University of Kassel

Faculty of Organic Agricultural Sciences

Department of Grassland Science and Renewable Plant Resources

Digital image analysis as a tool to estimate legume contributions in legume-grass swards

Thesis for obtaining the degree doctor of agricultural sciences

tendered by

Maike Himstedt

born in Hildesheim

This work has been accepted by the Faculty of Organic Agricultural Sciences of the University of Kassel as a thesis for acquiring the academic degree of Doktor der Agrarwissenschaften (Dr. agr.).

1. Supervisor: Prof. Dr. Michael Wachendorf (University of Kassel)
2. Supervisor: Prof. Dr. Andreas Bürkert (University of Kassel)
3. Referee: Prof. Dr. Eva Schlecht (University of Kassel)
4. Referee: Prof. Dr. Martin Kappas (Georg-August-University of Göttingen)

Defence day: 08. January 2010

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Dissertation selbständig und ohne unerlaubte Hilfe angefertigt und andere als die in der Dissertation angegebenen Hilfsmittel nicht benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen sind, habe ich als solche kenntlich gemacht. Kein Teil dieser Arbeit ist in einem anderen Promotions- oder Habilitationsverfahren verwendet worden.

Preface

This thesis is submitted to the Faculty of Organic Agricultural Sciences of the University of Kassel to fulfil the requirements for the degree Doktor der Agrarwissenschaften (Dr. agr.).

This dissertation is based on three papers as first author, which are published or submitted to international refereed journals. They are included in chapter 4, 5 and 6. Chapter 1 gives the introduction to all parts of the thesis. Chapter 2 contains the objectives of the work and chapter 3 gives basic information on digital image analysis. Chapter 7 considers the results of the chapters 4, 5 and 6 in a general discussion. A general conclusion and the summary are given in chapter 8 and chapter 9.

The following papers contribute to this thesis:

Chapter 4:

Himstedt, M., T. Fricke, and M. Wachendorf .2009. Determining the contribution of legumes in legume-grass mixtures using digital image analysis. Crop Science 49, 1910-1916.

Chapter 5:

Himstedt, M., T. Fricke, and M. Wachendorf .2009. The relationship between coverage and dry matter contribution of forage legumes in binary legume-grass mixtures. Crop Science (submitted).

Chapter 6:

Himstedt, M., T. Fricke, and M. Wachendorf .2009.An advanced image analysis procedure to estimate legume contents in legume-grass swards. Biosystems Engineering (submitted).

Table of contents

Tables

Chapter 4:

Chapter 5:

Chapter 6:

Figures

Abbrevations

CDS: combined dataset - dataset of all sward images

DIA: digital image analysis

DM: dry matter

LC_M: measured legume coverage (% of area)

LC_{DIA}: estimated legume coverage (% of area)

LD_M: measured legume contribution (% of DM)

LD_{DIA}: estimated legume contribution (% of DM)

LCG: dataset of sward images including pure grass swards, pure lucerne swards and lucerne-grass swards

RCG: dataset of sward images including pure grass swards, pure red clover swards and red clover-grass swards

RMSE: root mean squared error

SE: standard error

WCG: dataset of sward images including pure grass swards, pure white clover swards and white clover-grass swards

1 General introduction

Legumes with their ability to fix atmospheric nitrogen by legume-rhizobium symbiosis contribute considerably to the nutrient supply in forage production (Boller and Nösberger, 1987; Frame et al., 1985; Heichel and Henjum, 1991; Wachendorf et al., 2004). Besides their productivity the botanical composition of legume-grass swards is an important factor for successful arable farming in both organic and conventional farming systems. Many detailed studies investigating legume-grass mixtures have shown that high amounts of nitrogen (>300 kg N ha^{-1}) can be fixed by legumes (Loges, 1998; Schmidtke, 1997; Weißbach, 1995, Carlsson and Huss-Danell, 2003). Hence, forage legumes in mixture with grass are virtually self-sufficient for nitrogen and can concurrently transfer appreciable nitrogen to the companion grass (Heichel and Henjum, 1991).
Growth of legumes can vary strongly through spatial and temporal influences, e.g. cutting, frost, and drought damage. Field scale variations in legume contents are only partly explained by the level of seasonal disturbance itself (Schwinning and Parsons, 1996). Hence, continuous mapping of legume distribution in the field could help to understand the processes which affect the abundance of legumes in swards at field scale and to adapt the grassland management to these processes. Since yield and proportion of legumes are strongly related to the amount of fixed nitrogen (Høgh-Jensen et al., 2004), frequent information on the status of legume-grass swards could help to adjust fertilizer application and to predict the nitrogen supply of the soil for arable crops to be grown after the legume-grass mixtures. Besides, forage management could be optimized as forage quality is highly affected by the yield and legume contribution in the swards (Frame, 1992; Sleugh et al., 2000).

To predict spatial distribution of legumes on grassland, systematic manual plant sampling has been used (Gottardi, 2008). Although widely used in experiments, these procedures are not applicable to real farm situations because they are too labour intensive. A methodology is desirable which allows a quick and non-laborious assessment of the legume biomass in mixed swards. The indirect assessment of the sward biomass by field spectroscopy (Biewer et al. 2008, 2009; Numata et al., 2007; Schino et al., 2003) has produced reliable data but could not determine the legume DM contribution.

For this purpose, digital image analysis (DIA) of near-ground imagery is a prospective tool as it has been successfully applied to identify and estimate biomass and locate individual plants. In agriculture the analysis of digital images has been used to estimate soybean [*Glycine max* (L.) Merr.] canopy cover (Purcell, 2000) and turf grass cover (Richardson et al., 2001) using hue (H) and saturation (S) values of the image pixels to distinguish between crop and soil. Paruelo et al. (2000) used red-green-blue (RGB) colour information of images to estimate plant biomass, where the percentage of green pixels of the image and green grass biomass showed a correlation of 0.87. RGB images, transformed to intensity maps (greyscale values) by means of an equation including a particular constant, were used for textural analysis of soil images (Roy et al., 2006) and to discriminate between soil and crops (Hague et al., 2006). The resulting images exhibits good contrast between plant material and soil and are insensitive to the amplitude of the illuminant (Marchant and Onyango, 2002). Beside the transformed RGB values Onyango et al. (2005) successfully used morphological filtering to distinguish between crops and weed species. Petry and Kühbauch (1989) used image analysis techniques to identify six typical winter weed species using form parameters including area, perimeter, minimum and maximum diameter of the plants. Sökefeld et al. (2007) utilised a bispectral camera to distinguish between plants and soil background, in addition they determined weed species and crop shape parameters, as well as contour and skeleton features to calculate a classification algorithm (Weisa and Gerhards, 2007; Gerhards et al., 2002; Nordmeyer, 2006).

The use of DIA to distinguish between plant species of a heterogeneous grassland canopy may be more difficult than identifying plants against a uniform soil background. A canopy of diverse grassland plants creates difficulties for the DIA like overlapping of leaves and tillers, varied leaf colours and shapes, shadows on leaves and soil, and additionally different leaf appearances during the growing season. Up to today image analysis techniques have not been applied frequently to separate individual species in grassland or legume-grass swards. Gebhardt et al. (2006) and Gebhardt and Kühbauch (2007) have detected *R. obtusifolius* L. in mixed grassland swards using homogeneity thresholds in addition to shape, colour and texture features. Bonesmo et al. (2004) developed a semi-automatic image processing system to estimate the coverage of white clover in a legume-grass mixture, based on clover colour and morphological properties as a tool to analyse spatial dynamics especially in experimental grass-clover swards.

The objective of this study was to evaluate DIA as a tool to determine the legume dry matter contribution in legume-grass mixtures. To address this objective, a pot experiment was conducted in the greenhouse with several legume-grass mixtures of different sward ages using constant recording geometry and illumination.

2 Research objectives

The objective of this study was to evaluate if digital image analysis can be used to estimate the legume dry matter (DM) contribution of legume-grass swards across a wide range of legume species, legume proportions and growth stages.

A pot experiment was conducted in a greenhouse under controlled conditions to allow the potential of digital image analysis to be assessed for estimating the legume DM contribution of legume-grass swards using constant recording geometry and illumination. Images of a field experiment were used to consolidate objective i).

The specific objectives of this investigation were:

i) to evaluate if a relationship between legume coverage and legume DM contribution in legume-grass swards exists, as legume coverage is the achievable information of an image.

ii) to develop a digital image analysis procedure to estimate the legume coverage in images of legume-grass swards.

iii) to evaluate if with the findings of i) and ii) an estimation of legume DM contribution with images of legume-grass swards is feasible.

3 Basic information on digital image analysis

„The purpose of image analysis is to give the raw data of images a symbolic meaning that fits to a certain model of the real world and can be used for decisions or further processing. This means, the aim is to make a machine understand what is in an image." (Brox, 2005, p.1)

The field of digital image processing is the study of algorithms for the transformation of pixel values with the goal to get useful information out of digital images. Finally, image analysis is the effort to imitate the performance of the human visual system by means of a machine. The human's visual capabilities are extraordinary. Vision is an important part of the human creature; very large parts of the human brain are reserved only for the processing of the information provided by the eyes. This all makes clear that to imitate these capabilities using a machine is a challenging task; it also explains why progress made in this field often looks trivial at the first glance, as the same task is accomplished so easily by a human. There are many pixels in the image supplying their position and their grey value or colour, yet they give no information on the objects in the scene. There is a priori no relation between pixels and objects. The task of image segmentation is exactly to provide such a relationship (Brox, 2005).

An image is a visual representation of an object or group of objects. Digital images have a finite set of digital values, called picture elements or pixels. These pixels are the smallest individual elements in the image consisting of one or more quantities, e.g. colour or greyscale values, related to that specific position in the image. Digital images contain a fixed number of rows and columns of pixels and can be created by a variety of input devices and techniques, such as digital cameras or scanners.

3.1 Pixel depth and colour format

Pixel values for images, which contain only black and white, can be easily represented by a single bit (two colour possibilities): 0=black, 1=white. However, a colour image contains much more information and may be represented by 24 bits per pixel (2^{24} colour possibilities). Colour depth or bit depth, is a term which applies to computer graphics describing the number of bits used to represent the colour of a single pixel. Higher bit depth provides a broader range of distinct colours, e.g. 1-bit colour (2^1 = 2 colours), 2-bit colour (2^2 = 4 colours), 4-bit colour (2^4 = 16 colours), and so on. The bit depth tells how many unique colours an image can display, the image format gives information about what colours are actually contained within the image (Haberäcker, 1987). The colour format of the sward images used in this study was the 24-bit RGB (red, green, blue) image format. For the different purposes of image analysis this format was transformed into 8-bit greyscale images and 24-bit HSL (hue, saturation, lightness) images, resulting from 8-bit for H, S and L respectively.

3.2 Greyscale images

Greyscale pixel values represent the level of greyness or brightness (Y), ranging from completely black to completely white. In an 8-bit greyscale image, a pixel with a value of zero is black, a pixel with 255 white and a pixel with 127 a grey colour halfway between black and white. 8-bit greyscale is the most common greyscale format in use. For the transformation from RGB colour space for each pixel one value can be evaluated using Y= (R+G+B)/3.

3.3 Colour Models

A colour model is a standard way to represent colour in mathematical terms. There are many colour models in use; the RGB (red, green, blue), HSI (hue, saturation, intensity) and HSL (hue, saturation, lightness) models are most frequently used in digital image processing. The RGB colour model is used by most digital imaging devices such as monitors and cameras. In the RGB colour model every colour is represented as a mixture of varying levels of red, green and blue. In a 24-bit RGB image each pixel value is made up of three separate 8-bit samples representing the level of brightness of its respective colour red, green or blue. The HSL colour model describes a colour in terms of how it is perceived by the human eye. In the HSL model a hue (H) is specified by its position on a hexagon (Figure 3.1) as measured by its distance in degrees from the red axis (for example, a hue value of 120 would indicate green, which is 120^o from red). The HSL model is useful for comparing two colours, or for changing one colour to another, as only the H value needs to be changed. Changes of the saturation (S) value from highly saturated to reduced saturated appears as some white is added to the colour. The decrease of lightness (L) appears as some black was added. Pixel's lightness (L) is determined according to the following formula: L = [(max(R,G,B) + min(R,G,B) + 1] / 2 (Media Cybernetics, 1999).

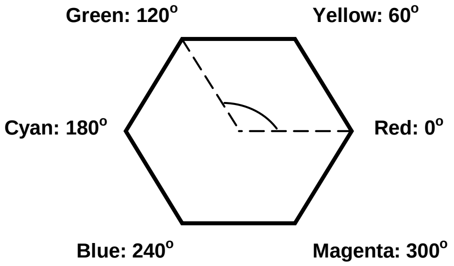


Figure 3. 1: Hue Hexagon; the hue value of 120 indicates green, as it is 120° from red (source: Media Cybernetics, 1999)

3.4 Threshold

Thresholds are used to assign pixels to either foreground or background. For the threshold settings histograms of the single HSL values or grey values were used respectively. The x axis of e.g. a grey value histogram is the brightness scale. The histogram spans from black (minimum value of the brightness range) to white (maximum value). The grey value of white is dependent upon the bit depth of the image. For example, an 8-bit image has a maximum grey value of 255. The y axis of the histogram is the number of the pixels found for each grey value in the image. Figure 3.2 depicts the threshold settings for bare soil in an HSL colour image of a red clover sward using H, S and L.

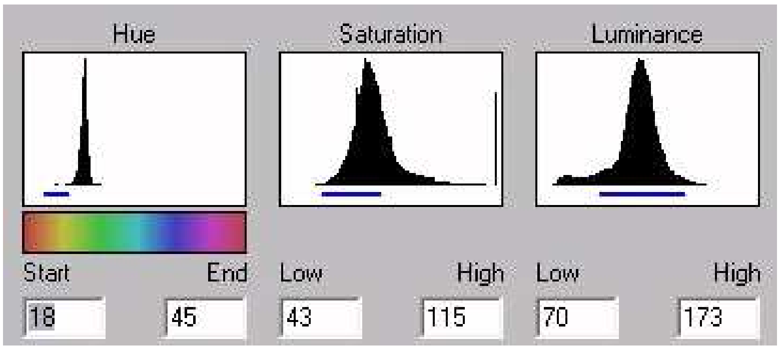


Figure 3. 2: Threshold settings for bare soil in an HSL colour image of a red clover sward using H, S and L. The blue lines below the histograms display the threshold from the start to the end values. Pixel values within this range becoming foreground pixels (source: Optimas 6.5, Media Cybernetics).

3.5 Morphological operations

Morphological operators are local operators or so called pixel-based filters. They are important tools for image processing as they enable a customized filtering of the image with the goal to highlight distinct regions. Below the operating mode for the application on greyscale images is described, this is applicable in the same manner to H, S, and L values of colour images.

Morphological operators change greyscale values of a pixel in relation to the grey values of the pixels in the defined neighbourhood. The neighbourhood of a pixel consists of the directly circumjacent pixel and is later called structuring element. The structuring element moves across the image pixel by pixel, applying defined operations to the central pixel.

The following morphological operators were used in this study:

i) Erode: This operation replaces each grey value of central pixel with the value of the darkest pixel of the neighbourhood as defined by the structuring element.

ii) Dilate: This operation replaces each grey value of the central pixel with the value of the brightest pixel of the structuring element.

iii) Opening: This operation is erosion followed by dilation.

The size of the structuring element is important for the effect and is defined by the user, e.g. sizes with 3*3, 5*5, or 7*7 pixels. The operation can be applied with different iterations. Also different shapes of the structuring element can be chosen like squares or diamonds. In the present study quadratic structuring elements were used.

3.6 Form parameters

Digital image processing often includes the analysis of form parameters of the foreground objects defined by e.g. threshold segmentation. In the present study the following form parameters were tested:

i) Area: A value which can be extracted from area screen objects giving the bounded area in calibrated units squared.

ii) Breadth: A value which can be extracted from area screen objects giving the sum of the maximum distance of the boundary from either side of the major axis in calibrated units. For symmetrical area boundaries this will be equivalent to the "minor axis" length.

iii) Rectangularity: A value which can be extracted from area screen objects giving the object's area divided by the area of an enclosing bounding box aligned with the major axis. This is a dimensionless number with a minimum value approaching 0 achieved only for narrow cross-like boundaries. The value is pi over 4 (0.79) for circular boundaries, 0.5 for square boundaries and approaches 1.0 for long and narrow rectangular boundaries.

iv) Circularity: A value which can be extracted from area screen objects giving the ratio of the area perimeter length squared divided by the object's area. This is a dimensionless number with a minimum value of four pi (12.57) achieved only for circular boundaries (the value is 16 for square boundaries and 20.78 for equilateral triangular boundaries) (Media Cybernetics, 1999).

3.7 Used Hardware and Software

The digital 24-bit RGB images of the swards were taken with a Canon Power Shot G6 Digital Camera, using flashlight for all pictures to ensure uniform illumination. The distance between camera and sward surface was constant 80 cm at the nadir position of the plot. A superfine compression was chosen and the pictures were saved in JPEG (Joint Photographic Experts Group) format at a resolution of 2592x1944 pixels. Using four concrete vertices, digital pictures were georeferenced with the program SAGA (System for Automated Geoscientific Analyses Version 2.0, 2005) to compensate distortions. The sward images were clipped to obtain the swards without rims of vessel. The resulting image size was 766x744 pixels, related to resolution and plot size. For image processing and analysis the software Optimas from Media Cybernetics (Silver Spring, MD) was applied.

4 Determining the contribution of legumes in legume-grass mixtures using digital image analysis

Abstract Digital image analysis could be a rapid and precise technique for estimating legume proportions in grass swards. In 2004, we conducted a pot study to evaluate a digital image analysis (DIA) system for estimation of legume dry matter (DM) contribution in legume-grass mixtures. Examination of pure swards and binary legume-grass mixtures of red clover (*Trifolium pratense* L.), white clover (*Trifolium repens* L.), lucerne (*Medicago sativa* L.), and perennial ryegrass (*Lolium perenne* L.) took place after 35, 49, and 63 days of growth. To estimate the cover percentage of legumes in the swards, a total of 64 digital pictures were taken. The DM contribution of legumes (% of total biomass) showed a significant relationship with the measured proportion of image area covered by legumes (% of total area) which was classified visually. A DIA system for greyscale images was developed with the software Optimas. We found that DIA could be used to accurately predict legume contribution in mature swards. Legume contribution, as estimated by DIA, was significantly correlated with DM contribution of red clover (R^2 0.87), white clover (R^2 0.85) and lucerne (R^2 0.79). Bare ground reduced the predictive ability of DIA in young or open swards. Use of DIA may be limited until we refine the method to deal with bare ground and different leaf shapes associated with various weed species.

4.1 Introduction

The botanical composition of legume-grass swards varies spatially and affects important processes and attributes like nitrogen (N_2) fixation and herbage yield (Loges, 1998; Loges et al., 2000; Neuendorff, 1996). The proportion of legume in mixtures is of particular importance due to their ability to fix atmospheric N (Ledgard and Steele, 1992; Wachendorf et al., 2004), which can reduce the need for industrial N fertilizers (Carlsson and Huss-Danell, 2003; Malhi et al., 2002). Carlsson and Huss-Danell (2003) reported rates of N_2 fixation in above ground plant tissues for red clover (373 kg N ha^{-1}), white clover (545 kg N ha^{-1}), and lucerne (350 kg N ha^{-1}) and found a strong correlation between N_2 fixation per hectare and year and the legume dry matter (DM) yield. Høgh-Jensen et al. (2004) also described an empirical model for quantifying symbiotic nitrogen fixation in legume-grass mixtures on the basis of the legume DM yield.

The productivity of legume-grass swards is influenced by many factors, such as soil N status, competition among legumes and associated grasses (Ledgard and Steele, 1992) and weather (Vinther, 2006). Once the site-specific yield of legumes is known, sward management practices like fertilization, reseeding, or the prediction of preceding crop effects, could be adapted more precisely. Recent results show that the DM yield of legume-grass mixtures can be predicted by field spectroscopy with a reasonable accuracy (Biewer et al., 2008, 2009), whereas the detection of legume DM contribution in the swards requires other sensors.

Digital image analysis (DIA) and machine vision technologies have been successfully applied to agriculture to identify and estimate phytomass and locate individual plants. For example, analysis of digital and photographic images has been used to estimate soybean [*Glycine max* (L.) Merr.] canopy cover (Purcell, 2000), turf grass cover (Richardson et al., 2001) and biomass in semiarid regions (Paruelo et al., 2000). Digital image analysis can be used to distinguish between crops and weed species (Hague et al., 2006; Onyango et al., 2005; Petry and Kühbauch, 1989). Sökefeld et al. (2007) used a bispectral camera to distinguish between plants and soil background, identifying weed species and crop shape parameters, contour and skeleton features to calculate a classification algorithm (Weisa and Gerhards, 2007; Gerhards et al., 2002; Nordmeyer, 2006).

The use of DIA of a heterogeneous grassland canopy may be more difficult than identifying plants against a uniform soil background as in arable crops. The image analysis technique has not been applied to separate individual species, but rather to identify prominent species like *Rumex obtusifolius* L. or categories such as legumes. A canopy of diverse grassland plants presents several difficulties to DIA, including the diversity of optical plant properties within a mixed sward, varied leaf colours and shapes, overlapping of leaves and tillers, shadows on leaves and soil, nonuniform soil background and different leaf appearances during the growing season. Others have detected *R. obtusifolius* L. in mixed grassland swards by recording images with a remote-controlled vehicle under constant geometric conditions in the field, segmenting the images using homogeneity threshold and defining objects and features describing shape, colour and texture (Gebhardt et al., 2006; Gebhardt and Kühbauch, 2007). Bonesmo et al. (2004) developed an image processing system to estimate the canopy cover of white clover in a legume-grass mixture, based on clover colour and morphological properties. The efficiency of the system was field tested and a close relationship was found between the canopy cover of clover estimated by image processing and the canopy cover estimated by manually marking the clover. Fransen et al. (1998) used DIA to quantify the horizontal vegetation pattern in savannah grasslands and applied the nearest neighbour and non-overlapping domain analysis.

Our objective was to evaluate DIA as a tool to determine the legume dry matter contribution in legume-grass mixtures. To address this objective, we conducted a pot experiment in the greenhouse with several legume-grass mixtures of different sward age using constant recording geometry and illumination.

4.2 Material and Methods

4.2.1 Sward Construction and Experimental Design

In winter 2004, we conducted a 9-wk pot experiment with pure swards and binary legume-grass mixtures of red clover (‘Tamara’), white clover (‘Klondyke’), lucerne (‘Daisy’), and perennial ryegrass (‘Lilora’). Eight experimental swards were investigated: monoculture swards of perennial ryegrass (G; sown at 20 kg ha^{-1}), red clover (R; sown at 8 kg ha^{-1}), white clover (W; sown at 4 kg ha^{-1}), and lucerne (L; sown at 16 kg ha^{-1}) and four binary mixtures with red clover-ryegrass (R8G; sown at 8 and 20 kg ha^{-1}; and R2G, sown at 2 and 20 kg ha^{-1}), white clover-ryegrass (WG; sown at 4 and 20 kg ha^{-1}) and lucerne-ryegrass (LG; sown at 16 and 20 kg ha^{-1}). The R2G and R8G treatments were added to achieve a greater cover range for this important legume species. All treatments were sown manually in four replicates in a distance of 12 cm between rows and at a sowing depth of

0.5 cm. Sward size was 0.119 m^2, length and width were 34 by 35 cm, including three rows of seedlings. Wooden pots were filled with 2 cm drainage substratum (Lavagrus) and about 16 cm homogenized loamy soil (sandy loam; 3.6 % sand, 73 % silt, 23.4 % clay and 2 % humus). Soil analysis indicated sufficient contents of phosphorus, magnesium and potassium and a pH-value of 6.7. No fertilizer was applied. The pots were randomly arranged in the greenhouse and illuminated with constant artificial daylight for 12 h d^{-1}. The average temperature in the greenhouse was 14.7°C, the minimum at night being 7.9°C and the maximum at daytime 24.0°C. To compare different aged swards, legume-grass mixtures were sown in calendar weeks 46, 48, 50, and 52 and harvest was performed at a common date, which was 21, 35, 49, or 63 d after sowing. Monoculture swards were harvested 63 d after sowing.

To determine the sward composition, total aboveground biomass was sorted to grass, legumes and unsown species. The samples were dried 48 h at 65°C. Nomenclature for the unsown species identified in the swards followed Rothmaler et al. (1996). Analysis of variance was conducted using the GLM procedure of SAS 9.1 (SAS Institute, Cary, NC). A fixed-effects model with a simple error structure was used, as factor levels of sward type and sward age were completely randomized in the experiment. For regression analysis the REG procedure in SAS was used.

4.2.2 Digital Image Analysis

One day before harvest, digital pictures of the swards were taken with a Canon (Tokyo, Japan) Power Shot G6 Digital Camera, using flashlight for all pictures to ensure uniform illumination. A superfine compression was chosen and the pictures were saved in JPEG (Joint Photographic Experts Group) format at a resolution of 2592x1944 pixels. Using four concrete vertices, digital pictures were georeferenced with the program SAGA (System for Automated Geoscientific Analyses Version 2.0, 2005) to compensate distortions. The sward images were clipped to obtain the swards without rims of vessel. The resulting image size was 766x744 pixels, related to resolution and plot size. For DIA the image processing software Optimas from Media Cybernetics (Silver Spring, MD) was applied.

Digital RGB colour images were imported in TIF (tagged image file) format and converted into 8-bit greyscale images, because greyscale pictures are less complex and speed processing.

Determination of DM contribution of legumes by DIA follows three steps (Figure 4.1):

Step 1. Quantifying the relation between coverage and contribution of legumes.

The determination of legume DM contribution requires the measurement of the area of legume cover. Therefore, the actual visible legume coverage in each image was manually encircled to measure the legume coverage (LC_M, % area). The measured DM contribution of legumes (LD_M) was regressed with the visible coverage (LC_M) to establish a relationship between both parameters.

Step 2. Processing of a digital image analysis procedure to estimate the legume coverage (% area).

Due to the pronounced differences in the shape of legume and grass leaves, the DIA procedure applies erosion and dilation filters to the greyscale images of the swards. This pixel-based filter combination removes small objects, like small grass leaves. A quadratic structuring element was defined, for example 9 x 9 pixels, and applied to the image. This element moved across the image pixel by pixel to first apply the erosion filter to the whole image and subsequently the dilation filter. The erosion replaced the grey value of the central pixel of the structuring element with the darkest grey value within the structuring element. So small areas in the image, like grass leaves, continuously get darker until they are dark grey or black. Major areas, like legume leaves, get darker, with light grey values persisting in the middle area of the leave. The subsequent dilation works in reverse, replacing the central pixel of the structuring element with the brightest grey value within the structuring element.

Implementing both the dilation and the erosion at the same frequency, the still-remaining light grey areas obtained their previous size. Both filters can be executed with different structuring element sizes and in different numbers of iteration. In the present study structuring elements of 3*3, 5*5, 7*7 and 9*9 pixels were tested and iteration varied between 1 and 13. Subsequently, a greyscale threshold of [71:255], which was determined by test series, was set for image segmentation. The aim of segmentation was to separate the light grey areas (the legume leaves) from the background (dark grey and black). All 52 procedure combinations of structuring element sizes and iteration were applied to the images. The optimal DIA procedure was identified by the lowest difference between reference value (LC_M) and the predicted legume coverage. Included were all swards except the 21-d swards as they turned out to be too low in biomass with only one leaf unfolded.

Step 3. Calculating DM contribution by DIA using the equation of Step 1 and the DIA procedure developed in Step 2.

The DM contribution of legumes (% of DM) was calculated by the combination of equation of Step 1 and the cover percentages of legumes (% area) estimated by DIA. Finally, the calculated DM contribution of legumes was related to the measured DM contribution of legumes to compare overall accuracy.

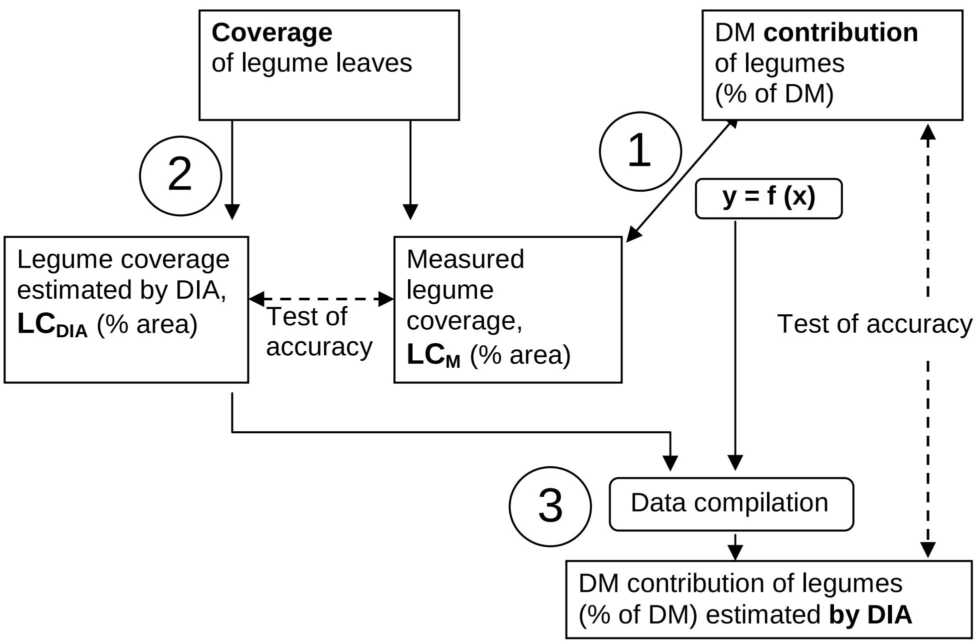


Figure 4. 1: The determination of dry matter (DM) contribution of legumes by digital image analysis (DIA). Numbers are indicating functional steps of the DIA Procedure - Step1: quantifying the relation between coverage and contribution of legumes in equation form; Step 2: processing of a digital image analysis procedure to estimate the coverage of legumes; Step 3: calculating DM contribution by DIA using the equation of Step1 and the DIA procedure of Step 2.

4.3 Results

4.3.1 Characterization of the swards

Regarding the 63-d swards, the total biomass of pure swards (mean 131.2 g m^{-2}) and mixtures (mean 129.9 g m^{-2}) were similar, whereas legume-grass mixtures outyielded pure sown legume swards when biomass only of legumes and grass was considered (mean 107.3 g m^{-2} and mean 83.5 g m^{-2}, respectively) (Figure 4.2). Both pure swards and mixtures of red clover and lucerne performed better than white clover. The R8G mixtures resulted in a clover content of 42 % (average of four replicates), whereas contents were 17 % in R2G swards, and 18 % in WG. Lucerne accounted for 36 % of the biomass in LG. In R2G and WG, the legumes contributed little to a total biomass due to presence of unsown species. The 63-d swards showed legume contents that are common for legume-grass swards in practice, so they were used as background for the image analysis procedure, whereas the 49- and the 35-d swards were included to ensure a broad application area.

Table 4. 1: Effect of seed mixture (T), sward age (A) and the interaction T*A on biomass (g DM [dry matter] m-2) and contribution of legumes, grass and unsown species (% of DM).

	Legumes		Grass		Unsown species		Total yield		Yield leg + grass	
	F	P	F	P	F	P	F	P	F	P
Biomass (g DM m^{-2})										
T	173.97	***	7.04	***	0.19	NS†	27.56	***	41.74	***
A	901.13	***	1367.61	***	68.58	***	710.72	***	630.70	***
T*A	81.81	***	15.83	***	1.22	NS	14.35	***	27.04	***
Species contribution (% of DM)										
T	65.26	***	25.1	***	2.69	NS				
A	6.95	***	8.00	***	12.27	***				
T*A	2.46	*	2.05	NS	2.39	*				

* P≤0.05,; ** P≤0.01; *** P≤0.001; †NS not significant.

Suppression of unsown species in the legume-grass mixtures and pure grass swards was stronger than in pure legume swards. This may be due to the higher DM yield of legumes and grass in these swards. The calculation for all 63-d swards resulted in a negative exponential relationship between total yield of sown species and the content of unsown species (R^2 0.58, SE 8.7). Predominant unsown species were *Alopecurus myosuroides* Huds., *Stellaria media* (L.) Vill., *Matricaria recutita* L., and *Capsella bursa-pastoris* (L.) Medik. As shown in Table 4.1, the different sown mixtures had no significant influence on the DM yield and proportion of unsown species, whereas the age of the swards did. Over the course of time, the legume species developed in different ways regarding the DM contribution, which is proved by a significant mixture x sward age interaction.

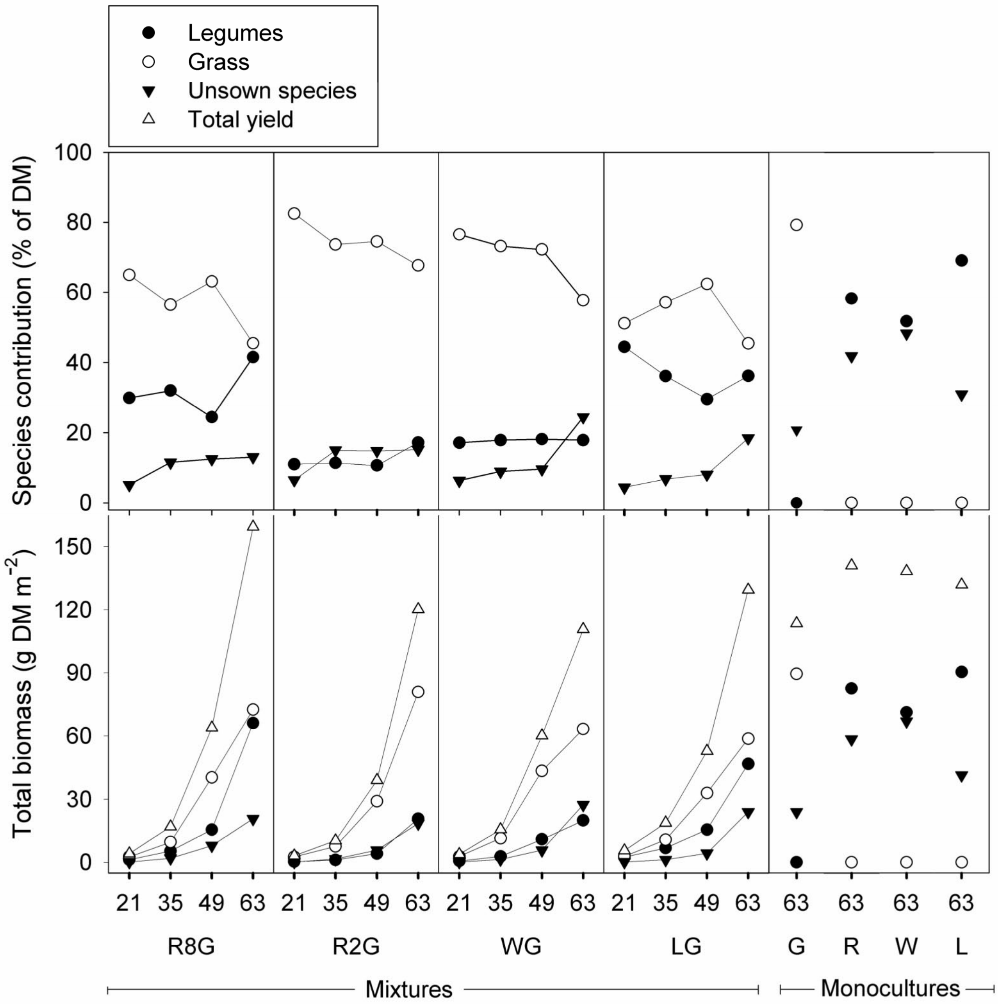


Figure 4. 2: Development of total biomass (g DM m^{-2}) and species contribution (% of DM) of legumes, grass, and unsown species. Mean values of four replicates 21, 35, 49, and 63 d after sowing, G= grass, R= red clover, W= white clover, L= lucerne, R8G= red clover (8 kg ha^{-1})-grass, R2G= red clover (2 kg ha-1)-grass, WG= white clover-grass, LG= lucerne-grass.

4.3.2 Digital Image Analysis

The percent cover of legume leaves (LC_M) was positively related to legume DM contribution (R^2 0.90, SE 5.9) across all legume species including the 35-, 49-, and 63-d swards (Figure 4.3). A better correlation was found (R^2 0.92, SE 5.8) across all legume species including only 49- and 63-d swards. Examining the three legumes species including 49- and 63-d pure swards and mixtures individually, better relationships could be found (red clover: R^2 0.94, SE 5.1; white clover: R^2 0.96, SE 4.3; and lucerne: R^2 0.97, SE 4.7) (Table 4.2).

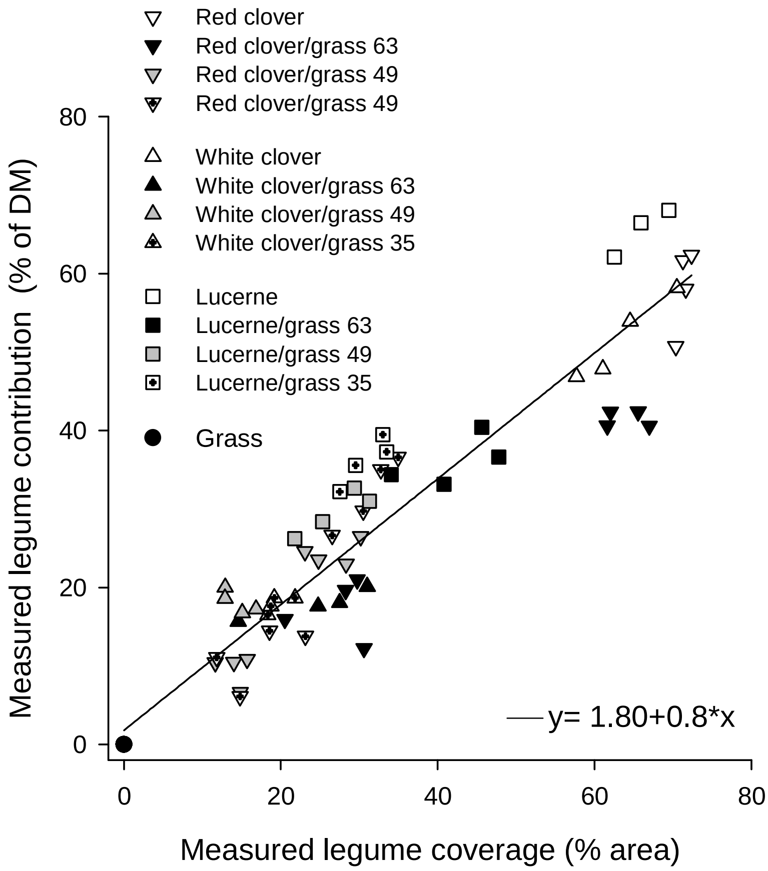


Figure 4. 3: Relationship between the measured legume coverages (% area) and the measured legume contributions (% of DM). Regression line includes 35-, 49-, and 63-d swards. Statistics are shown in Table 2.

Table 4. 2: Relationship between the measured legume coverage (x, % area) and the measured contribution of legumes (y, % of [dry matter] DM) related to different sward ages and legume species. The legume specific calculations included the mixtures and pure swards of the selected legume as well as pure grass swards.

	n	*Age of the swards included*	R^2	*SE*	*Equation*
		days		% of DM	
All swards	63	35, 49, 63	0.90***	5.90	1: y=1.80+0.80*x
All swards	47	49, 63	0.92***	5.76	2: y=1.36+0.80*x
Red clover	24	49, 63	0.94***	5.13	3: y=0.84+0.75*x
White clover	16	49, 63	0.96***	4.03	4: y=2.27+0.77*x
Lucerne	15	49, 63	0.97***	4.07	5: y=2.07+0.95*x

NS not significant; * P≤0.05; ** P≤0.01; *** P≤0.001

A structuring element of 5 x 5 and five iterations of erode and dilate turned out to be the optimal DIA procedure in terms of minimal computation effort and maximal accuracy of predictions. The use of a smaller structuring element of 3*3 pixel and 13 iterations improved the accuracy but the operation needed noticeably more computation time.

The legume coverage estimated by this DIA procedure showed good correlations with LC_M in the 49- and 63-d swards (R^2 0.87, SE 8.9). Coverage in the 35-d swards was not predicted satisfactorily because of 40 to 69 % visible bare grounds (Table 4.3). Overall accuracy was checked from the residuals between LC_M and LC_{DIA} (coverage predicted by the optimal DIA procedure) including all 35-, 49- , and 63-d swards (Figure 4.4). Negative values at low DM yield indicate that DIA poorly predicted legume coverage in young or open swards. For total DM yields above 30 g m^{-2} residuals varied between -20 and +20%, with slightly higher prediction errors for pure grass swards and swards containing lucerne than for swards containing red or white clover. The swards below 30 g m^{-2} were mostly represented by 35-d harvests, which were young and open swards. These harvests were eliminated from further DIA analysis because legume cover could not be predicted accurately.

Table 4. 3: Relationship between the measured legume coverage (LC_M, % area) and the estimated legume coverage (LC_{DIA}, % area) related to different sward ages and legume species.

	n	Age of the swards included	R^2	SE	Equation
		days		%	
All swards	63	35, 49, 63	0.55***	14.9	y=3.59+0.74*x
All swards	47	49, 63	0.87***	8.9	y=0.21+1.02*x
Red clover	24	49, 63	0.91***	7.8	y=4.29+1.05*x
White clover	16	49, 63	0.93***	6.7	y=0.75+1.01*x
Lucerne	15	49, 63	0.83***	10.1	y=1.58+1.19*x

*** P≤0.001

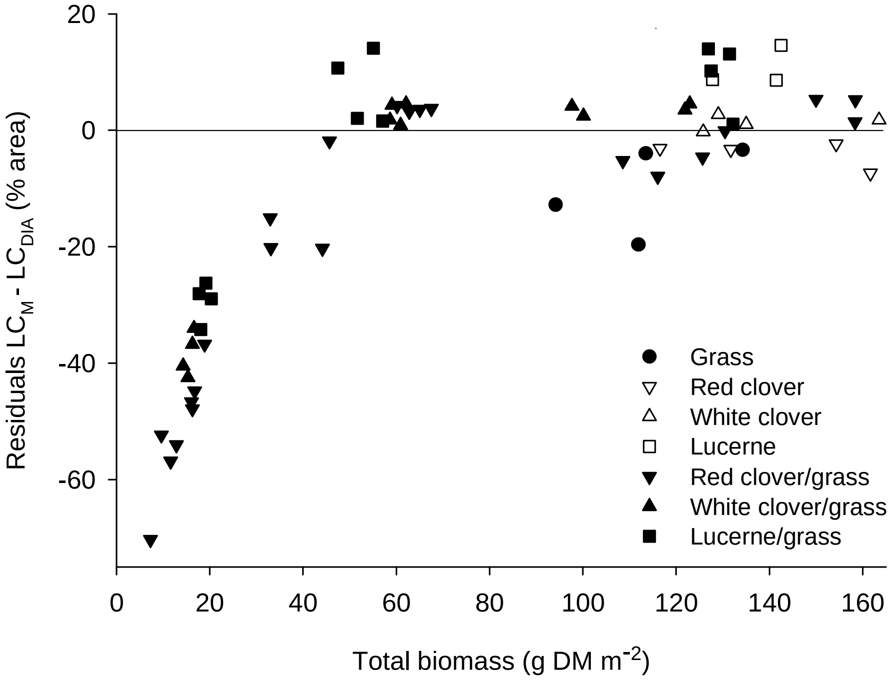


Figure 4. 4: Residuals of the measured legume coverage (LC_M, % area) and legume coverage estimated by the optimal DIA (LC_{DIA}, % area). Included are the 35-, 49- and 63-d swards.

Based on the relationship between the measure legume coverage (LC_M) and the measured contribution of legumes (% of DM), as shown in Figure 4.3, the contribution of legumes (% of DM) was calculated with the predicted legume coverage (LC_{DIA}) for all 49- and 63-d swards. The calculation was made for each legume species separately with the equations shown in Table 4.2. Comparing the measured and the estimated contributions of legumes a coefficient of determination of 0.81 (SE 8.9) was found (Figure 4.5). Neither the test for the regression coefficient, nor for the intercept suggested a deviation from the bisecting line. In mixtures that had few legumes, the DIA procedure slightly overestimated legume contribution (Figure 4.5), whereas lucerne contribution was frequently underestimated.

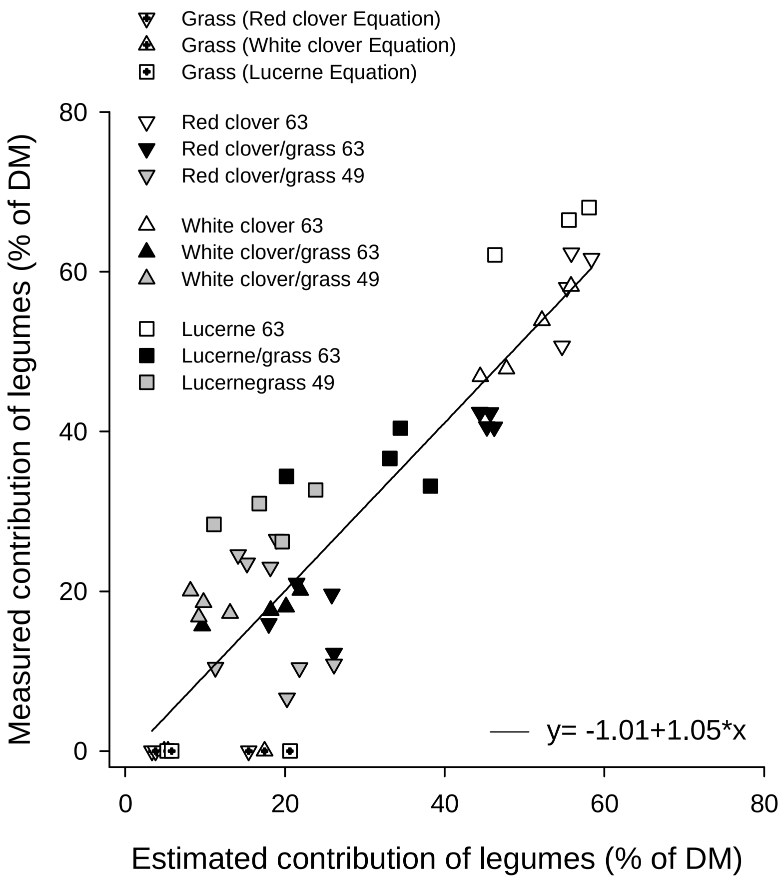


Figure 4. 5: Comparison of the measured legume contribution (% of DM) and the estimated legume contribution (% of DM) based on DIA and legume specific relations between coverage and DM contribution. Included are the 49 and 63-d swards.

4.4 Discussion

The main goal of the study was to evaluate a digital image analysis (DIA) procedure to determine the legume DM contribution of mixed legume-grass swards using pixel-based filters and threshold segmentation of the software package Optimas.

Because DIA only exploits information from the surface of the sward canopy, a prerequisite for the development of a complete procedure is to identify a relationship between cover percentage and the DM contribution of legumes. The accuracy of the common relationship based on all three legumes and sward ages of 35, 49, and 63 days (R^2 0.90, SE 5.9 % of DM) was improved by the consideration of including only 49 and 63-d swards (R^2 0.92, SE 5.8 % of DM). A further advancement can be achieved with a legume-specific approach of 49 and 63-days old swards with R^2 values of 0.94, 0.96, and 0.97 for red clover, white clover, and lucerne, respectively. These results suggest that structural attributes of the swards influence the relationship between the two parameters. Red clover and lucerne are tall legumes with leaves on an erect stem, whereas white clover is low growing, with leaves borne on petioles originating from the stolon. Given the same DM contribution of legumes, leaves of white clover and red clover can cover a larger area, as clover leaves are broad and horizontally oriented, whereas lucerne leaves are narrow.

Detection of legume DM contribution by DIA was weak in pure grass swards and in most swards at early growth stages (Figure 4.4), where biomass was low and areas of bare ground were visible. Since our approach did not include colour information, some of these bare areas were misclassified as legume leaves. Incorporating colour information may improve accuracy of DIA especially for predicting legume components in open swards. Our approach can be used for more productive swards, which are normally cut when the grass is at early head emergence (Frame, 1992) and only small areas of soil may be exposed.

Further misclassifications were caused by the occurrence of unsown species. As the experimental swards were not weeded, unsown species achieved proportions of up to 50% of the total dry matter yield. *Alopecurus myosuroides* and *Stellaria media* were the most common weeds in older swards. While the round leaves of *Stellaria media* were partly identified as legumes, *Alopecurus myosuroides* did not cause any misclassification. This is in accordance with results by Onyango et al. (2005), who found that the success of weed classification was strongly affected by the type of weed. Apart from the effects of visible bare soil, legume identification in pure grass swards may also have been caused by leaves of *S.*

media. Contrarily, underestimation of lucerne DM contribution in mixed swards was caused by the lanceolate leaf shape of lucerne.

The direct assessment of essential quantitative and qualitative traits of mixed forage swards in the field would be a major advance in an efficient and environmentally friendly management of legume-based farming systems. In low-input production systems, like organic agriculture, a synchronized determination of total yield and legume proportion by appropriate sensors in the field would allow a more accurate prediction of the amount of legume-derived nitrogen in soil in the pasture phase. In a parallel study with the same experimental swards, near infrared spectroscopy (NIRS) calibration of total sward biomass proved promising (Biewer et al., 2009). Thus, the assessment of legume dry matter yields by the combination of DIA and field spectroscopy methods could help to adjust the nitrogen management in arable systems.

4.5 Conclusion

A relationship exists between legume coverage (% area) and contribution (% of DM) in legume-grass mixtures. Thus, the DM contribution of legumes can be calculated determining the proportion of image area covered by legumes. For that purpose, a digital image analysis procedure created on the basis of greyscale sward images was successfully tested. The best results were obtained with individual predictions of red clover, white clover and lucerne. Digital image analysis is a promising approach to predict legume contribution in productive legume-grass swards but needs improvements to better characterize the influence of bare soil and weed species.

Acknowledgement

We are grateful to Matt Sanderson (USDA-ARS-Pasture Systems and Watershed Management Research Unit, University Park, PA) for pre-reviewing the manuscript.

5 The relationship between coverage and dry matter contribution of forage legumes in binary legume-grass mixtures

Abstract An efficient and accurate detection of legume dry matter (DM) contribution in legume-grass mixtures is of great importance for a targeted management of legume-based swards. Legume coverage may be an appropriate indicator for the contribution of legumes, as it can be assessed by digital image analysis. But this requires a detailed knowledge about the relationship between these sward attributes. A pot experiment under controlled conditions was conducted to examine this relationship across a wide range of legume species (*Trifolium repens* L., *T. pratense* L., *Medicago sativa* L.), legume proportion (0-80 % of DM), and growth stage (tillering to heading). Multiple regression analysis revealed a positive relationship between legume contribution and coverage for the separate legume species (R^2 0.98 to 0.99), as well as for the combined dataset including all legumes (R^2 0.98). Also the validation of the model with swards of a field experiment showed good results (R^2 0.98). Total biomass was related to clover coverage in a complex manner, reducing legume contribution by up to 20% (relative) with an increase of total biomass from 30 to 150 g DM m^{-2}.

5.1 Introduction

Legumes with their ability to fix nitrogen contribute considerably to the nutrient supply in forage production (Boller and Nösberger 1987; Frame et al. 1985; Heichel and Henjum 1991; Wachendorf et al. 2004), which is of particular importance in organic agriculture, where they contribute 20-50 % to the arable farm area. In northwest Europe legume-grass swards are usually grown as short-term grassland for 1-3 years in a crop rotation system and are managed by cutting and used for silage or fresh fodder. It is well known that the yield and proportion of legumes in a sward are strongly related to the amount of fixed nitrogen (Loges, 1998; Høgh-Jensen et al., 2004). Therefore, frequent information on the status of legume-grass swards could help to direct fertilizer applications and to predict the nitrogen supply of the soil for arable crops in the field to be grown after the legume-grass mixtures.

As manual clipping and separation of legume-grass swards, which is a common procedure in field experimentation, is not feasible in agricultural practice, a methodology is desirable which allows a quick and non-laborious assessment of the legume biomass in mixed swards. The indirect assessment of the sward biomass by field spectroscopy (Biewer et al., 2008, 2009) was shown to produce reliable data but could not determine the legume DM contribution. For this purpose, digital image analysis (DIA) produced promising results when applied to different forage legume-grass swards (Himstedt et al., 2009). While the detection of the legume coverage (% area) from digital images may be further improved by novel image analysis routines, there is a gap in knowledge on the relationship between the coverage and the DM contribution of legumes in mixtures. Yet, the precise knowledge of this relationship as well as on the effects of legume species and sward age is of great importance for the application of DIA procedures. Recent results from a pot experiment, including various binary legume-grass mixtures, suggest close legume-specific relationships between legume coverage and legume contribution (% of DM) for 49- and 63-d-old swards (red clover R^2 0.94, SE 5.1; white clover R^2 0.96, SE 4.0, and lucerne R^2 0.97, SE 4.1), whereas the accuracy for the common relationship was lower (R^2 0.92, SE 5.8). The integration of low-yielding 35-d-old swards further reduced the accuracy of the common relationship (R^2 0.90, SE 5.9) (Himstedt et al., 2009). Furthermore, some relationships lead to negative predictions for legume contribution when coverage was close to zero and heteroscedasticity was a frequent problem of the models, which is common for percentage data, as there is less scope for response at levels close to zero than at intermediate values.

Thus, the objective of the present study was to establish relationships between the coverage (% area) and the contribution (% of DM) of legumes in binary legume-grass mixtures, which are valid for a wide range of legume species and sward age and which produce reliable predictions even at low levels of legume DM contribution and coverage.

5.2 Material and methods

5.2.1 Pot experiment

In 2004 a 9-wk pot experiment with pure swards and binary legume-grass mixtures of white clover (*Trifolium repens* L.), red clover (*T. pratense* L.), lucerne (*Medicago sativa* L.), and perennial ryegrass (*Lolium perenne* L.) was conducted. Eight experimental swards were investigated: monoculture swards of ryegrass (20 kg of seed ha^{-1}), red clover (8 kg ha^{-1}), white clover (4 kg ha^{-1}) and lucerne (16 kg ha^{-1}) and the binary mixtures red clover-ryegrass (8/20 kg ha^{-1} and 2/20 kg ha^{-1}), white clover-ryegrass (4/20 kg ha^{-1}), and lucerne-ryegrass (16/20 kg ha^{-1}). To compare different aged swards, legume-grass mixtures were sown in calendar weeks 48, 50, and 52 and harvest was performed at a common date, which was 35, 49, or 63 d after sowing. Monoculture swards were harvested 63 d after sowing. All treatments were sown manually in four replicates at a distance of 12 cm between rows and a sowing depth of 0.5 cm. Sward size was 0.119 m^2, length and width were 34 by 35 cm, including three rows of seedlings. Wooden pots were filled with 2 cm drainage substratum (Lavagrus) and about 16 cm homogenised loamy soil (sandy loam; 3,6% sand, 73% silt, 23,4% clay, and 2% humus). Soil analysis indicated sufficient contents of phosphorus, magnesium and potassium and a pH-value of 6.7. No fertilizers were applied.

5.2.2 Field experiment

For validation of the models which were developed with swards grown under greenhouse conditions data and digital pictures of swards from a field experiment were used. The field experiment was conducted on the organic experimental farm Neu Eichenberg of the University of Kassel (5123'N, 954'E, 227 m a.s.l.). In addition to pure swards of perennial ryegrass (15 kg ha^{-1}), red clover (8 kg ha^{-1}), and white clover (4 kg ha^{-1}) binary mixtures of each legume with perennial ryegrass (8/15 kg ha^{-1} and 4/15 kg ha^{-1}, respectively) were tested. The soil characteristics were the same as for the pot experiment. The experimental treatments were established in four replicates on 2 June 2005. During the two-year experiment the average rainfall was 550 mm and the average temperature 9.9° C. In 2006, spring, summer and autumn growth were sampled at weekly intervals to determine effects of growth stages and harvested on 12 June, 25 July and 14 September, respectively. For validation only swards with a total biomass in the approximate scope of the model were used (biomass up to 280 g m^{-2}). This was appropriate to a set of 46 swards sampled on 3 and 8 May, 10 and 17 June, as well as on 23 and 28 August.

5.2.3 Measuring of legume leaf coverage and dry matter contribution

One day before the harvest digital pictures of the swards were taken with a Canon (Tokyo, Japan) Power Shot G6 Digital Camera, using flashlight for all pictures to ensure uniform illumination. On the basis of four fixed vertices the digital pictures were georeferenced with the program SAGA (System for Automated Geoscientific Analyses Version 2.0, 2005) to compensate for distortions. The actual legume coverage (% area) was measured manually tracing the legume covered areas in each picture. These areas were digitized and measured as the percentage of the green covered area of the picture. Visible bare ground was excluded from the calculation. To determine the sward composition, total aboveground biomass was fractionated into grass, legumes and unsown species. The samples were dried 24 hours at 65°C. As no weeding was conducted, some swards contained appreciable amounts of unsown species (Table 5.1). The following analysis was carried out legume specific for red clover-grass (RCG), white clover-grass (WCG), lucerne-grass (LCG), and additionally for the combined dataset (CDS) including all swards. Each model includes purely sown swards of grass and legumes and the respective binary mixtures. The combination of various levels of sowing rate and sward age resulted in a diverse data set, covering a wide range in the variables relevant in this study (Table 5.1). More details of sward characteristics are given elsewhere (Himstedt et al., 2009).

5.2.4 Transformation of variables, statistical analysis and presentation of results

Legume leaf coverage (% area) and DM contribution (% of DM) frequently showed values equal or close to zero. These values are at the minimum achievable and suggest that models in which the untransformed legume contribution was the response could lead to predictions below zero, and might also include interactions induced by scale rather than biology (Cox and Snell, 1989). Under conditions in which levels of legume contribution are close to zero, the absolute effects of treatments or other factors or variables tend to be smaller than at intermediate values, simply because there is less scope for response. For example, a factor that doubles low legume contribution will change 2 to 4% at very low levels but will increase 20 to 40%, a seeming interaction with legume level which obscures the simplicity of the doubling effect of the factor. Such a factor could not work in this way at legume contributions above 50%. To avoid these difficulties, legume leaf coverage (LC) and dry matter contribution (LD) were transformed to the logit-scale, defined as log[LC/(100-LC)] and log[LD/(100-LD)], before inclusion as independent and response variable. The logit transformation tends to give a scale that reduces the effect of interaction simply due to scale and also ensures that the model will lead to a predicted legume contribution lying within the range 0 to 100 (Connolly and Wachendorf, 2001). Moreover, heteroscedasticity was a frequent problem in modelling on the natural scale and various transformations (logarithm, inverse and square root) were tested. By far the best of these in terms of eliminating heteroscedasticity was the logit transformation.

Multiple regression analysis was conducted using the GLM procedure of SAS 9.1 (SAS institute, 2002-2003). In the calculations, 35-, 49-, and 63-d-old swards were included. The selection of terms for inclusion in the model depended on standard statistical model selection methods (Draper and Smith, 1998) and obeyed the rules of hierarchy and marginality (Nelder, 1994; Nelder and Lane, 1995). Effect terms were included if their significance exceeded the 5% level. The rules of hierarchy ensured that if an interaction was included in the model then the variables involved in the interaction were also included separately. The modelling obeyed the marginality principle (Nelder and Lane, 1995), which implies that if a term appears as part of a more complex element in the model then, in general, the term itself is not tested for significance. This is because the meaning of such terms is open to misinterpretation (Connolly and Wachendorf, 2001). Thus, t- and P-values for model coefficients where the term is involved in a higher order term in the model were generally omitted. The models, expressed as estimates of multiple regression coefficients (Table 5.3), may be difficult to appreciate and interpret, particularly as transformations of response and independent variables have been used. In the current analysis, predictions were back-transformed to the original scale (i.e. legume coverage in % area and legume contribution in % of DM), using

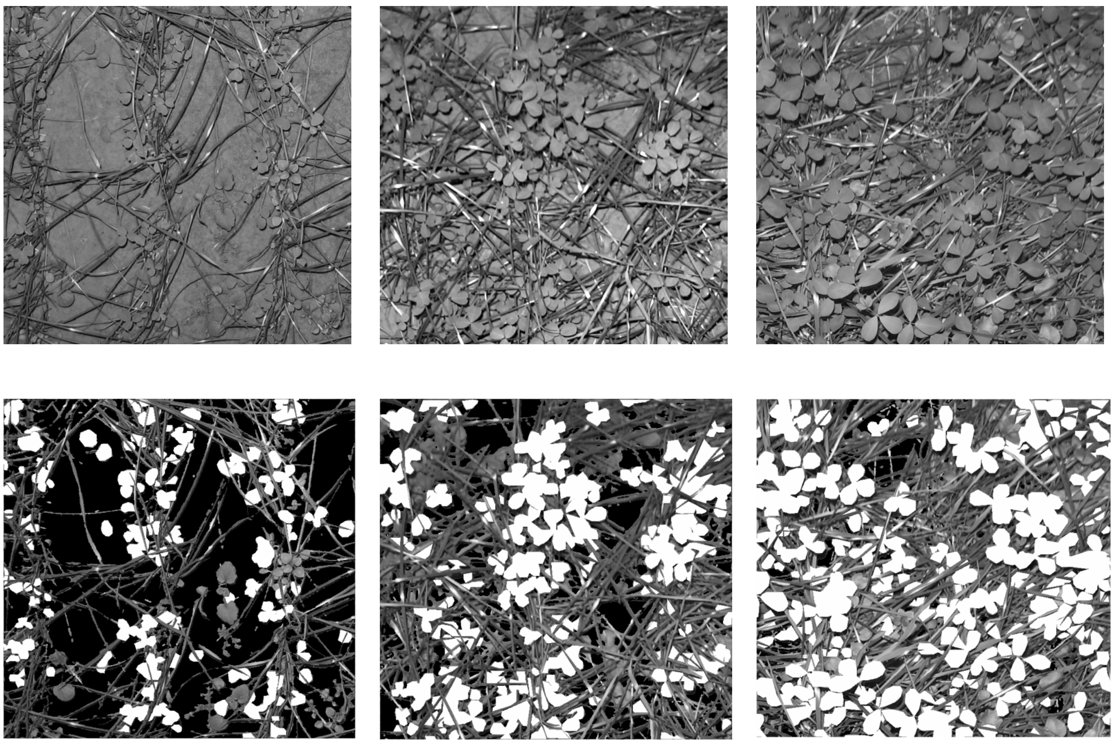
$$\text{Predicted} \quad \text{LC} = \frac{100 \times e^{\text{predicted logit(LC)}}}{1 + e^{\text{predicted logit(LC)}}} \qquad \text{Eq. [5.1]}$$

$$\text{Predicted} \quad \text{LD} = \frac{100 \times e^{\text{predicted logit(LD)}}}{1 + e^{\text{predicted logit(LD)}}} \qquad \text{Eq. [5.2]}$$

and a graphical presentation was used. Predictions for LC as a continuous variable are plotted as lines, one each for a different total biomass (approximately its mean up to its standard deviation). The range chosen for prediction from the independent variables was selected to exclude values close to the observed minimum and maximum of the variable (Table 5.1), and predictions outside the range of the observed data were excluded.

5.3 Results and discussion

From a previous study dealing with the determination of legume DM contribution in binary legume-grass mixtures using DIA it turned out that the underlying relationship between legume coverage and DM contribution is of fundamental importance for the DIA procedure (Himstedt et al., 2009). Figure 5.1 illustrates for LCG mixtures that with increasing total biomass the legume coverage increased from 27.5% to 40.8% of the green leaf area, whereas the legume contribution (approximately 33% of DM) nearly remained constant. Similar findings also occurred for the red clover- (RCG) and white clover-grass (WCG) mixtures.



Age	35 days	49 days	63 days
BM	20 g DM m^{-2}	52 g DM m^{-2}	132 g DM m^{-2}
LC	27.5% area	29.4% area	40.8% area
LD	32.2% of DM	32.7% of DM	33.2% of DM

Figure 5. 1: Greyscale digital images (top row) and pictures showing manually classified legume leaves (bottom row) of lucerne-grass swards of different ages (day after sowing), total biomass (BM, g DM m^{-2}) and leaf coverage (LC, % of green area) but similar legume contribution (LD, % of DM). (White overlay: legume covered area, black overlay: bare ground).

Table 5. 1: Characteristics of the swards included in the red clover-grass (RCG), white clover-grass (WCG), lucerne-grass (LCG), and combined dataset (CDS). Each model includes purely sown swards of grass and legumes and the respective binary mixtures. The legume coverage is measured as % area of the green area within the digital image.

Sward type	n	Total biomass				Grass contribution				Weed contribution				Legume contribution				Legume coverage			
		Mean	SD†	Min	Max	Mean	SD	Min	Max‡	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
		(g DM m^{-2})				(% of DM)												(% area)			
RCG	32	83.1	55.8	7.3	170.9	56.6	24.7	0	85.8	19.1	11.8	5.1	50.2	24.3	18.3	0	62.3	31.5	23.4	0	72.4
WCG	20	87.7	46.5	14.3	163.5	56.5	30.1	0	85.8	22.4	15.1	6.5	53.1	21.1	17.4	0	58.2	24.4	22.0	0	70.5
LCG	19	87.9	48.2	17.7	142.5	51.4	25.8	0	85.8	16.8	10.6	5.3	37.9	31.8	20.7	0	68.0	31.4	21.4	0	69.5
CDS	71	85.7	50.7	7.3	170.9	55.2	26.3	0	85.8	19.4	18.9	5.1	53.1	25.4	18.9	0	68.0	29.5	22.4	0	72.4

† SD: Standard deviation; ‡ The maximum grass contribution is identical for all sward types, as they include the same purely sown grass swards.

Table 5. 2: Relationship between measured legume contribution (g DM m^{-2}) and the measured legume coverage (% area) using models A-D for the red clover-grass (RCG), white clover-grass (WCG), lucerne-grass (LCG), and combined dataset (CDS).

Effects in the model	Sward age included (days)	RCG			WCG			LCG			CDS		
		R^2	Pr > F	RMSE†	R^2	Pr > F	RMSE	R^2	Pr > F	RMSE	R^2	Pr > F	RMSE
A: LC	49, 63	0.94	<0.0001	5.13	0.96	<0.0001	4.03	0.97	<0.0001	4.07	0.92	<0.0001	5.76
B: LC	35, 49, 63	0.91	<0.0001	5.54	0.96	<0.0001	3.58	0.96	<0.0001	4.47	0.90	<0.0001	5.90
C: LC, BM, LC*BM‡	35, 49, 63	0.94	<0.0001	4.88	0.97	<0.0001	3.25	0.98	<0.0001	3.29	0.93	<0.0001	5.16
D: LLC, BM, LLC*BM§¶	35, 49, 63	0.98	<0.0001	0.31	0.99	<0.0001	0.24	1.00	<0.0001	0.14	0.98	<0.0001	0.31

The independent variables were total biomass (BM), legume coverage (LC) and the logit-transformed legume coverage (LLC).
†RMSE, root mean square error of the model; ‡LC*BM interaction only significant for RCG and CDS; §LLC*BM interaction only significant for CDS;
¶The response variable in Model D was transformed to the logit scale.

The initial step in extracting useful information on the relationship between legume coverage and legume DM contribution from the broad and complex body of data, was to extend the original dataset as reported by Himstedt et al. (2009) (Table 5.2, Model A) through the inclusion of 35-d-old swards, which were more open and low-yielding. As a result, the coefficients of determination decreased and standard errors increased for red clover swards, lucerne swards and the combined dataset, indicating that within the legume species the pattern of the relationship differed among the different levels of total biomass (Table 5.2, Model B). Because of this the relationship between total biomass and legume coverage and legume contribution respectively was tested, but neither for any legume-specific nor for the combined dataset (CDS) a significant relationship was found (data not shown). Next, total biomass (BM) was included in the model and tested as single linear and quadratic term and in interaction with legume coverage. As none of the quadratic terms and complex interactions were significant, Model C was chosen to be the best in terms of R^2 ($\geq$ 0.94) and root mean square error (RMSE; $\leq$ 4.88 g DM m^{-2}) with a significant LC x BM interaction only for RCG and CDS. Although the accuracy of Model C was similar to that of the original Model A, but covered a wider range in terms of sward age, the residual plots indicated severe heteroscedasticity for the models and predictions for clover contribution at low levels of legume coverage were frequently below zero. To prevent these problems, both legume contribution (as the response variable) and coverage (as explanatory variable) were logit-transformed in Model D. Due to the transformation, the R^2 ($\geq$ 0.98) and RMSE ($\leq$ 0.31 g DM m^{-2}) values, although they indicate a high model accuracy, can not be compared to those of the Models A-C which are on the original scale.

The parameter estimates (Table 5.3) include a significant BM x LLC interaction for CDS, while for the single legume species the interaction was not significant. The diagrams show a similar shape for the three legume species (Figure 5.2) and illustrate that, at an intermediate level of legume coverage of 40% the DM contribution is reduced by 10% if the total biomass increased from 30 to 150 g DM m^{-2}. This may be because more non-leguminous biomass (i.e. grasses and weeds) can be covered by legume leaves at high levels of total biomass than at low levels. The BM x LLC interaction in the CDS model displays the legume specific characteristics an enables a common calculation. Furthermore, predicted legume contributions are close to zero at very low levels of legume coverage, without showing negative values. The relationships between the measured legume contribution and the calculated legume contribution resulting from the back-transformation of variables indicate high model accuracy for all legume species as well as for the common dataset (Figure 5.3):

RCG: $y = 0.16 + 1.01\,x$; R^2 0.93; SE 4.81

WCG: $y = 0.65 + 0.96\,x$; R^2 0.97; SE 3.22

LCG: $y = -0.50 + 1.01\,x$; R^2 0.98; SE 3.07

CDS: $y = 0.09 + 1.01\,x$; R^2 0.93; SE 5.00

with y = measured legume contribution (% of DM)

x = calculated legume contribution (% of DM).

Table 5. 3: Parameter estimates for model D of legume contribution (g m^{-2}) of the red clover-grass (RCG), white clover-grass (WCG), lucern-grass (LCG), and combined dataset (CDS).

Variable	Estimate	SE	t value	Pr > |t|
<u>RCG</u>				
INTERCEPT	-0.298	0.1109	-2.68	0.0119
LLC	0.923	0.0248	37.15	<0.0001
BM	-0.002	0.0010	-2.43	0.0214
<u>WCG</u>				
INTERCEPT	0.055	0.1258	0.44	0.6650
LLC	0.941	0.0216	43.56	<0.0001
BM	-0.004	0.0012	-3.03	0.0076
<u>LCG</u>				
INTERCEPT	0.275	0.0692	3.98	0.0011
LLC	0.990	0.0119	82.98	<0.0001
BM	-0.003	0.0007	-4.48	0.0004
<u>CDS</u>				
INTERCEPT	0.168	0.1346	1.25	0.2154
LLC	1.135	0.0942	-[†]	-
BM	-0.004	0.0011	-	-
LLC*BM	-0.002	0.0008	-2.04	0.0450

The response variable was transformed to the logit scale. The independent variables were biomass (BM, g DM m^{-2}) and the logit-transformed legume coverage (LLC).

[†]t-values and probabilities of main effects were omitted when the effect was included in a significant interaction.

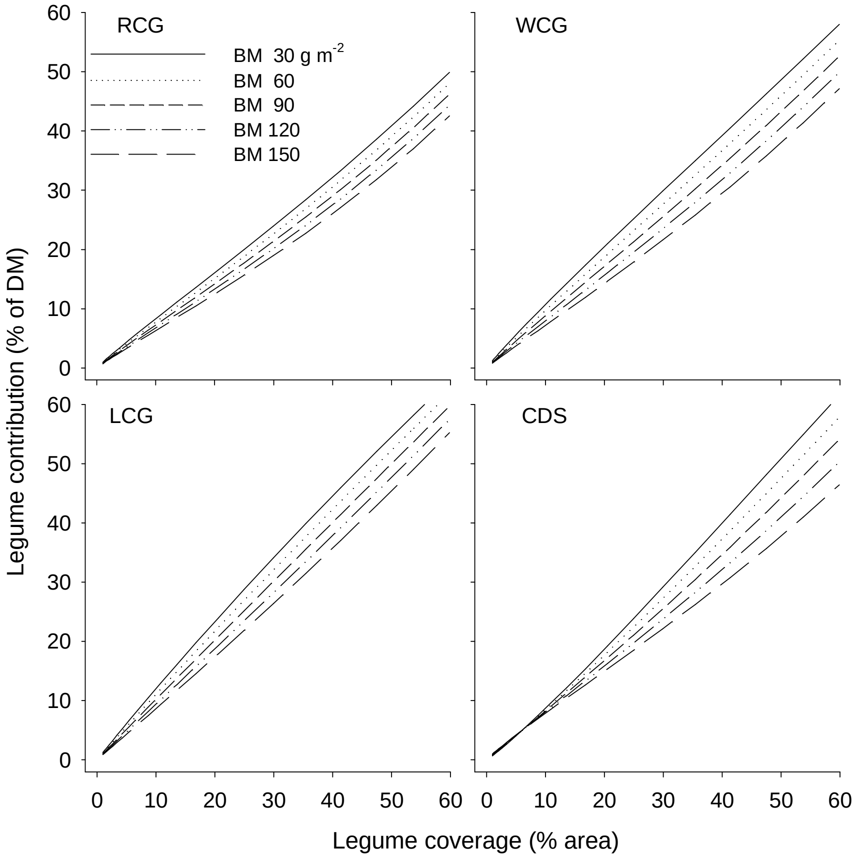


Figure 5. 2: Predictions of the legume contribution (% of DM) with the model D from Table 5. 2 for the RCG (Red clover-grass), WCG (White clover-grass), LCG (Lucerne-grass), and CDS (combined dataset) swards.

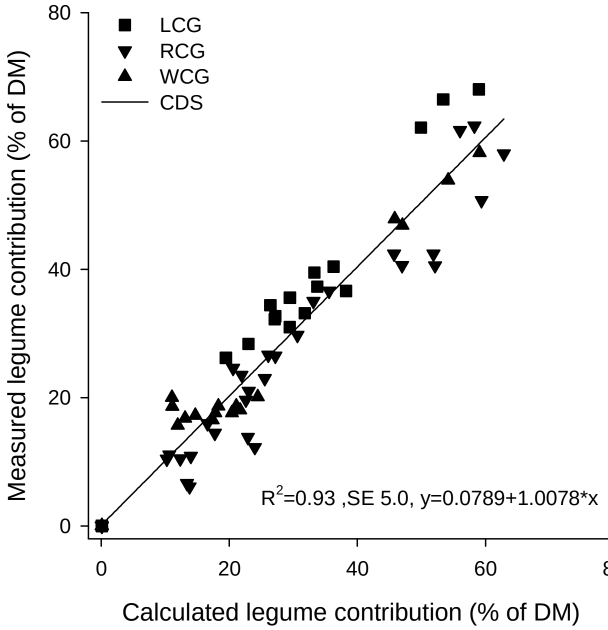


Figure 5. 3: Relationship between measured legume contribution (% of DM) and predictions from the model D (Table 5. 2) for the combined dataset (RCG: Red clover-grass, WCG: White clover-grass, LCG: Lucerne-grass, and CDS: Combined dataset).

The regression lines almost matched the bisecting line and residuals generally indicated a sufficient homogeneity of variance. For all datasets tests for intercept equal to zero and slopes equal to 1 did not show any significance. Accuracy of the model for the common dataset was lower than for WCG and LCG, but comparable to RCG. Apparently, for the range of total biomass occurring in the present study, which did not include mature swards with strongly elongated stems, the pattern of relationships among the different legume-grass swards was very similar.

The validation of the model with swards of a field experiment showed good results (R^2 0.98, SE 6.0; Figure 5.4). The biomass range of these swards was selected considering the scope of the model. Differing to the swards of the pot experiment grasses and legumes showed elongated stems and sward height was up to 45 cm. This shows, that the model can predict legume contribution for most practical legume-grass swards (Frame, 1992; Ledgard and Steele, 1992; Loges, 1998).

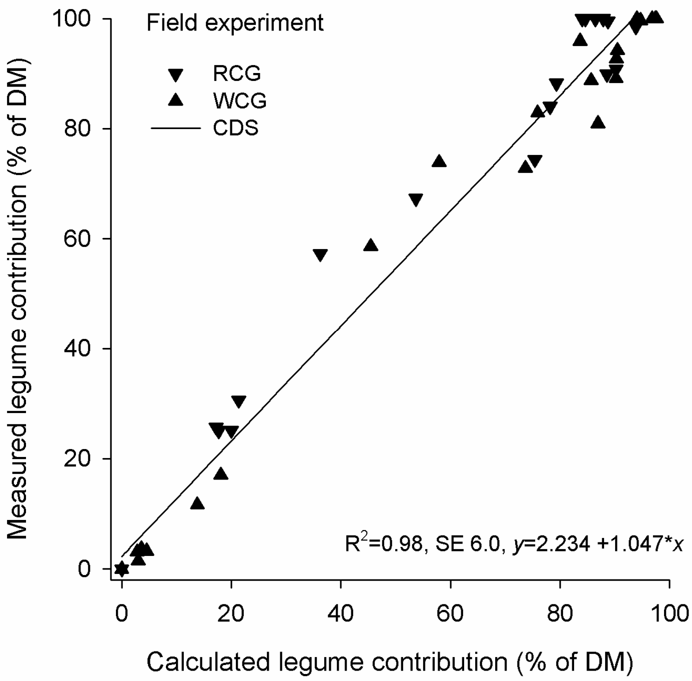


Figure 5. 4: Validation of the model with swards of a field experiment. Relationship between measured legume contribution (% of DM) and predictions from the model D (Table 5.2) for the combined dataset (RCG: Red clover-grass, WCG: White clover-grass, and CDS: Combined dataset).

Nevertheless, further research is necessary to adjust the relationships at higher levels of total biomass, where the differences among legume species (e.g. between the tall growing red clover and the more prostrate growing white clover) may become more prevalent.

The integration of total biomass into the model for determining legume contribution does not necessarily reduce its applicability in practice. Recent results from field experiments demonstrated the potential of near infrared field-spectroscopy for the prediction of total biomass of legume-grass mixtures (Biewer et al., 2009). Thus, a combined estimate of total biomass and legume coverage by field spectroscopy and DIA, respectively, may allow an accurate prediction of the legume contribution of legume-grass mixtures.

5.4 Conclusions

The modelling strategy was successful in developing biologically meaningful models giving insight into the relationship between the legume coverage in binary legume-ryegrass communities and the legume contribution. Problems with heteroscedasticity and negative predictions on the original scale did not occur when legume contribution and coverage were transformed to the logit-scale. Positive relationships between legume contribution and coverage were found for the separate legume species, as well as for the combined dataset. Given the same level of legume coverage, legume contribution decreased with increased total biomass. This phenomenon, which occurred for all legume species as well as for the combined dataset, may be because more non-leguminous biomass was covered by legume leaves at high levels of total biomass than at low levels. The relationships established between legume coverage and DM contribution provides the basis for a rapid and precise estimation of legume contribution by digital image analysis.

Acknowledgement

We are grateful to Matt Sanderson (USDA-ARS-Pasture Systems and Watershed Management Research Unit, University Park, PA) for pre-reviewing the manuscript.

6 An advanced image analysis procedure to estimate legume contents in legume-grass swards

Abstract An efficient and accurate detection of legume dry matter (DM) contribution in legume-grass mixtures is of great importance for a targeted management of legume-based swards. Legume coverage may be an appropriate indicator for the contribution of legumes, as it can be assessed by digital image analysis (DIA). A pot experiment under controlled conditions was conducted to examine the perspectives of DIA across a wide range of legume species [white clover (*Trifolium repens* L.), red clover (*T. pratense* L.), lucerne (*Medicago sativa* L.)], legume proportions (0-70 % of DM) and growth stages (begin of tillering to begin of heading). An advanced procedure for the determination of legume dry matter contribution by digital image analysis is suggested, which comprises the inclusion of colour information in the analysis of images and which applies an advanced function to predict legume dry matter contribution from legume coverage by considering total sward biomass. Bare soil areas in young and open swards could be determined very accurately which in turn allowed a precise estimation of legume coverage of green area across a wide range from 0 - 72.4 %. This resulted in a very accurate prediction of legume contribution (% of DM) with R^2 of 0.90, 0.94 and 0.93 for red clover, white clover and lucerne, respectively.

6.1 Introduction

Under moderate European conditions legume-grass swards are usually grown as short-term grassland for 1-3 years in a crop rotation system and are managed by cutting and used for silage or fresh fodder. The amount of fixed nitrogen in a sward is strongly related to the yield and contribution of legumes (Loges, 1998; Carlsson and Huss-Danell, 2003; Høgh-Jensen et al., 2004). Frequent information on the status of legume-grass swards could help to direct fertilizer applications and to predict the nitrogen supply of the soil for the subsequent crop. An indirect assessment of the sward biomass by field spectroscopy (Biewer et al., 2008, 2009) or imaging spectroscopy (Schut and Ketelaars, 2003a, 2003b) was shown to produce reliable data but could not determine the legume DM contribution. Based on greyscale images a digital image analysis (DIA) procedure was proposed by Himstedt et al. (2009) which allows the estimation of legume DM contribution with an acceptable accuracy for swards with more than 30 g DM m^{-2}. Legume contribution (% of DM) estimated by DIA was significantly correlated with the measured legume contribution (red clover: R^2 0.85, RMSE 7.8 %; white clover: R^2 0.87, RMSE 7.3 %; lucerne: R^2 0.79, RMSE 10.73 %). However, in younger and more open swards with less then 30 g DM m^{-2} misclassification of bare soil as legume was a serious problem. On average, 24 % of the area classified as legume was visible bare soil, with a range from 0.3 to 92 % (Himstedt et al., 2008).

Recent research showed that the differentiation in image analysis between bare soil and crop tissue was improved by the integration of colour information, such as hue (H), saturation (S) and lightness (L) (Himstedt et al., 2008). A threshold of H and S was also used by Richardson et al. (2001) and Karcher and Richardson (2005) to objectively measure turfgrass characteristics such as percent ground cover and turf colour. The HSL or HSI (hue, saturation, intensity) space is often preferred, as the RGB (red, green and blue) space suffers a strong degree of correlation among the three components. Furthermore, the three values are highly sensitive to the variation of lightness, whereas the classification based on HSL is less influenced by shadows and fluctuations of illumination (Cheng and Sun, 2000; Lock et al., 2004; Lee et al., 1999). Camargo and Smith (2009) also effectively segmented diseased from healthy plant areas with converting RGB images into H and special I colour transformations.

DIA was successfully introduced for weed identification in arable crop, where plants could be detected against a uniform soil background (Sökefeld et al., 2007). Under such conditions plant species identification worked well with the application of form parameters (Petry and Kühbauch, 1989; Sökefeld et al., 2000). Lee et al. (1999) stated that the recognition of plant species by the use of form parameters performed well when operated on a smooth surface with distinctly shaped and well separated objects. Gebhardt et al. (2006) identified dock (*Rumex obtusifolius* L.) in grasslands by the use of form parameters like area and perimeter, which was supported by the distinct difference between the broad-leaved dock leaves and the much smaller grass and legumes leaves and the infrequence of occlusions.

Morphological operators are important tools in DIA procedures. It is particularly the Erode and Dilate operators which support the differentiation of objects of different shape by shrinking and dilating objects (Soille, 1999). Onyango et al. (2005) used the erosion to detect cabbage plants against smaller unsown plants, which then faded out. Van Droogenbroeck and Buckley (2005) found that, although erosion and dilation were closely related they did not work exactly in reverse, i.e. one may end up with a smaller area than the original. When applied to images of legume-grass mixtures thin grass leaves were removed whereas rounder clover leaves were left (Himstedt et al. 2009). Bonesmo et al. (2004) used these operators successfully for semi-automatically mapping of legume coverage in smooth meadow grass (*Poa pratensis* L.) and white clover-dominated swards for experimental analysis.

An improved mathematical procedure was proposed by Himstedt and Wachendorf (2009) for the calculation of legume DM contribution from leaf coverage, comprising following features: i. transformation of legume contribution and coverage data to the logit-scale in order to prevent both problems with heteroscedasticity and negative predictions which frequently occur with models on the original scale. ii. inclusion of total biomass (BM) information in order to increase the model accuracy and to allow for a wider range in terms of sward age.

This paper gives results of an advanced DIA procedure for estimating legume DM contribution when all elements mentioned above (colour information; improved modelling) were fully implemented and compares them with those of a standard DIA procedure, as suggested by Himstedt et al. (2009). A special consideration was paid to young and open swards, where standard DIA procedures completely failed in producing reliable predictions for legume DM contribution.

6.2 Material and methods

In 2004 a 9-wk pot experiment with pure swards and binary legume-grass mixtures of white clover, red clover, lucerne, and perennial ryegrass (*Lolium perenne* L.) was conducted. Eight experimental swards were investigated: monoculture swards of ryegrass (20 kg of seed ha^{-1}), red clover (8 kg ha^{-1}), white clover (4 kg ha^{-1}) and lucerne (16 kg ha^{-1}) and the binary mixtures red clover-ryegrass (8/20 kg ha^{-1} and 2/20 kg ha^{-1}), white clover-ryegrass (4/20 kg ha^{-1}) and lucerne-ryegrass (16/20 kg ha^{-1}). To compare different aged swards, legume-grass mixtures were sown in calendar weeks 48, 50, and 52 and harvest was performed at a common date, which was 35, 49, or 63 d after sowing. Monoculture swards were harvested 63 d after sowing. All treatments were sown manually in four replicates at a distance of 12 cm between rows and a sowing depth of 0.5 cm. Sward size was 0.119 m^2, length and width were 34 by 35 cm, including three rows of seedlings. Wooden pots were filled with 2 cm drainage substratum (Lavagrus) and about 16 cm homogenised loamy soil (sandy loam; 3.6% sand, 73% silt, 23.4% clay, and 2% humus). Soil analysis indicated sufficient contents of phosphorus, magnesium and potassium and a pH-value of 6.7. No fertilizers were applied. To determine the sward composition, total aboveground biomass was sorted to grass, legumes and unsown species. The samples were dried 48 h at 65°C.

For the development of DIA procedures legume-specific datasets were used (red clover RCG, white clover WCG, lucerne LCG) with each including the mixtures and pure swards of the selected legume as well as pure grass swards. The averaged legume contribution of RCG was 24.3 % of DM ranging from 0 % of DM in pure grass swards to 62.3 % of DM in purely sown legume swards. In WCG average legume contribution was 21.1 % of DM (0 - 58.2 % of DM) and in LCG 31.8 % of DM on average (0 - 68.0 % of DM). As the experimental swards were not weeded, unsown species achieved averaged contributions of 19.3 % (5.1 to 53.1 %) of the total dry matter yield. Predominant unsown species were *Alopecurus myosuroides* Huds., *Stellaria media* (L.) Vill., *Matricaria recutita* L., and *Capsella bursa-pastoris* (L.) Medik.

6.2.1 Digital image analysis procedure

One day before the harvest, digital pictures of the swards were taken with a Canon (Tokyo, Japan) Power Shot G6 Digital Camera. To ensure uniform illumination flashlight was used for all pictures. Distance from the camera to the sward surface was kept constant at 80 cm at the nadir position of the plot. A superfine compression was chosen, and the pictures were saved in JPEG (Joint Photographic Experts Group) format at a resolution of 2592 x 1944 pixels. Using four concrete vertices, digital pictures were georeferenced with the program SAGA (System for Automated Geoscientific Analyses Version 2.0, 2005) to compensate distortions. The sward images were clipped to obtain the swards without rims of vessel. The resulting image size was 766 x 744 pixels in tiff (tagged image file format) related to resolution and plot size. For image analysis the image processing software Optimas of the Media Cybernetics Company was used. A total of 64 sward images were available.

Digital RGB images were converted into 24-bit HSL colour images (Hue, Saturation, Lightness) and discrimination between bare soil and plants was performed using histogram segmentation with legume-specific thresholds (Delon et al., 2005). The overall average of threshold values was used for automating the DIA procedure. Due to the pronounced differences in the shape of legume and grass leaves, the DIA procedure applies erosion and dilation filters. These filters were defined as quadratic structuring elements with a dimension of 7*7 pixels which scan the whole image pixel by pixel to remove small objects (e.g. small grass leaves) by applying first an erosion and subsequently a dilation procedure. The erosion replaced the H, S and L value of the central pixel of the structuring element by the minimum value within the element. Hereby small areas in the image, like grass leaves, continuously get darker, whereas in major areas, like legume leaves, light-coloured values persist in the central area. The subsequent dilation works in reverse by replacing the central pixel of the structuring element with the maximum H, S and L value which remained within the structuring element. Implementing both filters at the same frequency, the remaining light areas almost obtained their previous size. The optimum size of structuring element and number of iterations were determined by test series.

The detailed algorithm of the optimized image analysis procedure (DIA_C) was as follows:

1. Import of the images in tiff-format with 766x744 pixels.

2. Conversion of the digital 24-bit RGB colour images into 24-bit HSL images.

3. Morphological filtering:

Opening (structuring element size 7x7 pixels), Erode (structuring element size 7x7 pixels) (Iterations: RCG 3, WCG and LCG 2) and Dilate (structuring element size 7x7 pixels) (Iterations: RCG 3, WCG and LCG 2).

4. Colour segmentation:

Identification of bare soil with threshold H (16-45) S (40-154) L (41-173), Identification of legume covered areas with legume specific thresholds (RCG, H (54-91) S (35-200) L (75-200); WCG, H (55-75) S (52-188) L (80-200); LCG, H (54-85) S (39-211) L (81-152)).

5. Calculation of the legume contribution (% of DM; LD) from the legume covered area (% of the green area of the image, i.e. image area minus bare soil area; LC) with legume specific relationships. The relationships as used for the standard image analysis based on greyscale pictures (DIA_G; Himstedt et al., 2009) were:

RCG: LD=0.838+0.745*LC Eq. [6.1]

WCG: LD=2.274+0.773*LC Eq. [6.2]

LCG: LD=2.067+0.945*LC Eq. [6.3]

The relationships between LD and LC including total sward biomass (BM) as an additional effect (Himstedt and Wachendorf 2009) were:

RCG: logit (LD)=-0.298+0.923*logit (LC)-0.002*BM Eq. [6.4]

WCG: logit (LD)=0.055+0.941*logit (LC)-0.004*BM Eq. [6.5]

LCG: logit (LD)=0.275+0.990*logit (LC)-0.003*BM Eq. [6.6]

LC was transformed to the logit-scale (e.g. defined as log[LC/(100-LC)]). The predictions of LD were back-transformed to the original scale (legume contribution in % of DM), using

$$\text{Predicted} \quad \text{LD} = \frac{100 \times e^{\text{predicted logit(LD)}}}{1 + e^{\text{predicted logit(LD)}}} \qquad \text{Eq. [6.7]}$$

For regression analysis, the GLM procedure of SAS 9.1 (SAS Institute, Cary, NC) was used.

6.3 Results

6.3.1 Detection of bare soil and legume coverage

The use of the HSL colour space allowed an accurate detection of bare soil coverage (% area) (R^2 0.99; SE 2.82) (Figure 6.1). It is particularly in young swards where large areas of bare soil are visible and where an inaccurate detection would strongly confound the prediction results. Low residuals indicate a high accuracy for bare soil area prediction across the whole range of values between 0.5 % and 69.2 %. The average value of the 64 image-specific colour thresholds ranged for hue from 18 to 45, for saturation from 43 to 115 and for lightness from 70 to 173, respectively. For an automated DIA_C procedure thresholds for H, S and L were adopted to the changed colour values after erosion and dilation with H (16-45), S (40-154) and L (41-173).

A wide range of form parameters (i.e. size, breadth, rectangularity, and circularity of areas) was tested across all sward types, but none did improve prediction accuracy of legume coverage significantly (data not shown). Contrarily to other agronomic applications where form parameter proved successfully, the shape of leaves did not differ enough between the components in mixed legume-grass swards. With the legume specific DIA based on greyscale images (DIA_G) from 49- and 63-d old swards an overall correlation between estimated and measured legume coverage (LC) was achieved of R^2 0.87, SE 8.9 (Table 6.1). However, inclusion of 35-d old swards revealed the limitations of this approach: standard errors of prediction increased by 3 - 8.5 % area and the slope of the regression lines deviated significantly from one, indicating a severe overestimation of LC at higher levels of legume coverage. Based on a legume specific DIA_C procedure estimated legume coverage (LC) showed good correlations with the measured values across the whole range of sward ages (R^2=0.96, SE 4.7) (Figure 6.2). A somewhat lower precision occurred in the lucerne-specific model with errors of up to 12 % area for pure grass swards.

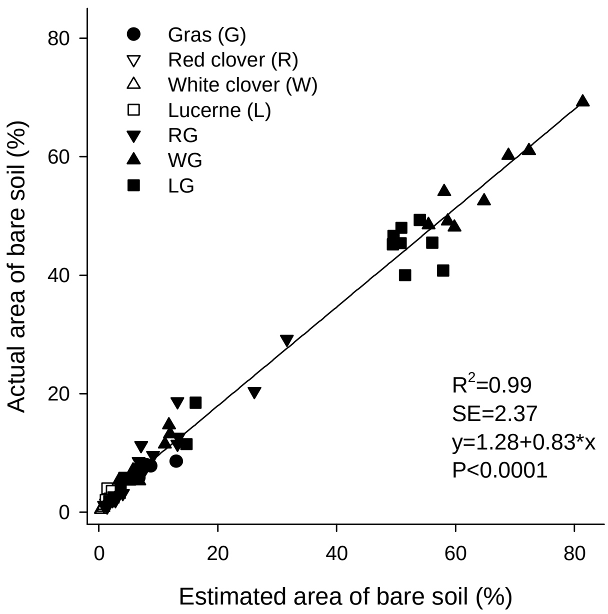


Figure 6. 1: Relationship between the actual area of bare soil and the area estimated by legume-specific digital image analysis (DIA, % area). Overall regression line includes 35-, 49-, and 63-d old swards. Slope and intercept were significant different from 1 and zero, respectively.

Compared to DIA_G inclusion of colour information improved the performance especially for swards with less than 50 g biomass m^{-2} (Figure 6.3), with only a slight underestimation. For swards > 100 g m^{-2} minor differences occurred between measured and estimated values for DIAC, whereas residuals for DIA_G remained large.

Table 6. 1: Relationship between the measured (y) and estimated (x) legume coverage (% area) and the measured (y) and estimated (x) legume contribution (% of DM), respectively, related to different digital image analysis (DIA) procedures (DIA_G greyscale images, DIA_C colour images). DIA was conducted for different datasets (RCG red clover-grass, WCG white clover-grass, LCG lucerne-grass). Every dataset includes pure swards of grass and legumes and the respective binary mixtures. Legume-specific equations were used for calculation of legume contribution (DIA_G: Eqs. 6.1-6.3 and DIA_C: Eqs. 6.4-6.6).

	Legume coverage (% area)			*Legume contribution (% of DM)*		
DIA procedure Sward age	DIA_G 49, 63-d	DIA_G 35, 49, 63-d	DIA_C 35, 49, 63-d	DIA_G 49, 63-d	DIA_G 35, 49, 63-d	DIA_C 35, 49, 63-d
RCG						
R^2	0.91	0.52[£]	0.97	0.85	0.51	0.90
RMSE	7.84	16.35	4.24	7.78	12.85	5.89
y=f(x)	y=-4.29+1.05x	y=0.73+0.73x	y=0.9+0.97x	y=-4.07+1.05x	y=-0.22+0.76x	y=2.07+0.98x
P-value	P<0.0001	P<0.0001	P<0.0001	P<0.0001	P<0.0001	P<0.0001
WCG						
R^2	0.93	0.66	0.96	0.87	0.63	0.94
RMSE	6.72	13.02	4.61	7.32	10.74	4.31
y=f(x)	y=0.75+1.01x	y=1.30+0.79x	y=0.09+0.93x	y=1.04+0.99x	y=1.90+0.78x	y=0.91+0.89x
P-value	P<0.0001	P<0.0001	P<0.0001	P<0.0001	P<0.0001	P<0.0001
LCG						
R^2	0.83	0.63	0.94	0.79	0.69	0.93
RMSE	10.05	12.95	4.91	10.73	11.36	5.52
y=f(x)	y=1.58+1.19x	y=4.23+0.93x	y=-2.12+1.04x	y=0.21+1.17x	y=2.38+0.99x	y=-3.54+1.08x
P-value	P<0.0001	P<0.0001	P<0.0001	P<0.0001	P<0.0001	P<0.0001
Overall regression						
R^2	0.87	0.55[£]	0.96	0.81	0.56[£]	0.91
RMSE	8.90	14.90	4.70	8.91	12.38	5.57
Y=f(x)	y=0.22+1.03	y=3.59+0.74x	y=0.63+0.98x	y=-1.01+1.05	y=1.41+0.82x	y=0.78+0.97x
P-value	P<0.0001	P<0.0001	P<0.0001	P<0.0001	P<0.0001	P<0.0001

Tests for intercepts were not significant

[£] Slopes are significantly different to 1.

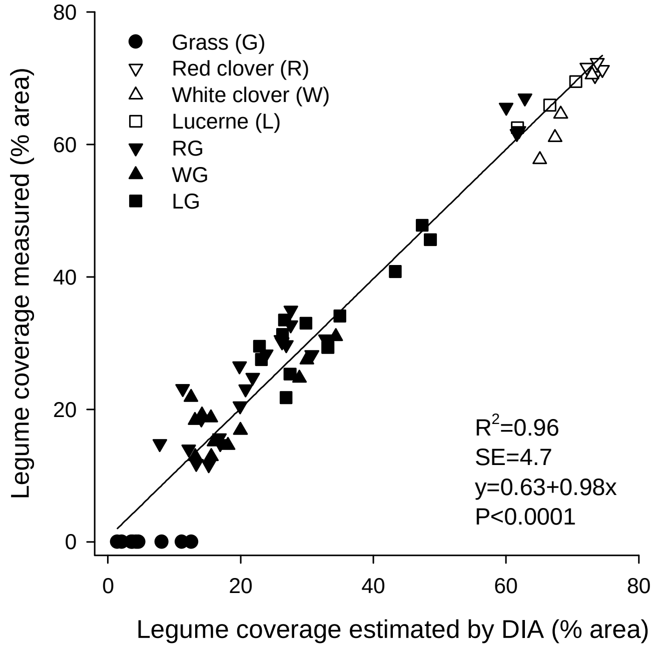


Figure 6. 2: Relationship between the measured legume coverage (% area) and the legume coverage estimated by legume-specific digital image analysis (DIA, % area). Overall regression line includes 35-, 49-, and 63-d old swards.

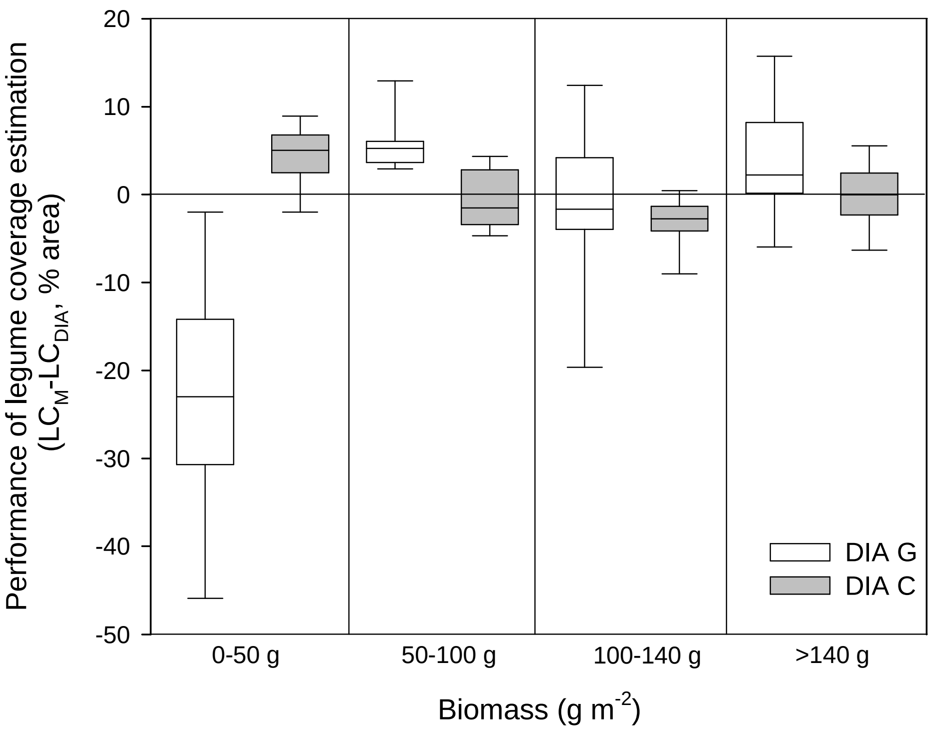


Figure 6. 3: Performance of legume coverage estimation by DIA_G (digital image analysis, greyscale images) and DIA_C (colour images). LC_M: measured legume coverage (% area). LC_{DIA}: legume coverage estimated by DIA (% area). Included are 35-, 49-, and 63-d old swards. The boundary of the box closest to zero indicates the 25th percentile, the boundary of the box farthest from zero indicate the 75th percentile. The line within the box marks the median. Error bars above and below the box indicate the 10th and 90th percentiles.

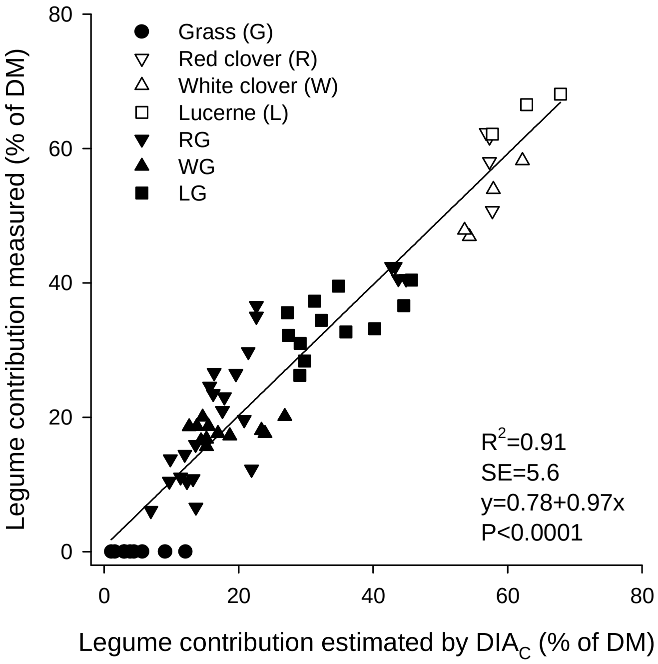


Figure 6. 4: Relationship between the measured legume contribution (% of DM) and the legume contribution estimated by legume specific digital image analysis DIA_C (colour images, % of DM). Overall regression line includes 35-, 49-, and 63-d old swards.

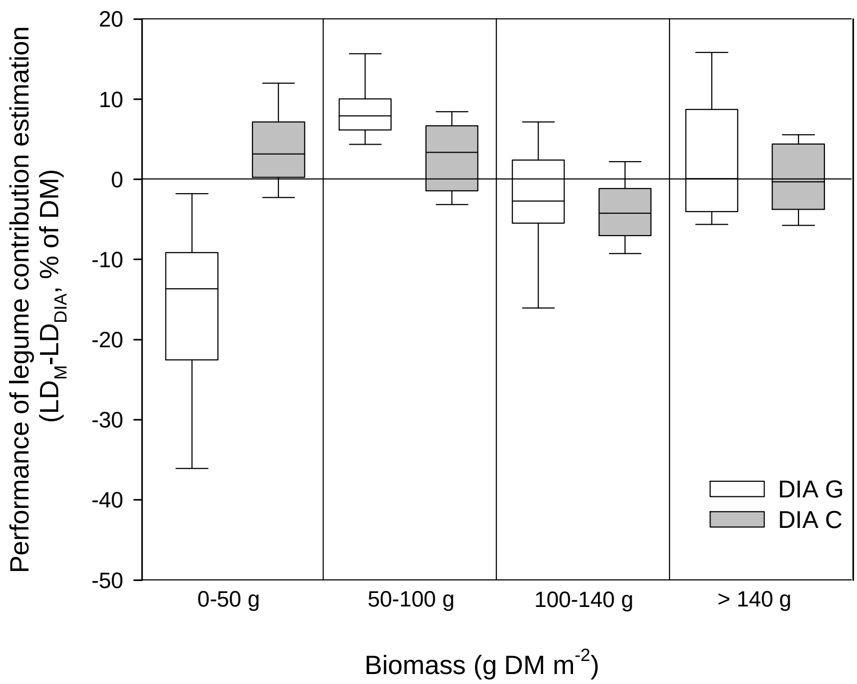


Figure 6. 5: Performance of legume contribution estimation by DIA_G (digital image analysis, greyscale images) and DIA_C (digital image analysis, colour images). LD_M: measured legume DM contribution (% of DM), LD_{DIA}: legume DM contribution based on DIA (% of DM). Included are 35-, 49-, and 63-d old swards. Legume specific equations were used for calculation of LD (DIA_G: Eqs.6.1-6.3 and DIA_C: Eqs. 6.4-6.6). The boundary of the box closest to zero indicates the 25th percentile, the boundary of the box farthest from zero indicate the 75th percentile. The line within the box marks the median. Error bars above and below the box indicate the 10th and 90th percentiles.

6.3.2 Determination of legume dry matter contribution

Legume contribution estimated by the DIA_C procedure and the legume specific equations (Eqs. 6.4 – 6.6) were closely related to the measured values (R^2 0.91, SE 5.6) (Figure 6.4), whereas somewhat higher standard errors were found for pure grass and red clover swards. For swards with a biomass of up to 100 g m^{-2} and more than 140 g m^{-2} residuals for DIA_C were closer to zero than for DIA_G (Figure 6.5). For swards with 100 to 140 g m^{-2} DIA_C showed a reduced variation, but a slight over-estimation was noticeable by both procedures.

6.4 Discussion

Inclusion of colour information into DIA significantly improved the identification of legume leaves in mixed swards. This particularly applied to young and more open swards with low biomasses, where plenty of visible bare soil occurred. The rate of misclassification of bare soil was significantly reduced by DIA_C which allows an extended application of the procedure across a wider range of swards. Furthermore, the results show that the HSL space is appropriate for the segmentation among legume and grass after the erode – dilate procedure. Bonesmo et al. (2004) suggests that, the larger the clover leaves are the higher number of erosions may be applied, common numbers of iterations in DIA_C seem to present a reasonable compromise which brought about a high accuracy of detection of legume coverage over the whole range of sward age (R^2 0.96, SE 4.7) (Figure 6.2). The underestimation of legume coverage in extremely young swards with less than 50 g m^{-2} points at the limit of the procedure, as small-sized legume leaves are occasionally eliminated by the erode - dilate procedure similar to narrow grass leaves. The fact that erosion and dilation does not necessarily work exactly in reverse (Van Droogenbroeck and Buckley, 2005) apparently does not reduce the applicability of these morphological operators in the present context.

Karcher and Richardson (2005) pointed out, that the appropriate HSL threshold settings will vary depending on factors such as turf species and variety, management practices, light conditions present when images were collected and camera model. They find it necessary to determine the precise hue and saturation levels that will select the pixel of interest within the images before executing the macro. In the present study HSL thresholds were determined with a wide scope of application across different sward ages allowing an accurate separation of green biomass and bare soil. Albeit soil and vegetation could be segregated well in young swards, DIA_C could not completely avoid misclassifications in swards with 100 – 140 g m^{-2}. Obviously, information from the HSL colour space does not allow a definite differentiation between plant species within dense canopies. However, the accuracy of +/ - 7 % is acceptable.

In the present study advanced models for the calculation of legume contribution from legume coverage were used, which were recently suggested by Himstedt and Wachendorf (2009) using the same experimental swards. These models suggest that at the same level of legume coverage, legume contribution decreases with increasing total sward biomass. This phenomenon, which occurred for all legume species, may be because more non-leguminous biomass was covered by legume leaves at high levels of total biomass than at low levels. Compared to the standard DIA_C procedure, inclusion of biomass data, increased R^2 by 0.06 (WCG 0.09, RCG 0.04, LCG 0.12) and reduced SE by 1.7 % DM (WCG 2.27, RCG 0.94, LCG 3.49) on average. Furthermore, problems with heteroscedasticity and negative predictions, which emerged from the standard models, did not occur as legume contribution and coverage were transformed to the logit-scale. The integration of total biomass into the model for determining legume contribution does not necessarily reduce its applicability in practice. Recent results from field experiments demonstrated the potential of near infrared field-spectroscopy for the prediction of total biomass of legume-grass mixtures (Biewer et al., 2009). Thus, a combined estimate of total biomass and legume coverage by field spectroscopy and DIA, respectively, may allow an accurate prediction of the legume contribution in legume-grass mixtures.

6.5 Conclusions

An advanced procedure for the determination of legume dry matter contribution by digital image analysis is suggested, which comprises the inclusion of morphological operators and colour information in the analysis of images and which applies an advanced function to predict legume dry matter contribution from legume coverage by considering total sward biomass. Bare soil areas in young and open swards could be determined very accurately which in turn allowed a precise estimation of legume coverage across a wide range from 11.8 - 72.4 %. Low residuals between measured and calculated values of legume dry matter contribution were found for the separate legume species, as well as for the combined dataset. The suggested digital image analysis procedure provides a rapid and precise estimation of legume dry matter contribution for different legume species across a wide range of sward age.

7 General discussion

The objective of this study was to evaluate, if digital image analysis (DIA) can be used to estimate the legume dry matter (DM) contribution of legume-grass swards across a wide range of legume species, legume proportions and growth stages. A pot experiment was conducted in a greenhouse under controlled conditions in order to assess the potential of digital image analysis for estimating the legume DM contribution in legume-grass swards using constant recording geometry and illumination. The DIA development has been carried out by means of 64 sward images. In addition 46 images of a field experiment were used to validate the relationship between legume coverage and legume DM contribution in legume-grass swards in order to test the applicability of DIA for practical purposes. In contrast to the pot experiment sward age of the investigated field plots showed a wider range which resulted in the inclusion of grasses and legumes with elongated stems and sward heights up to 45 cm.

7.1 The relationship between legume coverage and contribution

Information obtained by analysing a digital sward image are legume coverage related to the green area in the image. Hence, the underlying relationship between legume coverage and DM contribution is of fundamental importance for the estimation of legume DM contribution by means of digital image analysis. The used references were the measured legume contributions (% of DM) and the measured legume coverage (% of area) obtained by digital encircling the legume leaves in the images by hand.

A common relationship based on all three legumes and sward ages of 35, 49, and 63 days was found with R^2 0.90. This relationship was improved by the legume-specific approach of only 49- and 63-d old swards (R^2 0.94, 0.96, and 0.97 for red clover, white clover, and lucerne, respectively) since differing structural attributes of the legume species influence the relationship between these two parameters. The main aim of the study was to develop a DIA procedure for a wide range of swards. Furthermore, an advanced approach was the application of this model by including 35-d-old swards, which were characterised by more visible soil and low yields. Paruelo et al. (2000) described, that the relation between the percentage of green pixels and total green biomass changed during the growing season. Therefore, total biomass (BM) was included in the models. The resulting legume specific models suggest that at the same level of legume coverage, legume contribution decreases with increasing total sward biomass. This phenomenon, which occurred for all legume species, may be caused by the increasing amount of non-leguminous biomass, covered by legume leaves at high levels of total biomass. As the biomass effect differs for the legume species, legume specific models indicate high accuracies (R^2 0.94, 0.97, 0.98 with SE 4.9, 3.3, 3.3% of DM for red clover, white clover, and lucerne, respectively). For the common dataset, including all legume species, the results were similar to red clover (R^2 0.93, SE 5.16 % of DM). Subsequently, legume contribution and coverage was transformed to the logit-scale, in order to avoid problems with heteroscedasticity and negative predictions. This additionally enhanced the accuracy of the model.

The validation of the model by using digital images collected over field grown swards with biomass ranges considering the scope of the model showed promising results. The comparison of measured and calculated legume DM contribution resulted in a regression coefficient of R^2 0.98 (SE 6.0 % of DM). In this model grasses and legumes with elongated stems and sward heights up to 45 cm were included. This suggests that the model, developed in this study, is able to predict legume contribution for most common legume-grass swards (Frame, 1992; Ledgard and Steele, 1992; Loges, 1998). Nevertheless, further research is necessary in order to adjust the relationships at higher levels of total biomass, where the differences among legume species may become more prevalent.

7.2 Prospects and constraints of digital image analysis to estimate legume coverage and contribution

In a first approach greyscale images (DIA_G) were used to estimate the legume coverage in legume-grass sward images, receiving feasible results for dense swards only. Although good contrast exhibition between plant material and soil background is attested for greyscale images (Hague et al., 2006), the accuracy of the presented DIA_G decreased considerably with appearance of visible bare soil. This may be due to the different transformation from RGB colour space to greyscale images, as Hague et al. (2006) and Marchant and Onyango (2002) used a particular transformation not used in the present study. Literature often reports the use of HSL and HSI thresholds for successful segmentation among plant area and bare soil (Purcell, 2000; Karcher and Richardson, 2005; Lock et al., 2004). Therefore, this transformation into HSL colour images was desired. For DIA_C (digital image analysis, including colour information) HSL thresholds were determined with a wide scope of application across different sward ages allowing an accurate separation of green biomass and bare soil. With an ascertained HSL threshold the detection of bare soil coverage (% area) resulted in a prediction accuracy of R^2 0.99 (SE 2.82 % area).

This significantly improved the identification of legume coverage in mixed swards, especially in young and more open swards with low biomasses, where plenty of bare soil was visible. The identification of bare soil is important in order to allow for an extended application of the procedure across a wide range of swards. Furthermore, HSL thresholds are appropriate for the segmentation among legume and grass after the erode - dilate procedure. Especially, the underestimation of lucerne, as appeared using greyscale images for analysis, is improved using HSL colour thresholds. Also the misclassification of unsown species was reduced, e.g. *Stellaria media* in pure sown grass swards from maximum 20 % to maximum 12 %.

Preconditioned to the segmentation using HSL thresholds is an image processing with the morphological operators erode and dilate since these operators support the differentiation of objects of different shapes by shrinking and dilating objects (Soille, 1999). Onyango et al. (2005) and Bonesmo et al. (2004) used this capability to discriminate between plant species. The statement that dilation and erosion are not inverse operators, i.e. one may end up with a smaller area than the original (Van Droogenbroeck and Buckley, 2005), does not reduce the applicability of these morphological operators in the present context. Bonesmo et al. (2004) suggests that the larger the clover leaves, the higher is the number of erosions being applied. In the present DIA_C approach, common numbers of iteration present a reasonable compromise yielding in high accuracy of legume coverage detection over the whole range of sward ages (R^2 0.96, SE 4.7 %). The slight underestimation of legume coverage in extremely young swards with less than 50 g m^{-2} points at the limit of the procedure, as small-sized legume leaves are occasionally eliminated by the erode-dilate procedure similar to narrow grass leaves. In this context it is important to keep the distance between sward surface and camera at 80 cm constantly to ensure equal conditions.

The best results were obtained with individual predictions of red clover, white clover and lucerne, respectively. Different DIA_C procedures allow for different sizes, shapes and colours of the legume leaves. As red clover leaf sizes exceed white clover the DIA_C for red clover includes three iterations of the erode-dilate procedure. The lanceolate leaf shape of lucerne allows for only two iterations just like for white clover. The HSL thresholds vary for the different legumes as proposed by Karcher and Richardson (2005). In addition, they propose that an adjustment of the HSL thresholds is also necessary depending on factors such as management practices, light conditions and camera model. In the present pot experiment the defined HSL thresholds were applicable for the whole range of sward ages without further adjustments. For the implementation in practice on field scale the need for an HSL threshold adjustment has to be clarified, concerning both the identification of legumes and the separation of plant tissue and bare soil, since soil colours are locally different.

With the enhanced relationship between coverage and contribution and the improved estimation of legume coverage, digital image analysis is an appropriate tool to estimate legume DM contribution in swards.

The inclusion of HSL colour information and total biomass into DIA procedure resulted in a prediction accuracy for legume DM contribution of R^2 0.90, 0.94 and 0.93 with SE 5.89, 4.31, 5.52 % of DM for red clover, white clover and lucerne swards, respectively.

All parts of the procedure, detection of bare soil, the estimation of coverage, and the calculation of contribution, have default settings in store which can be adjusted to practical application. Additionally, the estimation of bare soil and legume coverage (% area) can be applied individually, if necessary.

The direct assessment of mixed forage swards in the field would be a major advance in an efficient and environmentally friendly management of legume-based farming systems. In low-input production systems, like organic agriculture, a synchronized determination of total yield and legume proportion by appropriate sensors in the field would allow for a more accurate prediction of the amount of legume-derived nitrogen in soil in the pasture phase. In a parallel study examining the same experimental swards, field spectroscopy proved to be promising for the detection of swards biomass at all growth stages (Biewer et al., 2009). Thus, the assessment of legume dry matter yields by the combination of DIA and field spectroscopy could help to adjust the nitrogen management in arable systems.

8 Conclusions

The following conclusions can be drawn from the greenhouse experiment and the field study with pure stands and binary mixtures of different forage legumes and perennial ryegrass:

iv) A relationship between legume coverage and legume DM contribution can be ascertained. The model with included biomass performed strongly, with the best relationships for the legume specific approaches (R^2 0.94, 0.97, 0.98 with SE 4.9, 3.3, 3.3% of DM for red clover, white clover and lucerne, respectively).

v) To avoid problems with heteroscedasticity and negative predictions legume contribution and coverage was transformed to the logit-scale. This additionally enhanced the accuracy of the model.

vi) The validation of the model on swards of the field experiment with biomass ranges considering the scope of the model showed promising results. Included were grasses and legumes with elongated stems and sward heights up to 45 cm. This shows, that the model can predict legume contribution for most practical legume-grass swards.

vii) The estimation of legume coverage using digital image analysis is feasible. HSL colour images are most suitable to discriminate between soil and plant tissue using threshold segmentation.

viii) The morphological operators erode and dilate support the differentiation of objects of different shape by shrinking and dilating objects. When applied to images of legume-grass mixtures thin grass leaves were removed whereas rounder clover leaves were left. With the subsequent legume specific HSL threshold the legume coverage can be estimated.

ix) With the inclusion of HSL colour information and biomass into the estimation of coverage and the calculation of contribution, respectively, the prediction accuracy for legume DM contribution was R^2 0.90, 0.94 and 0.93 with SE 5.89, 4.31, 5.52 % of DM for red clover, white clover and lucerne swards, respectively.

Further research is needed in order to evaluate the accuracy of the whole DIA procedure on swards of a field experiment on varying sites and vegetation periods to enhance the robustness of the models.

The integration of total biomass into the model for determining legume contribution does not necessarily reduce its applicability in practice. Particularly as for the calculation of legume DM yields it is also needed. A combination of total biomass and legume coverage estimation by field spectroscopy and DIA, respectively, may allow an accurate prediction of the legume contribution and legume dry matter yield of legume-grass mixtures.

9 Summary

Productivity and forage quality of legume-grass swards are important factors for successful arable farming in both organic and conventional farming systems. For these objectives the botanical composition of the swards is of particular importance, especially, the content of legumes due to their ability to fix airborne nitrogen. As it can vary considerably within a field, a non-destructive detection method while doing other tasks would facilitate a more targeted sward management and could predict the nitrogen supply of the soil for the subsequent crop.

This study was undertaken to explore the potential of digital image analysis (DIA) for a non destructive prediction of legume dry matter (DM) contribution of legume-grass mixtures. For this purpose an experiment was conducted in a greenhouse, comprising a sample size of 64 experimental swards such as pure swards of red clover (*Trifolium pratense* L.), white clover (*Trifolium repens* L.) and lucerne (*Medicago sativa* L.) as well as binary mixtures of each legume with perennial ryegrass (*Lolium perenne* L.). Growth stages ranged from tillering to heading and the proportion of legumes from 0 to 80 %.

Based on digital sward images three steps were considered in order to estimate the legume contribution (% of DM):

i) The development of a digital image analysis (DIA) procedure in order to estimate legume coverage (% of area).

ii) The description of the relationship between legume coverage (% area) and legume contribution (% of DM) derived from digital analysis of legume coverage related to the green area in a digital image.

iii) The estimation of the legume DM contribution with the findings of i) and ii).

i) In order to evaluate the most suitable approach for the estimation of legume coverage by means of DIA different tools were tested. Morphological operators

such as erode and dilate support the differentiation of objects of different shape by shrinking and dilating objects (Soille, 1999). When applied to digital images of legume-grass mixtures thin grass leaves were removed whereas rounder clover leaves were left. After this process legume leaves were identified by threshold segmentation. The segmentation of greyscale images turned out to be not applicable since the segmentation between legumes and bare soil failed. The advanced procedure comprising morphological operators and HSL colour information could determine bare soil areas in young and open swards very accurately. Also legume specific HSL thresholds allowed for precise estimations of legume coverage across a wide range from 11.8 - 72.4 %. Based on this legume specific DIA procedure estimated legume coverage showed good correlations with the measured values across the whole range of sward ages (R^2 0.96, SE 4.7 %). A wide range of form parameters (i.e. size, breadth, rectangularity, and circularity of areas) was tested across all sward types, but none did improve prediction accuracy of legume coverage significantly.

ii) Using measured reference data of legume coverage and contribution, in a first approach a common relationship based on all three legumes and sward ages of 35, 49 and 63 days was found with R^2 0.90. This relationship was improved by a legume-specific approach of only 49- and 63-d old swards (R^2 0.94, 0.96 and 0.97 for red clover, white clover, and lucerne, respectively) since differing structural attributes of the legume species influence the relationship between these two parameters. In a second approach biomass was included in the model in order to allow for different structures of swards of different ages. Hence, a model was developed, providing a close look on the relationship between legume coverage in binary legume-ryegrass communities and the legume contribution: At the same level of legume coverage, legume contribution decreased with increased total biomass. This phenomenon may be caused by more non-leguminous biomass covered by legume leaves at high levels of total biomass. Additionally, values of legume contribution and coverage were transformed to the logit-scale in order to avoid problems with heteroscedasticity and negative predictions. The resulting relationships between the measured legume contribution and the calculated legume contribution indicated a high model accuracy for all legume species (R^2

0.93, 0.97, 0.98 with SE 4.81, 3.22, 3.07 % of DM for red clover, white clover, and lucerne swards, respectively). The validation of the model by using digital images collected over field grown swards with biomass ranges considering the scope of the model shows, that the model is able to predict legume contribution for most common legume-grass swards (Frame, 1992; Ledgard and Steele, 1992; Loges, 1998).

iii) An advanced procedure for the determination of legume DM contribution by DIA is suggested, which comprises the inclusion of morphological operators and HSL colour information in the analysis of images and which applies an advanced function to predict legume DM contribution from legume coverage by considering total sward biomass. Low residuals between measured and calculated values of legume dry matter contribution were found for the separate legume species (R^2 0.90, 0.94, 0.93 with SE 5.89, 4.31, 5.52 % of DM for red clover, white clover, and lucerne swards, respectively).

The introduced DIA procedure provides a rapid and precise estimation of legume DM contribution for different legume species across a wide range of sward ages. Further research is needed in order to adapt the procedure to field scale, dealing with differing light effects and potentially higher swards.

The integration of total biomass into the model for determining legume contribution does not necessarily reduce its applicability in practice as a combined estimation of total biomass and legume coverage by field spectroscopy (Biewer et al. 2009) and DIA, respectively, may allow for an accurate prediction of the legume contribution in legume-grass mixtures.

10 Zusammenfassung

Sowohl in der ökologischen als auch in der konventionellen Landwirtschaft sind Produktivität, Bestandeszusammensetzung und Futterqualität von Leguminosengras-Beständen wichtige Parameter für einen erfolgreichen Feldfutterbau. Diese können jedoch innerhalb eines Feldes beachtlichen Schwankungen unterworfen sein, so dass eine nicht destruktive Erfassung der Bestandszusammensetzung während der Feldarbeit ein verbessertes Management der Bestände sowie der Düngung ermöglichen würde. In diesem Zusammenhang sind die Leguminosen durch ihre Fähigkeit Luftstickstoff zu binden von besonderem Interesse. Ist zum Beispiel der Leguminosen-Trockenmasse (TM)-Ertrag bekannt, wird eine Vorhersage des Stickstoffangebots für die Nachfrucht ermöglicht (Høgh-Jensen et al., 2004).

Die vorliegende Untersuchung evaluiert das Potenzial der Bildanalyse für die Erfassung des Leguminosen-Ertragsanteils in Leguminosengras-Beständen anhand eines Gewächshausversuches. Der Gewächshausversuch hatte einen Probenumfang von 64 Leguminosengras-Beständen. Untersucht wurden Reinsaaten und binäre Leguminosengras-Gemenge aus Rotklee (Trifolium pratense L.), Weißklee (Trifolium repens L.), Luzerne (Medicago sativa L.) und Deutschem Weidelgras (Lolium perenne L.) im Alter von 35, 49 und 63 Tagen. Der Anteil der Leguminosen in den Pflanzenbeständen schwankte zwischen 0 und 80%. Für die Validierung eines Aspektes der Analyse wurden 46 Bilder von Freiland-Beständen mit gleichen Leguminosen und ähnlichen Biomassen verwendet.

Da anhand digitaler Bilder von Beständen nur der Leguminosen-Deckungsgrad ermittelt werden kann, wurde im ersten Schritt ein Zusammenhang zwischen Leguminosen-Deckungsgrad (% Fläche) und Leguminosen-Ertragsanteil (% der Trockenmasse [TM]) ermittelt. Für eine multiple Regressionsanalyse wurden alle Altersstufen mit einbezogen und für Ertragsanteile und Deckungsgrade der Leguminosen (%) eine Logit-Transformation verwendet, da frühere Untersuchungen zeigten, dass Probleme der Relativzahlen (Varianzinhomogenität und negative Schätzwerte) dadurch vermieden werden können (Connolly and Wachendorf, 2001). Es hat sich gezeigt, dass die Einbeziehung der Biomasse notwendig ist, um den Einsatz für unterschiedlich entwickelte Bestände zu ermöglichen. Durch dieses Model kann ein grundsätzlicher Zusammenhang von Leguminosen-Deckungsgrad und -Ertragsanteil in Leguminosengras-Beständen beschrieben werden: Bei gleich bleibendem Leguminosen-Deckungsgrad sinkt der Leguminosen-Ertragsanteil mit steigender Biomasse. Der Anteil der Nicht-Leguminosen Biomasse, welcher durch die Leguminosen verdeckt wird, scheint mit der Biomasse zuzunehmen. Dieser Einfluss der Biomasse ist für die einzelnen Leguminosenarten unterschiedlich, so zeigen spezifische Berechnungen für die einzelnen Leguminosenarten die besten Ergebnisse. Die Beziehung zwischen den gemessenen und den ermittelten (anhand des Models und gemessener Deckungsgrade) Leguminosen-Ertragsanteilen zeigt eine hohe Güte (R^2 0.93, 0.97, 0.98 mit SE 4.81, 3.22, 3.04 jeweils für Rotklee, Weißklee und Luzerne). Die Validation des Modells anhand von Beständen eines Freilandversuches mit ähnlichen Biomassen (bis 28 dt ha^{-1}) zeigte, dass Ertragsanteile für die meisten praxisnahen Bestände abgeschätzt werden können.

Die Abschätzung des Leguminosen-Deckungsgrades mittels digitaler Bildanalyse konnte mit der höchsten Genauigkeit anhand von HSL Farbbildern (Hue, Saturation, Lightness) durchgeführt werden. Anhand von HSL-Schwellenwerten ist eine Trennung von Boden und Grünfläche möglich, was Voraussetzung ist für eine gute Abschätzung der Leguminosen-Fläche in Prozent der Grünfläche. Eine Trennung von Gras und Leguminosen wird aufgrund der unterschiedlichen Blattformen durch die Anwendung der morhologischen Operatoren Erode und Dilate ermöglicht. Durch die Erosion werden die schmalen Grasblätter so sehr geschrumpft (verdunkelt), dass kein heller Kern in der Mitte verbleibt um bei der folgenden Dilatation (Ausdehnung) wieder sichtbar zu werden. Die runderen Leguminosenflächen bewahren nach der Erosion einen hellen Kern und können so durch Dilatation wieder ausgedehnt werden. Durch die gleiche Anzahl von Erosion und Dilatation bleiben die ursprünglichen Flächengrößen der Leguminosen erhalten. Nach dieser Prozedur ist eine Trennung von Leguminosen- und Grasflächen aufgrund von HSL-Schwellenwerten möglich. Formparameter, wie Rundheit und Rechtwinkligkeit der Flächen, wurden getestet, konnten die Güte aber nicht verbessern. Anhand von leguminosenspezifischen Schwellenwerten kann mit der entwickelten Bildanalyse der Leguminosen-Deckungsgrad mit hoher Güte geschätzt werden (R^2 0.96, SE 4.7).

Anhand der bildanalytisch geschätzten Deckungsgrade und der ermittelten Beziehung (jeweils leguminosenspezifisch) kann der Leguminosen-Ertragsanteil berechnet werden. Das Ergebnis ist eine bildanalytische Ermittlung von Leguminosen-Ertragsanteilen mit einem hohen Bestimmtheitsmaß und vertretbaren Standardabweichungen (R^2 0.90, 0.94, 0.93 mit SE 5.89, 4.31, 5.52 % TM jeweils für Rotklee-, Weißklee- und Luzerne-Bestände).

Für eine Anwendung in der Praxis bleibt zu untersuchen, in wieweit die Bildanalyse an andere Lichtverhältnisse, unterschiedliche Bodenfarben, höhere und eventuell auch blühende Bestände angepasst werden muss. Eine Erfassung der Biomasse der Bestände ist notwendig, da dieser Parameter für die Umrechnung von Deckungsgrad zu Ertragsanteil benötigt wird. Aber auch für eine Berechnung des Leguminosen-TM-Ertrags zur Kalkulierung des eingebrachten Luftstickstoffs ist die Biomasse notwendig. Hierfür könnten feldspektroskopische Methoden in Frage kommen, die von Biewer et al. (2009) für die Biomasse Abschätzung von Leguminosengras-Beständen erfolgreich getestet wurden.

11 References

Biewer, S., S. Erasmi, T. Fricke, and M. Wachendorf. 2008. Prediction of yield and the contribution of legumes in legume-grass mixtures using field spectrometry. Precision Agriculture 10:128-144.

Biewer, S., T. Fricke, and M. Wachendorf. 2009. Determination of dry matter yield from legume-grass swards by field spectroscopy. Crop Science 49,1927-1936.

Boller, B.C., and J. Nösberger. 1987. Symbiotically fixed nitrogen from field-grown white and red clover mixed with ryegrasses at low levels of ^{15}N-fertilization. Plant and Soil 104:219-226.

Bonesmo, H., K. Kaspersen, and A. Bakken. 2004. Evaluating an Image Analysis System for Mapping White Clover Pastures. Acta Agriculturae Scandinavica, Section B - Soil and Plant Science 54 (2): 76-82.

Brox, T. 2005. From pixels to regions: Partial differential equations in image analysis (PhD thesis). Saarland University, Saarbrücken.

Camargo, A., and J.S. Smith. 2009. An image-processing based algorithm to automatically identify plant disease visual symptoms. Biosystems Engineering 102: 9-12.

Carlsson, G., and K. Huss-Danell. 2003. Nitrogen fixation in perennial forage legumes in the field. Plant and Soil 253: 353-372.

Cheng, H., and Y. Sun. 2000. A hierarchical approach to colour image segmentation using homogeneity. IEEE Transactions on image processing 9: 2071-2082.

Connolly, J., and M. Wachendorf. 2001. Developing multisite dynamic models of mixed species plant communities. Ann. Bot. 88:703-712.

Cox, D.R., and E.J. Snell. 1989. The analysis of binary data, 2nd edition. London. Chapman & Hall.

Delon, J., A. Desolneux, J. Lisani, and A. Petro .2005. Colour Image Segmentation Using Acceptable Histogramm Segmentation. Pattern Recognition and Image Analysis Part 2: 239-246

Draper, N.R., and H. Smith. 1998. Applied regression analysis, 3rd edn. New York. John Wiley & Sons.

Frame, J., R.D. Harkness, and I.V. Hunt. 1985. Effect of seed rate of red clover and of companion timothy or tall fescue on herbage production. Grass and Forage Science 40:459-465.

Frame, J. 1992. Improved grassland management. Farming Press Books. 234. Ipswich, United Kingdom.

Fransen, B., M. de Boer, M. Terlou, H. During, and W. Dijkman. 1998. Using image-analysis to quantify the horizontal vegetation pattern in two multi-species savanna grasslands. Plant Ecology 138: 231-237.

Gebhardt, S., J. Schellberg, R. Lock, and W. Kühbauch. 2006. Identification of broad-leaved dock (Rumex obtusifolius L.) on grassland by means of digital image processing. Precision Agriculture 7: 165-178.

Gebhardt, S. and W. Kühbauch. 2007. A new algorithm for automatic Rumex obtusifolius detection in digital images using colour and texture features and the influence of image resolution. Precision Agriculture 8: 1-13.

Gerhards, R., M. Sökefeld, C. Timmermann, W. Kühbauch, and M. Williams II. 2002. Site-Specific Weed Control in Maize, Sugar Beet, Winter Wheat and Winter Barley. Precision Agriculture 3: 25-35.

Gottardi, S. 2008. Flächenhafte und funktionale Analyse kleinräumiger standort-bedingter Variationen der Grünlandvegetation und -produktivität. PhD. Thesis. Univ. of Kassel, Germany.

Haberäcker, P. 1987. Digitale Bildverarbeitung – Grundlagen und Anwendungen. 2. Aufl. Hanser, München; Wien.

Heichel, G.H., and K.I. Henjum. 1991. Dinitrogen fixation, nitrogen transfer and productivity of forage legume-grass communities. Crop Science 31:202-208.

Himstedt. M., T. Fricke, and M. Wachendorf. 2008. The application of image analysis to estimate legume contents in legume - grass swards. Grassland Science in Europe 13.

Himstedt, M., T. Fricke, and M. Wachendorf.2009. Determining the contribution of legumes in legume-grass mixtures using digital image analysis. Crop Science 49: 1910-1916.

Himstedt, M., and M. Wachendorf. 2009. Der Zusammenhang von Leguminosen-Deckungsgrad und –Ertragsanteil als Grundlage für bildanalytische Ertragsanteilschätzungen. – The relationship between legume coverage and contribution as basis for the estimation of legume contribution by digital image analysis. Mitteilungen der Gesellschaft für Pflanzenbauwissenschaften 21:135-136.

Hague, T., N. Tillett, and H. Wheeler, H. 2006. Automated crop and weed monitoring in widely spaced cereals. Precision Agriculture 7: 21-32.

Høgh-Jensen, H., R. Loges, F. Jorgensen, F. Vinther, and E. Jensen. 2004. An empirical model for quantification of symbiontic nitrogen fixation in grass-clover mixtures. Agricultural Systems 82: 181-194.

Jensen, J.R. 2000. Remote sensing of the environment: An earth resource perspective. Prentice Hall, Upper Saddle River, New Jersey, USA.

Karcher, E., and M. Richardson. 2005. Batch Analysis of Digital Images to Evaluate Turfgrass Characteristics. Crop Science 45: 1536-1539.

Lee, W., D.C. Slaughter, and D. Giles. 1999. Robotic Weed Control System for Tomatoes. Precision Agriculture 1: 95-113.

Lock, R., I. Rademacher, H. Nonn, and W. Kühbauch. 2004. Methods of digital image processing to quantify ground cover of turf grass. Grassland Science in Europe 9: 790-792.

Ledgard, S. and K. Steele. 1992. Biological nitrogen fixation in mixed legume/ grass pastures. Plant and Soil 141: 137-153.

Loges, R. 1998. Yield, forage quality, nitrogen fixation and succeeding crop effect of red clover and red clover/ grass swards. (in German.) PhD thesis. Christian-Albrechts-University of Kiel, Germany.

Loges, R., A. Kaske, K. Ingwersen, and F. Taube. 2000. Yield, forage quality, residue nitrogen and nitrogen fixation of different forage legumes. IFOAM 2000-The World Grows Organic, Proceedings of the 13th International IFOAM Scientific Conference: 83.

Malhi, S., R. Zentner, and K. Heier. 2002. Effectiveness of alfalfa in reducing fertilizer N input for optimum forage yield, protein concentration, returns and energy performance of bromegrass-alfalfa mixtures. Nutrient Cycling in Agroecosystems 62: 219-227.

Marchant, J.A., and C.M. Onyango.2002. Shadow-invariant classification for scenes illuminated by daylight. Journal of Optical Society of America A 17 (11): 1952-1961.

Media Cybernetics. 1999. The Imaging Experts Optimas 6.5, From Images to Answers. TM User Guide and Technical Reference, 9. edition, 548. Silver Spring.

Nelder, J.A.1994. The statistics of linear models: back to the basics. Statistics and Computing 4: 221-234.

Nelder, J.A., and P.W. Lane. 1995. The computer analysis of factorial experiments. In memoriam - Frank Yates. The American Statistician 49: 382-385.

Neuendorff, J. 1996. White clover (Trifolium repens L.) and its contribution to yield production of grassland considering methods to quantify nitrogen fixation. (in German.) PhD thesis. University of Kassel.

Nordmeyer, H. 2006. Patchy weed distribution and site-specific weed control in winter cereals. Precision Agriculture 7: 219-231.

Numata, I., D.A. Roberts, O.A. Chadwick, J. Schimel, F.R. Sampaio, F.C. Leonidas and J.V. Soares. 2007. Characterization of pasture biophysical properties and the impact of grazing intensity using remotely sensed data. Remote Sensing of Environment 109: 314-327.

Onyango, C., J. Marchant, A. Grundy, K. Phelps, and R. Reader. 2005. Image Processing Performance Assessment Usind Crop Weed Competition Models. Precision Agriculture 6: 183-192.

Paruelo, J., W. Lauenroth, and P. Roset. 2000. Estimating aboveground plant biomass using a photographic technique. Journal of Range Management 53(2): 190-193.

Petry, W., and W. Kühbauch. 1989. Automatisierte Unterscheidung von Unkrautarten nach Formparametern mit Hilfe der quantitativen Bildanalyse - Automatic distinction of weed species using form parameters by means of digital image processing. Journal of Agronomy and Crop Science 163: 345-351.

Purcell, L. 2000. Soybean canopy coverage and light interception measurements using digital imagery. Crop Science 40: 834-837.

Richardson, M., D. Karcher, and L. Purcell. 2001. Quantifying Turfgrass Cover Using Digital Image Analysis. Crop Science 41: 1884-1888.

Rothmaler, W., M. Bäßler, E. Jäger, and K. Werner (ed.) 1996. Excursion flora of Germany. Gustav Fischer Verlag. Jena.

SAS Institute. 2002–2003. SAS version 9.1. Cary, NC, USA: SAS Institute.

Schmidtke, K. 1997. Einfluss von Rotklee (*Trifolium pratense* L.) in Reinsaat und Gemenge mit Poaceen auf symbiontische N_2-Fixierung, bodenbürtige N-Aufnahme und $CaCl_2$-extrahierbare N-Fraktionen im Boden. PhD. thesis Univ. of Gießen; Eigenverlag Schmidtke, Germany.

Schut, A., and J. Ketelaars. 2003a. Monitoring grass swards using imaging spectroscopy. Grass and Forage Science 58: 276-286

Schut, A., and J. Ketelaars. 2003b. Assessment of seasonal dry-matter yield and quality of grass swards with imaging spectroscopy. Grass and Forage Science 58: 385-396

Schwinning, S. and A.J. Parsons. 1996. A spatially explicit population model of stoloniferous N-fixing legumes in mixed pasture with grass. Journal of Ecology 84: 815-826.

Schino, G., F. Borfecchia, L. De Cecco, C. Dibari, M. Iannetta, S. Martini and F. Pedrotti. 2003. Satellite estimate of grass biomass in a mountainous range in central Italy. Agroforestry Systems 59: 157-162.

Sleugh, B., K.J. Moore, J.R. George and E.C. Brummer. 2000. Binary legume-grass mixtures improve forage yield, quality, and seasonal distribution. Agronomy Journal 92: 24-29.

Sökefeld, M., R. Gerhards, H. Oebel, and R. Therburg. 2007. Image acquisition for weed detection and identification by digital image analysis. *In:* Stafford, J.V. (ed.) Precision Agriculture: 523-528. Wageningen Academic Publishers.

Sökefeld, M., R. Gerhards, and W. Kühbauch. 2000. Site-specific weed control-from weed recording to herbicide application. Journal of Plant Diseases and Protection, Special Issue XVII: 227-233

Soille, P. 1999. Morphological image analysis. Springer, Berlin, Germany.

Van Droogenbroeck, M., and M.J. Buckley. 2005. Morphological Erosions and Openings: Fast Algorithms Based on Anchors. Journal of Mathematical Imaging and Vision, 22: 121-142.

Vinther, F. 2006. Effects of cutting frequency on plant production, N-uptake and N2 fixation in above- and below-ground plant biomass of perennial ryegrass-white clover swards. Grass and Forage Science 61: 154-163.

Wachendorf, M., M. Büchter, H. Trott, and F. Taube. 2004. Performance and environmental effects of forage production on sandy soils. II. Impact of defoliation system and nitrogen input on nitrate leaching losses. Grass and Forage Science 59: 56-68.

Weisa, M., and R. Gerhards. 2007. Feature extraction for the identification of weed species in digital images for purpose of site-specific weed control. *In:* Stafford, J.V. (ed.) Precision Agriculture: 537-544. Wageningen Academic Publishers.

Weißbach, F. 1995: Über die Schätzung des Beitrages der symbiontischen N_2-Fixierung durch Weißklee zur Stickstoffbilanz von Grünlandflächen. Landbauforschung Völkenrode, 45 (2): 67-74.