GIMA
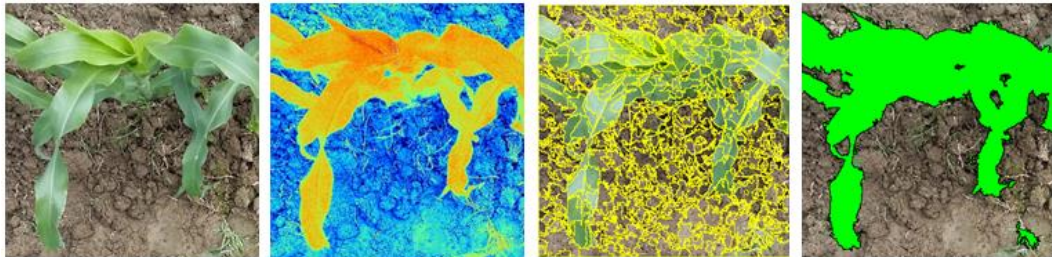## Geographical Information Management and Applications

Measuring green vegetation cover over agricultural fields: a multi-scale study using smartphones and UAV

Master's Thesis

Student: Nikolina Mileva
E-mail: mileva@student.utwente.nl
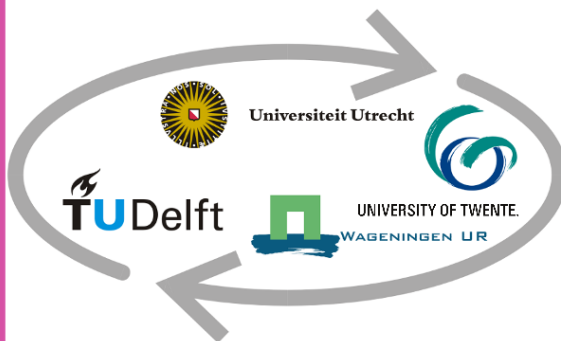Student numbers:
4291662 (UU)
s6020445 (ITC)
890110-572-130 (WUR)

Supervisor: Dr. Raul Zurita-Milla
Professor: Prof. Menno-Jan Kraak
Advisors: Dr. Rolf de By and
Dr. Emma Izquierdo-Verdiguier

Universiteit Utrecht
TUDelft
UNIVERSITY OF TWENTE.
WAGENINGEN UR

Summary

The fraction of green vegetation cover (fCover) is an important parameter having variety of applications in the field of agriculture and forestry. It refers to the percentage of ground covered with photosynthetically active green parts of a plant. Many methods can be used for the estimation of fCover such as statistical methods relying on vegetation indices, physical methods using the inversion of canopy reflectance models and other. fCover estimates can be retrieved using various instruments and platforms ranging from low cost, consumer grade cameras to professional spectrometers and from unmanned aerial vehicles to satellites.

The main objective of this thesis is to study the variation in fCover when estimated at various spatial resolutions, with various instruments and over different extents. Estimates of fCover derived from smartphones and from UAV images were compared. The various estimations were evaluated to better understand the impact of geographic scale on fCover as well as the added value of images with a NIR band. With this analysis, it was possible to estimate the extent with which field measurements agree with the average fCover of a field which can help to design improved fCover field sampling strategies.

Next to this, classification and regression methods to estimate fCover were compared. For the classification method, an object-based image analysis approach was adopted which is believed to outperform pixel-based classification techniques in the analysis of high resolution imagery (Blaschke et al., 2014). In the regression method, various vegetation indices and their statistics were used as predictors which allowed the most important color and NIR indices to estimate fCover to be identified.

The results showed that minor differences in the resolution of smartphone photos coming from the device models and camera lenses do not have major impact on the fCover estimates. But using too coarse spatial resolution as obtained from UAVs at 65 m altitude can lead to substantial over- or underestimation of fCover. When possible, it is better to use higher resolution photos as obtained from high-end smartphone models because these contribute to the more accurate estimation of fCover. The estimation of fCover based on UAV images also showed that color indices obtained at 12 m altitude were better predictors for fCover than vegetation indices in the near-infrared region obtained at 100 m height.

Comparing the two approaches for estimating fCover revealed that both regression and classification algorithms can poorly deal with crops whose leaves are not well defined in respect to the background. Both algorithms had some difficulties estimating fCover when the sun reflection was strong. But the regression algorithm could deal much better with changing illumination when no sun glint was involved.

The appropriate number of field samples to be taken depends on the accuracy of fCover sought. Surveying as little as 5% of a field, gave very good fCover estimate that did not differ from the actual fCover with more than 4%.

Acknowledgement

Contents

1. Introduction and research objectives

Estimating vegetation properties from remotely sensed images dates back to the Cold War when the USA used satellites to forecast wheat crop yield in the USSR and thus predict production shortfalls (Logsdon et al., 1998). Crop yield is still an important factor in national security because the failure of a country to produce sufficient quantities of agricultural products can lead to violent conflicts and mass migration (Campbell et al., 2007). This is especially true for developing countries in sub-Saharan Africa and South Asia not having the financial resources to import agricultural products and suffering from climate changes (Schmidhuber & Tobiello, 2007). Nowadays the access to high resolution imagery is not limited only to reconnaissance authorities. Satellite images with fairly high resolution such as Landsat-7 or Landsat-8 with 30m resolution in the visible and near-infrared bands and Sentinel-2 offering 10m resolution in the same bands are freely available. Farmers can use this information for predicting crop yield, monitoring diseases, planning irrigation and fertilizers application (Bastiaanssen et al., 2000; Seelan et al., 2003). Humanitarian organizations can also make use of it by identifying regions threatened by food shortages and thus planning in advance counter measures (Kaiser et al., 2003).

Besides these traditional remote sensing platforms, an increasing amount of data is being collected with unmanned aerial vehicles (UAVs) and smartphones making possible the estimation of vegetation properties for smaller agricultural fields and at better spatial resolution (Anderson & Gaston, 2013). UAVs do not only offer the possibility to collect higher resolution imagery than satellites, they have other advantages as well such as collecting data at different angles (Ren et al., 2017) and in cloudy weather conditions (Bondi et al., 2016). For the purposes of precision agriculture, topographic and atmospheric corrections are often not needed making the use of UAV images less laborious than the use of satellite images (Jakob et al., 2017). UAVs are able to carry various instruments ranging from hyperspectral to simple consumer grade RGB cameras (Salamí et al., 2014). Smartphones are more limited with respect to the cameras that they can carry but they might be equipped with sensors such as gyroscopes and inclinometers collecting information about the position, motion and surrounding environment (Pongnumkul et al., 2015). These sensors combined with the processing power of smartphones and dedicated applications can convert smartphones into effective tools for collecting vegetation properties information that might even replace more traditional field instruments. An example, the smartphone application PocketLAI can acquire LAI estimates comparable to those acquired with classical instruments such as LAI-2000 and digital cameras for hemispherical photography (Campos-Taberner et al., 2015).

Satellites, UAVs and smartphones allow the estimation of vegetation properties at different spatial scales. The spatial extent used in the domain of agriculture can vary widely – from 1 m^2 (the area that usually a digital camera or a smartphone photo covers when used to take nadir looking photos of vegetation at an operator height), to the size of a field (again varying widely is size) and beyond. And, while it is often assumed that the better spatial resolution enables the more accurate estimation of vegetation properties, there are many other factors that play a role when selecting the right scale (Woodcock & Strahler, 1987). When estimating vegetation properties of a field using field measurements, appropriate field sampling strategy has to be employed. In general, the more the samples within the field are, the more accurate the vegetation parameter estimations of a field are. But

a balance has to be found between the number of field measurements and the level of accuracy needed (Wang et al., 2015; Gharun et al., 2017).

With the increased amount and availability of data, new methods are needed for its interpretation. While simple statistical techniques such as linear regression are still common, more sophisticated machine learning algorithms have gained popularity in recent times because of their good performance. However, this abundance of instruments, data and methods can be overwhelming. Therefore, this Master's thesis aims at comparing vegetation properties estimated using various instruments (smartphones and UAV images taken at different heights) and analytical methods. This thesis focuses on the fraction of green vegetation cover (fCover) as it is an important vegetation parameter having variety of applications in agriculture as discussed in the next chapter.

1.1. Specific research objectives and research questions

1.   Analyze the effect of measuring fCover at various spatial resolutions

1.1 What are the differences in fCover estimates when using smartphones with different resolution (and camera lenses)?

1.2 What is the error of fCover estimates derived from low-end smartphones when compared to the estimates derived from a high-end smartphone model?

1.3 What are the differences in fCover estimates of a field when using a UAV operated at multiple heights?

2.   Study the influence of the method and of the available spectral features on the fCover values

2.1 What are the differences in fCover estimates when a regression approach is used?

2.2 Which of the color indices mentioned in the methodology chapter is the most important for the estimation of fCover?

2.3 Which of the NIR indices mentioned in the methodology chapter is most important for the estimation of fCover?

2.4 What are the differences in fCover estimates when a classification approach is used?

3.   Model the impact of field sampling on fCover estimates

3.1 How many smartphone photos, as simulated using UAV images, are needed for a statistically robust characterization of the fCover value of a field?

2. Literature review

The first section of this chapter contains an overview of the areas of application of fCover predominately in the field of agriculture. The next section outlines different methods used for estimating bio-physical variables with emphasize on machine learning methods and object-based image analysis. In the third section, the different types of platforms and instruments for collecting remotely sensed data are described. The last two sections introduce the problems of selecting proper scale and field sampling strategy.

2.1. Applications of fCover

fCover has various applications, and many of these applications are in the field of precision agriculture. Precision agriculture refers to agricultural practices that aim at maximizing yield, minimizing inputs and/or improving the quality of the agricultural products (McBratney et al., 2005). What is specific for precision agriculture is the use of technologies such as the global positioning system (GPS), geographic information systems (GIS), various sensors, remote sensing (Seelan et al., 2003). Remotely sensed images have a variety of applications in precision agriculture – they are used for mapping soil moisture content, determining the phenological state of crops, monitoring crop growth, nutrient deficiency and evapotranspiration rate, identifying crop diseases, assessing weed and insect infestation (Moran et al., 1997). Based on remotely sensed images an estimate of the radiation reflected by the crops can be made which can indirectly give us insights into different bio-physical properties of the vegetation such as chlorophyll and nitrogen content, biomass, etc. (Mulla, 2013). Reflectance can also be used for calculating various vegetation indices which do not only relate to the bio-physical properties of crops but serve as inputs in agroecosystem models too (Dorigo et al., 2007). One of these bio-physical parameters is fCover.

According to Zeng et al. (2000), fCover was first used to predict ground surface moisture content in the atmospheric model of Deardorff (1978) where it was called "fractional foliage surface". It refers to "the fraction of the green vegetation in the nadir direction" (Li et al., 2015). fCover has many applications in the field of agriculture - ranging from irrigation (Calera et al., 2001; Er-Raki et al., 2010; El Hajj et al., 2016) and crop residues management (Bannari et al., 2007; Daughtry et al., 2005) to yield estimation (Yang et al., 2010; Castaldi et al., 2015), but also in other domains such as forestry (Berberoglu et al., 2009; Wu, 2011; Pfeifer et al., 2016; Jing et al., 2011) and climatology (Zeng et al., 2000). It often outperforms classical vegetation indices such as NDVI as estimator for yield variation (Yang et al., 2010) and soil erosion (Zhongming et al., 2010).

One of the applications of fCover in agriculture is its use for the estimation of soil water evaporation. Soil water evaporation is negatively correlated to fCover because the lower the fCover the higher the evaporation (Er-Raki et al., 2010). fCover is essential for the calculation of the crop coefficient used in the methodology for estimating crop water requirements of the Food and Agriculture Organization of the United Nations (Calera et al., 2001; Er-Raki et al., 2010). In this methodology, fCover is originally referred to as ground cover (Doorenbos & Pruitt, 1977). fCover has also been used in models for soil moisture estimation (El Hajj et al., 2016).

fCover can also be used to estimate the yield variation of different crop types such as sorghum and cotton (Yang et al., 2010) or wheat (Castaldi et al., 2015). Yang et al. (2010) use linear spectral unmixing method with two endmembers – plant and soil, where the reflectance spectrum of each pixel can be represented as a linear function of the endmembers. Both the constrained (ranging between 0 and 1) and the unconstrained plant abundance fraction calculated using this methodology show more satisfactory results than NDVI as estimators for yield variation.

fCover can be calculated for post-harvest agricultural fields to quantify the amount of crop residues (Bannari et al., 2007, Daughtry et al., 2005). The tillage performed in agricultural fields accelerates soil erosion and decreases the content of carbon in soil. Crop residues on agricultural land can prevent these processes, thus crop residue fraction cover is an important measure in conservation tillage systems (Daughtry et al., 2005). fCover has been widely used in soil erosion models not only for agricultural fields but for diverse land cover types (Wang et al., 2002; de Asis & Omasa, 2007; Vrieling et al., 2008; Zhou et al., 2008; Zhongming et al., 2010; Panagos et al., 2016). The stratified vegetation cover index proposed by Zhongming et al. (2010) shows more satisfactory results in explaining soil erosion than NDVI. It measures the vegetation cover in relation to vegetation type and structure and is thus closely related to fCover.

Besides agriculture, fCover is also used in forestry. Berberoglu et al. (2009) used fCover together with other biophysical variables as predictors for the percentage tree cover in pure and mixed conifer forests - an important indicator for forests and woodlands assessment (Wu, 2011). Pfeifer et al. (2016) used fCover for assessing tropical forests degradation. Jing et al. (2011) used the relationship between vegetation coverage and precipitation to model natural vegetation restoration in mountainous areas. fCover can be applied also for ecological assessment and monitoring (Godínez-Alvarez et al., 2009) and in variety of models such as climate models (Zeng et al., 2000) and models for estimating grassland production (Roumiguié et al., 2016).

Despite the wide use of fCover, the multiple terms being used to describe it are not yet standardized which makes the comparison of studies using this bio-physical parameter difficult (Godínez-Alvarez et al., 2009). In literature, fCover can be often regarded to as (fractional/fraction of/percentage of) (green) foliage, ground, plant, vegetation or canopy cover. Depending on the purpose for which it is used, cover may refer only to certain types of vegetation such as tree and shrub canopies, only to green vegetation, to all types of vegetation including litter or in the case of ground cover also to non-vegetation material such as rocks (Bonham, 2013). fCover can be expressed in different ways - as percentage, proportion or using a categorical scale (Bonham, 2013). In this thesis, fCover is expressed as percentage and refers only to the photosynthetically active green parts of a plant.

2.2. Methods for estimating fCover

Various methods exist for estimating fCover. Some of them rely on visual interpretation while other use remote sensing observations (Bonham, 2013). In this thesis, the emphasis is put on methods using the latter approach. For visual interpretation methods the reader is referred to the book of Bonham (2013).

Remote sensing observations do not directly provide insights into the biophysical or biochemical characteristics of the surface being sensed, they only quantify the radiation field reflected or emitted by it. An intermediate step is required to convert the measured radiation into vegetation products (Baret & Buis, 2008). The methods used to perform this step can broadly be classified into three groups – physical, statistical, and hybrid (Liang, 2005).

2.2.1. Statistical methods

Statistical methods rely mainly on vegetation indices (VIs; Liang, 2005). VIs are used for describing the relationship between the reflectance measured by a sensor and a parameter of interest (North, 2002). Ideally, they have high sensitivity to the parameter of interest and are insensitive to perturbing factors such as soil color change and atmospheric effects (Verstraete & Pinty, 1996). Various VIs have been used to estimate fCover. While NDVI is the most popular of them (Samani Majd et al., 2013; Zhang et al., 2013; Adhikari et al., 2016), it is not always the best choice. For example, NDVI is suitable for estimating fCover in savannas (Wu, 2011), pecan orchards (Samani Majd et al., 2013) and arid and semi-arid areas (Zhang et al., 2013) but for tropical landscapes the reduced simple ratio index is better suited (Adhikari et al., 2016).

Some VIs are based only on the red, green and blue spectral bands. These color indices have the advantage that they are more intuitive to humans as human eyes can actually "see" these colors (Meyer & Neto, 2008). Color indices such as the Excess Green Index (ExG; Woebbecke, et al., 1995) and the Normalized Difference Index (NDI; Perez et al., 2000) are particularly suitable for estimating fCover because they allow to easily differentiate between plants and other background by using an appropriate threshold value. Color indices are less sensitive than NIR indices to the lighting conditions and can be adjusted to different backgrounds (Campbel, 1996 as cited in Meyer & Neto, 2008). Although vegetation is well distinguished in the NIR band due to the strong reflection in the latter, color indices still offer an alternative way to estimate fCover, when information in the NIR band is missing.

Statistical methods based on VIs are computationally efficient and suitable for application at a regional scale. However, as they are specific to the vegetation type, they are not always suitable for application at a larger scale (Jia et al., 2015).

The soil line concept is also a statistical method used to estimate bio-physical variables. Soil pixels plotted in the red and near infra-red spectral space are positively correlated and form a line referred to as soil line. When plotted together with vegetation, the soil line represents the base of a triangle with the points above it showing vegetation (Liang, 2005). The soil line depends on variety of factors such as soil type, moisture content, organic matter content and other. Thus, it is not possible to retrieve global soil line but it has to be calculated for each particular case (Fox et al., 2004).

Another method that can be associated with the group of statistical methods is spectral unmixing. Spectral unmixing is a technique where the spectral reflectance of a scene is presented as a weighted sum of the reflectance spectra of its components (North, 2002). At least two components are needed – for example, plant and soil (Yang et al., 2010). But it is possible to include more endmembers – soil, green vegetation and senescent vegetation (North, 2002) or plant, shaded soil and illuminated soil (Gilabert et al., 2000).

2.2.2. Physical and hybrid models

Physical models are based mainly on inverting canopy reflectance models but soil and snow reflectance models can be used as well (Liang, 2005). These models simulate the radiation field emitted or reflected by the surface for a given type of vegetation. To solve the opposite problem – i.e. to retrieve the biophysical variables from the remote sensing measurements of the radiation field, the model should be inverted. Traditional inversion methods rely on repeated model runs until the total deviation between actual data and model predictions is minimized (Kimes et al., 2000; Liang, 2005). Whether the inversion is successful heavily depends on the initial values provided to the reflectance model. If these values differ substantially from the "true" values, iterating the model may not lead to convergence. Thus, it is often necessary to provide several sets with initial values before the best solution can be selected increasing even more the computational time needed (Liang, 2005).

The main advantage of physical models is that training datasets simulated by inverse models cover more possible combinations of conditions than training datasets based on empirical data (Baret & Buis, 2008). But they also have drawbacks – radiative models have model uncertainties due to the simplifications of the canopy structure and uncertainties originating from the radiometric measurements; the distribution of the variables in the radiative models is poorly known (Baret & Buis, 20087). These uncertainties make the inversion problem ill-posed – a set of different solutions can correspond to similar reflectance values, and the use of prior information is often required (Combal et al., 2003). However, as such information is not always available other approaches based on contextual and temporal information are gaining popularity (Zurita-Milla et al., 2015). Physically based models are also computationally expensive and therefore not suitable for application on regional and global scale. This limitation can be overcome by employing a generic algorithm which mimics the natural selection during evolutionary processes when selecting suitable set of initial variables for the reflectance model (Liang, 2005). Another alternative are look-up tables which store precomputed radiative transfer functions and thus allow for fast and accurate radiometric processing (Rebordão, 1989; Mobley et al., 2005). Hybrid methods are also computationally more efficient than physical models (Liang, 2005). In hybrid methods, a radiative transfer model is used for the creation of a training dataset and a non-parametric statistical model is applied for describing the relationship between the reflectance and the bio-physical variable. Machine learning techniques such as artificial neural networks are often used (e.g. Baret et al., 2007; Jia et al., 2015).

2.2.3. Other methods

Machine learning techniques can be applied not only in the context of radiative transfer models but also directly for retrieving bio-physical variables from reflectances (e.g. Verrelst et al., 2012). The estimation of fCover can be considered as a classification problem with two classes – green vegetation and background, or as a regression problem where fCover is a dependent variable. Machine learning techniques are well suited for addressing both types of problems. In machine learning, it can be differentiated between various learning types such as rote learning, learning from instruction, learning by analogy and learning from example. Learning from example encompasses supervised and unsupervised learning each of them trying to find a general rule, which can explain the data based on a sample of it. In supervised learning, a training sample is available, in which both the dependent and independent variables are known (input-output pairs). Based on the training sample, a function is derived which can map an input to an output minimizing the error. Depending on the type of the output, supervised learning can be approached as regression or classification problem. In a classification problem, the output is a discrete class or category, while in a regression problem it is a value of a continuous variable (Camastra & Vinciarelli, 2015). Although machine learning algorithms are well suited to solve this type of problems, they are often considered to be black-box models (Lary et al., 2016).

Many of the methods mentioned so far and in particularly classification can be applied both on pixel and object level. When the estimation of bio-physical variables is carried out on a pixel level, information such as texture, context and shape is often neglected (Hay & Castilla, 2008). Object based image analysis (GEOBIA) techniques make use of this information by exploring the characteristics of features in the images and the relationships between them (Blaschke, 2004). Such techniques have already been used for the estimation of various vegetation parameters, fCover being one of them (e.g. Laliberte et al., 2007; Bauer & Strauss, 2014). In GEOBIA, it is argued that pixels are not always the best spatial unit when the mapping of landscape elements is concerned (Addink et al., 2012). In a typical GEOBIA workflow, an image is first segmented and then classified (Blaschke, 2004). During the segmentation step, similar pixels are grouped together into objects (Addink et al., 2012) or "objects candidates" – homogenous and semantically meaningful groups of pixels (Blaschke, 2010).

Segmentation techniques have been widely used in remote sensing long before GEOBIA emerged as a paradigm (Blaschke, 2010). They can be divided generally into four groups – point-based, edge-based, region-based, and combined (Schiewe, 2002), using the discontinuity or similarity of the pixel's grey level values (Addink et al., 2012). In the context of image analysis, point-based approaches use the homogeneity between pixels and by applying a threshold divide the image in two or more categories. They are not particularly suitable for remotely sensed data because the reflectance values of an object may vary depending on the location. In edge-based approaches, the segments are defined by their outlines. In region-based approaches, the similarity between the pixels or already-existing regions is compared resulting in splitting or merging existing segments (Schiewe, 2002).

The image objects created during the segmentation step may correspond to real world objects such as individual buildings and tree crowns but this is not always possible often due to the mismatch between the spatial resolution of the image and the size of the object of interest (Addink et al., 2012). Image segmentation can be performed on multiple levels and scales (Addink et al., 2012). The scale depends on

the object of the interest which can heavily range in size - from vegetation patches to more generalized classes such as forest and agricultural area (Benz et al., 2004). At each scale different features might be revealed (Blaschke, 2010). After the image has been segmented, it can be further processed by considering additional object properties such as size and shape allowing the interpretation and classification of the image (Addink et al., 2012) or in other words the derivation of "meaningful objects" (Blaschke, 2010).

The first commercially available software specially dedicated to GEOBIA is eCognition (Flanders et al., 2003; Blaschke, 2010). It was followed by other software developments such as Feature Analyst available as a plug-in for ArcGIS and ERDAS Imagine (Opitz & Blundell, 2008) and ENVI Feature Extraction (Hölbling & Neubert, 2008 as cited in Blaschke, 2010). Free open source solutions are available as well. Examples are the object-based image segmentation implemented in SAGA GIS (Böhner et al., 2006), the Orfeo Toolbox available as a plug-in for QGIS (Inglada & Christophe, 2009) and InterIMAGE (Pahl et al., 2008). There are also stand-alone scripts offering GEOBIA workflows which make use of dedicated R and Python packages such as RSGISLib (Clewly et al., 2014).

GEOBIA techniques often outperform traditional pixel-based classifiers when high resolution images are concerned not only in accuracy but also in computational efficiency (Hay & Castilla, 2008). However, segmentation is an ill-posed problem, so multiple solutions exist (Blaschke, 2004; Hay & Castilla, 2008).

2.3. Instruments and platforms for estimating fCover

fCover estimates can be retrieved using various instruments and platforms ranging from low cost, consumer grade cameras to professional spectrometers and from unmanned aerial vehicles to satellites.

2.3.1. Digital cameras and smartphones

Digital cameras are often used to take in-situ ground pictures usually only in the visible bands of the spectrum. Both hemispherical and nadir looking photographs are suitable as they all provide similar fCover estimates (Mougin et al., 2004). For taking nadir looking photographs simple consumer grade cameras can be used (e.g. Meyer et al., 2004). Hemispherical photographs require fish eye lenses (Jonckheere et al., 2004).

Smartphones can also be used for taking high quality photographs. They are useful for vegetation properties estimation not only because of their high resolution camera but also because of other embedded sensors collecting information about the position, motion and surrounding environment (Pongnumkul et al., 2015). Smartphone photos have been used in a variety of agricultural applications such as disease detection and diagnosis, soil study, determining ripeness of fruits, fertilizer application and irrigation (Pongnumkul et al., 2015). They are not only compact, accessible and computationally efficient (Aquino et al., 2015) but in combination with dedicated applications can also enable the direct estimation of bio-physical variables (Campos-Taberner et al., 2015; Bauer et al., 2016; Qu et al., 2016; Vesali et al., 2016) and even provide an all-round farm management information system (Welte et al., 2013). Vegetation properties such as the leaf area index (LAI) and chlorophyll content derived from smartphone photos are comparable to those estimated with specialized instruments (Campos-Taberner

et al., 2015; Orlando et al., 2016; Qu et al., 2016; Vesali et al., 2017). Smartphones are a promising tool that may replace traditional instruments for field measurements (Campos-Taberner et al., 2015). This is possible both due to the high resolution camera that many of them have and to other built-in sensors such as gyroscopes and inclinometers which allow the measurement of the angular velocity and the angles of slope (Campos-Taberner et al., 2015; Qu et al., 2016). However, more research is needed to evaluate the impact of the smartphone model and the type of vegetation when smartphone photos are used for the estimation of vegetation parameters (Campos-Taberner et al., 2015). Another concern is the location accuracy provided by smartphone devices. Many of the agricultural activities in which smartphones are involved are location based. However, the actual locational accuracy of the smartphone devices has to be carefully tested before they can be used for making any agricultural decisions (Seyyedhasani et al., 2016).

2.3.2. Satellites and their instruments

Satellite and areal images are replacing costly field surveys for collecting crop status information. The increasing spatial and temporal resolution of space-borne and airborne remote sensors enables the global monitoring of various crop characteristics in a timely manner (Somers et al., 2009). Global estimate of fCover is already provided by Copernicus Global Land Service at 1 km resolution based on images from the Proba-V satellite which is equipped with the Vegetation imager instrument. With the launch of Sentinel-2 in 2015 with the MultiSpectral Instrument (MSI) on board, images in the visible and near-infrared bands are freely available at 10 m resolution making possible the estimation of fCover at even better spatial scale. Multispectral images with near a meter ground sampling distance are available upon payment from commercial satellite images providers such as Digital Globe. However, when validating remotely sensed data, field surveys are inevitable (Korhonen et al., 2006). For example, in order to validate satellite products derived from medium resolution sensors such as MODIS, 40-50 photographs per 1km^2 are needed. Satellite observations still remain the only feasible way of obtaining fCover at regional and global scales (Xiao et al., 2016).

2.3.3. Unmanned Aerial Vehicles (UAVs)

An even higher spatial resolution (in the order of cm) can be derived by using unmanned aerial vehicles (UAVs). UAVs can offer better temporal frequency than conventional satellites and airborne platforms (Anderson & Gaston, 2013) because their operation is not affected by cloud coverage; UAVs are able to collect images at various viewing angles (Ren et al., 2017) and in various weather conditions such as complete cloud cover (Bondi et al., 2016). Another advantage over satellite and airborne platforms is that for the purposes of precision agriculture, topographic correction is often not needed because most agricultural areas are flat and have similar geographic features (Jakob et al., 2017). UAV images have been used in variety of agricultural applications such as modelling canopy structure, ripeness monitoring, water stress detection, estimation of nitrogen level, pathogen detection, aerobiological sampling, plant health monitoring, mapping invasive weeds, monitoring herbicide application (Salamí et al., 2014). For many of these applications, it is necessary to estimate various vegetation parameters (Salamí et al., 2014). UAV images used to estimate vegetation properties such as canopy height and aboveground biomass have shown results comparable to field measurements (Li et al., 2016). Some vegetation

properties such as nitrogen amount might be even more accurately estimated from UAV images rather than specialized field instruments due to the fact that soil and shadows are often eliminated in UAV images (Agüera et al., 2011).

The sensors mounted on UAVs range from simple RGB cameras (e.g. Burkart et al., 2017) to multispectral (e.g. Navia et al., 2016) and hyperspectral cameras (e.g. Zheng et al., 2016). However, for monitoring vegetation red, green and near-infrared channels are most commonly used (Salamí et al., 2014). Many of the cameras used are commercial off-the-shelf devices that could be easily upgraded so that information in the NIR band is collected (Salamí et al., 2014). The spatial resolution of the cameras usually ranges between 5 and 12 megapixels (Salamai et al., 2014). Commonly used vegetation indices can be calculated without additional correction to reflectance of the UAV images and with low signal to noise ratio (Jakob et al., 2017). However, when digital numbers are used instead of calibrated reflectance values, the results might be only qualitative and not quantitative (Candiago et al., 2015). UAV data also requires different and often more complex pre-processing than data collected by aerial and space platforms. Despite the wide use of UAVs in different domains, the geometric and radiometric correction of such data is still demanding (Jakob et al., 2017).

2.4. The importance of scale

The instruments and platforms listed in the previous sections allow the estimation of vegetation properties at various scales. Scale refers not only to the spatial resolution but also to the spatial extent (Goodchild & Procter, 1997) and between the two of them there is often a trade-off: when the extent increases, the resolution usually decreases (Glenn et al., 2014).

It is often assumed that better spatial resolution contributes to the more accurate estimation of bio-physical variables. However, multiple factors have to be taken into account when choosing the best spatial resolution (Woodcock & Strahler, 1987). Woodcock & Strahler (1987) suggest a method based on the spatial structure of the image for determining the optimal relationship between spatial resolution and the environmental characteristics of interest. Their method assesses the spatial structure of images by using local variance which is measured by the standard deviation averaged over a moving window. If the size of the pixel is similar or smaller to that of the object of interest, the variance is high. If on the other hand the object of interest is smaller than the pixel size, the variance is low. Thus, a rapid change in variance can show when the optimal scale is reached.

Several authors have pointed out that vegetation indices derived from UAV images are better estimates for vegetation parameters than Gaofen and Landsat images (Gao et al., 2016; Lukas et al., 2016) and comparable to field measurements (Rasmussen et al., 2016; Li et al., 2016) or even better than them (Agüera et al., 2011; Balota & Oakes, 2016). Higher resolution satellite images such as those provided by RapidEye have performance similar to ground-based optical sensors (Bu et al., 2017) and UAVs and aircrafts (Matese et al., 2015) when used for estimating vegetation properties. These results might be due to the fact that the pixel size provided by these platforms is similar to the size of the vegetation properties and thus better able to capture the variability in their values. However, the correlation between the vegetation indices and the bio-physical variables of interest depends not only on the pixel size but also on the bio-physical variables themselves, the vegetation indices used for their estimation,

the phenological stage and the vegetation type (Lukas et al., 2016; Matese et al., 2015). Thus for example, vegetation indices derived from Landsat images were better estimates for nitrogen content than indices derived from UAV images, while for other parameters such as biomass and nitrogen uptake the opposite was true (Lukas et al., 2016). Davidson & Csillag (2016) found stronger relationship between biomass and various vegetation indices at coarser spatial resolution. This phenomenon might be explained by the fact that at coarser scales the variance between the dependent and independent variables is lower, while the covariation between the two increases (Davidson & Csillag, 2001). Woodcock & Strahler (1987) argue that the technique used to derive information from images also plays an important role when selecting a suitable scale. So, the fact that better estimates of bio-physical variables are derived from coarser resolution imagery might be due also to the limited ability of the widely used pixel-based methods to deal with high resolution data (Hay & Castilla, 2008).

2.5. The importance of field sampling strategy

Comparing vegetation properties across different scales resembles the problem of validating vegetation products derived from satellite imagery with field measurements. In each of the cases, a sampling strategy has to be proposed as the extents of the study area at the different scales usually do not overlap. However, when designing a sampling strategy, the spatial resolution of the final product has to be taken into account (Woodgate et al., 2012). While at a coarser spatial resolution (1 km^2), various sampling strategies exist (e.g. Baret et al., 2005), this is not the case for higher spatial resolution products (Chen et al., 2016). Chen et al. (2016) studied the impact of sampling for estimating fCover over alpine grassland using UAV images of 30x30 m plots and varying number of 0.5x0.5 m quadrats within the plots. They suggested that the larger the number of quadrats is, the better the fCover estimate over the field is. Godínez-Alvarez et al. (2009) suggest 70 m long transects with at least 10 m spacing between them and fCover to be recorded every 1 m when line-point intercept method is used; when grid-point intercept method is used, fCover is to be recorded at 1 m^2 quadrats every 14 m. Though, their sampling strategy is developed for fCover studies based on visual interpretation and not imagery analysis. Balance should be achieved between the costs for field sampling and the level of precision sought (Wang et al., 2015; Gharun et al., 2017). However, more sampling plots are needed when the variance of the modeled parameter is high (Gharun et al., 2017). The optimal sampling size can be derived also based on statistical features or on some prior knowledge (Wang et al., 2015).

3. Study area, Data and Methods

3.1. Study area

The study area is located in Bangladesh which is characterized by sub-tropical humid climate (Fig. 1). There are four seasons – dry winter season lasting from December to February, pre-monsoon hot summer season lasting from March to May, rainy monsoon season lasting from June to September, and post-monsoon autumn season lasting from October to November (Shahid, 2010). Maize and wheat are the most popular crops after rice but while maize can be grown both during the winter as well as the summer season (Rahman et al., 2012), wheat can be grown only during the dry winter season (O'Callaghan et al., 1994). Mung beans on the other hand are usually grown in the post-monsoon dry season (Afzal et al., 2006). Various crops were sampled during the field surveys such as wheat, maize and mung beans (Fig. 2, Fig. 3). The field surveys were conducted during the cropping seasons of 2014, 2015 and 2016. For the purposes of this Master's thesis only data from 2016 was used. The field surveys took place in Barisal, Kalapara and Patuakhali located in the south part of the country (Fig. 1).
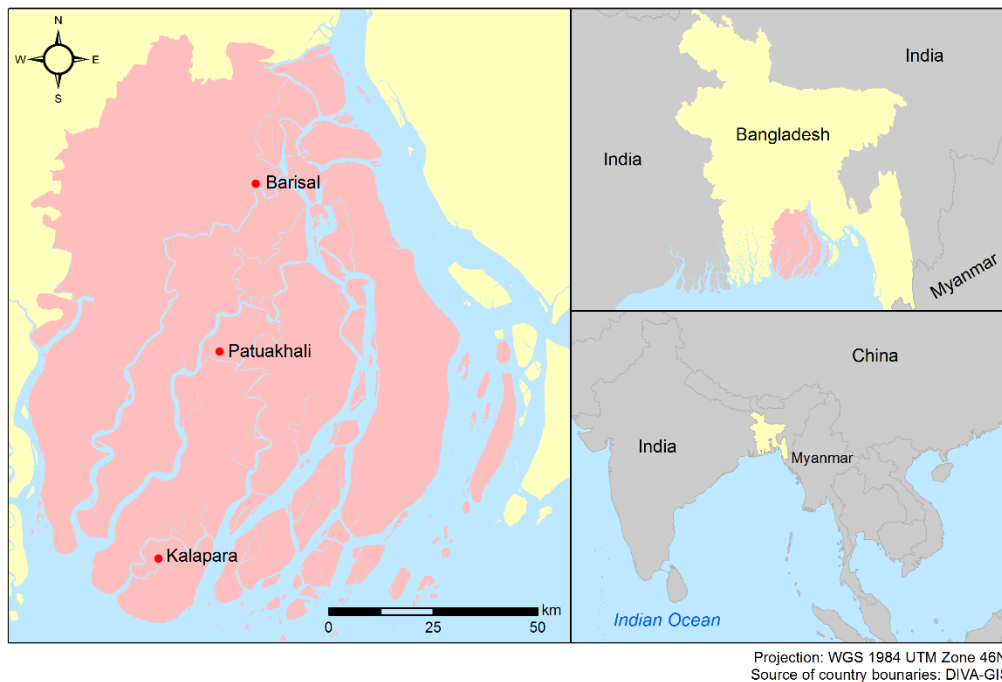
Fig. 1. Map showing the three study sites – Barisal, Kalapara and Patuakhali, located in the administrative division of Barisal (depicted in light red)

Projection: WGS 1984 UTM Zone 46N
Source of country bounaries: DIVA-GIS

3.2. Data

Several fields were visited and photos were collected using smartphones and GeoX-8000 octocopter. During the field campaign in 2014, only one smartphone model was used to collect RGB images. In 2015 and 2016, various devices were used – Lava, Okapia, Samsung Galaxy Grand Prime (shortly Samsung Galaxy), Symphony smartphones, a Samsung tablet and other (Fig. 2). However, some of the photos were taken only with certain devices and not with all smartphones available. These photos were excluded from the dataset so that each device type is represented by an equal number of photos. Thus, out of 292 photos taken in 2016 only 165 were used for the analysis. Each of the pictures was taken

vertically above ground and covered approximately 1 m^2 as measured by a square frame placed on the ground during the field surveys (Fig. 4).
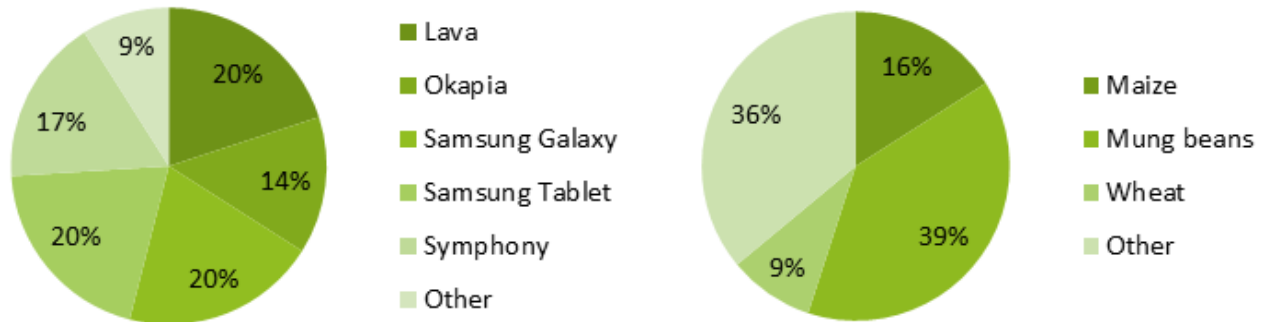
Fig. 2. Percentage of smartphone photos per device and crop type for the field campaign in 2016
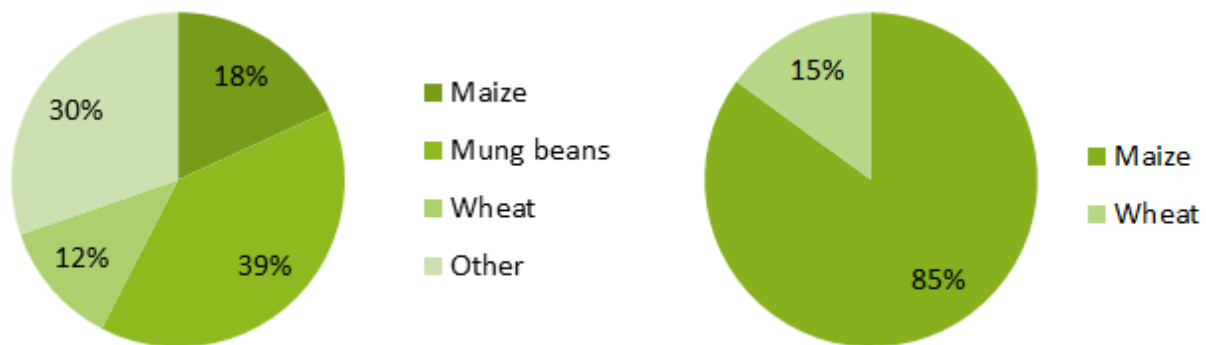
Fig. 3. Percentage of photos per crop type from smartphones (left) and UAV (right) used for analysis
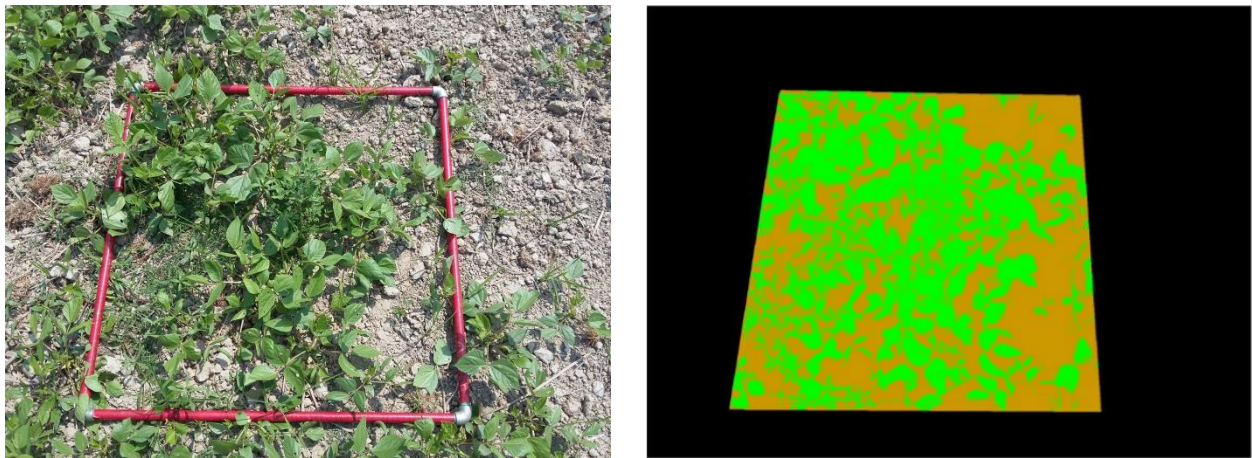
Fig. 4. The smartphone photos were collected vertically above ground using a square frame (left); the pictures were then manually cropped only to the extent of the red frame and annotated (right).

Each of the devices provided varying picture quality with Samsung Galaxy having the best resolution followed by Okapia, Lava, Symphony and the Samsung Tablet. The smartphone photos were not geo-referenced. They were collected following a regular grid and each photo was taken three times using the same device. The photos were processed with CAN-EYE - software that allows the extraction of fCover from RGB images. In CAN-EYE, the first step of processing is cropping the images in order to exclude the

14

parts of them that are not relevant for the analysis. The photos from the field surveys in Bangladesh were cropped so that only the area within the sampling frame was included. Additionally, an adjustment factor is applied brightening or darkening the image, so it can be better differentiated between green and background pixels. Next, the images are roughly classified using thresholding and the initial classification is then manually adjusted until a satisfactory result is reached. The output is annotated images having two classes – green vegetation and background (Weiss & Baret, 2016). These annotated images were used as ground truth. However, as the photos processed with CAN-EYE contained black background (Fig. 4) additional processing was needed to remove it before the images could be imported in SAGA GIS (Fig. 5). The black background was removed using the image processing software XnView Version 2.40 (XnSoft, 2017).
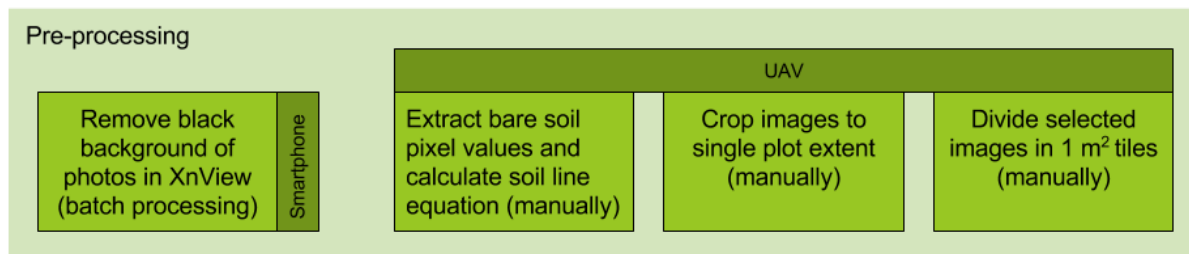
Fig. 5. *Main pre-processing steps for smartphone and UAV images*

RGB photos of the fields were collected with an octocopter at three different altitudes – 7 m, 12 m and 65 m. A SonyNex 7 camera was used to acquire these photos. Multispectral images were collected as well at an altitude of 100 m resulting in a different spatial resolution than the RGB images (Appendix A). For the collection of multispectral images, the Miniature Multiple Camera Array (Mini-MCA) by Tetracam (2017) was used. It covers five spectral bands – 530 nm, 680 nm, 710 nm, 740 nm and 800 nm, with 10 nm band width. Some of the UAV images were not processed due to a failure in the processing software or due to insufficient data. Thus, for some dates and test sites images were available only for certain altitudes and spectral bands. Only images which contain information both in the visible and NIR bands were used in the analysis (Appendix A). For all UAV images considered in the analysis, hundred bare soil pixels per image were manually identified; the values for these pixels in the red and NIR band were retrieved; a linear function was fitted to them in Excel (Microsoft, 2016), thus calculating the soil line equation parameters. As the UAV images usually include many different plots with different crops, the images were manually cropped according to the extent of each single plot (Fig. 5). This procedure resulted in 150 images at plot level; 30 of them being collected at more than two different altitudes. Selected plots were further divided into 1x1 m tiles to enable the analysis of the impact of sampling strategy on fCover (Fig. 5). All UAV images were geo-referenced and orthorectified with Agisoft (Agisoft LLC, 2014).

3.3. Methods

This section contains a detailed description of the methods used to answer the research questions. The first sub-section describes how the effect of spatial resolution on fCover is measured. The analysis is performed both for UAV and smartphone devices. The second sub-section describes the regression model used for calculating fCover. The various vegetation indices used as predictors in the regression model are outlined as well as the random forests algorithm. The next sub-section introduces an alternative classification method for estimating fCover relying on object-based image analysis. In the last section, it is elaborated on how an optimal number of smartphone photos can be determined so that the fCover of a field is accurately estimated.

3.3.1. The effect of spatial resolution

To study the effect of spatial resolution, first the differences in fCover estimates when using smartphones with different resolution (and camera lenses) were examined. fCover estimates based on annotated photos from various smartphone devices - Lava, Okapia, Samsung Galaxy, Symphony smartphones and a Samsung tablet, were used. The fCover estimates were compared for each photo and smartphone model. The differences were evaluated by subtracting the lowest estimate from the highest estimate for each photo. Thus, the photos and the corresponding smartphone models for which this difference was largest could be identified. Next, the error in fCover estimates derived from low-end smartphones when compared to the estimates derived from a high-end smartphone model was measured. This was done by comparing the fCover measurements based on annotated photos from the different smartphone devices to those derived from the annotated pictures from the Samsung Galaxy device, which is assumed to best represent ground truth. A linear model was used to visualize the relationships between them. The coefficient of determination (R^2) was calculated for each linear model in Excel as the squared Pearson's correlation coefficient (r) using the following formula:

$$r(x, y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})/(n-1)}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2/(n-1)}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2/(n-1)}}$$

where x_i and y_i are the variables, $\bar{x}$ and $\bar{y}$ are their sample means, and n is the sample size (Vittinghof et al., 2011).

fCover estimates were retrieved also by applying a regression algorithm on UAV images taken from two different altitudes – 12 m and 65 m. As no annotated UAV images were available to be used as ground truth, no statement was made about the error of these estimates. The emphasis was on the variation in the fCover values at different altitudes. The relationship between the different estimates was described using a linear model.

3.3.2. Regression algorithm

Two different algorithms for estimating fCover were proposed – a classification and a regression algorithm. For the regression algorithm, a machine learning technique was used. There is abundance of machine learning algorithms used in remote sensing – artificial neural networks, support vector machines, self-organizing map, decision trees, random forests, etc. (Lary et al., 2016). In the present research, a random forests (RF) regression algorithm was applied. The RF algorithm can be used not only

for prediction but also for assessing variable importance (Breiman, 2001). In the proposed methodology, the dependent variable fCover was estimated using a number of independent variables represented by various vegetation indices. The vegetation indices were first calculated on pixel level and then aggregated over the whole photo using their statistics (minimum, maximum, mean, and standard deviation). The fCover estimates obtained were compared to those based on the annotated photos using linear regression. The coefficient of determination was calculated. Using the ability of the RF algorithm to measure variable importance, the different vegetation indices were ranked according to their ability to predict fCover. Due to the lack of measurements in the NIR band for the smartphone photos, only color vegetation indices were used for the estimation of fCover from them. For the UAV images, both NIR and color indices were used as predictors. The processing steps are described in Fig. 6.
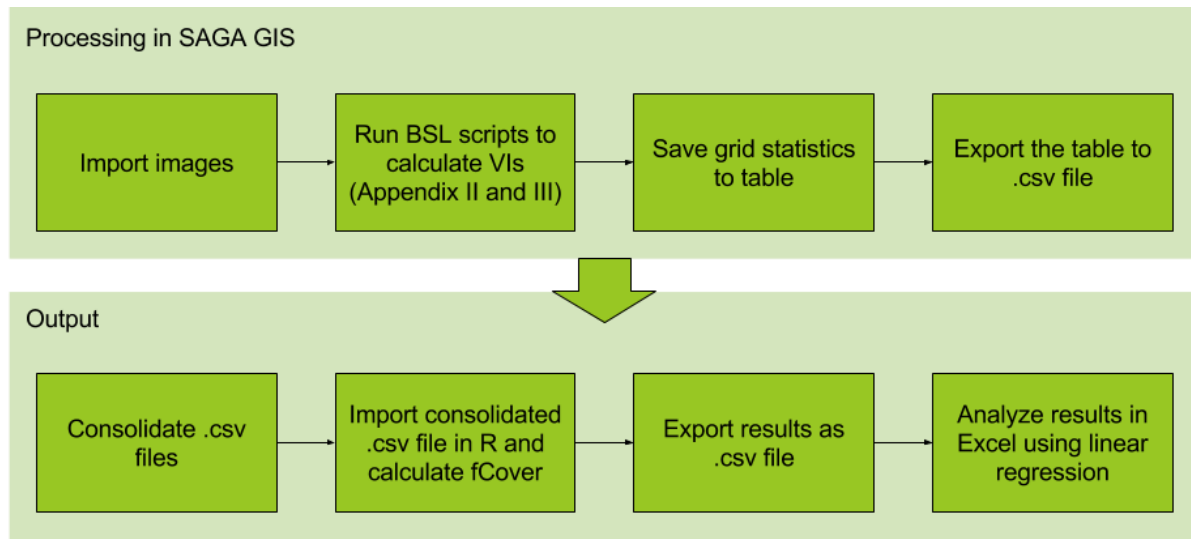
Fig. 6. *Processing steps for the regression approach*

3.3.3. Color indices

In total, ten color indices were calculated. The color indices included the Red Green Ratio Vegetation Index (RGVI) by Jordan (1969), the Excess Green Index (ExG) by Woebbecke et al. (1995), the Excess Red Index (ExR) by Meyer et al. (1999), the Normalized Difference Index (NDI) by Perez et al. (2000), the Excess Green minus Excess Red index (ExGR) and the Improved Normalized Difference Vegetation Index (INDI) by Meyer & Neto (2008), the Modified Excess Green index (MExG) by Burgos-Artizzu et al. (2011), the Color Index of Vegetation Extraction (CIVE) by Kataoka et al. (2003), the vegetation index (VEG) by Hague et al. (2006), the combined index (COM1) by Guijarro et al. (2011), and a modification of the combined index (COM2) by Guererro et al. (2012) These are listed in Table 1.

For calculating COM1, Guijarro et al. (2011) combined the ExG, ExGR, CIVE and VEG indices into one new index that can better segment green plants. The indices were weighted according to their relevance. The weights proposed were respectively $w_1=0.25$, $w_2=0.3$, $w_3=0.33$ and $w_4=0.12$. Guererro et al. (2012) proposed a similar weighted index - COM2. However, as they used only ExG, CIVE and VEG, the weights were slightly adjusted. They used the following weighting - $w_1=0.36$, $w_2=0.47$, $w_3=0.17$. The VEG index used $\alpha=0.667$ as a constant parameter (Hague et al., 2006). The CIVE, ExG, ExR, ExGR, MExG, VEG indices were calculated without applying a threshold value marking the difference between green and non-

green pixels. When using random forests, it is not needed to define a threshold as the algorithm automatically finds the values (nodes) in which the independent variables should be split so that the sum of squared error between the actual and the mean value in the node is minimized (Moisen, 2008). All indices were calculated using the Boehner's Simple Language (BSL) module in SAGA GIS Version 4.1.0 (Conrad et al., 2015). The VIs formulas were saved as SAGA GIS parameter files (Appendix B).

Index	Formula	Index	Formula
ExG	$2g - r - b$	INDI	$128 \times \left(\left(\frac{G - R}{G + R} \right) + 1 \right)$
ExR	$1.3R - G$	CIVE	$0.441R - 0.811G + 0.385B + 18.78745$
ExGR	$ExG - ExR$	COM1	$w_1 ExG + w_2 ExGR + w_3 CIVE + w_4 VEG$
MExG	$1.262G - 0.844R - 0.311B$	COM2	$w_1 ExG + w_2 CIVE + w_3 VEG$
NDI	$\dfrac{G - R}{G + R}$	VEG	$\dfrac{G}{R^\alpha \times B^{1-\alpha}}$

Table 1. List with color indices; R refers to red, G refers to green, and B refers to the blue spectral band; g, r and b refer to the percentage of red, green and blue pixels respectively; w refers to weights; α is a constant.

For the UAV images collected at 100 m height some of the color indices could not be calculated because no data was collected in the blue band at this altitude. Therefore, for the calculation of the color indices the data obtained at 12 m or if not available at 65 m height was used.

3.3.4. NIR indices

For the estimation of fCover from UAV images with a NIR band, additional indices were included. There was a special focus on indices accounting for soil as these are known to be more robust than the widely used NDVI (e.g. North, 2002). The indices used were the Difference Vegetation Index (DVI), the Normalized Difference Vegetation Index (NDVI), the Renormalized Difference Vegetation Index (RDVI) by Roujean & Breon (1995), the Green Normalized Difference Vegetation Index (GNDVI) by Gitelson et al. (1996), the Soil Adjusted Vegetation Index (SAVI) by Huete (1988), the Simple Ration index (SR) by Jordan (1969), the Green Chlorophyll Index (GCI) and the Red Edge Chlorophyll Index (RECI) by Gitelson et al. (2006), the Global Environmental Monitoring Index (GEMI) by Pinty & Verstraete (1992), the Modified Soil Adjusted Vegetation Index (MSAVI) by Qi et al. (1994), the Triangular Vegetation Index (TVI) by Broge & Leblanc (2001), the Adjusted Transformed Soil Adjusted Vegetation Index (ATSAVI) by Baret & Guyot (1991). These are listed in Table 2.

DVI and NDVI are one of the earliest vegetation indices and have been used in variety of domains (Roujean & Breon, 1995; Liang, 2005). DVI is less sensitive to the background reflectance but it is more sensitive to the spectral and directional canopy properties as compared to NDVI (Roujean & Breon, 1995). SAVI is similar to NDVI but it includes an adjustment index *L* which accounts for the influence of soil. The adjustment factor generally varies between 0 and 1 taking smaller values if the vegetation density is low and larger values (>0.5) if the vegetation density is high. For the estimation of SAVI, an adjustment factor *L=0.5* was used as it is proved to reduce soil noise for a range of different vegetation densities (Huete, 1988). In MSAVI, the constant *L* used in SAVI was replaced with a variable function thus minimizing further the sensitivity to soil (Qi et al., 1994). RDVI is also suggested to be less sensitive to soil reflectance than NDVI (Roujean & Breon, 1995). GEMI is less sensitive to atmospheric effects than

NDVI and SR (Pinty & Verstraete, 1992). GNDVI is as well similar to NDVI but instead of the red band it uses the green band; it has wider dynamic range and is more sensitive to chlorophyll compared to NDVI (Gitelson et al., 1996). The RECI index is calculated using both red edge bands at 710 and 740 nm.

Additionally, indices using the parameters from the soil line equation were included (Table 2). To calculate the soil line indices, the soil line equation intercept and slope values were averaged over all available NIR images containing bare soil pixels. One of the first indices to use these parameters is the Perpendicular Vegetation Index (PVI) by Richardson & Weigand (1977). The Transformed Soil Adjusted Vegetation Index (TSAVI) by Baret et al. (1989) also uses the soil line parameters and similarly to SAVI minimizes the influence of soil brightness. Baret et al. (1989) suggest that the intercept and the slope are related by a second order polynomial, however this assumption was not considered here as both parameters were empirically estimated. The Second Soil Adjusted Vegetation Index (SAVI2) by Major et al. (1990) is claimed to be superior to SR, PVI and SAVI. Here it was applied in its simplest version without iteration. The Adjusted Transformed Soil Adjusted Vegetation Index (ATSAVI) by Baret & Guyot (1991) is similar to TSAVI but it includes additional parameter minimizing soil effects (*X=0.08*). The VIs formulas were saved as SAGA GIS parameter files (Appendix C).

Index	Formula	Index	Formula
SR	$\dfrac{NIR}{R}$	**GCI**	$\dfrac{NIR}{G} - 1$
DVI	$NIR - R$	**TVI**	$60(NIR - G) - 100(R - G)$
NDVI	$\dfrac{NIR - R}{NIR + R}$	**RECI**	$\dfrac{NIR}{R_{edge}} - 1$
RDVI	$\dfrac{NIR - R}{\sqrt{NIR + R}}$	**PVI**	$(NIR - aR - b)/\sqrt{1 + a^2}$
GNDVI	$\dfrac{NIR - G}{NIR + G}$	**SAVI2**	$NIR / \left(R - \dfrac{b}{a}\right)$
SAVI	$\dfrac{NIR - R}{NIR + R + L}(1 + L)$	**TSAVI**	$\dfrac{a(NIR - aR - b)}{(aNIR + R - ab)}$
MSAVI	$\dfrac{2NIR + 1 - \sqrt{(2NIR + 1)^2 - 8(NIR - R)}}{2}$	**ATSAVI**	$\dfrac{a(-aR - b)}{aNIR + R - ab + X(1 + a^2)}$
GEMI	$\eta(1 - 0.25\eta) - (R - 0.125)/(1 - R)$ $\eta = \dfrac{2(NIR^2 - R^2) + 1.5NIR + 0.5R}{NIR + R + 0.5}$		

Table 2. *Indices including the near-infrared (NIR) band; L is parameter between 0 and 1 used for minimizing soil background effect; a and b refer to the slope and intercept in the soil line equation; X is parameter used for reducing background effects.*

3.3.5. Random forests

The RF algorithm is an ensemble method meaning that it is based on a large number of decision trees (Liaw & Wiener, 2002). In a decision tree, the data is split at multiple nodes. At each node the data is separated in two parts according to a value of a certain predictor variable. The predictor variable is chosen based on the impurity of the data at the node – this is a measure for the homogeneity of the

data. The process of splitting continues until it is no longer possible to create new nodes or when a certain number of nodes is reached as defined by the user (Moisen, 2008). Unlike decision trees, in RF the splits are based on the variable minimizing impurity most out of a set of randomly chosen predictor variables (Liaw & Wiener, 2002). The RF algorithm is insensitive to the number of independent variables, so it can handle a big number of predictors. It is also robust to overfitting and faster than other ensemble methods using bagging and boosting (Breiman, 2001). While RF can be referred to as "black box" model, it offers the possibility to estimate various metrics which facilitate its interpretation (Prasad et al., 2006). One of these metrics is the previously mentioned variable importance. Variable importance is estimated by measuring the prediction error when the data for the predictor variable is randomly permuted (Prasad et al., 2006). It has to be noted that while the variables importance measures may vary from run to run, their ranking is stable (Liaw & Wiener, 2002).

In this thesis, the randomForest package in R (Liaw & Wiener, 2002) was used to create the RF model. The R script is available in Appendix D.

3.3.6. Classification algorithm

As stated in the literature review chapter, the estimation of fCover can also be approached as a classification problem. GEOBIA models are suitable for classifying very high resolution images or images in which the object of interest is represented by several pixels (Blaschke et al., 2008). They have already been used to estimate ground cover and showed difference between 1 and 4% between the true and estimated cover (Luscier et al., 2006).

In this thesis, SAGA GIS Version 4.1.0 (Conrad et al., 2015) was used to implement a GEOBIA workflow. First, for each device, crop type and acquisition date one RGB photo and the corresponding annotated photo were selected for training resulting in 46% of all photos being used as training samples. For this purpose, photos which represent relatively equally both classes (green and non-green) were chosen. Both the annotated and original images were imported in SAGA GIS and aligned, so that they overlap spatially. The original RGB image was then segmented (Fig. 7).
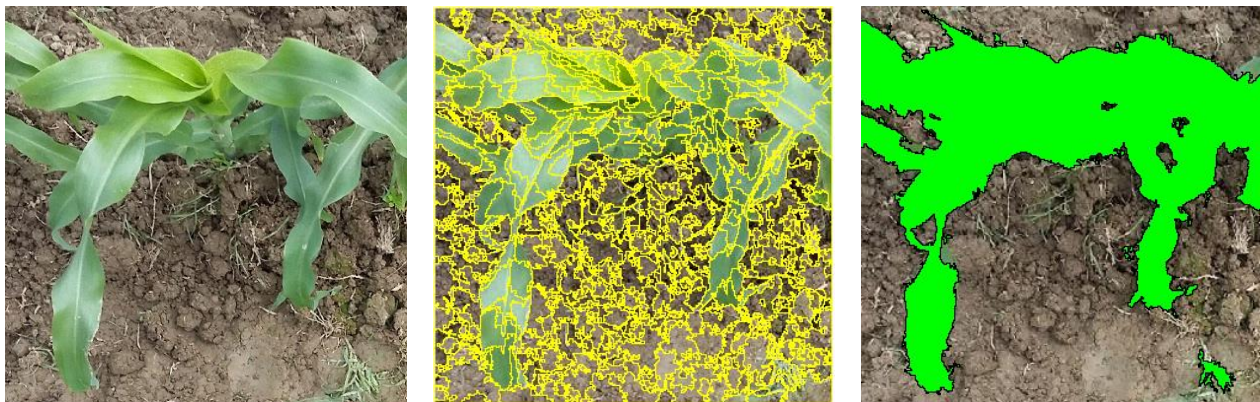
Fig. 7. *Supervised classification based on object-based image segmentation; the picture on the left shows a fragment of an original photo of maize taken with Samsung Galaxy smartphone; the picture in the middle shows the segmentation performed on the same photo; the picture on the right represents the original photo overlaid with the areas classified as green vegetation.*

The parameters for the segmentation were retrieved in a stepwise manner in which the segmented image is overlaid with the original image and the parameters are adjusted until a satisfactory grouping of the pixels is achieved as in Luscier et al. (2006). A grouping was perceived as satisfactory when the segments did not cover more than one class. After the segmentation was performed, for each segment in the training photos the class (green/non-green) from the corresponding annotated photo was extracted. These training samples were used for the creation of the classifiers. Once a classifier was trained for each combination of crops, smartphones and acquisition dates, the rest of the images were segmented and classified using the existing classifiers (Fig. 7). The method used for classifying the images was a winner-takes-all method combining the binary encoding, Mahalanobis distance, maximum likelihood and spectral angle mapping methods. The SAGA GIS tools used to implement this workflow are shown in Fig 8.
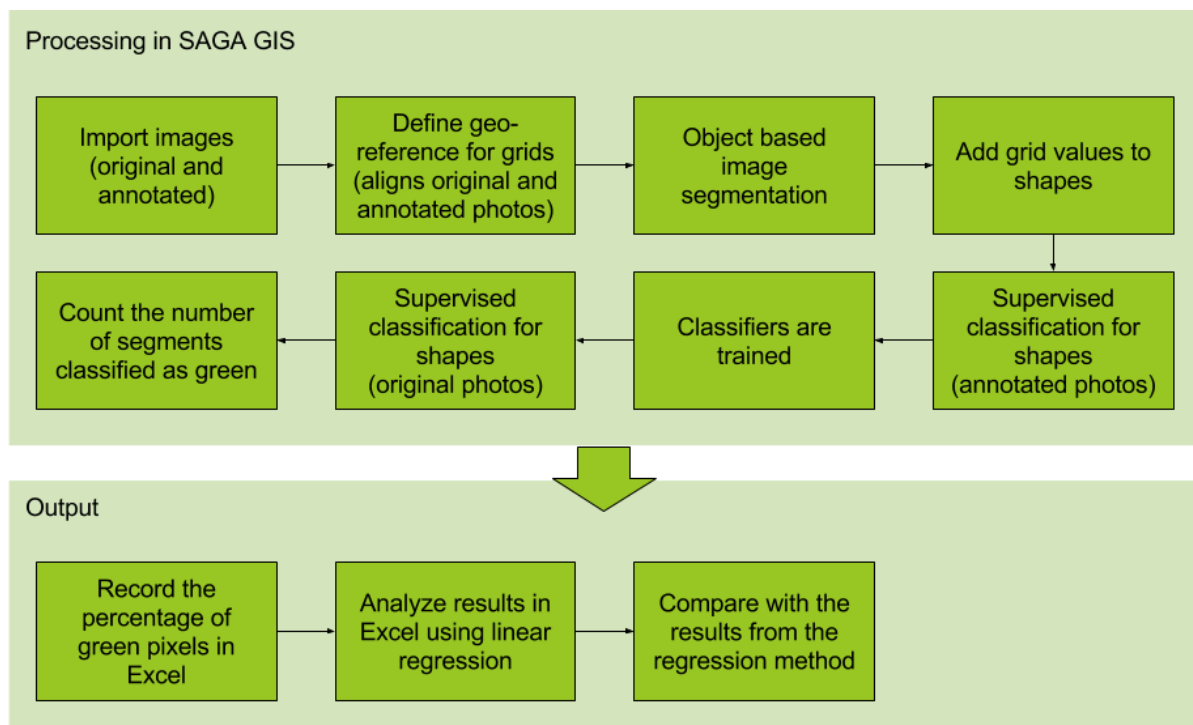
Fig. 8. *Processing steps for the classification approach*

The fCover estimates obtained using the classification approach were compared to those based on the annotated photos using linear regression. The coefficient of determination was calculated.

3.3.7. The impact of field sampling
When taking photos of a field using a smartphone, it is not feasible to cover the whole field. A sampling strategy has to be employed optimizing the number of photos taken so that there is a statistically robust relationship to the fCover value of the whole field. UAV images on the other hand are not as limited as smartphones in regards to their spatial extent and can easily capture many fields at once. In order to make a statement about the optimal number of pictures to be taken on ground, smartphone photos covering entire fields were simulated using high resolution UAV images. This simulation was performed by cropping parts of the UAV images to the same extent that smartphone photos have. Thus, each of the

21

study fields was divided into 1x1 m tiles. Tiles at the edges of the images were omitted during the analysis as those included areas from adjacent fields (Fig. 9). Next, fCover was calculated using the average of different number of simulated photos. The photos were selected randomly and each selection was repeated ten times (Appendix E). The distribution of the fCover measurements and the relation to the field fCover was examined. Due to size of the fields, the analysis was performed only for two of them - one wheat field with high fCover value (0,59) and one maize field with low fCover value (0,19).
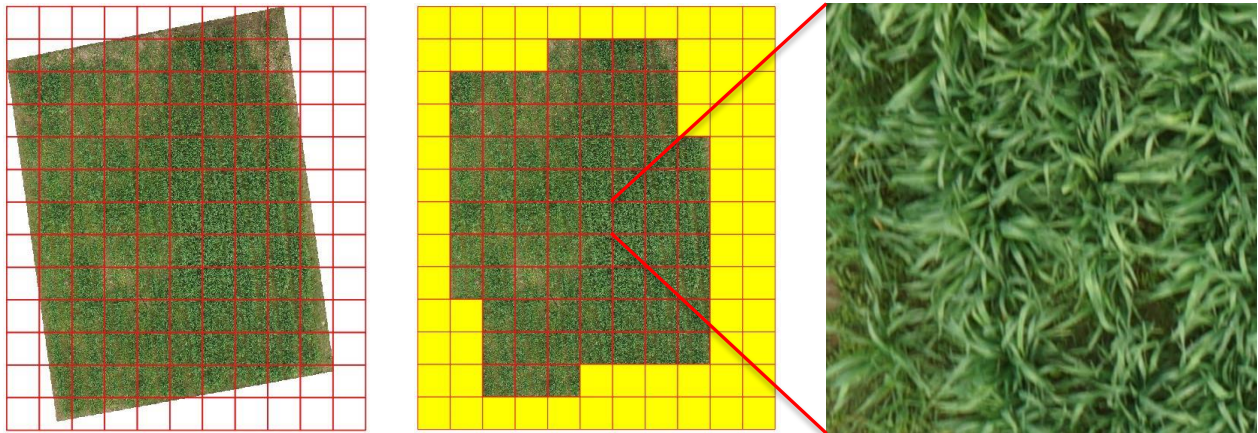
Fig. 9. The picture on the left shows a wheat field which is divided in 1x1 m tiles according to the grid cells; some of the tiles contain parts outside of the study field and are therefore removed from the analysis, these tiles are depicted in yellow in the second picture; the picture on the right shows a single tile used as a simulation of a smartphone photo.

4. Results, Discussion and Conclusion

The results of the analysis described in the previous chapter are presented. The structure of this chapter closely follows the one of the research questions outlined in the first part. The first section elaborates on the effect of spatial resolution on fCover for various types of devices. The next section describes the influence of the method used for fCover estimation. Two methods are compared – classification and regression, showing both their advantages and disadvantages as well as how they handle images of various conditions. In the third section, the problem of choosing an appropriate number of samples is reviewed linking the number of samples taken per field to the field fCover. The last two sections contain discussion of the results and conclusion.

4.1. The effect of spatial resolution

The effect of spatial resolution on fCover estimates was first examined by comparing the estimates based on annotated photos from different devices (Fig. 10). The analysis showed that the difference in fCover estimates ranges between 0 and 0,45 (Appendix F). The largest differences were recorded between Symphony and Samsung Tablet devices. These devices have the lowest resolution. What was typical for the cases with large fCover differences, was that they all depicted crops whose leaves were not well distinguished from the background (Fig. 11). However, for 70% of the samples the difference between the lowest estimate and the highest estimate of the same area was lower than 0,1 (Appendix F).
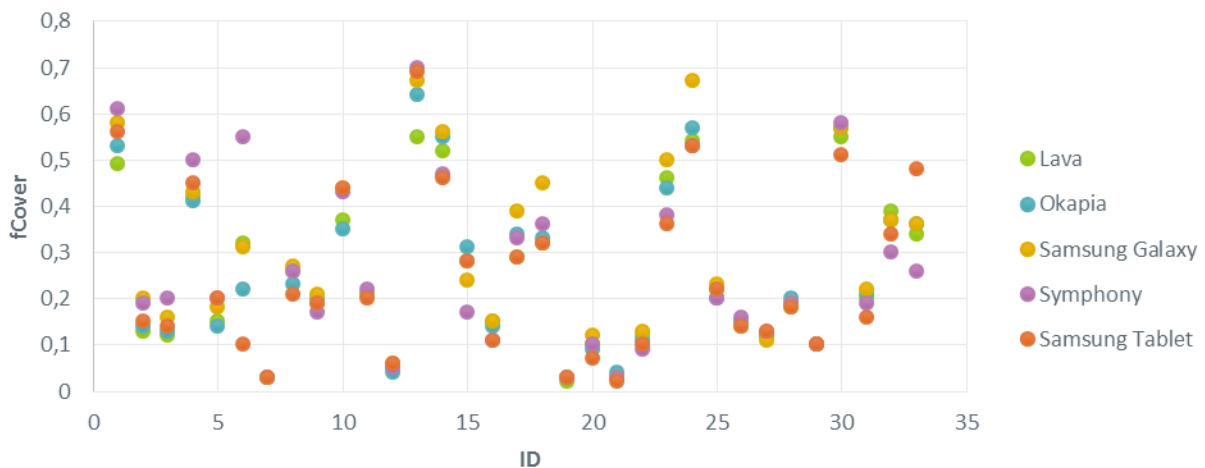
Fig. 10. Differences in the fCover estimates obtained from annotated photos from different devices

The fCover estimates based on annotated photos from Samsung Galaxy were considered to best represent the ground truth. Comparing the fCover estimates based on the annotated photos from other devices to the Samsung Galaxy estimates showed that fCover values derived from higher resolution photos were more accurate than those derived from lower resolution photos. The increase of the error with the deterioration of the resolution was clearly visible from the variance explained. Devices providing better resolution such as Okapia and Lava had larger coefficient of determination than devices with lower resolution such as Symphony and Samsung Tablet (Fig. 12). The percentage of variance

explained decreased with the diminishing resolution provided by these devices proving that the better resolution had positive impact on fCover accuracy.
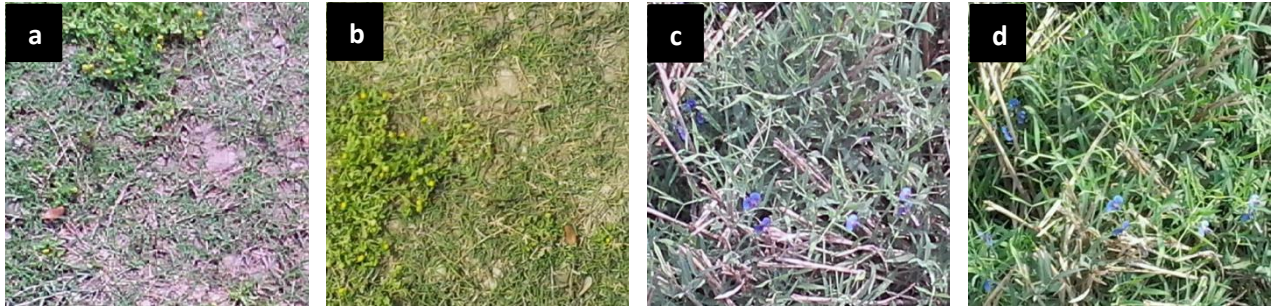
Fig. 11. *Parts of photos taken with Symphony (a, c) and Samsung Tablet (b, d) devices having large differences in the fCover estimates; the difference in the fCover estimates are 0,45 between a and b, and 0,22 between c and d.*
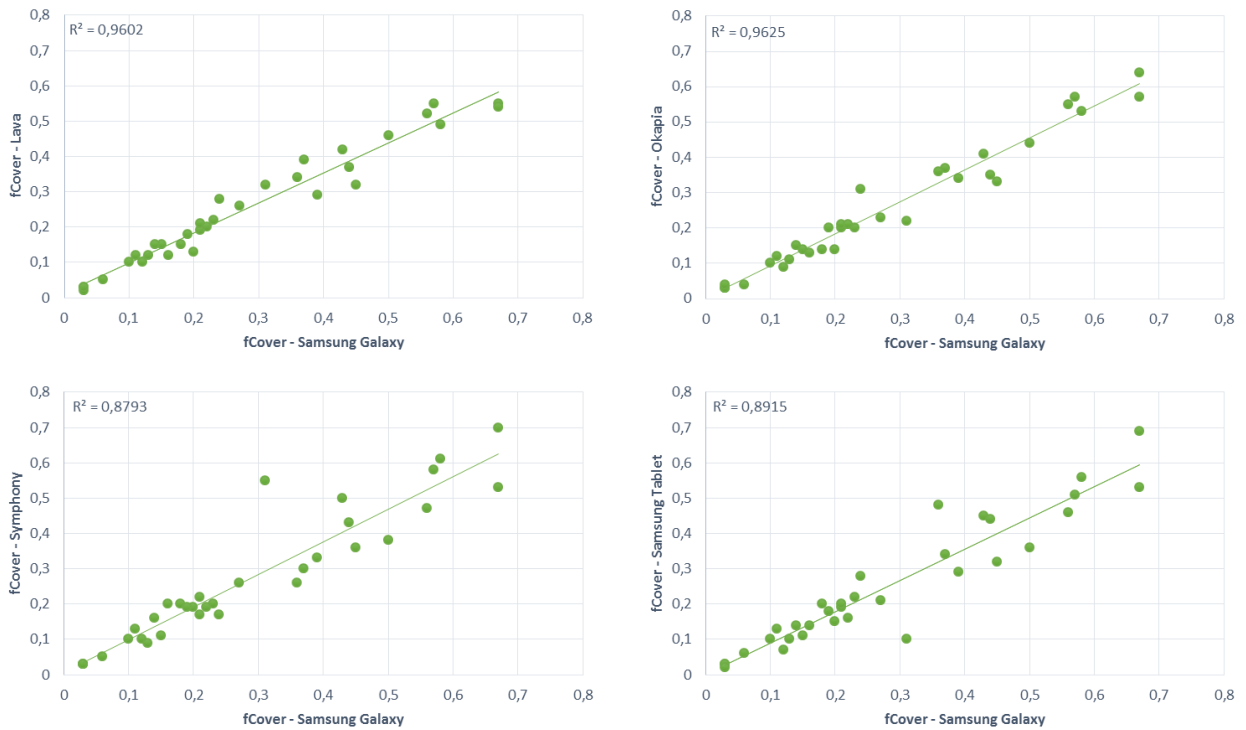
Fig. 12. *Scatterplots showing the linear relationship between the fCover measurements based on manually annotated photos; measurements from different devices are compared to Samsung Galaxy measurements (assumed to be closest to the ground truth).*

Lastly, the fCover estimates derived from UAV images were compared. Due to the lack of training data for the UAV images, no comparison could be made to ground truth data. In order to obtain fCover estimates, a regression algorithm was applied. The regression model was trained with smartphone photos and the predictors for the model were calculated at field level. The regression model is presented in more detail in the next section. fCover estimates derived from images at two different altitudes – 12 m and 65 m, were compared. There were in total 30 fields for which fCover was estimated at both heights.

All of the fields were planted with maize and showed very little variation in the estimated fCover values (Fig. 13).

Fig. 13. The scatterplot on the left shows the linear relationship between the fCover measurements as predicted by the regression algorithm using UAV images taken at an altitude of 12 m and 65 m; the boxplots on the right show the distribution of the fCover values derived from UAV images taken at an altitude of 12 m and 65 m.

The fCover estimates at an altitude of 12 m and 65 m were not correlated as evidenced by the low variance explained (Fig. 13). The fCover estimated at the height of 65 m was approx. twice higher than the fCover estimated at 12 m. No conclusion could be made which one of the estimates was over-/underestimated due to the lack of ground truth data. It was assumed that this outcome was related to the difference in the spatial resolution obtained at the two different altitudes. This outcome showed that substantial difference in the spatial resolution of UAV images can lead to underestimation or overestimation of fCover by as much as a factor of two.

4.2. Influence of the method on the estimated fCover values

First, the regression model is presented. The regression model trained with smartphone photos is referred to as RF.model, the model trained with both RGB and NIR UAV images is referred to as RF.model_NIR and the model trained only with NIR UAV images is referred to as RF.model_NIRO. All models provided a satisfactory fit to the data as it can be seen from the variance explained and the mean of squared residuals (Table 3). The models with NIR indices clearly outperformed the model relying only on color indices. However, it has to be noted that because of the lack of training data for the UAV images, the RF.model trained on smartphone photos was used to calculate fCover estimates from the RGB UAV images. As both RGB and multispectral UAV images were collected over the same areas, the RF.model based estimates were used as training data for the RF.model_NIR and the RF.model_NIRO. Thus, the variance explained for the latter models was artificially boosted and the fCover estimates retrieved from these models were identical to those obtained with the RF.model. The variance explained showed in Table 3 was calculated based on the out of bag (OOB) samples. These samples were not used during the training of the model and therefore give the opportunity to realistically test the ability of the regression model to predict fCover. The randomForest package in R uses "pseudo R-squared" to estimate the percentage of variance explained, which is calculated as follow:

$$R^2 = 1 - mse/Var(y)$$

where mse is the vector of mean square errors, and $Var(y)$ is the variance of the response variable (Breiman et al., 2015). It has to be differentiated between the variance explained based only on OOB samples as in Table 3 and the variance explained showed in Fig. 13 as in the latter both the OOB and the training samples were included, thus increasing the coefficient of determination.

Table 3. The percentage of variance explained and mean of squared residuals of out of bag (OOB) samples as per model; OOB samples are not used during the training of the models.

	% Variance explained	Mean of squared residuals
RF.model	66.73%	0.0104
RF.model_NIR	97.04%	0.0005
RF.model_NIRO	92%	0.0013

While the fCover estimates derived from manually annotated photos best aligned with those derived from the highest resolution images, the percentage of variance explained for all types of devices was high (>95%) (Fig. 14). Higher resolution photos however did not always outperform lower resolution photos when used for fCover estimation. For example, the fCover measurements derived from Samsung Tablet photos corresponded better to those derived from manually annotated photos taken with the same device compared to Symphony and Lava derived estimates. Thus, it can be concluded that minor differences in the photos resolution resulting from the device type did not have major impact on the fCover estimates.
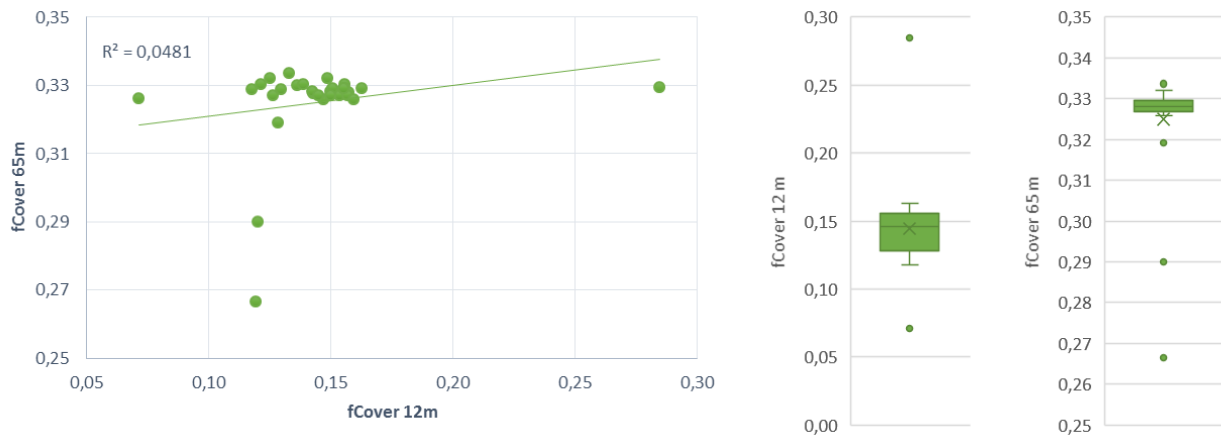
Samsung Tablet

$R^2 = 0,9702$
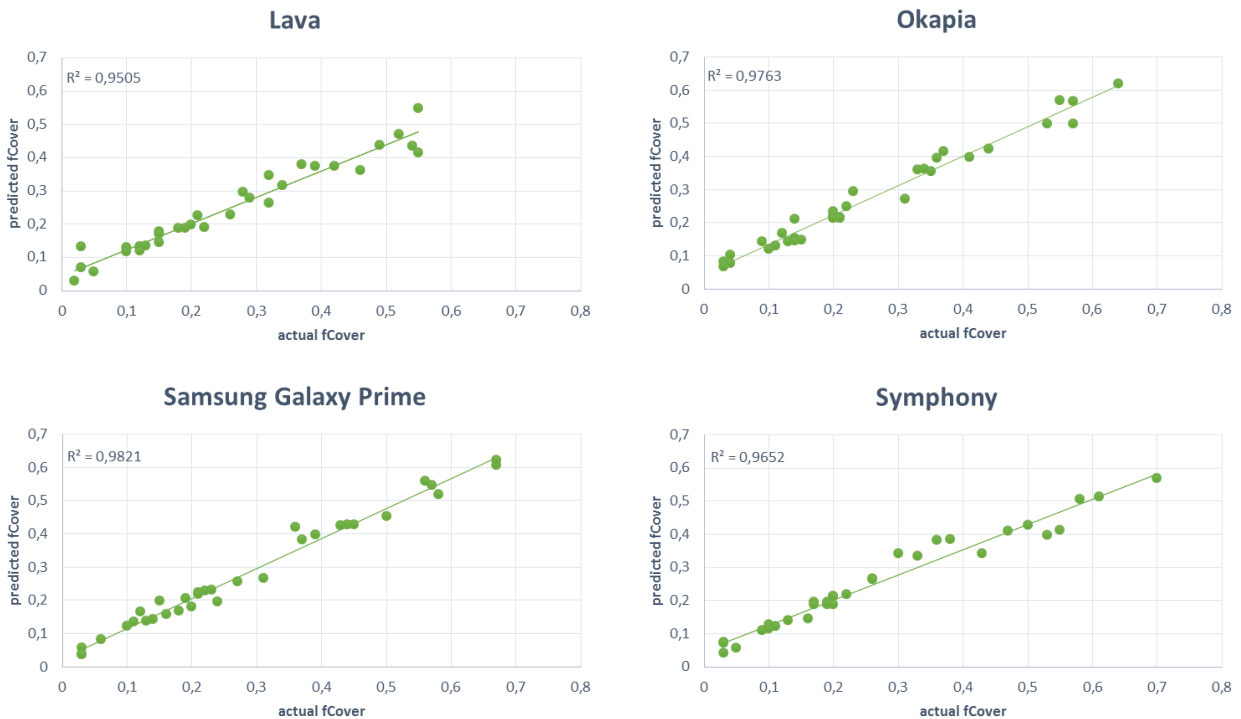
Fig. 14. *The charts show the linear relationship between the fCover measurements as predicted by the regression algorithm and the actual fCover measurements as calculated from the manually annotated photos.*

An analysis of the residuals showed that for all devices, there was a particular photo for which the observed fCover exceeded the estimated fCover with at least 0,1. What was peculiar about this photo was that it had strong reflectance of sun light (glint), so some parts of the plants appeared white on the photos (Fig. 15). This phenomenon might have led to the underestimation of fCover by the regression algorithm. All other photos where the absolute difference between the ground truth and the predicted fCover peaked (>0,1) were characterized by rather granular structure of the plants in contrast to wheat, maize and mung beans which have well distinguished leaves (Fig. 15).

Fig. 15. *Photos taken with Lava (a), Okapia (b), Samsung Galaxy (c), Symphony (d) and Samsung Tablet (e) devices showing the effect of glint; the last photo (f) shows the type of plant for which the residuals in the regression algorithm often peaked.*

In the RF.model, the most important parameter was the mean value of the ExR index. In this model, it was evident that the mean values of the indices were more important variables than any other statistics (minimum, maximum values and standard deviation). In contrast, in the RF.model_NIR both the mean and the standard deviation of the VIs played an important role. In the RF.model_NIR, the most important parameter was the mean value of the ExR index too. The most important parameter based on a NIR VI was the standard deviation of the RECI index using the red edge at 710 nm. The same held true also for the RF.model_NIRO (Table 4).

Table 4. The table shows the first ten most important variables in each model (RF.model, RF.model_NIR, RF.model_NIRO) as measured by node purity of the random forests trees.

	RF.model			RF.model_NIR			RF.model_NIRO	
№	VI	Importance		VI	Importance		VI	Importance
1	ExR_mean	1.01990145		ExR_mean	4.397639e-01		RECI1_std	0.4130000876
2	ExGR_mean	0.51256694		INDI_std	2.251080e-01		GCI_std	0.2333413393
3	INDI_mean	0.46265651		RECI1_std	1.870275e-01		RECI1_mean	0.1716849517
4	COM1_mean	0.34974767		NDI_std	1.627520e-01		SR_std	0.1583610722
5	NDI_mean	0.27896846		INDI_mean	1.622089e-01		TSAVI_std	0.1110948238
6	ExR_std	0.23431179		NDI_mean	1.289101e-01		ATSAVI_std	0.1095700862
7	ExGR_std	0.16522928		ExGR_mean	2.426570e-02		RECI2_std	0.0785645667
8	CIVE_mean	0.15114382		VEG_min	2.143247e-02		GEMI_std	0.0476674859
9	ExGR_max	0.12776804		GCI_std	1.855404e-02		RDVI_std	0.0328379150
10	ExG_mean	0.12513443		RECI2_std	1.684232e-02		DVI_std	0.0185556549

Lastly, results for the classification model are presented. The percentage of variance explained when the training samples were excluded amounted to 76% (Fig. 16). However, this measure cannot be directly compared to variance explained for the regression algorithm (Table 3) because different amount of training samples were used.
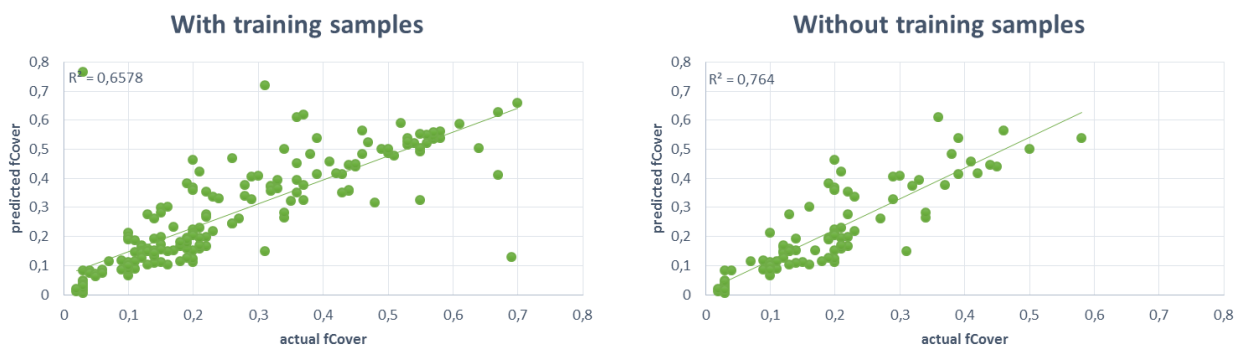
Fig. 16. These charts show the coefficients of determination for the classification model with and without including the training samples.

The fCover estimates derived from manually annotated photos best aligned with those derived from the images taken with Symphony, Samsung Galaxy and Okapia (Fig. 17). So, similarly to the regression approach, higher resolution photos did not always outperform lower resolution photos when used for

fCover estimation. However, in contrast to the regression approach for some devices this relationship was very weak. For example, for Lava pictures the variance explained amounted to only 0,44 (Fig. 17). It is unlikely that the low coefficient of determination was caused by the resolution of Lava smartphones because fCover estimates derived from other devices offering lower resolution (e.g. Symphony, Samsung Tablet) had stronger relationship to the actual fCover. It is assumed that this phenomenon is rather related to the characteristics of the photos. Thus, it can be concluded that minor differences in the photos resolution resulting from the device type were not strongly correlated to the fCover estimates when the latter were calculated using a classification approach and manually annotated photos from the same device. In both cases when a regression and when a classification approaches were used, there was no clear pattern showing that higher resolution photos were always better aligned to the manually annotated photos taken with the same device.
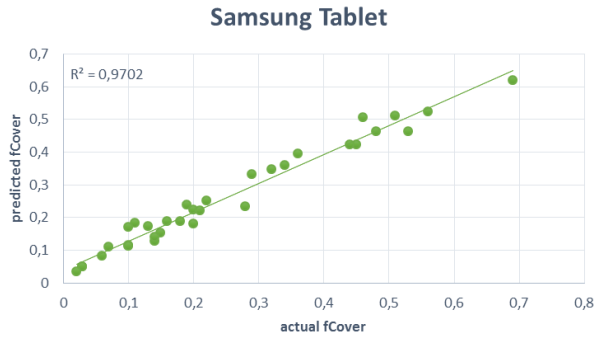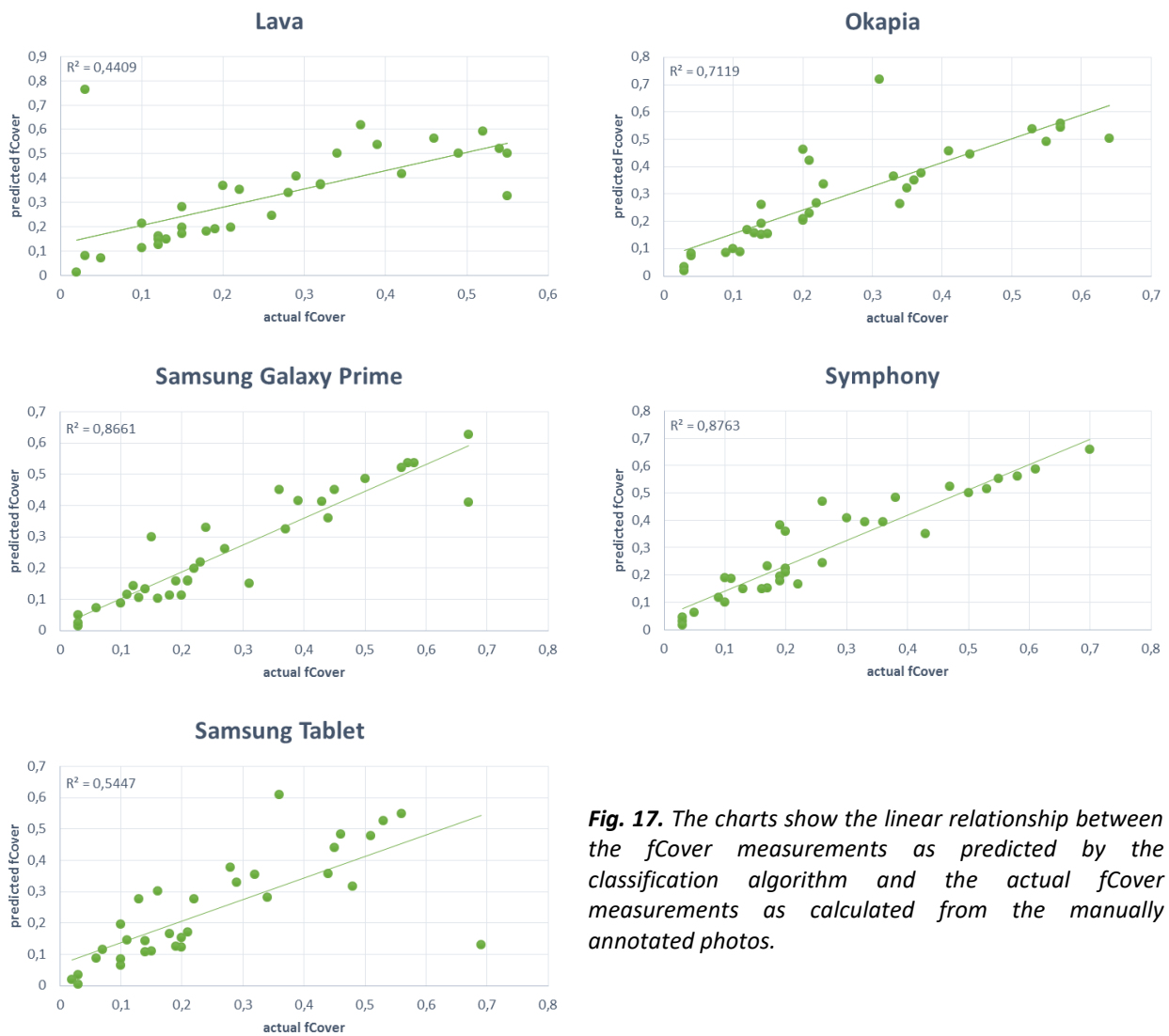

Fig. 17. The charts show the linear relationship between the fCover measurements as predicted by the classification algorithm and the actual fCover measurements as calculated from the manually annotated photos.

The overall percentage of variance explained including the training samples amounted to 66% opposed to 76% when the training samples were excluded (Fig. 16). These contradictory results could be

explained with the strategy adopted for selecting training photos and the limited amount of data. For each device, crop type and date of acquisition a separate classifier was trained resulting in 46% of data being used for training. However, for some crops there was only one photo per acquisition date. This often resulted in a classifier being trained for that particular photo but not being applied on any other photos. Mostly these photos contained crops other than mung beans, maize and wheat that did not have well distinguished leaves. Such photos were more challenging for classification and by training a classifier specific for them they were excluded from the validation dataset. Thus, the validation dataset included samples that were easier to classify (e.g. mung beans, maize, wheat). The classifiers trained on such photos performed worse in comparison to other classifiers for other crops (mung beans, maize, wheat) but because they were not applied on many other photos, they had limited negative influence on the overall accuracy of the fCover estimates. An example for this phenomenon is shown in Fig. 18; for this particular photo the predicted fCover was 0,72 while fCover from the annotated photo amounted to 0,31.
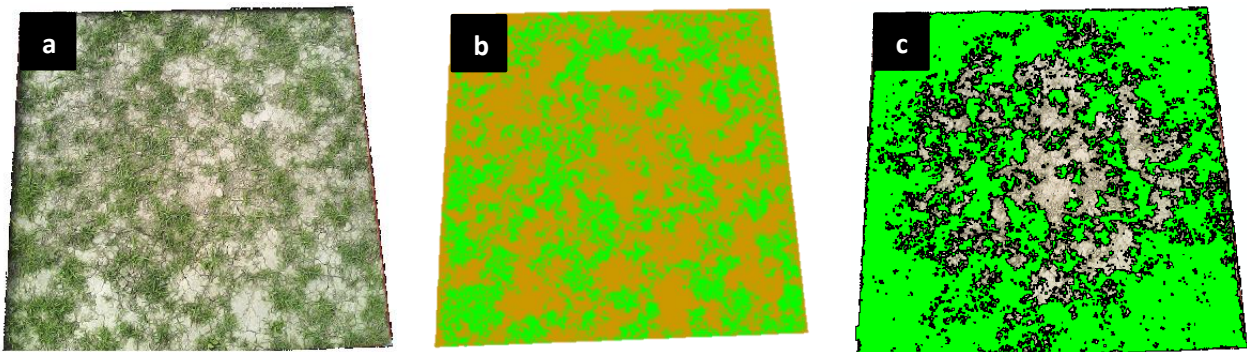
Fig. 18. The original RGB photo (a), the corresponding annotated photo (b), and classified photo (c); the classifier was trained using the annotated version of the photo.

Another common reason for misclassification was the difference in the illumination in the photos used for training the classifiers and the photos on which the same classifiers were applied (Fig. 19). The approach to use for the training of the classifiers only photos that were acquired in the same day did not manage to minimize the effect of changing illumination. Strong reflectance of sun light in some of the photos was also a common reason for the inaccurate estimation of fCover similarly to the regression algorithm (Fig. 15).

Fig. 19. The photo on the left was used for training a classifier applied on the photo on the right; the actual fCover on the right photo was 0,55, while the predicted fCover amounted to 0,32.

4.3. The impact of field sampling

The last research question concentrated on the effect of field sampling and namely the optimal number of smartphone photos to be taken, so that a statistically robust characterization of the fCover value of a field can be obtained. One wheat plot with denser vegetation cover (*fCover* = 0,59) and one maize plot with lower vegetation cover (*fCover* = 0,19) were examined (Fig. 20). For the estimation of fCover, the regression algorithm was used. The fields had different size, so dividing them in 1x1 m tiles resulted in having 75 tiles for the wheat field and 87 tiles for the maize field. For both plots, the average fCover was calculated by using an increasing number of randomly collected tiles with each selection being repeated ten times.
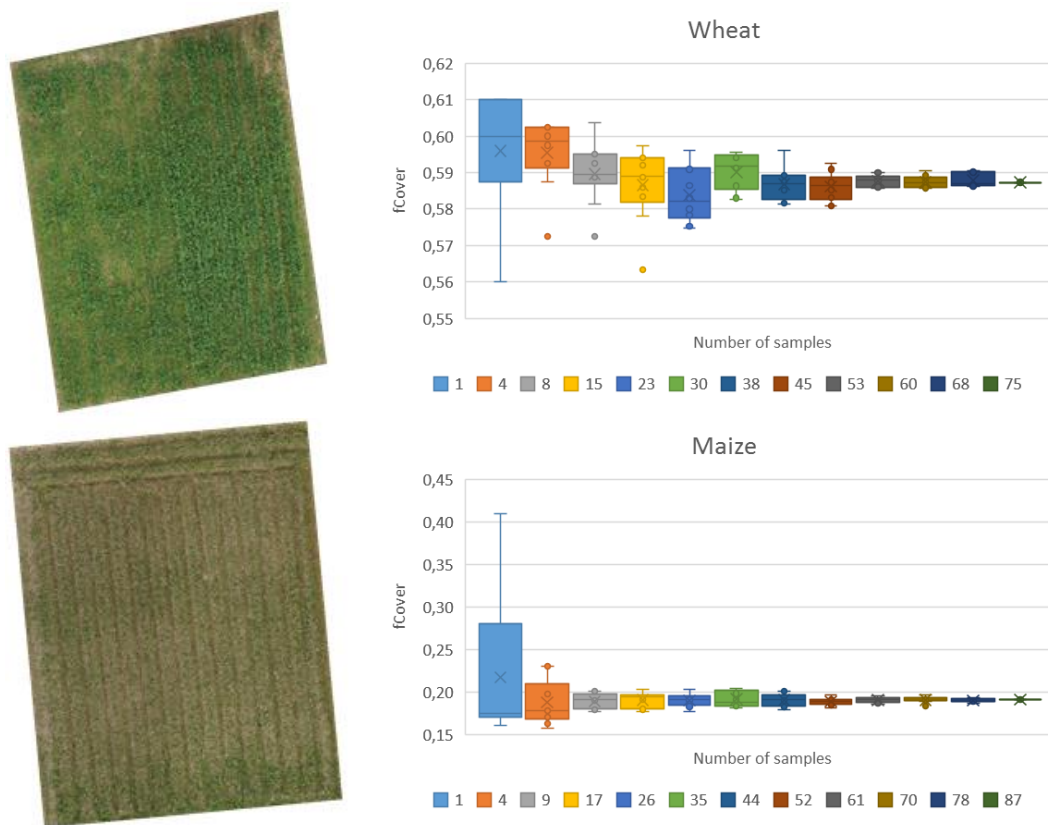
Fig. 20. *Distribution of the fCover measurements per field using an increasing number of randomly selected samples; the analysis is performed for one wheat and one maize field.*

For the wheat field, the analysis showed that fCover estimates based only on one sample could yield results very similar to the field fCover. The difference between the fCover estimates based on one sample and the field fCover for the wheat field ranged between -3% and 1%. For the maize field, this difference ranged between -3% and 22% probably due to its more heterogeneous pattern. Four samples were needed to achieve an fCover estimate similar to the field fCover (ca. ±4% difference).

4.4. Discussion

The images used in this thesis were not radiometrically corrected, only digital numbers were used for the analysis. The lack of radiometric correction has certain implications for the fCover estimates. First, the parameters of the soil line equation used for the calculation of some of the vegetation indices could be influenced. If the images used for its estimation were radiometrically corrected, the parameters of the equation would have been precisely calculated while using the digital numbers resulted in having diverse pixel values for bare soil. However, the vegetation indices using the soil line equation parameters were not among the most important indices, thus it can be assumed that the lack of radiometric correction does not have major impact on fCover estimates when applying the regression approach. Secondly, the lack of radiometric correction has an impact on the classification of images. When using a classification approach, differing sun reflectance can lead to under- or overestimation of fCover with more than a factor of two. Possible solutions for this problem are to apply a radiometric correction or to develop different classifiers adjusted to various levels of sun illumination and record the level of illumination for each photo taken. This is especially crucial for UAV images which while providing reliable base for the estimation of vegetation indices are prone to angular variation in reflectance and thus can lead to misleading results (Rasmussen et al., 2016).

Comparing both regression and classification approaches showed that regression algorithms are better suited for fCover estimation. The regression approach outperformed the classification approach with its better ability to deal with varying sun illumination. This was evidenced by the more accurate estimation of fCover for photos of the same crop where varying illumination conditions were present. The better performance might be also due to the fact that while the classification algorithm was relying only on the digital numbers of the pixels, in the regression algorithm vegetation indices were used as predictors. In their study about the estimation of cotton leaf area index, Bendig et al. (2015) concluded that models using vegetation indices rather than spectral reflectance as independent variables are better at predicting. Thus, a possible way of improving the classification approach is to use vegetation indices during the processes of segmentation and/or classification.

While this research makes use of different methods for estimating fCover, it is not possible to apply the whole spectrum of available methods. As shown in the literature review there is wide abundance of methods which makes a complete overview and comparison between all of them hardly possible. Priority is given to those which are computationally more efficient. Thus, physical models are excluded in favor of statistical and machine learning models. But only for the regression approach, there are plenty of different statistical and machine learning models that can be used for its implementation. In this thesis, a random forests model was used. Although it showed satisfactory results, it is important to cover wider range of machine learning methods (e.g. Support Vector Machines) before implementing it on a larger scale. The classification method also offers wide variety of different options for its implementation. While here an object-based approach was chosen, classification can be performed also on pixel level. More sophisticated software for object-based image analysis offer additional parameters to be adjusted besides the parameters for band width for seed point generation and generalization available in SAGA GIS. eCognition for example offers the opportunity to define classes based on the values for these classes for certain vegetation indices.

Analyzing the impact of field sampling on fCover showed that at least 5% of the area has to be studied in order to achieve an fCover estimate that does not differ from the actual fCover with more than 4%. A similar study of fCover over alpine grassland showed that sampling as little as 0,1% of the area can yield fCover accuracy of over 80% (Chen et al., 2016). However, Chen et al. (2016) outlined that these results largely depend on the homogeneity of the area as measured by the coefficient of variation.

Due to the lack of smartphone photos and UAV images of the same area when studying the impact of field sampling simulated photos were used. While these photos cover approximately the same area as smartphone photos, they are not a perfect substitute for smartphone photos as they do not offer the same picture quality. Optimally, real smartphone photos have to be used.

While programming scripts were used to calculate the vegetation indices and to train the regression model, for both the regression and classification algorithm a higher degree of automation could be achieved. This would include batch processing for the images in SAGA GIS and automated coupling with the R script without the need to export and preprocess the data in Excel. SAGA GIS does offer the opportunity for batch processing and also has Python API which as well can be used to automate the processes.

4.5. Conclusion

The main objective of this thesis was to study the impact of various factors on fCover. First, the effect of spatial resolution on the fCover estimates was examined by using different devices such as smartphones and UAV providing various level of resolution. Next, the influence of the method used for fCover estimation was studied by comparing two types of approaches – regression and classification. Lastly, the impact of field sampling was analyzed by relating the fCover obtained from varying number of samples to the fCover of the field, where these samples were collected.

The analysis of smartphone photos showed that the differences in the fCover estimates obtained from the various devices are minor. Comparing the fCover estimates derived from low-end smartphones to those derived from a high-end smartphone model revealed an increase of the error in fCover estimates when photos with lower resolution were used. When fCover estimates are based on UAV images, using too coarse spatial resolution as obtained at 65 m altitude can lead to substantial over- or underestimation of fCover. The fCover estimates based on UAV images acquired at 12 m and 65 m altitude differed by as much as a factor of two.

The difference between the predicted fCover retrieved with the regression method and the actual fCover as calculated from the annotated photos was low. This was evidenced by the high percentage of variance explained between the two estimates, which amounted to more than 95% for all types of devices. The most import color index was the ExR index, while from the vegetation indices in the near-infrared region this was the RECI index based on the red edge at 710 nm. The difference between the predicted fCover retrieved with the classification method and the actual fCover as calculated from the annotated photos was low. However, the percentage of variance explained for the different devices varied greatly.

The number of smartphone photos needed in order to achieve statistically robust characterization of the fCover of a field depends on the heterogeneity of the particular field. For the fields studied, surveying

four photos per field yielded an fCover estimate that did not differ from the actual fCover with more than 4%. Four photos amounted to approx. 5% of the field area.

Future research could concentrate on developing a routine for radiometric correction of smartphone photos and UAV images or on developing classifiers accounting for various levels of sun illumination. Exploring the application of vegetation indices for the purpose of segmentation and/or classification could improve the results obtained with the object-based image analysis technique. Additional machine learning methods for regression could be investigated.

References

Addink, E.A., Van Coillie, F.M.B. & de Jong, S.M. 2012, "Introduction to the GEOBIA 2010 special issue: From pixels to geographic objects in remote sensing image analysis", *International Journal of Applied Earth Observation and Geoinformation,* vol. 15, no. 1, pp. 1-6.

Adhikari, H., Heiskanen, J., Maeda, E.E. & Pellikka, P.K.E. 2016, "The effect of topographic normalization on fractional tree cover mapping in tropical mountains: An assessment based on seasonal Landsat time series", *International Journal of Applied Earth Observation and Geoinformation*, vol. 52, pp. 20-31.

Afzal, M. A., Bakr, M. A., Hamid, A., Haque, M. M., & Aktar, M. S. 2006, "Adoption and seed production mechanisms of modern varieties of mungbean in Bangladesh", *Improving income and nutrition by incorporating mungbean in cereal fallows in the Indo-Gangetic Plains of South Asia, DFID Mungbean Project for 2002–2004*.

AgiSoft PhotoScan, 2014, St. Petersburg, Russia: Agisoft LLC.

Agüera, F., Carvajal, F. & Pérez, M. 2011, "Measuring sunflower nitrogen status from an unmanned aerial vehicle-based system and an on the ground device", International Archives of the Photogrammetry, *Remote Sensing and Spatial Information Sciences - ISPRS Archives*, pp. 33-37.

Anderson, K., & Gaston, K. J. 2013, "Lightweight unmanned aerial vehicles will revolutionize spatial ecology", *Frontiers in Ecology and the Environment*, vol. 11, no. 3, pp. 138-146.

Aquino, A., Millan, B., Gaston, D., Diago, M. & Tardaguila, J. 2015, "vitisFlower®: Development and testing of a novel android-smartphone application for assessing the number of grapevine flowers per inflorescence using artificial vision techniques", *Sensors (Switzerland)*, vol. 15, no. 9, pp. 21204-21218.

Balota, M. & Oakes, J. 2016, "Exploratory use of a UAV platform for variety selection in peanut", *Proceedings of SPIE - The International Society for Optical Engineering.*

Bannari, A., Staenz, K. & Khurshid, K.S. 2007, "Remote sensing of crop residue using Hyperion (EO-1) data", *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 2795-2799.

Baret, F. & Buis, S. 2008, "Estimating canopy characteristics from remote sensing observations: Review of methods and associated problems", *Advances in Land Remote Sensing: System, Modeling, Inversion and Application*, pp. 173-201.

Baret, F. & Guyot, G. 1991, "Potentials and limits of vegetation indices for LAI and APAR assessment", *Remote Sensing of Environment*, vol. 35, no. 2-3, pp. 161-173.

Baret, F., Guyot, G. & Major, D.J. 1989, "TSAVI: A vegetation index which minimizes soil brightness effects on LAI and APAR estimation", *Digest - International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 1355-1358.

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J.L. & Leroy, M. 2007, "LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. Part 1: Principles of the algorithm", *Remote Sensing of Environment*, vol. 110, no. 3, pp. 275-286.

Baret, F., Weiss, M., Allard, D., Garrigues, S., Leroy, M., Jeanjean, H., ... & Bohbot, H. 2005, "VALERI: a network of sites and a methodology for the validation of medium spatial resolution land satellite products", *Remote Sensing of Environment*, vol. 76, no. 3, pp. 36-39.

Bastiaanssen, W.G., Molden, D.J., & Makin, I.W. 2000, "Remote sensing for irrigated agriculture: examples from research and possible applications", *Agricultural water management*, vol. 46, no. 2, pp. 137-155.

Bauer, J., Siegmann, B., Jarmer, T. & Aschenbruck, N. 2016, "Smart fLAIr: A smartphone application for fast LAI retrieval using Ambient Light Sensors", *SAS 2016 - Sensors Applications Symposium, Proceedings*, pp. 401-406.

Bauer, T. & Strauss, P. 2014, "A rule-based image analysis approach for calculating residues and vegetation cover under field conditions", *Catena*, vol. 113, pp. 363-369.

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M.L. & Bareth, G. 2015, "Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley", *International Journal of Applied Earth Observation and Geoinformation*, vol. 39, pp. 79-87.

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I. & Heynen, M. 2004, "Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 58, no. 3-4, pp. 239-258.

Berberoglu, S., Satir, O. & Atkinson, P.M. 2009, "Mapping percentage tree cover from Envisat MERIS data using linear and nonlinear techniques", *International Journal of Remote Sensing*, vol. 30, no. 18, pp. 4747-4766.

Blaschke, T. 2004, "Object-based contextual image classification built on image segmentation", *2003 IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, pp. 113-119.

Blaschke, T. 2010, "Object based image analysis for remote sensing", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 1, pp. 2-16.

Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Queiroz Feitosa, R., van der Meer, F., van der Werff, H., van Coillie, F. & Tiede, D. 2014, "Geographic Object-Based Image Analysis - Towards a new paradigm", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 87, pp. 180-191.

Böhner, J., Selige, T., & Ringeler, A. 2006, "Image segmentation using representativeness analysis and region growing", *Göttinger GeographischeAbhandlungen*, vol. 115, pp. 29-38.

Bondi, E., Salvaggio, C., Montanaro, M. & Gerace, A.D. 2016, "Calibration of UAS imagery inside and outside of shadows for improved vegetation index computation", *Proceedings of SPIE - The International Society for Optical Engineering*.

Bonham, C. D. 2013, *Measurements for terrestrial vegetation (Second edition),* Chichester, West Sussex, UK: John Wiley-Blackwell.

Breiman, L. 2001, "Random forests", *Machine learning*, vol. 45, no. 1, pp. 5-32.

Breiman, L., Cutler, A., Liaw, A. & Wiener, M. 2015, "Breiman and Cutler's Random Forests for Classification and Regression", [online] Available at: https://cran.r-project.org/web/packages/randomForest/randomForest.pdf [accessed 11 August 2017].

Broge, N.H. & 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*, vol. 76, no. 2, pp. 156-172.

Bu, H., Sharma, L.K., Denton, A. & Franzen, D.W. 2017, "Comparison of satellite imagery and ground-based active optical sensors as yield predictors in sugar beet, spring wheat, corn, and sunflower", *Agronomy Journal*, vol. 109, no. 1, pp. 299-308.

Burgos-Artizzu, X.P., Ribeiro, A., Guijarro, M. & Pajares, G. 2011, "Real-time image processing for crop/weed discrimination in maize fields", *Computers and Electronics in Agriculture*, vol. 75, no. 2, pp. 337-346.

Burkart, A., Hecht, V.L., Kraska, T. & Rascher, U. 2017, "Phenological analysis of unmanned aerial vehicle based time series of barley imagery with high temporal resolution", *Precision Agriculture*, pp. 1-13.

Calera, A., Martínez, C. & Meliá, J. 2001, "A procedure for obtaining green plant cover: Relation to NDVI in a case study for barley", *International Journal of Remote Sensing*, vol. 22, no. 17, pp. 3357-3362.

Camastra, F. & Vinciarelli, A. 2015, *Machine learning for audio, image and video analysis: Theory and applications.* 2nd ed. London: Springer.

Campbell, K. M., Gulledge, J., McNeill, J. R., Podesta, J., Ogden, P., Fuerth, L., Woolsey, R. J., Lennon, A. T., Smith, J. & Weitz, R. 2007, *The age of consequences: the foreign policy and national security implications of global climate change.* Washington DC: Center for strategic and international studies.

Campos-Taberner, M., Garcia-Haro, F.J., Confalonieri, R., Martinez, B., Moreno, A., Sanchez-Ruiz, S., Gilabert, M.A., Camacho, F., Boschetti, M. & Busetto, L. 2015, "Intercomparison of instruments for measuring leaf area index over rice", *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3389-3392.

Candiago, S., Remondino, F., De Giglio, M., Dubbini, M. & Gattelli, M. 2015, "Evaluating multispectral images and vegetation indices for precision farming applications from UAV images", *Remote Sensing*, vol. 7, no. 4, pp. 4026-4047.

Castaldi, F., Casa, R., Pelosi, F. and Yang, H. 2015, "Influence of acquisition time and resolution on wheat yield estimation at the field scale from canopy biophysical variables retrieved from SPOT satellite data", *International Journal of Remote Sensing*, vol. 36, no. 9, pp. 2438-2459.

Chen, J., Yi, S., Qin, Y. & Wang, X. 2016, "Improving estimates of fractional vegetation cover based on UAV in alpine grassland on the Qinghai–Tibetan Plateau", *International Journal of Remote Sensing*, vol. 37, no. 8, pp. 1922-1936.

Clewley, D., Bunting, P., Shepherd, J., Gillingham, S., Flood, N., Dymond, J., Lucas, R., Armston, J. & Moghaddam, M. 2014, "A python-based open source system for Geographic Object-Based Image Analysis (GEOBIA) utilizing raster attribute tables", *Remote Sensing*, vol. 6, no. 7, pp. 6111-6135.

Combal, B., Baret, F., Weiss, M., Trubuil, A., Mace, D., Pragnere, A., Myneni, R., Knyazikhin, Y. & Wang, L. 2003, "Retrieval of canopy biophysical variables from bidirectional reflectance: Using prior information to solve the ill-posed inverse problem", *Remote sensing of environment*, vol. 84, no. 1, pp. 1-15.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V. & Böhner, J. 2015, "System for Automated Geoscientific Analyses (SAGA) v. 2.1.4", *Geoscientific Model Development*, vol. 8, no. 7, pp. 1991-2007.

Daughtry, C.S.T., Hunt, E.R., Doraiswamy, P.C. & McMurtrey, J.E. 2005, "Remote sensing the spatial distribution of crop residues", *Agronomy Journal*, vol. 97, no. 3, pp. 864-871.

Davidson, A. & Csillag, F. 2001, "The influence of vegetation index and spatial resolution on a two-date remote sensing-derived relation to C4 species coverage", *Remote Sensing of Environment*, vol. 75, no. 1, pp. 138-151.

de Asis, A.M. & Omasa, K. 2007, "Estimation of vegetation parameter for modeling soil erosion using linear Spectral Mixture Analysis of Landsat ETM data", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, no. 4, pp. 309-324.

Deardorff, J. W. 1978, "Efficient prediction of ground temperature and moisture with inclusion of a layer of vegetation", *J. Geophys. Res.*, vol. 83, pp. 1889–1903.

Doorenbos, J. & Pruitt, W.D. 1977, *Guidelines for predicting crop water requirements: Irrigation and Drainage Paper No. 24*, Rome: Food and Agriculture Organization of the United Nations.

Dorigo, W.A., Zurita-Milla, R., de Wit, A.J., Brazile, J., Singh, R. & Schaepman, M.E. 2007, "A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling", *International journal of applied earth observation and geoinformation*, vol. 9, no. 2, pp. 165-193.

El Hajj, M., Baghdadi, N., Zribi, M., Belaud, G., Cheviron, B., Courault, D. & Charron, F. 2016, "Soil moisture retrieval over irrigated grassland using X-band SAR data", *Remote Sensing of Environment*, vol. 176, pp. 202-218.

Er-Raki, S., Chehbouni, A. & Duchemin, B. 2010, "Combining satellite remote sensing data with the FAO-56 dual approach for water use mapping in irrigated wheat fields of a semi-arid region", *Remote Sensing*, vol. 2, no. 1, pp. 375-387.

Excel, 2016, Microsoft Office.

Flanders, D., Hall-Beyer, M. & Pereverzoff, J. 2003, "Preliminary evaluation of eCognition object-based software for cut block delineation and feature extraction", *Canadian Journal of Remote Sensing,* vol. 29, no. 4, pp. 441–452.

Fox, G.A., Sabbagh, G.J., Searcy, S.W. & Yang, C. 2004, "An automated soil line identification routine for remotely sensed images", *Soil Science Society of America Journal*, vol. 68, no. 4, pp. 1326-1331.

Gao, L., Li, C., Wang, B., Yang, G., Wang, L. & Fu, K. 2016, "Comparison of precision in retrieving soybean leaf area index based on multi-source remote sensing data", *Chinese Journal of Applied Ecology*, vol. 27, no. 1, pp. 191-200.

Gharun, M., Possell, M., Jenkins, M.E., Poon, L.F., Bell, T.L. & Adams, M.A. 2017, "Improving forest sampling strategies for assessment of fuel reduction burning", *Forest Ecology and Management*, vol. 392, pp. 78-89.

Gilabert, M. A., Garcıa-Haro, F. J. & Melia, J. 2000, "A mixture modeling approach to estimate vegetation parameters for heterogeneous canopies in remote sensing" *Remote Sensing of Environment*, vol. 72, no. 3, pp. 328-345.

Gitelson, A.A., Kaufman, Y.J. & Merzlyak, M.N. 1996, "Use of a green channel in remote sensing of global vegetation from EOS-MODIS", *Remote Sensing of Environment*, vol. 58, no. 3, pp. 289-298.

Gitelson, A.A., Vina, A., Ciganda, V., Rundquist, D.C. & Arkebauer, T.J. 2005, "Remote estimation of canopy chlorophyll content in crops", *Geophysical Research Letters*, vol. 32, no. 8.

Glenn, E.P., Nagler, P.L. & Huete, A.R. 2014, "Change Detection Using Vegetation Indices and Multiplatform Satellite Imagery at Multiple Temporal and Spatial Scales" in *Scale Issues in Remote Sensing*, pp. 79-107.

Godínez-Alvarez, H., Herrick, J.E., Mattocks, M., Toledo, D. & Van Zee, J. 2009, "Comparison of three vegetation monitoring methods: their relative utility for ecological assessment and monitoring", *Ecological indicators*, vol. 9, no. 5, pp. 1001-1008.

Goodchild, M. F., & Proctor, J. 1997, "Scale in a digital geographic world", *Geographical and environmental modelling,* 1, 5-24.

Guerrero, J.M., Pajares, G., Montalvo, M., Romeo, J. & Guijarro, M. 2012, "Support Vector Machines for crop/weeds identification in maize fields", Expert Systems with Applications, vol. 39, no. 12, pp. 11149-11155.

Guijarro, M., Pajares, G., Riomoros, I., Herrera, P.J., Burgos-Artizzu, X.P. & Ribeiro, A. 2011, "Automatic segmentation of relevant textures in agricultural images", *Computers and Electronics in Agriculture*, vol. 75, no. 1, pp. 75-83.

Hague, T., Tillett, N.D. & Wheeler, H. 2006, "Automated crop and weed monitoring in widely spaced cereals", *Precision Agriculture*, vol. 7, no. 1, pp. 21-32.

Hay, G.J. & Castilla, G. 2008, Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. In: *Blaschke, T., Lang, S. & Hey, G. J. (eds.), Object-based image analysis*. Springer Berlin Heidelberg, pp. 75-89.

Huete, A.R. 1988, "A soil-adjusted vegetation index (SAVI)", *Remote sensing of environment*, vol. 25, no.3, pp. 295-309.

Inglada, J. & Christophe, E. 2009, "The orfeo toolbox remote sensing image processing software*", International Geoscience and Remote Sensing Symposium (IGARSS)*.

Jakob, S., Zimmermann, R. & Gloaguen, R. 2017, "The Need for Accurate Geometric and Radiometric Corrections of Drone-Borne Hyperspectral Data for Mineral Exploration: MEPHySTo—A Toolbox for Pre-Processing Drone-Borne Hyperspectral Data", *Remote Sensi*ng, vol. 9, no. 1, pp. 88-105.

Jia, K., Liang, S., Liu, S., Li, Y., Xiao, Z., Yao, Y., Jiang, B., Zhao, X., Wang, X., Xu, S. & Cui, J. 2015, "Global land surface fractional vegetation cover estimation using general regression neural networks from MODIS surface reflectance", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 9, pp. 4787-4796.

Jing, X., Yao, W.Q., Wang, J.H. & Song, X.Y. 2011, "A study on the relationship between dynamic change of vegetation coverage and precipitation in Beijing's mountainous areas during the last 20 years", *Mathematical and Computer Modelling,* vol. 54, no. 3, pp. 1079-1085.

Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., Weiss, M. & Baret, F. 2004, "Review of methods for in situ leaf area index determination: Part I. Theories, sensors and hemispherical photography", *Agricultural and forest meteorology*, vol. 121, no. 1, pp. 19-35.

Jordan, C.F. 1969, "Derivation of leaf-area index from quality of light on the forest floor", *Ecology*, vol. 50, no. 4, pp. 663-666.

Kaiser, R., Spiegel, P. B., Henderson, A. K., & Gerber, M. L. 2003, "The application of geographic information systems and global positioning systems in humanitarian emergencies: lessons learned, programme implications and future research", *Disasters*, vol. 27, no. 2, pp. 127-140.

Kataoka, T., Kaneko, T., Okamoto, H. & Hata, S. 2003, "Crop growth estimation system using machine vision", *IEEE/ASME International Conference on Advanced Intelligent Mechatronics, AIM*, pp. 1079-1083.

Kimes, D.S., Knyazikhin, Y., Privette, J.L., Abuelgasim, A.A. & Gao, F. 2000, "Inversion methods for physically-based models", *Remote Sensing Reviews*, vol. 18, no. 2, pp. 381-439.

Korhonen, L., Korhonen, K.T., Rautiainen, M. & Stenberg, P. 2006, "Estimation of forest canopy cover: a comparison of field measurement techniques", *Silva Fennica*, vol. 40, no. 4, pp. 577-588.

Laliberte, A.S., Rango, A., Herrick, J.E., Fredrickson, E.L. & Burkett, L. 2007, "An object-based image analysis approach for determining fractional cover of senescent and green vegetation with digital plot photography", *Journal of Arid Environments*, vol. 69, no. 1, pp. 1-14.

Lary, D.J., Alavi, A.H., Gandomi, A.H. & Walker, A.L. 2016, "Machine learning in geosciences and remote sensing", *Geoscience Frontiers*, vol. 7, no. 1, pp. 3-10.

Li, W., Niu, Z., Chen, H., Li, D., Wu, M. & Zhao, W. 2016, "Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system", *Ecological Indicators*, vol. 67, pp. 637-648.

Li, W., Weiss, M., Waldner, F., Defourny, P., Demarez, V., Morin, D., Hagolle, O. and Baret, F. 2015, "A generic algorithm to estimate LAI, FAPAR and FCOVER variables from SPOT4_HRVIR and landsat sensors: evaluation of the consistency and comparison with ground measurements", *Remote Sensing*, vol. 7, no. 11, pp. 15494-15516.

Liang, S. 2005, *Quantitative remote sensing of land surfaces*. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Liaw, A., & Wiener, M. 2002, "Classification and regression by randomForest", *R news*, vol. 2, no. 3, pp. 18-22.

Logsdon, J.M., Launius, R.D., Onkst, D.H., & Garber, S.J. 1998, *Exploring the Unknown: Selected Documents in the History of the US Civilian Space Program. Volume 3: Using Space*.

Lukas, V., Novák, J., Neudert, L., Svobodova, I., Rodriguez-Moreno, F., Edrees, M. & Kren, J. 2016, "The combination of UAV survey and Landsat imagery for monitoring of crop vigor in precision agriculture", *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, pp. 953-957.

Luscier, J.D., Thompson, W.L., Wilson, J.M., Gorham, B.E. & Dragut, L.D. 2006, "Using digital photographs and object-based image analysis to estimate percent ground cover in vegetation plots", *Frontiers in Ecology and the Environment*, vol. 4, no. 8, pp. 408-413.

Major, D.J., Baret, F. & Guyot, G. 1990, "A ratio vegetation index adjusted for soil brightness", *International Journal of Remote Sensing*, vol. 11, no. 5, pp. 727-740.

Matese, A., Toscano, P., Di Gennaro, S.F., Genesio, L., Vaccari, F.P., Primicerio, J., Belli, C., Zaldei, A., Bianconi, R. & Gioli, B. 2015, "Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture", *Remote Sensing*, vol. 7, no. 3, pp. 2971-2990.

McBratney, A., Whelan, B., Ancev, T. & Bouma, J. 2005 "Future directions of precision agriculture", *Precision agriculture*, vol. 6, no. 1, pp. 7-23.

Meyer, G.E. & Neto, J.C. 2008, "Verification of color vegetation indices for automated crop imaging applications", *Computers and Electronics in Agriculture*, vol. 63, no. 2, pp. 282-293.

Meyer, G.E., Hindman, T. & Laksmi, K. 1999, "Machine vision detection parameters for plant species identification", *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 3543, pp. 327-335.

Meyer, G.E., Hindman, T.W., Jones, D.D. & Mortensen, D.A. 2004, "Digital camera operation and fuzzy logic classification of uniform plant, soil, and residue color images", *Applied Engineering in Agriculture*, vol. 20, no. 4, pp. 519-529.

Meyer, G.E., Neto, J.C., Jones, D.D. & Hindman, T.W. 2004, "Intensified fuzzy clusters for classifying plant, soil, and residue regions of interest from color images", *Computers and Electronics in Agriculture*, vol. 42, no. 3, pp. 161-180.

Miniature Multiple Camera Array (Mini-MCA), 2017, Chatsworth, CA: Tetracam Inc.

Mobley, C.D., Sundman, L.K., Davis, C.O., Bowles, J.H., Downes, T.V., Leathers, R.A., Montes, M.J., Bissett, W.P., Kohler, D.D.R., Reid, R.P., Louchard, E.M. & Gleason, A. 2005, "Interpretation of hyperspectral remote-sensing imagery by spectrum matching and look-up tables", *Applied Optics*, vol. 44, no. 17, pp. 3576-3592.

Moisen, G. G. 2008, Classification and Regression Trees. In: *Jørgensen, S. E. & Fath, B. D. (eds.), Encyclopedia of Ecology*. Amsterdam: Elsevier, pp. 582-588.

Moran, M.S., Inoue, Y., & Barnes, E.M. 1997, "Opportunities and limitations for image-based remote sensing in precision crop management" *Remote sensing of Environment*, vol. 61, no. 3, pp. 319-346.

Mougin, E., Demarez, V., Diawara, M., Hiernaux, P., Soumaguel, N. & Berg, A. 2014, "Estimation of LAI, fAPAR and fCover of Sahel rangelands (Gourma, Mali) ", *Agricultural and Forest Meteorology*, vol. 198, pp. 155-167.

Mulla, D.J. 2013, "Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps", *Biosystems engineering*, vol. 114, no. 4, pp. 358-371.

Navia, J., Mondragon, I., Patino, D. & Colorado, J. 2016, "Multispectral mapping in agriculture: Terrain mosaic using an autonomous quadcopter UAV", *2016 International Conference on Unmanned Aircraft Systems, ICUAS 2016*, pp. 1351-1358.

North, P.R. 2002, "Estimation of f APAR, LAI, and vegetation fractional cover from ATSR-2 imagery", *Remote sensing of Environment*, vol. 80, no. 1, pp. 114-121.

O'Callaghan, J.R., Hossain, A.H.M.S. & Wyseure, G.C.L. 1994, "Some factors limiting the yield of wheat in Bangladesh", *Computers and Electronics in Agriculture*, vol. 11, no. 4, pp. 337-345.

Opitz, D. & Blundell, S. 2008, Object recognition and image segmentation: the Feature Analyst® approach. In: *Blaschke, T., Lang, S. & Hey, G. J. (eds.), Object-based image analysis*. Springer Berlin Heidelberg, pp. 153-167.

Orlando, F., Movedi, E., Coduto, D., Parisi, S., Brancadoro, L., Pagani, V., Guarneri, T. & Confalonieri, R. 2016, "Estimating leaf area index (LAI) in vineyards using the pocketLAI smart-app", *Sensors (Switzerland),* vol. 16, no. 12.

Pahl, M., Kux, H.J.H. & Feitosa, R.Q. 2008, Arquitetura de um sistema baseado em conhecimento para a interpretaçao de dados de sensoriamento remoto de múltiplos sensores (Doctoral dissertation, INPE Sao José dos Campos).

Panagos, P., Borrelli, P., Meusburger, K., Alewell, C., Lugato, E. & Montanarella, L. 2015, "Estimating the soil erosion cover-management factor at the European scale", *Land Use Policy*, vol. 48, pp. 38-50.

Perez, A.J., Lopez, F., Benlloch, J.V., Christensen, S. 2000, "Color and shape analysis techniques for weed detection in cereal fields", *Computer and Electronics in Agriculture*, vol. 25, pp.197–212.

Pfeifer, M., Kor, L., Nilus, R., Turner, E., Cusack, J., Lysenko, I., Khoo, M., Chey, V.K., Chung, A.C. & Ewers, R.M. 2016, "Mapping the structure of Borneo's tropical forests across a degradation gradient", *Remote Sensing of Environment*, vol. 176, pp. 84-97.

Pinty, B. & Verstraete, M.M. 1992, "GEMI: a non-linear index to monitor global vegetation from satellites", *Plant ecology*, vol. 101, no. 1, pp. 15-20.

Pongnumkul, S., Chaovalit, P., & Surasvadi, N. 2015, "Applications of smartphone-based sensors in agriculture: a systematic review of research", *Journal of Sensors*.

Prasad, A.M., Iverson, L.R. & Liaw, A. 2006, "Newer classification and regression tree techniques: bagging and random forests for ecological prediction", *Ecosystems*, vol. 9, no. 2, pp. 181-199.

Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H. & Sorooshian, S. 1994, "A modified soil adjusted vegetation index", *Remote Sensing of Environment*, vol. 48, no. 2, pp. 119-126.

Qu, Y., Meng, J., Wan, H. & Li, Y. 2016, "Preliminary study on integrated wireless smart terminals for leaf area index measurement", *Computers and Electronics in Agriculture*, vol. 129, pp. 56-65.

Rahman, S., Rahman, M.S. & Rahman, M.H. 2012, "Joint determination of the choice of growing season and economic efficiency of maize in Bangladesh", *Journal of the Asia Pacific Economy*, vol. 17, no. 1, pp. 138-150.

Rasmussen, J., Ntakos, G., Nielsen, J., Svensgaard, J., Poulsen, R.N. & Christensen, S. 2016, "Are vegetation indices derived from consumer-grade cameras mounted on UAVs sufficiently reliable for assessing experimental plots?", *European Journal of Agronomy*, vol. 74, pp. 75-92.

Rebordão, J.M. 1989, "Lookup Table Loadings for Image Processing with Controlled Knots", *Computer Vision, Graphics and Image Processing*, vol. 47, no. 2, pp. 189-202.

Ren, D.D., Tripathi, S. & Li, L.K. 2017, "Low-cost multispectral imaging for remote sensing of lettuce health", *Journal of Applied Remote Sensing*, vol. 11, no. 1.

Richardson, A. J. & Weigand, C. 1977, "Distinguishing Vegetation from Soil Background Information", *Photogrammetric Engineering and Remote Sensing*, vol. 43, no. 12, pp. 1541–1552.

Roujean, J.L. & Breon, F.M. 1995, "Estimating PAR absorbed by vegetation from bidirectional reflectance measurements", *Remote Sensing of Environment*, vol. 51, no. 3, pp. 375-384.

Roumiguié, A., Sigel, G., Poilvé, H., Bouchard, B., Vrieling, A. & Jacquin, A. 2016, "Insuring forage through satellites: testing alternative indices against grassland production estimates for France", *International Journal of Remote Sensing,* vol. 38, no. 7. pp. 1912-1939.

Salamí, E., Barrado, C. & Pastor, E. 2014, "UAV flight experiments applied to the remote sensing of vegetated areas", *Remote Sensing*, vol. 6, no. 11, pp. 11051-11081.

Samani Majd, A.M., Bleiweiss, M.P., DuBois, D. & Shukla, M.K. 2013, "Estimation of the fractional canopy cover of pecan orchards using Landsat 5 satellite data, aerial imagery, and orchard floor photographs", *International journal of remote sensing*, vol. 34, no. 16, pp. 5937-5952.

Schiewe, J. 2002, "Segmentation of high-resolution remotely sensed data-concepts, applications and problems", *International Archives of Photogrammetry Remote Sensing And Spatial Information Sciences*, vol. 34, no. 4, pp. 380-385.

Schmidhuber, J. & Tubiello, F.N. 2007, "Global food security under climate change", *Proceedings of the National Academy of Sciences*, vol. 104, no. 50, pp. 19703-19708.

Seelan, S.K., Laguette, S., Casady, G.M. & Seielstad, G.A. 2003, "Remote sensing applications for precision agriculture: A learning community approach", *Remote Sensing of Environment*, vol. 88, no. 1, pp. 157-169.

Seyyedhasani, H., Dvorak, J.S., Sama, M.P. & Stombaugh, T.S. 2016, "Mobile Device-Based Location Services Accuracy", *Applied Engineering in Agriculture*, vol. 32, no. 5, pp. 539-547.

Shahid, S. 2010, "Recent trends in the climate of Bangladesh", *Climate Research*, vol. 42, no. 3, pp. 185-193.

Somers, B., Delalieux, S., Verstraeten, W.W. & Coppin, P. 2009, "A conceptual framework for the simultaneous extraction of sub-pixel spatial extent and spectral characteristics of crops", *Photogrammetric Engineering & Remote Sensing*, vol. 75, no. 1, pp. 57-68.

Verrelst, J., Muñoz, J., Alonso, L., Delegido, J., Rivera, J.P., Camps-Valls, G. & Moreno, J. 2012, "Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and-3", *Remote Sensing of Environment*, vol. 118, pp. 127-139.

Verstraete, M.M. & Pinty, B. 1996, "Designing optimal spectral indexes for remote sensing applications", *IEEE Transactions on Geoscience and Remote Sensing,* vol. 34, no. 5, pp. 1254-1265.

Vesali, F., Omid, M., Mobli, H. & Kaleita, A. 2017, "Feasibility of using smart phones to estimate chlorophyll content in corn plants", *Photosynthetica*, vol. 55, no. 4, pp. 603-610.

Vittinghoff, E., Glidden, D. V., Shiboski, S. C. & McCulloch, C. E. 2011, "Regression methods in biostatistics: linear, logistic, survival, and repeated measures models", *Springer Science & Business Media*, pp. 69-131.

Vrieling, A., De Jong, S.M., Sterk, G. & Rodrigues, S.C. 2008, "Timing of erosion and satellite data: A multi-resolution approach to soil erosion risk mapping", *International Journal of Applied Earth Observation and Geoinformation*, vol. 10, no. 3, pp. 267-281.

Wang, C., Meng, Q., Zhan, Y., Peng, J., Wei, X., Yang, J. & Li, J. 2015, "Ground sampling methods for surface soil moisture in heterogeneous pixels", *Environmental Earth Sciences*, vol. 73, no. 10, pp. 6427-6436.

Wang, G., Wente, S., Gertner, G.Z. & Anderson, A. 2002, "Improvement in mapping vegetation cover factor for the universal soil loss equation by geostatistical methods with Landsat Thematic Mapper images", *International Journal of Remote Sensing*, vol. 23, no. 18, pp. 3649-3667.

Weiss, B. & Baret, F. 2016, "CAN_EYE V6.4.6 User Manual", [online] Available at: https://www6.paca.inra.fr/can-eye/Documentation/Documentation [accessed 4 August 2017].

Welte, J., Ault, A., Bowman, C., Ellis, S., Buckmaster, D., Ess, D. & Krogmeier, J. 2013, "An approach to farm management information systems using task-specific, collaborative mobile apps and cloud storage services", *ASAE Annual International Meeting 2013, ASABE 2013*, pp. 258-281.

Woebbecke, D.M., Meyer, G.E., Von Bargen, K., Mortensen, D.A. 1995, "Color indices for weed identification under various soil, residue and lighting conditions", *Transactions of the ASAE*, vol. 38, no. 1, pp. 259–269.

Woodcock, C. E., & Strahler, A. H. 1987, "The factor of scale in remote sensing", *Remote sensing of Environment*, vol. 21, no. 3, pp. 311-332.

Woodgate, W., Soto-Berelov, M., Suarez, L., Jones, S., Hill, M., Wilkes, P., Axelsson, C., Haywood, A. & Mellor, A. 2012, "Searching for the optimal sampling design for measuring LAI in an upland rainforest", *CEUR Workshop Proceedings*.

Wu, W. 2011, "Derivation of tree canopy cover by multiscale remote sensing approach", *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, pp. 142-149.

Xiao, Z., Wang, T., Liang, S. & Sun, R. 2016, "Estimating the fractional vegetation cover from GLASS leaf area index product", *Remote Sensing*, vol. 8, no. 4, pp. 337-356.

XnView Version 2.40, 2017, Reims, France: XnSoft.

Yang, C., Everitt, J.H. & Du, Q. 2010, "Applying linear spectral unmixing to airborne hyperspectral imagery for mapping yield variability in grain sorghum and cotton fields", *Journal of Applied Remote Sensing*, vol. 4, no. 1.

Zeng, X., Dickinson, R. E., Walker, A., Shaikh, M., DeFries, R. S. & Qi, J. 2000, "Derivation and evaluation of global 1-km fractional vegetation cover data for land modeling", *Journal of Applied Meteorology*, vol. 39, no. 6, pp. 826-839.

Zhang, X., Liao, C., Li, J. & Sun, Q. 2013, "Fractional vegetation cover estimation in arid and semi-arid environments using HJ-1 satellite hyperspectral data", *International Journal of Applied Earth Observation and Geoinformation*, vol. 21, pp. 506-512.

Zheng, H., Zhou, X., Cheng, T., Yao, X., Tian, Y., Cao, W. & Zhu, Y. 2016, "Evaluation of a UAV-based hyperspectral frame camera for monitoring the leaf nitrogen concentration in rice", *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 7350-7353.

Zhongming, W., Lees, B.G., Feng, J., Wanning, L. & Haijing, S. 2010, "Stratified vegetation cover index: A new way to assess vegetation impact on soil erosion", *Catena*, vol. 83, no. 1, pp. 87-93.

Zhou, P., Luukkanen, O., Tokola, T. & Nieminen, J. 2008, "Effect of vegetation cover on soil erosion in a mountainous watershed", *Catena*, vol. 75, no. 3, pp. 319-325.

Zurita-Milla, R., Laurent, V.C.E. & van Gijsel, J.A.E. 2015, "Visualizing the ill-posedness of the inversion of a canopy radiative transfer model: A case study for Sentinel-2", *International Journal of Applied Earth Observation and Geoinformation*, vol. 43, pp. 7-18.

Appendix A

Place	Date	RGB	Multispectral	Resolution RGB (cm)	Resolution multispectral (cm)
Barisal	20160113	0	100	n/a	5,14
Barisal	20160113	0	100	n/a	5,11
Barisal	20160123	65	0	1,15	n/a
Barisal	20160203	12	100	0,45	5,34
Barisal	20160302	65	0	1,20	n/a
Barisal	20160331	0	100	n/a	5,44
Barisal	20160413	0	100	n/a	5,11
Barisal	20160424	0	100	n/a	5,07
Kalapara	20151229	12	0	1,21	n/a
Kalapara	20160120	12	100	0,20	5,21
Kalapara	20160120	12	100	0,20	5,43
Kalapara	20160131	12	100	0,23	5,28
Kalapara	20160131	65	100	1,23	5,28
Kalapara	20160209	7	0	0,12	n/a
Kalapara	20160209	12	0	0,21	n/a
Kalapara	20160303	0	100	n/a	5,57
Kalapara	20160323	0	100	n/a	5,09
Kalapara	20160405	0	100	n/a	5,34
Kalapara	20160420	0	100	n/a	5,21
Patuakhali	20160106	65	0	1,17	n/a
Patuakhali	20160117	12	0	0,30	n/a
Patuakhali	20160117	65	0	1,21	n/a
Patuakhali	20160127	12	100	0,32	5,29
Patuakhali	20160127	12	100	0,32	5,05
Patuakhali	20160127	65	100	8,83	5,29
Patuakhali	20160127	65	100	8,83	5,05
Patuakhali	20160207	12	100	0,30	5,09
Patuakhali	20160207	12	100	0,30	5,07
Patuakhali	20160304	65	100	1,31	5,53
Patuakhali	20160316	65	0	1,32	n/a
Patuakhali	20160407	0	100	n/a	5,36
Patuakhali	20160418	0	100	n/a	5,08
Patuakhali	20160418	0	100	n/a	5,25

Appendix B

```
Matrix R(), G(), B(), ExG, ExR, ExGR, MExG, NDI, INDI, CIVE, COM1, COM2, VEG;
Point p;

ExG  = R;
ExR = R;
ExGR = R;
MExG = R;
NDI = R;
INDI = R;
CIVE = R;
COM1 = R;
COM2 = R;
VEG = R;

foreach p in R do
{
  ExG[p]  = 2*G[p] - R[p] - B[p];
  ExR[p] = 1.3*R[p] - G[p];
  ExGR[p] = 3*G[p] - 2.3*R[p] - B[p];
  MExG[p] = 1.262*G[p] - 0.844*R[p] - 0.311*B[p];
  NDI[p] = (G[p] - R[p])/(G[p] + R[p]);
  INDI[p] = 128*((G[p] - R[p])/(G[p] + R[p]) + 1);
  CIVE[p] = 0.441*R[p] - 0.811*G[p] + 0.385*B[p] + 18.78745;
  COM1[p] = 0.25*(2*G[p] - R[p] - B[p]) + 0.3*(3*G[p] - 2.3*R[p] - B[p]) + 0.33*(0.441*R[p] - 0.811*G[p] +
0.385*B[p] +18.78745) + 0.12*(G[p]/(R[p]^0.667 + B[p]^0.333));
  COM2[p] = 0.36*(2*G[p] - R[p] - B[p]) + 0.47*(0.441*R[p] - 0.811*G[p] + 0.385*B[p] +18.78745) +
0.17*(G[p]/(R[p]^0.667 + B[p]^0.333));
  VEG[p] = G[p]/(R[p]^0.667 + B[p]^0.333);
}

showMatrix(ExG);
showMatrix(ExR);
showMatrix(ExGR);
showMatrix(MExG);
showMatrix(NDI);
showMatrix(INDI);
showMatrix(CIVE);
showMatrix(COM1);
showMatrix(COM2);
showMatrix(VEG);
```

Appendix C

```
Matrix R(), G(), NIR(), R1(), R2(), SR, DVI, NDVI, RDVI, GNDVI, SAVI, MSAVI, GEMI, GCI, TVI, RECI1, RECI2, PVI,
SAVI2, TSAVI, ATSAVI;
Point p;

SR  = R;
DVI = R;
NDVI = R;
RDVI = R;
GNDVI = R;
SAVI = R;
MSAVI = R;
GEMI = R;
GCI = R;
TVI = R;
RECI1 = R;
RECI2 = R;
PVI = R;
SAVI2 = R;
TSAVI = R;
ATSAVI = R;


foreach p in R do
{
  SR[p]  = NIR[p]/R[p];
  DVI[p] = NIR[p] - R[p];
  NDVI[p] = (NIR[p] - R[p])/(NIR[p] + R[p]);
  RDVI[p] = (NIR[p] - R[p])/((NIR[p] + R[p])^(1/2));
  GNDVI[p] = (NIR[p] - G[p])/(NIR[p] + G[p]);
  SAVI[p] = ((NIR[p] - R[p]) *1.5)/(NIR[p] + R[p] + 0.5);
  MSAVI[p] = (2*NIR[p] + 1 -((2*NIR[p] + 1)^2 - 8*(NIR[p] - R[p]))^(1/2))/2;
  GEMI[p] = ((2*(NIR[p]^2 - R[p]^2) + 1.5*NIR[p] + 0.5*R[p])/(NIR[p] + R[p] + 0.5))*(1 - 0.25*((2*(NIR[p]^2 -
R[p]^2) + 1.5*NIR[p] + 0.5*R[p])/(NIR[p] + R[p] + 0.5))) -((R[p] - 0.125)/(1 - R[p]));
  GCI[p] = NIR[p]/R[p] - 1;
  TVI[p] = 60*(NIR[p] - G[p]) - 100*(R[p] - G[p]);
  RECI1[p] = NIR[p]/R1[p] - 1;
  RECI2[p] = NIR[p]/R2[p] - 1;
  PVI[p] = (NIR[p] - 0.85*R[p] - 27.06)/1.31;
  SAVI2[p] = NIR[p]/(R[p] - 31.84);
  TSAVI[p] = (0.85*NIR[p] - 0.72*R[p] - 23)/(0.85*NIR[p] + R[p] - 23);
  ATSAVI[p] = (-0.72*R[p] - 23)/(0.85*NIR[p] + R[p] -22.86);
}

showMatrix(SR);
showMatrix(DVI);
showMatrix(NDVI);
showMatrix(RDVI);
showMatrix(GNDVI);
showMatrix(SAVI);
showMatrix(MSAVI);
showMatrix(GEMI);
showMatrix(GCI);
showMatrix(TVI);
showMatrix(RECI1);
showMatrix(RECI2);
showMatrix(PVI);
showMatrix(SAVI2);
showMatrix(TSAVI);
showMatrix(ATSAVI);
```

Appendix D

```
# import data
dataset = read.csv("C:/Users/nikolina/Desktop/GIMA/Module 7 - Master thesis/Smartphone_dataset.csv", header =
TRUE, sep = ";", dec = ",")

#assign names to variables
ID<-dataset$ID
date_1<-dataset$Date_1
date_2<-dataset$Date_2
device<-dataset$Device
fNumber<-dataset$FNumber
fCover<-dataset$fCover

ExG_mean<-dataset$ExG_mean
ExG_min<-dataset$ExG_min
ExG_max<-dataset$ExG_max
ExG_std<-dataset$ExG_std

ExR_mean<-dataset$ExR_mean
ExR_min<-dataset$ExR_min
ExR_max<-dataset$ExR_max
ExR_std<-dataset$ExR_std

ExGR_mean<-dataset$ExGR_mean
ExGR_min<-dataset$ExGR_min
ExGR_max<-dataset$ExGR_max
ExGR_std<-dataset$ExGR_std

MExG_mean<-dataset$MExG_mean
MExG_min<-dataset$MExG_min
MExG_max<-dataset$MExG_max
MExG_std<-dataset$MExG_std

NDI_mean<-dataset$NDI_mean
NDI_min<-dataset$NDI_min
NDI_max<-dataset$NDI_max
NDI_std<-dataset$NDI_std

INDI_mean<-dataset$INDI_mean
INDI_min<-dataset$INDI_min
INDI_max<-dataset$INDI_max
INDI_std<-dataset$INDI_std

CIVE_mean<-dataset$CIVE_mean
CIVE_min<-dataset$CIVE_min
CIVE_max<-dataset$CIVE_max
CIVE_std<-dataset$CIVE_std

COM1_mean<-dataset$COM1_mean
COM1_min<-dataset$COM1_min
COM1_max<-dataset$COM1_max
COM1_std<-dataset$COM1_std

COM2_mean<-dataset$COM2_mean
COM2_min<-dataset$COM2_min
COM2_max<-dataset$COM2_max
COM2_std<-dataset$COM2_std

VEG_mean<-dataset$VEG_mean
VEG_min<-dataset$VEG_min
VEG_max<-dataset$VEG_max
VEG_std<-dataset$VEG_std

# Random Forest
library(randomForest)
RF.model<-randomForest(fCover~ExG_mean+ExG_min+ExG_max+ExG_std+ExR_mean+ExR_min+ExR_max+ExR_std+
                       ExGR_mean+ExGR_min+ExGR_max+ExGR_std+MExG_mean+MExG_min+MExG_max+MExG_std+
                       NDI_mean+NDI_min+NDI_max+NDI_std+INDI_mean+INDI_min+INDI_max+INDI_std+
```

```
                         CIVE_mean+CIVE_min+CIVE_max+CIVE_std+COM1_mean+COM1_min+COM1_max+COM1_std+
                         COM2_mean+COM2_min+COM2_max+COM2_std+VEG_mean+VEG_min+VEG_max+VEG_std, data=dataset)
importance(RF.model)
save(RF.model, file = "RF.rda")
plot(RF.model)
print(RF.model)


# Predicted fCover - training dataset
dataMatrix<-cbind(ExG_mean, ExG_min, ExG_max, ExG_std, ExR_mean, ExR_min, ExR_max, ExR_std, ExGR_mean,
                  ExGR_min, ExGR_max, ExGR_std, MExG_mean, MExG_min, MExG_max, MExG_std, NDI_mean, NDI_min,
                  NDI_max, NDI_std, INDI_mean, INDI_min, INDI_max, INDI_std, CIVE_mean, CIVE_min, CIVE_max,
                  CIVE_std, COM1_mean, COM1_min, COM1_max, COM1_std, COM2_mean, COM2_min, COM2_max, COM2_std,
                  VEG_mean, VEG_min, VEG_max, VEG_std)
dataSmartphone<-data.frame(dataMatrix)
pred_fCover<-predict(RF.model, dataSmartphone)
write.csv(pred_fCover, file = "C:/Users/nikolina/Desktop/GIMA/SM_fCover_pred.csv")


# Predicted fCover - UAV dataset
UAVdataset = read.csv("C:/Users/nikolina/Desktop/GIMA/Module 7 - Master thesis/UAV_dataset.csv", header = TRUE,
sep = ";", dec = ",")

ExG_mean<-UAVdataset$ExG_mean
ExG_min<-UAVdataset$ExG_min
ExG_max<-UAVdataset$ExG_max
ExG_std<-UAVdataset$ExG_std

ExR_mean<-UAVdataset$ExR_mean
ExR_min<-UAVdataset$ExR_min
ExR_max<-UAVdataset$ExR_max
ExR_std<-UAVdataset$ExR_std

ExGR_mean<-UAVdataset$ExGR_mean
ExGR_min<-UAVdataset$ExGR_min
ExGR_max<-UAVdataset$ExGR_max
ExGR_std<-UAVdataset$ExGR_std

MExG_mean<-UAVdataset$MExG_mean
MExG_min<-UAVdataset$MExG_min
MExG_max<-UAVdataset$MExG_max
MExG_std<-UAVdataset$MExG_std

NDI_mean<-UAVdataset$NDI_mean
NDI_min<-UAVdataset$NDI_min
NDI_max<-UAVdataset$NDI_max
NDI_std<-UAVdataset$NDI_std

INDI_mean<-UAVdataset$INDI_mean
INDI_min<-UAVdataset$INDI_min
INDI_max<-UAVdataset$INDI_max
INDI_std<-UAVdataset$INDI_std

CIVE_mean<-UAVdataset$CIVE_mean
CIVE_min<-UAVdataset$CIVE_min
CIVE_max<-UAVdataset$CIVE_max
CIVE_std<-UAVdataset$CIVE_std

COM1_mean<-UAVdataset$COM1_mean
COM1_min<-UAVdataset$COM1_min
COM1_max<-UAVdataset$COM1_max
COM1_std<-UAVdataset$COM1_std

COM2_mean<-UAVdataset$COM2_mean
COM2_min<-UAVdataset$COM2_min
COM2_max<-UAVdataset$COM2_max
COM2_std<-UAVdataset$COM2_std

VEG_mean<-UAVdataset$VEG_mean
VEG_min<-UAVdataset$VEG_min
VEG_max<-UAVdataset$VEG_max
VEG_std<-UAVdataset$VEG_std
```

```
dataMatrix1<-cbind(ExG_mean, ExG_min, ExG_max, ExG_std, ExR_mean, ExR_min, ExR_max, ExR_std, ExGR_mean,
                   ExGR_min, ExGR_max, ExGR_std, MExG_mean, MExG_min, MExG_max, MExG_std, NDI_mean, NDI_min,
                   NDI_max, NDI_std, INDI_mean, INDI_min, INDI_max, INDI_std, CIVE_mean, CIVE_min, CIVE_max,
                   CIVE_std, COM1_mean, COM1_min, COM1_max, COM1_std, COM2_mean, COM2_min, COM2_max, COM2_std,
                   VEG_mean, VEG_min, VEG_max, VEG_std)

dataUAV<-data.frame(dataMatrix1)
pred_fCover_UAV<-predict(RF.model, dataUAV)
write.csv(pred_fCover_UAV, file = "C:/Users/nikolina/Desktop/GIMA/UAV_fCover_pred.csv")

# Train model with UAV dataset (NIR indices included)
UAVdataset_NIR = read.csv("C:/Users/nikolina/Desktop/GIMA/Module 7 - Master thesis/UAV_dataset_rgb12.csv",
header = TRUE, sep = ";", dec = ",", blank.lines.skip = FALSE)

fCover<-UAVdataset_NIR$fCover_rgb

ExG_mean<-UAVdataset_NIR$ExG_mean
ExG_min<-UAVdataset_NIR$ExG_min
ExG_max<-UAVdataset_NIR$ExG_max
ExG_std<-UAVdataset_NIR$ExG_std

ExR_mean<-UAVdataset_NIR$ExR_mean
ExR_min<-UAVdataset_NIR$ExR_min
ExR_max<-UAVdataset_NIR$ExR_max
ExR_std<-UAVdataset_NIR$ExR_std

ExGR_mean<-UAVdataset_NIR$ExGR_mean
ExGR_min<-UAVdataset_NIR$ExGR_min
ExGR_max<-UAVdataset_NIR$ExGR_max
ExGR_std<-UAVdataset_NIR$ExGR_std

MExG_mean<-UAVdataset_NIR$MExG_mean
MExG_min<-UAVdataset_NIR$MExG_min
MExG_max<-UAVdataset_NIR$MExG_max
MExG_std<-UAVdataset_NIR$MExG_std

NDI_mean<-UAVdataset_NIR$NDI_mean
NDI_min<-UAVdataset_NIR$NDI_min
NDI_max<-UAVdataset_NIR$NDI_max
NDI_std<-UAVdataset_NIR$NDI_std

INDI_mean<-UAVdataset_NIR$INDI_mean
INDI_min<-UAVdataset_NIR$INDI_min
INDI_max<-UAVdataset_NIR$INDI_max
INDI_std<-UAVdataset_NIR$INDI_std

CIVE_mean<-UAVdataset_NIR$CIVE_mean
CIVE_min<-UAVdataset_NIR$CIVE_min
CIVE_max<-UAVdataset_NIR$CIVE_max
CIVE_std<-UAVdataset_NIR$CIVE_std

COM1_mean<-UAVdataset_NIR$COM1_mean
COM1_min<-UAVdataset_NIR$COM1_min
COM1_max<-UAVdataset_NIR$COM1_max
COM1_std<-UAVdataset_NIR$COM1_std

COM2_mean<-UAVdataset_NIR$COM2_mean
COM2_min<-UAVdataset_NIR$COM2_min
COM2_max<-UAVdataset_NIR$COM2_max
COM2_std<-UAVdataset_NIR$COM2_std

VEG_mean<-UAVdataset_NIR$VEG_mean
VEG_min<-UAVdataset_NIR$VEG_min
VEG_max<-UAVdataset_NIR$VEG_max
VEG_std<-UAVdataset_NIR$VEG_std

SR_mean<-UAVdataset_NIR$SR_mean
SR_min<-UAVdataset_NIR$SR_min
SR_max<-UAVdataset_NIR$SR_max
SR_std<-UAVdataset_NIR$SR_std
```

```
DVI_mean<-UAVdataset_NIR$DVI_mean
DVI_min<-UAVdataset_NIR$DVI_min
DVI_max<-UAVdataset_NIR$DVI_max
DVI_std<-UAVdataset_NIR$DVI_std

NDVI_mean<-UAVdataset_NIR$NDVI_mean
NDVI_min<-UAVdataset_NIR$NDVI_min
NDVI_max<-UAVdataset_NIR$NDVI_max
NDVI_std<-UAVdataset_NIR$NDVI_std

RDVI_mean<-UAVdataset_NIR$RDVI_mean
RDVI_min<-UAVdataset_NIR$RDVI_min
RDVI_max<-UAVdataset_NIR$RDVI_max
RDVI_std<-UAVdataset_NIR$RDVI_std

GNDVI_mean<-UAVdataset_NIR$GNDVI_mean
GNDVI_min<-UAVdataset_NIR$GNDVI_min
GNDVI_max<-UAVdataset_NIR$GNDVI_max
GNDVI_std<-UAVdataset_NIR$GNDVI_std

SAVI_mean<-UAVdataset_NIR$SAVI_mean
SAVI_min<-UAVdataset_NIR$SAVI_min
SAVI_max<-UAVdataset_NIR$SAVI_max
SAVI_std<-UAVdataset_NIR$SAVI_std

MSAVI_mean<-UAVdataset_NIR$MSAVI_mean
MSAVI_min<-UAVdataset_NIR$MSAVI_min
MSAVI_max<-UAVdataset_NIR$MSAVI_max
MSAVI_std<-UAVdataset_NIR$MSAVI_std

GEMI_mean<-UAVdataset_NIR$GEMI_mean
GEMI_min<-UAVdataset_NIR$GEMI_min
GEMI_max<-UAVdataset_NIR$GEMI_max
GEMI_std<-UAVdataset_NIR$GEMI_std

GCI_mean<-UAVdataset_NIR$GCI_mean
GCI_min<-UAVdataset_NIR$GCI_min
GCI_max<-UAVdataset_NIR$GCI_max
GCI_std<-UAVdataset_NIR$GCI_std

TVI_mean<-UAVdataset_NIR$TVI_mean
TVI_min<-UAVdataset_NIR$TVI_min
TVI_max<-UAVdataset_NIR$TVI_max
TVI_std<-UAVdataset_NIR$TVI_std

RECI1_mean<-UAVdataset_NIR$RECI1_mean
RECI1_min<-UAVdataset_NIR$RECI1_min
RECI1_max<-UAVdataset_NIR$RECI1_max
RECI1_std<-UAVdataset_NIR$RECI1_std

RECI2_mean<-UAVdataset_NIR$RECI2_mean
RECI2_min<-UAVdataset_NIR$RECI2_min
RECI2_max<-UAVdataset_NIR$RECI2_max
RECI2_std<-UAVdataset_NIR$RECI2_std

PVI_mean<-UAVdataset_NIR$PVI_mean
PVI_min<-UAVdataset_NIR$PVI_min
PVI_max<-UAVdataset_NIR$PVI_max
PVI_std<-UAVdataset_NIR$PVI_std

SAVI2_mean<-UAVdataset_NIR$SAVI2_mean
SAVI2_min<-UAVdataset_NIR$SAVI2_min
SAVI2_max<-UAVdataset_NIR$SAVI2_max
SAVI2_std<-UAVdataset_NIR$SAVI2_std

TSAVI_mean<-UAVdataset_NIR$TSAVI_mean
TSAVI_min<-UAVdataset_NIR$TSAVI_min
TSAVI_max<-UAVdataset_NIR$TSAVI_max
TSAVI_std<-UAVdataset_NIR$TSAVI_std

ATSAVI_mean<-UAVdataset_NIR$ATSAVI_mean
```

```
ATSAVI_min<-UAVdataset_NIR$ATSAVI_min
ATSAVI_max<-UAVdataset_NIR$ATSAVI_max
ATSAVI_std<-UAVdataset_NIR$ATSAVI_std

RF.model_NIR<-randomForest(fCover~ExG_mean+ExG_min+ExG_max+ExG_std+ExR_mean+ExR_min+ExR_max+ExR_std+
                           ExGR_mean+ExGR_min+ExGR_max+ExGR_std+MExG_mean+MExG_min+MExG_max+MExG_std+
                           NDI_mean+NDI_min+NDI_max+NDI_std+INDI_mean+INDI_min+INDI_max+INDI_std+
                           CIVE_mean+CIVE_min+CIVE_max+CIVE_std+COM1_mean+COM1_min+COM1_max+COM1_std+
                           COM2_mean+COM2_min+COM2_max+COM2_std+VEG_mean+VEG_min+VEG_max+VEG_std+
                           SR_mean+SR_min+SR_max+SR_std+DVI_mean+DVI_min+DVI_max+DVI_std+
                           NDVI_mean+NDVI_min+NDVI_max+NDVI_std+RDVI_mean+RDVI_min+RDVI_max+RDVI_std+
                           GNDVI_mean+GNDVI_min+GNDVI_max+GNDVI_std+SAVI_mean+SAVI_min+SAVI_max+SAVI_std+
                           MSAVI_mean+MSAVI_min+MSAVI_max+MSAVI_std+GEMI_mean+GEMI_min+GEMI_max+GEMI_std+
                           GCI_mean+GCI_min+GCI_max+GCI_std+TVI_mean+TVI_min+TVI_max+TVI_std+
                           RECI1_mean+RECI1_min+RECI1_max+RECI1_std+RECI2_mean+RECI2_min+RECI2_max+RECI2_std+
                           PVI_mean+PVI_min+PVI_max+PVI_std+SAVI2_mean+SAVI2_min+SAVI2_max+SAVI2_std+

TSAVI_mean+TSAVI_min+TSAVI_max+TSAVI_std+ATSAVI_mean+ATSAVI_min+ATSAVI_max+ATSAVI_std, data=UAVdataset_NIR,
na.action=na.omit)

importance(RF.model_NIR)
save(RF.model_NIR, file = "RF_NIR.rda")
plot(RF.model_NIR)
print(RF.model_NIR)

dataMatrix2<-cbind(ExG_mean, ExG_min, ExG_max, ExG_std, ExR_mean, ExR_min, ExR_max, ExR_std, ExGR_mean,
                   ExGR_min, ExGR_max, ExGR_std, MExG_mean, MExG_min, MExG_max, MExG_std, NDI_mean, NDI_min,
                   NDI_max, NDI_std, INDI_mean, INDI_min, INDI_max, INDI_std, CIVE_mean, CIVE_min, CIVE_max,
                   CIVE_std, COM1_mean, COM1_min, COM1_max, COM1_std, COM2_mean, COM2_min, COM2_max, COM2_std,
                   VEG_mean, VEG_min, VEG_max, VEG_std, SR_mean,SR_min,SR_max,SR_std,DVI_mean,DVI_min,DVI_max,
                   DVI_std,NDVI_mean,NDVI_min,NDVI_max,NDVI_std,RDVI_mean,RDVI_min,RDVI_max,RDVI_std,
                   GNDVI_mean,GNDVI_min,GNDVI_max,GNDVI_std,SAVI_mean,SAVI_min,SAVI_max,SAVI_std,
                   MSAVI_mean,MSAVI_min,MSAVI_max,MSAVI_std,GEMI_mean,GEMI_min,GEMI_max,GEMI_std,
                   GCI_mean,GCI_min,GCI_max,GCI_std,TVI_mean,TVI_min,TVI_max,TVI_std,
                   RECI1_mean,RECI1_min,RECI1_max,RECI1_std,RECI2_mean,RECI2_min,RECI2_max,RECI2_std,
                   PVI_mean,PVI_min,PVI_max,PVI_std,SAVI2_mean,SAVI2_min,SAVI2_max,SAVI2_std,
                   TSAVI_mean,TSAVI_min,TSAVI_max,TSAVI_std,ATSAVI_mean,ATSAVI_min,ATSAVI_max,ATSAVI_std)

dataMatrix3<-cbind(SR_mean,SR_min,SR_max,SR_std,DVI_mean,DVI_min,DVI_max,
                   DVI_std,NDVI_mean,NDVI_min,NDVI_max,NDVI_std,RDVI_mean,RDVI_min,RDVI_max,RDVI_std,
                   GNDVI_mean,GNDVI_min,GNDVI_max,GNDVI_std,SAVI_mean,SAVI_min,SAVI_max,SAVI_std,
                   MSAVI_mean,MSAVI_min,MSAVI_max,MSAVI_std,GEMI_mean,GEMI_min,GEMI_max,GEMI_std,
                   GCI_mean,GCI_min,GCI_max,GCI_std,TVI_mean,TVI_min,TVI_max,TVI_std,
                   RECI1_mean,RECI1_min,RECI1_max,RECI1_std,RECI2_mean,RECI2_min,RECI2_max,RECI2_std,
                   PVI_mean,PVI_min,PVI_max,PVI_std,SAVI2_mean,SAVI2_min,SAVI2_max,SAVI2_std,
                   TSAVI_mean,TSAVI_min,TSAVI_max,TSAVI_std,ATSAVI_mean,ATSAVI_min,ATSAVI_max,ATSAVI_std)


dataUAV_NIR<-data.frame(dataMatrix2)
pred_fCover_UAV_NIR<-predict(RF.model_NIR, dataUAV_NIR)
write.csv(pred_fCover_UAV_NIR, file = "C:/Users/nikolina/Desktop/GIMA/UAV_fCover_pred_NIR.csv")

# training model only with NIR indices
RF.model_NIRO<-randomForest(fCover~SR_mean+SR_min+SR_max+SR_std+DVI_mean+DVI_min+DVI_max+DVI_std+
                            NDVI_mean+NDVI_min+NDVI_max+NDVI_std+RDVI_mean+RDVI_min+RDVI_max+RDVI_std+
                            GNDVI_mean+GNDVI_min+GNDVI_max+GNDVI_std+SAVI_mean+SAVI_min+SAVI_max+SAVI_std+
                            MSAVI_mean+MSAVI_min+MSAVI_max+MSAVI_std+GEMI_mean+GEMI_min+GEMI_max+GEMI_std+
                            GCI_mean+GCI_min+GCI_max+GCI_std+TVI_mean+TVI_min+TVI_max+TVI_std+
                            RECI1_mean+RECI1_min+RECI1_max+RECI1_std+RECI2_mean+RECI2_min+RECI2_max+RECI2_std+
                            PVI_mean+PVI_min+PVI_max+PVI_std+SAVI2_mean+SAVI2_min+SAVI2_max+SAVI2_std+

TSAVI_mean+TSAVI_min+TSAVI_max+TSAVI_std+ATSAVI_mean+ATSAVI_min+ATSAVI_max+ATSAVI_std, data=UAVdataset_NIR,
na.action=na.omit)

importance(RF.model_NIRO)
save(RF.model_NIRO, file = "RF_NIRO.rda")
plot(RF.model_NIRO)
print(RF.model_NIRO)

dataUAV_NIRO<-data.frame(dataMatrix3)
pred_fCover_UAV_NIRO<-predict(RF.model_NIOR, dataUAV_NIRO)
```

```
write.csv(pred_fCover_UAV_NIRO, file = "C:/Users/nikolina/Desktop/GIMA/UAV_fCover_pred_NIRO.csv")


# obtain fCover values for simulated smartphone photos
dataset_simulated = read.csv("C:/Users/nikolina/Desktop/GIMA/Module 7 - Master
thesis/Smartphone_dataset_simulated.csv", header = TRUE, sep = ";", dec = ",")

ExG_mean<-dataset_simulated$ExG_mean
ExG_min<-dataset_simulated$ExG_min
ExG_max<-dataset_simulated$ExG_max
ExG_std<-dataset_simulated$ExG_std

ExR_mean<-dataset_simulated$ExR_mean
ExR_min<-dataset_simulated$ExR_min
ExR_max<-dataset_simulated$ExR_max
ExR_std<-dataset_simulated$ExR_std

ExGR_mean<-dataset_simulated$ExGR_mean
ExGR_min<-dataset_simulated$ExGR_min
ExGR_max<-dataset_simulated$ExGR_max
ExGR_std<-dataset_simulated$ExGR_std

MExG_mean<-dataset_simulated$MExG_mean
MExG_min<-dataset_simulated$MExG_min
MExG_max<-dataset_simulated$MExG_max
MExG_std<-dataset_simulated$MExG_std

NDI_mean<-dataset_simulated$NDI_mean
NDI_min<-dataset_simulated$NDI_min
NDI_max<-dataset_simulated$NDI_max
NDI_std<-dataset_simulated$NDI_std

INDI_mean<-dataset_simulated$INDI_mean
INDI_min<-dataset_simulated$INDI_min
INDI_max<-dataset_simulated$INDI_max
INDI_std<-dataset_simulated$INDI_std

CIVE_mean<-dataset_simulated$CIVE_mean
CIVE_min<-dataset_simulated$CIVE_min
CIVE_max<-dataset_simulated$CIVE_max
CIVE_std<-dataset_simulated$CIVE_std

COM1_mean<-dataset_simulated$COM1_mean
COM1_min<-dataset_simulated$COM1_min
COM1_max<-dataset_simulated$COM1_max
COM1_std<-dataset_simulated$COM1_std

COM2_mean<-dataset_simulated$COM2_mean
COM2_min<-dataset_simulated$COM2_min
COM2_max<-dataset_simulated$COM2_max
COM2_std<-dataset_simulated$COM2_std

VEG_mean<-dataset_simulated$VEG_mean
VEG_min<-dataset_simulated$VEG_min
VEG_max<-dataset_simulated$VEG_max
VEG_std<-dataset_simulated$VEG_std

dataMatrix3<-cbind(ExG_mean, ExG_min, ExG_max, ExG_std, ExR_mean, ExR_min, ExR_max, ExR_std, ExGR_mean,
                   ExGR_min, ExGR_max, ExGR_std, MExG_mean, MExG_min, MExG_max, MExG_std, NDI_mean, NDI_min,
                   NDI_max, NDI_std, INDI_mean, INDI_min, INDI_max, INDI_std, CIVE_mean, CIVE_min, CIVE_max,
                   CIVE_std, COM1_mean, COM1_min, COM1_max, COM1_std, COM2_mean, COM2_min, COM2_max, COM2_std,
                   VEG_mean, VEG_min, VEG_max, VEG_std)

dataSmartphone_simulated<-data.frame(dataMatrix3)
pred_fCover_sim<-predict(RF.model, dataSmartphone_simulated)
write.csv(pred_fCover_sim, file = "C:/Users/nikolina/Desktop/GIMA/SM_sim_fCover_pred.csv")
```

Appendix E

```r
# import data
dataset_w = read.csv("C:/Users/nikolina/Desktop/GIMA/Module 7 - Master thesis/Wheat_simulated.csv", header =
TRUE, sep = ";", dec = ",")
dataset_m = read.csv("C:/Users/nikolina/Desktop/GIMA/Module 7 - Master thesis/Maize_simulated.csv", header =
TRUE, sep = ";", dec = ",")

# assign names to variables
fCover_wheat <- dataset_w$fCover
fCover_maize <- dataset_m$fCover

# estimate fCover for wheat using different combinations of samples
fCover_fieldW <- replicate(10, sample(fCover_wheat, 1, replace = FALSE))
mean1 <- mean(fCover_fieldW[1])
mean2 <- mean(fCover_fieldW[2])
mean3 <- mean(fCover_fieldW[3])
mean4 <- mean(fCover_fieldW[4])
mean5 <- mean(fCover_fieldW[5])
mean6 <- mean(fCover_fieldW[6])
mean7 <- mean(fCover_fieldW[7])
mean8 <- mean(fCover_fieldW[8])
mean9 <- mean(fCover_fieldW[9])
mean10 <- mean(fCover_fieldW[10])
result1 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldW <- replicate(10, sample(fCover_wheat, 4, replace = FALSE))
mean1 <- mean(fCover_fieldW[,1])
mean2 <- mean(fCover_fieldW[,2])
mean3 <- mean(fCover_fieldW[,3])
mean4 <- mean(fCover_fieldW[,4])
mean5 <- mean(fCover_fieldW[,5])
mean6 <- mean(fCover_fieldW[,6])
mean7 <- mean(fCover_fieldW[,7])
mean8 <- mean(fCover_fieldW[,8])
mean9 <- mean(fCover_fieldW[,9])
mean10 <- mean(fCover_fieldW[,10])
result2 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)


fCover_fieldW <- replicate(10, sample(fCover_wheat, 8, replace = FALSE))
mean1 <- mean(fCover_fieldW[,1])
mean2 <- mean(fCover_fieldW[,2])
mean3 <- mean(fCover_fieldW[,3])
mean4 <- mean(fCover_fieldW[,4])
mean5 <- mean(fCover_fieldW[,5])
mean6 <- mean(fCover_fieldW[,6])
mean7 <- mean(fCover_fieldW[,7])
mean8 <- mean(fCover_fieldW[,8])
mean9 <- mean(fCover_fieldW[,9])
mean10 <- mean(fCover_fieldW[,10])
result3 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldW <- replicate(10, sample(fCover_wheat, 15, replace = FALSE))
mean1 <- mean(fCover_fieldW[,1])
mean2 <- mean(fCover_fieldW[,2])
mean3 <- mean(fCover_fieldW[,3])
mean4 <- mean(fCover_fieldW[,4])
mean5 <- mean(fCover_fieldW[,5])
mean6 <- mean(fCover_fieldW[,6])
mean7 <- mean(fCover_fieldW[,7])
mean8 <- mean(fCover_fieldW[,8])
mean9 <- mean(fCover_fieldW[,9])
mean10 <- mean(fCover_fieldW[,10])
result4 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldW <- replicate(10, sample(fCover_wheat, 23, replace = FALSE))
mean1 <- mean(fCover_fieldW[,1])
mean2 <- mean(fCover_fieldW[,2])
```

```
mean3 <- mean(fCover_fieldW[,3])
mean4 <- mean(fCover_fieldW[,4])
mean5 <- mean(fCover_fieldW[,5])
mean6 <- mean(fCover_fieldW[,6])
mean7 <- mean(fCover_fieldW[,7])
mean8 <- mean(fCover_fieldW[,8])
mean9 <- mean(fCover_fieldW[,9])
mean10 <- mean(fCover_fieldW[,10])
result5 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldW <- replicate(10, sample(fCover_wheat, 30, replace = FALSE))
mean1 <- mean(fCover_fieldW[,1])
mean2 <- mean(fCover_fieldW[,2])
mean3 <- mean(fCover_fieldW[,3])
mean4 <- mean(fCover_fieldW[,4])
mean5 <- mean(fCover_fieldW[,5])
mean6 <- mean(fCover_fieldW[,6])
mean7 <- mean(fCover_fieldW[,7])
mean8 <- mean(fCover_fieldW[,8])
mean9 <- mean(fCover_fieldW[,9])
mean10 <- mean(fCover_fieldW[,10])
result6 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldW <- replicate(10, sample(fCover_wheat, 38, replace = FALSE))
mean1 <- mean(fCover_fieldW[,1])
mean2 <- mean(fCover_fieldW[,2])
mean3 <- mean(fCover_fieldW[,3])
mean4 <- mean(fCover_fieldW[,4])
mean5 <- mean(fCover_fieldW[,5])
mean6 <- mean(fCover_fieldW[,6])
mean7 <- mean(fCover_fieldW[,7])
mean8 <- mean(fCover_fieldW[,8])
mean9 <- mean(fCover_fieldW[,9])
mean10 <- mean(fCover_fieldW[,10])
result7 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldW <- replicate(10, sample(fCover_wheat, 45, replace = FALSE))
mean1 <- mean(fCover_fieldW[,1])
mean2 <- mean(fCover_fieldW[,2])
mean3 <- mean(fCover_fieldW[,3])
mean4 <- mean(fCover_fieldW[,4])
mean5 <- mean(fCover_fieldW[,5])
mean6 <- mean(fCover_fieldW[,6])
mean7 <- mean(fCover_fieldW[,7])
mean8 <- mean(fCover_fieldW[,8])
mean9 <- mean(fCover_fieldW[,9])
mean10 <- mean(fCover_fieldW[,10])
result8 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldW <- replicate(10, sample(fCover_wheat, 53, replace = FALSE))
mean1 <- mean(fCover_fieldW[,1])
mean2 <- mean(fCover_fieldW[,2])
mean3 <- mean(fCover_fieldW[,3])
mean4 <- mean(fCover_fieldW[,4])
mean5 <- mean(fCover_fieldW[,5])
mean6 <- mean(fCover_fieldW[,6])
mean7 <- mean(fCover_fieldW[,7])
mean8 <- mean(fCover_fieldW[,8])
mean9 <- mean(fCover_fieldW[,9])
mean10 <- mean(fCover_fieldW[,10])
result9 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldW <- replicate(10, sample(fCover_wheat, 60, replace = FALSE))
mean1 <- mean(fCover_fieldW[,1])
mean2 <- mean(fCover_fieldW[,2])
mean3 <- mean(fCover_fieldW[,3])
mean4 <- mean(fCover_fieldW[,4])
mean5 <- mean(fCover_fieldW[,5])
mean6 <- mean(fCover_fieldW[,6])
mean7 <- mean(fCover_fieldW[,7])
mean8 <- mean(fCover_fieldW[,8])
```

```
mean9 <- mean(fCover_fieldW[,9])
mean10 <- mean(fCover_fieldW[,10])
result10 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldW <- replicate(10, sample(fCover_wheat, 68, replace = FALSE))
mean1 <- mean(fCover_fieldW[,1])
mean2 <- mean(fCover_fieldW[,2])
mean3 <- mean(fCover_fieldW[,3])
mean4 <- mean(fCover_fieldW[,4])
mean5 <- mean(fCover_fieldW[,5])
mean6 <- mean(fCover_fieldW[,6])
mean7 <- mean(fCover_fieldW[,7])
mean8 <- mean(fCover_fieldW[,8])
mean9 <- mean(fCover_fieldW[,9])
mean10 <- mean(fCover_fieldW[,10])
result11 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldW <- replicate(10, sample(fCover_wheat, 75, replace = FALSE))
mean1 <- mean(fCover_fieldW[,1])
mean2 <- mean(fCover_fieldW[,2])
mean3 <- mean(fCover_fieldW[,3])
mean4 <- mean(fCover_fieldW[,4])
mean5 <- mean(fCover_fieldW[,5])
mean6 <- mean(fCover_fieldW[,6])
mean7 <- mean(fCover_fieldW[,7])
mean8 <- mean(fCover_fieldW[,8])
mean9 <- mean(fCover_fieldW[,9])
mean10 <- mean(fCover_fieldW[,10])
result12 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

result <- cbind(result1, result2, result3, result4, result5, result6, result7, result8, result9, result10,
result11, result12)
write.csv(result, file = "C:/Users/nikolina/Desktop/GIMA/Wheat_simulated_sampled.csv")

# estimate fCover for maize using different combinations of samples

fCover_fieldM <- replicate(10, sample(fCover_maize, 1, replace = FALSE))
mean1 <- mean(fCover_fieldM[1])
mean2 <- mean(fCover_fieldM[2])
mean3 <- mean(fCover_fieldM[3])
mean4 <- mean(fCover_fieldM[4])
mean5 <- mean(fCover_fieldM[5])
mean6 <- mean(fCover_fieldM[6])
mean7 <- mean(fCover_fieldM[7])
mean8 <- mean(fCover_fieldM[8])
mean9 <- mean(fCover_fieldM[9])
mean10 <- mean(fCover_fieldM[10])
result1 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldM <- replicate(10, sample(fCover_maize, 4, replace = FALSE))
mean1 <- mean(fCover_fieldM[,1])
mean2 <- mean(fCover_fieldM[,2])
mean3 <- mean(fCover_fieldM[,3])
mean4 <- mean(fCover_fieldM[,4])
mean5 <- mean(fCover_fieldM[,5])
mean6 <- mean(fCover_fieldM[,6])
mean7 <- mean(fCover_fieldM[,7])
mean8 <- mean(fCover_fieldM[,8])
mean9 <- mean(fCover_fieldM[,9])
mean10 <- mean(fCover_fieldM[,10])
result2 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldM <- replicate(10, sample(fCover_maize, 9, replace = FALSE))
mean1 <- mean(fCover_fieldM[,1])
mean2 <- mean(fCover_fieldM[,2])
mean3 <- mean(fCover_fieldM[,3])
mean4 <- mean(fCover_fieldM[,4])
mean5 <- mean(fCover_fieldM[,5])
mean6 <- mean(fCover_fieldM[,6])
mean7 <- mean(fCover_fieldM[,7])
mean8 <- mean(fCover_fieldM[,8])
```

```
mean9 <- mean(fCover_fieldM[,9])
mean10 <- mean(fCover_fieldM[,10])
result3 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldM <- replicate(10, sample(fCover_maize, 17, replace = FALSE))
mean1 <- mean(fCover_fieldM[,1])
mean2 <- mean(fCover_fieldM[,2])
mean3 <- mean(fCover_fieldM[,3])
mean4 <- mean(fCover_fieldM[,4])
mean5 <- mean(fCover_fieldM[,5])
mean6 <- mean(fCover_fieldM[,6])
mean7 <- mean(fCover_fieldM[,7])
mean8 <- mean(fCover_fieldM[,8])
mean9 <- mean(fCover_fieldM[,9])
mean10 <- mean(fCover_fieldM[,10])
result4 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldM <- replicate(10, sample(fCover_maize, 26, replace = FALSE))
mean1 <- mean(fCover_fieldM[,1])
mean2 <- mean(fCover_fieldM[,2])
mean3 <- mean(fCover_fieldM[,3])
mean4 <- mean(fCover_fieldM[,4])
mean5 <- mean(fCover_fieldM[,5])
mean6 <- mean(fCover_fieldM[,6])
mean7 <- mean(fCover_fieldM[,7])
mean8 <- mean(fCover_fieldM[,8])
mean9 <- mean(fCover_fieldM[,9])
mean10 <- mean(fCover_fieldM[,10])
result5 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldM <- replicate(10, sample(fCover_maize, 35, replace = FALSE))
mean1 <- mean(fCover_fieldM[,1])
mean2 <- mean(fCover_fieldM[,2])
mean3 <- mean(fCover_fieldM[,3])
mean4 <- mean(fCover_fieldM[,4])
mean5 <- mean(fCover_fieldM[,5])
mean6 <- mean(fCover_fieldM[,6])
mean7 <- mean(fCover_fieldM[,7])
mean8 <- mean(fCover_fieldM[,8])
mean9 <- mean(fCover_fieldM[,9])
mean10 <- mean(fCover_fieldM[,10])
result6 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldM <- replicate(10, sample(fCover_maize, 44, replace = FALSE))
mean1 <- mean(fCover_fieldM[,1])
mean2 <- mean(fCover_fieldM[,2])
mean3 <- mean(fCover_fieldM[,3])
mean4 <- mean(fCover_fieldM[,4])
mean5 <- mean(fCover_fieldM[,5])
mean6 <- mean(fCover_fieldM[,6])
mean7 <- mean(fCover_fieldM[,7])
mean8 <- mean(fCover_fieldM[,8])
mean9 <- mean(fCover_fieldM[,9])
mean10 <- mean(fCover_fieldM[,10])
result7 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldM <- replicate(10, sample(fCover_maize, 52, replace = FALSE))
mean1 <- mean(fCover_fieldM[,1])
mean2 <- mean(fCover_fieldM[,2])
mean3 <- mean(fCover_fieldM[,3])
mean4 <- mean(fCover_fieldM[,4])
mean5 <- mean(fCover_fieldM[,5])
mean6 <- mean(fCover_fieldM[,6])
mean7 <- mean(fCover_fieldM[,7])
mean8 <- mean(fCover_fieldM[,8])
mean9 <- mean(fCover_fieldM[,9])
mean10 <- mean(fCover_fieldM[,10])
result8 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldM <- replicate(10, sample(fCover_maize, 61, replace = FALSE))
mean1 <- mean(fCover_fieldM[,1])
```

```
mean2 <- mean(fCover_fieldM[,2])
mean3 <- mean(fCover_fieldM[,3])
mean4 <- mean(fCover_fieldM[,4])
mean5 <- mean(fCover_fieldM[,5])
mean6 <- mean(fCover_fieldM[,6])
mean7 <- mean(fCover_fieldM[,7])
mean8 <- mean(fCover_fieldM[,8])
mean9 <- mean(fCover_fieldM[,9])
mean10 <- mean(fCover_fieldM[,10])
result9 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldM <- replicate(10, sample(fCover_maize, 70, replace = FALSE))
mean1 <- mean(fCover_fieldM[,1])
mean2 <- mean(fCover_fieldM[,2])
mean3 <- mean(fCover_fieldM[,3])
mean4 <- mean(fCover_fieldM[,4])
mean5 <- mean(fCover_fieldM[,5])
mean6 <- mean(fCover_fieldM[,6])
mean7 <- mean(fCover_fieldM[,7])
mean8 <- mean(fCover_fieldM[,8])
mean9 <- mean(fCover_fieldM[,9])
mean10 <- mean(fCover_fieldM[,10])
result10 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldM <- replicate(10, sample(fCover_maize, 78, replace = FALSE))
mean1 <- mean(fCover_fieldM[,1])
mean2 <- mean(fCover_fieldM[,2])
mean3 <- mean(fCover_fieldM[,3])
mean4 <- mean(fCover_fieldM[,4])
mean5 <- mean(fCover_fieldM[,5])
mean6 <- mean(fCover_fieldM[,6])
mean7 <- mean(fCover_fieldM[,7])
mean8 <- mean(fCover_fieldM[,8])
mean9 <- mean(fCover_fieldM[,9])
mean10 <- mean(fCover_fieldM[,10])
result11 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

fCover_fieldM <- replicate(10, sample(fCover_maize, 87, replace = FALSE))
mean1 <- mean(fCover_fieldM[,1])
mean2 <- mean(fCover_fieldM[,2])
mean3 <- mean(fCover_fieldM[,3])
mean4 <- mean(fCover_fieldM[,4])
mean5 <- mean(fCover_fieldM[,5])
mean6 <- mean(fCover_fieldM[,6])
mean7 <- mean(fCover_fieldM[,7])
mean8 <- mean(fCover_fieldM[,8])
mean9 <- mean(fCover_fieldM[,9])
mean10 <- mean(fCover_fieldM[,10])
result12 <- rbind(mean1, mean2, mean3, mean4, mean5, mean6, mean7, mean8, mean9, mean10)

result <- cbind(result1, result2, result3, result4, result5, result6, result7, result8, result9, result10,
result11, result12)
write.csv(result, file = "C:/Users/nikolina/Desktop/GIMA/Maize_simulated_sampled.csv")
```

Appendix F

ID	Lava	Okapia	Samsung Galaxy	Symphony	Samsung Tablet	Difference
1	0,49	0,53	0,58	0,61	0,56	0,12
2	0,13	0,14	0,2	0,19	0,15	0,07
3	0,12	0,13	0,16	0,2	0,14	0,08
4	0,42	0,41	0,43	0,5	0,45	0,09
5	0,15	0,14	0,18	0,2	0,2	0,06
6	0,32	0,22	0,31	0,55	0,1	0,45
7	0,03	0,03	0,03	0,03	0,03	0
8	0,26	0,23	0,27	0,26	0,21	0,06
9	0,19	0,2	0,21	0,17	0,19	0,04
10	0,37	0,35	0,44	0,43	0,44	0,09
11	0,21	0,21	0,21	0,22	0,2	0,02
12	0,05	0,04	0,06	0,05	0,06	0,02
13	0,55	0,64	0,67	0,7	0,69	0,15
14	0,52	0,55	0,56	0,47	0,46	0,1
15	0,28	0,31	0,24	0,17	0,28	0,14
16	0,15	0,14	0,15	0,11	0,11	0,04
17	0,29	0,34	0,39	0,33	0,29	0,1
18	0,32	0,33	0,45	0,36	0,32	0,13
19	0,02	0,03	0,03	0,03	0,03	0,01
20	0,1	0,09	0,12	0,1	0,07	0,05
21	0,03	0,04	0,03	0,03	0,02	0,02
22	0,12	0,11	0,13	0,09	0,1	0,04
23	0,46	0,44	0,5	0,38	0,36	0,14
24	0,54	0,57	0,67	0,53	0,53	0,14
25	0,22	0,2	0,23	0,2	0,22	0,03
26	0,15	0,15	0,14	0,16	0,14	0,02
27	0,12	0,12	0,11	0,13	0,13	0,02
28	0,18	0,2	0,19	0,19	0,18	0,02
29	0,1	0,1	0,1	0,1	0,1	0
30	0,55	0,57	0,57	0,58	0,51	0,07
31	0,2	0,21	0,22	0,19	0,16	0,06
32	0,39	0,37	0,37	0,3	0,34	0,09
33	0,34	0,36	0,36	0,26	0,48	0,22