

Modeling and Mapping Leaf Area Index Using Spectral and Spatial information of High-Resolution Multispectral Images: Case Study of Bavarian Forest National Park

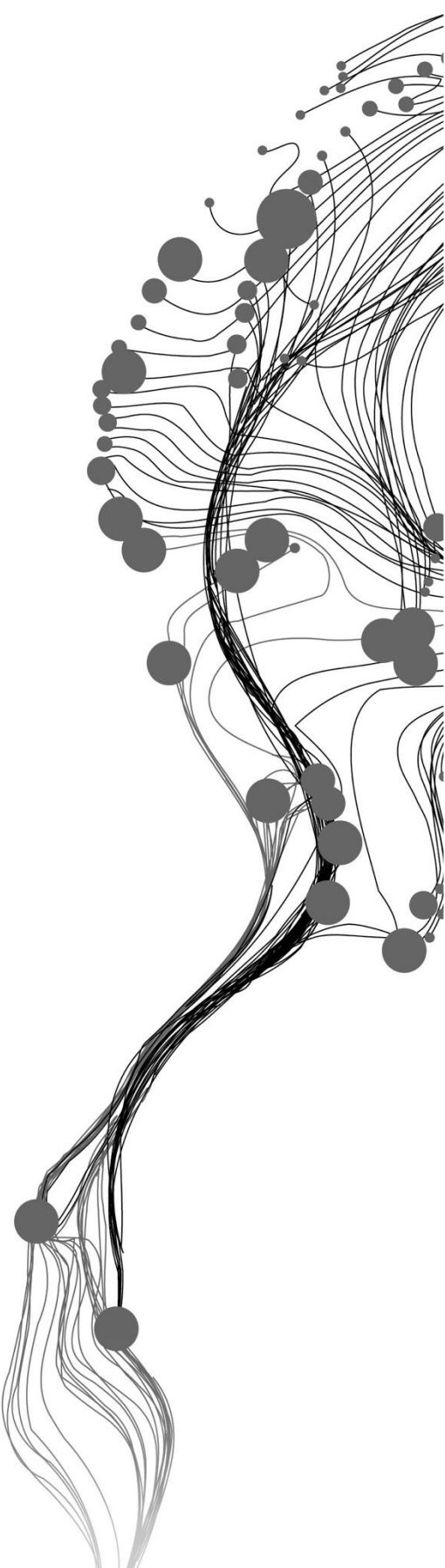
SOODABEH AMIRINEZHAD

February 2016

SUPERVISORS:

Dr. (Roshanak) Darvishzadeh

Prof. dr. (Andrew) Skidmore



Modeling and Mapping Leaf Area Index Using Spectral and Spatial information of High-Resolution Multispectral Images: Case Study of Bavarian Forest National Park

SOODABEH AMIRINEZHAD

Enschede, The Netherlands, February 2016

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Natural Resource Management

SUPERVISORS:

Dr. (Roshanak) Darvishzadeh

Prof. dr. (Andrew) Skidmore

THESIS ASSESSMENT BOARD:

Dr.Ir. C.A.J.M.(Kees) de Bie (Chair)

Dr. J. Clevers (External examiner, University of Wageningen)

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

ABSTRACT

Leaf Area Index (LAI) is important biophysical parameters in the forest. Having accurate maps of LAI spatial distribution is a necessity for forest management because they help to monitor changes in trees canopy. Optical remote sensing has provided valuable source of information in vast areas for LAI estimation. Since the new satellite image of RapidEye which contains the red edge band (690-730nm) has been introduced, it is possible to extract additional information from the reflectance of target object on the image. As such, high-resolution multispectral images of RapidEye have been utilized in this study. Field measurements of LAI were obtained in August 2015 in Bavarian Forest National Park, Germany, concurrent with the time of the RapidEye image. Spectral and spatial information extracted from the RapidEye images have been utilized for estimating LAI.

Three empirical methods have been investigated: (i) Vegetation Indices (VI) which involved calculation of ratio based and red-edge indices, (ii) Gray Level Co-occurrence Matrix (GLCM) approach which used to obtain different Texture Variables (TVs) and (iii) Random Forest which utilized the combination of spectral bands, VIs and TVs. Regression models were used to indicate the relationship between the dependent variable (measured LAI) and independent variable (estimated LAI from VIs or GLCM). Cross validation also was used to evaluate the regression model to assess how the model could estimate LAI.

The results revealed that VIs did not yield reliable LAI with the cross-validated RMSE= 1.23 and $R^2 = 0.27$ and also the result of GLCM returned cross-validated RMSE = 1.23 and $R^2 = 0.28$. However the use of the Random Forest model which benefits from building independent decision trees raised the accuracy of LAI estimation to RMSE = 0.85 and $R^2 = 0.43$. The results demonstrated that the accuracy of LAI retrieval from multispectral remote sensing data can be improved when multivariate methods such as Random Forest are utilized.

Keywords: Leaf Area Index, Vegetation Indices, GLCM, Random Forest, RapidEye, Bavarian Forest National Park.

ACKNOWLEDGEMENTS

First of all, I am grateful to the God for the good health and well-being that were necessary to complete this thesis.

I wish to express my sincere thanks to Excellent UTS-ITC scholarship committee, for providing me this opportunity to pursue my abilities and interests.

I cannot express enough thanks to my first supervisor Dr. Roshanak Darvishzadeh for her continued support and encouragement for accomplishing this thesis. The door of her office was always open whenever I had a question about my research or writing. She consistently allowed this thesis to be my own work, but steered me in the right direction whenever she thought I needed it.

I would like to have a very special thanks to my second supervisor Prof. Dr. Andrew Skidmore. Without his companionship and elaborate comments, finishing present thesis was not possible. I was always inspired by his knowledge, outstanding comments and very kind smile. Thanks.

I must express my very profound gratitude to my friend *Hamed* who was always supportive and standing next to me whenever I was exhausted and in difficult moments during the thesis. I would like to thank other friends, *Samaneh*, *John*, *Elnaz* and *Asliban* for their kindness and encouragement. Your being was so valuable to me. I also would like to thank *Alex Zvoleff*, the maintainer of GLCM code in R and *Willem Nieuwenhuis* ICT developer of Natural Resources department, ITC faculty, for their supports by guiding and responding to the questions.

Last but not least, I take this opportunity to show my deep appreciation to my father *Hamid*, my mother *Touba* and my sister *Hanieh* for providing me with unfailing support and continuous encouragement from different continents in the world. THANK YOU AND MISS YOU!

Soodابه Amirinezhad

Enschede,

The Netherlands, February 2016

TABLE OF CONTENTS

1.	Introduction.....	1
1.1.	Background.....	1
1.1.1.	Importance of forest structure and LAI.....	1
1.1.2.	Remote sensing and LAI estimation.....	2
1.1.3.	Statistical and physical approaches.....	2
1.1.4.	Spectral information: Vegetation Indices.....	3
1.1.5.	Spatial information: Textural Information.....	4
1.2.	Literature review:	4
1.2.1.	Parametric models.....	4
1.2.2.	Machine learning based models.....	5
1.3.	Research problem	6
1.3.1.	Research objectives.....	7
1.3.2.	Research questions.....	7
1.3.3.	Research Hypothesis	7
2.	Materials:	8
2.1.	Study area	8
2.2.	Field measurement of LAI	9
2.3.	RapidEye Imagery.....	10
3.	Method	12
3.1.	Image Pre-processing.....	12
3.2.	Vegetation Indices generation.....	13
3.3.	Texture Variables generation	14
3.4.	Linear Regression Analysis.....	17
3.5.	Random Forest modeling.....	17
3.6.	Cross-validation.....	18
3.7.	LAI mapping.....	18
3.8.	Flowchart.....	18
4.	Results.....	20
4.1.	The characteristics of the measured LAI in Bavarian Forest National Park.....	20
4.2.	Correlation between Vegetation Indices and LAI.....	20
4.3.	Correlation between Texture Variables and LAI.....	22
4.4.	Random Forest.....	24
5.	Discussion.....	28
5.1.	The influence of atmospheric errors.....	29
6.	Conclusion and Recommendation.....	31
6.1.	Brief answer to the research questions:	31
6.2.	Suggestions for further work:	32

LIST OF FIGURES

Figure 1. The location of the study area, Bavarian Forest National Park	9
Figure 2. Landscape of Bavarian Forest National Park.....	9
Figure 3. Measuring LAI in the plot using LAI 2200	10
Figure 4. Distribution of the sample plots in Bavarian Forest National Park	10
Figure 5. Gray Level Co-occurrence Matrix operation	15
Figure 6. The methodology using multispectral satellite images to model and map LAI.....	19
Figure 7. The correlation between Vegetation Indices and measured LAI.....	21
Figure 8. Plot of cross-correlation between the measured and estimated cross-validated LAI.....	22
Figure 9. The correlation between Textural Variables of GLCM band3 and the measured LAI	23
Figure 10. The correlation between Textural Variables of GLCM band 5 and the measured LAI	23
Figure 11. The correlation between Textural Variables of GLCM RRI1 and the measured LAI.....	23
Figure 12. Plot of cross-correlation between measured and estimated cross-correlated LAI	24
Figure 13. Leave One Out error for each plot by Random Forest model.....	25
Figure 14. Plot of cross-correlation between the measured and estimated cross-correlated LAI.....	26
Figure 15. LAI map of Bavarian Forest National Park created from Random Forest model	27

LIST OF TABLES

Table 1. Summary of specifications of RapidEye sensor	11
Table 2. Selected Vegetation Indices for LAI estimation	14
Table 3. 2 nd -order Texture Variables and equations	16
Table 4. Summary statistics of the measured LAI in Bavarian Forest National Park	20
Table 5. Final applied window sizes of GLCM for two bands and raster layer	22
Table 6. The correlation between combination of band and window sizes and measured LAI.	24
Table 7. 10 top-ranked important predictor variables obtained from Random Forest.....	25

“Education is not preparation for life, education is life itself.”

*John Dewey,
1859-1952*

1. INTRODUCTION

In the following chapter the importance of forest ecosystem, remote sensing of vegetation and the approaches for LAI estimation using image spectral and spatial information are discussed. Further, previous studies regarding LAI estimation in forest ecosystems are discussed. Finally at the end of the chapter research problem, objectives as well as research questions and hypothesis are presented.

1.1. Background

In this section we review the importance of forests and LAI as well as current LAI measurement techniques. We also review spectral and spatial characteristics of the remotely sensed satellite images and improvement in computational processing which motivated us to test a methodology for LAI estimation.

1.1.1. Importance of forest structure and LAI

Forests are origin of valuable timber and non-timber products that we use every day in our life. For example, forests are responsible for production of large amount of oxygen and they are places for vital processes such as photosynthesis and evapotranspiration (Maass, Vose, Swank, & Martínez-Yrizar, 1995). Moreover, forests cover about 31% world's surface (Food and Agriculture Organization of the United Nations, 2006), and they provide habitat for a vast variety of flora and fauna. Therefore, characterization of the forest biophysical parameters is very important. There are several biophysical parameters related to plant canopies. Quantitative measurements of these parameters are needed to study the productivity of forest and to monitor the fluxes of energy and matters (Yu, Yang, & Fan, 2015).

In this study, Leaf Area Index (LAI), which is widely used to quantify the structure of the tree canopy in the forest, is studied. LAI is defined as the ratio of the one-sided area of the leaf per unit ground surface area, and it is a dimensionless quantity (Watson, 1947). It is used to determine interactions between atmosphere and tree leaves surface (Monteith and Unsworth, 1990) such as radiation uptake, precipitation interception, conversion of energy and gas exchange (e.i., Leaves emit O_2 to the atmosphere and absorb CO_2 from it). Leaf area is an essential element of photosynthetic carbon assimilation, and, therefore, LAI is used as an indicator of trees growth (Barclay, 1998).

For deciduous trees, LAI increases during the growing season and the minimum LAI is in winter (leaf-off) season, however, LAI of coniferous trees shows less variability in the year (Scurlock, Asner, & Gower, 2001). LAI is widely used to monitor forests over time because it is an indicator of changes and damaged to trees caused by human activities and natural hazards such as frost, storm, defoliation and drought (Maass et al., 1995; Bréda, 2003; Liu, Chen, Jin, & Qi, 2015). Thus, measured LAIs is a relevant source of data for management and conservation of forests.

There are direct and indirect methods to measure LAI. The first group of methods called direct method includes harvesting and non-harvesting methods. In harvesting method, leaf area and vertical distribution of leaf area is measured leaf by leaf (Jonckheere et al., 2004). In non-harvesting method, litter leaves are collected in litter traps during leaf-off season to measure the one-sided area of leaves by using planimetric and gravimetric methods (Zheng & Moskal, 2009). If leaves in latter trap overlap, the measured LAI value will be underestimated. Although direct methods are accurate, they are labour-intensive and costly (Barclay, 1998) and they are often used to measure references for the indirect methods. The second method include indirect contact and non-contact methods. In this method optical measurment such as hemispherical photography, sunfleck centimeters and optical instruments are used (e.g., TRAC, LAI-2000, LI-COR) (Jonckheere et al., 2004). Estimating LAI from remote sensing images is the most common indirect non-contact method (Zheng & Moskal, 2009) which is not labor involved and it is preferable for estimating LAI at large spatial and temporal scales (Chen et al., 1997).

1.1.2. Remote sensing and LAI estimation

Remote sensing is defined as the science of gathering information without any direct contact with the object on the earth (Curran, 1980). In particular, the application of satellite images is remarkably increasing because they are applicable to non-destructively acquiring data from large-scale areas (Kayitakire et al., 2006). Moreover, recent developments in remote sensing have improved the spatial and spectral resolution of satellite images (Peddle et al., 1999). In satellite image, target objects on the earth (e.g. canopy of the trees in a forest) are identified and characterized based on their unique spectral signature. According to the type of the sensors, reflected radiation from the target objects are allocated to high spectral resolution of hyperspectral images (narrowband sensors), or to a few broadband wavelengths in multispectral images(broadband sensors).

Although hyperspectral images provide a very detailed information based on the objects reflectance, there are drawbacks of using them. First, the cost of acquiring this data is relatively high. Second, processing hyperspectral data has technical complexity (He et al., 2011). However, processing multispectral images is easier and more cost-effective than hyperspectral images. Multispectral images offer few broad spectral bands to retrieve a wide range of biophysical, biochemical and structural parameters of plants including: LAI (Shen, Li, & Guo., 2014), crown closure (Lee et al., 2006), biomass (Güneralp, Filippi, & Randall, 2014), Nitrogen (Bagheri, Ahmadi, Alavipanah, & Omid, 2013), water content (Wang, Wang, Zhou, Liu, & Wang, 2011), chlorophyll (Croft et al., 2015), specific leaf area (Johnson, Roczen, Youkhana, Nemani, & Bosch, 2003) and tree height (Sprintsin, Berliner, Cohen, & Karnieli, 2013). In particular, for LAI, analyzing the relationship between canopy characteristics and their corresponding reflected electromagnetic radiation on a multispectral images is used to estimate LAI in large scale vegetated area such as forest (Pasolli et al., 2015).

1.1.3. Statistical and physical approaches

For estimating LAI from remote sensing images, there are two general approaches. The first one is physical approach describing the transfer and interaction of radiation inside the canopy. This approach provides an explicit connection between the LAI and the canopy reflectance (Houborg, Soegaard, & Boegh, 2007). The

second approach, so called statistical approach, creates a statistical relationship between measured biophysical parameters and indices and measures based on surface reflectance (Ehammer, Fritsch, Conrad, & Lamers, 2010).

Both the physical and statistical approaches have some advantages and disadvantages. With the physical approach, different strategies are proposed for the inversion of physical models including numerical optimization method, look-up tables and Artificial Neural Network (NN) (Combal et al., 2003; Fang & Liang, 2003). Although physical approach is robust for assessing the effect of biophysical variables on canopy reflectance (Yuan et al., 2015), it is unfeasible to be used in large areas for estimating LAI. Physical approach also has ill-posed problem which is the nature of model inversion (Combal, Baret & Weiss, 2002; Houborg et al., 2007).

The statistical approach emphasizes on the correlation between in-situ LAI measurements and their corresponding spectral-driven information. The use of statistical approach is more common compared to physical approach because it is quicker to implement than physical approach, and often, an comparable estimation is obtained (Yang, Zhao, Liu, Huang, & Wang, 2011). Several research studies have been done using statistical approach for estimating either crop LAI (Johnson, Roczen, Youkhana, Nemani, & Bosch, 2003; Xie et al., 2014) or forest LAI (Chen & Cihlar, 1996; Wulder, LeDrew, Franklin, & Lavigne, 1998; Cohen, Maersperger, Gower, & Turner, 2003; Schlerf, Atzberger, & Hill, 2005; Wang, Adiku, Tenhunen, & Granier, 2005; Soudani, François, le Maire, Le Dantec, & Dufrêne, 2006) (For more details refer to appendix A).

In the statistical approach two type of models is used which are univariate and multivariate models. Univariate models for estimating LAI is confined to the use of only one variable extracted from spectral or spatial characteristics of remote sensing image (Fassnacht, Gower, MacKenzie, Nordheim, & Lillesand, 1997; Gong, Pu, Biging, & Larrieu, 2003). There are several univariate-based studies that use Vegetation Indices (VIs) and Texture Variables (TVs) to estimate LAI (Curran, Dungan, & Gholz, 1992). VIs are calculated from spectral reflectance of plants and TVs extracted from spatial information of remote sensing images. However, such univariate LAI estimation is often site-specific and it is influenced by variation of atmospheric condition (Bannari, Morin, Bonn, & Huete, 1995). These drawbacks of univariate models motivated researchers to make use of multivariate models such as multiple stepwise linear regression and partial least square regression that have two or more independent variables (Pu, 2012).

In particular, the statistical multivariate models based on Machine Learning Algorithm (MLA) are getting more and more popular. Machine learning is a scientific disciplines that explore the construction of the algorithm which can learn from the data (Brownlee, 2013). MLA have the potential to handle a large number of input variables (Kashyap, Ahmed, Hoque, Roy, & Bhattacharyya, 2015) to estimate LAI in large-scale areas such as forest. Advantage of MLA is that they often can rank the importance of inputs variables (i.e. predictors) which is also crucial for understanding the effectiveness of variables for estimating LAI (Prasad, Iverson, & Liaw, 2006). Considering the advantages of multivariate models that use MLA, there are two potential groups of input variables that can be used to estimate LAI. These groups are reviewed in more detail in the following sub-sections.

1.1.4. Spectral information: Vegetation Indices

VIs are mathematical transformations to measure the spectral contribution of the vegetation on the satellite images (Elvidge & Chen, 1995). Considering spectral properties of the plants, different spectral bands are used in VIs equations (i.e. combination of visible bands (e.g. Blue, Green, Red) and NIR band). The reflectance in Red band coming from chlorophyll content and reflectance in NIR is based on the internal

structure of the plant cell (Mróz & Sobieraj, 2004). Red edge wavelength, which previously used to be obtained by calculating specific wavelength in hyperspectral images, have been introduced as a completely separate band in a few satellite images such as RapidEye image. Less sensitivity of Red edge band to the atmospheric errors results in a better estimation of LAI using VIs. The old common VI, Normalized Difference Vegetation Index (NDVI) is a well-known index for estimating various parameters in different ecosystems (Turner, Cohen, Kennedy, Fassnacht, 1999). In this VI, Red and NIR bands are involved. Since Red edge band has been introduced it can be replaced by any band in VI formulas.

It is important to consider, VIs might be affected by different factors such as soil background, atmospheric scattering and also tree structural factors such as leaf angle and canopy structure. For example, in some cases, soil color influences the value of VI, such as dark soil caused higher values of VIs and conversely, bright soil results in a smaller value in NDVI and higher values in Perpendicular Vegetation Index (PVI) (Schmidt & Karnieli, 2001). But exploitation of some bands of the image in VIs formula can minimize or eliminate these effects. For example, Blue band is used to separate the effect of atmosphere from the vegetation reflectance (Mróz & Sobieraj, 2004).

1.1.5. Spatial information: Textural Information

Tree species structural properties such as difference in crown closure, shadows, or tree density can be detected in the measures of texture calculated on high resolution remote sensing images (Wulder, LeDrew, Franklin, & Lavigne, 1998). Texture is a complex parameter to be extracted, and it depends on the target objects that are studied (in this case trees in the forest), physiographic condition, and the window size (Zhou et al., 2014). Texture provides structural information as a low-cost additional source of information which improves LAI estimation.

Spatial features describe how tonal variation spatially distributed within a band (Haralick, Shanmugan, & Dinstein, 1973). The limited number of bands in the multispectral image made it possible to extract useful information from one single band, especially the one which is sensitive to the vegetation reflectance such as NIR band. Based on the texture pattern of the forest which caused by vegetation coverage, heterogeneity and tree species, obtained values of TVs differ. Obviously, the most heterogeneous an area, the more spatially variable and the more homogeneous an area, the less spatially variable it is (Song & Dickinson, 2008).

1.2. Literature review:

Considering the explanations in background, in this section we are going to first highlight some research studies that applied parametric models (i.e. using VIs and texture) and next we will review studies that have been done using MLA for estimating LAI.

1.2.1. Parametric models

In many studies, it has been proven that utilization of VIs is a direct and effective method for LAI estimation. Although VIs have been developed in the 1970s, their importance can be understood when they are still calculated from satellite images for different scales and purposes such as predicting vegetation productivity and estimating biophysical parameters (Steven, Malthus, Baret, Xu, & Chopping, 2003). Some examples of using VIs and texture and their combination in the previous studies are as follow:

Curran et al., (1992) investigated the linear relationship between NDVI extracted from the bands of TM data with LAI measured from the field in the Slash pine forest in northern Florida. Wang et al., (2005) also

reviewed NDVI and Enhanced Vegetation Index (EVI) with the data coming from three satellites, AVHRR, SPOT and MODIS using linear regression in the deciduous forest. Eckert, Kellenberger, & Lencinas, (2005) estimated forest LAI using mid-infrared Vegetation Index (MVI) and modified Normalized Difference Vegetation Index (mNDVI) using multilinear regression. This study conducted in a forest in Argentina and VIs derived from ASTER and Landsat TM+ data. Ehammer, Fritsch, Conrad, & Lamers, (2010) and Ali et al. (2015) indicated that selection of indices depended on the spectral characteristics of the sensor which in these cases data provided from RapidEye using linear regression. These studies were undertaken to show the importance of red edge band of RapidEye in cropland, in the first study NDVIrededge and Red Edge Ratio Index and in second study Soil adjusted red edge index (SARE) were selected as indices.

Since the necessity of using red edge band of RapidEye mentioned in different literature, in the present study we take advantage of using Red edge and other spectral bands to investigate the relationship between VIs and LAI measurements. The use of VIs as a single feature for estimating LAI is one of the suitable calculations as long as the correlation between measured LAI and extracted VI is strong, otherwise probably LAI estimation is not reliable. It is also should be taken into account that for obtaining an accurate LAI estimation, VIs are not the only source of input data, and spatial information such as texture could improve the accuracy of LAI estimation.

Several studies have shown the potential of high-resolution imageries for mapping LAI (Song & Dickinson, 2008; Zhou et al., 2014). The textural information extracted from these high-resolution image is an essential element for estimating and mapping canopy LAI (Pu & Gong, 2004). In most of the cases, texture information is a supplementary source of information for estimating LAI. For instance, Kraus, Schmidt, Dech, & Samimi (2009), investigated the use of spectral and spatial information of the high-resolution SPOT-4 and ASTER images for estimating LAI using regression analysis in the rainforest in east Africa . In another study, Zhou et al. (2014) used Quickbird satellite data to describe the relationship between extracted spectral and spatial information and estimate LAI. Three different techniques, simple linear and non-linear regression were applied for estimating LAI: VIs, textural information extraction and their combination using simple linear and non-linear regression in the semi-arid region of the Loess Plateau, China. Pu & Cheng (2015) also studied that the combination of VIs and extracted textural information from WorldView-2 sensor increases the accuracy of estimating LAI. They used Gray-level co-occurrence matrix (GLCM) for eight 2-order statistical TVs and finally linear regression analysis to predict and map Forest LAI. The study area was composed of floodplain wetlands associated with Cypress Creek.

Referring to the literature, the combination of both VIs and the textural information is an appropriate method for LAI estimation, and it results in high modelling accuracy. Then it proves that utilization of TVs besides the VIs is important using high-resolution multispectral image.

1.2.2. Machine learning based models

Development of multivariate models which are almost new to the field of estimating forest LAI, results in exploiting many predictor variables affecting LAI estimation and also using MLA. Different multivariate statistical models such as NN, Support Vector Machine and Random Forest (RF) have been involved in many of forestry applications and modelling LAI, which use extracted spectral and textural information from the image. Obviously, the optimal performance of MLA obtained when it is utilized in the right application with the correct and complete tuning. There are examples of previous studies that considered Machine Learning based models:

Bacour, Baret, Béal, Weiss, & Pavageau, (2006) used NN for estimating LAI in 6 biome classes: 1) grasses and cereal crops, 2) shrubs, 3) broadleaf crops, 4) savannas, 5) broadleaf forests, and 6) needle leaf forests

in different countries using data from MERIS and MODIS. In another study, Heiskanen, Rautiainen, Korhonen, Möttö, & Stenberg, (2011) investigated LAI by selecting homogeneous forest patches in Finland, using SPOT, HRVIR, and Landsat ETM+ satellite data. Two methods of NN and k-Nearest Neighbor Algorithm (kNN) were compared.

Regarding comparing two MLAs such as Random Forest with NN, Robust to overfitting, easy to use, and yielding better test error at the start of training time can be mentioned. However, NN is one of the models that is hard to be trained from a limited data, also, complex architecture optimization, low calculation robustness and enormous training time (Nitze, Schulthess, & Asche, 2012). There are some comparing studies between Random Forest algorithm and other MLA, for instance, Latifi & Koch (2012) compared two methods, Most Similar Neighbour (MSN) and Random Forest, for estimating forest inventory variables using airborne optical and laser scanner in forest sited southwestern, Germany.

In most recent studies, Random Forest was applied to assess and model the biomass and Carbon in the forests. Examples are as follow: Bunn (2005) modelled photosynthetic activity predict the trends to the year 2050 with the use of general circulation model predictors. Powell et al. (2010) used Landsat satellite imagery to quantify forest biomass using Random Forest Algorithm in a study area in Arizona and Minnesota. Other forest attributes such as, canopy height using MLA with the integration of Lidar and Landsat data (Hudak, Lefsky, Cohen, & Berterretche, 2002; Wulder & Seemann, 2003), succession stages using LIDAR data (Falkowski, Evans, Martinuzzi, Gessler, & Hudak, 2009) and tree crown damage using digital aerial photo and other geographical data (Thompson & Spies, 2009) also revealed good results using Random Forest model.

As a summary, it can be concluded that all the applications for estimating LAI tended to obtain an accurate prediction using spectral, spatial or the combination of these two methods. However, still there is not any invariant method for estimating forest LAI from satellite images, there are modelling approaches that produce high accuracies in most of the applications such as Random Forest. As it is mentioned, the use of Random Forest model is spread to other cases such as quantify forest biomass, monitoring carbon stock, and no study focused on using Random Forest to estimate and map forest LAI from high-resolution multispectral images.

1.3. Research problem

As the above literature review reveals, accurate estimation of LAI particularly, using high-resolution multispectral images is still a challenge in heterogeneous forests. This is because spectral data which is stored in each image pixel is composed of radiation of more than one tree species (López-Serrano, Landete-Castillejos, Martínez-Millán, & Cerro-Barja, 2000; Darvishzadeh, Skidmore, Schlerf, & Atzberger, 2008). It is not clear that how such heterogeneity affects the use of reflectance calculated from few spectral bands of multispectral image for LAI estimation.

The reflectance of the green understory is often generalized to the reflectance of upper level canopy of trees in remote sensing images. Moreover, The angular distribution of foliage elements and canopy structure affect the reflectance calculated from remote sensing images (Rhoul, Goulas, Daumard, Ounis, & Moya, 2014). It is not known how such generalization affect the estimation of LAI using spectral data that comes from a few number of broad bands of multispectral image. This issue need to be analysed while estimating LAI to verify the reliability of the calibrated model.

Using purely statistical method when few input data exist, makes LAI estimation more challenging. These methods often consider a parametric relationship between LAI and predictor variables such as VIs and TVs.

For example, many studies used linear methods to estimate LAI. However, calibration of parametric methods with few input data is not always accurate. Moreover, the parametric relationship between LAI and predictor variables can be site-specific. This means that a calibrated model for specific study area might not work for another study area.

1.3.1. Research objectives

This study aims to test a methodology for LAI estimation in a heterogeneous forest from high-resolution multispectral images. The spectral and spatial information obtained from the RapidEye image are utilized to retrieve LAI. The main objectives of this study are as follows:

1. To estimate LAI using NDVI, RVI and Red edge vegetation indices derived from the RapidEye image in a heterogeneous forest.
2. To evaluate the use of textural information extracted from the high-resolution images for LAI estimation in a heterogeneous forest.
3. To evaluate the use of MLA in combining textural and VIs information to estimate LAI.

1.3.2. Research questions

Answering the following research questions helps to achieve the sub-objectives as mentioned above:

1. Among vegetation indices studied, which one will provide a more accurate estimate (Highest R^2 and lowest RMSE) for estimating LAI?
2. In a comparison of using VIs, which degree of accuracy, textural information will offer for estimating LAI based on the high-resolution multispectral image?
3. How effective Random Forest algorithm estimate LAI by combining VIs and textural based information from a RapidEye image?

1.3.3. Research Hypothesis

1. Using Red edge VI will provide a more accurate estimate of LAI (regarding R^2 and RMSE) compared to other VIs.
2. LAI estimated using the textural information from the RapidEye image, has lower accuracy in comparison to the one estimated using VIs.
3. Using Random Forest algorithm makes a significant improvement in accuracy of estimated LAI compared to using VI and texture information alone.

In the next five chapters, we will explain the methodology which is used in this study for estimation of LAI in Bavarian Forest National Park, Germany. The outcomes are validated using the field measurement of LAI obtained parallel to the time of the image and are presented in the result section. Further, the discussion will elaborate and discuss the obtained results and finally in the last chapter the conclusion and recommendation for future works are presented.

*“The more that you read, the more things you will know,
the more that you learn the more places you’ll go.”*

*Dr. Seuss,
1904- 1991*

2. MATERIALS:

In the following chapter, the study area is described in terms of geographical characteristics, weather condition, and flora and fauna species. In the next section, field measurements method and satellite image specifications are described in details.

2.1. Study area

Bavarian Forest National Park (BFNP) is one of the important forests in Europe which located between (49 ° 3' 19" N, 13 ° 12' 9" E), on the border of Germany and Czech Republic (See figures 1 and 2). It covers the area of 24, 218 ha. The elevation of the area ranges from 600m above sea level with its lowest valley, to the ridges, with the highest elevation at 1,453m above sea level. The climate of the region is temperate, and the mean annual temperature varies between 6.5 °C in the valleys and 2 °C at higher elevations. The high annual precipitation is (1200-1800 mm) without a strong dry period, and usually, a considerable amount of precipitation occurs as snowfall (Ali, Darvishzadeh, Skidmore, & Duren, 2015). The snow cover remains for 7-8 months in higher elevations and 5-6 months in the valleys.

The natural forest ecosystem of BFNP varies with altitude: there are alluvial spruce forests e.g. Norway Spruce in the valleys, mixed Silver fir and Norway European beech mountain forests on the hillsides, and mountain spruce forest in the high areas. The dominant species in the forest are European beech (*Fagus sylvatica* L.), fir (*Abies alba*), and Norway spruce (*Picea abies*). In the mixed forests located on the mountain, other species such as Sycamore maple (*Acer pseudoplatanus* L). Mountain ash (*Sorbus aucuparia* L.), a Goat Willow (*Salix caprea*) also can be found (Heurich & Neufanger, 2005).

In the BFNP, 7,300 species have been identified which is composed of 3,850 animal species, 1,860 types of fungus, 490 species of moss, 340 lichens and 760 vascular plant species (“Bavarian Forest National Park,” 2015). BFNP is of exceptional important for the protection of variety of wildlife species such as lynx (*Lynx lynx*), capercaillie (*Tetrao urogallus*) and large ungulates for instance, red deer (*Cervus elaphus*) and roe deer (*Capreolus capreolus*) (Cailleret, Heurich, & Bugmann, 2014). Wildlife management of BFNP is instituted in several management zones such as a wildlife population control, a reduction of disturbances by restricting public accessibility to certain areas, or a reduction of winter feeding (Heurich, Baierl, Günther, & Sinner,

2011). Zoning of the forest is because of better protecting the BFNP from wildlife damages, preserving the natural diversity of the flora and fauna from human interference and activities.

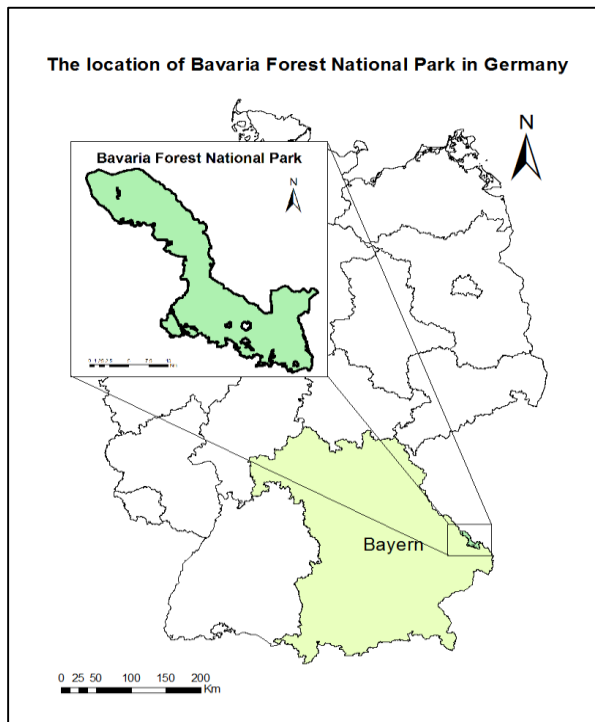


Figure 1. The location of the study area, Bavarian Forest National Park



Figure 2. Landscape of Bavarian Forest National Park
taken by Elnaz Neinavaz
(Ph.D. candidate of NRS department, ITC faculty)

2.2. Field measurement of LAI

A field campaign was conducted in BFNP by staff and Ph.D. students from Natural Resources Department, Faculty of ITC, University of Twente in August 2015. The area was stratified into forest and non-forest using the land cover map provided by the management of the park. Sample's coordinates were randomly generated in the forest stratum to select plots. A total of 37 plots of 30 m x 30 m were generated, and a GPS (Global Positioning System) was used to locate them in the field. In each sample plot, several biochemical and biophysical parameters including LAI were measured. In each plot, LAI was measured using the Plant Canopy Analyser LAI-2200 (LICOR Inc., Lincoln, NE, USA). For each sample plot, reference samples of above-canopy radiation were taken by measuring incoming radiation outside of the sample plot in an open spaces area. Five below-canopy samples were collected within each plot (See figures 3 and 4), from which the average LAI was calculated. LAI measured using LAI-2200 corresponds to plant area index (PAI), including the photosynthetic and non-photosynthetic components (Chen et al., 1997). However, in our study, the measurements corresponds to effective plant area index and in the following sections, these measurements are abbreviated as LAI.



Figure 3. Measuring LAI in the plot using LAI 2200
by Haidi Abdullah
(Ph.D. candidate of NRS department, ITC Faculty)

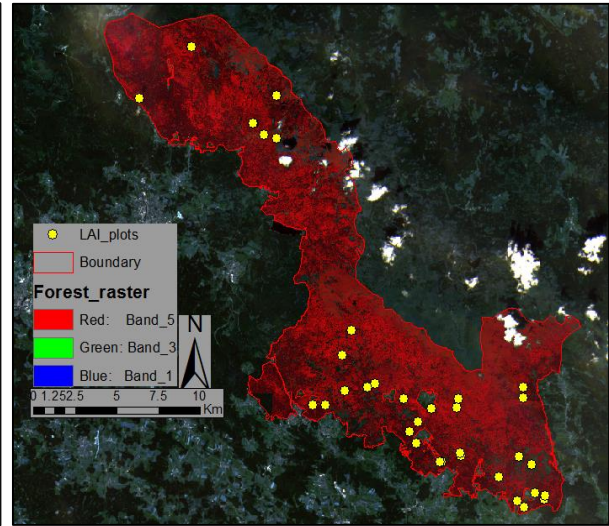


Figure 4. Distribution of the sample plots in Bavarian Forest National Park, August 2015.

2.3. RapidEye Imagery

The RapidEye constellation launched from August 2008 in Kazakhstan, has five identical equally calibrated satellite. It is expected to be active until at least 2017 (RapidEye, 2011b). The type of the focal plane is a multi-spectral push broom. RapidEye satellite flies in the orbit altitude of 630 km in sun-synchronous orbit, and it crosses the equator at 11:00 AM local time (Kramer, 2002). According to RapidEye imaging system, it has five spectral bands of Blue(440-510nm), Green(520-590nm), Red(630-685nm), Red edge(690-730nm), NIR(760-850nm). Ground sampling distance at Nadir is 6.5 meter, and orthorectified spatial resolution is 5 meter. The off-nadir temporal resolution of the sensor is daily and at nadir is 5.5 days. In terms of collection capabilities, the swath path covers of 77 km at nadir. The maximum geometry collection of RapidEye is up to 1 million km^2 per day, for each satellite. The specification of RapidEye satellite is presented in Table 1.

In this study, four high-resolution multispectral images of RapidEye 5 acquired to cover whole the area of the BFNP, which captured on 8th August 2015. Some of the advantages of using RapidEye image for this study are as follow: First, because the images used for this study captured concurrent at the time of field measurements the results are expected to match completely and because the data are quite new the output will be very up-to-date and valuable. Secondly, RapidEye image has the Red edge wavelength as a separate band which is sensitive to the vegetation reflectance, and it can help for better extraction of information from VIs. Lastly, as it mentioned above, RapidEye images have 5-meter spatial resolution, which is very appropriate for implementing texture analysis on the image.

Table 1. Summary of specifications of RapidEye sensor.
Obtained from RapidEye, (2011b)

Satellite Number	5
Life Span	7 Years
Orbital Altitude	630 km
Equatorial crossing time	11 a.m.
Sensor Type	Pushbroom
Spectral Range	Band 1: Blue (440-510 nm) Band 2: Green (520-590 nm) Band 3: Red (630-685 nm) Band 4: Red Edge (690-730 nm) Band 5: Near Infrared (760-850 nm)
Pixel size (Orthorectified)	5 m
GSD	6.5 m
Swath Width	77 km
Global Revisit Time	1 day
Dynamic Range	12 bit

“The whole purpose of education is to turn mirrors into windows”
Sydney J. Harris
1917-1986

3. METHOD

3.1. Image Pre-processing

Four separate images of RapidEye were attached to each other to have a complete raster of the BFNP using ArcGIS 10.3.1 and then, the shapefile of the forest boundary was used to extract the same extent of the boundary from the raster image. In between, Since the goal of the research is to estimate LAI in the forest, other land-uses such as agricultural fields, urban areas, roads, and water bodies were excluded from the image using Land-use shapefile of the region. So the rest of the map is consist of lying dead woods, coniferous forest, deciduous forest and their mixture.

Image pre-processing consists of correction of radiometric and geometric errors. Geometric errors are distortions between the actual image coordinate and the ideal image coordinate (Nasda, 1999). In this study, the RapidEye images were already geometrically corrected (by image provider). Radiometric errors are any unwanted errors in the image data which is due to the limitation of sensing, signal digitization, and data recording processes (Varshney & Arora, 2004). First radiometric source is internal errors such as calibration source and detector response, and the second is the external error such as atmospheric errors.

Since, first step for locating or identifying features in an image is to calculate the reflectance spectra of the features on the image. In this study, regarding radiometric corrections we converted pixel radiation to reflectance spectra. Reflectance is the radiation proportion of a surface which is striking to the radiation reflected from it (Shippert, 2013). The pixel values in the raw image were the (amount of) radiation coming from the objects on the ground. However, we needed the image reflectance to start processing and extracting VIs and TVs from the image. Converting radiation to reflectance help to eliminate the cloud or other atmospheric components that are effecting on the target object spectra (Shippert, 2013). Reflectance values extract quantitative information about the target object.

For each pixel we calculated its reflectance (values) using following formula (RapidEye, 2011a):

$$REF(i) = RAD(i) / (\pi * \sin^2(SunDist) / (EAI(i) * \cos(SolarZenith)))$$

Where:

REF: reflectance value

n: Number of the spectral bands

RAD: Radiance value

SunDist: Earth-Sun Distance at the day of acquisition in Astronomical Units (This is not a fixed value, and it varies between 0.983 289 8912 AU and 1.016 710 3335 AU. It has to be calculated for the image acquisition point in time.

EAI: Exo-Atmospheric Irradiance, EAI in Blue band is 1997.8 W/m²μm, in Green band is 1863.5 W/m²μm, in Red band is 1560.4 W/m²μm, in Red edge is 1395.0 W/m²μm and Near Infrared is 1124.4 W/m²μm.

SolarZenit: Solar Zenith angle in degrees (= 90° - sun elevation)

The calculation was done using 'Band.Math' tools of the ENVI software in which the formula parameters were reserved for RapidEye image.

The conversion of the coordinate system of the measured LAI plots to the image coordinate system was another pre-processing. This was because the coordinate systems and datum of these two data sources were not the same. The coordinate system of the satellite image was WGS-1984-UTM-Zone-33N and the coordinate system of measured LAI plots was the German coordinate of DHDN_3_Degree_Gauss_Zone_4. We used ArcGIS 10.3.1 Software to convert the coordinate system of plots to the coordinate system of the satellite image.

The average reflectance value of the plots was calculated on the image using a window size of 5 by 5 pixels. For this, the NRS module developed by Willem Nieuwenhuis in Natural Resource department, Faculty of ITC were used. The module offers a set of functions and allows users to extract an aggregated spectral profile with a user specified window size. It can read the coordinate of point features from either a text table or a shape file to get the spectral profile from the image. The resulted profiles, for all locations, are then written in a text table. The header line of each column contains the coordinates of that location.

3.2. Vegetation Indices generation

We used the spectral bands of RapidEye images to calculate several indices that have potential for LAI estimation (Broge & Leblanc, 2001; Schlerf et al., 2005). In Table 2 the investigated vegetation indices are presented. The widely used vegetation indices were directly computed from the reflectance data and their relationships with LAI were analysed using the Pearson correlation coefficient (*r*). Taking the ratio of two Red and NIR bands reduces the noise in the reflectance (Chen, 1996) (e.g. in NDVI, RVI, etc.) and ARVI, SAVI, and PVI are three indices which are using for decreasing the effects of background and atmospheric effects in LAI mapping (Colombo, Bellenger, Fasolinic, Marino, 2003).

Moreover, as it is shown in table 2, PVI required site specific soil line slopes (*a*) and intercepts (*b*). As no soil spectral data was available, soil line parameters were constant to arbitrary values (*a*=0.9; *b*=0.1) (Schlerf et al., 2005). For the SAVI, *L* is a canopy background adjustment factor set as 0.5 (Soudani et al., 2006). For EVI, *G*, *C*₁, *C*₂, *L* are coefficients to correct scattering of aerosol, absorption, and background brightness, they were accordingly set to 2.5, 6, 7.5, 1, respectively, and also *γ* in ARVI equation equals to 1.0 based on the study by Soudani et al., (2006).

Table 2. Selected Vegetation Indices for LAI estimation

Vegetation indices	Reference
$NDVI = \frac{NIR - R}{NIR + R}$	(Rouse et al., 1974)
$NDVI_{rededge} = \frac{Rededge - R}{Rededge + R}$	
$RVI = \frac{NIR}{R}$	
$RRI1 = \frac{NIR}{Rededge}$	
$RRI2 = \frac{Rededge}{R}$	
$EVI = \frac{NIR - R}{NIR + C_1 * R - C_2 * B + L} * G$	(Huete et al., 1994)
$SAVI = \frac{NIR - R}{NIR + R + L} (1 + L)$	
$PVI = \frac{NIR - aR - c}{\sqrt{a^2 + 1}}$	
$ARVI = \frac{NIR - rb}{NIR + rb}; rb = R - \gamma(B - R)$	(Kaufman & Tanre, 1992)
$REDNDVI = \frac{NIR - Rededge}{NIR * Rededge}$	
$GRNDVI = \frac{Rededge - Green}{Rededge * Green}$	
$TVI = \sqrt{NDVI + 0.5}$	

3.3. Texture Variables generation

Textural content information highly depends on the spatial resolution of the used image and also depends on the direction, shape and distribution of the target object (Kayitakire et al., 2006). Amongst all methods for extracting texture information, GLCM proved to be an efficient method for classification and analysing surface texture of an image (Pathak & Barooah, 2013). GLCM characterizes texture by the co-occurrence of the gray tones, which in course resolution texture distribution of tonal variation varies very slightly while the variation of fine texture is so rapid within a distance (Haralick, 1979). Thus, we used the 2nd-order texture measure or GLCM to generate texture predictor variables. GLCM is calculated from the spatial dependency gray level co-occurrence matrix which describes the probability of each pair of pixel values co-occurring in a defined direction and distance (Haralick, Shanmugam, & Dinstein, 1973; Pu & Cheng, 2015). The GLCM operation will be easier to explain using figure bellow:

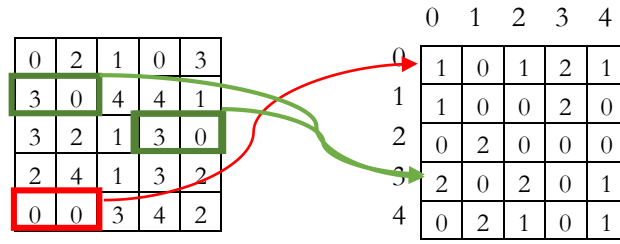


Figure 5. Gray Level Co-occurrence Matrix operation

GLCM can read the value of pair adjusted pixels and find all the similar sequence within a fixed window size, then record it in the output. For example, in figure 5 the element (0,0) contains 1 because there is only one instance sequence in the original window. Or the element (3,0) indicates 2 similar sequences found in the original window.

Four known GLCM parameters that control the variation of the output of GLCM (Albregtsen, 2008) were considered for extracting TVs. They included: (1) The size of the moving window that can vary from the small size (3 by 3 pixels) to larger size (13 by 13 pixels), it is usually selected based on the plot size on the ground and pixel size on the image; (2) The parameter displacement of the window (δ) which ranges from 1,2 to 10 (Gadkari, 2004), the larger displacement results in the GLCM that does not contain detailed textural information. (3) Direction of the pair pixels (θ) is; for instance angle of 0° , 45° , 90° , 135° , 180° , 225° , 270° , 315° that GLCM can read (Gadkari, 2004). (4) The gray level of the matrix (G), which is determined by the maximum gray level of a pixel, greater gray level values show more accuracy with a higher cost of texture computation (Albregtsen, 2008). In figure 5, the size of window size, δ , θ and G are respectively 5 by 5 pixels, 1, 0° and 5.

Four main directions (0° , 45° , 90° , 135°) and three window sizes: small (3by3 pixels), medium (5by5 pixels), and large (7by7 pixels), have been applied to all bands of the RapidEye image. In this case, it was not possible to apply bigger window sizes of 9 or 11, as it would have exceeded the size of the measured sample plots. Referring to the code maintainer of GLCM in R, the displacement and direction were encoded to the same parameter in GLCM package; meaning that by testing different directions for a window size, displacement was tested as well. In this study, first we examined the correlation between TVs and LAI based on different window sizes with constant direction 0° , and then we tested the correlation between TVs and LAI based on different directions with a fixed window size

Haralick et al. (1973) introduced 14 TVs that can be derived from GLCM. Among those there are seven variables (Angular Second Moment(ASM), contrast, variance, homogeneity, correlation, entropy, and mean) are the most relevant to remote sensing image analysis (Baraldi & Parmiggiani, 1995). Homogeneity, contrast, and entropy are variables related to the specific textural characteristics of the image, while, dissimilarity, mean and standard deviation characterize the complexity and the nature of gray level transitions defined in the co-occurrence matrix (Wulder et al., 1998). Table 3. lists the examined TVs that used for extraction of texture information.

Table 3. 2nd-order Texture Variables and equations

Texture feature	Formula
Mean	$\sum_{ij=0}^{N-1} ip(i,j)$
Variance	$\sum_{ij=0}^{N-1} p(i,j) (i - \mu_i)^2$
Homogeneity	$\sum_{ij=0}^{N-1} \frac{p(i,j)}{1 + (i,j)^2}$
Contrast	$\sum_{ij=0}^{N-1} p(i,j)(i - j)^2$
Dissimilarity	$\sum_{ij=0}^{N-1} p(i,j) i - j $
Entropy	$\sum_{ij=0}^{N-1} p(i,j)(-\ln p(i,j))$
Angular Second Moment (Energy)	$\sum_{ij=0}^{N-1} p(i,j)^2$

Where :

Each element in (i, j) represents the number of times that the pixel with value i occurred adjusted to a pixel with value j . μ_i in Variance equation is the GLCM-Mean (Baraldi & Parmiggiani, 1995).

The functionality of TVs explained by the Gadkari (2004): Variance is the heterogeneity measurement which is correlated with standard deviation and when the gray level differ from their mean, Variance increases. Homogeneity measures the spatial frequency of the image, and a larger value will be obtained for smaller gray tone difference in pair pixels. Contrast, measures the highest and the lowest value of the continuous set of pixels, thus low values of contrast, result in a low spatial frequency of the image. Entropy measures the disorder or complexity of the image, therefore, the larger values of Entropy, the more texturally complexity in the image. Angular Second Moment also measures the textural uniformity. When the gray level distribution has a periodic form the high energy values will be obtained. Dissimilarity same as Contrast measures the spatial frequency of the image but with a linear relationship (Ozdemir, Norton, Ozkan, Mert, & Senturk, 2008).

An innovative experiment of GLCM was done by calculating TVs for the raster layer produced by VI to test if it can improve the accuracy of LAI estimation or not. Moreover, the ratios of Mean and Variance from the combination of different bands and different window sizes were also calculated. For example, for each band two different window sizes and for each window size four different probabilities of mean-mean, mean-variance, variance-variance and variance-mean existed that we calculated related ratios.

So, the final output of the texture was made up of 63 TVs (extracted from two bands and one raster layer when for every band and layer there are three window sizes and for each window size there are seven TVs (3*3*7)) and 16 texture ratios (extracted from two bands when for each band there are two window sizes and for each window size there are four TVs(2*2*4)), in total 79 texture predictor variables.

3.4. Linear Regression Analysis

Regression is an empirical approach to analyse and model the relationship between dependent and independent variables (Cohen et al., 2003). In this study, the linear regression is used to analyse the correlation between the measured LAI and the extracted predictor variables from VIs and GLCM. To show the accuracy of the linear models, three statistical measures: Pearson correlation coefficient (r), Coefficient of determination (R^2) and Root Mean Squared Error (RMSE) are utilized. The strength and direction between the two variables (e.g. VI and LAI) are defined by Pearson correlation coefficient (r); R^2 shows how close is the data with the fitted regression model, and RMSE is the difference between observed and predicted values by a predictor or estimator (Chai & Draxler, 2014).

3.5. Random Forest modeling

Random Forest (RF) a machine learning method was used to calibrate and model LAI estimation. In particular, RF uses an ensemble of independent decision trees (i.e. tree-like model or graph of decisions; (Pal, 2005)) for classification and regression (Breiman, 2001). This method has potentials that make it useful for estimation of response variable when there is a large number of predictor variables (Díaz-Urriarte & Alvarez de Andrés, 2006). Moreover, RF has a predictive performance even when most of the predictor variables are not significantly correlated with response variable (Horning, 2010).

RF has no assumption of distribution of input data, either they have a linear or non-linear relationship as well as high order interactions effect between predictor variables (Strobl, Boulesteix, Zeileis, & Hothorn, 2007). This is because RF builds and averages a large number of decorrelated decision trees to predict the response variable (Hastie, Tibshirani, & Jerome, 2009). As the number of trees in RF becomes large the generalization error (i.e., the accuracy of an algorithm to predict outcome values) reaches a constant and minimum level (Breiman, 2001). Thus, RF does not need complex tuning and often leads higher prediction accuracy than other modeling methods.

RF modeling method can rank predictor variables based on their effect and importance in response variable estimation. The importance of each predictor variable is obtained by randomly permuting the variables values and the difference between modified and original predictor variable divided by standard error (Cutler et al., 2007).

The extracted VIs and GLCM variables as well as (bands) reflectances of the image (see subsection 3.2 and 3.3) were used as 96 predictor variables to train regression forest for LAI estimation. The measured LAIs were introduced as a response variable. We decided to initialize three necessary parameters of RF into recommended ones in the literature. The first parameter was the number of considered predictor variables in each split of the decision trees or *mtry* which was set to one-third of the number of predictor variables (Liaw & Wiener, 2002). The second parameter was the number of trees in RF, which was set to 1000 where the generalization error is minimized, and accuracy is improved (Breiman, 2001). The third parameter called Node size was the minimum number of predictor variables at the end of the tree, and it set to 1 which results in the largest possible trees (Roli, Kittler & Windeatt, 2004).

RF model was trained in R software using “RandomForest” package (Liaw & Wiener, 2015). We used predictor variables correspond to 37 ground observations to train RF. To find most important predictor variables we repeated the training and calculated the number of times that each predictor variables showed up in top-10. We repeated between 10 and 100 times in steps of 10 and chose the number of times that the top-10 predictor variables are stated not going to change afterward.

3.6. Cross-validation

The Cross-validation (CV) validates the result of model outputs with an independent dataset (Schneider, 1997). There are well-known types of CV such as K-fold and Leave-One-Out (LOO) method. If the number of folds of k-fold is equal to sample size, then the CV type will be LOO. Since the number of our samples (i.e., the measured LAI) were relatively small (37) we use the LOO method. In this method, each sample is estimated by the remaining samples, one sample point is taken out from the dataset as the test set and with the remaining samples, a model (here linear regression) is created. Then the model is used to predict the sample that was left out. Then RMSE is reported from the calculation of the difference between estimated values based on the models and measured values.

LOO is done for all the 37 measured LAI base on VIs regression model. Every time the constructed model is trained on 36 measured LAI as the training set and one sample as the test set. Then the trained model was tested on one sample left out of the test set in order to retrieve LAI for it. The same procedure is applied to the second fold, and the subsequent 35 folds one by one. Then the Cross validated RMSE and cross validated R^2 of the model can be acquired between the estimated LAI (obtained for each and every measured point based on the model) and actual measured LAI. Similarly, the same procedure is also applied to the most correlated TV with measured LAI.

It is normal to train RF with 2/3 of data and compare the predicted and measured LAI values of remaining 1/3 data to evaluate model accuracy. However, as above-mentioned, the number of samples data is relatively small. Thus, we used LOO method to evaluate the accuracy of RF model. This also could help to compare the accuracy of RF with explored linear models in subsection 4.2 and 4.3. RMSEs were calculated for each of 37 samples, and the average of RMSEs reported as the accuracy of RF model (see section 4.4).

3.7. LAI mapping

By comparing the accuracy of the created models, we used the most accurate one (in terms of high R^2 and low RMSE) to map LAI in BFNP. For this purpose raster of predictor variables were generated using RapidEye image. The calibrated model was applied to each pixel values of raster associated with predictor variables. The map of LAI was explored and discussed considering available land-cover data. In particular “raster.predict” R built in function was used to input raster of predictor variables and to output raster of LAI.

3.8. Flowchart

Figure 5 presents an overview of the methodology used in this study, starting with pre-processing of the high-resolution multispectral images of RapidEye, and extraction of predictor variables. Accuracy assessment using linear models and cross-validation and finally the use the most accurate model for estimating LAI and generation of final LAI map.

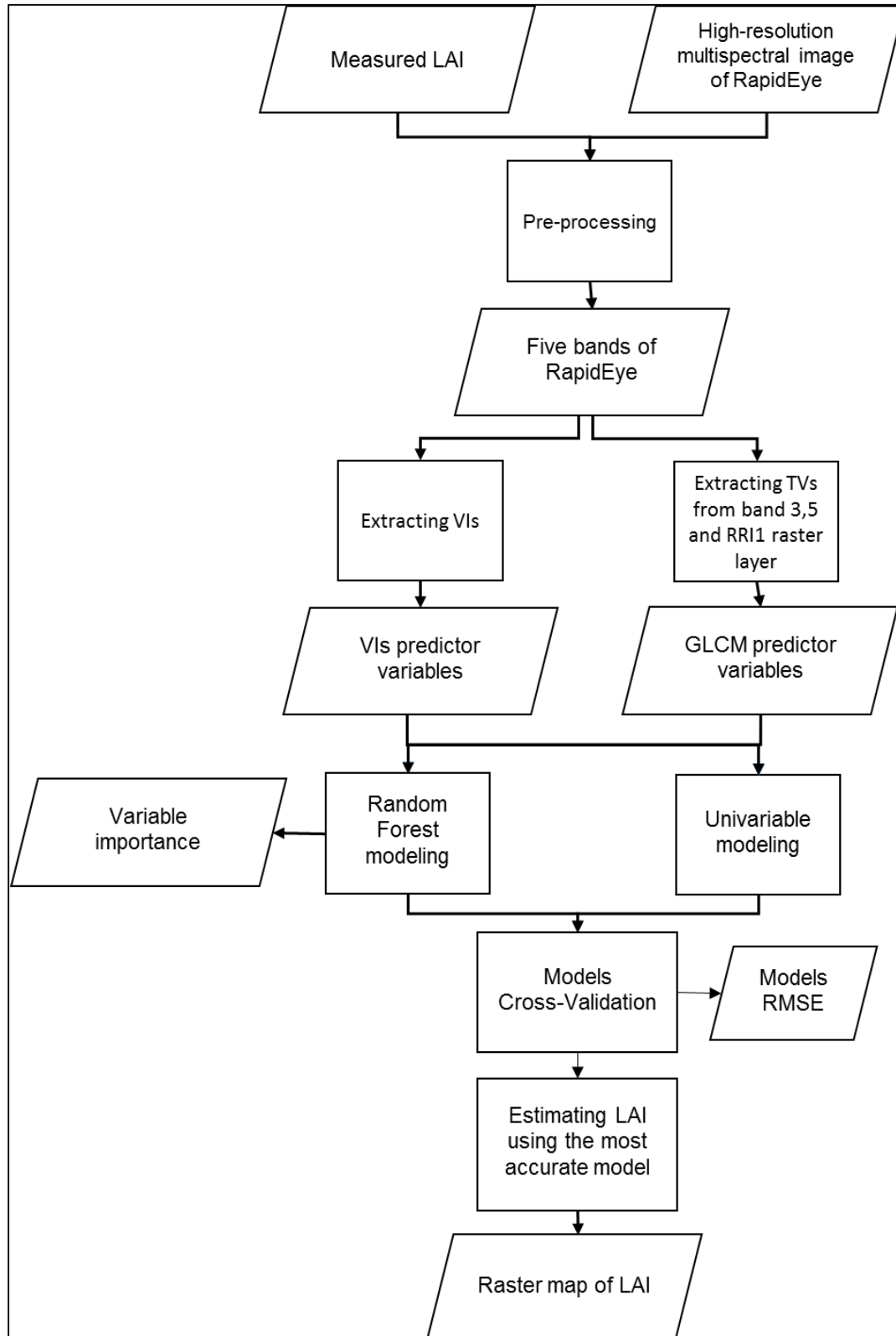


Figure 6. The methodology using multispectral satellite images to model and map LAI.

“Education is the most powerful weapon which you can use to change the world.”

Nelson Mandela

1918-2013

4. RESULTS

Spectral and spatial data, as well as their combinations, were used to estimate LAI in Bavarian Forest National Park using RapidEye images. The assortment of the data and results of the study are presented in this chapter in the following sequence: First, characteristics of the collected field data will be discussed. Next, the results of extracted spectral information and their correlation with the measured LAI using VIs will be discussed. Then, the results of all TVs extracted from texture analysis of the image and the correlation with the measured LAI will be presented followed by the results of RF. Finally, the RMSE and R^2 obtained from different models for estimating LAI in BFNP are compared in details.

4.1. The characteristics of the measured LAI in Bavarian Forest National Park

Summary statistics of the measured LAI are presented in Table 5. As can be observed in the table, the range of the measured LAI is relatively large indicating that the measured data covers relatively a wide variety of low and high LAI values in the forest. The measured LAI had a maximum of 5.86 and a minimum value of 0.50. The collected LAI had a coefficient of variation (the ratio of standard deviation and mean) of 0.44 which shows the frequency distribution of measured LAI in our dataset.

Table 4. Summary statistics of the measured LAI in Bavarian Forest National Park

Mean	Standard Deviation	Range	Coefficient of variation	Minimum	Maximum	Count
3.35	1.46	5.35	0.44	0.50	5.86	37

4.2. Correlation between Vegetation Indices and LAI

In this study, selection of the VIs was influenced by three factors (i) indices which have shown high a correlation with LAI in previous studies, (ii) the limited number of spectral bands in RapidEye and (iii) the existence of the red edge band. The mean spectral reflectance for each plot was used to calculate the VIs. Next, the correlation between the calculated VI and the measured LAI were investigated. Figure 7 present the correlation between measured LAI and studied VIs. As can be observed from this figure, the ratio index which composed of NIR and the red-edge bands RRI1 indicated the highest correlation with the measured LAI ($R=0.58$). Other indices including NDVI, TVI, RVI and EVI had returned a lower correlation ($R=0.5$). The lowest correlation value was for NDVI index calculated using the green band i.e. GRNDVI with R

equal to 0.07. In general indices which contain the NIR band, had higher correlation with the measured LAI compared to the ones that did not use NIR in their calculation. Soil based indices also did not demonstrate a high correlation with LAI. Due to unavailability of soil spectral data, the soil parameters were fixed to constant values and, therefore, could have affected the calculation of these indices.

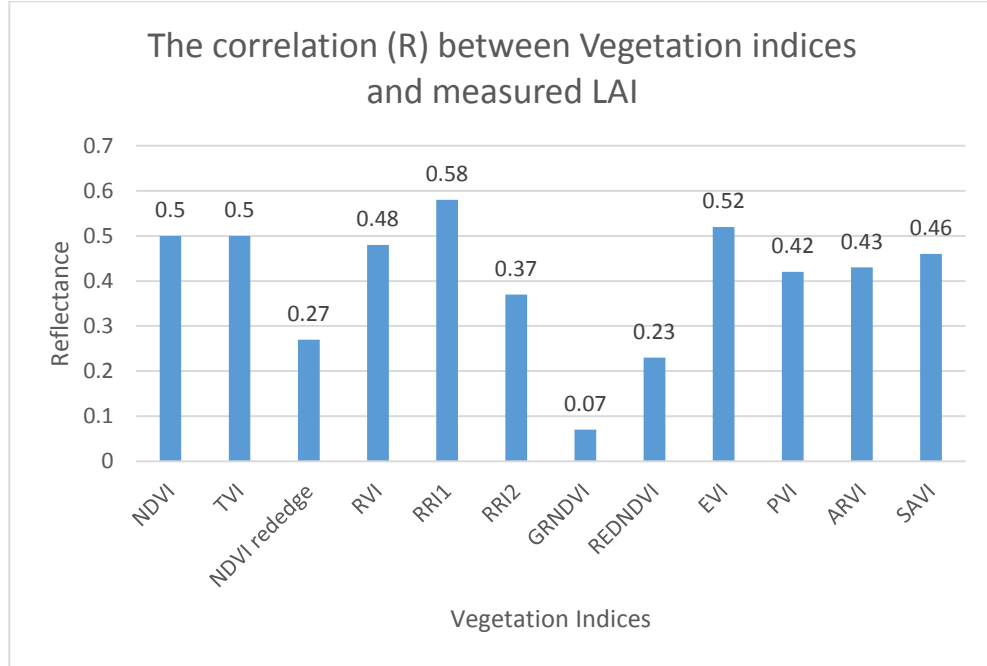


Figure 7. The correlation (R) between Vegetation Indices and measured LAI

In the next step of the analysis, for the best performing VIs, cross-validated R^2 and RMSE were computed using linear regression. As it can be observed from figure 7 the RRI1 index had the highest cross-validated R^2 and the lowest RMSE. Figure 8 illustrates the relationship between measured and estimated LAI for RRI1 index using the linear regression model. It seems that for many sample plots the LAI is mainly underestimated.

For each point in figure 8, there is a predicted LAI value from the linear model. To be more visually clear, we drew a $X=Y$ line in the graph to determine how predicted values were calculated by the model. About 43.24% of the plots are overestimated, and 56.76% are underestimated based on the model. There are also three plots that do not follow the linear model compared to other plots. These plots are located far from the regression line.

The first hypothesis of the research is accepted, as we could demonstrate that VI which made use of Red edge band of RapidEye image could estimate LAI with the Lower cross-validated RMSE and higher R^2 compare to other VIs (see section 1.3.3).

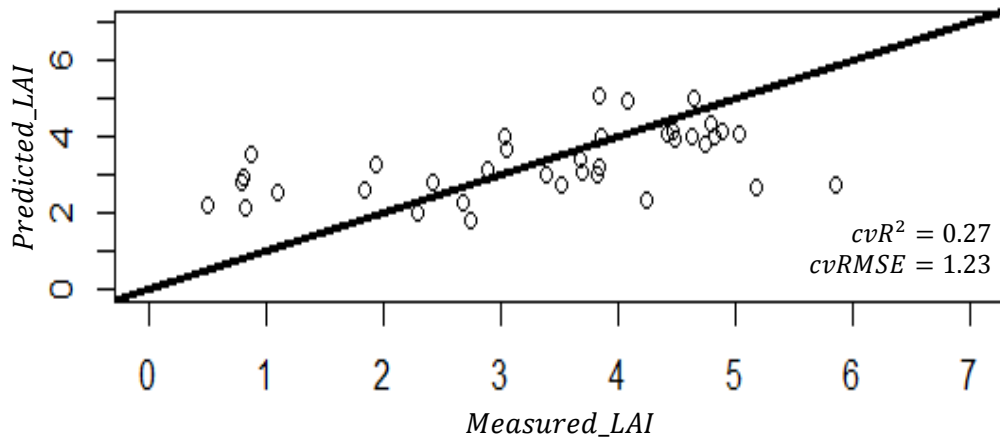


Figure 8. Plot of cross-correlation between the measured and estimated cross-validated LAI from RRI1 vegetation index

4.3. Correlation between Texture Variables and LAI

To analyse how GLCM texture analysis performed in estimating LAI, the correlation between TVs and measured LAI have been investigated. Seven TVs were calculated based on the formula of GLCM for all the bands and correlated with the measured LAI, but only two Red and NIR bands returned the best correlation with LAI, and then we continued further calculations with these two bands. The use of different direction with fixed window sizes (see section 3.3) proved that there are no higher correlations than fixed direction with different window sizes. Therefore, we excluded the different calculations of directions from the study. Then the final applied window sizes to each band and raster layer are as shown in the table below:

Table 5. Final applied window sizes of GLCM for two bands and raster layer

Layer	Window size 0 ° degree	Abbreviation
3(Red)	3 by 3	GLCM 33
3(Red)	5 by 5	GLCM 35
3(Red)	7 by 7	GLCM 37
5(NIR)	3 by 3	GLCM 53
5(NIR)	5 by 5	GLCM 55
5(NIR)	7 by 7	GLCM 57
RRI1	3 by 3	GLCM-RRI1-3
RRI1	5 by 5	GLCM-RRI1-5
RRI1	7 by 7	GLCM-RRI1-7

The experiment of extracting TVs from RRI1 led to a good correlation with the window size of 3 and 5. The results of the correlation analysis of GLCM band 3 and band 5, as well as raster RRI1, have been demonstrated in the figures 9, 10 and 11. The most significant correlation belongs to GLCM-mean and GLCM-variance TVs calculation, especially when they are extracted from the RRI1 raster layer. Moreover, the correlation between GLCM-mean for RRI1 and the measured LAI for window sizes 5 by 5 and 7 by 7 slightly descended in order. In the second place, GLCM-entropy and GLCM-second-moment have the highest correlation, but the R values do not exceed more than 0.42 in any of them. The minimum correlations between all the TVs and LAI are seen in GLCM-homogeneity, GLCM-contrast and GLCM-dissimilarity TVs for all the bands and window sizes.

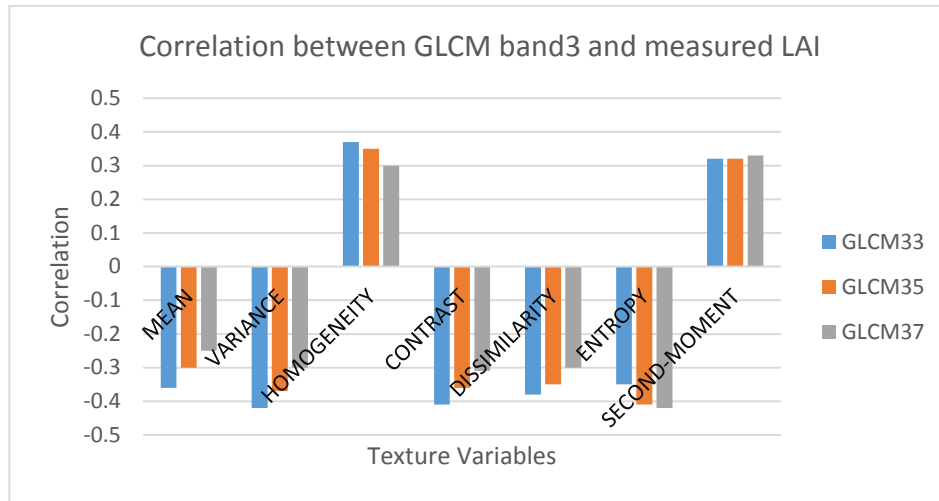


Figure 9. The correlation between Textural Variables of GLCM band3 and the measured LAI

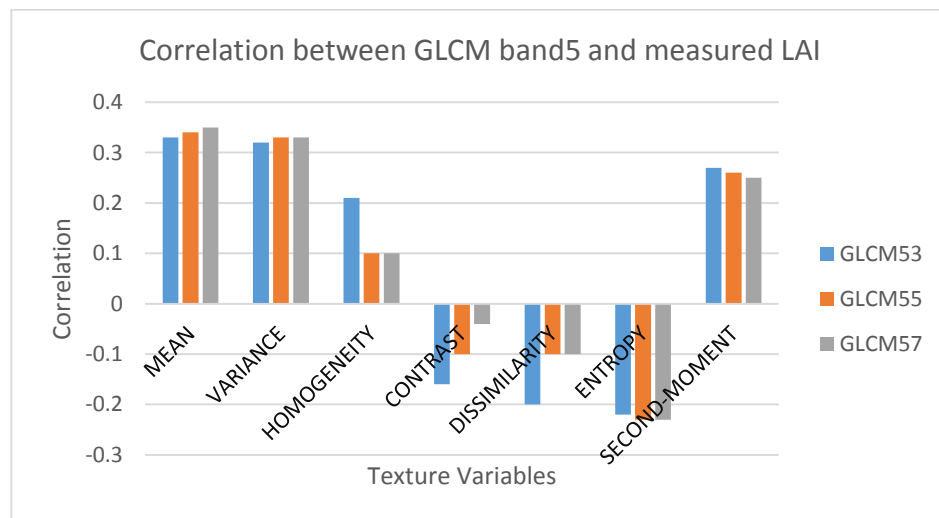


Figure 10. The correlation between Textural Variables of GLCM band 5 and the measured LAI

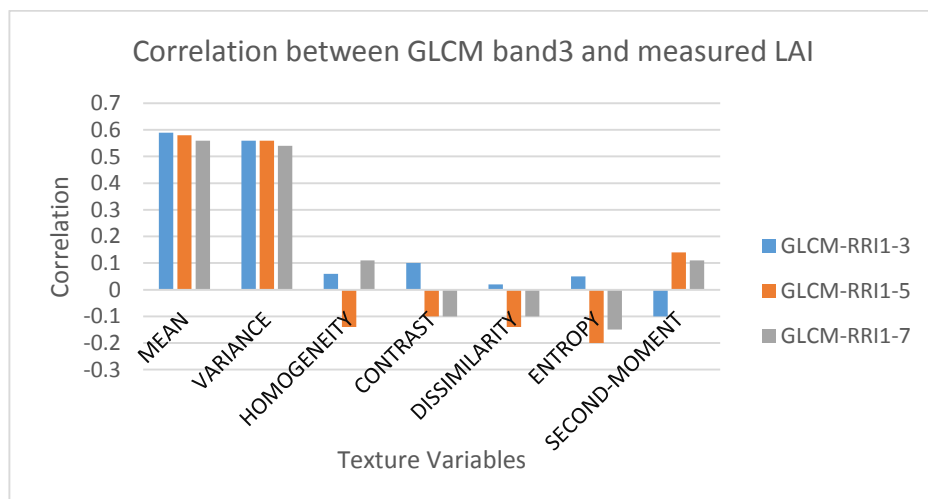


Figure 11. The correlation between Textural Variables of GLCM RRI1 and the measured LAI

Table 6 demonstrates the correlation between the measured LAI and the ratios of TVs from Red and NIR band with window sizes 3 and 5. As can be realised from the table, the highest correlation does not exceed more than 0.53 for the ratio 353mm and 353vv. The lowest correlation (0.36) belongs to the ratio 533mv.

Table 6. The correlation between combination of band and window sizes and measured LAI. The first and second digits show the band numbers and the third digit belongs to the window size; “m” stands for GLCM-mean and “v” stands for GLCM-variance, respectively.

Ratio (combination)	353 mm	353 mv	353 vm	353 vv	533 mm	533 mv	533 vm	533 vv	355 mm	355 mv	355 vm	355 vv	535 mm	535 mv	535 vm	535 vv
R (correlation)	-0.53	-0.46	-0.51	-0.53	0.48	0.36	0.42	0.41	-0.51	-0.46	-0.49	-0.52	0.47	0.37	0.43	0.42

As it can be observed from figure 11, the most correlated TV is the GLCM-mean of pair pixels in RRI1 raster with the window size of 3 by 3 pixels. Consequently, cross-validated R^2 and RMSE were computed for this TV using linear regression. Results are presented in figure 12 as can be observed from the figure R^2 and RMSE were respectively 0.28 and 1.23 which interestingly has a similar performance as using VI RRI1 linear model. Again $X=Y$ line was drawn in the graph to determine how predicted values were calculated by the model. Further analysis of the predicted LAI values revealed that 43.24% of the plots are over-extrapolated, 48.65% are underestimated and 8.11% estimated perfectly.

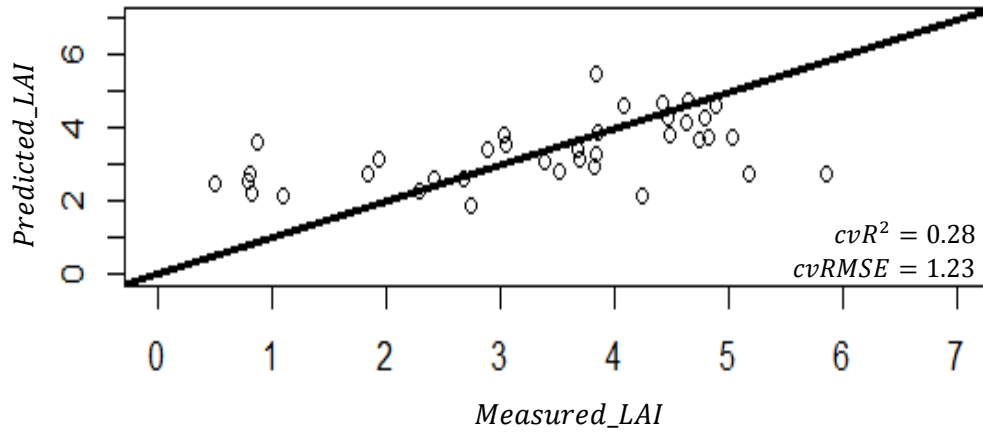


Figure 12. Plot of cross-correlation between measured and estimated cross-correlated LAI from GLCM RRI1-mean

Covering the second research question, there is a stronger correlation between TVs and LAI compared to VIs and LAI. Therefore, the second hypothesis is also accepted (see section 1.3.3).

4.4. Random Forest

Random Forest model conducted for LAI estimation in BFNP using 96 predictor variables extracted VIs and TVs from the image. Because of the number of VIs and TVs, their importance for estimating LAI was found using “variable importance” of Random Forest in R. Totally, 17 predictor variables are ranked as important predictor variables, however, 10 top ranked important predictor variables are illustrated in the table 7 based on their modes. This table also shows the number of repetitions of each predictor variables in total 50 iterations. The number of repetitions of last two predictor variables considerably drops to 29 and 25 times. This means that for other 7 important predictor variables of RF model that are not shown in the table below the number of repetitions are less than 25 repetitions.

Table 7. 10 top-ranked important predictor variables obtained from Random Forest

No	Ranked important predictor variables for estimating LAI in RF model	Repetition time in 50 repetitions
1	$VI - RRI1$	50
2	$GLCM_{RRI13} - Mean$	50
3	$GLCM_{RRI15} - Mean$	50
4	$GLCM_{RRI17} - Mean$	50
5	$GLCM_{RRI13} - Variance$	50
6	$GLCM_{RRI15} - Variance$	50
7	$GLCM_{RRI17} - Variance$	50
8	$GLCM57 - Second\ moment$	50
9	$VI - EVI$	29
10	$GLCM_{RRI15} - Second\ moment$	25

As mentioned in section 3.6, to have comparable results leave one out cross validation was performed for accuracy assessment of the RF model. Figure 13 shows the RMSE of the 37 plots of measured LAI. In this figure, the highest RMSE belongs to the plot number 4 and then plot number 31.

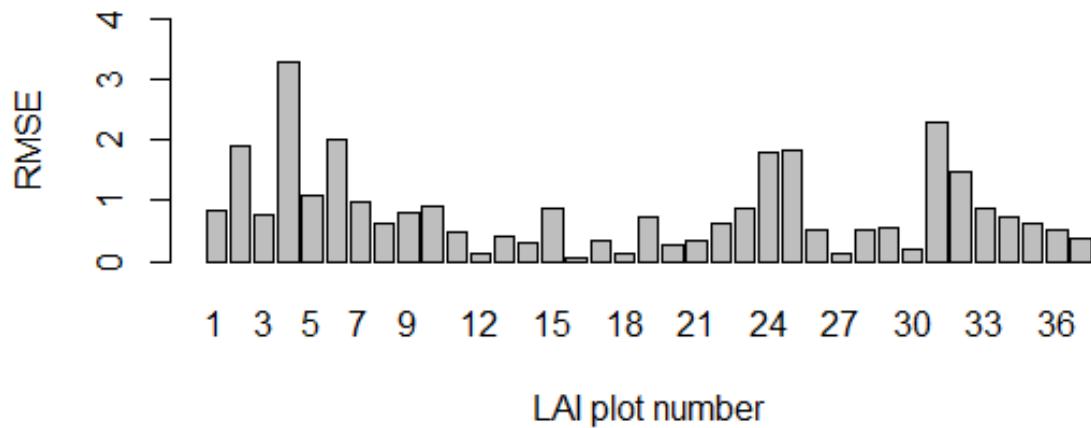


Figure 13. Leave One Out error for each plot by Random Forest model

In figure 14, cross-validated R^2 and RMSE are also originated from the relation between estimated LAI from RF model and the measured LAI. Referring to the graph, RF model has the best predictive performance with the cross-validated R^2 of 0.43 and cross-validated RMSE of 0.85. As it can be seen, the predicted values of LAI are so close to the optimal values which are close to $X=Y$ line.

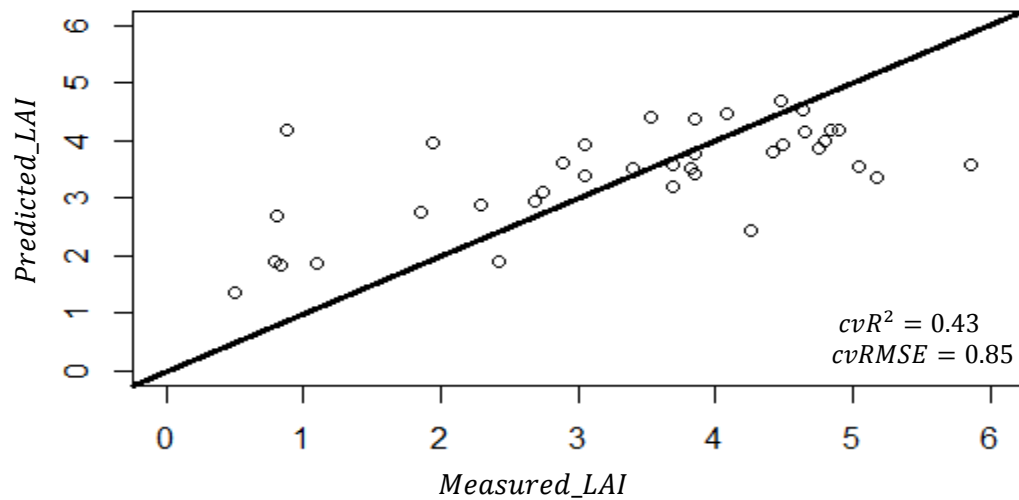


Figure 14. Plot of cross-correlation between the measured and estimated cross-correlated LAI using Random Forest model.

The third hypothesis and question (see section 1.3.3) is respectively accepted and answered, using Random Forest model yield a more accurate LAI estimation compare to two linear models of VIs and GLCM.

It should be mentioned that, we tested all the models after removing three measured LAI which did not regularly follow the prediction models (Plot number 4, 25, 31). Although removing them resulted in higher correlations for all the models, we decided to make use of all the measured LAIs.

Figure 15 demonstrates the final map of LAI in BFNP that was obtained through RF model using all 96 predictor variables and their raster layers. In the map of LAI using RF model, the statistics of the LAI values were very close to the real values in the measured plots. The total range of values in LAI data covers from 0.51 to 5.86 and the range in the map is from 0.94 to 4.5. Overlaying the land-cover map of the forest on the map of LAI, it was observed that the areas where to have smaller values of LAI belonged to the dead lying wood or coniferous stands in all young, medium or mature stage in the south-east and some areas in the northern part. Higher values of LAI belong to the vast areas covered by deciduous stands in the central areas in north and south part of BFNP.

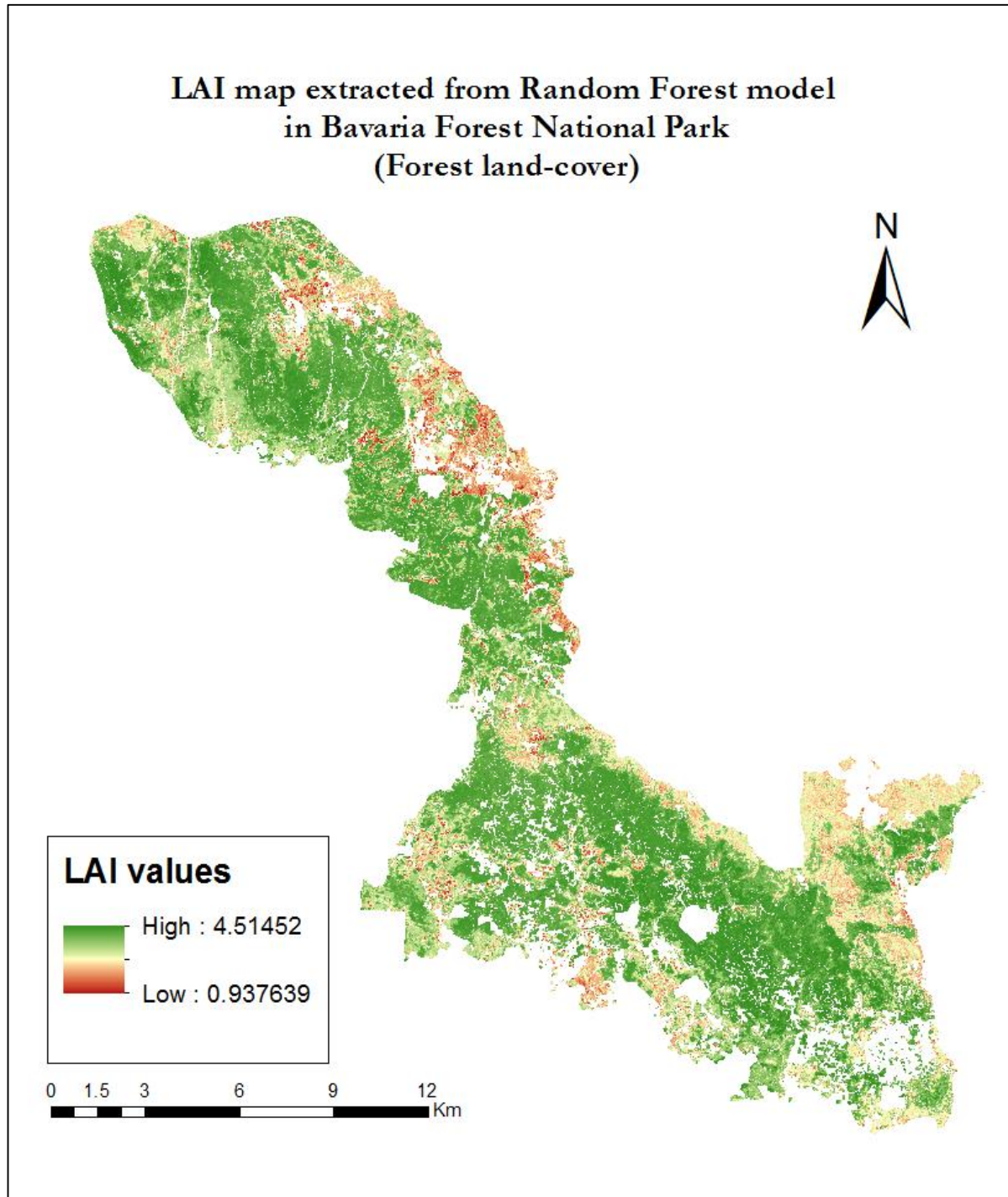


Figure 15. LAI map of Bavarian Forest National Park created from Random Forest model

“You have to go wholeheartedly into anything in order to achieve anything worth having.”
Frank Lloyd Wright,
1867- 1959

5. DISCUSSION

In the following chapter, the final results obtained in this study are discussed:

VIs were extracted from the bands of the RapidEye image and were correlated to the measured LAI using linear regression model. Low sensitivity of NIR and red edge band to the water vapour in the atmosphere and soil reflectance resulted in best correlation between RRI1 (ratio of NIR and red edge band) and measured LAI. In general, the correlation between VIs and the measured LAI for all the VIs were not significantly high, where for RRI1, R^2 was 0.27. In the linear model of RRI1, two measured LAIs experienced a significant difference when they were estimated. For example, plot number 31 with the measured LAI value of 5.86 (maximum LAI) was estimated 2.75 and LAI value of 0.5 (minimum LAI) was estimated 2.25. The possible explanation for this might be the understory cover which has not been considered in this study, but could have caused a high reflection in plots with low LAI values (above canopy). The results obtained in this study are lower than those obtained by Asam et al., (2013) which derived the R^2 values of 0.57 to 0.85 using VIs extracted from RapidEye image. This can be explained probably by the influence of atmospheric errors which were not corrected in this study (see also 5.1).

Besides VIs, GLCM-mean and GLCM-variance TVs derived from RRI1 raster also were ranked as very important predictor variables in variable importance of the RF model for estimating LAI. However, there were weak correlations between GLCM TVs and the measured LAI, it is rather clear that TVs extracted from RapidEye image had the potential to improve the accuracy of LAI map. Wulder, LeDrew, Franklin, & Lavigne (1998) also demonstrated the importance of using TVs from airborne CASI imagery for LAI estimation. However, identification of proper TVs, suitable window sizes and image bands seemed to be very essential (Chen, Stow, & Gong, 2004). The innovation of analysing TVs from raster produced by RRI1 as a VI fulfilled in this study and it was identified to play very important role for estimating LAI. Red and NIR bands were suitable for texture analysis using the available data, with the used window sizes 3, 5 and 7. Texture analysis will differ with the characteristics of the study area and the type of image, and then future users of the same methodology for other regions should be aware of these settings. Again in GLCM RRI1 linear model, two minimum and maximum measured LAIs were estimated respectively, 2.47, 2.73 which are rather significant differences between estimated and measured LAI.

Although all the models could not accurately estimate LAI values less than 1.5, there was a major improvement in the result of non-parametric, non-linear RF model (RMSE = 0.85) comparing to other two regression models (RMSE = 1.23). The coefficient of determination of RF model did improve ($R^2 = 0.43$), and it is likely due to the use of 96 predictor variables and decision trees in the RF model. Random selection of 32 predictor variables in each node and matching the best model for estimating LAI resulted in the higher accuracy of LAI estimation using the RF model. In terms of comparing the RF and linear models it can be highlighted that RF models can handle either linear or non-linear regression while linear models (even multiple linear regressions) only perform well when we know the relations are definitely linear. As can be seen in figure 11 and figure 12 the relation between the LAI estimated from RRI1 as well as the GLCM-RRI1-mean and measured LAI appears not to be linear.

It can be highlighted that two linear models extrapolated the LAI in a wider range of values compare to RF model, in RRI1 from -1.30 to 9.30 and in texture from -0.30 to 8.13 (LAI mapped using linear models corresponding RRI1 VI and GLCM RRI1 can be found in appendix D). This last sentence clarifies two problems with estimating LAI from linear models. First, the large LAI values estimated by these models show that the input reflectance data are biased with reflectance of green understory which, can overestimate LAI. Second, negative LAI values estimated by linear models that were located in the areas with very low LAI, and in reality there is no negative LAI value.

A high-resolution map of estimated LAI was obtained using the Random Forest model by utilizing RapidEye spectral and spatial information and field measurements of LAI as inputs. This map can be applied in various forestry and park management activities including forest productivity, evapotranspiration and understanding physiological processes (Tang et al., 2012) over BFNP. Since BFNP covers a large area and consists various tree species, the result of current study provides a source of information for upcoming projects such as observing plant dynamics caused by climatic warming (Krebs, Boutin, & Boonstra, 2001) in this region.

The implemented methodology for LAI estimation using the RF model can be used as a reference method for other geographical regions with different vegetation covers. Map uncertainty and errors will be different for other regions. Therefore, the user should consider the specifications of the target area for another project. As the mean LAI value for this study was 3.35, in other geographical regions depending on the height, latitude and longitude, type of forest, heterogeneity and homogeneity the range of the LAI increases or decreases. For example, in the study of Heiskanen, Rautiainen, Korhonen, Möttöus, & Stenberg (2011) the LAI in boreal forest varies between 1.5 to 2.4 and in another study Eckert et al., (2005) measured LAI values between 1.0 to 2.6. Therefore, it should be kept in mind that all the aspects that considered in the presented study should be transferred to another area specifications. In terms of comparing the literature and their results, it should be taken into account that for each study there is a specific collected data, method and environmental condition and these differences impede the comparability (Asner, Scurlock, & A. Hicke, 2003). Moreover, seasonal difference of LAI, which is related to the amount of rainfall or other climatic conditions and also the growth stage level of the trees are other reasons that influence the results of different studies (Bréda, 2003). In our case, LAI values were measured when all the trees were in the greenest stage of life and deciduous trees covered by the maximum number of the leaves.

5.1. The influence of atmospheric errors

Eliminating the atmospheric distortions from the image would help in extraction of values from pixels for any purposes. In the present study, because of some problems using atmospheric correction tool, FLAASH in ENVI, we assumed that there is no atmospheric error on the image. Therefore, for improving the

accuracy of the models and more accurate calculation of the average reflectance for each plot on the image, atmospheric correction has to be done.

Since red edge band is insensitive to the soil cover percentage, atmospheric errors, canopy structure and solar zenith angle, the VIs which made use of Red edge band are not under the influence of atmospheric effects (Xie et al., 2014a). But other predictor variables such as TVs are directly affected by the atmospheric condition such as fog and cloud. Therefore, the TVs calculated from the bands of the image were not accounted as the top-ranked predictor variables in the variable importance of the RF model for estimating LAI.

In summary, our results indicated that implementing RF model on RapidEye image is very suitable for LAI estimation in particular for the forest ecosystem. Estimating forest LAI is very valuable for understanding different processes in the forest, as it is one of the critical structure variables. The utilization of RapidEye image for modelling LAI in the forest also can be expanded to agricultural areas but with different transformation settings (Vuolo, Atzberger, Richter, & Dash, 2010).

“Unless you try to do something beyond what you have already mastered, you will never grow.”
Ralph Waldo Emerson
1803-1882

6. CONCLUSION AND RECOMMENDATION

The most important conclusions of present study are as follow:

- Since Red edge and NIR band contain useful information regarding vegetation cover, VI that made use of these two bands made a better correlation with measured LAI compares to other VIs.
- Texture information is a supplementary source of information for estimating LAI from the high-resolution image such as RapidEye. Despite the TVs extracted from reflectance (bands) of the RapidEye image, the TVs extracted from a highly correlated VI could estimate LAI with relatively the same accuracy as VI itself.
- In this study, the ratio of different TVs extracted from the bands and different window sizes did not end up with a strong correlation with measured LAI.
- In this study, an accurate LAI estimation was not possible using one single independent variable and linear model, when we do not know if the relation between dependent and independent variables is linear or non-linear.
- A better LAI estimation obtained using RF model. Multi variability and complexity of the RF model results in a more accurate LAI estimation compare to simple linear models.
- Utilization of the spectral and spatial information extracted from RapidEye image and using Random Forest model was a progressive method for estimating LAI.

6.1. Brief answer to the research questions:

Q1: Among vegetation indices studied, which one provide a more accurate estimate (Highest R^2 and lowest RMSE) for estimating LAI?

Answer: Amongst all the VIs, RRI1 that calculated from the ratio of NIR and Red edge bands provided the best correlation between measured and estimated LAI using the linear model. The RMSE and R^2 of the model are respectively 1.23 and 0.27.

Q2: In a comparison of using VIs, which degree of accuracy of textural information offer for estimating LAI based on the high-resolution multispectral image?

Answer: The low correlation between TVs with measured LAI indicated that texture information extracted from the RapidEye image can be used as a source of information for estimating LAI. Although, the GLCM TVs of RRI1 showed the highest correlation, especially GLCM-mean has almost a similar RMSE and R^2 with RRI1 as VI (RMSE = 1.23, R^2 = 0.28).

Q3: How effective Random Forest algorithm estimate LAI by combining VIs and textural based information from a RapidEye image?

Answer: Considering RMSE and R^2 resulted from RF model respectively, 0.85 and 0.43 there is a strong evidence that RF model accurately estimate LAI using all the predictor variables coming from spectral and spatial information of the RapidEye image.

6.2. Suggestions for further work:

1. Extracting VIs and GLCM TVs from the image are under the influence of atmospheric noises which eventuate in low correlation and higher RMSE for LAI estimation. Thus, applying atmospheric correction tools such as ATCOR or FLAASH for eliminating atmospheric errors is highly recommended.
2. Other parameters such as canopy cover, slope and elevation might be affecting in RF model for LAI estimation. Thus, for further studies, it is recommended to involve these parameters into the model.
3. A large number of observations result in better understanding of the relations between dependent and independent variable. Moreover, the linear relationship with a few observation might change to nonlinear when the number of collected samples enlarge. Thus, it is recommended to increase the number of collected samples.
4. TVs ratios calculated from band GLCM in this study were not recognized as important predictor variables for estimating LAI. Finding the importance of TVs ratios extracted from GLCM of the VI raster layers is also recommended.

LIST OF REFERENCES

- Albrechtsen, F. (2008). Statistical Texture Measures Computed from Gray Level Cooccurrence Matrices. *Image Processing Laboratory, Department of Informatics, University of Oslo*, 1–14. doi:10.2307/302397
- Ali, A. M., Darvishzadeh, R., Skidmore, A. K., & Duren, I. Van. (2015). Effects of Canopy Structural Variables on Retrieval of Leaf Dry Matter Content and Specific Leaf Area From Remotely Sensed Data, *PP(99)*, 1–12.
- Ali, M., Montzka, C., Stadler, A., Menz, G., Thonfeld, F., & Vereecken, H. (2015). Estimation and Validation of RapidEye-Based Time-Series of Leaf Area Index for Winter Wheat in the Rur Catchment (Germany). *Remote Sensing*, 7(3), 2808–2831. doi:10.3390/rs70302808
- Asam, S., Fabritius, H., Klein, D., Conrad, C., & Dech, S. (2013). Derivation of leaf area index for grassland within alpine upland using multi-temporal RapidEye data. *International Journal of Remote Sensing*, 34(23), 8628–8652. doi:10.1080/01431161.2013.845316
- Asner, G. P., Scurlock, J. M. O., & A. Hicke, J. (2003). Global synthesis of leaf area index observations: implications for ecological and remote sensing studies. *Global Ecology and Biogeography*, 12(3), 191–205. doi:10.1046/j.1466-822X.2003.00026.x
- Bacour, C., Baret, F., Béal, D., Weiss, M., & Pavageau, K. (2006). Neural network estimation of LAI, fAPAR, fCover and LAI×Cab, from top of canopy MERIS reflectance data: Principles and validation. *Remote Sensing of Environment*, 105(4), 313–325. doi:10.1016/j.rse.2006.07.014
- Bagheri, N., Ahmadi, H., Alavipanah, S. K., & Omid, M. (2013). Multispectral remote sensing for site-specific nitrogen fertilizer management. *Pesquisa Agropecuária Brasileira*, 48(10), 1394–1401. doi:10.1590/S0100-204X2013001000011
- Bannari, A., Morin, D., Bonn, F., & Huete, A. R. (1995). A review of vegetation indices. *Remote Sensing Reviews*, 13(1), 95–120. doi:10.1080/02757259509532298
- Baraldi, A., & Parmiggiani, F. (1995). Investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters. *IEEE Transactions on Geoscience and Remote Sensing*, 33(2), 293–304. doi:10.1109/36.377929
- Barclay, H. J. (1998). Conversion of total leaf area to projected leaf area in lodgepole pine and Douglas-fir. *Tree Physiology*, 18(3), 185–193. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12651388>
- Bavarian Forest National Park. (2015). Retrieved from http://www.nationalpark-bayerischer-wald.de/english/nationalpark/nature_protection/index.htm
- Bréda, N. J. J. (2003). Ground-based measurements of leaf area index: A review of methods, instruments and current controversies. *Journal of Experimental Botany*, 54(392), 2403–2417. doi:10.1093/jxb/erg263
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. doi:10.1023/A:1010933404324
- Broge, N. ., & Leblanc, E. (2001). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment*, 76(2), 156–172. doi:10.1016/S0034-4257(00)00197-8
- Brownlee, J. (2013). A Tour of Machine Learning Algorithms. Retrieved from <http://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/>
- Bunn, A. G. (2005). Observed and predicted responses of plant growth to climate across Canada. *Geophysical Research Letters*, 32(16), L16710. doi:10.1029/2005GL023646

- Cailleret, M., Heurich, M., & Bugmann, H. (2014). Reduction in browsing intensity may not compensate climate change effects on tree species composition in the Bavarian Forest National Park. *Forest Ecology and Management*, 328, 179–192. doi:10.1016/j.foreco.2014.05.030
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250. doi:10.5194/gmd-7-1247-2014
- Chen, J. M. (1996). Evaluation of Vegetation Indices and a Modified Simple Ratio for Boreal Applications. *Canadian Journal of Remote Sensing*, 22(3).
- Chen, J. M., & Cihlar, J. (1996). Retrieving leaf area index of boreal conifer forests using Landsat TM images. *Remote Sensing of Environment*, 55(2), 153–162. doi:10.1016/0034-4257(95)00195-6
- Chen, J. M., Rich, P. M., Gower, S. T., Norman, J. M., & Plummer, S. (1997). Leaf area index of boreal forests: Theory, techniques, and measurements. *Journal of Geophysical Research*, 102(D24), 29429. doi:10.1029/97JD01107
- Chen*, D., Stow, D. A., & Gong, P. (2004). Examining the effect of spatial resolution and texture window size on classification accuracy: an urban environment case. *International Journal of Remote Sensing*, 25(11), 2177–2192. doi:10.1080/01431160310001618464
- Cohen, W. B., Maersperger, T. K., Gower, S. T., & Turner, D. P. (2003). An improved strategy for regression of biophysical variables and Landsat ETM+ data. *Remote Sensing of Environment*, 84(4), 561–571. doi:10.1016/S0034-4257(02)00173-6
- Colombo, R., Bellenger, D., Fasolin, D., Marino, C. M. (2003). Retrieval of leaf area index in different vegetation types using high resolution satellite data. *Remote Sensing of Environment*, 86(1), 120–131. doi:10.1016/S0034-4257(03)00094-4
- Combal, B., Baret, F., Weiss, M., Trubuil, A., Macé, D., Pragnère, A., Wang, L. (2003). Retrieval of canopy biophysical variables from bidirectional reflectance. *Remote Sensing of Environment*, 84(1), 1–15. doi:10.1016/S0034-4257(02)00035-4
- Combal, B., Baret, F., & Weiss, M. (2002). Improving canopy variables estimation from remote sensing data by exploiting ancillary information. Case study on sugar beet canopies. *Agronomie*, 22(2), 205–215.
- Croft, H., Chen, J. M., Zhang, Y., Simic, A., Noland, T. L., Nesbitt, N., & Arabian, J. (2015). Evaluating leaf chlorophyll content prediction from multispectral remote sensing data within a physically-based modelling framework. *ISPRS Journal of Photogrammetry and Remote Sensing*, 102, 85–95. doi:10.1016/j.isprsjprs.2015.01.008
- Curran, P. (1980). Multispectral remote sensing of vegetation amount. *Progress in Physical Geography*, 4(3), 315–341.
- Curran, P. J., Dungan, J. L., & Gholz, H. L. (1992). Seasonal LAI in slash pine estimated with Landsat TM. *Remote Sensing of Environment*, 39, 3–13. doi:10.1016/0034-4257(92)90136-8
- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random Forests for Classification in Ecology. *Ecology*, 88(11), 2783–2792. doi:10.1890/07-0539.1
- Darvishzadeh, R., Skidmore, A., Schlerf, M., & Atzberger, C. (2008). Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland. *Remote Sensing of Environment*, 112(5), 2592–2604.

- Díaz-Uriarte, R., & Alvarez de Andrés, S. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, 7, 3. doi:10.1186/1471-2105-7-3
- Eckert, S., Kellenberger, T. W., & Lencinas, J. D. (2005). Classification and Forest Parameter Extraction of Patagonian Lenga Forests With Aster and Landsat Etm + Data. *October*, 23–27.
- Ehammer, A., Fritsch, S., Conrad, C., & Lamers, J. P. a. (2010). Statistical derivation of fPAR and LAI for irrigated cotton and rice in arid Uzbekistan by combining multi-temporal RapidEye data and ground measurements. *Proceedings of SPIE* 7824, 7824(0), 1–10. doi:10.1117/12.864796
- Elvidge, C., & Chen, Z. (1995). Comparison of broad-band and narrow-band red and near-infrared vegetation indices. *Remote Sensing of Environment*, 54(1), 38–48. doi:10.1016/0034-4257(95)00132-K
- Falkowski, M. J., Evans, J. S., Martinuzzi, S., Gessler, P. E., & Hudak, A. T. (2009). Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. *Remote Sensing of Environment*, 113(5), 946–956. doi:10.1016/j.rse.2009.01.003
- Fang, H. L., & Liang, S. L. (2003). Retrieving leaf area index with a neural network method: Simulation and validation. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(9), 2052–2062. doi:10.1109/TGRS.2003.813493
- Fassnacht, K. S., Gower, S. T., MacKenzie, M. D., Nordheim, E. V., & Lillesand, T. M. (1997). Estimating the leaf area index of North Central Wisconsin forests using the landsat thematic mapper. *Remote Sensing of Environment*, 61(2), 229–245. doi:10.1016/S0034-4257(97)00005-9
- Food and Agriculture Organization of the United Nations. (2006). *Global Forest Resources Assessment 2005: Progress towards sustainable forest management. FAO Forestry Paper*. doi:ISBN 92-5-105481-9
- Gadkari, D. (2004). *Image Quality Analysis Using GLCM*. University of Central Florida.
- Gong, P., Pu, R., Biging, G. S., & Larrieu, M. R. (2003). Estimation of forest leaf area index using vegetation indices derived from Hyperion hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 41(6), 1355–1362. doi:10.1109/TGRS.2003.812910
- Güneralp, İ., Filippi, A. M., & Randall, J. (2014). Estimation of floodplain aboveground biomass using multispectral remote sensing and nonparametric modeling. *International Journal of Applied Earth Observation and Geoinformation*, 33, 119–126. doi:10.1016/j.jag.2014.05.004
- Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5), 786–804. doi:10.1109/PROC.1979.11328
- Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6), 610–621. doi:10.1109/TSMC.1973.4309314
- Haralick, R., Shanmugan, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*. doi:10.1109/TSMC.1973.4309314
- Hastie, T., Tibshirani, R., & Jerome, F. (2009). *Springer Series in Statistics The Elements of. The Mathematical Intelligencer* (Second., Vol. 27). Stanford, California: Springer. doi:10.1007/b94608
- He, K. S., Rocchini, D., Neteler, M., & Nagendra, H. (2011). Benefits of hyperspectral remote sensing for tracking plant invasions. *Diversity and Distributions*, 17(3), 381–392. doi:10.1111/j.1472-4642.2011.00761.x
- Heiskanen, J., Rautiainen, M., Korhonen, L., Möttö, M., & Stenberg, P. (2011). Retrieval of boreal forest LAI using a forest reflectance model and empirical regressions. *International Journal of Applied Earth Observation and Geoinformation*, 13(4), 595–606. doi:10.1016/j.jag.2011.03.005

- Heurich, M., Baierl, F., Günther, S., & Sinner, K. F. (2011). Management and conservation of large mammals in the Bavarian Forest National Park. *Silva Gabreta*, 17(1), 1–18. Retrieved from http://www.npsumava.cz/gallery/17/5219-sg17_1-heurich.pdf
- Heurich, M., & Neufanger, M. (2005). *Die Wälder des Nationalparks Bayerischer Wald*. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Die+W?lder+des+Nationalpar ks+Bayerischer+Wald#0>
- Horning, N. (2010). Random Forests: An algorithm for image classification and generation of continuous fields data sets. In *Proceeding of International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences* (pp. 9–11). Hanoi, Vietnam.
- Houborg, R., Soegaard, H., & Boegh, E. (2007). Combining vegetation index and model inversion methods for the extraction of key vegetation biophysical parameters using Terra and Aqua MODIS reflectance data. *Remote Sensing of Environment*, 106(1), 39–58. doi:10.1016/j.rse.2006.07.016
- Hudak, A. T., Lefsky, M. A., Cohen, W. B., & Berterretche, M. (2002). Integration of lidar and Landsat ETM+ data for estimating and mapping forest canopy height. *Remote Sensing of Environment*, 82(2-3), 397–416. doi:10.1016/S0034-4257(02)00056-1
- Huete, A., Justice, C., & Liu, H. (1994). Development of vegetation and soil indices for MODIS-EOS. *Remote Sensing of Environment*, 49(3), 224–234. doi:10.1016/0034-4257(94)90018-3
- Johnson, L. F., Roczen, D. E., Youkhana, S. K., Nemani, R. R., & Bosch, D. F. (2003). Mapping vineyard leaf area with multispectral satellite imagery. *Computers and Electronics in Agriculture*, 38(1), 33–44. doi:10.1016/S0168-1699(02)00106-0
- Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., & Weiss, M. (2004). Methods for Leaf Area Index Determination Part I: Theories , Techniques and Instruments. *Area*, 121, 1–42. doi:10.5897/AJAR2013.7346
- Kashyap, H., Ahmed, H. A., Hoque, N., Roy, S., & Bhattacharyya, D. K. (2015). Big Data Analytics in Bioinformatics : A Machine Learning Perspective. *LATEX CLASS FILES*, 13(9), 1–20.
- Kaufman, Y. J., & Tanre, D. (1992). Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *Geoscience and Remote Sensing, IEEE Transactions on*. doi:10.1109/36.134076
- Kayitakire, F., Hamel, C., & Defourny, P. (2006). Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. *Remote Sensing of Environment*, 102(3-4), 390–401. doi:10.1016/j.rse.2006.02.022
- Kramer, H. . (2002). RapidEye Earth Observation Constellation. Retrieved from <https://directory.eoportal.org/web/eoportal/satellite-missions/r/rapideye>
- Kraus, T., Schmidt, M., Dech, S. W., & Samimi, C. (2009). The potential of optical high resolution data for the assessment of leaf area index in East African rainforest ecosystems. *International Journal of Remote Sensing*, 30(19), 5039–5059. doi:10.1080/01431160903022878
- Krebs, C. J., Boutin, S., & Boonstra, R. (2001). *Ecosystem Dynamics Of The Boreal Forest: The Kluane Project*. Oxford University Press. doi:10.2307/3071790
- Latifi, H., & Koch, B. (2012). Evaluation of most similar neighbour and random forest methods for imputing forest inventory variables using data from target and auxiliary stands. *International Journal of Remote Sensing*, 33(21), 6668–6694. doi:10.1080/01431161.2012.693969
- Lee, K. S., Park, Y. I., Kim, S. H., Park, J. H., Woo, C. S., & Jang, K. C. (2006). Remote sensing estimation of forest LAI in close canopy situation. *Korean Journal of Remote Sensing*, 22(5), 305–311.

- Liaw, a., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, 2(December), 18–22. doi:10.1177/154405910408300516
- Liaw, A. ., & Wiener, M. (2015). Title Breiman and Cutler's Random Forests for Classification and Regression Description Classification and regression based on a forest of trees using random inputs. Retrieved from <https://www.stat.berkeley.edu/~breiman/RandomForests/>
- Liu, Z., Chen, J. M., Jin, G., & Qi, Y. (2015). Estimating seasonal variations of leaf area index using litterfall collection and optical methods in four mixed evergreen–deciduous forests. *Agricultural and Forest Meteorology*, 209-210, 36–48. doi:10.1016/j.agrformet.2015.04.025
- López-Serrano, F. R., Landete-Castillejos, T., Martínez-Millán, J., & Cerro-Barja, A. Del. (2000). LAI estimation of natural pine forest using a non-standard sampling technique. *Agricultural and Forest Meteorology*, 101(2-3), 95–111. doi:10.1016/S0168-1923(99)00171-9
- Maass, J., Vose, J. M., Swank, W. T., & Martínez-Yrizar, A. (1995). Seasonal changes of leaf area index (LAI) in a tropical deciduous forest in west Mexico. *Forest Ecology and Management*, 74(1-3), 171–180. doi:10.1016/0378-1127(94)03485-F
- Monteith, J. L. and Unsworth, M. H. (1990). *Principles of Environmental Physics* (second edi.). London: Cambridge University Press. Retrieved from <http://dx.doi.org/10.1017/S0014479700001381> (About DOI)
- Mróz, M., & Sobieraj, A. (2004). Comparison of several vegetation indices calculated on the basis of a seasonal SPOT XS time series, and their suitability for land cover and agricultural crop identification. *Technical Sciences / University of Warmia and Mazury in Olsztyn*, 7(7), 39–66.
- National Space Development Agency of Japan, NASDA, & Remote Sensing Technology Center of Japan, RESTEC. (1999). *Remote Sensing and GIS*. (J. A. of R. Sensing, Ed.). Asian Institute of Technology (AIT). Retrieved from <http://wtlab.iis.u-tokyo.ac.jp/~wataru/lecture/rsgis/index.htm>
- Nitze, I., Schulthess, U., & Asche, H. (2012). Comparison of machine learning algorithms random forest, artificial neural network and support vector machine to maximum likelihood for supervised crop type classification. *Proceedings of the 4th GEOBLA*, 35–40.
- Ozdemir, I., Norton, D. A., Ozkan, U. Y., Mert, A., & Senturk, O. (2008). Estimation of Tree Size Diversity Using Object Oriented Texture Analysis and Aster Imagery. *Sensors*, 8(8), 4709–4724. doi:10.3390/s8084709
- Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1), 217–222. doi:10.1080/01431160412331269698
- Pasolli, L., Asam, S., Castelli, M., Bruzzone, L., Wohlfahrt, G., Zebisch, M., & Notarnicola, C. (2015). Retrieval of Leaf Area Index in mountain grasslands in the Alps from MODIS satellite imagery. *Remote Sensing of Environment*, 165, 159–174. doi:10.1016/j.rse.2015.04.027
- Pathak, B., & Barooah, D. (2013). Texture analysis based on the gray-level co-occurrence matrix considering possible orientations. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 2(9), 4206–4212. Retrieved from http://www.ijareeie.com/upload/2013/september/7_-TEXTURE.pdf
- Peddle, D. R., Hall, F. G., & Ledrew, E. F. (1999). Spectral mixture analysis and geometric-optical reflectance modeling of boreal forest biophysical structure. *Remote Sensing of Environment*, 67(3), 288–297. doi:10.1016/S0034-4257(98)00090-X

- Powell, S. L., Cohen, W. B., Healey, S. P., Kennedy, R. E., Moisen, G. G., Pierce, K. B., & Ohmann, J. L. (2010). Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sensing of Environment*, 114(5), 1053–1068. doi:10.1016/j.rse.2009.12.018
- Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems*, 9(2), 181–199. doi:10.1007/s10021-005-0054-1
- Pu, R. (2012). Comparing Canonical Correlation Analysis with Partial Least Squares Regression in Estimating Forest Leaf Area Index with Multitemporal Landsat TM Imagery. *GIScience & Remote Sensing*, 49(1), 92–116. doi:10.2747/1548-1603.49.1.92
- Pu, R., & Cheng, J. (2015). Mapping forest leaf area index using reflectance and textural information derived from WorldView-2 imagery in a mixed natural forest area in Florida, US. *International Journal of Applied Earth Observation and Geoinformation*, 42, 11–23. doi:10.1016/j.jag.2015.05.004
- Pu, R., & Gong, P. (2004). Wavelet transform applied to EO-1 hyperspectral data for forest LAI and crown closure mapping. *Remote Sensing of Environment*, 91(2), 212–224. doi:10.1016/j.rse.2004.03.006
- RapidEye. (2011a). From the Corporate RapidEye Website – Frequently Asked Questions, 1. Retrieved from https://apollomapping.com/wp-content/user_uploads/2011/09/RapidEye_irradiance_values.pdf
- RapidEye. (2011b). Satellite imagery product specifications, (November), 1–48. Retrieved from http://www.flyby.it/images/brochure/rapideye/eng/re_product_specifications_eng.pdf
- Rhou, C., Goulas, Y., Daumard, F., Ounis, A., & Moya, I. (2014). Effect of the Canopy Structure on the Chlorophyll Fluorescence Yield At Canopy Scale. *5th International Workshop on Remote Sensing of Vegetation Fluorescence*, (1), 1–9. Retrieved from http://congrexprojects.com/Custom/14C04/14C04_index.htm
- Roli, F. Kittler, J., & Windeatt, T. (2004). *Multiple Classifier Systems*. Berlin: Springer-Verlag Berlin Heidelberg.
- Rouse, J. W. J., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring Vegetation Systems in the Great Plains with ERTS. *Third Earth Resources Technology Satellite-1 Symposium*, 1, 309.
- Schlerf, M., Atzberger, C., & Hill, J. (2005). Remote sensing of forest biophysical variables using HyMap imaging spectrometer data. *Remote Sensing of Environment*, 95(2), 177–194. doi:10.1016/j.rse.2004.12.016
- Schmidt, H., & Karnieli, a. (2001). Sensitivity of vegetation indices to substrate brightness in hyper-arid environment: The Makhtesh Ramon Crater (Israel) case study. *International Journal of Remote Sensing*, 22(17), 3503–3520. doi:10.1080/01431160110063779
- Schneider, J. (1997). Cross Validation. Retrieved from <https://www.cs.cmu.edu/~schneide/tut5/node42.html#SECTION00092000000000000000>
- Scurlock, J. M. O., Asner, G. P., & Gower, S. T. (2001). Worldwide Historical Estimates and Bibliography of Leaf Area Index, 1932-2000. ORNL *Technical Memorandum* TM-2001/268, 27(December).
- Shen, L., Li, Z., & Guo, X. (2014). Remote Sensing of Leaf Area Index (LAI) and a Spatiotemporally Parameterized Model for Mixed Grasslands, 4(1), 46–61.

- Shippert, P. (2013). Digital Number, Radiance, and Reflectance. Retrieved from <http://www.exelisvis.com/Home/NewsUpdates/TabId/170/ArtMID/735/ArticleID/13592/Digital-Number-Radiance-and-Reflectance.aspx>
- Song, C., & Dickinson, M. B. (2008). Extracting forest canopy structure from spatial information of high resolution optical imagery: tree crown size versus leaf area index. *International Journal of Remote Sensing*, 29(19), 5605–5622. doi:10.1080/01431160802060904
- Soudani, K., François, C., le Maire, G., Le Dantec, V., & Dufrêne, E. (2006). Comparative analysis of IKONOS, SPOT, and ETM+ data for leaf area index estimation in temperate coniferous and deciduous forest stands. *Remote Sensing of Environment*, 102(1-2), 161–175. doi:10.1016/j.rse.2006.02.004
- Spritsin, M., Berliner, P., Cohen, S., & Karnieli, A. (2013). Using Multispectral Spaceborne Imagery to Assess Mean Tree Height in a Dryland Plantation. *ISRN Forestry*, 2013, 1–8. doi:10.1155/2013/485264
- Steven, M. D., Malthus, T. J., Baret, F., Xu, H., & Chopping, M. J. (2003). Intercalibration of vegetation indices from different sensor systems. *Remote Sensing of Environment*, 88(4), 412–422. doi:10.1016/j.rse.2003.08.010
- Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics*, 8, 25. doi:10.1186/1471-2105-8-25
- Thompson, J. R., & Spies, T. A. (2009). Vegetation and weather explain variation in crown damage within a large mixed severity wildfire. *Forest Ecology and Management*, 258, 1684–1694. doi:10.1016/j.foreco.2009.07.031
- Turner, D. P., Cohen, W. B., Kennedy, R. E., Fassnacht, K. S., & Briggs, J. M. (1999). Relationships between Leaf Area Index and Landsat TM Spectral Vegetation Indices across Three Temperate Zone Sites. *Remote Sensing*, 68(April 1998), 52–68.
- Varshney, P. K., & Arora, M. K. (2004). *Advanced image processing techniques for remotely sensed hyperspectral data : with 30 tables*. Berlin [u.a.]: Springer.
- Vuolo, F., Atzberger, C., Richter, K., & Dash, J. (2010). Retrieval of Biophysical Vegetation Products From Rapideye Imagery. *Symposium A Quarterly Journal In Modern Foreign Literatures*, XXXVIII, 281–286. Retrieved from http://www.isprs.org/proceedings/XXXVIII/part7/a/pdf/281_XXXVIII-part7A.pdf
- Wang, L. T., Wang, S. X., Zhou, Y., Liu, W. L., & Wang, F. T. (2011). [Vegetation water content retrieval and application of drought monitoring using multi-spectral remote sensing]. *Guang Pu Xue Yu Guang Pu Fen Xi= Guang Pu*, 31(10), 2804–2808.
- Wang, Q., Adiku, S., Tenhunen, J., & Granier, A. (2005). On the relationship of NDVI with leaf area index in a deciduous forest site. *Remote Sensing of Environment*, 94(2), 244–255. doi:10.1016/j.rse.2004.10.006
- Watson, D. J. (1947). Comparative Physiological Studies on the Growth of Field Crops: I. Variation in Net Assimilation Rate and Leaf Area between Species and Varieties, and within and between Years. *Annals of Botany*, 11(41), 41–76 CR – Copyright © 1947 Oxford University. doi:10.2307/42907002
- Wulder, M. A., LeDrew, E. F., Franklin, S. E., & Lavigne, M. B. (1998). Aerial Image Texture Information in the Estimation of Northern Deciduous and Mixed Wood Forest Leaf Area Index (LAI). *Remote Sensing of Environment*, 64(1), 64–76. doi:10.1016/S0034-4257(97)00169-7

- Wulder, M. a., & Seemann, D. (2003). Forest inventory height update through the integration of lidar data with segmented Landsat imagery. *Canadian Journal of Remote Sensing*, 29(5), 536–543. doi:10.5589/m03-032
- Xie, Q., Huang, W., Liang, D., Chen, P., Wu, C., Yang, G., & Zhang, J. (2014a). Indices Derived From Airborne Hyperspectral Images in Winter Wheat. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(8), 3586–3594.
- Xie, Q., Huang, W., Liang, D., Chen, P., Wu, C., Yang, G., & Zhang, J. (2014b). Indices Derived From Airborne Hyperspectral Images in Winter Wheat. *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING*, 7(8), 3586–3594.
- Yang, G. J., Zhao, C. J., Liu, Q., Huang, W. J., & Wang, J. H. (2011). Inversion of a Radiative Transfer Model for Estimating Forest LAI From Multisource and Multiangular Optical Remote Sensing Data. *Ieee Transactions on Geoscience and Remote Sensing*, 49(3), 988–1000. doi:10.1109/tgrs.2010.2071416
- Yu, Y., Yang, X., & Fan, W. (2015). Estimates of forest structure parameters from GLAS data and multi-angle imaging spectrometer data. *International Journal of Applied Earth Observation and Geoinformation*, 38, 65–71. doi:10.1016/j.jag.2014.12.013
- Yuan, H., Ma, R., Atzberger, C., Li, F., Loiselle, S., & Luo, J. (2015). Estimating Forest fAPAR from Multispectral Landsat-8 Data Using the Invertible Forest Reflectance Model INFORM. *Remote Sensing*, 7(6), 7425–7446. doi:10.3390/rs70607425
- Zheng, G., & Moskal, L. M. (2009). Retrieving Leaf Area Index (LAI) Using Remote Sensing: Theories, Methods and Sensors. *Sensors*, 9(4), 2719–2745. doi:10.3390/s90402719
- Zhou, J.-J., Zhao, Z., Zhao, J., Zhao, Q., Wang, F., & Wang, H. (2014). A comparison of three methods for estimating the LAI of black locust (*Robinia pseudoacacia* L.) plantations on the Loess Plateau, China. *International Journal of Remote Sensing*, 35(1), 171–188. doi:10.1080/01431161.2013.866289

Appendix

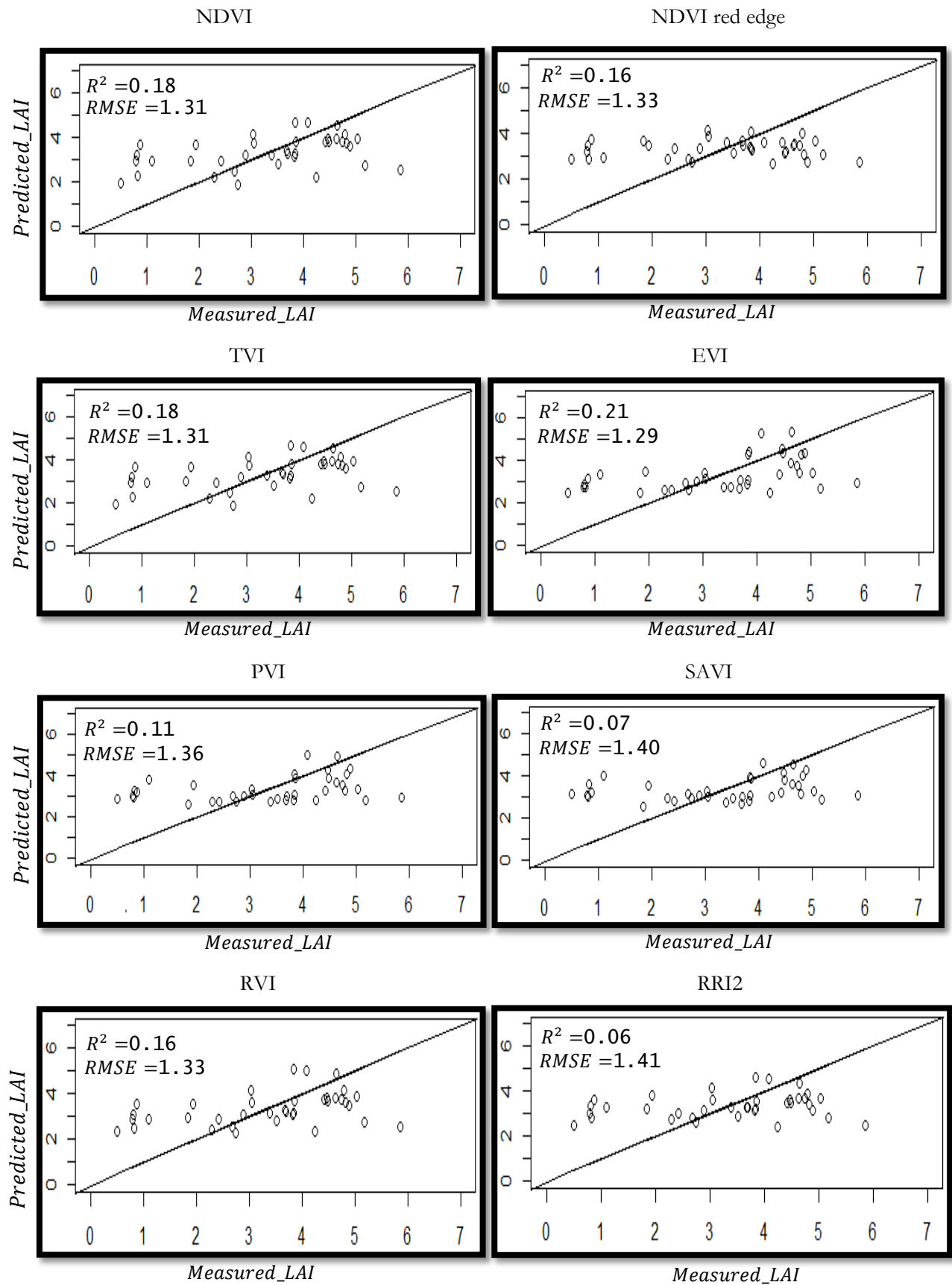
A. The table of literature review for estimating LAI in the Forests using univariate and multivariate models.

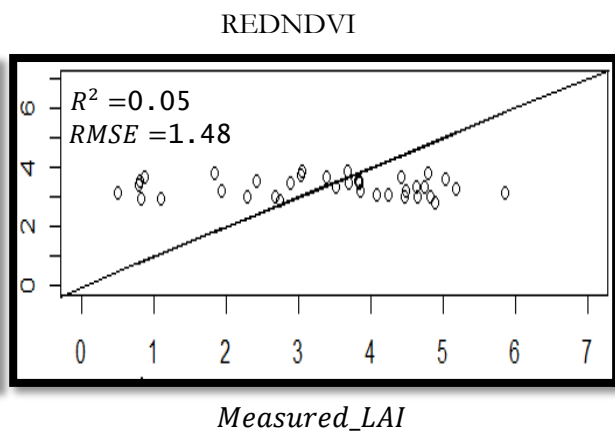
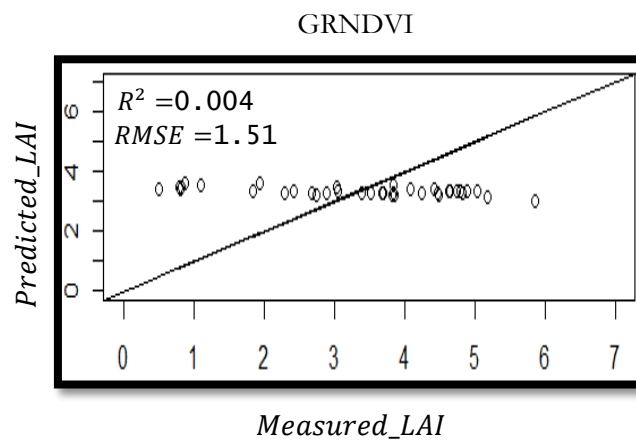
No	Reference	Image data	Study area	Modelling method	Retrieval accuracy
1	(P. J. Curran et al., 1992)	Landsat thematic mapper (TM)	Slash pine	Regression analysis (VIs)	RMSE=0.74
2	(Fassnacht et al., 1997)	Landsat Thematic mapper (TM)	Mixed conifer hard wood forest	Linear regression (VIs)	$R^2=0.69$, 0.73 for different reports
3	(Wulder et al., 1998)	CASI imagery	Broad leaf deciduous and coniferous forest	Regression models (Texture and VIs)	NDVI+ semi variance+ second order $R^2=0.51$
4	(Colombo, R. Bellinger, D. Fasolinic, D. Marino, 2003)	IKONOS	Different vegetation type	Multivariable Regression analysis	R^2 NDVI= 0.33 , R^2 NDVI+dissimilarity = 0.62
5	(Cohen et al., 2003)	Landsat ETM+	Black spruce	Regression analysis, Canonical correlation analysis (VIs)	For different models RMSE = 10.41-12.68
6	(Johnson et al., 2003)	IKONOS	vineyard	Regression analysis	$R^2= 0.72$
7	(Pu & Gong, 2004)	Eo1-Hyperion	Mixed coniferous forest	3methods of Texture analysis: BS, PCA, WT ¹	RMSE of BS = 0.82 RMSE of PCA = 0.80 RMSE of WT= 0.41
8	(Wang et al., 2005)	AVHRR, SPOT, MODIS	Deciduous forest	Regression Analysis (VIs)	Reports on different years for AVHRR $R^2= 0.74 - 0.63$
9	(Schlerf et al., 2005)	Hymap	Homogeneous Norway spruce	Regression Analysis (VIs)	$R^2= 0.79$ RMSE=0.52

¹ Bands selection(BS), Principle Component Analysis(PCA)s, Wavelength transform (WT)

10	(Soudani et al., 2006)	IKONOS, SPOT, ETM+	Coniferous, deciduous	Regression Analysis, PROSAIL, PROSPECT	$R^2 = 0.77$ RMSE=0.17
11	(Song & Dickinson, 2008)	IKONOS imagery	Coniferous hardwood forest	Regression Analysis (Texture and VIs)	Best combination $R^2 = 0.56$
12	(Ehammer et al., 2010)	RapidEye	Cotton and Rice field	Linear and non-linear regression	Multiple results of LAI in each field, optimal R^2 rice =0.92, optimal Cotton $R^2 = 0.81$
13	(Pu, 2012)	Landsat TM	Wetland associated with Cypress Creek	Partial Least Square and Canonical Correlation Analysis	PLS RMSE=0.67 CCA RMSE=0.45
14	(Xie et al., 2014b)	Pushbroom Hyperspectral Imager (PHI)	Winter Wheat	Simple Regression model (VIs)	Multiple results, the best $R^2 NDVI - like = 0.77$
15	(Zhou et al., 2014)	Quickbird image	Black locus	Regression analysis (Texture and VIs)	$R^2 = 0.84$
16	(Pu & Cheng, 2015)	WorldView-2	Mixed forest	Linear regression (Canonical variables) Spectrum based features and texture based features	Optimal RMSE = 0.36

B. The figures show the correlation between VIs and measured LAI





C. Example of raster layers produced by GLCM for raster RRI1, window 3.

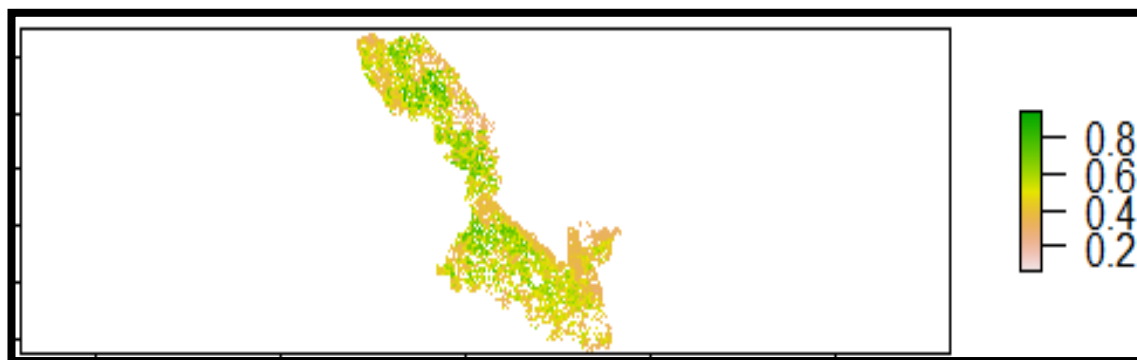


Figure c.1. Mean-GLCM of raster layer RRI1

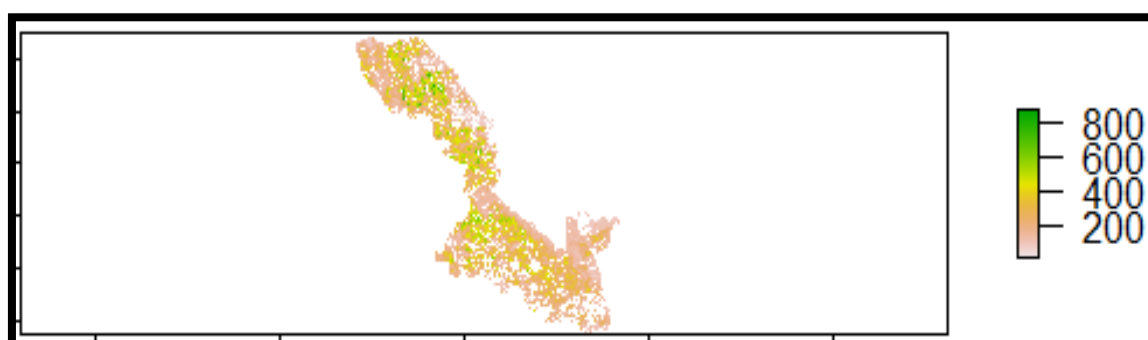


Figure c.2. Variance-GLCM of raster layer RRI1



Figure c.3. Homogeneity-GLCM of raster layer RRI1

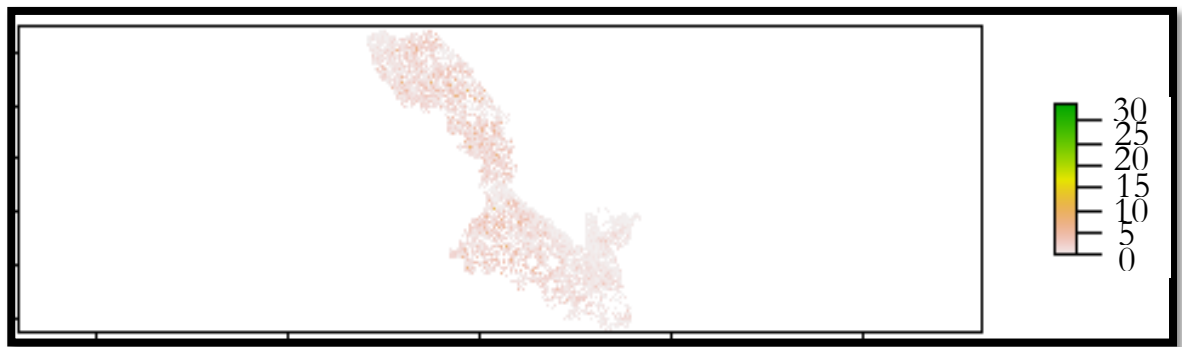


Figure c.4. Contrast-GLCM of raster layer RRI1

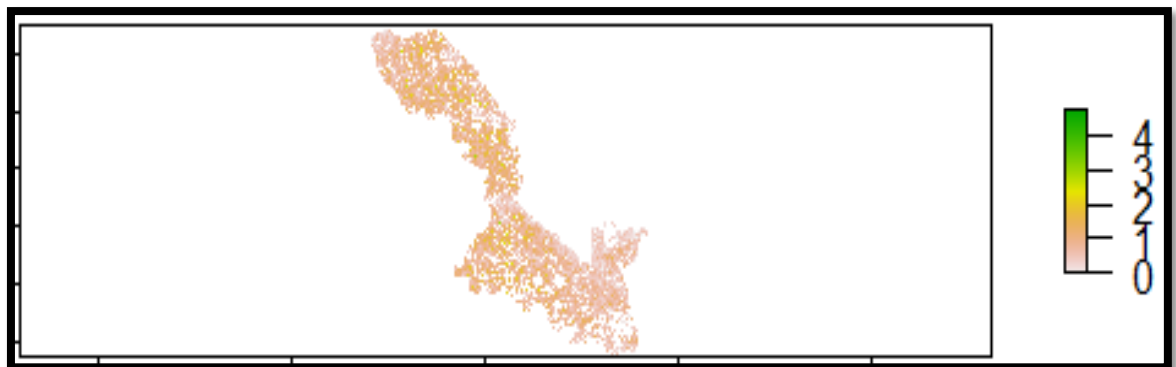


Figure c.5. Dissimilarity-GLCM of raster layer RRI1

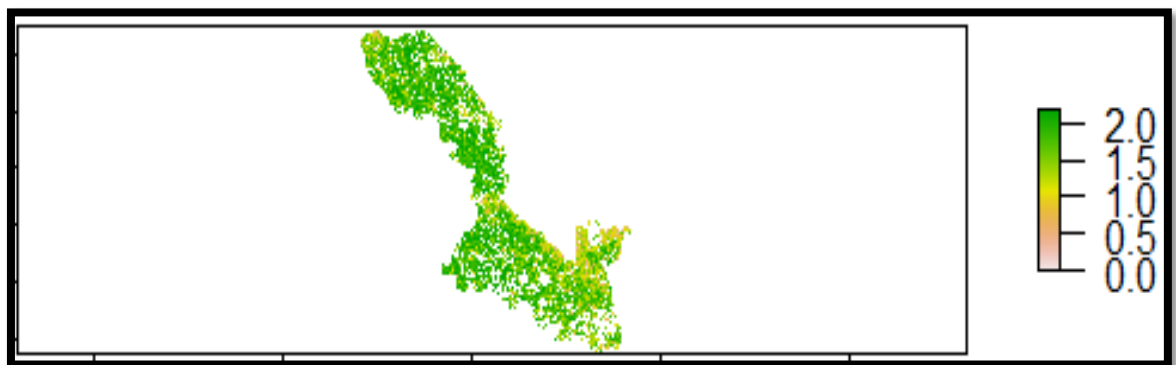


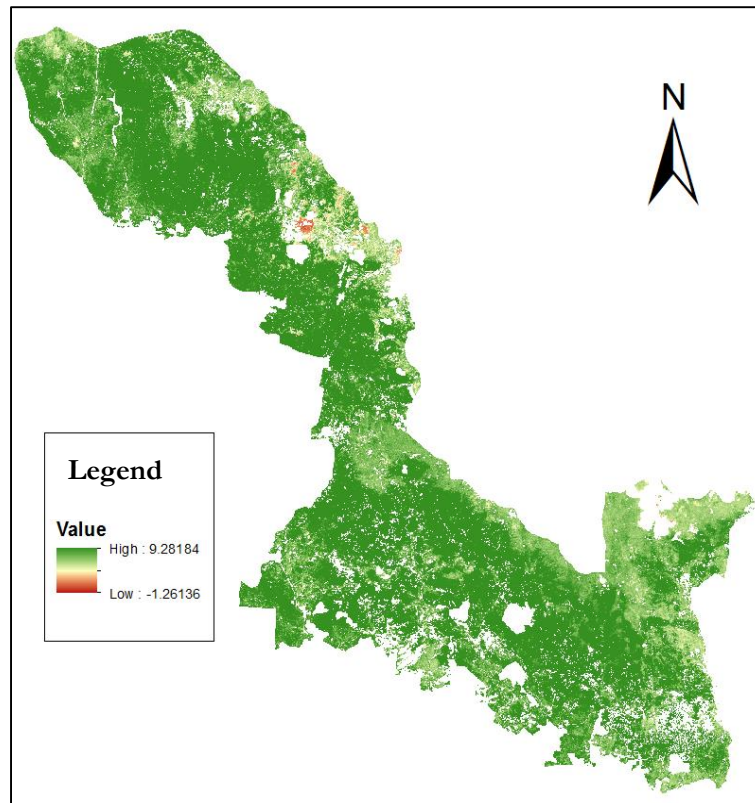
Figure c.6. Entropy-GLCM of raster layer RRI1



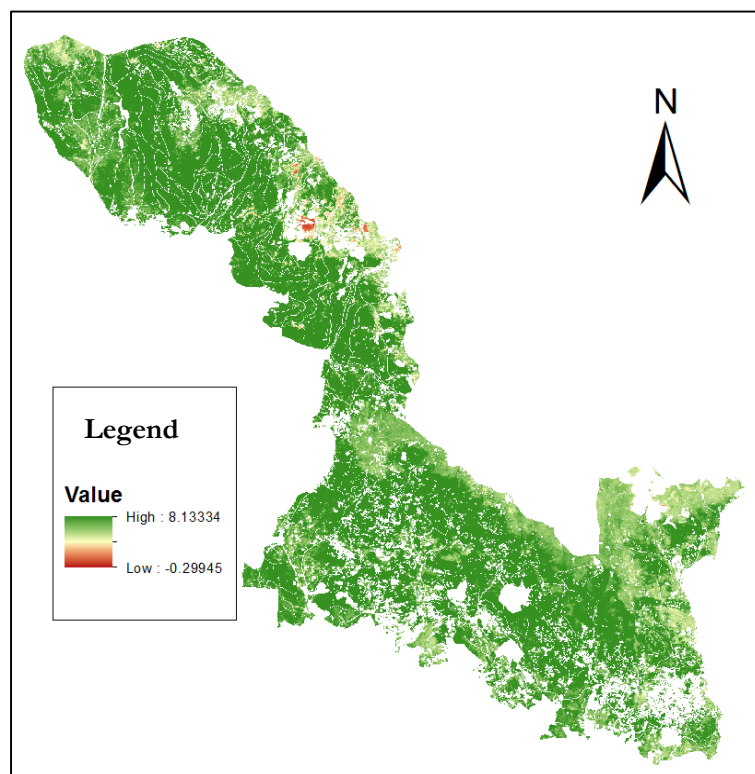
Figure c.7. Second moment-GLCM of raster layer RRI1

D. The LAI maps created by linear models RRI1 (a) and GLCM RRI1 (b).

(a)



(b)



E. R script used:

```
#####library#####
library(raster)
library(rgdal)
library(glcm)
library(maptools)
library(cvTools)
library(Metrics)
library(graphics)
library(randomForest)
library(data.table)
library(stats)
library(ggplot2)
#####Bands#####
band1<-
raster("D:\\MSc\\Imagedata\\Bavaria_boundary_raster_corrected\\Forest_Boundary\\Forest_noclou
d.tif", band=1, values=TRUE)
band2<-
raster("D:\\MSc\\Imagedata\\Bavaria_boundary_raster_corrected\\Forest_Boundary\\Forest_noclou
d.tif", band=2, values=TRUE)
band3<-
raster("D:\\MSc\\Imagedata\\Bavaria_boundary_raster_corrected\\Forest_Boundary\\Forest_noclou
d.tif", band=3, values=TRUE)
band4<-
raster("D:\\MSc\\Imagedata\\Bavaria_boundary_raster_corrected\\Forest_Boundary\\Forest_noclou
d.tif", band=4, values=TRUE)
band5<-
raster("D:\\MSc\\Imagedata\\Bavaria_boundary_raster_corrected\\Forest_Boundary\\Forest_noclou
d.tif", band=5, values=TRUE)
Image_Forest = stack(band1, band2, band3, band4, band5)
#####
ref1=Shapefile_attributes$band1
ref2=Shapefile_attributes$band2
ref3=Shapefile_attributes$band3
ref4=Shapefile_attributes$band4
ref5=Shapefile_attributes$band5
#####polygon and plots#####
d=readShapePoints(file.choose())
Shapefile_attributes =
data.frame(band1=d$band1,band2=d$band2,band3=d$band3,band4=d$band4,band5=d$band5,
LAI=d$LAI, N=d$N, E=d$E)
#####NDVI#####
ndvi=(band5-band3)/(band5+band3)
NDVI_values = (ref5 - ref3)/(ref5 + ref3)
Table = data.frame(Shapefile_attributes,NDVI_values)
```



```

LAI= Table$LAI
NDVI = Table$NDVI_values
cor(LAI,NDVI)
#####RVI#####
rvi<- band5/band3
RVI_values = ref5/ref3
Table <- data.frame(Shapefile_attributes,NDVI_values,RVI_values)
X <- Table$LAI
Y <- Table$RVI_values
cor(Y,X)
#####NDVIrededge#####
ndvirededge <- (band4-band3)/(band4*band3)
NDVIrededge_values = (ref4-ref3)/(ref4*ref3)
Table <- data.frame(Shapefile_attributes,NDVIrededge_values,NDVI_values,RVI_values)
X <- Table$LAI
Y <- Table$NDVIrededge_values
cor(Y,X)
#####RRI1#####
RVIrededge <- band5/band4
RVIrededge_values = ref5/ref4
Table <-
data.frame(Shapefile_attributes,NDVI_values,RVI_values,NDVIrededge_values,RVIrededge_values)
X <- Table$LAI
Y <- Table$RVIrededge_values
cor(Y,X)

RRI1_cvfit = Table$RVIrededge_values
LAI_Measured = Table$LAI
Predicted_LAI=c()
slope=c()
intercept=c()
for(i in 1:37){
  model_RRI1= lm(LAI_Measured[-i]~RRI1_cvfit[-i])
  slope[i]= model_RRI1$coefficients[2]
  intercept[i] = model_RRI1$coefficients[1]
  Predicted_LAI[i] = model_RRI1$coefficients[2]*RRI1_cvfit[i]+model_RRI1$coefficients[1]
}
RRI1_predictions = data.frame(slope, intercept, Predicted_LAI, LAI_Measured)
cor(LAI_Measured,Predicted_LAI)^2
rmse(LAI_Measured,Predicted_LAI)
plot(LAI_Measured,Predicted_LAI, xlim=c(0,7), ylim=c(0,7))
model=lm(LAI_Measured~Predicted_LAI)
lines(LAI_Measured,LAI_Measured)

```

```

writeRaster(RVIrededge, filename="D:\\MSc\\Imagedata\\LAI_37\\test.tif", format="GTiff",
overwrite=TRUE)
RRI1_raster <- raster("D:\\MSc\\Imagedata\\LAI_37\\test.tif", values=TRUE)
#####crating map of LAI using linear model#####
#####
image_RVIred=stack(RVIrededge)

Table3 = data.frame(LAI=d$LAI,
                    layer=RVIrededge_values)

model_linear = lm(LAI ~ layer, data=Table3)
Map_linear = predict(image_RVIred, model_linear)
writeRaster(Map_linear, filename="D:\\MSc\\Imagedata\\37LAI_Forest_R\\RRI1_Map.tif",
format="GTiff", overwrite=TRUE)
#####RRI2#####
RRI2<- band4/band3
RRI2_values = ref4/ref3
Table <-
data.frame(Shapefile_attributes,NDVI_values,RVI_values,NDVIrededge_values,RVIrededge_values,RRI2
_values)
X <- Table$LAI
Y <- Table$RRI2_values
cor(Y,X)
#####Green Rededge NDVI#####
GRNDVI <- (band4-band2)/(band4*band2)
GRNDVI_values = (ref4-ref2)/(ref4*ref2)
Table <-
data.frame(Shapefile_attributes,NDVI_values,RVI_values,NDVIrededge_values,RVIrededge_values,RRI2
_values,GRNDVI_values)
X <- Table$LAI
Y <- GRNDVI_values
cor(Y,X)
#####RedNDVI#####
REDNDVI <- (band5-band4)/(band5*band4)
REDNDVI_values = (ref5-ref4)/(ref5*ref4)
Table <-
data.frame(Shapefile_attributes,NDVI_values,RVI_values,NDVIrededge_values,RVIrededge_values,RRI2
_values,GRNDVI_values,REDNDVI_values)
X <- Table$LAI
Y <- Table$REDNDVI_values
Cor(Y,X)
#####TVI#####
TVI <- sqrt(ndvi+0.5)
TVI_values = sqrt(NDVI_values+0.5)

```

```

Table <-
data.frame(Shapefile_attributes,NDVI_values,RVI_values,NDVIrededge_values,RVIrededge_values,RR12
_values,GRNDVI_values,REDNDVI_values,TVI_values)
X <- Table$LAI
Y <- Table$TVI_values
cor(Y,X)
#####Enhanced vegetation index(EVI)#####
EVI <- (2.5*((band5-band3)/(band5+6.0*band3-7.5*band1+1)))
EVI_values = (2.5*((ref5-ref3)/(ref5+6.0*ref3-7.5*ref1+1)))
Table <-
data.frame(Shapefile_attributes,NDVI_values,RVI_values,NDVIrededge_values,RVIrededge_values,RR12
_values,GRNDVI_values,REDNDVI_values,TVI_values,EVI_values)
X <- Table$LAI
Y <- Table$EVI_values
cor(Y,X)
#####Perpendicular vegetation index(PVI)#####
PVI <- ((band5-(band3*0.9)-0.1)/((sqrt(1+(0.81))))))
PVI_values = ((ref5-(ref3*0.9)-0.1)/(sqrt(1+(0.81))))
Table <-
data.frame(Shapefile_attributes,NDVI_values,RVI_values,NDVIrededge_values,RVIrededge_values,RR12
_values,GRNDVI_values,REDNDVI_values,TVI_values,EVI_values,PVI_values)
X <- Table$LAI
Y <- Table$PVI_values
cor(Y,X)
#####Atmospherically resistance(ARVI)#####
rb <- (band3-1(band1-band3))
ARVI <- ((band5-rb)/(band5+rb))

rb <- (ref3-1(ref1-ref3))
ARVI_values = ((ref5-rb)/(ref5+rb))
Table <-
data.frame(Shapefile_attributes,NDVI_values,RVI_values,NDVIrededge_values,RVIrededge_values,RR12
_values,GRNDVI_values,REDNDVI_values,TVI_values,EVI_values,PVI_values,ARVI_values)
X <- Table$LAI
Y <- Table$ARVI_values
cor(Y,X)
#####Soil adjusted vegetation index(SAVI)#####
SAVI <- (((band5-band3)/(band5+band3+0.5))*1.5)
SAVI_values = (((ref5-ref3)/(ref5+ref3+0.5))*1.5)
Table <-
data.frame(Shapefile_attributes,NDVI_values,RVI_values,NDVIrededge_values,RVIrededge_values,RR12
_values,GRNDVI_values,REDNDVI_values,TVI_values,EVI_values,PVI_values,ARVI_values,SAVI_val
ues)
X <- Table$LAI
Y <- Table$SAVI_values
cor(Y,X)

```

```
#####GLCM BAND3#####
tex33 <- glcm(band3, n_grey = 32, window = c(3,3), shift=list(c(0,1)), statistics = c("mean", "variance",
"homogeneity", "contrast", "dissimilarity", "entropy", "second_moment"))
LAI33 <- data.frame(extract(tex33,data.frame(d$N,d$E)))
X <- Table$LAI
Y <- LAI33$glcm_mean
cor(Y,X)

tex35 <- glcm(band3, n_grey = 32, window = c(5,5), shift=list(c(0,1)), statistics = c("mean", "variance",
"homogeneity", "contrast", "dissimilarity", "entropy", "second_moment"))
LAI35 <- data.frame(extract(tex35,data.frame(d$N,d$E)))
X <- Table$LAI
Y <- LAI35$glcm_mean
cor(Y,X)

tex37 <- glcm(band3, n_grey = 32, window = c(7,7), shift=list(c(0,1)), statistics = c("mean", "variance",
"homogeneity", "contrast", "dissimilarity", "entropy", "second_moment"))
LAI37 <- data.frame(extract(tex37,data.frame(d$N,d$E)))
X <- Table$LAI
Y <- LAI37$glcm_mean
cor(Y,X)

#####GLCM BAND5#####

tex53 <- glcm(band5, n_grey=32, window=c(3,3), shift=list(c(0,1)), statistics = c("mean", "variance",
"homogeneity", "contrast", "dissimilarity", "entropy", "second_moment"))
LAI53 <- data.frame(extract(tex53,data.frame(d$N,d$E)))
X <- LAI
Y <- LAI53$glcm_mean
cor(Y,X)

tex55 <- glcm(band5, n_grey=32, window=c(5,5), shift=list(c(0,1)), statistics = c("mean", "variance",
"homogeneity", "contrast", "dissimilarity", "entropy", "second_moment"))
LAI55 <- data.frame(extract(tex55,data.frame(d$N,d$E)))
X <- LAI
Y <- LAI55$glcm_variance
cor(Y,X)

tex57 <- glcm(band5, n_grey = 32, window = c(7,7), shift=list(c(0,1)), statistics = c("mean", "variance",
"homogeneity", "contrast", "dissimilarity", "entropy", "second_moment"))
LAI57 <- data.frame(extract(tex57,data.frame(d$N,d$E)))
X <- Table$LAI
Y <- LAI57$glcm_mean
cor(Y,X)

#####GLCM RRI1#####
RRI1_raster3 <- glcm(RRI1_raster, n_grey=32, window=c(3,3), shift=list(c(0,1)), statistics =
c("mean", "variance", "homogeneity", "contrast", "dissimilarity", "entropy", "second_moment"))
```

```

RRI13 <- data.frame(extract(RRI1_raster3,data.frame(d$N,d$E)))
X <- Table$LAI
Y <- RRI13$glcm_mean
cor(Y,X)
RRI13_cvfit = RRI13$glcm_mean
LAI_Measured = Table$LAI
Predicted_LAI=c()
slope=c()
intercept=c()
for(i in 1:37){
  model_RRI13= lm(LAI_Measured[-i]~ RRI13_cvfit [-i])
  slope[i]= model_RRI13$coefficients[2]
  intercept[i] = model_RRI13$coefficients[1]
  Predicted_LAI[i] = model_RRI13$coefficients[2]* RRI13_cvfit [i]+ model_RRI13$coefficients[1]
}
RRI13_predictions = data.frame(slope, intercept, Predicted_LAI, LAI_Measured)
cor(LAI_Measured,Predicted_LAI)^2
rmse(LAI_Measured,Predicted_LAI)
plot(LAI_Measured,Predicted_LAI)
model=lm(Predicted_LAI~LAI_Measured)
#####crating map of LAI using txt linear model#####
#####
image_txt=stack(RRI1_raster3$glcm_mean)

TableRRI1 = data.frame(LAI=d$LAI,
                        glcm_mean=RRI13$glcm_mean)

model_linear_txt = lm(LAI ~ glcm_mean, data=TableRRI1)
Map_linear_txt = predict(image_txt, model_linear_txt)
plot(Map_linear_txt)
writeRaster(Map_linear_txt, filename="D:\\MSc\\Imagedata\\37LAI_Forest_R\\texture_Map.tif",
format="GTiff", overwrite=TRUE)
#####RRI15#####
RRI1_raster5 <- glcm(RVI_raster, n_grey=32, window=c(5,5), shift=list(c(0,1)), statistics = c("mean",
"variance", "homogeneity", "contrast", "dissimilarity", "entropy", "second_moment"))
RRI15 <- data.frame(extract(RRI1_raster5,data.frame(d$N,d$E)))
X <- LAI
Y <- RVI15$glcm_variance
cor(Y,X)

RRI1_raster7 <- glcm(RVI_raster, n_grey = 32, window = c(7,7), shift=list(c(0,1)), statistics = c("mean",
"variance", "homogeneity", "contrast", "dissimilarity", "entropy", "second_moment"))
RRI17 <- data.frame(extract(RRI1_raster7,data.frame(d$N,d$E)))
X <- LAI
Y <- RVI17$glcm_mean
cor(Y,X)

```

```
#####Ratio texture (raster)#####
#####353#####
ratio353mm = tex33$glcm_mean/tex53$glcm_mean
ratio353mv = tex33$glcm_mean/tex53$glcm_variance

ratio353vm = tex33$glcm_variance/tex53$glcm_mean
ratio353vv = tex33$glcm_variance/tex53$glcm_variance

#####533#####

ratio533mm = tex53$glcm_mean/tex33$glcm_mean
ratio533mv = tex53$glcm_mean/tex33$glcm_variance

ratio533vm = tex53$glcm_variance/tex33$glcm_mean
ratio533vv = tex53$glcm_variance/tex33$glcm_variance
#####355#####

ratio355mm = tex35$glcm_mean/tex55$glcm_mean
ratio355mv = tex35$glcm_mean/tex55$glcm_variance

ratio355vm = tex35$glcm_variance/tex55$glcm_mean
ratio355vv = tex35$glcm_variance/tex55$glcm_variance

#####535#####

ratio535mm = tex55$glcm_mean/tex35$glcm_mean
ratio535mv = tex55$glcm_mean/tex35$glcm_variance

ratio535vm = tex55$glcm_variance/tex35$glcm_mean
ratio535vv = tex55$glcm_variance/tex35$glcm_variance

#####Ratio texture (table)#####
#####353#####
tratio353mm = LAI33$glcm_mean/LAI53$glcm_mean
tratio353mv = LAI33$glcm_mean/LAI53$glcm_variance

tratio353vm = LAI33$glcm_variance/LAI53$glcm_mean
tratio353vv = LAI33$glcm_variance/LAI53$glcm_variance

#####533#####
tratio533mm = LAI53$glcm_mean/LAI33$glcm_mean
tratio533mv = LAI53$glcm_mean/LAI33$glcm_variance

tratio533vm = LAI53$glcm_variance/LAI33$glcm_mean
```

```
tratio533vv = LAI53$glcm_variance/LAI33$glcm_variance
#####355#####
```

```
tratio355mm = LAI35$glcm_mean/LAI55$glcm_mean
tratio355mv = LAI35$glcm_mean/LAI55$glcm_variance
```

```
tratio355vm = LAI35$glcm_variance/LAI55$glcm_mean
tratio355vv = LAI35$glcm_variance/LAI55$glcm_variance
#####535#####
```

```
tratio535mm = LAI55$glcm_mean/LAI35$glcm_mean
tratio535mv = LAI55$glcm_mean/LAI35$glcm_variance
```

```
tratio535vm = LAI55$glcm_variance/LAI35$glcm_mean
tratio535vv = LAI55$glcm_variance/LAI35$glcm_variance
#####
```

```
#####Random Forest without importance ranking#####
```

```
image_stack <- stack(band1,
  band2,
  band3,
  band4,
  band5,
  ndvi,
  rvi,
  ndvirededge,
  RVirededge,
  RRI2,
  GRNDVI,
  REDNDVI,
  TVI,
  EVI,
  PVI,
  ARVI,
  SAVI,
  tex33$glcm_mean,
  tex33$glcm_variance,
  tex33$glcm_homogeneity,
  tex33$glcm_contrast,
  tex33$glcm_dissimilarity,
  tex33$glcm_entropy,
  tex33$glcm_second_moment,
  tex35$glcm_mean,
  tex35$glcm_variance,
  tex35$glcm_homogeneity,
  tex35$glcm_contrast,
```

tex35\$glcm_dissimilarity,
 tex35\$glcm_entropy,
 tex35\$glcm_second_moment,
 tex37\$glcm_mean,
 tex37\$glcm_variance,
 tex37\$glcm_homogeneity,
 tex37\$glcm_contrast,
 tex37\$glcm_dissimilarity,
 tex37\$glcm_entropy,
 tex37\$glcm_second_moment,
 tex53\$glcm_mean,
 tex53\$glcm_variance,
 tex53\$glcm_homogeneity,
 tex53\$glcm_contrast,
 tex53\$glcm_dissimilarity,
 tex53\$glcm_entropy,
 tex53\$glcm_second_moment,
 tex55\$glcm_mean,
 tex55\$glcm_variance,
 tex55\$glcm_homogeneity,
 tex55\$glcm_contrast,
 tex55\$glcm_dissimilarity,
 tex55\$glcm_entropy,
 tex55\$glcm_second_moment,
 tex57\$glcm_mean,
 tex57\$glcm_variance,
 tex57\$glcm_homogeneity,
 tex57\$glcm_contrast,
 tex57\$glcm_dissimilarity,
 tex57\$glcm_entropy,
 tex57\$glcm_second_moment,
 RVI_raster3\$glcm_mean,
 RVI_raster3\$glcm_variance,
 RVI_raster3\$glcm_homogeneity,
 RVI_raster3\$glcm_contrast,
 RVI_raster3\$glcm_dissimilarity,
 RVI_raster3\$glcm_entropy,
 RVI_raster3\$glcm_second_moment,
 RVI_raster5\$glcm_mean,
 RVI_raster5\$glcm_variance,
 RVI_raster5\$glcm_homogeneity,
 RVI_raster5\$glcm_contrast,
 RVI_raster5\$glcm_dissimilarity,
 RVI_raster5\$glcm_entropy,
 RVI_raster5\$glcm_second_moment,
 RVI_raster7\$glcm_mean,


```

RVI_raster7$glcm_variance,
RVI_raster7$glcm_homogeneity,
RVI_raster7$glcm_contrast,
RVI_raster7$glcm_dissimilarity,
RVI_raster7$glcm_entropy,
RVI_raster7$glcm_second_moment,
ratio353mm,
ratio353mv,
ratio353vm,
ratio353vv,
ratio355mm,
ratio355mv,
ratio355vm,
ratio355vv,
ratio533mm,
ratio533mv,
ratio533vm,
ratio533vv,
ratio535mm,
ratio535mv,
ratio535vm,
ratio535vv)

```

```

Table = data.frame(LAI=d$LAI,
  Forest_nocloud.1=ref1,
  Forest_nocloud.2=ref2,
  Forest_nocloud.3=ref3,
  Forest_nocloud.4=ref4,
  Forest_nocloud.5=ref5,
  layer.1=NDVI_values,
  layer.2=RVI_values,
  layer.3=NDVIrededge_values,
  layer.4=RVIrededge_values,
  layer.5=RRI2_values,
  layer.6=GRNDVI_values,
  layer.7=REDNDVI_values,
  layer.8=TVI_values,
  layer.9=EVI_values,
  layer.10=PVI_values,
  layer.11=ARVI_values,
  layer.12=SAVI_values,
  glcm_mean.1 = LAI33$glcm_mean,
  glcm_variance.1 = LAI33$glcm_variance,
  glcm_homogeneity.1 = LAI33$glcm_homogeneity,
  glcm_contrast.1 = LAI33$glcm_contrast,
  glcm_dissimilarity.1 = LAI33$glcm_dissimilarity,

```

```

glcm_entropy.1 = LAI33$glcm_entropy,
glcm_second_moment.1 = LAI33$glcm_second_moment,
glcm_mean.2 = LAI35$glcm_mean,
glcm_variance.2 = LAI35$glcm_variance,
glcm_homogeneity.2 = LAI35$glcm_homogeneity,
glcm_contrast.2 = LAI35$glcm_contrast,
glcm_dissimilarity.2 = LAI35$glcm_dissimilarity,
glcm_entropy.2 = LAI35$glcm_entropy,
glcm_second_moment.2 = LAI35$glcm_second_moment,
glcm_mean.3 = LAI37$glcm_mean,
glcm_variance.3 = LAI37$glcm_variance,
glcm_homogeneity.3 = LAI37$glcm_homogeneity,
glcm_contrast.3 = LAI37$glcm_contrast,
glcm_dissimilarity.3 = LAI37$glcm_dissimilarity,
glcm_entropy.3 = LAI37$glcm_entropy,
glcm_second_moment.3 = LAI37$glcm_second_moment,
glcm_mean.4 = LAI53$glcm_mean,
glcm_variance.4 = LAI53$glcm_variance,
glcm_homogeneity.4 = LAI53$glcm_homogeneity,
glcm_contrast.4 = LAI53$glcm_contrast,
glcm_dissimilarity.4 = LAI53$glcm_dissimilarity,
glcm_entropy.4 = LAI53$glcm_entropy,
glcm_second_moment.4 = LAI53$glcm_second_moment,
glcm_mean.5 = LAI55$glcm_mean,
glcm_variance.5 = LAI55$glcm_variance,
glcm_homogeneity.5 = LAI55$glcm_homogeneity,
glcm_contrast.5 = LAI55$glcm_contrast,
glcm_dissimilarity.5 = LAI55$glcm_dissimilarity,
glcm_entropy.5 = LAI55$glcm_entropy,
glcm_second_moment.5 = LAI55$glcm_second_moment,
glcm_mean.6 = LAI57$glcm_mean,
glcm_variance.6 = LAI57$glcm_variance,
glcm_homogeneity.6 = LAI57$glcm_homogeneity,
glcm_contrast.6 = LAI57$glcm_contrast,
glcm_dissimilarity.6 = LAI57$glcm_dissimilarity,
glcm_entropy.6 = LAI57$glcm_entropy,
glcm_second_moment.6 = LAI57$glcm_second_moment,
glcm_mean.7 = RVI3$glcm_mean,
glcm_variance.7 = RVI3$glcm_variance,
glcm_homogeneity.7 = RVI3$glcm_homogeneity,
glcm_contrast.7 = RVI3$glcm_contrast,
glcm_dissimilarity.7 = RVI3$glcm_dissimilarity,
glcm_entropy.7 = RVI3$glcm_entropy,
glcm_second_moment.7 = RVI3$glcm_second_moment,
glcm_mean.8 = RVI5$glcm_mean,
glcm_variance.8 = RVI5$glcm_variance,

```

```

glcm_homogeneity.8 = RVI5$glcm_homogeneity,
glcm_contrast.8 = RVI5$glcm_contrast,
glcm_dissimilarity.8 = RVI5$glcm_dissimilarity,
glcm_entropy.8 = RVI5$glcm_entropy,
glcm_second_moment.8 = RVI5$glcm_second_moment,
glcm_mean.9 = RVI7$glcm_mean,
glcm_variance.9 = RVI7$glcm_variance,
glcm_homogeneity.9 = RVI7$glcm_homogeneity,
glcm_contrast.9 = RVI7$glcm_contrast,
glcm_dissimilarity.9 = RVI7$glcm_dissimilarity,
glcm_entropy.9 = RVI7$glcm_entropy,
glcm_second_moment.9 = RVI7$glcm_second_moment,
layer.13=tratio353mm,
layer.14=tratio353mv,
layer.15=tratio353vm,
layer.16=tratio353vv,
layer.17=tratio355mm,
layer.18=tratio355mv,
layer.19=tratio355vm,
layer.20=tratio355vv,
layer.21=tratio533mm,
layer.22=tratio533mv,
layer.23=tratio533vm,
layer.24=tratio533vv,
layer.25=tratio535mm,
layer.26=tratio535mv,
layer.27=tratio535vm,
layer.28=tratio535vv)

```

```
##### LEAVE ONE OUT OF RF#####
```

```

rmse_one = c()
for(i in 1:37){
  rf <- randomForest(Table$LAI[-i] ~ ., data=Table [-i,2:97],importance=TRUE, ntree=1000, mtry=32,
nodesize=1)
  LAI_predicted[i] <- predict(rf, Table [i,2:97])
  rmse_one[i] = rmse(LAI_predicted[i],Table$LAI[i])
}
mean(rmse_one)
cor(LAI_predicted,Table$LAI)^2
#####training with 37 points and calculation of importance and R2#####
important_vars = c()
for (i in 1:50) {
  rf_predict <- randomForest(LAI ~ ., data=Table,importance=TRUE, ntree=1000, mtry=32,
nodesize=1)
  dd_rf_predict = data.frame (rf_predict$importance)

```

```

first_10_importannt_var = head(row.names(data.frame(dd_rf_predict[order(-
dd_rf_predict$X.IncMSE),])),10)
important_vars = c(first_10_importannt_var,important_vars)
}
sort(table(important_vars))

prediction_37point = predict(rf_predict, Table [2:97])
cor(prediction_37point,Table$LAI)^2
##### creating map #####
Map_LAI = predict(image_stack, rf_predict, progress='window', type ='response')
plot(Map_LAI)

writeRaster(Map_LAI, filename="D:\\MSc\\Imagedata\\37LAI_Forest_R\\LAI_Map.tif",
format="GTiff", overwrite=TRUE)

```