the analyst (Mather and Magaly, 1999; Al-Doski et al., 2013). Furthermore, it should be considered that a thematic class decision is not required if a proportional or fractional evaluation is appropriate. Several major sub-pixel classification techniques are available such as Fuzzy Classification, Neural Networks, Regression Modelling, Regression Tree Analysis, and Spectral Mixture Analysis (Al-Doski et al., 2013; Lillesand et al., 2014; Abburu and Golla, 2015; Al-Doski et al., 2013). Soft classifiers are often used for mapping approaches at finer resolution than the pixel size of a sensor. With hard classifiers such sub-pixel targets, for example fractions of land cover classes within one pixel, remain undetected (Radoux et al., 2016).

2.4. Parametric, Non-metric and Non-parametric Methods

Another kind of distinction in image classification is defined by the statistics they are based on. For most of the classifiers assumptions are made ”concerning the shape and distribution of decision volumes in multispectral space” (Kasischke et al., 2004). Algorithms based on parametric statistics, assume image data to be normally distributed, like Maximum Likelihood, Mahalanobis Distance and various others as summarized in table 2.1. However the training data for supervised image classification do not always fulfil these assumptions (Jensen et al., 2009; Kasischke et al., 2004). Therefore, non-parametric algorithms are needed, which are not based on any data distribution premise. Most common non-parametric algorithms are artificial neural networks, support vector machine and various sub-pixel based classifiers (Jensen et al., 2009; Al-Doski et al., 2013). Alternatives

are Non-metric algorithms which combine both approaches and can operate on real valued data and nominal scaled data. The most common non-metric algorithms are based on decision tree approaches also including the random forest algorithm (Jensen et al., 2009; Kasischke et al., 2004).

Within this thesis forest mapping approach the usual parametric classification algorithm like maximum likelihood may not be appropriate regarding the forest type and tree cover density differentiation. The ground truth data is primarily assessed for validation purposes, consequently, the 100 by 100 m plots are characterised by heterogeneous signatures of non-forest and forest areas. Additionally, the Non-Forest areas are more frequent within the original ground truth data set, representing an unbalanced trainings data set. Accordingly, Non-parametric and non-metric classifiers, operating on real valued data and nominal scaled data, are more suitable for the problem statement of this work, than parametric approaches. Thus, to overcome problem of heterogeneous trainings data machine learning techniques are considered within this thesis as they do not rely on data distribution assumptions, can handle noisy observations and can easily be applied to large datasets.

3. Automated Forest Mapping Approaches

As mentioned in chapter 1, remotely sensed imagery can be exploited to estimate forest characteristics like type or density and detailed forest attributes like heigh, volume or biomass, producing maps for large regions of interest (Laurin et al., 2016). Lately, several different approaches for forest mapping have been developed that are based on different scale, classifiers, forest attributes as well as forest definitions, ancillary data and a wide variety of sensor data. The choice of a suitable approach depends on the study objectives and the conditions mentioned above, as well on, the level-of-detail in training data, if available (Campbell, 2002).

For the selection of the appropriate classification method, mapping scale respectively the geometric resolution of the data should be taken into consideration, as object-based approaches work better with high-spatial resolution data, due to the visibility of meaningful objects in high resolution images. Consequently, in this case the use of pixel-based classification concepts are suboptimal. To overcome this problem segments are created and classified instead of pixels (Ohmann et al., 2014; Lillesand et al., 2014; Campbell, 2002; Rees, 2001).

Ohmann et al. (2014) tested a canopy cover estimation with LANDSAT imagery and inventory plot data using the k Nearest Neighbours classifier. The Canopy Cover defined as percentage of all living trees is mapped considering different aspects of spatial scale like the k value or the mapping unit. Across all scaling options a map accuracy for canopy cover with an average RMSE of 0.24 could be achieved (Ohmann et al., 2014). Furthermore, the eleven discrete vegetation classes sparse, open, broadleaf (small, medium/large), mixed conifer-broadleaf (small, medium, large/very large) and conifer forest (small, medium, large, very large) are classified with an overall accuracy ranging from 39 - 52 % across the scaling options. Ohmann et al. (2014) propose that the accuracy assessment resolution strongly influenced the map accuracy. In other words, map reliability assessed on larger areas is more accurate than on plot-level. Furthermore, the value of k strongly influenced the map accuracy. Nevertheless, the optimal k value is depended on the objective of the study. On the contrary, the effects of the spatial resolution and heterogeneous plots on the overall accuracies are small,

however, the exclusion of heterogeneous plots from the training data slightly improved the map accuracies (Ohmann et al., 2014).

Furthermore, with high resolution data using textural information within the classification process can be taken into account. Several studies have proven that adding spatio-contextual information in the classification process can improve the accuracy for forest mapping approaches (M. Li et al., 2014; Ferreira et al., 2016). The supervised machine learning classifiers, Linear Discriminant Analysis, Support Vector Machine and Random Forest are compared by Ferreira et al. (2016), using hyperspectral and simulated multispectral data from the WorldView-3 sensor to map tree species in Brazil. The authors claim that the Linear Discriminant Analysis outperformed the other methods and produced the best result with the inclusion of the SWIR bands, providing an accuracy of 74,8 %. Moreover, an object-based approach was tested to identify the crowns, providing an accuracy of 84,9 % based on the multispectral data. Therefore, it seems that for tree species mapping at small scale with high resolution data object-based approaches work better than pixel-based approaches (Ferreira et al., 2016).

In contrast other studies have proven that object-based approaches applied on coarse resolution data, do not increase the accuracy results (Thessler et al., 2008; Dorren et al., 2003). Forest type mapping in rain forest regions, achieving good results, is demonstrated by Thessler et al. (2008), applying k-nearest-neighbour and Linear Discriminant Analysis approaches on LANDSAT TM and SRTM elevation model. 104 field observation plots are used to classify three different forest types with an overall accuracy ranging from 82 - 91 %. However, the discrimination of old growth forests does not work adequately and results in low kappa values. One of the more significant findings to emerge from this study is that a prior segmentation process does not increase the classification accuracy significantly (Thessler et al., 2008).

Nevertheless, pixel-based classifiers still play a major role due to the fact that they are almost fully automatic, easy to apply and widely available in remote sensing image processing software programs or packages. Several studies focused on the integration of coarse resolution satellite data with forest inventory plots, achieving good results for forest attribute mapping at small scale (Xiao et al., 2002; Baffetta et al., 2012; Dorren et al., 2003; Ohmann et al., 2014; X. Zhu and Liu, 2014). The potential of SPOT-4 VEGETATION data for forest mapping, using temporal composite data datasets is demonstrated by Xiao et al. (2002) using 1 km spatial resolution SPOT data regarding forest type discrimination in Northeaster China. The unsupervised Iterative Isodata Clustering classifier and cluster labelling through experts are used to estimate forest types "including evergreen needleleaf forest, deciduous needleleaf forest, deciduous broadleaf forest, and mixed forests" (Xiao et al., 2002). Between derived forest maps and a national land cover dataset used as reference, a spatial agreement in forest

pixels about 70 % (Error of omission: 30 %, Error of Commission: 10 %) is achieved. This study has shown that automated unsupervised techniques also produce moderate results and should be taken into account regarding forest mapping approaches. Furthermore, Immitzer et al. (2016) have observed a clear trend between accuracy and spatial resolution. The model based on spatial high resolution data achieved with 74 % the best accuracy results in a tree species classification process. In high resolution image data the mixed-pixel problem decreases and furthermore, textural information is available to supplement the classification process (Immitzer et al., 2016).

Besides, the spectral resolution is one factor to be considered regarding forest mapping approaches. Hyperspectral data allows the retrieval of biophysical parameters which can be associated with vegetation functions for detailed forest type mapping (Laurin et al., 2016; Masek et al., 2015). Maximum likelihood and Support Vector Machine approaches are compared by Laurin et al. (2016) regarding tree species discrimination in tropical forest using simulated Sentinel-2 data. They conclude that the forest type discrimination works well with the simulated Sentinel-2 data, adding texture information to the classification process. The overall accuracy could slightly be improved by using the Support Vector Machine approach (92.34 %) over the Maximum Likelihood approach (90,73 %). Another dominant tree species mapping is performed by Gjertsen (2007), using field plots, land cover maps and LANDSAT TM images. The inventory plots are used to train the non-parametric Knn-classifier, weighting the plots according to their similarity of spectral pixels values in the inventory area. The accuracy of the dominant tree species classification has been tested on two different data sets at municipal scale, with five trials resulting in a accuracy ranging from 58,3 % to 63,1 %. Gjertsen (2007) implies that an accuracy improvement could be achieved by using different band ratios and correction for terrain illuminations. The spectral resolution of remotely sensed image data should be considered, since various studies used multispectral and hyperspectral data for vegetation mapping achieving good results (Gjertsen, 2007).

Beside the spatial and spectral resolution it is worthwhile to consider ancillary data as a factor to improve the classification process. Various studies have shown that adding vertical information can improve forest mapping results (Pippuri et al., 2016; Gjertsen, 2007; Dorren et al., 2003). Forest attribute mapping using laser scanning (Optech ALTM), satellite images (LANDSAT 5 TM) and sample plot data to estimate forest types and forest land attributes like volume, mean height and diameter is tested by Pippuri et al. (2016). Multinomial Logistic Regression Analyses is used for the classification process resulting in an overall accuracy of 0.90 % (Kappa: 0.37) for the main forest type classification. These findings suggest that in general forest attributes can be accurately classified using Lidar Data, remotely sensed images and forest inventory plots and results can

be improved by additional vertical information. Object-based and pixel-based approaches for forest cover mapping are compared by Dorren et al. (2003), using winter and summer LANDSAT TM 5 scenes and inventory plots. The main aim of the study is to compare Maximum Likelihood as pixel-based approach with Region Growing as segmentation method for a object-based approach. Another aim of the study is to estimate the impact on topographic normalisation and the usage of a digital elevation model on the mapping accuracy. Including a digital elevation model produced the highest overall accuracy with 73 %. "The importance of the DEM for forest mapping in steep mountainous terrain could be explained by the fact that the distribution of forest stand types on regional scale is mainly determined by altitude" (Dorren et al., 2003). The input inventory plots differentiate between following forest types: open coniferous forest, dense coniferous forest, broadleaved coniferous forest and mixed forest. The object based method achieved the best results (Overall accuracy: 70 %) with the usage of the digital elevation model, whereas the pixel-based methods achieved an overall accuracy of 73 %. The slightly higher accuracies for the object-based approach imply that there is no significant improvement through the prior segmentation process (Dorren et al., 2003).

Moreover, with large and free accessible archives like LANDSAT time series can be build up. There is evidence that time series have a positive effect on classification accuracy due to the fact that considering phenological information of forests can improve the discrimination of forest types sharing similar spectral behaviour over time (X. Zhu and Liu, 2014). The authors demonstrate a method for forest type discrimination based on LANDSAT time series and topographic information. The forest types pine forest, oak forest and mixed forest are discriminated. For the training and validation process ground plot data is used. The forest type discrimination works with an overall accuracy of 90,52 %, using Support Vector Machine approach. To improve the accuracy a Digital Elevation Model is integrated achieving an accuracy of 92,63 %. This study has shown that with the usage of the digital elevation model the accuracy can be slightly improved for mono-temporal image data, whereas the usage of dense seasonal time series can largely improve the accuracy (X. Zhu and Liu, 2014).

With the newly launched Sentinel missions another free data source is available. Considerably more work will have to be done to determine the potential of Sentinel data regarding detailed forest attribute mapping and large area forest characteristics mapping. Further research might explore the potential of the new red edge band of Sentinel-2 for forest mapping approaches.

4. Data Source

4.1. SENTINEL-2A

SENTINEL-2A images obtained from the Copernicus Scientific Data Hub are used in this thesis. The SENTINEL-2 Missions is part of the European environmental monitoring system COPERNICUS and will provide two passive optical earth observation satellites. SENTINEL-2A has been launched in June 2015 and SENTINEL-2B has been launched in March 2017 (ESA Earth Online, 2016; Pflug et al., 2016; Malenovsky et al., 2012). The satellites are equipped with the high resolution Multispectral Instrument (MSI) with 13 spectral bands in visible, near-infra-red (VNIR) and short waved infra-red (SWIR). Furthermore, three new bands in the red edge region located in the transition between red and near infra red, B5 (698 - 713 nm), B6 (733 - 748 nm) and B7 (773 – 793 nm), are introduced. The mission aim is to support different land monitoring services with a high revisit frequency, a systematic global land surface coverage of satellite images with high spatial resolution and enhanced spectral range (Fletcher K., 2012; Main-Knorn et al., 2015; Suhet, 2015; Gatti A., 2013; Delegido et al., 2011; P. Martimort et al., 2007; Drusch et al., 2012).

4.1.1. Characteristics

Both satellites, SENTINEL-2A and B operate in a sun-synchronous polar orbit at 786 km altitude. Radiances are detected at the Multispectral Instrument (MSI),

Table 4.1.: SENTINEL-2 characteristics

Characteristics	**SENTINEL 2**
Temporal resolution	10 days one satellite 5 days two satellites
Radiometric resolution	12 bit
Spectral resolution	13 bands
Spatial resolution	10, 20, 60 m

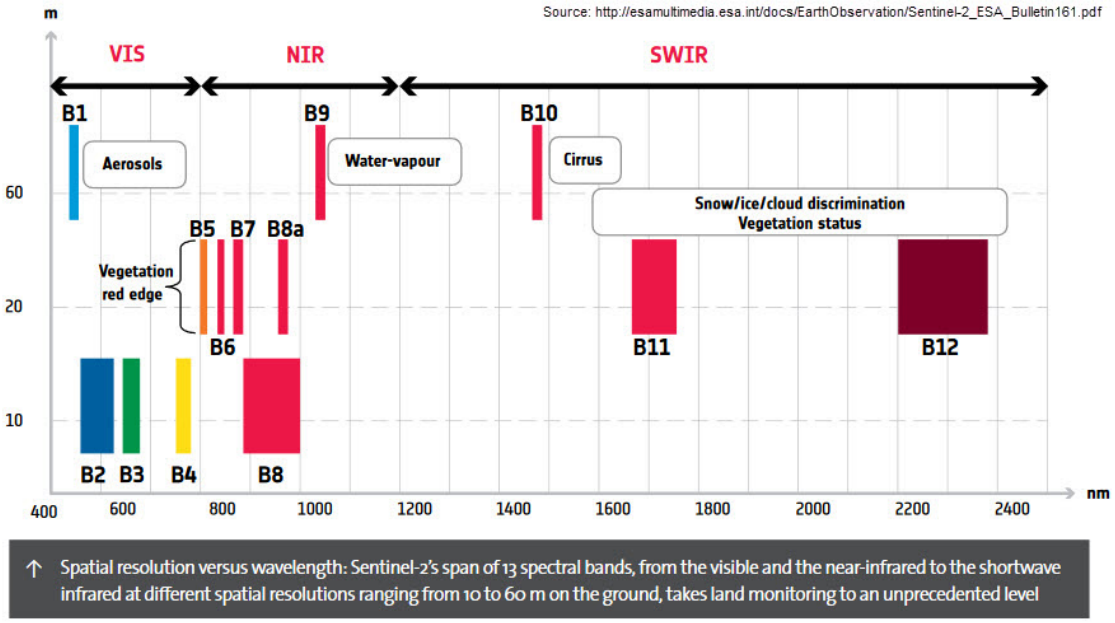


Figure 4.1.: SENTINEL-2 bands. Reprinted from: ESA, European Space Agency, 2015

as the satellite moves along path (P. Martimort et al., 2007; Suhet, 2015; Drusch et al., 2012). The characteristics for both satellites are summarised in table 4.1.

With both satellites on track each area will be revisited every five days under the same viewing conditions. For the time period only SENTINEL-2A is operational optical images are generated for the same area every 10 days. With different viewing conditions, such as overlap between swaths and rolling or tilting the satellite, the revisit frequency can be increased to 1 to 3 days (P. Martimort et al., 2007; Suhet, 2015; Drusch et al., 2012; Main-Knorn et al., 2015; Malenovsky et al., 2012).

The sensors radiometric resolution represents the level of reflectance that can be distinguished. A high radiometric resolution allows the sensor to distinguish more reflectance differences. The SENTINEL-2 sensor with 12-bit radiometric resolution enables the image to store 4096 (2^{12}) potential light intensity values (Suhet, 2015; Drusch et al., 2012; Main-Knorn et al., 2015).

Figure 4.1 shows the spectral and the spatial resolution from SENTINEL 2. The latter depends on the spectral band. SENTINEL-2A is provided with four bands extend from Visible to NIR with a spatial resolution of 10 m (B2, B3, B4, B8). Six bands are available with 20 m spatial resolution, four in the vegetation red edge domain and two large SWIR bands (B11, B12). The red edge bands, centred at 705 (B5), 740 (B6) and 783 (B7) nm, can be used for assessing vegetation status and types through monitoring parameters like green leaf area index and chlorophyll

content. Three bands at 60 m are provided, which are dedicated to improve the atmospheric correction and cloud screening. One is for the water vapour retrieval (B9), one for aerosol screening (B1) and one for cirrus cloud detection (B10) (P. Martimort et al., 2007; Fletcher K., 2012; Delegido et al., 2011; Main-Knorn et al., 2015).

4.1.2. Product Overview

The data itself is characterized by granules and tiles. "A granules is a minimum indivisible partition of a user product (containing all possible spectral bands)"(Suhet, 2015). The granule size is fixed and depends on the different product levels provided. The granules for the Level-0, Level-1A and the Level-1B products have a size of 25 km across track and 23 km along track. The granule size for the provided orthorectified products, such as Level-1C and Level-2A, is 100 by 100 km. All scenes are distributed to users as SAFE (Standard Archive Format for Europe) format (Suhet, 2015; Main-Knorn et al., 2015).

The Level-1C product is the first product level available in cartographic geometry. The ortho-images are represented in UTM/WGS84 projection. The product is derived from the Level-1B products, using a Digital Elevation Model to project the image in cartographic coordinates. Pixels represent the Top-of-Atmosphere reflectance and parameters for the radiance transformation are provided. Furthermore, cloud, land and water masks are provided. Pixel coordinates no longer refer to the pixel center, but to the upper left corner (Suhet, 2015; Drusch et al., 2012).

The Level-2a product represents the Bottom Of Atmosphere reflectance values, which are derived from the Level-1C product through an absolute calibration. The tiles have the same size and cartographic geometry as the associated Level-1C product. These products are not systematically generated at the ground segment. The generation is performed by users, supported through the Sen2Cor processor included in the SENTINEL-toolbox, using the downloaded Level-1C data as input (Suhet, 2015). Additional outputs are Scene Classifications, Aerosol Optical Thickness Map and a Water Vapor Map. All output images will be resampled to one resolution for all bands, which can be chosen by the user (Main-Knorn et al., 2015; Suhet, 2015; Muller-Wilm et al., 2013).

4.2. Forest Validation Plots

The Pan-European High Resolution (HR) Forest Layers from 2011 are provided by the Copernicus program formerly Global Monitoring for Environment and

Table 4.2.: Overview of the main forest HRLs. Reprinted from: Büttner et al., 2015

Forest	
Tree cover density (20m and 100m)	Tree cover density values from 1-100 %
Forest Type (20m) consisting of two grids	Dominant leaf type. Binary product: coniferous and broadleaved, MMU of 0.5 ha and 10 % tree cover density threshold applied
	Support layer: Maps trees under agricultural use and in urban context from CLC 2012 and imperviousness 2012
Forest Type (100m x 100m)	Coniferous, broadleaved and mixed. Trees under agricultural use and urban context from additional support grid removed

Security (GMES), apart from other HR Layers such as water bodies and wetlands. The HR layers are produced "through a combination of automatic processing and interactive rule based classification" (Sannier et al., 2015) for the reference year 2012 mainly based on 2011/2012 20 m geometric resolution satellite images, covering 39 European countries (Langanke et al., 2015). The HR forest layers are provided as original 20 m product and validated 100 m product consisting of two different layers. Both the tree cover density layer and the forest type layer, are available on the Copernicus Land Portal and described in table 4.2 (Büttner et al., 2015; Bruzzone and Demir, 2014; Sannier et al., 2015).

For the forest HR validation process a blind approach has been applied. In other words, the validation has been undertaken without considering the map information. In addition to the high resolution reference data the production data is used in the validation process to ensure the geometrical quality and reference period (Sannier et al., 2015).

Production Data:

- HR Image 2012 Coverage 1 CIR (25m)
- HR Image 2012 Coverage 2 RGB (20m)
- VHR Image 2012 CIR (2.5m)
- VHR Image 2012 RGB (2.5m)

(Sannier et al., 2015)

Table 4.3.: TCD and FTY omission and commission substrata. Source: (Sannier et al., 2015)

Stratification	**TCD**	**FTY**
Commission Low Probability	tree cover density 1-100 % & CLC tree cover classes	Forest Type 1-2-3 & CLC forest classes
Commission High Probability	tree cover density 1-100 % & CLC non tree cover classes	Forest Type 1-2-3 & CLC non forest classes
Omission Low Probability	tree cover density 0 % & CLC tree cover classes	Forest Type 0 & CLC forest classes
Omission High Probability	Rest of the area	Rest of the area

Reference Data:

- Bing maps (ArcGIS Basemap layer, RGB imagery)
- Arc2Earth imagery (Google commercial ArcGIS plugin, RGB imagery)
- National and regional web map services (RGB and/or CIR imagery)
- LUCAS field survey photos

(Sannier et al., 2015)

The stratified systematic validation sampling approach is conducted in two levels based on the LUCAS (Land Use Area frame statistical survey) points, which are located every 2 km on a regular grid covering Europe. With in the validation framework 23 strata are formed, defining a single country or groups of countries with an area greater than 90 000 km^2. A minimum number of 50 sample points per stratum has been defined to ensure that also small strata are represented in the sample. Regarding those strata the LUCAS samples are densified by creating every 200 m one point. Additionally, sample points have been added representing difficult classes. Therefore, sample points are stratified according on omission/commission strata, to get more samples for problematic areas, which are difficult to map and have a higher probability that errors will be found (substrata: Commission and Omission High). The Commission and Omission strata are created based on the differences between the HRL (High resolution Layer) 2012 product and Corine Land Cover (CLC). Consequently, sampling plots within those strata are weighted higher. The differences between FTY and TCD omission and commission strata are represented in table 4.3. Overall 23 390 plots are prepared across 39 European countries (FTY: 12 152, TCD 12 152). Those points are interpreted based on available image data listed above (Sannier et al., 2015).

Table 4.4.: Including and excluding decisions in TCD products. Reprinted from: Büttner et al., 2015

Included Features	
Evergreen/non-evergreen broadleaved, sclerophyllous and coniferous trees	
Orchards, olive groves, fruit and other tree plantations, agro-forestry areas, forest nurseries, regeneration and transitional woodlands	
Alleys, wooded parks and gardens	
Groups of trees within urban areas	
Forest management/usefeatures inside forests (forest roads, fire-breaks, thinning, etc.)	Included if tree cover can be detected from the 20m imagery
Forest damage features inside forests (partially burnt areas, storm damage, insect-infested damage, etc.)	
Excluded Features	
Open areas within forests (roads, permanently open vegetated areas, clear cuts, fully burnt areas, other severe forest damage areas, etc.)	Excluded if no tree cover can be detected from the 20m imagery
Shrub land	
Mediterranean bush land (macchia, guarrigue, etc.)	
Dwarf pine/green alder in high/mountainous areas	

4.2.1. Copernicus Tree Cover Density High resolution layer (TCD)

The tree cover density layer is a binary product containing values from 1 – 100 %. Only tree patches with a minimum width of 20 m are considered but no minimum mapping unit for the tree patches is given. The product is provided with 20 m square pixels and as 100 m aggregated product (Sannier et al., 2015; Büttner et al., 2015).

Regarding the validation process the same aggregation rule as mentioned above is used. Additionally the classes 1 to 10 % are interpreted depending on the visual estimation of trees in the HR 20 m satellite imagery. The definitions for the tree cover density validation are given in table 4.4 (Sannier et al., 2015).

Table 4.5.: FAO Forest Definition, Source: FAO, Food and Agriculture Organisation of the United Nations, 2015

FAO Forest Definition	
includes	forest nurseries and seed orchards that constitute an integral part of the forest
	forest roads, cleared tracts, firebreaks and other small open areas
	forest in national parks, nature reserves and other protected areas such as those of specific scientific, historical, cultural or spiritual interest
	windbreaks and shelterbelts of trees with an area of more than 0.5 ha and width of more than 20 m
	plantations primarily used for forestry purposes, including rubber wood plantations and cork oak stands
excludes	Land predominantly used for agricultural practices

In regard to the validation purpose the Tree Cover Density 100 m by 100 m product in European projection LAEA is used. The interpreter has to check whether the centre points of the 25 cell gird are located above tree crowns and then assign the tree cover density values of each 20 by 20 cell represented in figure 4.2a. Each plot is flagged, regarding the quality of the reference data. If no significant interpretation reference is available, due to clouds ect., the plot is flagged "LOW". Those plots are excluded from the validation process. The same procedure is valid for plots with a significant geometry difference between production and reference data (Sannier et al., 2015).

4.2.2. Copernicus Forest Type High resolution layer (FTY)

The forest definition for the Forest Type layer is as closely following the forest definition from the FAO, Food and Agriculture Organisation of the United Nations (2015) shown in table 4.5.

The Forest Type 20 by 20 m Layer is produced based on the tree cover density layer and consists of two sub-products, one representing dominant leaf types excluding patches smaller than 0.5 ha and with a tree cover density above 10 %, the other one representing dominant leaf type for the excluded features like trees under agricultural use and in urban context. Regarding the aggregated 100 x 100 m product trees under agricultural and urban context are excluded, additionally containing the mixed forest class despite to the 20 x 20 layer (Sannier et al., 2015; Langanke et al., 2015).

The 100 by 100 m product is validated according to the forest definition shown in table 4.5. Furthermore, aggregation rules are considered within the validation process. Forest patches are only considered in a 4-cell connectivity mode. In this context it has to be mentioned that only vertical and horizontal connected pixels are accounted. The minimum number of pixels to form a patch is 0,5 ha (MMU), corresponding to 13 cells in the 25 cell array. If more than one forest patch is located in the cell array tree stocked cells belonging to a patch with less than 0.5 ha (including the surrounding) are excluded from forest cover. The same is true for plots with less than 10 % tree cover. Further, the Minimum Mapping Width of 20 m has to be considered. Summarizing, if a plot has 13 cells of trees in a 4 cell connectivity mode and each patch has more than 0,5 ha including the surroundings and is at least 20 m wide, then FTY is assessed represented in figures 4.2b, 4.2c and 4.2d. Regarding the validation process the classes non-forest, broadleaved, coniferous and mixed forest are assigned by the interpreter as follows (Sannier et al., 2015; Langanke et al., 2015):

- non-forest
- Broadleaved Forest (0 - 25 % coniferous)
- Coniferous Forest (75 - 100 % coniferous)
- Mixed Forest (26 - 74 % coniferous)

(Sannier et al., 2015)

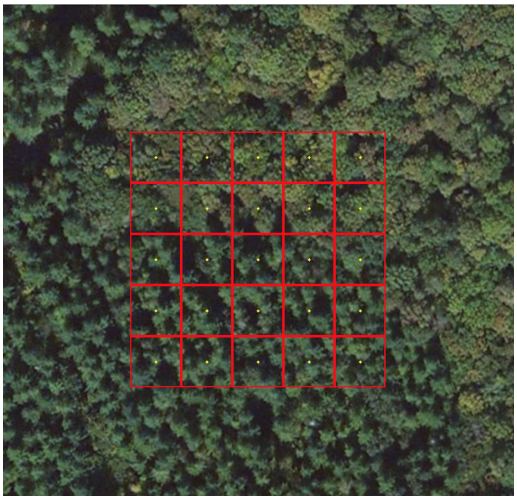
4.3. Digital Elevation Model

Important ancillary information is provided by a digital elevation model representing altitude information based the bare ground surface without vegetation or buildings. The digital elevation model used within the classification approach is derived from several local elevation models and the SRTM 1 arc-second model (U.S. Department of the Interior, U.S. Geological Survey, 2016) as well as the EU-DEM (European Environment Agency (EEA), 2017).

> *"SRTM 1 Arc-Second Global elevation data offer worldwide coverage of void filled data at a resolution of 1 arc-second (30 meters) and provide open distribution of this high-resolution global data set. Some tiles may still contain voids. [...] Please note that tiles above* 50° *north and below* 50° *south latitude are sampled at a resolution of 2 arc-second by 1 arc-second."(U.S. Department of the Interior, U.S. Geological Survey, 2016)*

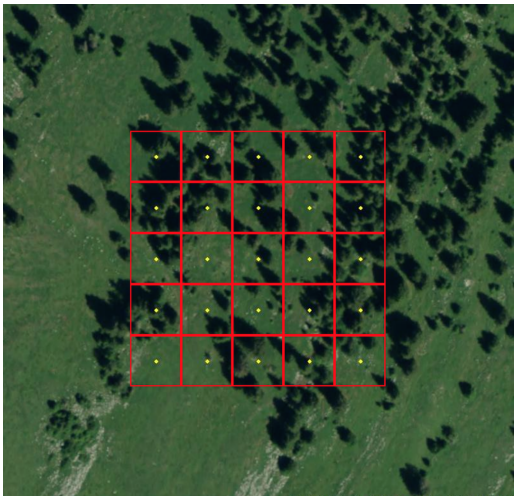
After the downloaded SRTM model is visually inspected still some small gaps are found within alpine areas, consequently additionally the EU-DEM with a ground

(a) TCD validation example



(b) FTY validation example 2

(c) FTY validation example 1



(d) FTY validation example 3

Figure 4.2.: TCD and FTY sampling plot data according to above definitions are provided by Joanneum Research.

resolution of 25 m in ETRS89-LAEA has been used to fill those gaps (European Environment Agency (EEA), 2017).

Considering the ground resolution of the SENTINEL-2A image data with 10 m or 20 m, local DEM's with similar ground resolution are used to provide more detailed altitude information. The first local DEM covers north Italy and provided by Tarquini et al. (2012) with 10 m ground resolution in UTM WGS 84 zone 32 projection system. The second local DEM covers the Austrian country and is provided by Geoland (2017) with 10 m ground resolution derived from airborne laser scanner data in UTM WGS 84 zone 32 projection system.

All derived elevation models are projected to ETRS89-LAEA and resampled to 10 m ground resolution. Furthermore, the models are mosaicked and filtered to smooth edges.

4.4. Corine Land Cover

The Corine Land Cover database 2012 obtained from the Copernicus website (Copernicus Land Monitoring Service, 2017) is used within this thesis to derive a current forest mask.

> *"The CORINE Land Cover (CLC) inventory was initiated in 1985 (reference year 1990). Updates have been produced in 2000, 2006, and 2012. It consists of an inventory of land cover in 44 classes. CLC uses a Minimum Mapping Unit (MMU) of 25 hectares (ha) for areal phenomena and a minimum width of 100 m for linear phenomena. [...] CLC is produced by the majority of countries by visual interpretation of high resolution satellite imagery. In a few countries semi-automatic solutions are applied, using national in-situ data, satellite image processing, GIS integration and generalisation. [...] "(Copernicus Land Monitoring Service, 2017)*

Three different forest classes are represented with in the Corine Land Cover database in vector format listed below. Those classes are used to derive a current Forest/Non-Forest mask with 100 x 100 m pixel size to exclude Non-Forest pixel for analysis and classification purposes.

"

- 3.1.1 Broad-leaved forest
- 3.1.2 Coniferous forest
- 3.1.3 Mixed forest

" (Copernicus Land Monitoring Service, 2017)

5. Study Area

In order to test the effectiveness of an independent survey points based classification a study case located in central Europe has been chosen. The study area covers about 33.000 square kilometres and the landscape is characterised by the presence of temperate forests, large plains, mountainous regions and hilly and lake regions.

Rigo et al. (2016) and the authors of the Motive Project (2016) differentiate between five ecological zones in Europe depending on the average temperature and rainfall. The boreal zone located in the northern part of Europe and characterised by short summers and long and cold winters resulting in dense coniferous forests dominated by spruce and fir. "The Atlantic climate in the western part turns to a continental climate with decreasing humidity and higher temperature variation in the eastern part of Europe" (Spiecker, 2003), therefore temperate forests located in the mid-latitude area of Europe vary from oceanic to continental influenced forests covering "[...] floodplain to mountain forests up to the alpine timberline."(Spiecker, 2003). As a result to the different climatic conditions resulting in 4 - 6 months growing season, temperate forests vary within the continent. Temperate forests are further differentiated between oceanic temperate forests, occurring in regions with high rainfall and high winter temperatures and temperate continental forest in regions with constant moderate rainfall which are mainly dominated by a variety of broadleaved and conifer tree species (Lindner et al., 2010; Spiecker, 2003, qtd. in: Mitscherlich, 1978). Europe is further characterized by sclerophyllous evergreen forests, occurring in the Mediterranean regions, which is characterized by hot and dry summers and mild winters (Motive Project, 2016; Rigo et al., 2016). The test site region ranges in altitude from 10 m above sea level in the Po Valley up to 4500 m above sea level in the Swiss Alps, containing approximately two third from the Alps. Accordingly, forest are widely distributed over mountainous terrain, reaching up to about 1800 m above sea level and having large variations in species across the elevation gradients. The canopy density generally decreases with increasing altitude due to the changing climate conditions (Spiecker, 2003, qtd. in: Reichle, 2013).

The conifer species are dominated in countries like Austria, while in Croatia, France, Hungary more than 60 % of the forest area are broadleaved dominated. In other central European countries like France and the Principality of Liechtenstein

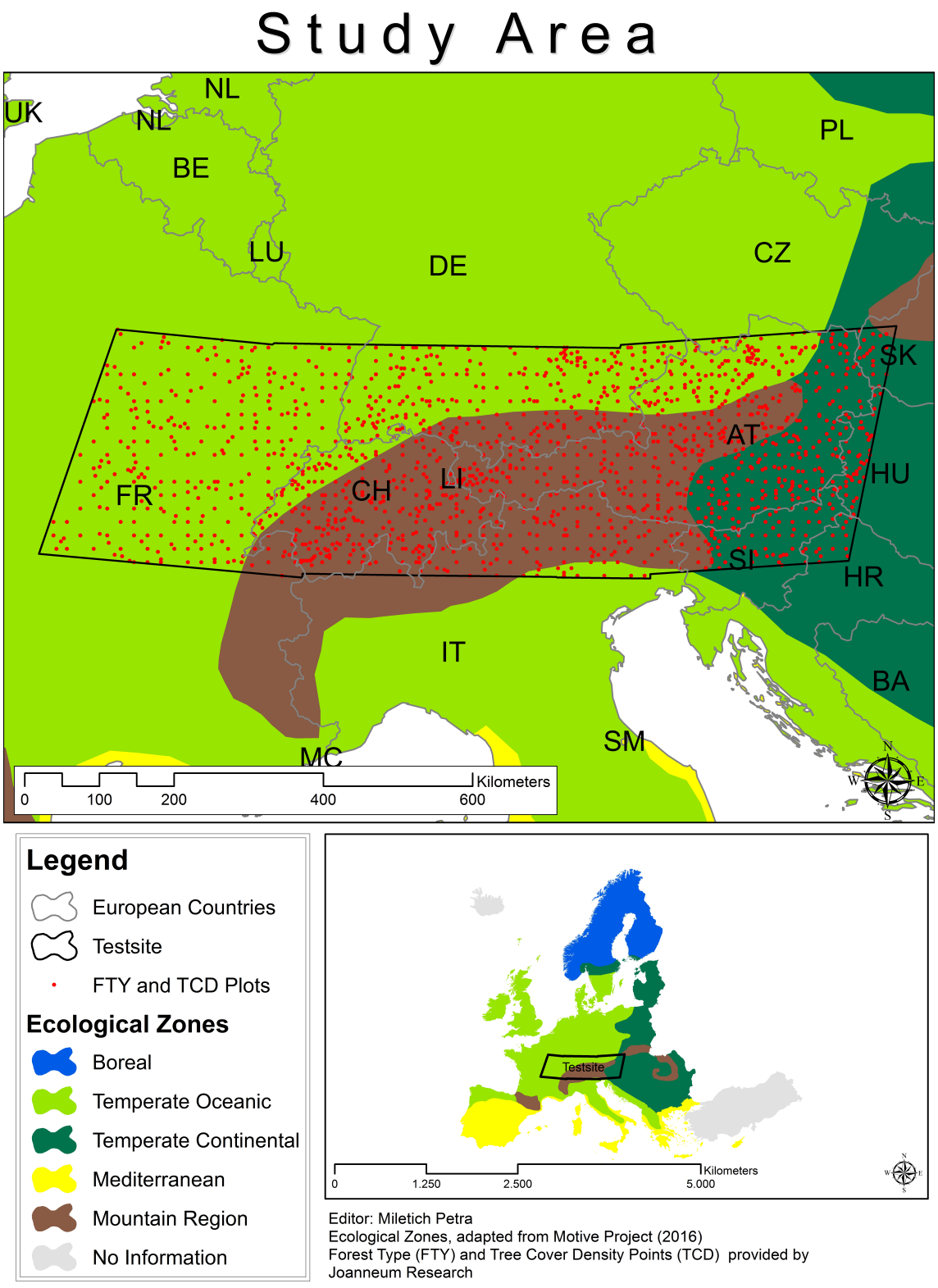


Figure 5.1.: Study area for TCD and FTY mapping. Sampling plot locations are shown as red dots.

Table 5.1.: Used SENTINEL-2A scenes

Date	Scene	Initials
04.07.2015	S2A_OPER_PRD_MSIL1C_PDMC_20160315T070646_R008_V20150822T104654_20150822T104654	Sa1
10.08.2015	S2A_OPER_PRD_MSIL1C_PDMC_20160514T142737_R122_V20150810T100716_20150810T100716	Sa2
13.08.2015	S2A_OPER_PRD_MSIL1C_PDMC_20160309T041851_R022_V20150813T101657_20150813T101657	Sa3
22.08.2015	S2A_OPER_PRD_MSIL1C_PDMC_20160408T105430_R108_V20150829T103705_20150829T103705	Sa4
26.08.2015	S2A_OPER_PRD_MSIL1C_PDMC_20160315T115050_R065_V20150826T102655_20150826T102655	Sa5
29.08.2015	S2A_OPER_PRD_MSIL1C_PDMC_20160408T105430_R108_V20150829T103705_20150829T103705	Sa6
30.08.2015	S2A_OPER_PRD_MSIL1C_PDMC_20160309T032233_R122_V20150830T100547_20150830T100547	Sa7

coppice contributes to more than 40 % to the forest area and countries like Croatia, Germany, Principality of Liechtenstein, Slovakia, Slovenia and Switzerland have more than 15 % unevenly aged forests. Information on the spatial coverage of the three main forest types broadleaved forest, coniferous forest and mixed forest are needed for multifunctional sustainable forest management (Spiecker, 2003).

The spatial extent of the study area is shown in figure 5.1 covering various parts of ten different European countries including France, Switzerland, Italy, Slovenia, Austria, Croatia, Hungary, Slovakia, Czech Republic, Germany and the Principality of Liechtenstein. The test site is represented by seven SENTINEL-2a scenes from July 2015 till August 2015 see table 5.1. Those scenes are chosen because they cover a latitudinal gradient throughout the continent. Accordingly, the test site area covers only three of the five ecological zones, with similar forest types. The size of the study case is chosen depending on the number of available plots within the SENTINEL-2A scene mosaic. After combining five scenes in east or west direction 1556 forest type and 1482 tree cover density plots are available within those scenes, represented by the red points in figure 5.1. Two more scenes covering the same area are included for gap-filling reasons (Sa1, Sa2).

6. Forest Mapping Workflow

This chapter provides a brief overview of the thesis workflow, the single processes will be explained in detail in the flowing chapters. Figure 6.1 represents a overview of all processes required within the classification framework.

In this thesis the thematic class definitions are adopted from the validation framework for the GIO HR Forest Layers, differentiating between continuous tree cover density classes and discrete forest type classes described in chapter 4.2.1 (TCD) and 4.2.2 (FTY). The SENTINEL-2A scenes are downloaded from the Copernicus Data Hub and exported into tiff format with the SENTINEL Toolbox including all layer with surface information and the scene classification. The image preprocessing is done with the Sen2Cor package and the use of the ARCGIS and ERDAS software including the following four steps: Cloud and shadow masking, terrain correction, atmospheric correction and no data handling. The corrected images are mosaicked to provide consistent coverage over a large region. Cloud and shadow masks are adopted from the scene classification and extracted from the image data. Afterwards the preprocessed mosaic is re-projected in the Lambert Azimuthal Equal Area (LAEA) coordinate system suitable for European data requiring true area representations.

The validation samples used as training areas need to be prepared for the classification process including a visual interpretation of changes between plot validation date and image acquisition date. Within the classification process only sample plots inside the study area are extracted and divided into trainings samples (85 %) and validation samples (15 %) or (20 %). To find the most suitable band combinations regarding the classification process signature analysis including regression analysis and separability analysis based on the Jefferies-Matusita distance measure have been evaluated. As mentioned in chapter 2 metric supervised hard classifiers are considered to be valuable for both continuous tree cover density and categorical forest type classification. Accordingly random forest and k-nearest-neighbours classifiers are applied and evaluated.

As represented the first section of figure 6.1 the data preparation steps for both classifications are identical, whereas within the classification process, post-processing process and the validation process continuous and discrete data are handled differently.

In regards to the tree cover density prediction regression models are trained with both classifiers and applied to predict the tree cover density over the whole mosaic. The regression models are characterized by using the mean value as decision rule within the random forest and the knn classifier training process. Fitting parameters are estimated if necessary with regression analysis within the post-processing stage. Those are applied if the resulting classified product is strongly underestimated to achieve continuous values from 0 to 100. After the calibration step the classification result is evaluated with scatter plots and regression coefficients to get a general overview about commission and omission errors. Thresholds are applied to group the density values in order to compute confusion matrices. The thematic accuracy for each product is determined by confusion matrices calculated to estimate the users and producers accuracies as well as the overall accuracy. As represented in the yellow section of figure 6.1 the tree cover density prediction is performed twice. The first approach includes also Non-Forest samples in comparison to the second approach which uses a forest mask to exclude Non-Forest areas. According to the tree cover density definition given by HR TCD Layer each tree without restrictions is considered. Consequently, it makes sense to include Non-Forest areas because trees also occur outside of forest areas. Nevertheless, forest applications have particular interest in tree cover density values inside forest areas. Therefore, both approaches are examined within this thesis.

In the case of the forest type distinction only one approach is examined, differentiating between three forest types within forest areas. The forest definition of the FAO excludes areas under urban and agricultural use. Accordingly, only forest areas can be considered as it does not make sense to classify Non-Forest areas within a forest product differentiating between forest types. Regarding the forest type distinction classification models are trained with both classifiers and applied on the whole mosaic. The classification models are characterized by using the majority vote as decision rule within the random forest or the knn classifier training process. The majority vote is used to deal with categorical data values and picks the most common class among all considered possibilities. In the post-processing stage the Minimum Mapping Unit of 0,5 ha needs to be considered, due to forest type definition given by the FAO. After the aggregation process using the majority vote over all forest types confusion matrices are calculated.

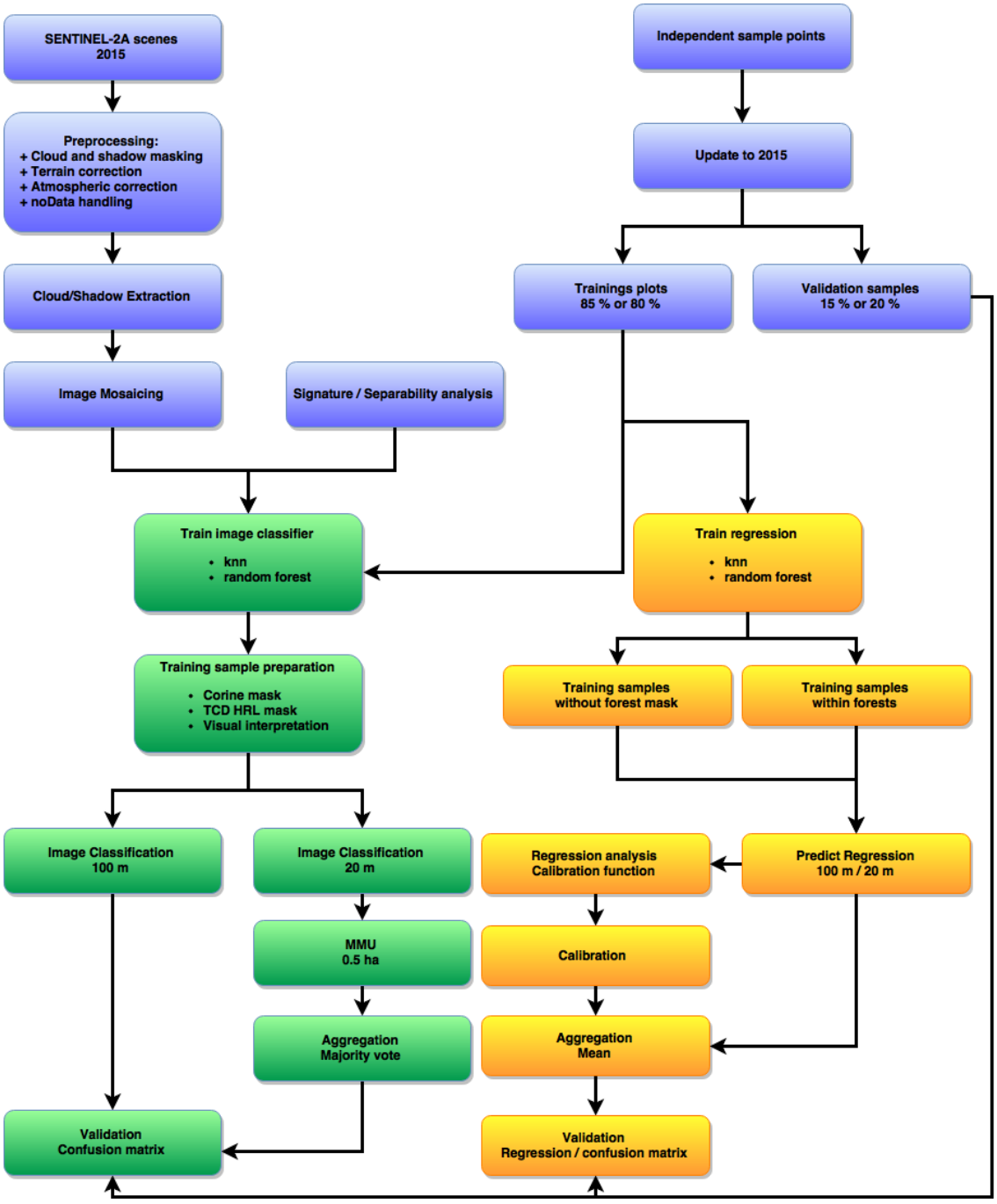


Figure 6.1.: Workflow regarding the TCD and FTY classification process

6.1. SENTINEL-2A Preprocessing

For SENTINEL-2 data product preprocessing the Sen2Cor processor is recommended, which can be installed in the SENTINEL-2 toolbox. It performs several tasks from cloud screening, scene classification to the Bottom of Atmosphere (BOA) reflectance calculation including terrain correction, adjacency correction and an empirical bidirectional reflectance distribution function (BRDF) correction. As input Level-1C products are required. The output product is identified as Level-2A product and includes the orthorectified bottom of atmosphere reflectance, aerosol optical thickness map, water vapour map, a scene classification map and various quality indicators like probabilistic cloud or snow masks (Main-Knorn et al., 2015; Muller-Wilm et al., 2013).

6.1.1. Atmospheric and Terrain Correction with Sen2Cor

In this thesis, SENTINEL-2A scenes with different acquisition dates are used, therefore an atmospheric correction of the image data is necessary, since the state of the atmosphere varies with time. The image preprocessing is done with the Sen2Cor package, thus it is compatible with the provided SAFE format, can be operated in batch mode and is compatible to the SENTINEL-2 toolbox. All bands are corrected and resampled with 10 m resolution. In order to the classification process the four 10 m bands are aggregated to 20 m and stacked to the six bands with 20 m spatial resolution.

How the processor works in detail will be explained in the following section. Sen2Cor is composed of several state-of-the-art techniques performing the task of correcting the reflectance values (Muller-Wilm et al., 2013).

> *"The [atmospheric correction] AC processing consists of a set of four different subtasks, (AOT, WV and terrain retrieval (optional with terrain and cirrus correction, having three different user products as output: AOT and WV tables on pixel level and the BOA corrected reflectance images for all bands measured"(Müller-Wilm, 2015)*

Parameters like atmospheric visual transparency, aerosol type and the water vapour column are needed for the atmospheric correction process and derived through the associated bands (Muller-Wilm et al., 2013). The Dense Dark Vegetation algorithm performs the task of the **Optical Aerosol Thickness** detection and aerosol type estimation, requiring areas of known reflectance behaviour. How the process works in detail is explained by R Richter et al. (2006):

> *"The method is based on a stable relationship between surface reflectance r in the short-wave infrared (SWIR; 2100–2200nm) and reflectance in the blue (480 nm) and red (660 nm) wavelengths. It starts with the approximation* $\rho_{SWIR} = \rho^{TOA}_{SWIR}$ *where* ρ^{TOA} *is the top-of-atmosphere reflectance. This approximation is fairly accurate, because of the small path radiance and high atmospheric transmittance* $\tau > 0.9$ *in the SWIR region [..]. Then, the relationships* $\rho_{blue} = \rho_{SWIR}/4$ *and* $\rho_{red} = \rho_{SWIR}/2$ *are employed and the differences between the TOA apparent reflectance and the reflectance from the above relationships are used to calculate the path radiance and corresponding aerosol optical depth. The spectral reflectance correlation between the red and SWIR wavelengths is sufficiently stable for an aerosol retrieval."(R Richter et al., 2006)*

Pflug et al. (2016) states after a first validation process comparing with AERONET sunphotometer measurements, that the algorithm only works with dark dense vegetation pixels within each granule.

Furthermore, the **Water Vapour** information is required in order to the correction of the reflectance values, gathered through the Atmospheric Pre-corrected Differential Absorption (APDA) algorithm, using the bands B8a and B9. Band 8a provides reflection values in an atmospheric window region, whose values can be seen as reference for the water vapour. Band 9 presents reflectance values in a absorption region. The offset between those two bands is a measure for absorption depth, which dominates the lower atmosphere, assuming that the surface reflectance for B9 is the same as for the B8a. In this case the absorption depth represents water vapour column content (Muller-Wilm et al., 2013). Pflug et al. (2016) observes after a first validation process with AERONET sunphotometer measurements that the algorithm is less influenced by clouds and missing dark vegetation pixels in order to the water vapour retrieval than for the aerosol thickness detection. Since the water vapour column in the upper atmosphere and the ozone absorption content cannot be estimated from image data, seasonal climatological values are taken, differentiating between summer and winter atmosphere (Pflug et al., 2016).

Optionally a **cirrus correction** can be performed using Band 10, which is located in the 1380 nm spectral region, where ground reflected signal is totally absorbed receiving only scattered cirrus signal at the sensor (Rudolf Richter et al., 2011). The Correlation between cirrus Band (B10) and other bands in NIR and SWIR leads to the cirrus contribution, which can be removed from the radiance signal (Muller-Wilm et al., 2013; Pflug et al., 2016). An established cirrus removal methods exists for visible and near infra-red bands. The Sen2Cor processor uses an extended technique removing cirrus during the atmospheric reflectance calculation to avoid reflectance artefacts and improving the accuracy of surface reflectance retrievals. Rudolf Richter et al. (2011) describes the standard cirrus removal method:

> *"A 'virtual' surface is defined consisting of the earth's surface and atmosphere beneath cirrus clouds. It includes the effects of surface reflection and atmospheric scattering and absorption processes. Above cirrus there is the remaining part of the atmosphere with a residual water vapour column. [...] The cirrus reflectance in the VNIR (from 0.4 to 1 µm) is linearly related to the cirrus reflectance at 1.38 µm [and] [...] the cirrus single scattering albedo (SSA) is close to 1 in this part of the spectrum which means that scattering dominates. It does not hold in the SWIR spectrum (1.6 µm, 2.2 µm) where cirrus absorption is higher [...] (Rudolf Richter et al., 2011)."*

This means that standard cirrus removal method only considers channels in NIR window regions, excluding the 0.94 m water vapour region. Hence, an enhanced cirrus removal method is introduced by Rudolf Richter et al. (2011), accounting for the water vapour absorption above cirrus clouds, during the atmospheric correction process using the B10 band. "For instruments with similar channels as SENTINEL-2 the water vapour is retrieved with a band in a window region (around 0.87 μm) and one in the 0.94 μm absorption region (Rudolf Richter et al., 2011)." The transmittance between sun-cirrus and cirrus-satellite is initially unknown, therefore radiative transfer codes using climatological atmosphere profiles for different seasons and areas adopted from MODTRAN models are used (Rudolf Richter et al., 2011).

Sen2Cor uses a combined approach for the atmospheric and **topographic correction**, eliminating the topographic effects during the BOA reflectance calculation. In order to the terrain correction processing a digital elevation model can automatically be downloaded by the processor using the SRTM 3 arcs model. The 90 m SRTM Digital Elevation Database required by the algorithm for rugged terrain, will be downloaded from CGIAR-CSI. In order to the topographic correction processing parameters like slope, aspect and hill shadow are calculated. The digital elevation model will be automatically resampled and referenced to the SENTINEL scene (Main-Knorn et al., 2015; Muller-Wilm et al., 2013).

In this thesis the SENTINEL-2A scenes are processed with 90 m resolution SRTM Digital Elevation Model (DEM) provided by CGIAR-CSI. The look-up table selections are set within the user configuration file L2A_GIPP.xml provided by the Sen2Cor processor. In the configuration file the three entries Aerosol Type, Mid Latitude and Ozone Content can be set (Müller-Wilm, 2015).

> *"The atmospheric model of SEN2COR (L2A_AtmCorr) is dependent on the calculation of radiative transfer functions for different sensor and solar geometries, ground elevations, and atmospheric parameters. [...] The processor reads the parameter in form of Look Up Tables (LUTs) pertaining to this parameter space and interpolates, if required. The LUTs have been generated*

via libRadtran, a library for the calculation of solar and thermal radiation in the Earth's atmosphere."(Müller-Wilm, 2015)

There are 24 different look-up tables provided by the processor including tables regarding rural and maritime aerosols, mid latitude summer and winter atmosphere profiles and various types of ozone concentration, which can be configured in the L2A_GIPP.xml file (Pflug et al., 2016).

The **Aerosol type** is selected depending on the climatology from the scene acquisition area and time, using built-in MODTRAN aerosols for example rural or maritime aerosol types are differentiated. The ozone content is provided by ancillary data for each scene individually (R Richter et al., 2006; Muller-Wilm et al., 2013). Regarding the test site location in mid Europe extending from France to Hungary the rural aerosol type is chosen. All used SENTINEL 2A scenes are acquired between July 2015 and August 2015 therefore look-up tables for atmospheric summer conditions are chosen.

The ozone content concentration, measured in Dobson Units (DU), should be chosen depending on the season (summer or winter). The standard ozone content for a latitude summer atmosphere is 331 DU and in winter it is about 330 DU (Müller-Wilm, 2015). Therefore a value of 331 is chosen for the atmospheric correction process.

After all required parameters regarding the atmospheric correction are estimated the BOA reflectance calculation is performed using a set of the available and calculated LUTs, chosen by the user according to geographic location and climatology (Main-Knorn et al., 2015; Muller-Wilm et al., 2013). Output images are automatically resampled to a user defined spatial resolution based on the three resolutions available in the input image. The three bands with 60 m spatial resolution (Band 1,9 and 10) will be omitted in the level-2a output, since they are not needed for land cover applications (Pflug et al., 2016; Muller-Wilm et al., 2013; Fletcher K., 2012). In a first validation process again using AERONET as reference data Pflug et al. (2016) observed a high spectral agreement between several example spectra like forest, bright soil ect. and the SENTINEL image data. The reference image is calculated running the atmospheric correction on the SENTINEL scene with an AOT constant measured by the AERONET instrument (Pflug et al., 2016).

Another major performance of the processor is the **scene classification**, which detects 12 different land cover classes (Muller-Wilm et al., 2013). "The algorithm is based on a series of threshold tests that use as input top-of-atmosphere reflectance from the SENTINEL-2 spectral bands. In addition, thresholds are applied on band ratios and indexes like the Normalized Difference Vegetation - and Snow Index (NDVI, NDSI)" (Muller-Wilm et al., 2013). In a first validation approach the scene classification reached an overall accuracy of 78 %. As reference random stratified

Table 6.1.: Corresponding bands regarding each spectral region across MSI, OLI and MODIS

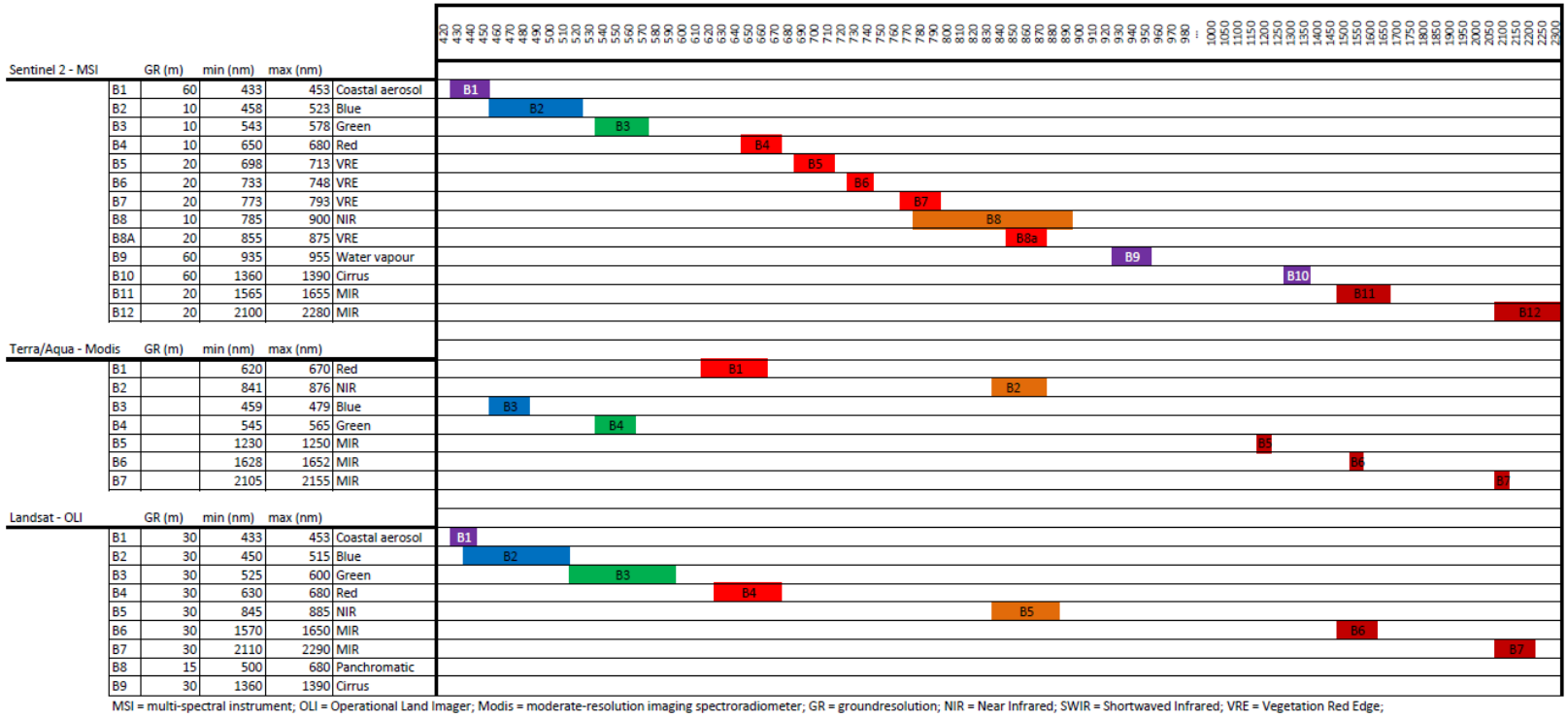
	MSI	MODIS	OLI
RED	B4	B1	B4
NIR	B8a	B2	B5
SWIR	B11	B6	B6
Resolution	20 m	463 m	30 m

sampling points are visually interpreted, with at least 50 points per class (Pflug et al., 2016).

6.1.2. MSI Comparison with OLI and MODIS

To check the consistency of the SENTINEL Level-2a products derived from the Sen2Cor processor (Version 2.2.1), SENTINEL-2 MSI is compared with two other sensors with similar spectral resolution (LANDSAT 8 OLI and TERRA/AQUA MODIS). Figure 6.2 highlights the different bands for all three sensors and also the differences in bandwidths for corresponding spectral bands of each sensor. It also shows which bands are comparable with each other over all three sensors. With regards to OLI the spatial resolution of MSI is finer except regarding the bands used for the atmospheric correction. These bands have been discarded from the analysis because B10 do not contain any surface information and the B9 and B10 are not useful for land cover analysis. The major spectral difference between OLI and MSI is the presence of red-edge bands in MSI and the thermal bands in OLI. Regarding MODIS the spatial resolution of MSI and OLI is much finer. In order to compare the sensors following bands are chosen and resampled to 920 m representing almost twice the MODIS resolution (see table 6.1).

MODIS and OLI sensors products adopted for this analysis are downloaded already containing reflectance values. The SENTINEL-2a scenes are processed with the Sen2Cor processor as explained in chapter 6.1.



Sentinel 2 - MSI	GR (m)	min (nm)	max (nm)	
B1	60	433	453	Coastal aerosol
B2	10	458	523	Blue
B3	10	543	578	Green
B4	10	650	680	Red
B5	20	698	713	VRE
B6	20	733	748	VRE
B7	20	773	793	VRE
B8	10	785	900	NIR
B8A	20	855	875	VRE
B9	60	935	955	Water vapour
B10	60	1360	1390	Cirrus
B11	20	1565	1655	MIR
B12	20	2100	2280	MIR

Terra/Aqua - Modis	GR (m)	min (nm)	max (nm)	
B1		620	670	Red
B2		841	876	NIR
B3		459	479	Blue
B4		545	565	Green
B5		1230	1250	MIR
B6		1628	1652	MIR
B7		2105	2155	MIR

Landsat - OLI	GR (m)	min (nm)	max (nm)	
B1	30	433	453	Coastal aerosol
B2	30	450	515	Blue
B3	30	525	600	Green
B4	30	630	680	Red
B5	30	845	885	NIR
B6	30	1570	1650	MIR
B7	30	2110	2290	MIR
B8	15	500	680	Panchromatic
B9	30	1360	1390	Cirrus

MSI = multi-spectral instrument; OLI = Operational Land Imager; Modis = moderate-resolution imaging spectroradiometer; GR = groundresolution; NIR = Near Infrared; SWIR = Shortwaved Infrared; VRE = Vegetation Red Edge;

Figure 6.2.: Comparison between the MSI, MODIS and OLI sensors. Based on: U.S. Geological Survey (2016), Suhet (2015), and NASA Official (2016).

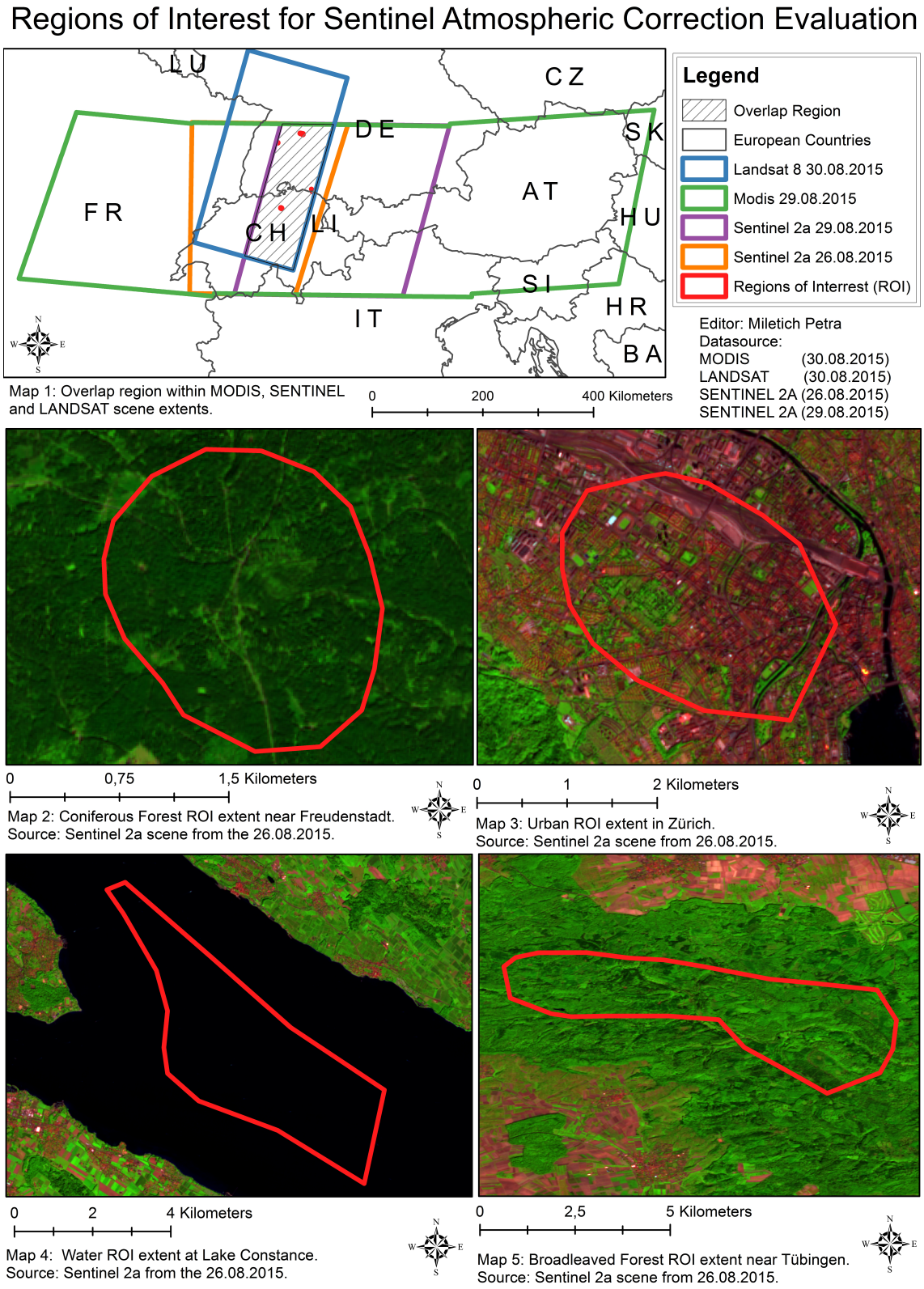


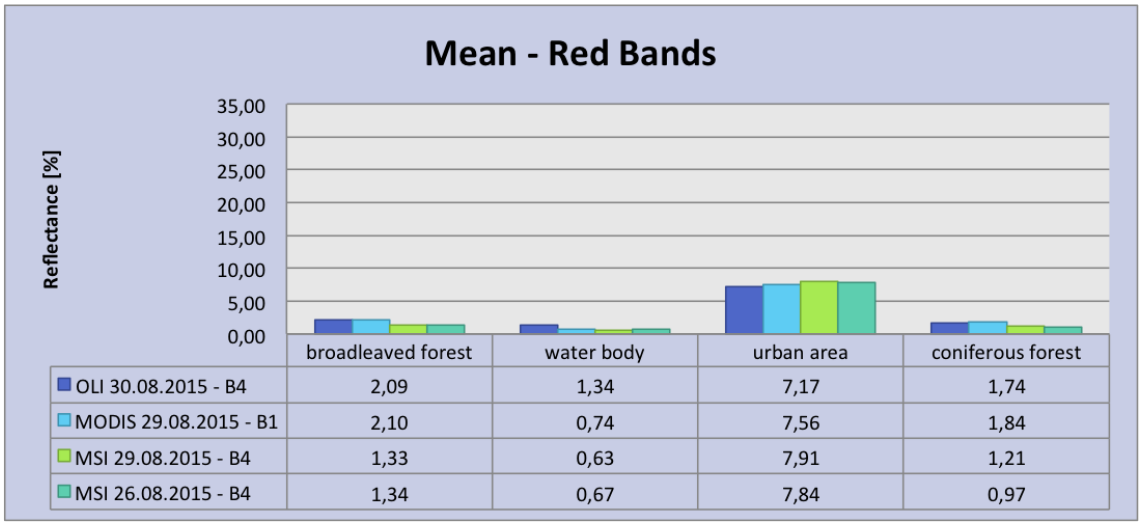
Figure 6.3.: Atmospheric correction evaluation

For the comparison, first, regions of interest are defined for the thematic classes water bodies, broadleaved forest, coniferous forest and urban area in the overlap area of all images to be compared including SENTINEL-2A MSI 26.08.2015, SENTINEL-2A MSI 29.08.2015, TERRA/AQUA MODIS, LANDSAT 8 OLI. Several example spectra are extracted within the reference images and in the images to be compared, represented in figure 6.3. One region of interest for each class is defined containing homogeneous pixels in the resampled products, and centered in the land cover feature to avoid adjacent land cover pixels influencing the comparison. The chosen classes can be considered as steady and therefore changes during short time intervals can be neglected.

Figure 6.4 shows the differences between the mean reflectance values of each region in the RED channel, showing similar results over all three sensors systems for each example area. Especially between MODIS and OLI nearly no differences is observed. In relation to the other sensors the reflectance values for OLI within the water body area are twice as high. Those differences can be explained with the presents of thin cirrus clouds. Nevertheless, it should be considered that in relation to the other classes or in absolute terms the differences are minor. Regarding broadleaved and coniferous forests the MSI sensor shows slightly lower reflectance values in both scenes, indicating a minor under correction for the Sen2Cor processor reflectance values. However, as shown in the statistics the reflectance differences in the red region are less than one percent.

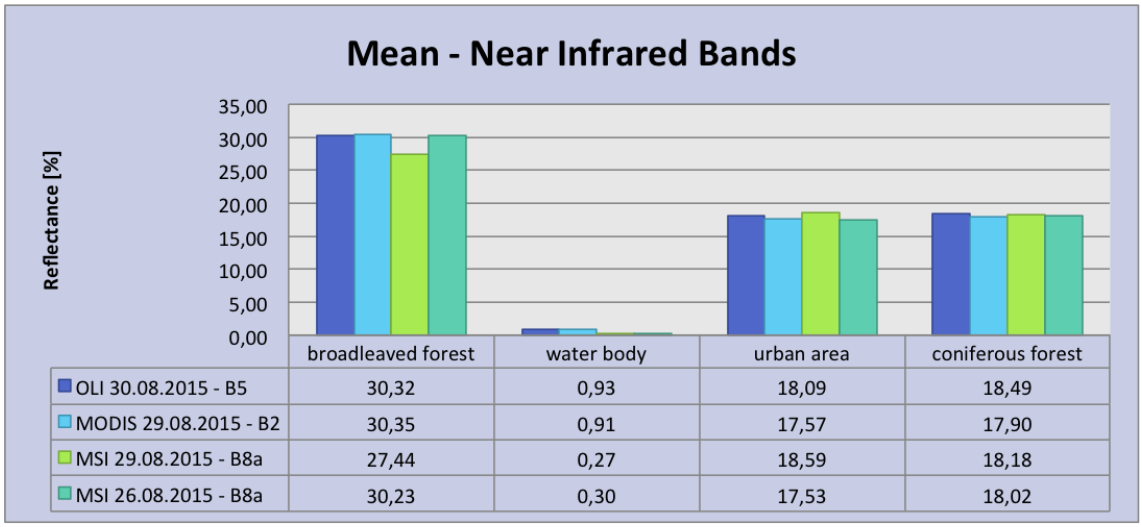
Figure 6.5 represents the differences between the mean reflectance values of each region in the NIR channel, again showing similar results across all three sensors systems with one exception. In the case of the broadleaved region the MSI sensor (29.08.2015) shows extreme low reflectance values in relation to the other sensors and the MSI scene acquired on the 26.08.2015. A visual inspection of the images shows a difference between the processed granules within the SENTINEL scenes. The test sites are located in two different granules. The reflectance values of the upper granule, including the broadleaved and coniferous forest test sites are under-corrected due to haze. Since broadleaved and coniferous forest are located in the same granule, it would be expected that the coniferous forest reflectance values would also be underestimated. Nevertheless, the reflectance values of coniferous forest are at least 20 % lower than the broadleaved reflectance values, accordingly the differences are smaller.

Figure 6.6 shows the differences between the mean reflectance values for the SWIR channel showing more and higher differences over all three sensors systems. In comparison with the RED and NIR channels the SENTINEL scenes also seem over-corrected in relation to the other two sensors. Regarding all example spectra the SENTINEL scenes appear slightly over corrected expect for the broadleaved forest area, whereas the LANDSAT scene seems slightly under-corrected. Nevertheless, it should be noted that the absolute reflectance differences are less than 2 %.



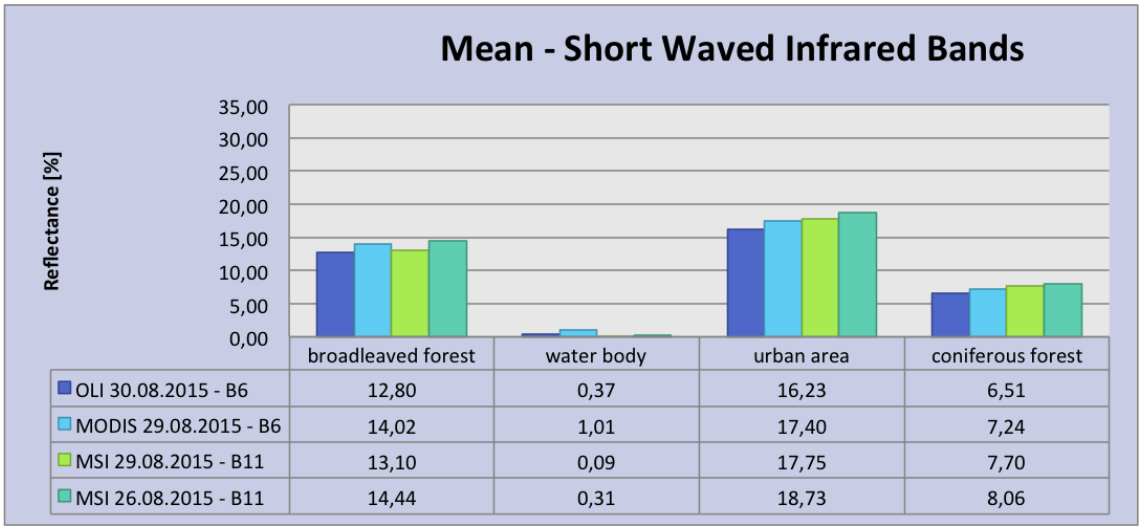
	broadleaved forest	water body	urban area	coniferous forest
OLI 30.08.2015 - B4	2,09	1,34	7,17	1,74
MODIS 29.08.2015 - B1	2,10	0,74	7,56	1,84
MSI 29.08.2015 - B4	1,33	0,63	7,91	1,21
MSI 26.08.2015 - B4	1,34	0,67	7,84	0,97

Figure 6.4.: Mean reflectance - RED band



	broadleaved forest	water body	urban area	coniferous forest
OLI 30.08.2015 - B5	30,32	0,93	18,09	18,49
MODIS 29.08.2015 - B2	30,35	0,91	17,57	17,90
MSI 29.08.2015 - B8a	27,44	0,27	18,59	18,18
MSI 26.08.2015 - B8a	30,23	0,30	17,53	18,02

Figure 6.5.: Mean reflectance - NIR band



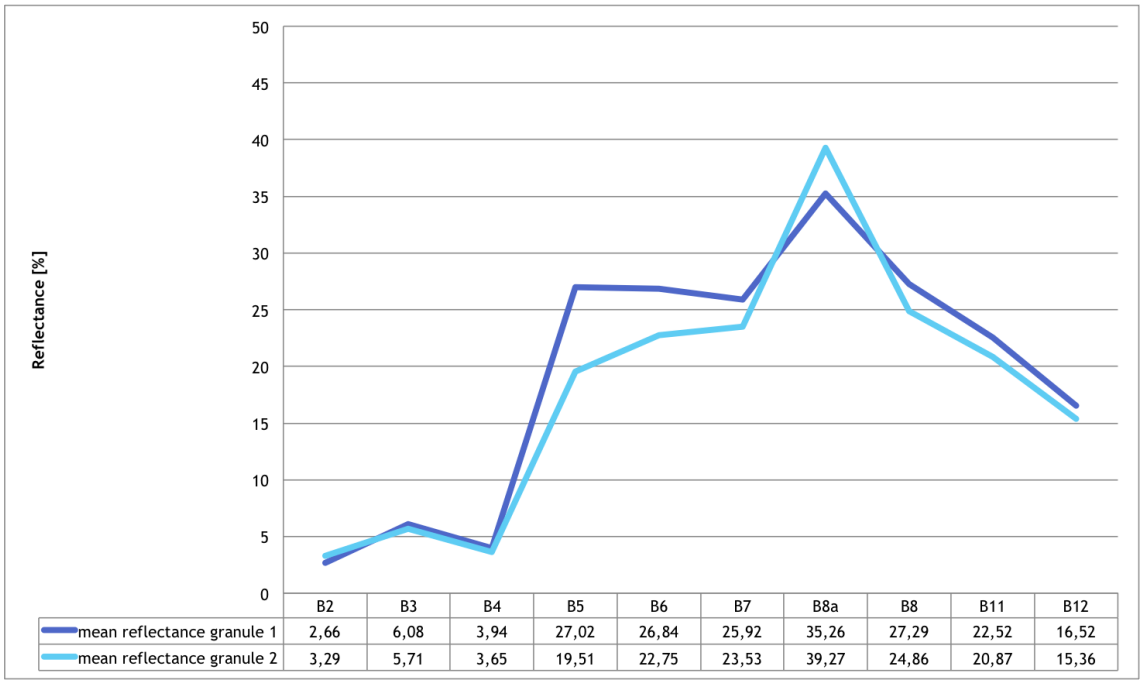
	broadleaved forest	water body	urban area	coniferous forest
OLI 30.08.2015 - B6	12,80	0,37	16,23	6,51
MODIS 29.08.2015 - B6	14,02	1,01	17,40	7,24
MSI 29.08.2015 - B11	13,10	0,09	17,75	7,70
MSI 26.08.2015 - B11	14,44	0,31	18,73	8,06

Figure 6.6.: Mean reflectance - SWIR bands

In general it appears that the example spectra for coniferous forest, broadleaved forest, water and urban area show the expected spectral agreement between reference spectra and images spectra to be compared. For the most part it seems that atmospheric correction with Sen2Cor works consistent like Pflug et al. (2016) have shown in their early validation approach of the Sen2Cor processor. Nevertheless, difficulties arise, when no dark dense vegetation pixels are within the granule. Small overcorrection in comparison with the AERONET sunphotometer measurements are observed by Pflug et al. (2016), which may result from false aerosol parameters, because Sen2Cor estimated higher aerosol optical thickness compared to the reference data. Nevertheless, it should be noted that there is still a possibility that the reference data obtained from AERONET sunphotometer measurements can be under-corrected. The major problem is the granule wise processing of Sen2Cor, which results in granule borders within the scenes in cases of missing dark dense vegetation pixels. The differences between two corresponding granules are shown in figure 6.7, indicating that the differences are higher for bands in the red-edge and NIR region. To get statistically more relevant results, more test sites within different overlap regions, additionally in scenes without dark dense vegetation should be extracted.

6.1.3. Cloud and Shadow Masking with SENTINEL Scene Classification Map

The identification and exclusion of clouds and clouds shadows is necessary due to their influence on spectral reflectance values and characteristics of the thematic classes. Two main cloud types are differentiated: Opaque clouds which block



	B2	B3	B4	B5	B6	B7	B8a	B8	B11	B12
mean reflectance granule 1	2,66	6,08	3,94	27,02	26,84	25,92	35,26	27,29	22,52	16,52
mean reflectance granule 2	3,29	5,71	3,65	19,51	22,75	23,53	39,27	24,86	20,87	15,36

Figure 6.7.: Mean reflectance comparison between two granules for broadleaved forests

almost all information from the surface and cirrus clouds which are partially transparent and share spectral signatures with the land cover beneath. Since with SENTINEL-2 no bands in the thermal spectrum are available, a cirrus band (B10 (1375 nm) is introduced in order to the cloud detection, which is most suitable regarding thin cirrus detection. Cloud pixels are characterized by lower temperature and higher reflectance than land cover pixels, therefore many cloud detection algorithms depend on the thermal band (Z. Zhu et al., 2015; Hagolle et al., 2010). Z. Zhu et al. (2015) demonstrate that the cirrus band is even more important than the thermal band for cloud detection processes, comparing LANDAT 4-7, LANDSAT 8 and simulated SENTINEL 2 images with the Fmask cloud detection algorithm.

Cloud Mask

SENTINEL-2a products are provided with different already available cloud masks depending on the product level. Within the Level-2A product the scene classification algorithms generates a classification map based on spectral threshold tests applied on the cirrus band and band ratios like NDVI and NDSI. The scene classification map provides four different classes of clouds (thin cirrus clouds, high, medium and low probability clouds), four land cover classes (vegetation, soil/deserts, water and snow) and two extra classes for shadows (differentiating

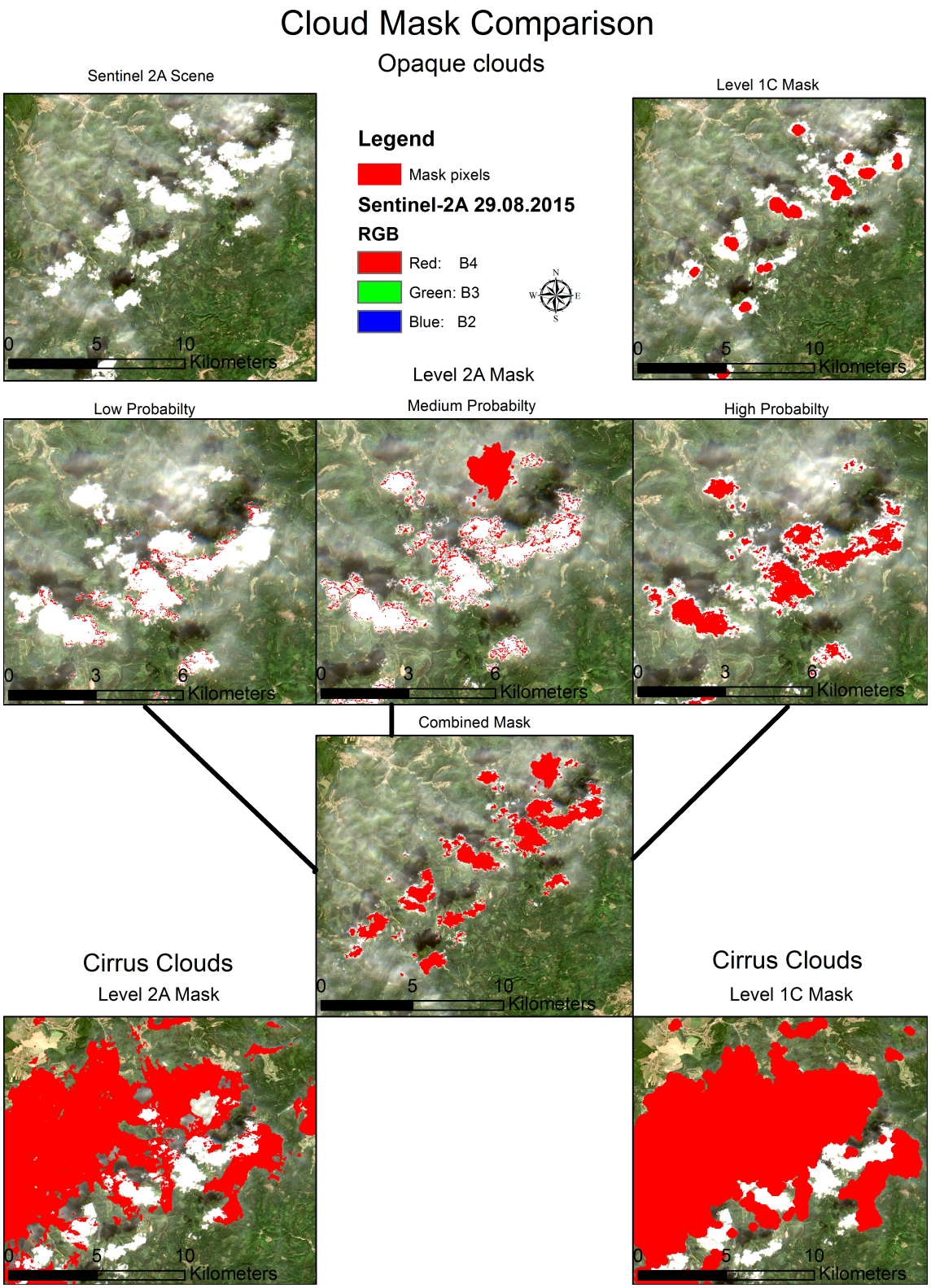


Figure 6.8.: Cloud mask: Comparison between Level 1C and Level 2A product

between cloud shadows and terrain shadows) (Fletcher K., 2012; Muller-Wilm et al., 2013). Furthermore, Level 1C products are provided with spatial filtered cloud masks differentiating between opaque and cirrus clouds based on the scene classification map with 60 m spatial resolution.

Apart from the scene classification map a probabilistic cloud map and a snow probability map, are produced during the processing with Sen2Cor. Within several iterations with different thresholds, the cloud probability map is updated during the Sen2Cor processing. The refined threshold values for the cloud detection algorithm are derived from LANDSAT 7 Automatic Cloud Cover Assessment. The opaque cloud detection algorithm uses the water vapour band (B1) and blue band (B2) in combination with two SWIR bands (B11, B12) to differentiate between snow and clouds. The algorithm first defines potential cloudy pixels within the red part of the spectrum through a certain threshold. Further, on those pixels undergo a filtering process based on spectra thresholds, band ratios and the indices normalized difference snow index (NDSI), normalized difference vegetation index (NDVI). Pixels with values under -0,1 in the NDSI are representing snow and are excluded as potential clouds. Vegetation pixels are excluded from the cloud map deriving information from NDVI and a NIR/Green ratio. Bare soil and water is also excluded through threshold with a blue/NIR ratio. A median filter (3x3 or 5x5) is applied on the three cloud classes to reduces false classifications occurring on border region with high contrast like river contours or shorelines (Louis et al., 2010).

Further, the cirrus band (B10) is used to detect ice high-altitude clouds, which are represented by high reflectance values in B10 and low reflectance values in B1 and B2 (Fletcher K., 2012; Muller-Wilm et al., 2013). The strength of the cloud band lies in the strong water vapour absorption within the mentioned spectral region. Thresholds are used for separating clear sky pixels from cloud pixels (Louis et al., 2010).

If a correlation of the cirrus signal in the cirrus band and the other wavelengths in the NIR region can be found, then the cirrus contribution can be removed. Rudolf Richter et al. (2011) introduced an enhanced cirrus removal method for multispectral images including channels in in the NIR and SWIR region as well as a water vapour band and a cirrus band. During the water vapour and surface reflectance calculation within the atmospheric correction process water vapour above cirrus clouds derived from the cirrus channel is considered, leading to more accurate assessments of the water vapour map. Hence, cirrus artefacts on the SWIR region can be avoided. As mentioned in chapter 6.1.1 radiative transfer look-up tables are used for the water vapour and reflectance retrieval. Regarding the cirrus correction further look-up tables containing above cirrus transmittances for different solar and viewing conditions are used (Rudolf Richter et al., 2011).

Figure 6.8 shows the different cloud classes provided by the Level-1C and Level-2A masks. The mask quality is subjectively examined by visual inspection and not statistically assessed. Level-1C cloud mask contributes cirrus and opaque clouds. In the Level-1C product the mask is already further processed to reduce misclassification and fill gaps. The Level 1C product opaque cloud mask does not cover the full cloud extent. Level-2A scene classification presents three different classes for opaque clouds based on probabilities (High, Medium, Low). Affecting all three probabilities classes misclassification are observed in urban areas and dry mountainous terrain. The high probabilities cloud class is calculated with a high threshold applied on cloud probability mask missing some thin cirrus clouds but reducing misclassification. With higher thresholds only the cloud centres are detected, missing cloudy pixels at the borders. The lower the probability cloud class calculated with a low threshold suffers from false detection for bright surfaces in dry environments in urban areas or on mountainous terrain.
In addition Level-1C and Level-2A products contribute thin cirrus cloud classes taking advantage of the new cirrus band B10. In the Level-1C product the mask is already further processed to reduce misclassification and fill gaps (Muller-Wilm et al., 2013). Both masks are screened reasonably well, but some thin clouds (mostly thinner plane contrails) are missed by the detection in Sen2Cor.

Finally, the cirrus clouds are adopted from the Level-1C product adding some missing contrails. Although Sen2Cor cloud mask tends to over detect clouds, high, medium and low probability clouds and the cirrus are adopted and further processed, due to the fact the low probabilities cloud mask are best at detecting the actual clouds. Clouds with an area less than 400 m^2 (corresponding to four 10 x 10 m pixels) are excluded to reduce the amount of false detections. Due to the too small cloud area a buffer of 200 m was assigned to the clouds. The misclassification in the urban and high altitude areas, can be neglected for forest mapping purposes.

Cloud Shadow and Terrain Shadow Mask

The classes cloud and terrain shadows are only provided by the Level-2A product represented in figure 6.9. Cloud shadows are often misclassified because they share similar spectral signatures with topographic shadows, water and wetlands (Z. Zhu et al., 2015). The cloud and mountainous shadows in the scene classification are classified too small, therefore a buffer of 200 m is applied to enlarge the shadow areas. Some water pixels with low reflectance area are masked as cloud shadow pixels due to similar reflectance values in the NIR and SWIR bands. To reduce the misclassification in the shadow mask patches with a size less than 400 m (corresponding to four 10 x 10 m pixels) are eliminated automatically.
In most cases cloud and cloud shadows are located next to each other depending

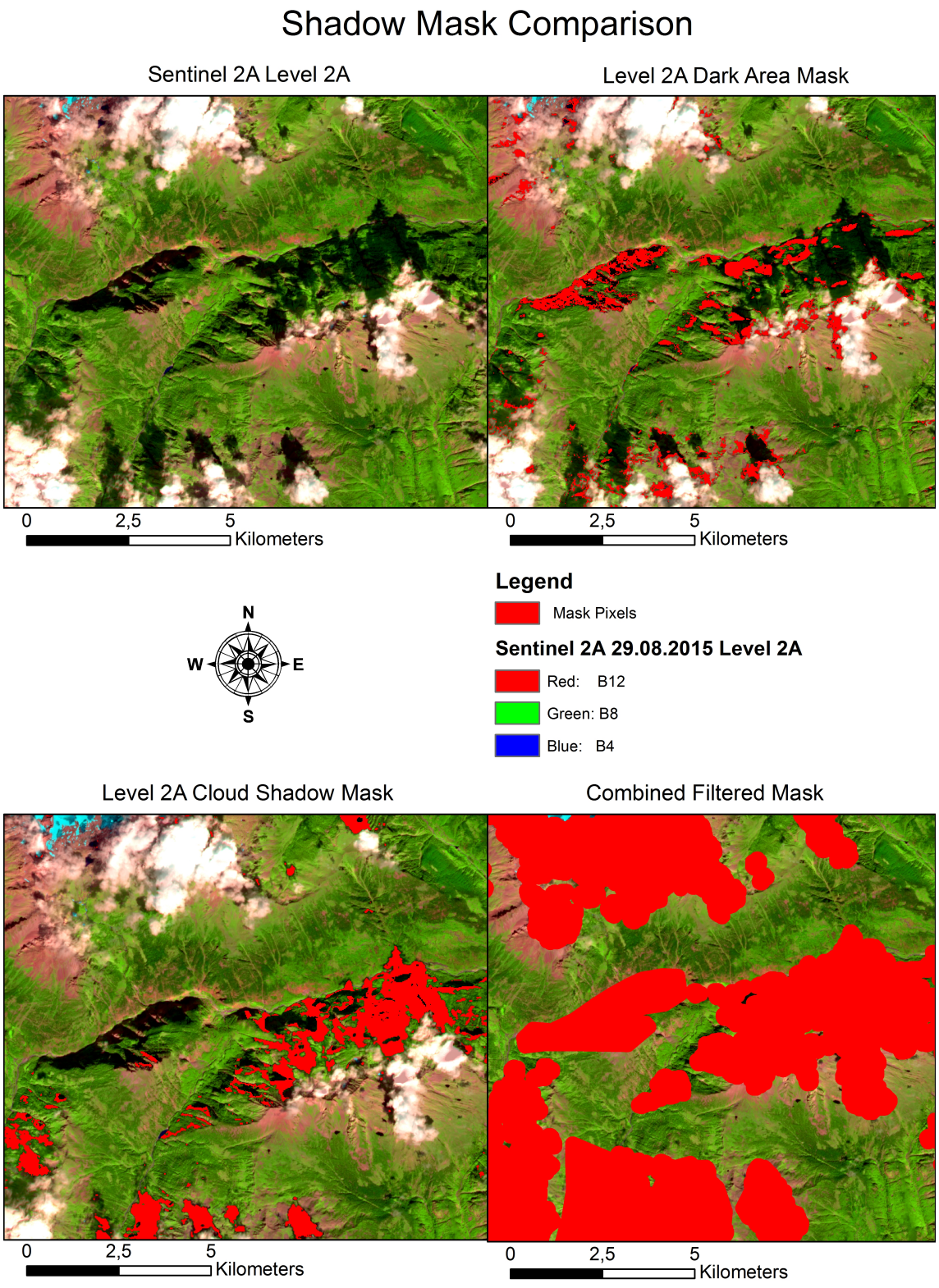


Figure 6.9.: Shadow mask: Comparison between terrain and cloud shadows.

on the sun location and image acquisition time and the shadow size should be nearly the same size as the cloud. To reduce the cloud shadow misclassification geometric relationships between clouds and cloud shadows could be used (Jin et al., 2013). Nevertheless, those misclassification can be neglected for the forest mapping purpose. Generally it should be considered that the used Sen2Cor version has been in a early tuning phase, and will be enhanced with next versions.

Finally, clouds, cloud shadows and terrain shadows are merged and excluded from the mosaic. The final product is represented in figure 6.9. Afterwards the gaps are restored with clean pixel values from overlapping regions or redundant scenes (sa1, sa2). Due to overlapping cloud regions and regions without second data source there are still no data values within the mosaic. Nevertheless, automated cloud screening is important to spend less time on data preparation especially regarding time series.

6.2. Trainings Data Preparation

As mentioned at the beginning the trainings data is collected through visual image interpretation at European scale for 2012 in the frame of the EEAVAL project provided by Joanneum Research. Since image data and plot data have different acquisition years, see chapter 4.2, it has been necessary in a first step to identify and relabel plots with changes between validation date and image acquisition date. All plots are screened for disturbances occurring when the spectral signal is no longer representing the actual label regarding changes due to clear cuts or partial clear cuts changes. Very few of these changes are found due to the close time and stable forest conditions. Those plots are updated according to the SENTINEL 2015 image data, consequently 1482 tree cover density and 1556 forest type plots are available within the whole mosaic. Furthermore, plots lying within cloudy or shadow areas are excluded, leaving 1013 tree cover density and 1108 forest type plots.

Figure 6.10 shows the final tree cover density value distribution within the trainings data set, showing that the non-forest class is overrepresented in relation to other tree cover density values. Consequently, another approach is tested only including tree cover density samples extracted by the forest mask obtained from Corine Land Cover. Figure 6.11 represents the distribution of the derived training samples, showing a more balanced trainings data set. It should be noted that the number of plots is decreased from 1013 to 475, if only plots within the Corine Land Cover forest mask are considered. This can be explained by the high percentage of Non-Forest plots within the EEA HR Layer validation data set. Moreover, 20 % of the plots are excluded due to accuracy assessment purposes

leaving only 380 points for the training process. Another interesting point is that according to figure 6.11 the tree cover density values corresponding to 4 % steps are more represented than others. This is explained by (Sannier et al., 2015):

> *"For estimating the tree crown cover, for each of the 25 sampling points it was interpreted if the point is located above a tree crown. Further, for each of the sampling units, a reliability estimate was assigned by the interpreter. The [validation] tool also provided the possibility to record a tree cover density value directly, which explains why validation density values are mostly, but not always in 4 % steps." (Sannier et al., 2015).*

Finally, the updated and extracted plot data is randomly divided in training areas and validation areas. Considering the accuracy assessment 15 % (regarding tree cover density classification with Non-Forest areas) or 20 % (regarding all classifications within forest areas) are extracted and excluded from the training process. It should be noted that the validation samples cannot be considered as fully independent, since they are gathered within the same framework and have to deal with the same interpretation error.

On the contrary, the forest type training plots are influenced by several different plot conditions as detected after the visual inspection of the plot data. One difficult case occurs for example if two thirds of the plot is covered with dense coniferous trees and the rest is covered with grassland mixing with the forest type signature. Combining the spectral signature of grassland and coniferous may cause a misclassification as broadleaved forest. Another case shown in figure 4.2d represents a coniferous forest type with low density. The underlying land cover influences the spectral signature, but still represents the proper class. The last case represented in figure 4.2b shows a plot fully covered by trees representing an almost pure signature for the forest type. This plot types are better suitable regarding a forest type estimation.

Under these circumstances it needs to be considered to exclude Non-Forest pixels from the trainings data to achieve more accurate forest type classification results. Usually researches in remote sensing avoid this issue by selecting homogeneous training observations or field plots. Theoretically, quantitative methods can be applied to identify heterogeneous plots for example those who exceed a heterogeneity threshold but in practice it has been proven to be extremely difficult to achieve desired results (Ohmann et al., 2014).

Alternatively, four options are tested within this thesis, to provide representative training data sets. The first option includes the identification of homogeneous training areas through visual interpretation. Consequently, homogeneous plots are identified by reviewing the plots in the JR validation tool using reference data listed in chapter 4.2. Since this involves a lot of human work and reduces

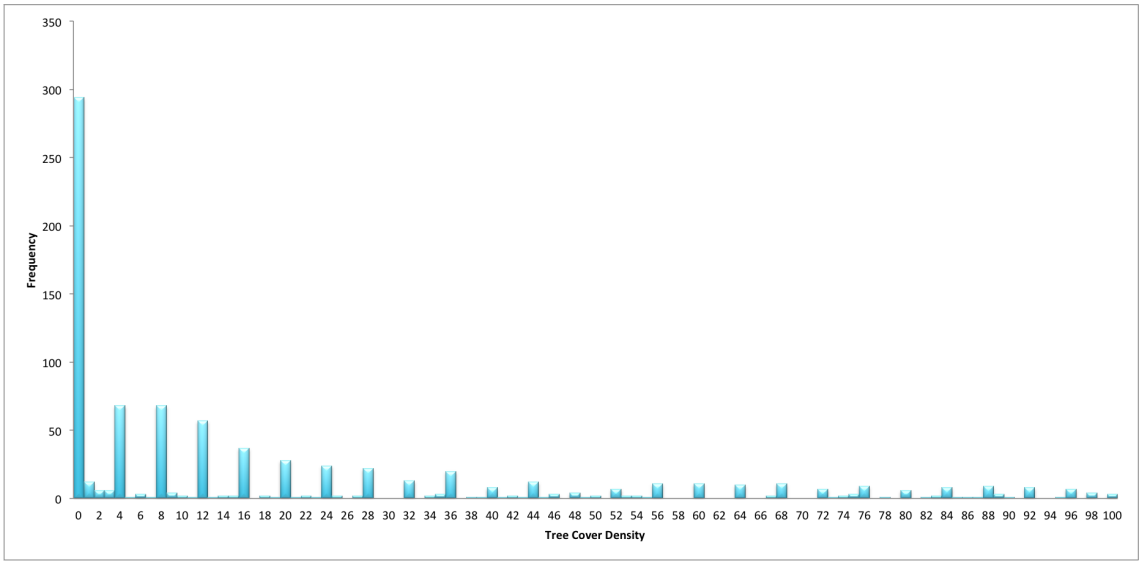


Figure 6.10.: Tree Cover Density training sample distribution

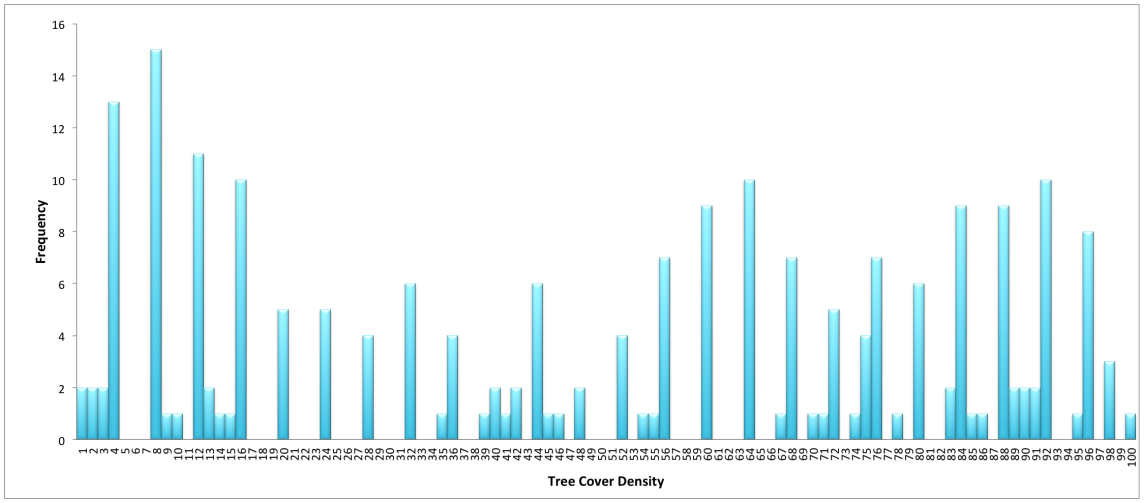
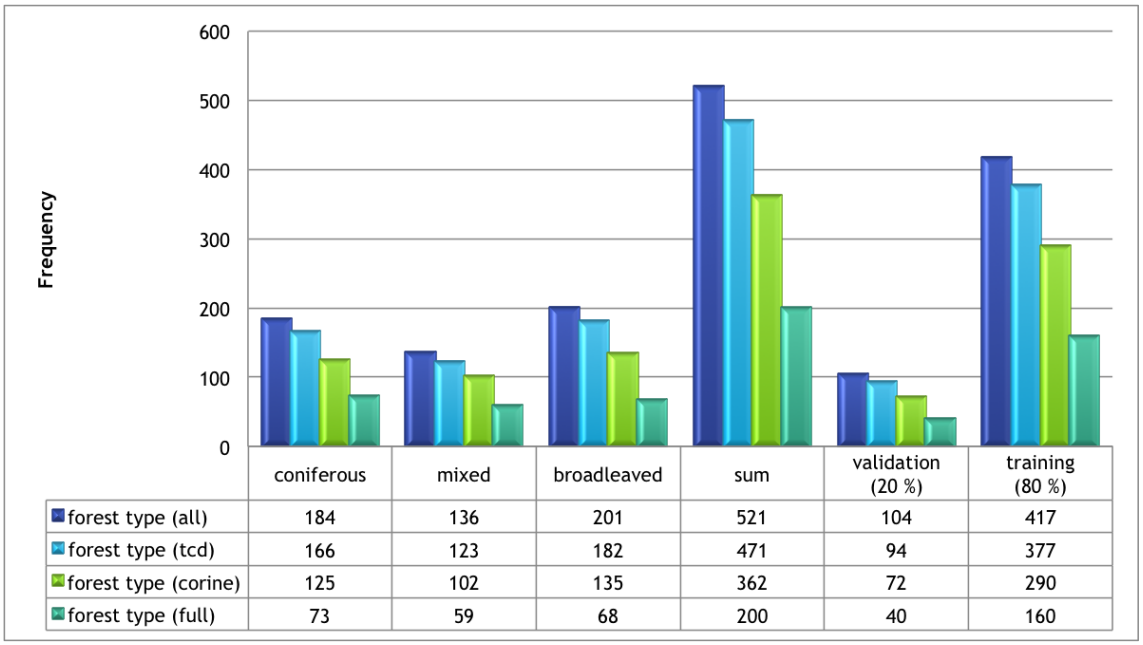


Figure 6.11.: Tree Cover Density training samples distribution within the Corine Land Cover forest mask



	coniferous	mixed	broadleaved	sum	validation (20 %)	training (80 %)
forest type (all)	184	136	201	521	104	417
forest type (tcd)	166	123	182	471	94	377
forest type (corine)	125	102	135	362	72	290
forest type (full)	73	59	68	200	40	160

Figure 6.12.: Forest Type training sample distribution, where forest type (all) includes all training samples, forest type (tcd) includes trainings samples extracted with the TCD forest mask, forest type (corine) includes trainings samples extracted with the Corine forest mask and forest type (full) includes trainings samples covered to 90 % with trees.

the amount of plots for training immensely a second method is tested including pixel extraction through a given forest mask. The second method extracts all 20 by 20 m pixels from the 100 by 100 m plots which do not overlap with the forest mask derived from the 2012 product. This process provides an increased automation degree but allows more inhomogeneous training samples resulting from different acquisitions dates between mask and image data. Considering the tree cover density product derived forest mask represents a tree non-tree mask instead of a forest mask a further approach is tested. Within this approach training samples are extracted with a forest mask derived from 2012 Corine Land Cover. The advantage of the Corine Land Cover frost mask is that wooded land and reforestations are included within the forest mask, which is not the case with the HRL Tree Cover Density product.

As a result four different trainings data sets are created. Figure 6.12 represents the class distribution within those training samples showing balanced data sets. The number of plots decreases from 1108 to 521, after plots located in cloud or shadow areas are excluded. Moreover, 20 % of the plots are excluded due to accuracy assessment purposes. The first data set (Forest Type (all)) includes all 521 plots spatially well distributed over the whole test site. After 20 % of the plots are excluded, 417 are left regarding the training process. The second trainings data set (Forest Type (tcd)) is based on the first data set and changes each plot size based

on the derived TCD mask. Within the training samples the plot areas are pruned with the forest mask resulting in plot areas only covering tree pixels reducing the plots size. This approach decreases the sample size to 377, eliminating forest labelled samples that are not overlapping with the tree cover density product of 2012. The third trainings data set (Forest Type (corine)) includes only plots within the Corine Land Cover forest mask resulting in 290 training samples. Those plots are fully located within forest areas including stocked areas and reforestations. The fourth trainings data set (Forest Type (full)) includes only plots which are 90 % covered by forest areas decreasing the trainings sample size to 160. Those number of those plots can be considered as pure.

6.2.1. TCD Reflectance Spectra Analysis

The major advantage using signature analysis is to investigate the basic spectral properties regarding different tree cover density values, which are related to the chlorophyll content. Within the SENTINEL scenes the spectral reflectance of tree cover density values are measured in 10 different bands. Since the classifiers only use spectral information, the spectral resolution refers to the ability to separate classes based on their spectral signature, setting the upper limit for the classification (Radoux et al., 2016).

Correlation analysis and linear regression analysis are used to estimate the relationship between tree cover density values and reflectance values of each individual band. Accordingly, the mean reflectance value is calculated for each TCD plot in each spectral band. Afterwards, the training areas are plotted against the reflectance values as represented with scattergrams in figure 6.13 and figure 6.14. Those represent the relationship between tree cover density values and reflectance values in each MSI sensor band. The tree cover density samples are plotted against reflectance in band 2 (blue), 3 (green), 4 (red), 5 (red-edge), 6 (red-edge), 7 (red-edge), 8 (NIR), 8a (NIR), 11 (SWIR) and 12 (SWIR) represented in figure 6.13 (a),(b),(c),(d),(e),(f) and figure 6.14 (a),(b),(c),(d), respectively.

The bands within the visible region of the spectrum show a decline in reflectance with increasing tree cover density values (figure 6.13a, 6.13b, 6.13c, 6.13d). The blue (B2) and red (B4) reflectance is weaker correlated to the tree cover density than the green part (B3) of the visible spectrum. The low R^2 values (see table 6.2 and 6.3) in the visible and NIR bands can be explained by the spectral characteristics of different tree species masking the spectral response on tree cover density. Pure broadleaved forest plots show higher reflectance values than mixed or coniferous forest plots. In the visible region the reflectance ranges 0 and 20 %.

Table 6.2.: Regression statistics regarding all TCD training samples

bands	R	R2	RMSE	slope	offset	p-value
B2	-0,26	0,07	0,04	-0,04	4,81	5,6032e-14
B3	-0,37	0,14	0,04	-0,06	7,81	8,3030e-28
B4	-0,36	0,13	0,05	-0,07	7,40	4,9978e-27
B5	-0,44	0,19	0,05	-0,09	12,45	9,7564e-40
B6	-0,21	0,04	0,07	-0,06	24,68	8,9661e-10
B7	-0,13	0,02	0,08	-0,04	25,44	2,6301e-04
B8	-0,14	0,02	0,08	-0,05	29,65	3,9263e-05
B8a	-0,13	0,02	0,09	-0,04	31,10	1,0827e-04
B11	-0,45	0,20	0,06	-0,12	23,11	9,8730e-42
B12	-0,46	0,21	0,05	-0,10	14,40	2,6068e-43

Table 6.3.: Regression statistics regarding TCD training samples within forest areas

bands	R	R2	RMSE	slope	offset	p-value
B2	-0,49	0,24	0,01	-0,03	3,49	8,5825e-15
B3	-0,60	0,36	0,02	-0,04	6,14	1,3188e-22
B4	-0,59	0,35	0,02	-0,04	5,30	3,1481e-22
B5	-0,63	0,39	0,02	-0,06	10,35	7,7134e-25
B6	-0,21	0,04	0,06	-0,04	23,35	0,0023
B7	-0,09	0,01	0,07	-0,02	26,57	0,0015
B8	-0,09	0,01	0,08	-0,02	27,82	7,3970e-05
B8a	-0,08	0,01	0,08	-0,02	29,07	1,407e-04
B11	-0,46	0,22	0,05	-0,08	19,67	3,8097e-13
B12	-0,57	0,33	0,03	-0,07	11,37	1,9964e-20

The first red-edge Band B5 shows similar characteristics than the visible channels while the other red-edge bands share characteristic with the NIR channels of the spectrum (see table 6.2 and 6.3). However, in the NIR and the red-edge part of the spectrum only a weak relationship can be observed regarding the tree cover density data with R^2 values reaching from 0,02 to 0,04 (figure 6.13e, 6.13f, 6.14a, 6.14b and table 6.2). In comparison to the bands in the visible region of the spectrum the R and R^2 (see table 6.2 and 6.3) in red-edge and NIR are lower indicating that reflectance differences regarding the tree species are larger.

Furthermore, due to the chlorophyll sensitivity of the NIR region the offset values are higher. Concerning, all bands from the visible, red-edge and the NIR region the slope is near to zero, implying that within those spectral regions tree cover density values are more difficult to differentiate. Moreover, nearly no difference between band 7, 8 and 8a is observed. In this context it should be considered that a dependency between the tree cover density and the reflectance values could be

overlapped by the different forest types broadleaved or coniferous. As table 6.2 shows B6 is the optimal band in the NIR region for capturing the tree cover density differences based on the trainings data set.

In the SWIR wavelength the training samples show the highest R and R^2 values throughout all bands, indicating a stronger relationship between reflectance values and tree cover density values (figure 6.14c, 6.14d and table 6.2). The high slope values show that the differences between the tree cover density values are quite higher than in all other bands. Moreover, the RMSE values are lower in the SWIR region than in the NIR region but still higher than in the visible wavelength, indicating that the correlation is influenced by the residuals. Furthermore, it should be noted that the R^2 shows the second best result, although the RMSE is quite high. With a R^2 of 0.33 and high RMSE values no high classification accuracy would be expected.

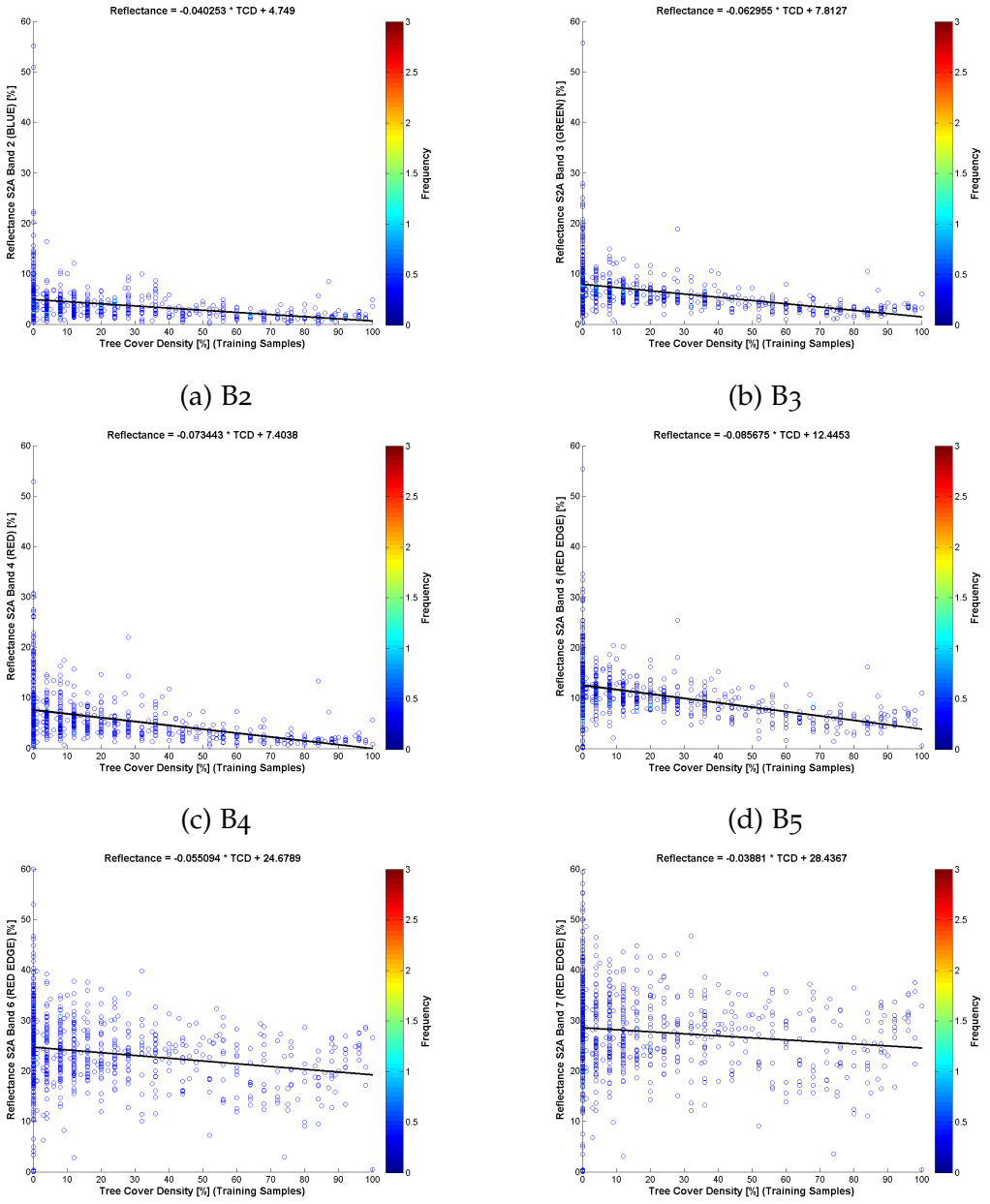
Regarding all different bands the R^2 values are quite low ranging from 0.02 to 0.21 indicating the relationship between the TCD and the reflectance variables is weak. Further, it should be noted that low tree cover density reflectance values are influenced by the underlying land cover. In most cases the underlying land cover is meadow, which has relative high reflectance values compared to forests. Nevertheless, the RMSE, which measures the average deviation from the tree cover density reflectance values, is quite low for all bands ranging from 0.04 to 0.08. This reinforces the assumption that the spectral signature differences regarding the tree cover density values are influenced by the different forest types. The correlation patterns are similar for all bands due to a weak negative correlation between the two parameters reflectance and tree cover density. Band 12, corresponding to the SWIR reflectance, is the band showing the strongest correlation with tree cover density, but the with an quite high RMSE value, indicating that no high classification accuracy can be expected. Higher accuracies are expected from bands with high R^2 and low RMSE values, for example B3 and B5. Bands 7, 8, 8a and 6 show weaker correlations and higher RMSE values.

Generally, some individual points are detected that fall outside the overall pattern of the scatter plot. Those outliers might have an influence on correlation but not on the classification using the random forest algorithm. However, the huge variations in the class zero (No-Trees) can be explained by the large class definition. Class zero includes several different landscape objects with different spectral attributes like trees, meadows, water bodies, urban areas and many more. The spectral variability (see table 6.2 and 6.3) regarding this class is bigger than with all other classes distorting the evaluation.

In order to evaluate the relationship between tree cover density and reflectance values without the distortion of non-tree samples, the regression is demonstrated regarding training samples only within the forest areas. Consequently, all training

samples outside the forest mask are excluded. Figure 6.15 and figure 6.16 show the results of the new approach. Removing the non-tree plots from the analysis shows a slight strengthening of the negative relationship in all bands. Under these circumstances the slope of each regression decreases.

As shown in table 6.3 the R and R^2 values rise by removing trainings samples outside the forest areas, implying a stronger correlation without those plots. However, nearly no change in the corresponding slope value can be observed after the sample removal. The low p - values in table 6.2 and table 6.3 imply that all spectral bands significantly correlate ($p < 0.05$).

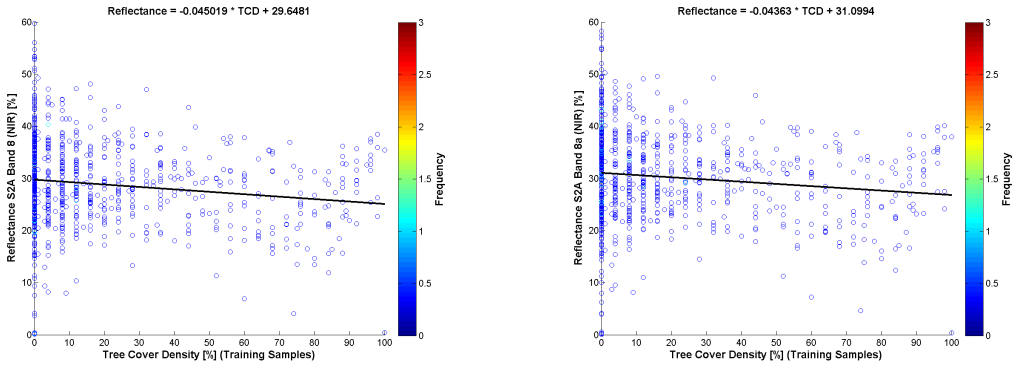


(a) B2 (b) B3

(c) B4 (d) B5

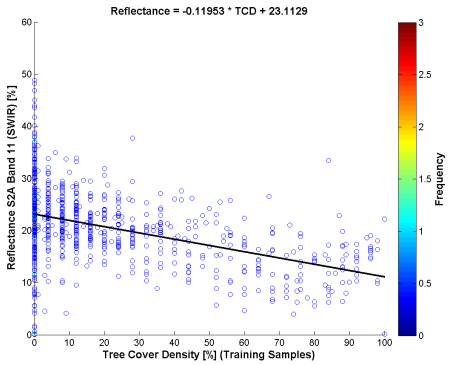
(e) B6 (f) B7

Figure 6.13.: Scatter plots of all TCD training samples and each SENTINEL band. B2 - B7

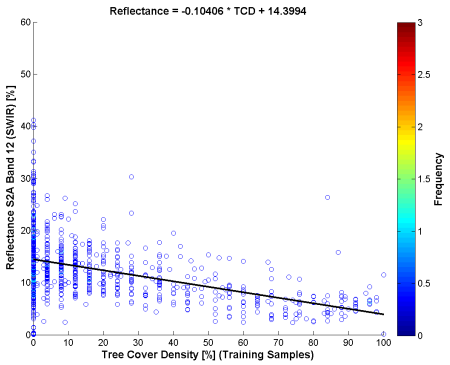


(a) B8

(b) B8a

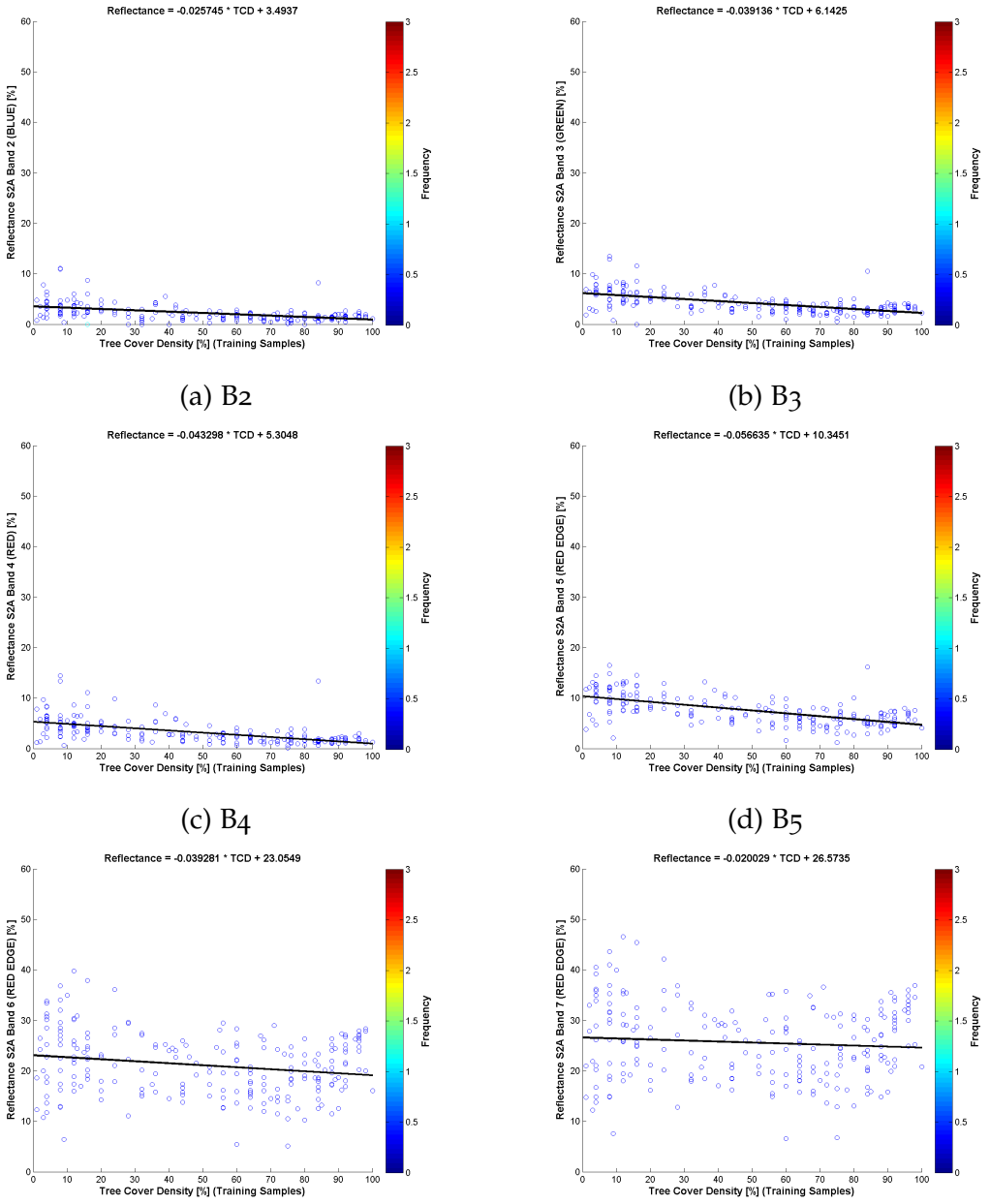


(c) B11



(d) B12

Figure 6.14.: Scatter plots of all TCD training samples and each SENTINEL band. B8 - B12

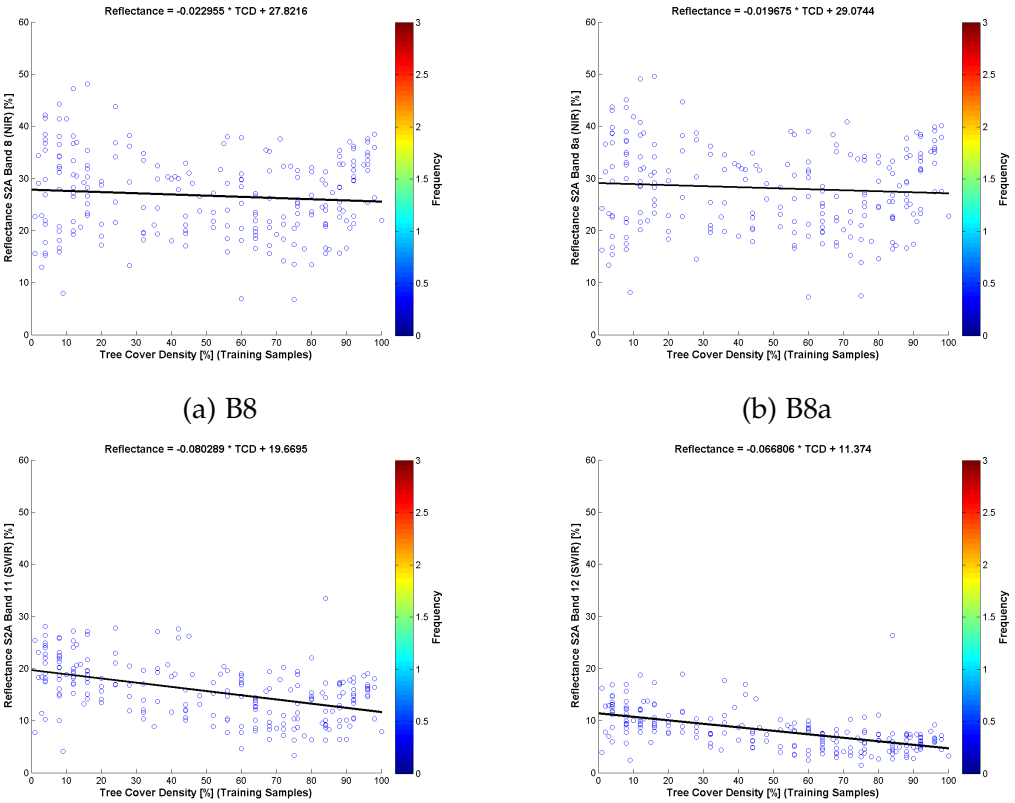


(a) B2 (b) B3

(c) B4 (d) B5

(e) B6 (f) B7

Figure 6.15.: Scatter plots of TCD training samples within forest areas and each SENTINEL band. B2 - B7



(a) B8 (b) B8a

(c) B11 (d) B12

Figure 6.16.: Scatter plots of TCD training samples within forest areas and each SENTINEL band. B8 - B12

6.2.2. FTY Separability and Signature Analysis

The potential of forest type discrimination of the Sentinel-2 bands is analysed by signature evaluations and separability using the Jefferies-Matusita separability metric and feature space representations. The separability evaluation is performed twice. In the analyses both, each individual MSI band and selected band combinations are considered. As table 6.4 shows B4 in the visible region of the spectrum, B8a in the red-edge / NIR region and B11 in the SWIR region achieved the best separation results regarding all three forest types throughout all individual bands provided by the MSI sensor. Therefore B11, B8a and B4 are chosen in figure 6.17 for further analysis.

Figure 6.17 shows the differences between the forest type classes within different tree cover density groups regarding three selected feature spaces of band B4, B8a and B11. In this figure the ellipses are generated with a standard deviation of one, representing the mean reflectance values and the variance throughout the

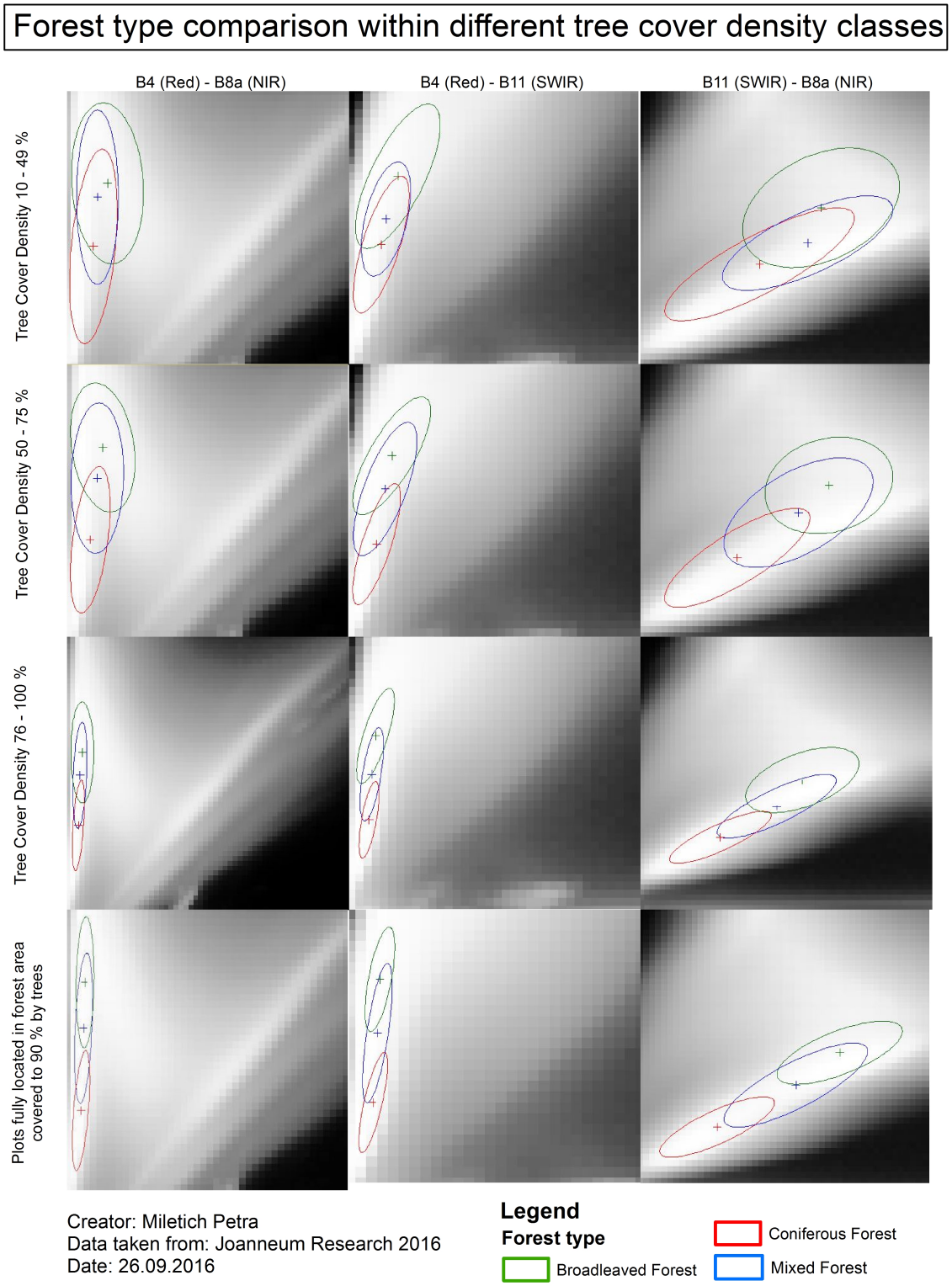


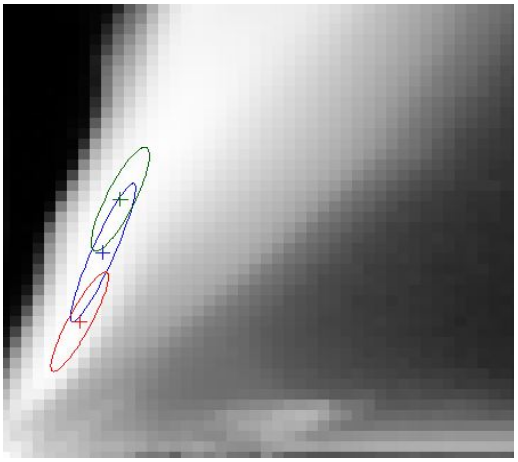
Figure 6.17.: Forest Type and Density feature space comparison

Table 6.4.: Separability analysis regarding each individual band with Jefferies-Matusita distance measure

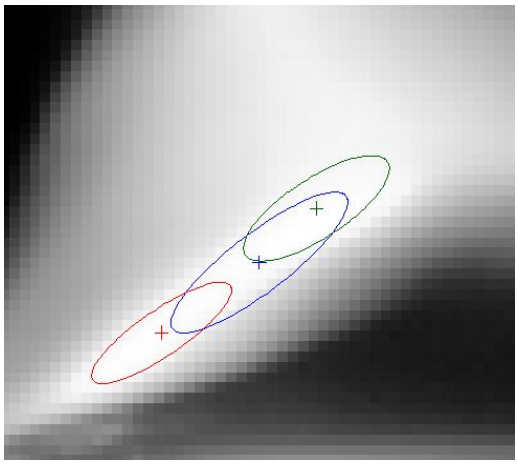
Sentinel-2 Band	Spectral Region	Average	Minimum
B2	Blue	209	116
B3	Green	208	155
B4	Red	250	128
B5	red-edge	255	182
B6	red-edge	370	231
B7	red-edge	374	207
B8	NIR	353	202
B8a	red-edge	382	213
B11	SWIR	407	308
B12	SWIR	327	211

training samples merged based on their forest type and tree cover density class membership. Higher density values result in smaller ellipses, considering that less training samples are available with higher density values. The underlying land cover types make the discrimination of the forest types with lower density values more challenging. Accordingly, the separation of broadleaved and coniferous tress is not always unambiguous. The same is true for the mixed class, sharing reflectance characteristics from both coniferous and broadleaved forests. The training set with plots fully located within forest area and covered to 90 % with trees, gathered through visual inspection (explained in chapter 6.2), is found to be better suited regarding the forest type discrimination, but it should be noted that with this training set only 200 training and 40 validation plots are left representing those conditions. Nevertheless, those plots more suitable regarding the forest type discrimination, due to the minimization of the negative influence of other land cover types on the signatures.

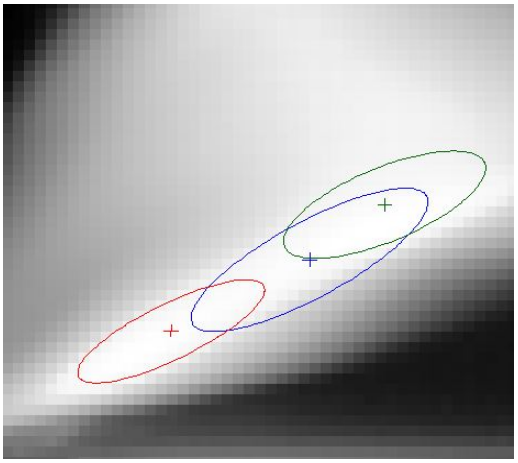
Legend:
O Broadleaved Forest
O Coniferous Forest
O Mixed Forest
\+ Mean



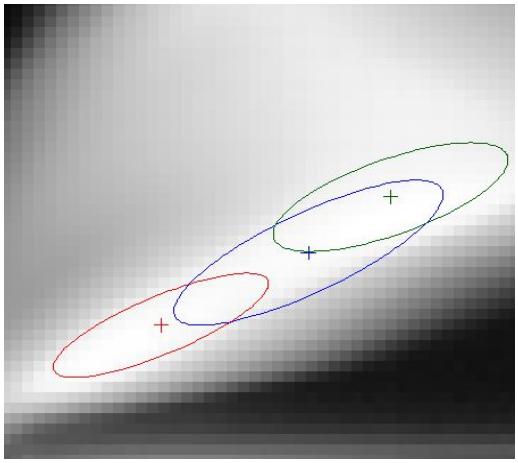
(a) Feature space: x-axis: B11 (SWIR), y-axis: B5 (RED-EDGE)



(b) Feature space: x-axis: B11 (SWIR), y-axis: B6 (RED-EDGE)



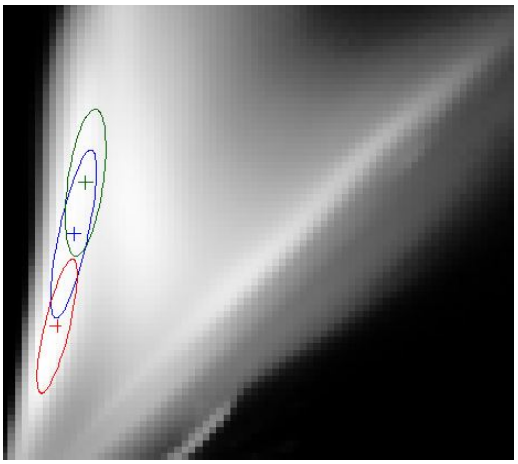
(c) Feature space: x-axis: B11 (SWIR), y-axis: B7 (RED-EDGE)



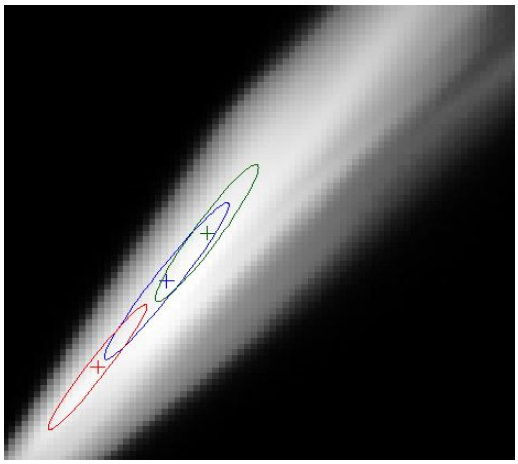
(d) Feature space: x-axis: B11 (SWIR), y-axis: B8a (RED-EDGE)

Figure 6.18.: Sentinel-2A feature spaces with FTY trainings data (full trainings data set) considering B11 combined with B5, B6, B7 and B8a.

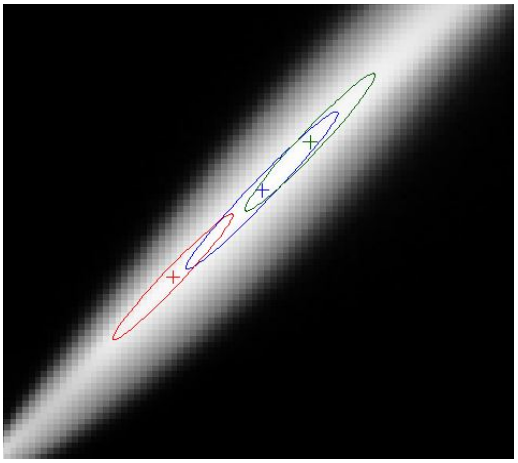
Legend:
O Broadleaved Forest
O Coniferous Forest
O Mixed Forest
+ Mean



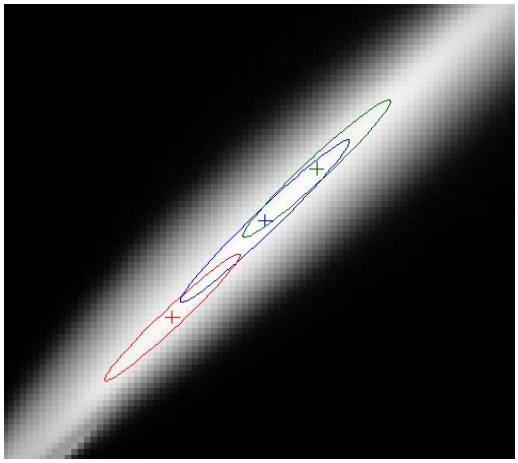
(a) Feature space: x-axis: B8 (NIR), y-axis: B5 (RED-EDGE)



(b) Feature space: x-axis: B8 (NIR), y-axis: B6 (RED-EDGE)



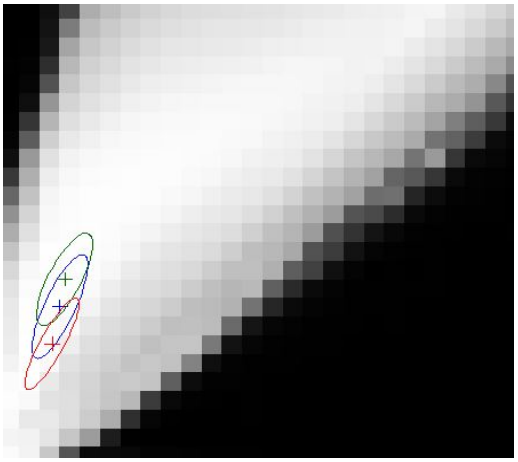
(c) Feature space: x-axis: B8 (NIR), y-axis: B7 (RED-EDGE)



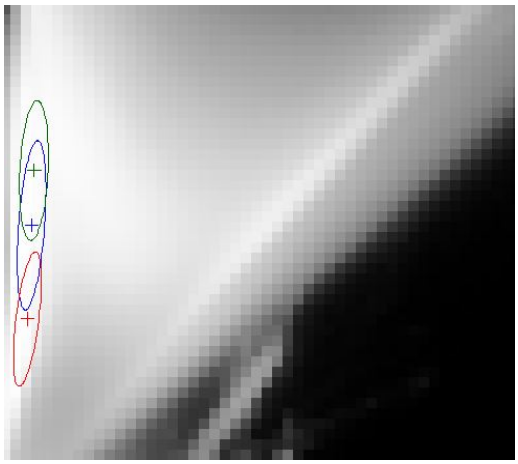
(d) Feature space: x-axis: B8 (NIR), y-axis: B8a (RED-EDGE)

Figure 6.19.: Sentinel-2A feature spaces with FTY trainings data (full trainings data set) considering B8 combined with B5, B6, B7 and B8a.

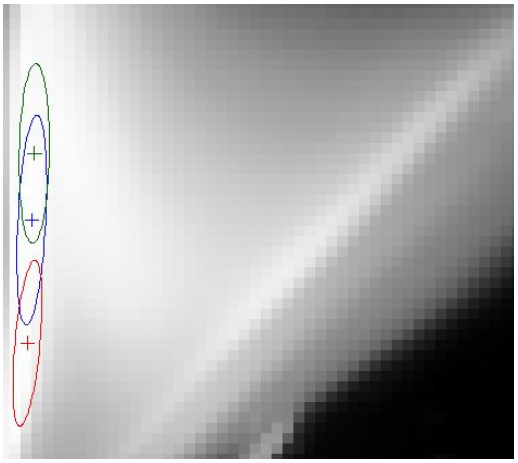
Legend:
O Broadleaved Forest
O Coniferous Forest
O Mixed Forest
+ Mean



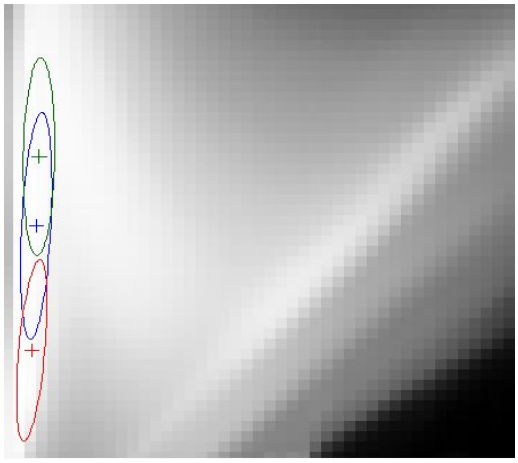
(a) Feature space: x-axis: B4 (RED), y-axis: B5 (RED-EDGE)



(b) Feature space: x-axis: B4 (RED), y-axis: B6 (RED-EDGE)



(c) Feature space: x-axis: B4 (RED), y-axis: B7 (RED-EDGE)



(d) Feature space: x-axis: B4 (RED), y-axis: B8a (RED-EDGE)

Figure 6.20.: Sentinel-2A feature spaces with FTY trainings data (full trainings data set) considering B4 combined with B5, B6, B7 and B8a.

Since Sentinel-2A is supported with four new bands in the red-edge region located in the spectral region between RED and NIR, particular intention is given to the benefit of those new bands regarding forest type separation. In accordance to the separability evaluation the bands B8a and B11 and the full trainings data set are considered since they show significant better results represented in figure 6.17. As mentioned in chapter 6.2 the full trainings data set consist only of plots fully located within forest areas covered to 90 % with trees, derived through visual interpretation.

The first red-edge band B4 is located close to the red region of the spectrum whereas the other three bands B5, B6 and B7 correlating strongly with the NIR band (see figure 6.20). Comparing the new red edge bands with B8 corresponding to the NIR wavelength, indicates that no addition information can be obtained using those bands in combination regarding the forest type discrimination (see figure 6.19).

Figure 6.18 shows the forest type separation differences between the red-edge bands plotted against B11. Between B7 and B8a nearly no difference can be observed in combination with B11, both providing the same information regarding the forest type discrimination as shown in figure 6.18c and 6.18d. Between B6 and B7 a small difference can be observed, as shown in figure 6.18b and 6.18c, since the bands are located farther away in the spectrum. B5 which is located near to the red wavelength region provides different information. The best separation between the forest types broadleaved forest and coniferous forest can be achieved with B7 in combination with B11 (see figure 6.18 and table 6.4).

As shown in figure 6.20 the B5 correlates strong with B4 due to their relative near location in the spectrum consequently, the forest types are worse separable wit B5 than with the other red-edge bands. The best separation can be achieved again with B7 in combination with B4. These results are also substantiated by the separability analysis of each individual band represented in table 6.4 showing higher separability values.

The results are similar to a study presented by Radoux et al. (2016) comparing the separability for the land cover classes broadleaved, needle leafed forest and grassland/pasture among others. Their study shows that several spectral bands can be used to discriminate broadleaved forest and grasslands or pastures. Radoux et al. (2016) state that most efficient bands are Green and first red-edge band (B5). Regarding the discrimination of different forest types the SWIR bands are considered as useful whereas B12 can be rated better than B11. Further, the authors point out that Within all infra-red bands the smaller B8a band is better regarding discrimination forest types than the wide infra-red band B8 (Radoux et al., 2016).

Table 6.5.: Separability analysis regarding different band combinations with Jefferies-Matusita distance measure

Abbreviation	Band Combination	Average	Minimum
all	all bands	814	686
wb	all bands without B2	786	669
wbg	all bands without B2 and B3	737	578
2L	B5, B6, B7, B8a, B11, B12	700	545
3L	B4, B5, B7, B8a, B11, B12	719	570
4L	B4, B5, B6, B8a, B11, B12	713	571
5L	B5, B6, B12	588	462
6L	B4, B8a, B11	600	467
7L	B3, B5, B6, B8a, B11, B12	710	577

The separability is assessed using Jefferies-Matusita calculating the separability regarding different band combinations to determine which set of bands is the most useful within the classification process. Table 6.5 contains the average divergence and the minimum divergence for each band combination set. The distance measure ranges from zero to a maximum of 1414. The higher the value the better the signatures are separable in the bands being chosen. A calculated Jeffries-Matusita distance of zero indicates that the signatures are inseparable (Appiah et al., 2015).

As mentioned above and represented in figure 6.18, 6.20 and 6.19 some bands contain redundant informations regarding forest type differentiation for example B4 and B5 in the red region, B6 and B7 in the red-edge region as well as B8 and B8a. Accordingly, several band combinations are tested by removing redundant bands. First, all bands are tested achieving a result of 803 as shown in table 6.5. As the visible region is highly influenced by errors in the atmospheric correction process it has been decided to discard the blue band from the classification process. Furthermore, it is tested if removing both blue and green have a severe impact on the separability resulting in even worse measures. Accordingly, the band combination L7 is tested including the green band, achieving a measure of 705. In addition L2 evaluates the impact on the separability discarding all visible bands but including the first red-edge band. Again no improvement can be achieved. The band combinations L3 and L6 are tested in respect to the forest type feature space evaluation containing bands with high individual separability measures. One combination provided with 6 bands, two in each spectral region and the other combination with 3 bands only containing those bands which achieve the best results. Same for the combinations L4 and L5 in respect to the tree cover density correlation evaluation.

It follows that the best average separability can be achieved by using all bands, in-

Table 6.6.: Separability measure with all bands

Signature Name	Coniferous	Mixed	Broadleaved
Coniferous Forest	0	725	1030
Mixed Forest	725	0	689
Broadleaved Forest	1030	689	0

dicating that no band can be eliminated without losing any information. Table 6.6 shows the separability regarding all different classes with all bands in combination. As expected coniferous forest and broadleaved forest are better separable from each other than with the mixed forest class. The minimum separability is achieved between mixed forest and broadleaved forest sharing similar spectral characteristics. Considering these findings confusions between broadleaved forest and mixed forest pixels in the classification results can be expected.

6.3. Classification Process

Regarding the classification and the prediction process the two classifiers random forest and k Nearest Neighbours are chosen and will be explained in detailed in the following sections. To estimate the impact of the different parameters on the classification accuracy various parameter values based on literature recommendations are applied on the 100 m spatial resolution mosaic. Within the continuous TCD prediction regression analysis are obtained regarding all parameters tested, examining the product accuracy with determination coefficients and regression coefficients. Therefore, both data sets, including Non-Forest areas or not, are used. As far as the FTY classification is concerned the categorical classification products are evaluated with overall accuracies and kappa values calculated from confusion matrices. Both trainings data sets include plots, which are located within a derived 2012 Corine Land Cover forest mask. Further, it needs to be mentioned that several aggregation rules, have been applied within the validation process, which need to be considered before accessing the accuracy regarding the forest type product. Those aggregation rules need to be considered due to the forest definition given by the FAO. In contrast to the forest type product no aggregation rules where applied regarding the tree cover density product throughout the validation process. According to the tree cover density layer description no specific definitions are given.

Table 6.7.: Random forest parameter tests using TCD samples. Default values for each parameter: t = 100, $n = \sqrt{N}$, s = 10, d = 5; where N is the number of training samples

trees (t)	R	R^2	RMSE	slope	offset
50	0.8053	0.6485	17.7883	0.6187	15.9838
100	0.8066	0.6506	17.7347	0.6179	16.3818
300	0.8040	0.6464	17.8406	0.6073	17.3040
500	0.8012	0.6420	17.9528	0.6056	17.3897
700	0.7984	0.6375	18.0656	0.6030	17.4114
nodes (n)	R	R^2	RMSE	slope	offset
5	0.7990	0.6384	18.0417	0.6109	16.8875
20	0.7838	0.6143	18.6330	0.6221	16.7066
40	0.7838	0.6143	18.6330	0.6221	16.7066
60	0.7838	0.6143	18.6330	0.6221	16.7066
samples (s)	R	R^2	RMSE	slope	offset
2	0.8041	0.6466	17.8374	0.6129	17.2107
4	0.7603	0.5781	17.1025	0.6789	15.4893
5	0.7896	0.6235	18.4114	0.5968	18.0867
10	0.8066	0.6506	17.7347	0.6179	16.3818
depth (d)	R	R^2	RMSE	slope	offset
3	0.7913	0.6261	18.3462	0.5887	17.9372
5	0.8066	0.6506	17.7347	0.6179	16.3818
10	0.7850	0.6162	18.5882	0.6103	16.9500
30	0.7846	0.6156	18.6036	0.6083	16.4960

Table 6.8.: Random forest parameter tests using FTY samples. Default values for each parameter: t = 100, $n = \sqrt{N}$, s = 10, d = 5; where N is the number of training samples

trees (t)	Overall Accuracy	Kappa
50	0,595238	0,347051
100	0.6071	0.3671
300	0.6071	0.3671
500	0.6071	0.3671
700	0.607143	0.365239
nodes (n)	Overall Accuracy	Kappa
5	0.6071	0.3671
10	0.6071	0.3671
20	0.6071	0.3671
30	0.6071	0.3671
40	0.6071	0.3671
samples (s)	Overall Accuracy	Kappa
2	0,619048	0,385038
4	0,619048	0,385038
5	0,607143	0,365239
10	0,607143	0,365239
20	0,607143	0,365239
depth (d)		
3	0,547619	0,269398
5	0,607143	0,365239
10	0,583333	0,336792
20	0,535714	0,265141
30	0,571429	0,316765

6.3.1. Random Forest Classifier

The random forest algorithm first proposed by BREIMAN 2001 belongs along with other boosting and bagging methods as well as classification trees in general to the ensemble learning methods, which generate many classifiers and aggregate their results to calculate their response (Liaw and Wiener, 2002; Horning et al., 2010; T. Li et al., 2016). The random forest algorithm learns the relationship between predictor and response data and can handle continuous, categorical and binary data sets (Ali et al., 2012; Horning et al., 2010; Grinand et al., 2013)

Since the random forest algorithm generates multiple decision trees it is necessary to explain how the decision tree algorithm works to generate one single tree. Decision trees use a set of binary rules to evaluate the target class. However, two types of trees are differentiated, classification trees for categorical data sets and regression trees for continuous data sets (Horning et al., 2010). "The tree is a set of binary decisions and terminal nodes connected by branches"(Horning et al., 2010). The binary decision tree uses 'greater or less than' rules at each node to decide for a class label. In other words the feature space is partitioned into n dimensional rectangles depending on the number of bands available. Those rectangles are rarely effective to partition the feature space and tend to over-fit the training data, in cases where the tree grows too large and the nodes represent only small subsets of the training data. To achieve better results and reduce over-fitting effect trees can be pruned, by removing nodes that are linked to noise in the training data (Horning et al., 2010).

As mentioned above the random forest algorithm generates multiple decision trees with random drawn subsets, instead of using all variables from the available data. The subsets are drawn with replacement, meaning that one sample can be selected several times, while others may not be selected at all (Belgiu and Drăguţ, 2016; Ali et al., 2012). Regarding each random sample a classification or regression tree is grown to the largest possible extent without pruning. At each node a random sample of predictor variable is extracted, among those the best split is chosen. To predict new data the prediction among all trees are aggregated using majority votes regarding the classification trees and the average for the regression trees (Liaw and Wiener, 2002; Ali et al., 2012). The class with the maximum vote over all decision trees is the one selected for the output product. Generally two thirds of the samples are used for training and one third for errors estimation calculating the out-of-bag (OOB) error (Belgiu and Drăguţ, 2016; Ali et al., 2012; Horning et al., 2010). In comparison to decision tree algorithms over-fitting is less of an issue, due to two steps of random selection used by the random forest algorithm. First each tree is trained with a randomly drawn subset. Second a splitting variable is chosen at each node from randomly selected attributes (T. Li et al., 2016). Using the randomly selected variables for splitting

the nodes results in a smaller correlation between the different trees, but also results in less predictive power. Accordingly, the number of nodes used, needs to be tested for each classification process individually (Horning et al., 2010).

The random forest algorithm offers a good prediction performance and is computationally effective but sensitive to the sample design. Colditz (2015) tested the impact on several sampling designs on decision tree algorithms and recommends the area-proportional allocation to achieve the best classification results, because classes occupying larger areas need more training samples. Mellor et al. (2015) and Ali et al. (2012) found out that the random forest algorithm is less sensitive to outlier training samples or noisy data. Furthermore, Dalponte et al. (2013) proposed that the algorithm fails to cope with imbalanced training data tending to favour the most representative class at the expense of the minority class. Thus, at each sample selection at each node during the tree construction fewer samples of the minority class are chosen. Also the size of the trainings data has an impact on the classification accuracy. Colditz (2015) recommend a sample size of 0,25 % of the whole study area.

Horning et al. (2010) pointed out the limitations using the random forest algorithm especially with regression models. The algorithm cannot predict values beyond the range of the response values in the data. It is important that each training subset covers the entire range of response values. Another issue is the algorithms tendency to overestimate the low values and underestimate the high values, because the response value of all trees is calculate by the average of all single trees.

Belgiu and Drăguţ (2016) point out that the trainings samples need to fulfil following requirements: ”

1. *training and validation data must be statistically independent*
2. *training samples must be class balanced*
3. *training samples must be representative of the target classes*
4. *and the training sample needs to be large enough to accommodate the increasing number of data dimensions.” (Belgiu and Drăguţ, 2016)*

The training samples regarding the tree cover density prediction fulfil most of the requirements mentioned by Belgiu and Drăguţ (2016), but the given training data set is imbalanced. As mentioned in chapter 6.2 there is a huge amount of Non- Forest training areas available within the original data set but less for high tree cover density values as shown in figure 6.10. The exclusion of Non-Forest samples using the Corine Land Cover forest mask shows a more balanced data set (see figure 6.11). Turning to the forest type training samples figure 6.12 again shows more balanced trainings data sets.

One of the major advantages of the random forest algorithm is that only a small amount of parameters need to be set (Liaw and Wiener, 2002). To estimate the impact of the different parameters on the forest type classification accuracy various values based on literature recommendations are applied on the 100 m spatial resolution mosaic. The amount of trees calculated, the nodes, the samples and the maximum depth of the trees are evaluated in combination with the default values of the Oreo Toolbox. Although a range of parameter values are tested including values suggested by literature to find the best parameter for classification it is possible that the parameter selected are still arbitrary regarding completely different trainings data sets.

The first parameter is the number of decision trees generated (tree parameter [t]). Belgiu and Drăguţ (2016) states that the selection of a large number of trees result in higher accuracies and less over-fitting but increases the prediction time linearly. Table 6.7 presenting the tree cover density results shows that increasing the number of trees increases R^2 minimally till reaching the turning point at 500 trees. In general, it seems that the number of trees has no severe impact on the correlation between the predicted values and observed tree cover density values. The RMSE is stable regarding all different tree parameter values. The offset value shows more fluctuations than other coefficients.Further, forest type results represented in table 6.8 show no difference in the overall accuracies between higher tree values.
Since the algorithm is robust against over fitting the amount of calculated trees can be as large as possible. However, it needs to be considered that a large number of trees increase computation time (Belgiu and Drăguţ, 2016). Belgiu and Drăguţ (2016) also mention that the majority of studies propose a tree parameter of 500, since the error stabilize at this point.

The second parameter is the number of variables to be selected for splitting (node parameter [n]), which influences the correlation between trees and the predictive power and is set to the square root of the number of input variable by default (Belgiu and Drăguţ, 2016). Increasing the node parameter has no effect on the R^2 values within the tree cover density prediction. Setting the node parameter to 5 achieved a slight higher R^2 value. In general, it seems that the node parameter has no significant impact on any of the regression coefficients. Further, the forest type results represented in table 6.8 show no difference between the amount of node at all. Even Belgiu and Drăguţ (2016) mention that the classification accuracy is less sensitive to the amount of trees calculated than to the amount of variable randomly chosen for gathering the best split at each node. Regarding the node parameter it also needs to be considered that a higher number of nodes will increase the computation time since the algorithm needs to calculate the information gain for each variable at each node (Belgiu and Drăguţ, 2016; Horning et al., 2010).

Nevertheless, with the statistics for the tree cover density classification presented in table 6.7 and the forest type classification in table 6.8 no significant difference between accuracy of the node and the tree parameter can be observed.

As third parameter the minimum number of samples in each node can be chosen (samples parameter [s]). If the number of sample in the node is smaller than this parameter the node will not be split. The Orfeo toolbox recommend a small percentage of the total data (ORFEO Toolbox, 2016b). Table 6.7 referring to the tree cover density prediction does not show a trend between different sample parameter as both the R^2 and the regression coefficients seems to vary arbitrary. The different R^2 value suggest to use the default sample size of 10. The results of the confusion matrices in table 6.8 show that the sample parameter also has almost no impact on the overall accuracy and kappa values. Choosing a lower sample value as recommend by the Orfeo toolbox results in slightly higher overall accuracy and kappa values, therefore the node parameter is set to 4 regarding further classification.

The last parameter which can be chosen by the user is the depth of the trees showing more significant fluctuations in the overall accuracy and kappa statistics (table 6.8) similar as the R^2 and regression coefficients (table 6.7). Therefore, the impact of the depth parameter on the correlation and accuracy is quite higher than the previous parameter. As the node value decreases the accuracy and the R^2 values increases till reaching the turning point at 5.

In general the R^2 value and the regression coefficients similar to the overall accuracy and kappa values are less sensitive to the amount of trees and the size of the randomly selected subset of features at each tree node than to the minimum number of sample in each node and the depth of the trees.

6.3.2. K - Nearest Neighbours Classifier

The k Nearest Neighbours (kNN) classifier is a non-parametric classification method based on similarity functions, this means no assumption regarding the underlying data distribution are made. It uses a similarity measure to find the nearest neighbours among the training samples. Within the kNN classifier no generalization on the trainings-data is performed. This means the decisions are based on the entire training data set or a subset. The classifier works equally well with arbitrary number of classes (Hassanat et al., 2014; Blanzieri and Melgani, 2008; Baffetta et al., 2012; Thessler et al., 2008; Gjertsen, 2007).

The prediction class is assigned among the training points with the most common amongst its k nearest neighbours measured by a distance function. This means that the similarity depends on a specific distance measure, while the number of

Table 6.9.: K Nearest Neighbours parameter tests with TCD training samples

k	R	R^2	RMSE	slope	offset
1	0.5781	0.3342	24.48	0.6255	13.90
5	0.7652	0.5924	17.19	0.6378	19.45
10	0.7797	0.6079	16.48	0.6785	17.67
20	0.7643	0.5912	17.60	0.6396	20.04
30	0.7623	0.5818	19.24	0.6152	21.88
40	0.7487	0.5556	20.02	0.3346	37.18

the most similar neighbours considered are depending on the k value, which is an integer chosen by the user. Finally, the pixel is classified by the majority of vote of its neighbours regarding the classification process. However, regarding the regression process the average is calculated instead of a majority vote including all considered neighbours (Hassanat et al., 2014; Thessler et al., 2008; Gjertsen, 2007; Thessler et al., 2008).

Usually the k value is chosen empirically depending on the target of the study. Another approach includes testing and evaluating several different k values on the data. High k values decrease the over-fitting problem by reducing the overall noise, so the classifier is less affected by a single change of data point. With k increasing to infinity every plots is considered as nearest neighbour. Accordingly, the classification output is represented by one class depending on the total majority. Nevertheless, if the k value is chosen to small the classifier is more sensitive to outliers and tends to over-fit the data (Hassanat et al., 2014; Gjertsen, 2007; Thessler et al., 2008).

However, the algorithm has some limitations such as the time complexity, which is linear to the number of training samples. The performance also depends on the k value, using high k values is more time consuming, because more features are considered. Another drawback is that the algorithm can only predict inside the range of trainings samples (Hassanat et al., 2014; Gjertsen, 2007).

The advantage of the kNN method is that it does not require learning, works well in low dimensions and it estimates all variables simultaneously (Hassanat et al., 2014; Gjertsen, 2007; Baffetta et al., 2012). "Another advantage is that the method better preserves the relationship or covariance structure between variables than methods where predictions are made separately for each variable" (Gjertsen, 2007).

The k - Nearest Neighbours algorithm can only be influenced through one parameter the k - value, which defines how many neighbours are taken in consideration. To estimate the impact of k value on the TCD prediction accuracy, different parameters based on literature recommendations are applied on the 100

Table 6.10.: K Nearest Neighbours parameter tests with FTY training samples

k- value	Overall Accuracy	Kappa
1	0.5538	0.2439
5	0.6043	0.2954
10	0.6127	0.3223
20	0.6528	0.3937
30	0.6378	0.3621
40	0.6226	0.3338

m spatial resolution mosaic. Within the training process with the knn classifier the training samples are composed of centred and reduced pixel values by computing the global mean and standard deviation for each band (ORFEO Toolbox, 2016a).

Although, (Hassanat et al., 2014; Gjertsen, 2007; Baffetta et al., 2012) achieves good classification result with low k values, the best forest type classification result in this case is achieved with a k-value of 20 (table 6.10). Significantly worse results are achieved only considering less than 5 neighbours. Even Hassanat et al. (2014) point out that at k = 1, the boundaries are over fitted and the accuracy reaches a minima. Further, the k value of 20 is used regarding all further forest type classification tests.

Results in table 6.9 show that the predicted tree cover density values are significantly and positively correlated with the reference samples. Setting the k value to 10 achieves the highest R^2 values and regression coefficients. As the k value decreases or increases the correlation get weaker.

6.3.3. Band Combinations Experiments

Furthermore, the impact of different band combinations on the classification accuracies are investigated with 100 m spatial resolution images and both classifiers. Results of the regression analysis in table 6.11 show that the association between the predicted products and the reference data are significant across all band combinations, suggesting that all bands are reliable measures to predict tree cover density with the random forest classifier. However, the prediction with the combination of all available Sentinel-2 bands with surface information has a significant stronger association with the tree cover density reference data using the random forest classifier than the other band combination. Both the determination and the regression coefficients show that the prediction result is significant associated with the reference data. All band combinations except for 4L and 4LD reached moderate good accuracy results ($R^2 > 0.50$). Further, the impact of adding a digital elevation model (10 m ground resolution) as additional

Table 6.11.: Random forest (t = 100, n = 5, s = 10, d = 5) band combination tests using TCD samples. Band combination abbreviations can be found in table 6.5

band combinations	R	R^2	RMSE	slope	offset
all	0.6553	0.8034	16.95	0.6028	21.2685
all + DEM	0.7534	0.5675	17.3148	0.6822	15.34
wb	0.7559	0.5713	17.2390	0.6739	15.70
wb + DEM	0.7490	0.5660	18.1439	0.6203	16.93
4L	0.7248	0.5253	18.1403	0.4759	22.05
4L + DEM	0.6608	0.4667	19.7612	0.3122	33.52
5L	0.6537	0.4375	24.9900	0.3223	7.58
5L + DEM	0.6043	0.3651	26.4300	0.2334	14.21
7L	0.7686	0.5907	16.8448	0.7332	14.05
7L + DEM	0.7620	0.5806	17.0512	0.7100	13.83

Table 6.12.: K Nearest Neighbours (k = 10) band combination tests using TCD samples. Band combination abbreviations can be found in table 6.5

band combinations	R	R^2	RMSE	slope	offset
all	0.7797	0.6079	16.4870	0.6785	17.67
all + DEM	0.7958	0.6333	15.9435	0.7249	15.02
wb	0.7766	0.6030	16.5890	0.6707	18.14
wb + DEM	0.7898	0.6238	16.1486	0.7184	15.18
4L	0.6871	0.4721	19.1297	0.5723	32.17
4L + DEM	0.7459	0.5564	17.5369	0.6729	22.49
5L	0.5637	0.3178	27.3900	0.3394	12.17
5L + DEM	0.7119	0.5068	23.2900	0.4187	9.700
7L	0.7786	0.6062	16.5233	0.7368	14.49
7L + DEM	0.8021	0.6461	14.7331	0.7785	11.43

Table 6.13.: Random forest (t = 300, n = 20, s = 5) band combination tests using FTY samples. Band combination abbreviations can be found in table 6.5

band combinations	Overall Accuracy	Kappa
all	0,7119	0,3827
all + DEM	0,6852	0,4328
wb	0,6852	0,4263
wb + DEM	0,6733	0,4172
3L	0,7028	0,4642
3L + DEM	0,6852	0,4263
6L	0.6528	0.3791
6L + DEM	0.6453	0.3146
7L	0,6943	0,4434
7L + DEM	0,6852	0,4328

band is investigated and the results obtained appear in table 6.11. Both R^2 values and regression coefficients show that including a digital elevation model does not improve the correlation between the predicted and the reference values using the random forest classifier. On the contrary the knn classifier shows improvement including the digital elevation model. It should be considered that the difference between the band combinations with digital elevation model and those without are not stable throughout all band combinations tested. Sometimes significant differences can be noticed while others only show slightly differences.

In table 6.13 and table 6.14 the band combination experiments with the both classifiers are presented regarding the forest type classification. The random forest classifier again works slightly better including all bands available with surface information. In contrast to the random forest classifier the knn classifier works slightly better with a band combination choosing only two band of each spectral region (3LD). Thus taking all bands result in a over representation of the NIR/red-edge region of the spectrum with 5 bands available. Nevertheless, all band combinations reached moderate good accuracy results (OA $>$ 0.60) regarding both classifiers. Conversely to the random forest classifier the overall accuracies obtained in each experiment are significant better with the additional band containing attitudinal information. Surprisingly, excluding the blue band which contains the most artefacts after the atmospheric correction does not achieved better result than taking all bands.

Table 6.14.: KNN (k = 20) band combination tests using FTY samples. Band combination abbreviations can be found in table 6.5

band combinations	Overall Accuracy	Kappa
all	0,6528	0,3937
all + DEM	0,6733	0,4125
wb	0,6528	0,3937
wb + DEM	0,6852	0,4328
3L	0,6384	0,3583
3L + DEM	0,6943	0,4562
6L	0.6128	0.3146
6L + DEM	0.6294	0.3791
7L	0,6384	0,3607
7L + DEM	0,6852	0,4434

7. Accuracy Assessment

7.1. Tree Cover Density

7.1.1. TCD product based on Forest samples

The accuracy assessment regarding the tree cover density product without forest mask is conducted in two levels. First, regression and correlation analysis are used to evaluate the correlation between reference data and map values. Ideally the regression coefficients slope and offset should not vary significant from 1 and 0, respectively. Second, the density values are aggregated creating ranged attributes for confusion error matrices and kappa statistics calculations to access the thematic accuracy and better understand classification error causes. Consequently, the tree cover density values are grouped in 20 % steps to create fewer classes for the confusion matrix computation.

The analysis of the scatter plots in figure 7.1 and figure 7.2 show a strong relationship between validation data and the classification result. However, there seems to be a general trend to systematically underestimate the actual tree cover density values. Figure 7.2 represents the classification result regarding the random forest classifier with the best result of all band combinations and shows that the map density values appear to underestimate actual tree cover density values. The high tree cover density values are underestimated, due to the aggregation effect within the training process, whereas the low tree cover density values are overestimated, considering the effect of mixing signatures due to different underlying land cover types. Figure 7.1 similar shows a strong relationship between validation and map data and an underestimation of the tree cover density values.

Table 7.1 represents the results of the regression analysis regarding four band combinations at 100 and 20 m resolution. The determination coefficient R^2 ranges from 0.74 to 0.80 over the different band combinations tested, meaning that the model accuracy is accurate. Furthermore, table 7.1 shows that regression and determination coefficients based on the classified 20 m product aggregated to 100 m do not vary markedly from the classified 100 m product. The differences between both products are significantly higher regarding the random forest

Table 7.1.: Comparison between 100 m and 20 m products regarding the TCD results with 95 validation samples.

	100 m product					20 m product				
band	R	R^2	RMSE	slope	offset	R	R^2	RMSE	slope	offset
knn classifier										
7LD	0.8021	0.6461	14.73	0.7785	11.43	0.8101	0.6563	15.44	0.6438	21.52
allD	0.7958	0.6333	15.94	0.7249	15.02	0.7904	0.6247	16.13	0.5973	22.02
wbD	0.7898	0.6238	16.15	0.7184	15.18	0.7900	0.6240	16.14	0.5975	21.98
4LD	0.7459	0.5564	17.54	0.6729	22.49	0.7467	0.5576	17.51	0.5379	33.13
random forest classifier										
all	0.6553	0.8034	16.95	0.6028	21.27	0.8185	0.6699	15.12	0.6166	20.26
7L	0.7686	0.5907	16.84	0.7332	14.05	0.8107	0.6573	15.41	0.6457	19.79
wb	0.7559	0.5713	17.24	0.6739	15.70	0.8137	0.6621	15.51	0.5904	22.54
4L	0.7248	0.5253	18.14	0.4759	22.05	0.7043	0.4960	18.69	0.3003	34.39

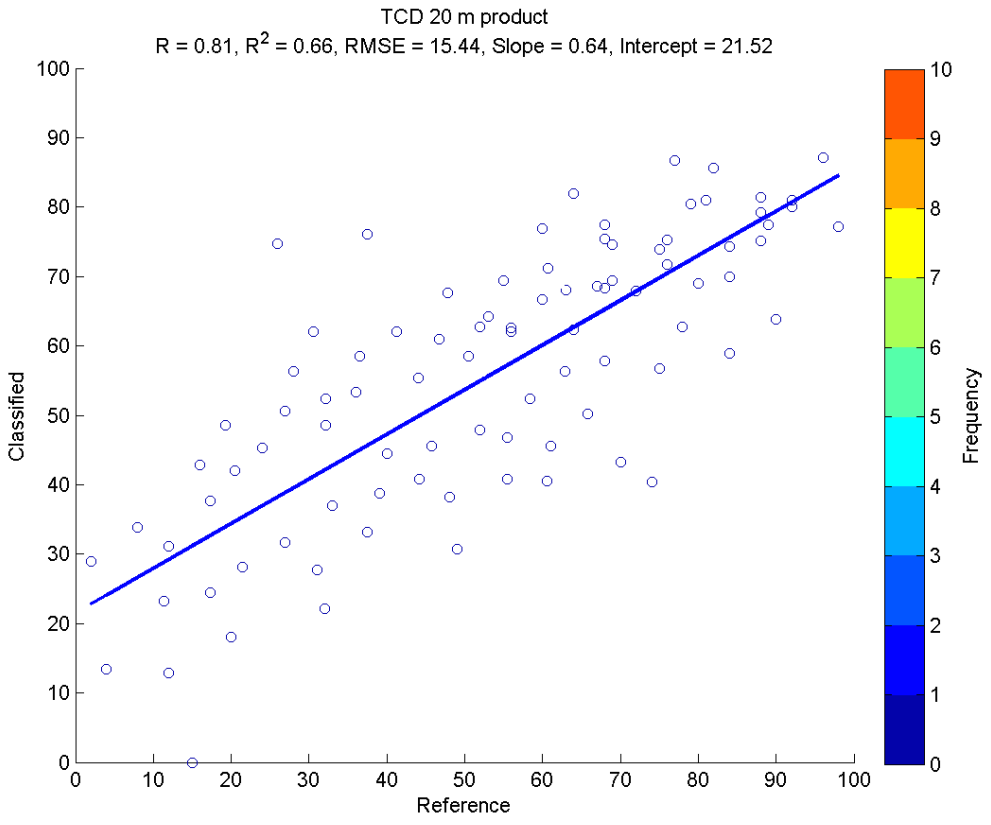


Figure 7.1.: Scatter plot presenting predicted results with the random forest classifier with 7LD band combination based on the 20 m resolution image data

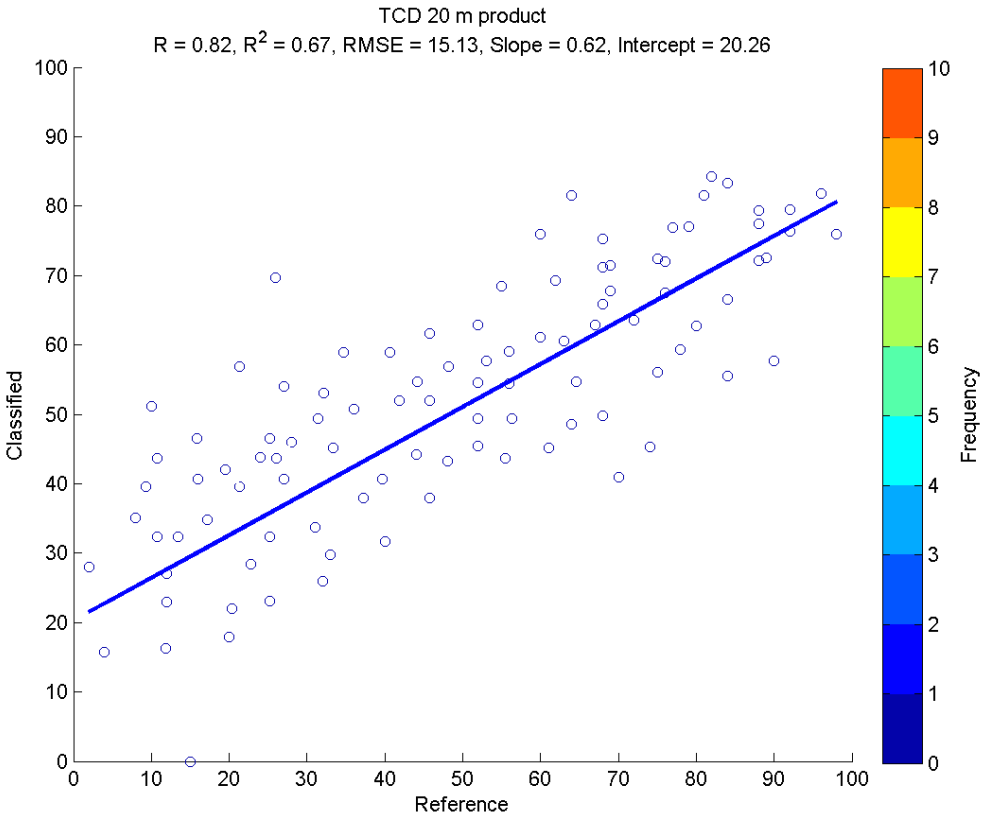


Figure 7.2.: Scatter plot presenting predicted results with the knn classifier and all bands based on the 20 m resolution image data

classier, except when including all bands. The regression analysis of the classified 20 m product also indicates that the band combination with all bands works better with the random forest classifier and the combination 7LD works best with the knn classifier achieving the highest R^2 values and regression coefficients. Overall, there is general trend that the tree cover density classification underestimates or overestimates the actual tree cover density values, despite the relationship between map and reference data is strong for example where R^2 is close to 0.60. This has been expected, since regression model prediction tends to overestimate the low values and under estimate the high values due to the average calculation within the training process (Horning et al., 2010). In the case of the tree cover density prediction with the random forest classifier more low tree cover density trainings samples are observed within the trainings data set than samples with high tree cover density values. Consequently, there is a possibility that after the random selection of subsets for each decision tree that the maximum range does not reach the higher density values.

As far as the thematic accuracy is concerned it is difficult to get a clear idea of the actual quality of the of the TCD product prediction without knowing where misclassification are present. Accordingly, the density values are aggregated in 10 % intervals and confusion matrices are calculated. Additionally confusion matrices with 20 % intervals are generated to get a general idea for a better understanding of the classification error causes. Figure 7.3 shows the distribution of the correctly classified elements in the confusion matrix as function of the tree cover density in 10 % intervals to estimate where the classification errors are higher. As the tree cover density increases, the number of right classified values increases, concerning the knn classifier. Similar but slightly better result are represented in figure 7.4 based on the random forest classifier, showing a wider range of right classified density values. In general it seems that the classification of lower density values is more complicated due to the underlying land cover and mixing spectral signatures than the higher density values. Furthermore, it is important to note that the low right classified frequency regarding the class with 100 % tree cover density can be explained by a low frequency of 100 % values in the classification result due to the general underestimation of TCD values. Table 7.2 representing the confusion matrix for the random forest classifier in combination with all bands, shows that the lower density values are more often underestimated than the higher values.

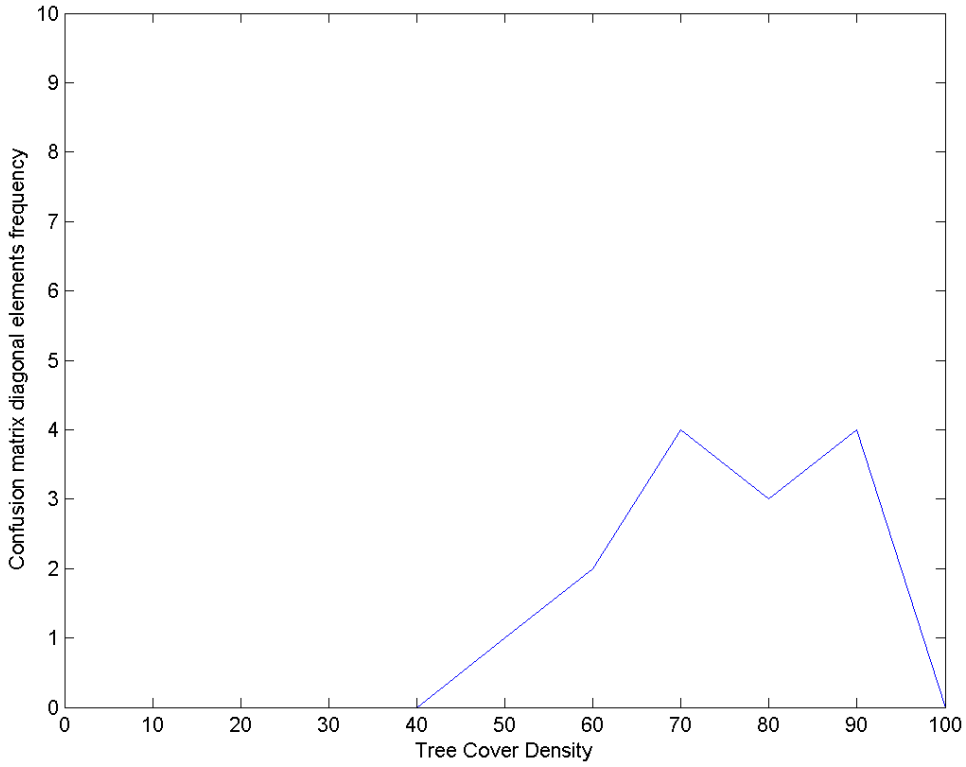


Figure 7.3.: Distribution of right classified values in function of TCD regarding the knn classified product based on the 7LD band combination.

Table 7.2.: TCD Error Matrix: Random forest classifier applied on all bands.

TCD	1-20%	21-40%	41-60%	61-80%	81-100%	Total	User Accuracy
1-20%	2	0	0	0	0	2	1
21-40%	18	4	0	0	0	22	0.7804
41-60%	10	4	9	8	2	33	0.7073
61-80%	0	5	4	15	9	33	0.7464
81-100%	0	0	0	1	4	5	0.9875
Total	30	13	13	24	15	95	
Prodcuer Accuracy	0.0666	0.3076	0.6923	0.6250	0.2666		0.35789

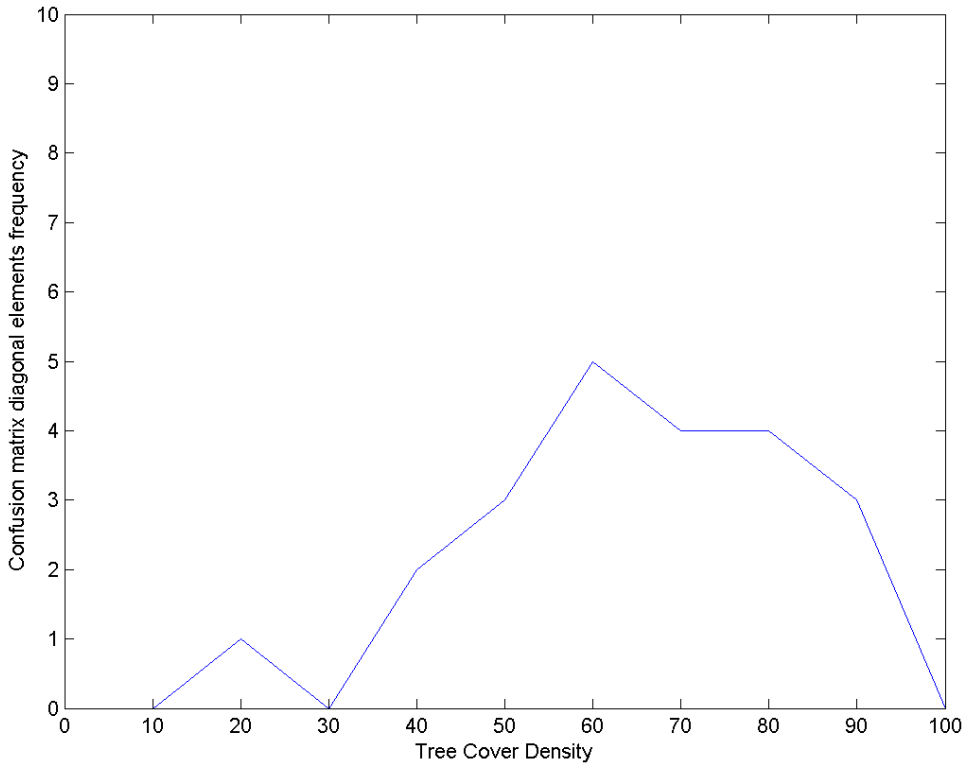


Figure 7.4.: Distribution of right classified values in function of TCD regarding the random forest classified product based on all bands.

Table 7.3.: TCD prediction results comparison before and after calibration.

	before calibration					after calibration				
band	R	R^2	RMSE	slope	offset	R	R^2	RMSE	slope	offset
knn classifier										
7LD	0.757	0.573	21.60	0.407	10.56	0.757	0.573	21.60	0.789	6.01
allD	0.771	0.592	21.36	0.465	10.42	0.771	0.592	21.36	0.785	5.05
wbD	0.758	0.574	21.58	0.453	10.35	0.758	0.574	21.58	0.779	5.86
4LD	0.753	0.568	21.75	0.443	10.33	0.753	0.568	21.75	0.777	5.79
random forest classifier										
all	0.821	0.675	18.86	0.487	9.95	0.821	0.675	18.86	0.800	5.51
7L	0.789	0.622	20.33	0.499	9.84	0.789	0.622	20.33	0.754	6.09
wb	0.788	0.621	20.37	0.520	10.08	0.788	0.621	20.37	0.760	6.41
4L	0.779	0.607	20.74	0.506	10.01	0.779	0.607	20.74	0.772	6.03

7.1.2. TCD product based on Forest and Non-Forest samples

To compare the tree cover density prediction results with the Copernicus HRL Layer also Non-Forest samples need to be included into the classification process. Consequently, a training set including all available samples from the interpretation process is used within the training process. The accuracy assessment regarding tree cover density product without forest mask is also conducted in two levels. First, regression and correlation analysis to evaluate the correlation between reference and map values are used. Second, the density values are grouped creating binary attributes for confusion error matrices and kappa statistics calculations to access the thematic accuracy. Since the weights cannot be included into the scatter plot analysis the second validation procedure is needed. Two different thresholds are suggested by Sannier et al. (2015) and applied forming two strata, with each containing forest or non-forest pixels. The first threshold groups all values from 0 to 9 and from 10 to 100, corresponding to the FAO forest definition. The second threshold groups values from 0 to 29 and from 30 to 100 corresponding to the evaluation approach suggested by Sannier et al. (2015).

Detailed regression coefficients like slope and intercept obtained from regression analysis without forest mask using all available trainings samples appear in table 7.3. The regression coefficients do not vary considerably with the different band combinations tested. Nevertheless, the best results with both classifiers are obtained with the band combination including all available bands containing surface information (excluding B1 (aerosol), B9 (water vapour) and B10 (cirrus)). Furthermore, the regression slopes are consistently below 0.6, where 1 would be the ideal slope value. Further, the underestimation of the actual tree cover density values appear higher when the Non-Forest class is included in the classification process. Although the relationship between classification product and

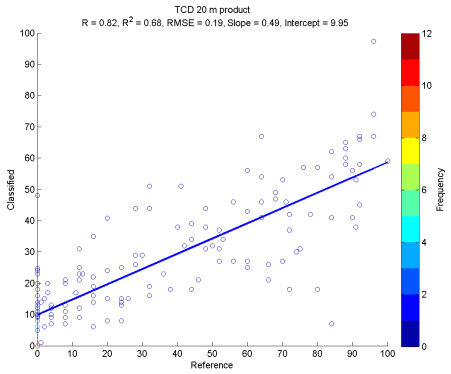
the reference data seems similar strong compared to the classification approach with the Corine Land Cover forest mask, with an R^2 value close to 0.6. Such a strong underestimation has not been expected, but can be explained by the huge amount of non-tree training samples in relation to the tree training samples.

Figure 7.5a shows the statistics based on the best random forest classification result and figure 7.5b based on the best knn classification result. Both demonstrate in detail how the map density values underestimate actual TCD values and showing a large concentrations of values close to 0 due to the over representation of the No-Tree class in the training samples. Under these circumstances a calibration of the tree cover density products is considered. Therefore, the classification product is calibrated using fitting parameters creating classification products covering the full TCD range (0 - 100 %). Accordingly, the training samples are plotted against the 100 m classification result to obtain the regression coefficients slope and offset (scatter plots and fitting parameters for each classification result are provided in annex C). Afterwards, following formula is applied to the corresponding 20 m resolution product, where *Acal* stand for after calibration and *BCal* stands for before calibration:

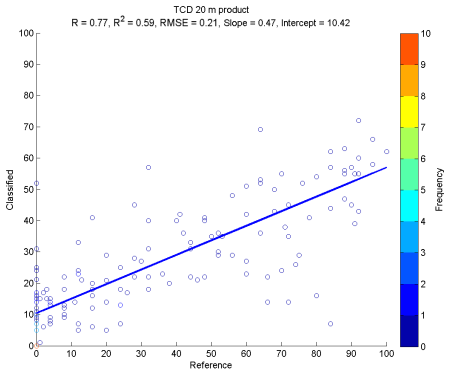
$$Classification(ACal) = \frac{Classification(BCal) - offset}{slope} \tag{7.1}$$

The same approach is applied to all different tree cover density classification with 20 m resolution with the corresponding fitting parameters. The corresponding scatter plots and values for the slope and offset parameter are provided in the annex part C.

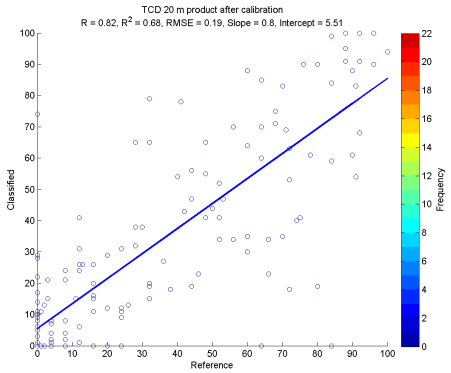
The results represented in table 7.3 show that the regression coefficients slope and offset are generally relatively lower regarding the classifications without calibration than after the calibration process, whereas the coefficient R, R^2 and RMSE do not vary at all. This can be explained by the formula 7.1 applied on the classification product, which works as a system displacement, only transforming the points on the diagonal without changing the correlation between the data. However, it is important to note that this calibration approach changes the classification result and the validation is not independent any more.



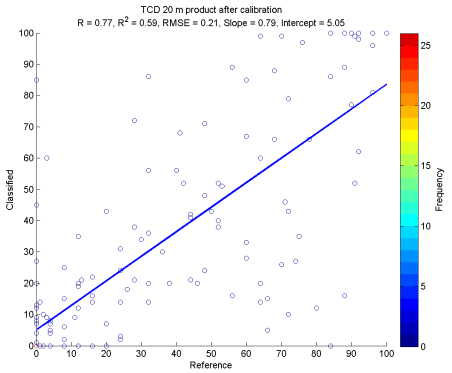
(a) Scatter plot before calibration with random forest classifier and all bands.



(b) Scatter plot before calibration with knn classifier and all bands.



(c) Scatter plot after calibration with random forest classifier and all bands.



(d) Scatter plot after calibration with knn classifier and all bands.

Figure 7.5.: Scatter plots of TCD prediction results before and after calibration.

As far as the thematic accuracy is concerned it is difficult to get a a clear idea of the actual quality of the of the TCD product prediction without the weights. Consequently, the density values are regrouped according to the thresholds $>= 10\%$ and $>= 30\%$ in non-tree and tree classes and confusion matrices are calculated. Following formula is adopted with modified inputs from Gallaun et al. (2015) and used to calculate the weighted confusion matrices where N is the number of validation samples, p the probability and w is the estimation weight, i indicates the row and j the column of the matrix:

$$\widehat{p_{ij}} = (\frac{1}{N}) * \sum_{x \epsilon (i.j)} \frac{1}{w} \tag{7.2}$$

To calculate the proportions of area for each cell of the error matrix (p_{ij}) formula 7.2 is applied. The overall accuracy (OA) 7.3 and the error of omission (EO) 7.5

and commission (EC) 7.4 are calculated as suggested by Gallaun et al. (2015) and listed below:

$$OA = \sum_{j=1}^{j} \widehat{p_{jj}} \tag{7.3}$$

$$\widehat{EC_j} = 1 - \frac{\widehat{p_{jj}}}{\sum_{i=1}^{j} \widehat{p_{ij}}} \tag{7.4}$$

$$\widehat{EO_i} = 1 - \frac{\widehat{p_{ii}}}{\sum_{j=1}^{i} \widehat{p_{ij}}} \tag{7.5}$$

Table 7.1 represents the results of the overall accuracies. From the statistics it can be depicted that calculating the error matrix with cumulated inclusion probabilities achieves significant higher results than not considering the weights. Furthermore, there seems to be only slight differences between the different band combinations, indicating that the classifiers are not that sensible to the band combinations. Regarding the not weighted overall accuracies there seem to be no difference between the different tree cover density thresholds. On the other hand regarding the weighted overall accuracies the differences between the tree cover density thresholds are significant higher. However, one problem with this approach is that it fails to take the limited usage of the validation points into account. It is important to note that the weights are calculated based on different European strata including over 12.000 points spread across the whole European country. Sannier et al. (2015) also mentions that the weights need to be recalculated if sample points are removed. The impact of removing only a small number of points might be minimal considering that over 12.000 sample points are interpreted.

Difficulties arise, however, when an attempt is made to compare the Copernicus HRL 2012 Tree Cover Density product with the current classification, thus different strata are evaluated within the Copernicus product. To compare both products the stratification weights within the Copernicus HRL validation points are used within the confusion matrix calculation. A serious weakness regarding this approach is that the subset used within the current tree cover density evaluation comprises only parts of the evaluation strata listed in table 7.5. Data from this table can be compared with caution with the data in table 7.4 regarding the weighted results which show the overall accuracies for the current tree cover density classification are equal or slightly better. There are no significant differences between the HRL Layer and the current classification results, considering the classification error based on the visual interpretation error. Furthermore, the

Table 7.4.: TCD overall accuracy - no-tree / tree product without forest mask.

	number of samples = 152		not weighted		weighted	
CL	parameter	BC	tcd <= 10	tcd <=30	tcd <=10	tcd <=30
knn	k = 10	allD	0.82	0.82	0.90	0.84
knn	k = 10	7LD	0.84	0.82	0.90	0.83
knn	k = 10	wbD	0.82	0.82	0.90	0.83
knn	k = 10	4LD	0.83	0.81	0.90	0.82
rf	d = 5, n = 5 s = 3, t = 100	all	0.85	0.85	0.92	0.84
rf	d = 5, n = 5 s = 3, t = 100	7L	0.84	0.84	0.93	0.85
rf	d = 5, n = 5 s = 3, t = 100	wb	0.86	0.86	0.93	0.85
rf	d = 5, n = 5 s = 3, t = 100	4L	0.86	0.84	0.93	0.84

Table 7.5.: Comparison with Copernicus HRL 2012 TCD product.

	Regression Analysis			Overall Accuracy	
Stratum	R^2	slope	offset	tcd <10 %	tcd <30 %
AT + CH + LI	0.7881	0.8822	0.2581	0.90	0.94
DE	0.7834	0.8589	0.5783	0.88	0.92
FR	0.6975	0.7864	0.2325	0.83	0.90
HU	0.7601	0.6861	0.6408	0.87	0.93
IT	0.6758	0.8789	4.2312	0.85	0.89

data from this table can be compared with the data in table 7.3 which shows that R^2 values for the current classification are higher than for the Copernicus product but on the other hand the slope values are higher for the Copernicus HRL product, which indicates that the high tree cover density values in the current classification are still underestimated (also see figure 7.6).

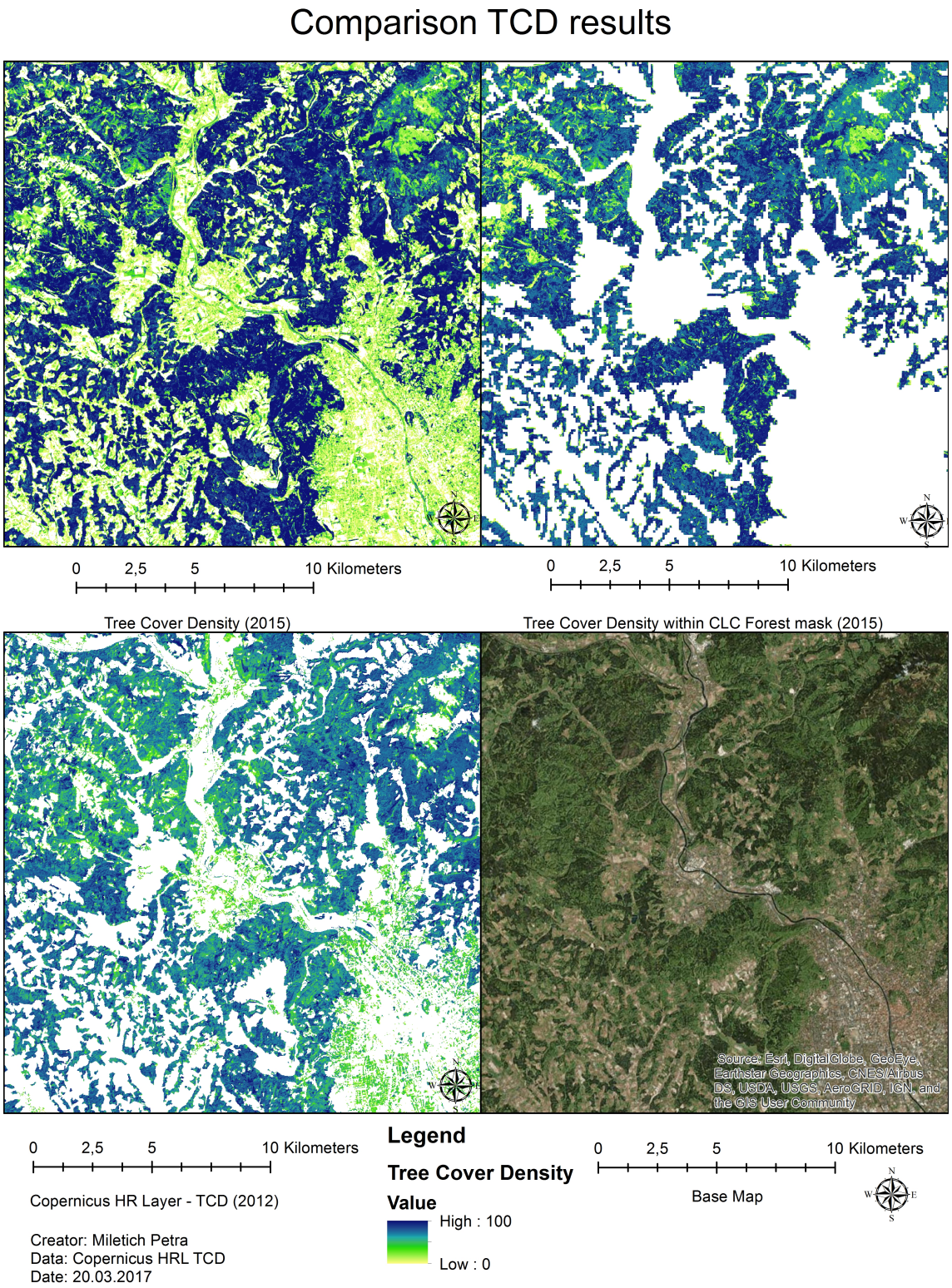


Figure 7.6.: Mapping product comparison regarding the tree cover density.

Table 7.6.: FTY Overall Accuracies 20 m product comparison between weighted and not weighted overall accuracies

				100 m product	20 m product	
classifier	parameter	bands	training samples	OA not weighted	OA not weighted	OA weighted
knn	k = 20	3LD	all	0.6943	0.7308	0.7739
knn		7LD	all	0.6852	0.7115	0.8062
knn		wbD	all	0.6852	0,7308	0.8128
knn		allD	all	0.6733	0.7308	0.8309
knn	k = 20	3LD	full	0.6554	0.6750	0.7464
knn		3LD	tcd	0.7054	0.7340	0.7937
knn		3LD	corine	0.7143	0.7361	0.8418
classifier	parameter	bands	training samples	OA not weighted	OA not weighted	OA weighted
rf	d = 5	all	all	0.7119	0.7308	0.8044
rf	$n = \sqrt{N}$	3L	all	0.7028	0.7308	0.7727
rf	s = 3	7L	all	0.6943	0.7308	0.7760
rf	t = 100	wb	all	0.6852	0.7404	0.8163
rf	d = 5, s = 3	all	full	0.6876	0.7250	0.7790
rf	$n = \sqrt{N}$	all	tcd	0.7241	0.7447	0.7790
rf	t = 100	all	corine	0.7389	0.7500	0.8967

7.2. Forest Type

The accuracies of forest type classification are assessed on the basis of confusion error matrices and error matrices based on inclusion probabilities to quantify the agreement between target pixel predictions and observations from the reference plots. Regarding the 20 m classification product four different trainings data sets are tested and evaluated to take into account input data limitations. Furthermore, four band combinations are tested with two spatial resolutions to estimate the influence of the aggregation effect.

In table 7.6, the results obtained with both classifiers and different trainings data sets are demonstrated. The highest overall accuracies are obtained with the random forest classifier using all available MSI sensor bands. The accuracy assessment is based on the classification product with 20 m ground resolution, which is aggregated to 100 by 100 m pixels using the majority vote. The classification results based on the 20 m mosaic performed slightly better than the results based on the 100 m product comparing the overall accuracies calculated without weights. Due to the aggregation effect the pixel show less outliers and variations within forest areas. The results in table 7.6 show that the overall performance of both classifiers does not show evident improvement compared to each other. In some cases the random forest algorithm works slightly better achieving higher

Table 7.7.: Best FTY results regarding both classifiers based on the 20 m product. Comparison between user and producer accuracies.

CL	BC	TS	N	OA	P-B	P-C	P-M	U-B	U-C	U-M
rf	all	all	104	0,73	0.87	0,78	0,33	0,76	0,81	0,47
rf	all	tcd	94	0,75	0,83	0,90	0,44	0,72	0,87	0,61
rf	all	full	40	0,73	1	0,88	0,36	0,60	0,83	0,71
rf	all	corine	72	0,74	0,91	0,77	0,50	0,56	0,89	0,91
knn	3LD	all	104	0,73	0,94	0.76	0,24	0,72	0,82	0,50
knn	3LD	tcd	94	0,73	0,88	0,80	0,42	0,76	0,83	0,53
knn	3LD	full	40	0,68	0.78	0.88	0,36	0,50	0,83	0.63
knn	3LD	corine	72	0,75	0.91	0,81	0,50	0,70	0,78	0,77

P = producer, U = user, B = broadleaved, C = coniferous, M = mixed.

overall accuracies.

Considering the four band combination experiments, the classification based on all bands shows slightly better results in both cases. In comparison to the 100 m resolution results the knn classifier performs better with the 3LD band combination, indicating that the difference is not significant and both band combinations are useful. Other band combinations perform as well except for the band combination with all bands without the blue band, implying that the blue band contains valuable information regarding the forest type discrimination.

The random forest algorithm provides higher producer accuracies regarding the broadleaved class than the knn classifier. The 'tcd' and 'corine' trainings data sets improve the overall accuracies of the both classifiers. Regarding the random forest producer accuracies both improved the broadleaved and the coniferous forest class. In combination with the knn classifier both show lower producer accuracies for broadleaved forest but higher results regarding coniferous and mixed forests. The classification results produced with the 'corine' trainings data set achieved the highest overall and producer accuracies with both classifiers with one exception. The k nearest neighbour producer accuracies regarding coniferous and broadleaved forests is significant lower than the classified product produced with 'all' trainings sample set. The highest random forest user accuracies for broadleaved and coniferous forest are achieved with 'full' trainings samples, whereas regarding mixed forests with the 'corine' data set. Broadleaved and coniferous reached the highest producer accuracy of 95 % and 88 % with the different trainings data set and the random forest classifier. In general for broadleaved forests accuracies higher than 78 % regarding both classifiers and all trainings data sources are achieved. The classification of coniferous forests also reached good accuracies over 76 %. The mixed forest class shows the lowest accuracies, with a maximum producer's accuracy of only 33 % considering the random forest

Table 7.8.: Confusion matrix: Random forest classifier trained with the 'corine' data set and applied on all bands.

	Broadleaved	Coniferous	Mixed	Total	User Accuracy
Broadleaved	19	3	5	27	0.704
Coniferous	2	25	5	32	0.781
Mixed	0	3	10	13	0.769
Total	21	31	20	72	
Producer Accuracy	0.905	0.806	0.5		0.750

Table 7.9.: Confusion matrix: knn classifier trained with the 'corine' data set and applied on the 3LD band mosaic.

	Broadleaved	Coniferous	Mixed	Total	User Accuracy
Broadleaved	19	6	9	34	0.559
Coniferous	2	24	1	27	0.889
Mixed	0	1	10	11	0.909
Total	21	31	20	72	
Producer Accuracy	0.905	0.774	0.5		0.736

algorithm and 24 % regarding the knn algorithm.

Difficulties arise, however, when an attempt is made to compare the analysis of the results for FTY classification with different trainings data set, due to the different validation set sizes represented by the column N in table 7.7. It is worth noting that the validation samples of the 'full' training data set with 40 samples are low.

Regarding the confusion matrices in figure 7.8 and 7.9, both show that the main issue of the confusion between the forest types seems to be caused by the mixed class which is particular difficult to differentiate from the others, even with the additional red edge bands providing information on the vegetation status. The confusion matrices for each classification product differentiating between the band combinations and trainings data sets are provided in annex part A.

To estimate the thematic accuracy the unequal sampling intensity appearing through the stratified systematic sampling approach has to be considered by applying a weight factor to each sample unit. The results obtained with the estimation weights are presented in table 7.6 and show higher accuracies regarding all experiments if weights are considered within the confusion matrix computation process, regarding both classifiers random forest and knn.

The quantitative classification accuracy assessment has shown the same results with visual observation shown in figure 7.7. In the borders of forest regions forest is always classified as mixed due to small proportions of other vegetation types.

It is noticeable that in areas, where mixed forest is dominated by one forest type (broadleaved or coniferous) which is homogeneously distributed over the trainings sample a confusion between mixed and broadleaved/coniferous forest is caused. Furthermore, in region with low tree cover densities the forest types are difficult to estimate due to the effect of the underlying land cover and coniferous forest pixels are misclassified as broadleaved forest.

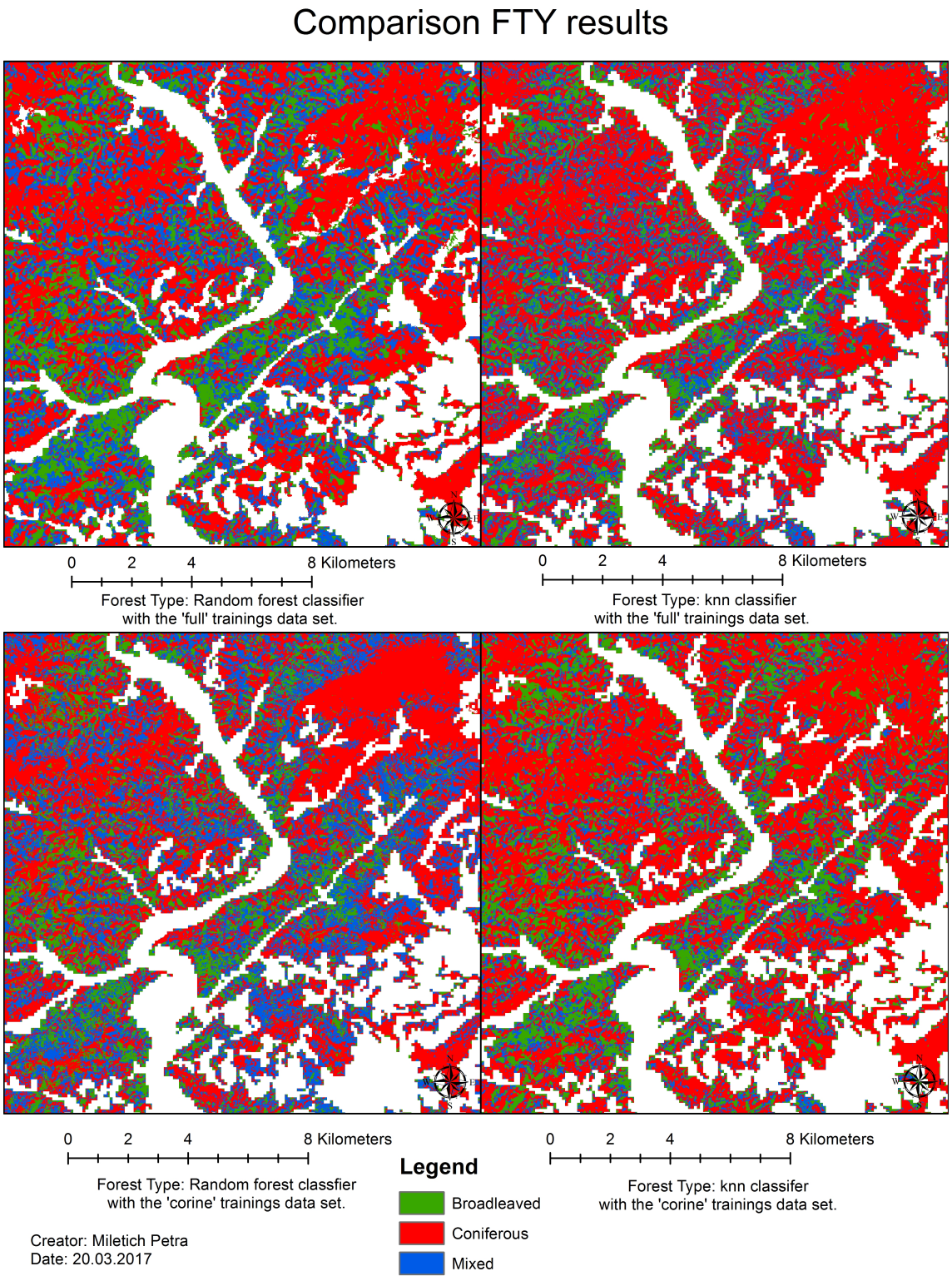


Figure 7.7.: Mapping product comparison regarding forest types.

8. Conclusion

This thesis introduces a methodological framework regarding forest type and tree cover density classification methods using SENTINEL-2A data and already available trainings data collected through visual image interpretation carried out in the frame of the EEAVAL project. The HR Layer validation plots are reused as training samples within the classification process. The results confirm that the those plots have the potential for mapping tree cover density values in large areas with an acceptable accuracy, while the produced forest type classifications reach moderate accuracies.

The proposed method avoids the time-consuming efforts to collect training areas within each satellite image reaching a higher automation degree. The training plots are available at European scale. Regarding the tree cover density classification no further preparation has been needed. In contrary, the heterogeneous forest type training samples appear to be a major problem regarding the training process. As mentioned above the trainings data initially has been collected for product validation reasons. Consequently, many plots are sampled within areas difficult to map for example forest borders. Accordingly, mixed land cover types, influencing plot signatures, bias the training process. To overcome these problem only trainings samples within forest areas are considered, extracting the training samples as well as the classification results with a forest mask derived from the Corine Land Cover 2012 product.

To estimate tree cover density and forest type classes over large areas a subset reaching from east France to west Hungary is chosen depending on the availability of the SENTINEL-2A scenes. With only SENTINEL-2a in the orbit at this point in time it has been difficult to acquire enough cloud free images, so only monotemporal images could be used. The usage of time series could increase the accuracy significantly, fully exploiting the potential of the SENTINEL-2a sensor data.

The SENTINEL-2A image preprocessing has been accomplished with the new introduced SENTINEL processor Sen2Cor. The results of the Sen2Cor atmospheric correction approach is compared with two other sensors, (MODIS on board the AQUA/TERRA satellite and OLI on board the LANDSAT 8 satellite) to check the consistency. The results regarding the mean spectral value comparison between

the three sensors indicate that the Sen2Cor processor works consistent, except with granules without dark dense vegetation pixels or with haze consistency.

As mentioned above clouds are a major problem regarding mapping approaches in general based on satellite images. Accordingly, this study provides an illustration of fully automatic generated cloud masks produced by the Sen2Cor processor. Several different cloud masks available within different product levels are presented and visually inspected. The illustration shows that the clouds borders are under-represented in all products available and partially clouds are still missing. Consequently, the different cloud masks have been combined and manually reviewed to exploit the full potential of each mask.

The tree cover density classification with all available training samples and without mask restrictions shows a general trend to underestimate the actual tree cover density values resulting in low accuracies. To adjust the classification product the training samples are plotted against the predicted values to estimate the impact of the underestimation, caused by large over-representations of the lower tree cover density value and aggregation effects within the training process. The regression coefficients are applied to the classified product to stretch the tree cover density values reaching the acquired range (0 - 100). This method of analysis has a number of limitations since the classification results are no longer independent, thus the impact of the coefficient need to be considered within the validation process. Perhaps the most serious disadvantage of this method is that the classification accuracies are no longer consistently comparable with other products. Nevertheless, the fitting approach performed in order to achieve better user accuracies.

As an alternative the training data and the classification results are extracted with a forest mask only considering sample points within forest areas. On one hand this approach results in less training and validation plots, but on the other hand the overrepresented non-forest class is removed. The results produced with this approach suggest that the impact of the non-forest class on the underestimation is significant high. Results regarding both approaches show nearly the same R^2 values but stronger differences in the regression coefficients slope and offset. Using the forest mask approach still results in an underestimated prediction of tree cover density values, however, the impact is lower.

The forest type classification shows significant lower classification accuracies compared to the results of other studies using medium resolution satellite images (Dorren et al., 2003; Xiao et al., 2002; X. Zhu and Liu, 2014; Dalponte et al., 2013; Laurin et al., 2016). The random forest classifier using all available Sentinel-bands with surface information (excluding B1 (aerosols), B9 (water vapour) and B10 (cirrus)) achieved the highest overall accuracy in combination with the 'corine' training data set considering the weights. In contrast to the not

weighted confusion matrices, weighted error matrices are calculated based on the inclusion probabilities. Furthermore, the tested random forest classifier attained slightly higher classification accuracies, than the knn classifier considering both, weighted and not weighted overall accuracies. The aggregation effect shows a slight improvement regarding the forest type classification results. It appears that after a visual product inspection the random forest classifier using all available bands, identifies fewer coniferous forest pixels than the knn classifier. On the contrary, the knn classification result shows some more 'salt and pepper' effects than the random forest classification. Considering the results of other studies (Dorren et al., 2003; Xiao et al., 2002; X. Zhu and Liu, 2014; Dalponte et al., 2013; Laurin et al., 2016) these accuracies might not be satisfactory for further applications, but this approach allows large area forest mapping with moderate accuracies.

Furthermore, no significant difference is found between the several band combination experiments including FTY and TCD results with on exception. Considering band combinations including only three promising bands significant lower accuracies are achieved. Moreover, the findings suggest that the knn classifier works better including a digital elevation model, whereas the random forest algorithm combined with topographic information may not necessarily helps to largely improve the tree cover density classification accuracy. Perhaps an improvement can be achieved by including topographic variables derived from the digital elevation model such as slope information. The potential of the new introduced red-edge band has been evaluated with signature and regression analysis, indicating that some red-edge bands include similar information regarding both, forest type and tree cover density plots. Nevertheless, using all band achieved better results, than excluding one or more of the redundant red-edge bands. It has been expected that the red-edge band help to differentiate between tree cover areas and agricultural land like arable land with annual crops, pastures and herbaceous vegetation associations. After the visual inspection of the product misclassification of arable lands are observed, depending on the underlying SENTINEL-2A scene. In that regard it needs to be considered that the classification accuracy is related to the association between spectral signatures and particular vegetation types, and consequently depends on the atmospheric correction accuracy. Since Sen2Cor fails to process granules without dense dark vegetation pixels those granules distort the training process and lead to misclassification.

The most important limitation lies in the fact that the accuracy of the classification based on point sampling is closely linked to the quality of the sampling design. Since the sampling points are created for validation purposes a random sampling approach has been used with additional points in critical areas. Nevertheless, regarding forest mapping aims only using homogeneous training plots could achieve better results. The actual tree cover density R^2 values might be higher

than the calculated values due to the over representation of problematic areas in the sampling approach. Further work needs to be done finding approaches to calculate actual R^2 with uneven inclusion probabilities.

In conclusion it can be stated the proposed methodological framework presents acceptable efficiency for the monitoring of forest attributes over large areas with reasonable accuracy. The framework is generic and can be applied on other scenes with different acquisition dates as long as survey points are available. The forest mapping methods applied in the 20 m resolution SENTINEL-2a image data demonstrate the relevance of SENTINEL-2 data providing information on both forest type and tree cover density. Nevertheless, there is still a need for developing classification processes with increased automation degree, with the potential in combining remotely sensed data correction and classification methods within one software package compatible with SENTINEL-2A data formats.

Finally, a number of important limitations need to be considered. First, the training samples are mainly gathered for validation purposes, thus more samples are allocated within problematic areas and less samples represent homogeneous forest areas. Furthermore, the atmospheric correction with Sen2Cor processor (version 2.2.1) has some difficulties with granules without dense dark vegetation pixels and haze. Nevertheless, it should be considered that Sen2Cor is still in his early tuning phases, and next versions will bring enhancements. Finally, the validation samples are not fully independent, since they are gathered within the same framework as the training samples and therefore have to deal with the same interpretation error.

Further work needs to be done to investigate whether there are still possibilities to improve the performance of forest type and tree cover density classification by including topographic variables derived from the digital elevation model such as slope information. Furthermore, time series of Sentinel-2 data should be considered in order to increase the classification accuracy. Multi-temporal images provide phenological information of forest, which can improve the discrimination of forest types sharing similar spectral characteristics (X. Zhu and Liu, 2014). Moreover, it needs to be investigated if the accuracies might be improved by a sub sampling approaches to reduce the number of samples within critical areas.

Bibliography

Abburu, Sunitha and Suresh Babu Golla (2015). "Satellite image classification methods and techniques: A review." In: *International journal of computer applications* 119.8 (cit. on p. 12).

Ali, Jehad et al. (2012). "Random forests and decision trees." In: *International Journal of Computer Science Issues (IJCSI)* 9.5 (cit. on pp. 71, 72).

Appiah, Divine Odame et al. (2015). "Application of geo-information techniques in land use and land cover change analysis in a Peri-Urban District of Ghana." In: *ISPRS International Journal of Geo-Information* 4.3, pp. 1265–1289 (cit. on p. 67).

Baffetta, Federica, Piermaria Corona, and Lorenzo Fattorini (2012). "A matching procedure to improve k-NN estimation of forest attribute maps." In: *Forest ecology and management* 272, pp. 35–50 (cit. on pp. 3, 5, 15, 74–76).

Belgiu, Mariana and Lucian Drăguţ (2016). "Random forest in remote sensing: A review of applications and future directions." In: *ISPRS Journal of Photogrammetry and Remote Sensing* 114, pp. 24–31 (cit. on pp. 71–73).

Blanzieri, Enrico and Farid Melgani (2008). "Nearest neighbor classification of remote sensing images with the maximal margin principle." In: *IEEE Transactions on Geoscience and Remote Sensing* 46.6, pp. 1804–1811 (cit. on p. 74).

Bruzzone, Lorenzo and Begüm Demir (2014). "A review of modern approaches to classification of remote sensing data." In: *Land Use and Land Cover Mapping in Europe*. Springer, pp. 127–143 (cit. on pp. 6, 21).

Büttner, G.and Dufourmont H et al. (2015). "GIO land (GMES/Copernicus initial operations land) High Resolution Layers (HRLs) – summary of product specifications." In: ed. by Tobias LANGANKE. URL: `https://cws-download.eea.europa.eu/pan-european/hrl/HRL_Summary_for_publication_v11.pdf` (cit. on pp. 21, 23).

Campbell, James B. (2002). *Introduction to remote sensing*. English. 3. London [u.a.]: Taylor & Francis. ISBN: 9780415282949;0415282942; (cit. on pp. 9–11, 14).

Colditz, René Roland (2015). "An evaluation of different training sample allocation schemes for discrete and continuous land cover classification using decision tree-based algorithms." In: *Remote Sensing* 7.8, pp. 9655–9681 (cit. on p. 72).

Copernicus Land Monitoring Service (2017). *CORINE Land Cover*. visited on 2017-02-30. URL: http://land.copernicus.eu/pan-european/corine-land-cover (cit. on p. 27).

Dalponte, Michele et al. (2013). "Tree species classification in boreal forests with hyperspectral data." In: *IEEE Transactions on Geoscience and Remote Sensing* 51.5, pp. 2632–2645 (cit. on pp. 72, 98, 99).

Delegido, Jesús et al. (2011). "Evaluation of sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content." In: *Sensors* 11.7, pp. 7063–7081 (cit. on pp. 18, 20).

Dorren, Luuk KA, Bernhard Maier, and Arie C Seijmonsbergen (2003). "Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification." In: *Forest Ecology and Management* 183.1, pp. 31–46 (cit. on pp. 15–17, 98, 99).

Al-Doski, Jwan, Shattri B Mansorl, and Helmi Zulhaidi Mohd Shafri (2013). "Image Classification in Remote Sensing." In: *Department of Civil Engineering, Faculty of Engineering, University Putra, Malaysia* (cit. on pp. 9, 10, 12).

Drusch, M et al. (2012). "Sentinel-2: ESA's optical high-resolution mission for GMES operational services." In: *Remote Sensing of Environment* 120, pp. 25–36 (cit. on pp. 18–20).

ESA Earth Online (2016). *Sentinel 2 Mission News*. visited on 2016-09-01. URL: https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/sentinel-2 (cit. on p. 18).

ESA, European Space Agency (2015). *The story of Sentinel-2, ESA Bulletin 161*. Tech. rep. ESA (cit. on p. 19).

ESA, European Space Agency (2015). *Landmonitoring with Sentinel*. visited on 2016-06-14. URL: https://earth.esa.int/web/sentinel/user-guides/sentinel-1-sar/applications/land-monitoring (cit. on p. 7).

ESA, European Space Agency (2016). *GlobCover*. visited on 2016-06-14. URL: http://due.esrin.esa.int/page_globcover.php (cit. on p. 4).

European Environment Agency (EEA) (2017). *Digital Elevation Model over Europe (EU-DEM)*. visited on 2017-01-30. URL: http://www.eea.europa.eu/data-and-maps/data/eu-dem (cit. on pp. 25, 27).

EUROSTAT (2015). *LUCAS - Land use and land cover survey*. visited on 2016-07-23. URL: http://ec.europa.eu/eurostat/statistics-explained/index.php/LUCAS_-_Land_use_and_land_cover_survey (cit. on p. 6).

FAO, Food and Agriculture Organisation of the United Nations (2015). "FAO's Global Forest Resources Assessment." In: (cit. on p. 24).

Ferreira, Matheus Pinheiro et al. (2016). "Mapping tree species in tropical seasonal semi-deciduous forests with hyperspectral and multispectral data." In: *Remote Sensing of Environment* 179, pp. 66–78 (cit. on p. 15).

Fletcher K. Editor ; European Space Agency, Editor, ed. (2012). *Sentinel-2 : ESA's optical high-resolution mission for GMES operational services*. European Space

Agency. URL: https://isulibrary.isunet.edu/opac/index.php?lvl=notice_display&id=8403 (cit. on pp. 18, 20, 37, 46).

Gallaun, Heinz et al. (2015). "Remote Sensing Based Two-Stage Sampling for Accuracy Assessment and Area Estimation of Land Cover Changes." In: *Remote Sensing* 7.9, pp. 11992–12008 (cit. on pp. 88, 89).

Gatti A., Bertolini A. (2013). *Sentinel-2 Products Specification Document*. Tech. rep. esa, Thales, Thales Alenia Space, acs (cit. on p. 18).

Geoland (2017). *Digital Elevation Model over Austria*. visited on 2017-01-30. URL: https://www.data.gv.at/katalog/dataset/d88a1246-9684-480b-a480-ff63286b35b7 (cit. on p. 27).

Gjertsen, Arnt Kristian (2007). "Accuracy of forest mapping based on Landsat TM data and a kNN-based method." In: *Remote Sensing of Environment* 110.4, pp. 420–430 (cit. on pp. 5, 16, 74–76).

Grinand, Clovis et al. (2013). "Estimating deforestation in tropical humid and dry forests in Madagascar from 2000 to 2010 using multi-date Landsat satellite images and the random forests classifier." In: *Remote Sensing of Environment* 139, pp. 68–80 (cit. on pp. 12, 71).

Hagolle, Olivier et al. (2010). "A multi-temporal method for cloud detection, applied to FORMOSAT-2, VENμS, LANDSAT and SENTINEL-2 images." In: *Remote Sensing of Environment* 114.8, pp. 1747–1755 (cit. on p. 44).

Hassanat, Ahmad Basheer et al. (2014). "Solving the problem of the K parameter in the kNN classifier using an ensemble learning approach." In: *International Journal of Computer Science and Information Security* 12.8 (cit. on pp. 12, 74–76).

Hirschmugl, Manuela et al. (2017). "Methods for Mapping Forest Disturbance and Degradation from Optical Earth Observation Data: a Review." In: *Current Forestry Reports* 3.1, pp. 32–45 (cit. on p. 7).

Horning, Ned et al. (2010). "Random Forests: An algorithm for image classification and generation of continuous fields data sets." In: *Proceedings of the International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences, Osaka, Japan*. Vol. 911 (cit. on pp. 71–73, 83).

Hyyppä, Juha et al. (2000). "Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes." In: *Forest Ecology and Management* 128.1, pp. 109–120 (cit. on pp. 3, 4).

Immitzer, Markus et al. (2016). "Verwendung von multispektralen Sentinel-2 Daten für die Baumartenklassifikation und Vergleich mit anderen Satellitensensoren." In: (cit. on p. 16).

Jensen, John R et al. (2009). "Image classification." In: *The SAGE handbook of remote sensing*, pp. 269–281 (cit. on pp. 9–13).

Jia, Kun et al. (2014). "Forest cover classification using Landsat ETM+ data and time series MODIS NDVI data." In: *International Journal of Applied Earth Observation and Geoinformation* 33, pp. 32–38. ISSN: 0303-2434. DOI: http://dx.doi.org/10.1016/j.jag.2014.04.015 (cit. on pp. 3, 5).

Jin, Suming et al. (2013). "Automated cloud and shadow detection and filling using two-date Landsat imagery in the USA." In: *International Journal of Remote Sensing* 34.5, pp. 1540–1560 (cit. on p. 49).

JRC, Joint Research Centre (2015). *Forest Cover Map 2000*. visited on 2016-06-14. URL: http://forest.jrc.ec.europa.eu/activities/forest-mapping/forest-cover-map-2000/ (cit. on p. 5).

Kasischke, Eric S. et al. (2004). "Remote sensing for natural resource management and environmental monitoring." English. In: ed. by Susan L. Ustin. Vol. 4. Hoboken, NJ: Wiley. Chap. Temperate and Boreal Forests. ISBN: 0471317934; 9780471317937; (cit. on pp. 9, 12, 13).

Langanke, Tobias et al. (2015). *GIO land (GMES/Copernicus initial operations land) High Resolution Layers (HRLs) – summary of product specifications*. Tech. rep. ESA (cit. on pp. 5, 7, 21, 24, 25).

Laurin, Gaia Vaglio et al. (2016). "Discrimination of tropical forest types, dominant species, and mapping of functional guilds by hyperspectral and simulated multispectral Sentinel-2 data." In: *Remote Sensing of Environment* 176, pp. 163–176 (cit. on pp. 3, 5–7, 14, 16, 98, 99).

Li, Miao et al. (2014). "A review of remote sensing image classification techniques: The role of spatio-contextual information." In: *European Journal of Remote Sensing* 47, pp. 389–411 (cit. on pp. 11, 12, 15).

Li, Teng et al. (2016). "On random hyper-class random forest for visual classification." In: *Neurocomputing* 172, pp. 281–289 (cit. on p. 71).

Liaw, Andy and Matthew Wiener (2002). "Classification and regression by randomForest." In: *R news* 2.3, pp. 18–22 (cit. on pp. 9, 12, 71, 73).

Lillesand, Thomas, Ralph W Kiefer, and Jonathan Chipman (2014). *Remote sensing and image interpretation*. John Wiley & Sons (cit. on pp. 9–12, 14).

Lindner, Marcus et al. (2010). "Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems." In: *Forest Ecology and Management* 259.4, pp. 698–709 (cit. on p. 28).

Louis, Jerome, Alexandre Charantonis, and Beatrice Berthelot (2010). "Coud Detection for Sentinel-2." In: *ESA Special Publication*. Vol. 686 (cit. on p. 46).

Main-Knorn, M et al. (2015). "Calibration and Validation Plan For The L2A Processor And Products Of The SENTINEL-2 MISSION." In: *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 40.7, p. 1249 (cit. on pp. 18–20, 34, 36, 37).

Malenovsky, Zbyněk et al. (2012). "Sentinels for science: Potential of Sentinel-1,-2, and-3 missions for scientific observations of ocean, cryosphere, and land." In: *Remote Sensing of Environment* 120, pp. 91–101 (cit. on pp. 18, 19).

Martimort, P. et al., eds. (2007). *Sentinel-2 Optical High Resolution Mission for GMES operational services*. Vol. 1. 54. ESA. URL: http://www.esa.int/esapub/bulletin/bulletin131/bul131b_martimort.pdf (cit. on pp. 18–20).

Masek, Jeffrey G et al. (2015). "The role of remote sensing in process-scaling studies of managed forest ecosystems." In: *Forest Ecology and Management* 355, pp. 109–123 (cit. on pp. 3, 6, 16).

Mather, Paul M and K Magaly (1999). *Computer Processing of Remotely Sensed Images—An Introduction* (cit. on pp. 9–12).

Mellor, Andrew et al. (2015). "Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin." In: *ISPRS Journal of Photogrammetry and Remote Sensing* 105, pp. 155–168 (cit. on p. 72).

Michalak, Roman (2014). "Forest monitoring in Europe and its importance to clean air policies and sustainable forest management." In: *Forests under pressure: Local responses to global issues*, p. 411 (cit. on pp. 3–5).

Mitscherlich, Gerhard (1978). *Wald, Wachstum und Umwelt. v. 1, Form und Wachstum von Baum und Bestand.* (Cit. on p. 28).

Moreno, Adam, Mathias Neumann, and Hubert Hasenauer (2016). "Optimal resolution for linking remotely sensed and forest inventory data in Europe." In: *Remote Sensing of Environment* 183, pp. 109–119 (cit. on pp. 3–6).

Motive Project (2016). *Temperate Oceanic Region*. visited on 2016-09-01. URL: http://motive-project.net/img/uplf/MOTIVE_factsheet_temperate%20oceanic%20region_final.pdf (cit. on p. 28).

Mountrakis, Giorgos, Jungho Im, and Caesar Ogole (2011). "Support vector machines in remote sensing: A review." In: *ISPRS Journal of Photogrammetry and Remote Sensing* 66.3, pp. 247–259. ISSN: 0924-2716. DOI: http://dx.doi.org/10.1016/j.isprsjprs.2010.11.001 (cit. on p. 9).

Müller-Wilm, Uwe (2015). "Sentinel-2 MSI-Level-2A Prototype Processor Installation and User Manual." In: *Telespazio VEGA Deutschland GmbH: Darmstadt, Germany* (cit. on pp. 34, 36, 37).

Muller-Wilm, Uwe et al. (2013). "Sentinel-2 Level 2A Prototype Processor: Architecture, Algorithms And First Results." In: *ESA Special Publication*. Vol. 722, p. 98 (cit. on pp. 20, 34–37, 46, 47).

Musaoglu, N and C Örmeci (2000). "Monitoring of forest change by using multitemporal satellite data." In: *EARSeL symposium on remote sensing in the 21st century*, pp. 41–45 (cit. on p. 6).

NASA Official (2016). *MODIS Specifications*. visited on 2016-11-17. URL: https://modis.gsfc.nasa.gov/about/specifications.php (cit. on p. 39).

Ohmann, Janet L, Matthew J Gregory, Heather M Roberts, et al. (2014). "Scale considerations for integrating forest inventory plot data and satellite image data for regional forest mapping." In: *Remote sensing of environment* 151, pp. 3–15 (cit. on pp. 5, 14, 15, 50).

ORFEO Toolbox (2016a). *K-Nearest Neighbours OpenCV 2.4.13.2 documentation*. visited on 2016-11-30. URL: http://docs.opencv.org/2.4/modules/ml/doc/k_nearest_neighbors.html (cit. on p. 76).

ORFEO Toolbox (2016b). *Random Trees OpenCV 2.4.13.2 documentation*. visited on 2016-11-30. URL: http://docs.opencv.org/2.4/modules/ml/doc/random_trees.html (cit. on p. 74).

Pflug, Bringfried et al. (2016). "Early Validation of Sentinel-2 L2A Processor and Products." In: *ESA Special Publication* (cit. on pp. 18, 35, 37, 38, 43).

Pippuri, Inka et al. (2016). "Classification of forest land attributes using multi-source remotely sensed data." In: *International Journal of Applied Earth Observation and Geoinformation* 44, pp. 11–22 (cit. on pp. 5, 16).

Probeck, Markus et al. (2014). "European forest monitoring approaches." In: *Land Use and Land Cover Mapping in Europe*. Springer, pp. 89–114 (cit. on pp. 4–7).

Radoux, Julien et al. (2016). "Sentinel-2's Potential for Sub-Pixel Landscape Feature Detection." In: *Remote Sensing* 8.6, p. 488 (cit. on pp. 12, 53, 66).

Rees, William G. (2001). *Physical principles of remote sensing*. English. 2., 1. publ. Cambridge [u.a.]: Cambridge Univ. Press (cit. on pp. 9, 10, 14).

Reichle, David E (2013). *Analysis of temperate forest ecosystems*. Vol. 1. Springer Science & Business Media (cit. on p. 28).

Richards, John A. and Xiuping Jia (1999). *Remote sensing digital image analysis: an introduction*. English. 3;3., rev. and enlarg; Berlin [u.a.]: Springer. ISBN: 3540648607;9783540648604; (cit. on p. 10).

Richter, R, D Schläpfer, and A Müller (2006). "An automatic atmospheric correction algorithm for visible/NIR imagery." In: *International Journal of Remote Sensing* 27.10, pp. 2077–2085 (cit. on pp. 34, 35, 37).

Richter, Rudolf et al. (2011). "Correction of cirrus effects in Sentinel-2 type of imagery." In: *International journal of remote sensing* 32.10, pp. 2931–2941 (cit. on pp. 35, 36, 46).

Rigo, D. de et al. (2016). "European Atlas of Forest Tree Species." In: ed. by J San-Miguel-Ayanz et al. Luxembourg. Chap. European forests: an ecological overview (cit. on p. 28).

Sannier, Christphe et al. (2015). *GMES Initial Operations / Copernicus Land monitoring services – Validation of products, PRELIMINARY HRL FOREST 2012 VALIDATION REPORT*. Tech. rep. SIRS, JR, EEA (cit. on pp. 7, 21–25, 50, 86, 89).

Schuck, Andreas et al. (2002). "Compilation of a calibrated European forest map derived from NOAA-AVHRR data." In: (cit. on p. 4).

Spiecker, Heinrich (2003). "Silvicultural management in maintaining biodiversity and resistance of forests in Europe—temperate zone." In: *Journal of Environmental Management* 67.1, pp. 55–65 (cit. on pp. 28, 30).

Suhet (2015). *Sentinel-2 User Handbook*. 1.2. European Space Agency (cit. on pp. 18–20, 39).

Tarquini, S. et al. (2012). *Release of a 10-m-resolution DEM for the Italian territory: Comparison with global-coverage DEMs and anaglyph-mode exploration via the web*. visited on 2017-01-30 (cit. on p. 27).

Thessler, Sirpa et al. (2008). "Using k-nn and discriminant analyses to classify rain forest types in a Landsat TM image over northern Costa Rica." In: *Remote Sensing of Environment* 112.5, pp. 2485–2494 (cit. on pp. 5, 15, 74, 75).

U.S. Department of the Interior, U.S. Geological Survey (2016). *Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global*. visited on 2016-11-30. URL: https://lta.cr.usgs.gov/SRTM1Arc (cit. on p. 25).

U.S. Geological Survey (2016). *Band Designations Landsat Satellites*. visited on 2016-01-15. URL: http://landsat.usgs.gov/band_designations_landsat_satellites.php (cit. on p. 39).

Xiao, Xiangming et al. (2002). "Characterization of forest types in Northeastern China, using multi-temporal SPOT-4 VEGETATION sensor data." In: *Remote Sensing of Environment* 82.2, pp. 335–348 (cit. on pp. 15, 98, 99).

Zhu, Xiaolin and Desheng Liu (2014). "Accurate mapping of forest types using dense seasonal Landsat time-series." In: *ISPRS Journal of Photogrammetry and Remote Sensing* 96, pp. 1–11. ISSN: 0924-2716. DOI: http://dx.doi.org/10.1016/j.isprsjprs.2014.06.012 (cit. on pp. 3, 7, 15, 17, 98–100).

Zhu, Zhe, Shixiong Wang, and Curtis E Woodcock (2015). "Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images." In: *Remote Sensing of Environment* 159, pp. 269–277 (cit. on pp. 44, 47).

Appendix A.

Appendix A.

Table A.1.: Confusion matrix - classifier: knn, training samples: all, band combination: 3LD

	Broadleaved	Coniferous	Mixed	Total	User Accuracy
Broadleaved	43	5	12	60	0.717
Coniferous	2	28	4	34	0.824
Mixed	1	4	5	10	0.5
Total	46	37	21	104	0
Producer Accuracy	0.935	0.757	0.238	0	0.731

Table A.2.: Confusion matrix - classifier: knn, training samples: full, band combination: 3LD

	Broadleaved	Coniferous	Mixed	Total	User Accuracy
Broadleaved	7	1	6	14	0.5
Coniferous	0	15	3	18	0.833
Mixed	2	1	5	8	0.625
Total	9	17	14	40	0
Producer Accuracy	0.778	0.882	0.357	0	0.675

Table A.3.: Confusion matrix - classifier: knn, training samples: tcd, band combination: 3LD

	Broadleaved	Coniferous	Mixed	Total	User Accuracy
Broadleaved	35	1	10	46	0.761
Coniferous	1	24	4	29	0.828
Mixed	4	5	10	19	0.526
Total	40	30	24	94	0
Producer Accuracy	0.875	0.8	0.417	0	0.734

Table A.4.: Confusion matrix - classifier: knn, training samples: corine, band combination: 3LD

	Broadleaved	Coniferous	Mixed	Total	User Accuracy
Broadleaved	19	3	5	27	0.704
Coniferous	2	25	5	32	0.781
Mixed	0	3	10	13	0.769
Total	21	31	20	72	0
Producer Accuracy	0.905	0.806	0.5	0	0.750

Table A.5.: Confusion matrix - classifier: rf, training samples: all, band combination: all

	Broadleaved	Coniferous	Mixed	Total	User Accuracy
Broadleaved	40	5	8	53	0.755
Coniferous	1	29	6	36	0.806
Mixed	5	3	7	15	0.467
Total	46	37	21	104	0
Producer Accuracy	0.87	0.784	0.333	0	0.731

Table A.6.: Confusion matrix - classifier: rf, training samples: full, band combination: all

	Broadleaved	Coniferous	Mixed	Total	User Accuracy
Broadleaved	9	0	6	15	0.6
Coniferous	0	15	3	18	0.833
Mixed	0	2	5	7	0.714
Total	9	17	14	40	0
Producer Accuracy	1	0.882	0.357	0	0.725

Table A.7.: Confusion matrix - classifier: rf, training samples: tcd, band combination: all

	Broadleaved	Coniferous	Mixed	Total	User Accuracy
Broadleaved	33	1	12	46	0.717
Coniferous	2	26	2	30	0.867
Mixed	5	2	11	18	0.611
Total	40	29	25	94	0
Producer Accuracy	0.825	0.897	0.44	0	0.745

Table A.8.: Confusion matrix - classifier: rf, training samples: corine, band combination: all

	Broadleaved	Coniferous	Mixed	Total	User Accuracy
Broadleaved	19	6	9	34	0.559
Coniferous	2	24	1	27	0.889
Mixed	0	1	10	11	0.909
Total	21	31	20	72	0
Producer Accuracy	0.905	0.774	0.5	0	0.736

Table A.9.: Error matrix based on inclusion probabilities - classifier: knn, training samples: all, band combination: 3LD

	Broadleaved	Coniferous	Mixed	Total	Commission Error
Broadleaved	0.382	0.005	0.0866	0.474	0.193
Coniferous	0.0225	0.308	0.0662	0.397	0.224
Mixed	0.000649	0.0457	0.083	0.129	0.358
Total	0.405	0.359	0.236	1	0
Omission Error	0.0571	0.141	0.648	0	0.773

Table A.10.: Error matrix based on inclusion probabilities - classifier: knn, training samples: full, band combination: 3LD

	Broadleaved	Coniferous	Mixed	Total	Commission Error
Broadleaved	0.197	0.00315	0.0254	0.226	0.126
Coniferous	0	0.337	0.185	0.522	0.354
Mixed	0.0373	0.00315	0.212	0.253	0.16
Total	0.235	0.343	0.422	1	0
Omission Error	0.159	0.0184	0.497	0	0.746

Table A.11.: Error matrix based on inclusion probabilities - classifier: knn, training samples: tcd, band combination: 3LD

	Broadleaved	Coniferous	Mixed	Total	Commission Error
Broadleaved	0.408	0.00086	0.0828	0.492	0.17
Coniferous	0.00086	0.238	0.0562	0.295	0.194
Mixed	0.0282	0.0374	0.148	0.213	0.308
Total	0.437	0.276	0.287	1	0
Omission Error	0.0665	0.139	0.485	0	0.794

Table A.12.: Error matrix based on inclusion probabilities - classifier: knn, training samples: corine, band combination: 3LD

	Broadleaved	Coniferous	Mixed	Total	Commission Error
Broadleaved	0.335	0.00347	0.0397	0.378	0.114
Coniferous	0.00335	0.327	0.00844	0.339	0.0348
Mixed	0	0.103	0.18	0.283	0.365
Total	0.339	0.434	0.228	1	0
Omission Error	0.00989	0.246	0.211	0	0.842

Table A.13.: Error matrix based on inclusion probabilities - classifier: rf, training samples: all, band combination: all

	Broadleaved	Coniferous	Mixed	Total	Commission Error
Broadleaved	0.334	0.00403	0.0765	0.415	0.194
Coniferous	0.000649	0.334	0.023	0.358	0.0662
Mixed	0.0706	0.0209	0.136	0.228	0.402
Total	0.405	0.359	0.236	1	0
Omission Error	0.176	0.0694	0.422	0	0.804

Table A.14.: Error matrix based on inclusion probabilities - classifier: rf, training samples: all, band combination: all

	Broadleaved	Coniferous	Mixed	Total	Commission Error
Broadleaved	0.235	0	0.0254	0.26	0.0977
Coniferous	0	0.332	0.185	0.517	0.357
Mixed	0	0.011	0.212	0.223	0.0494
Total	0.235	0.343	0.422	1	0
Omission Error	0	0.0321	0.497	0	0.779

Table A.15.: Error matrix based on inclusion probabilities - classifier: rf, training samples: tcd, band combination: all

	Broadleaved	Coniferous	Mixed	Total	Commission Error
Broadleaved	0.38	0.00086	0.0743	0.455	0.165
Coniferous	0.00199	0.269	0.00172	0.272	0.0136
Mixed	0.0559	0.00301	0.214	0.273	0.216
Total	0.437	0.273	0.29	1	0
Omission Error	0.132	0.0142	0.262	0	0.862

Table A.16.: Error matrix based on inclusion probabilities - classifier: rf, training samples: corine, band combination: all

	Broadleaved	Coniferous	Mixed	Total	Commission Error
Broadleaved	0.336	0.00661	0.0715	0.414	0.189
Coniferous	0.00242	0.398	0.00138	0.402	0.00945
Mixed	0	0.0287	0.155	0.183	0.156
Total	0.339	0.434	0.228	1	0
Omission Error	0.00716	0.0814	0.32	0	0.889

Appendix B.

Table B.1.: TCD Confusion Matrix: knn classifier, band combination 4LD

TCD	1-20%	21-40%	41-60%	61-80%	81-100%	Total	User Accuracy
1-20%	1	0	0	0	0	1	1
21-40%	16	1	0	0	0	17	0.80
41-60%	12	5	3	4	1	25	0.73
61-80%	1	7	9	16	7	40	0.66
81-100%	0	0	1	4	7	12	0.94
Total	30	13	13	24	15	95	
Producer Accuracy	0.03	0.08	0.23	0.67	0.47		0.29

Table B.2.: TCD Confusion Matrix: knn classifier, band combination 7LD

TCD	1-20%	21-40%	41-60%	61-80%	81-100%	Total	User Accuracy
1-20%	3	0	0	0	0	3	1
21-40%	16	3	1	1	0	21	0.78
41-60%	11	5	4	4	1	25	0.74
61-80%	0	5	8	17	9	39	0.69
81-100%	0	0	0	2	5	7	0.97
Total	30	13	13	24	15	95	
Producer Accuracy	0.1	0.23	0.31	0.71	0.33		0.33684

Table B.3.: TCD Confusion Matrix: knn classifier, band combination allD

TCD	1-20%	21-40%	41-60%	61-80%	81-100%	Total	User Accuracy
1-20%	3	0	0	0	0	3	1
21-40%	13	3	1	2	0	19	0.80
41-60%	11	5	6	5	1	28	0.73
61-80%	0	5	7	18	13	43	0.64
81-100%	0	0	0	1	1	2	0.07
Total	27	13	14	26	15	95	
Producer Accuracy	0.11	0.23	0.43	0.69	0.07		0.32632

Table B.4.: TCD Confusion Matrix: knn classifier, band combination wbD

TCD	1-20%	21-40%	41-60%	61-80%	81-100%	Total	User Accuracy
1-20%	3	0	0	0	0	3	1
21-40%	16	3	1	2	0	22	0.77
41-60%	11	5	5	5	1	27	0.73
61-80%	0	5	7	16	13	41	0.65
81-100%	0	0	0	1	1	2	0.99
Total	30	13	13	24	15	95	
Producer Accuracy	0.1	0.23	0.38	0.67	0.07		0.29474

Table B.5.: TCD Confusion Matrix: random forest classifier, band combination 4L

TCD	1-20%	21-40%	41-60%	61-80%	81-100%	Total	User Accuracy
1-20%	1	0	0	0	0	1	1
21-40%	13	3	0	0	0	16	0.84
41-60%	13	9	10	18	9	59	0.40
61-80%	0	3	3	6	6	18	0.83
81-100%	0	0	0	0	1	1	1
Total	27	15	13	24	16	95	
Producer Accuracy	0.04	0.20	0.77	0.25	0.06		0.22105

Table B.6.: TCD Confusion Matrix: random forest classifier, band combination 7L

TCD	1-20%	21-40%	41-60%	61-80%	81-100%	Total	User Accuracy
1-20%	3	0	0	0	0	3	1
21-40%	17	4	0	0	0	21	0.79
41-60%	10	4	8	8	2	32	0.70
61-80%	0	5	5	14	7	31	0.76
81-100%	0	0	0	2	6	8	0.98
Total	30	13	13	24	15	95	
Producer Accuracy	0.1	0.31	0.62	0.58	0.4		0.36842

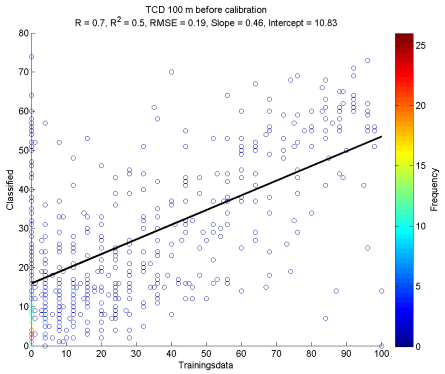
Table B.7.: TCD Confusion Matrix: random forest classifier, band combination all

TCD	1-20%	21-40%	41-60%	61-80%	81-100%	Total	User Accuracy
1-20%	2	0	0	0	0	2	1
21-40%	18	4	0	0	0	22	0.78
41-60%	10	4	9	8	2	33	0.71
61-80%	0	5	4	15	9	33	0.75
81-100%	0	0	0	1	4	5	0.99
Total	30	13	13	24	15	95	
Producer Accuracy	0.07	0.31	0.69	0.63	0.27		0.35789

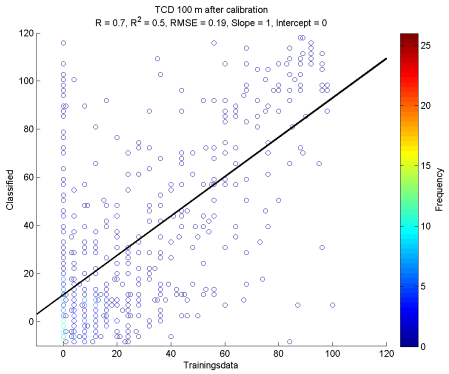
Table B.8.: TCD Confusion Matrix: random forest classifier, band combination wb

TCD	1-20%	21-40%	41-60%	61-80%	81-100%	Total	User Accuracy
1-20%	2	0	0	0	0	2	1
21-40%	19	4	0	0	0	23	0.77
41-60%	9	4	10	9	2	34	0.71
61-80%	0	5	3	14	8	30	0.77
81-100%	0	0	0	1	5	6	0.99
Total	30	13	13	24	15	95	
Producer Accuracy	0.07	0.31	0.77	0.58	0.33		0.36842

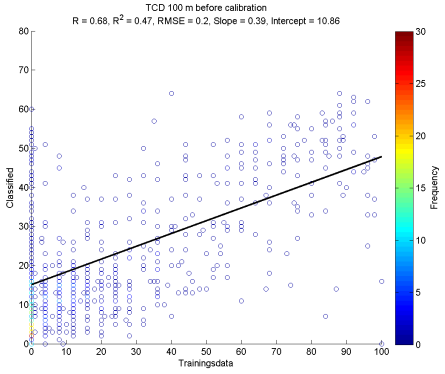
Appendix C.



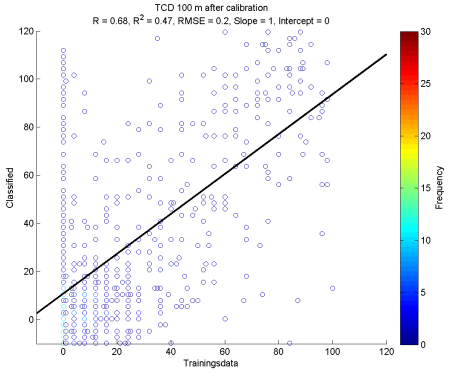
(a) CL: knn, BC: 7LD, before calibration



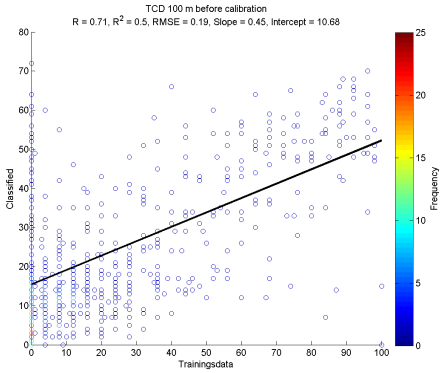
(b) CL: knn, BC: 7LD, after calibration



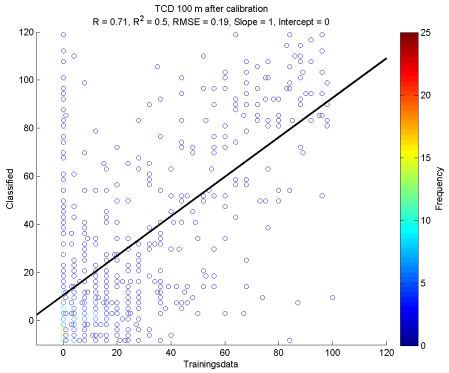
(c) CL: knn, BC: allD, before calibration



(d) CL: knn, BC: allD, after calibration

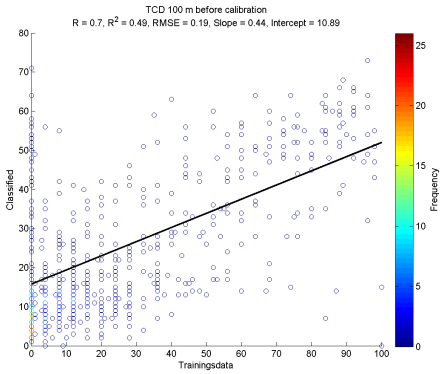


(e) CL: knn, BC: wbD, before calibration

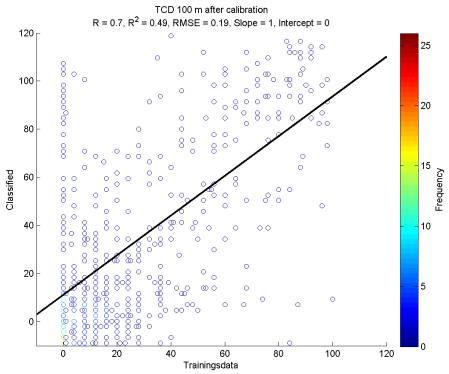


(f) CL: knn, BC: wbD, after calibration

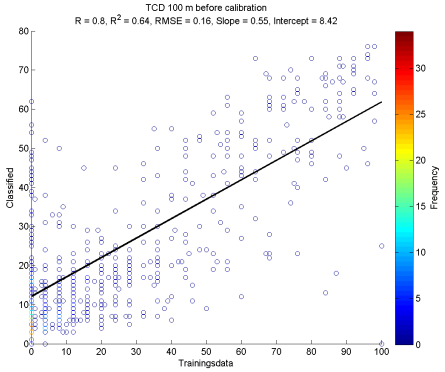
Figure C.1.: Scatter plot of predicted and trainings TCD data based on the 100 m mosaic.



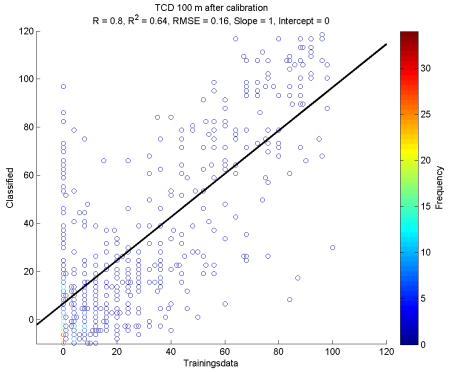
(a) CL: knn, BC: 4LD, before calibration



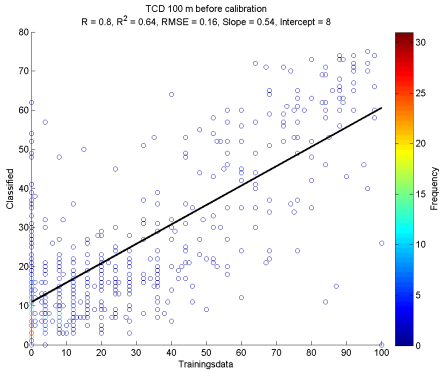
(b) CL: knn, BC: 4LD, after calibration



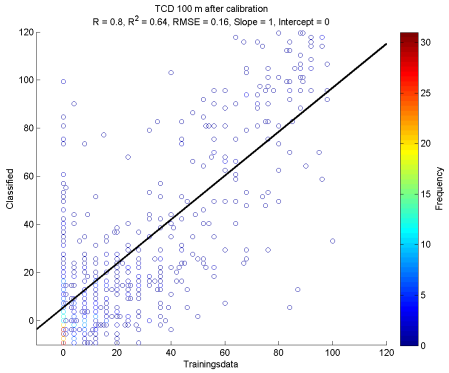
(c) CL: rf, BC: 7L, before calibration



(d) CL: rf, BC: 7L, after calibration

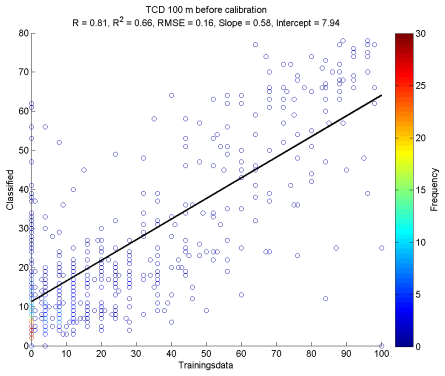


(e) CL: rf, BC: all, before calibration

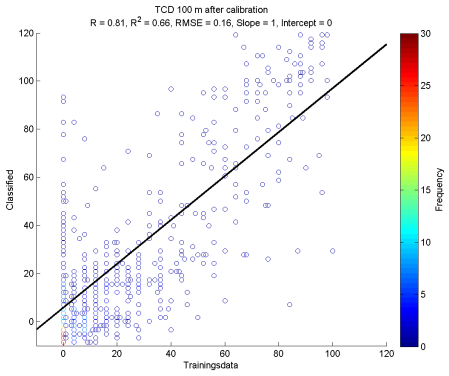


(f) CL: rf, BC: all, after calibration

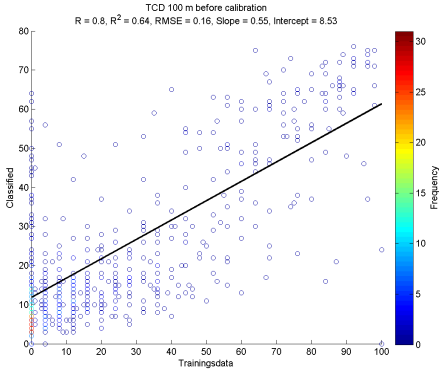
Figure C.2.: Scatter plot of predicted and trainings TCD data based on the 100 m mosaic.



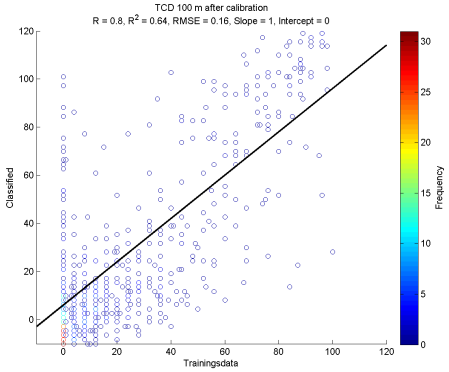
(a) CL: rf, BC: wb, before calibration



(b) CL: rf, BC: wb, after calibration



(c) CL: rf, BC: 4L, before calibration



(d) CL: rf, BC: 4L, after calibration

Figure C.3.: Scatter plot of predicted and trainings TCD data based on the 100 m mosaic.