INTRODUCTION

Grasslands deliver important economic and ecological services. They provide a substantial fraction of the ruminant feed requirements, prevent soil erosion, build soil fertility, require minimal pesticide use and are important for carbon sequestration (Lugato, Montanarella, & Jones, 2015; O’Mara, 2012; Wilkins & Humphreys, 2003). In Northwest Europe, Lolium perenne L. (perennial ryegrass, Lp) is one of the most frequently sown forage grass species because of its rapid establishment and growth, high yield potential and excellent feed quality (Humphreys, Feuerstein, Vandewalle, & Baert, 2010). These valuable (agro) services must be safeguarded.

Canopy height measurements and non-destructive biomass estimation of Lolium perenne swards using UAV imagery

Irene Borra-Serrano1,2 | Tom De Swaef1 | Hilde Muylle1 | David Nuyttens3 | Jürgen Vangeyte3 | Koen Mertens3 | Wouter Saeys4 | Ben Somers2 | Isabel Roldán-Ruiz1,5 | Peter Lootens1

1Plant Sciences Unit, Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Melle, Belgium
2Division of Forest, Nature and Landscape, KU Leuven, Leuven, Belgium
3Technology and Food Science Unit, Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Merelbeke, Belgium
4Department of Biosystems, MeBios, KU Leuven, Leuven, Belgium
5Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium

Correspondence
Peter Lootens, Plant Sciences Unit, Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Caritasstraat 39, 9090 Melle, Belgium.
Email: peter.lootens@ilvo.vlaanderen.be

Abstract

In perennial ryegrass breeding programmes, dry-matter yield (DMY) of individual plots is monitored destructively at the different cuts or derived from non-destructive canopy height measurements using devices like rising plate meters (RPM). These approaches both have constraints. Destructive sampling implies low temporal resolution, restraining the study of dry-matter accumulation rates, while RPM measurements are influenced by the canopy structure and limit intra-field variability identification. We present a phenotyping methodology, based on the use of an affordable RGB camera mounted on an unmanned aerial vehicle (UAV), to monitor the spatial and temporal evolution of canopy height and to estimate DMY. Weekly flights were carried out from April to October above a field comprising a diverse set of accessions. To test the capacity of UAV imagery to estimate canopy height, 8 ground control points and 28 artificial height references were placed at different locations. Accurate flights with an RMSE as low as 0.94 cm were achieved. In addition, canopy height was recorded using an RPM and destructive biomass samples were collected. Different models (linear, multiple linear, principal components, partial least squares regression and random forest) were used to predict DMY, and their performance was evaluated. The best estimations were obtained by combining variables including canopy height, vegetation indices and environmental data in a multiple linear regression ($R^2 = .81$). All models built using UAV data obtained a lower RMSE than the one using RPM data. The approach presented is a possibility for breeders to incorporate new information in their selection process.

KEYWORDS

biomass, drone, dry-matter yield, forage, RGB sensor
Breeding goals in perennial ryegrass are therefore to improve the yield of cultivated grasslands as well as their resilience to abiotic and biotic stresses and sustained production of high-quality forage over multiple years. In the last decades, an average yearly improvement of 0.3%–0.4% of dry-matter yield (DMY) has been achieved in this crop (Chaves, Vliegher, Waes, Carlier, & Marynissen, 2009; Laidig, Piepho, Drobek, & Meyer, 2014; McDonagh, O’Donovan, McEvoy, & Gilliland, 2016) while for other crops an average annual gain of 0.8%–1.2% is reached (Li, Rasheed, Hickey, & He, 2018). These values are still far from the estimated necessary annual gain of 2% in crop productivity that are required to achieve the food and bioresource demand for the projected global population in 2050 (Li et al., 2018). Genomics-enabled breeding approaches in combination with precision phenotyping tools are expected to contribute substantially to enhance the prediction accuracy and to accelerate genetic gains by shortening breeding cycles (Crossa et al., 2017). Indeed, the use of precision phenotyping tools to obtain a detailed and accurate characterization of the genetic value of an accession can result in a higher selection accuracy and a higher gain per breeding cycle (Cobb, DeClerck, Greenberg, Clark, & McCouch, 2013). In a recent review by Gebremedhin, Badenhorst, Wang, Spangenberg, and Smith (2019), phenotyping is presented as the bottleneck to implement genomic strategies in forage species. Furthermore, the adoption of precision phenotyping tools by breeders is frequently lagging behind scientific-technological advances in this area due to the lack of versatile and affordable approaches to screen large collections of breeding strains in a high-throughput manner (Araus & Cairns, 2014; Araus, Kefauver, Zaman-Allah, Olsen, & Cairns, 2018; Desta & Ortiz, 2014).

Dry-matter yield (DMY) is one of the most important traits in forage grass breeding (Barre et al., 2017). Biomass yield data can be obtained by visual assessments, measurements of canopy height (on its own or in combination with estimates of ground cover) or from destructive samplings. Visual assessments are carried out by experienced operators. Height measurements are non-destructive and can be done using rulers or devices such as rising plate meters (RPM), which measure plant height and canopy density. When using an RPM, the grass canopy is (gently) compressed by the plate weight, and the device measures the distance between the plate and the ground. This value is then converted into a value of DMY using a linear relationship (Holshof, Stienezen, & Galama, 2015). Destructive measurements involve cutting, drying and weighing of aboveground samples and are related to sward management (typically 4–5 times per year depending on management and the characteristics of the growing season). Despite their widespread use, all three approaches present constraints. Visual assessments are relatively cheap but are subjected to subjectivity of the operator. In addition, they can perform poorly when differences among the accessions are limited. Destructive measurements have limited temporal resolution. Using RPM, a high temporal resolution can be achieved but RPM measurements are not accurate in sparse swards or in tall, non-uniform canopies (Viljanen et al., 2018). In addition, with the exception of RPM when several measurements are carried out in the same plot, none of the three approaches are suitable to ascertain intra-plot spatial variability in an accurate manner, rendering a low spatial resolution.

As summarized by Ali, Cawkwell, Dwyer, Barrett, and Green (2016), remote sensing from satellites equipped with optical, radar detection and ranging (radar) or light detection and ranging (LiDAR) sensors can be used to estimate temperate grassland crop biomass (Wachendorf, Fricke, & Möckel, 2017). These methods are non-invasive and address some of the issues encountered by the ground-based monitoring approaches described in the previous paragraph, such as the labour intensity, the temporal limitations (and also spatial, depending on the required resolution) and human bias (Ali et al., 2016). However, satellite remote sensing has limitations regarding grassland monitoring and management, such as (a) imagery acquisition and delivery in a timely manner, (b) low spatial resolution (ranging from 5–100 m depending on the spectral band, sensor and altitude of the orbit), (c) image interpretation and data extraction and (d) dependence on weather conditions (cloud cover obstructing free sight) (Zhang & Kovacs, 2012). Although these limitations of satellite remote sensing have been formulated in the context of grassland management and monitoring, they are also applicable in a breeding context, especially if we consider the typical small plot sizes handled in breeding programmes (on the order of 1-10 m²). There are several options of phenotypic platforms, and the choice is dependent on the scale of the work and sensor chosen (Araus et al., 2018). For the application considered in this manuscript, close remote sensing using unmanned aerial vehicles (UAV) was the more appropriate choice.

Indeed, sensors mounted on a UAV allow flights to be scheduled in function of the particular needs of the breeding programme considered and they are relatively easy to perform. It is possible to acquire data throughout the whole growing season to obtain a time series of images at high temporal resolution, delivering dynamic information of crop growth and status (Zhang & Kovacs, 2012), while allowing the accurate screening of spatial heterogeneity in even relatively small plots. UAVs equipped with a high-resolution consumer digital camera can capture the visible part of the reflected electromagnetic energy in a red, green and blue (RGB) band via imagery with cm resolution (Bendig et al., 2015). Processing of these images, using Structure from Motion (SfM) based software (Dandois & Ellis, 2010), enables the generation of digital elevation models (DEM) derived from 3D point clouds and highly detailed orthophotos that deliver centimetre resolution (Bendig et al., 2014; Tilly et al., 2014). Compared with multispectral, hyperspectral or thermal sensors, RGB sensors have limitations in spectral resolution, but they have the advantages of lower cost and higher spatial resolution, and it is possible to calculate vegetation indices and estimate plant height from the same set of photographs. Spectral and spatial information can then be combined to estimate crop biomass (Bendig et al., 2015).

RGB sensors mounted on UAV (drones) have been used to estimate DMY based on DEM in maize (Li et al., 2016), winter wheat (Yue et al., 2017), barley (Näsi et al., 2018), corn (Geipel, Link, & Claupein, 2014), soybean (Raymond, Cavigelli, Daughtry, Mcmurtrey, & Walthall, 2005) and grasslands, among others. Specifically for grasslands, Possoch et al. (2016), Näsi et al. (2018),
Viljanen et al. (2018) and Rueda-Ayala, Peña, Höglin, Bengochea-Guevara, and Andújar (2019) have demonstrated the use of regression models involving plant height and/or vegetation indices derived from data obtained using RGB or multispectral sensors in combination with UAV to estimate DMY. Possoch et al. (2016) used a simple linear regression to relate canopy height and biomass yield of 18 plots on 11 dates (in total 196 samples) obtaining an $r^2$ of .63. Näsi et al. (2018) used simple linear regression and random forest (RF) to estimate DMY in a mixture of timothy and meadow fescue, but they considered only one flight and a relatively small number of samples ($n = 36$) which may limit the applicability of their conclusions to other situations. Viljanen et al. (2018) used data from four harvesting dates (all in June) with 96 samples of a mixture of timothy and meadow fescue to generate multiple linear regression (MLR) and RF models. Finally, Rueda-Ayala et al. (2019) compared an on-ground system with an UAV in a limited number of plots ($n = 10$) in two different species mixtures using simple linear regression models to estimate DMY and the UAV-based system showed a lower capability in the estimation ($R^2 < .6$). While these studies illustrate the potential of RGB and multispectral sensors to provide estimations of DMY of crops in general and forage grasses in particular, they were rather limited in the number of plots, conditions tested and/or growth stages considered, which may limit the general applicability of the results found. Therefore, to the best of our knowledge, estimation of DMY from canopy height over an entire growing season for perennial ryegrass using high-resolution UAV-RGB images has not been reported yet. This is especially challenging as the height of perennial ryegrass canopies are typically lower (0.2 and 0.4 m) than those of other forage crop species such as timothy–meadow fescue mixtures (0.6–0.7 m) (Viljanen et al., 2018) and bermudagrass/alfalfa mixtures (0.8–1.1 m) (Pittman, Arnall, Interrante, Moffet, & Butler, 2015).

The aim of this study was to develop a methodology for monitoring the spatial and temporal dynamics of biomass accumulation of perennial ryegrass plots throughout the growing season in an affordable, easy-to-use, reliable and non-destructive way using an RGB camera mounted on a UAV. First, we evaluated the accuracy of canopy height models (based on the structure from motion (SfM) and the derived digital elevation model) to determine canopy height, using measurements taken using a rising plate meter as reference. Second, correlations between different variable sets (canopy height, vegetation indices and temperature as environmental driver) were evaluated. Third, different regression models (MLR, PCR, PLS and RF) were built to estimate DMY. Finally, the performance of the different models was compared.

2 | MATERIALS AND METHODS

2.1 | Field trial

The data were acquired in 2017 in an experimental field (Figure 1) of 0.5 ha located in Melle (Belgium, lat. 50.98 N, long 3.78 E). The trial was sown on a sandy loam soil at the beginning of October 2015 at a density of 2 g seeds/m$^2$ and comprised a total of 420 perennial ryegrass accessions from Europe, North Africa and West Asia, as part of the FACCE-JPI ERA-NET + Grasslandscape project (https://www.faccejpi.com/Research-Themes-and-Achievements/Climate-Change-Adaptation/ERA-NET-Plus-on-Climate-Smart-Agriculture/Grasslandscape). This set of accessions displays a very broad range of genetic and phenotypic diversity regarding growth characteristics. To facilitate the management of the trial, the accessions were grouped according to heading date (124 late, 132 intermediate, 123 early and 41 very early accessions). Within each group, the accessions were organized in a randomized block design with three repetitions, rendering a total of 1,350 plots ($420 - 3$
repetitions) of 0.6 × 1.75 m (1.05 m²). The trial was managed under a cutting regime of four or five cuts per year at a mowing height of 5 cm. We select 39 representative plots of accessions from the different groups. In the following text, we refer to these plots as “target plots” (Lp-TPs).

In addition, 90 plots (0.6 × 1.75 m, 1.05 m²), named “Control plots” (Lp-CPs) below, were sown with cultivar Olano in the border areas of the different groups. The Lp-CPs, organized in three blocks (Figure 1), were used for additional measurements and for destructive sampling between cuts to provide extra data for method development.

The field trials were mown five times in 2017 (Figure 2), but the flight campaign included only four growth periods (GP1-GP4), as described in the following sections.

2.2 | Data acquisition

2.2.1 | UAV flights

In 2017, between April 25 and October 3, a weekly flight was performed when weather conditions allowed (Figure 2) using a dodecopter UAV model Oruxstar HYDRA-12 (AltiGator) equipped with a RGB (visible spectral range) camera (α6000, Sony Corporation), with 6,000 × 4,000 pixels, equipped with a 12-mm-wide-angle lens (ZEISS Touit 2.8/12, Carl Zeiss AG). Per flight a set of nadir images of the experimental field was collected from a height of 30 m above ground level, ensuring an overlap of 80% in both directions (forward-longitudinal-lap and side-lap). The gimbal on which the sensor was fixed compensated for UAV movements during the flight and guaranteed nadir image collection. The flight route was programmed in MikroKopter-Tool V2.14b (HiSystems GmbH) and was automated to ensure whole coverage of the trial. The camera was triggered automatically according to the designed flight route. The APS-C sensor was set to record images in manual mode to avoid different settings in successive images. Shutter time (exposure), aperture (F-stop) and sensitivity to light (ISO) were adjusted in the field before the start of each flight in the field based on the light conditions at that moment and auto-focus activated. All 22 flights were executed around 13 hr (local time). Each flight delivered 100 images. Seven minutes were needed to capture them (duration of the flight was less than 10 min from take-off to landing) which allowed light conditions to remain unchanged. During each flight, a grey card of 15 cm × 20 cm (18% reference grey, Novoflex Präzisionstechnik GmbH) was placed on the ground in the middle of the field, to be used as a colour neutral reference to correct for white balance and exposure (it was visible in 15% of the images) and to check whether differences in light conditions occurred. The raw images were converted to lossless tiff files in Lightroom v6.7 (Adobe Systems Incorporated). During each flight, artificial height references (AHR) of known height were placed in the field. They consisted of four sets of seven surfaces of 0.04 m² (20 cm × 20 cm) of different heights ranging from 5 to 40 cm (5, 10, 15, 20, 25, 30 and 40 cm). The sets were placed close to the borders and in the middle part of the field (Figure 1) and were used to estimate the altimetric accuracy.

The SFM software Agisoft Photoscan v1.2.6 Professional Edition (Agisoft LLC) was used to process the 22 sets of images and build georeferenced orthophotos and digital elevation models (DEMs). Thirty-five Ground Control Points (GCPs) georeferenced using a RTK GPS (Stonex S10 GNSS, Stonex SRL) were given in the software (Figure 1). Twenty-seven of these GCPs were used in the Agisoft workflow (calibration GCP, cGCP), and the remaining 8 GCPs were used for method validation (validation GCPs, vGCPs). The resulting ortho-mosaicked images had a mean planimetric spatial resolution of 0.6 cm, and the DEMs had a mean altimetric resolution of 1.5 cm. Both resolutions were defined automatically by the software based on the camera parameters and flight altitude.

Digital elevation models (DEMs) were used to build a canopy height model (CHM) of the field for each date. For this purpose, we built two types of DEMs: 1) digital terrain models (DTMs), corresponding to the ground at the time of mowing, and 2) a digital surface model (DSM) derived from the imagery collected in the consecutive weeks with the presence of canopy on the terrain. Two types of DTM were built: (a) DTM derived from the flight carried out as close as possible to each mowing event (with a correction of 5 cm due to
the remaining stubble after the mowing) or (b) DTM based on a triangulated irregular network (TIN) interpolation of GCPs distributed all over the trial. The CHM was derived by subtracting the DTM from the DSM. In total, 40 CHMs were built, 18 of which were produced using the DTM based on the stubble (a) and 22 using the TIN interpolation (b).

2.2.2 | Ground measurements

With the same frequency as the flights (Figure 2), a single non-destructive measurement of canopy height (CH) was taken in the central part of each Lp-TP using a rising plate meter (RPM, HerboMETRE, ARVALIS-Institut du Végétal, France) consisting of a square rising plate of 0.09 m² (30 × 30 cm). In the 90 Lp-CPs, three CH measurements per plot were recorded, consistently at the same RTK GPS georeferenced locations. These locations were marked on the ground with plastic labels.

In addition, three Lp-CP plots were harvested weekly to determine the DMY. Each biomass sample was weighed fresh, and a subsample of approximately 250 g was dried at 70°C for 48 hr to determine the dry-matter weight. To obtain the DMY, this weight was extrapolated to kg/ha. The harvests were done immediately after the UAV flight or the subsequent day. On three occasions (May 31, August 24 and October 4), DMY was determined in a similar way on the same set of 39 Lp-TP plots.

Air temperature during the experiment was recorded using a weather station located at a distance of 2.25 km from the trial and was used to calculate growing degree days (GDD) considering a base temperature of 0°C.

2.3 | Data handling and analysis

The workflow followed in this study is presented schematically in Figure 3. First, the accuracy of the UAV products was evaluated. Second, the canopy heights estimated from the UAV and measured with an RPM were compared. Third, DMY was estimated using different models and results were compared. QGIS v2.14.16-Essen (QGIS Geographic Information System; Open Source Geospatial Foundation Project) was used to extract information from the images and to build canopy height models. The statistical analyses were carried out in R v3.5.1 using RStudio v1.1.456 (RStudio: Integrated Development Environment for R, RStudio Inc.) with the packages FactoMineR, pls and randomForest.

2.3.1 | UAV product quality

As one of the aims of this study was to measure variables such as canopy height linked to planimetric and altimetric spatial resolution, the quality of the products derived from the flights was evaluated. First, the 8 vGCP described in 2.2.1 were used. Variance and root mean square error (RMSE) were calculated between the coordinates (X-easting, Y-northing and Z-altitude) measured with the RTK GPS and the coordinates derived from the orthophoto and the DEM. RMSE is then equivalent to the distance between the measured and the estimated coordinates. To investigate further the altimetric resolution, AHR in the expected range of the canopy considered in this experiment (0–40 cm) described in 2.2.1 were used. We calculated the height of the central part (5 × 5 cm) of each object using the CHM of each UAV flight (see Figure 4) and compared this value with its known height.

2.3.2 | Canopy height and vegetation indices

We compared the RPM measurements and canopy height data extracted from the CHM for 18 of the 22 flights; the other four flights were executed just after a cut, and no canopy height was measured. We used the Lp-CPs and considered only the RTK GPS locations at which the RPM measurements had been carried out.

FIGURE 3 Schematic overview of the workflow followed in this manuscript. AHR, artificial height references; AIC, Akaike’s information criterion; CHM, canopy height model; CHMv, canopy height model variables; DEM, digital elevation model; DSM, digital surface model; DTM, digital terrain model; GDDv, environmental variables; Lp-CPs, control plots; Lp-TPs, target plots; LR, linear regression; MLR, multiple linear regression; NRMSE, normalized RMSE; PCR, principal component regression; PLS, partial least squares; RF, random forest; RMSE, root mean square error; RPM, rising plate meter; vGCPs, validation ground control points; Viv, vegetation index variables
(3 per plot). Georeferenced polygons with the dimension of the plate were defined in QGIS to ensure that data extraction was at the exact same location. In these polygons, the mean height values were calculated. Correlations between the CHM values and the RPM data were determined on three levels: per flight, per growth period (all flights between two cuts) and for the entire growing season. The mean CH per plot, calculated from three measurements carried out in each control plot, was used for correlation analysis. Thus, linear regressions were built per flight, growth period and growing season.

Finally, for all whole plots considered for DMY estimation (45 Lp-CPs and 39 Lp-TPs) minimum, maximum, mean, P50 (median), P90, standard deviation and coefficient of variation values were derived for the canopy height from the CHMs and for two vegetation indices: Excess Green (ExG = (2G−(R + B))/(R + G+B)) (Woebbecke, Meyer, Bargen, & Mortensen, 1995) and Excess Green minus Excess Red (ExGR = ExG−ExR = ExG−(1.4R−G)/(R + G+B)) (Camargo Neto, 2004). These indices were selected because they are normalized in indices and reported as robust (less variable) with changes in lighting conditions (Woebbecke et al., 1995). In Borra-Serrano et al. (2018), 13 vegetation indices were tested and these two indices rendered consistent and stable results.

2.3.3 Prediction of dry-matter yield from UAV and ground-derived information

Different types of regression models were built to estimate DMY for the Lp-CPs and Lp-TPs selected using different sets of variables. An overview of all variables used in the modelling step is presented in Table 1. Previous to the model building, Pearson correlation was calculated among the variables. First, we built models incorporating canopy height variables (CHMv) only. Vegetation index variables (VIv) and environmental information (GDDv) were incorporated in two further steps. Because the RPM is a mixed measure combining both the height of the canopy and its density, we also tested the performance of VIv extracted from the UAV-RGB images. By combining CHMv and VIv, it was investigated whether the VIs describing the greenness of the canopy could contribute complementary information. A linear regression (LR) was used as reference. Backward multiple linear regression (MLR) using different variables from the three sets and a principal component regression (PCR) with all variable sets were carried out. In addition, partial least squares regression (PLS) and the machine learning technique random forest (RF) were tested. All variables

FIGURE 4 Schematic presentation of the calculation of the canopy height model (CHM, [a]). Side and top view of a set of artificial height references (AHR, [b]). Accordance of real height and the height of the twenty-eight AHR determined in four different zones of the field trial for five flights during growth period 3 (c–g). Due to adverse weather conditions, any flight was possible between 14/07 and 01/08. The grey band around the line indicates the 95% confidence interval [Colour figure can be viewed at wileyonlinelibrary.com]
except RPM were used, since we aimed to estimate DMY based on UAV data and replace RPM measurements. Only in one MLR model, RPM data were added to evaluate the complementarity with other variables.

To estimate the performance of the different models, the complete plot data set was divided into a training (70%) and validation (30%) set. The derived regression models were applied to the validation data set, and results (Table 2) were analysed and compared based on the root mean squared error (RMSE) and normalized RMSE (NRMSE) between the observed and predicted DMY, using Akaike’s information criterion (AIC) and the coefficient of determination ($R^2$). For the MLR models, after the backward selection process,
procedure (using the training data set), the results were evaluated in terms of significance of the predictor variables and multicollinearity using the variance inflation factor (VIF). The model was manually adjusted by removing predictor variables with \( p > .10 \) or with a VIF > 10. Since the MLR models were hierarchical, the significance of the variables was evaluated at each stage using ANOVA.

3 | RESULTS

3.1 | Accuracy of UAV products

The accuracy test explained in section 2.3.1 (performed for both types of CHMs) demonstrated that better results were achieved with the CHMs built using the TIN interpolated DTM. Only data derived from those CHMs are presented.

3.1.1 | Accuracy of the ground level references

Based on the eight validation GCPs (vGCPs), variance and RMSE were estimated for \( X \), \( Y \) and \( Z \) coordinates. The variance values ranged from \(-1.10\) cm to \(0.20\) cm, from \(0.04\) cm to \(0.60\) cm and from \(-2.25\) cm to \(-0.41\) cm for the three axes (\(X\), \(Y\) and \(Z\) respectively). RMSE values ranged from \(0.63\) cm to \(1.49\) cm (for \(X\) and \(Y\)) and from \(0.94\) cm to \(1.89\) cm for \(Z\).

3.1.2 | Accuracy of the height references

The variance was calculated for each AHR, field zone and flight (data not shown), rendering an average variance value of \(-0.56\) cm. In Figure 4, the height estimates for the AHR are shown for five representative CHM subsets.

3.2 | Comparison between the UAV-derived canopy height and the canopy height measured with the rising plate meter

Correlations were established between the CHM and RPM per flight and per growth period (4 to 5 flights/GP, GP1–GP4, Figure 5). The correlation between canopy height values derived from the CHM (CH_mean, \( n = 3 \) per plot) and the RPM values were 0.96, 0.80, 0.81 and 0.80 (\( p < .001 \), Figure 5 b–e) respectively. The correlation for the

FIGURE 5 Representation of a plot (dashed lines) and the positions measured with the rising plate meter (RPM, grey squares) in an orthophoto and canopy height model (a). The positions of the RPM measurements were RTK GPS referenced (black dots, yellow markers were placed in the field) to enable extracting canopy height from the same location from the canopy height models (CHM). Comparison of the canopy height using the rising plate meter (RPM) and the mean canopy height (CH_mean) derived from the different canopy height models for the four growth periods (GP, [b–e]). A regression was applied per flight. The correlation of all data for each GP is calculated and shown in the upper left corner of each figure and the (dot dashed) regression line with the 95% confidence interval (grey band). Comparison for the entire growing season with a linear and a loess regression and their 95% confidence interval between CHM and RPM (f) [Colour figure can be viewed at wileyonlinelibrary.com]
complete set of flights (17 in total, as one flight had to be discarded due to an error in the measurement with the RPM) was 0.92 (p < .001 Figure 5f). In general, higher canopy height values were obtained from the CHM than from the RPM. The lower canopy height values (underestimation) in the RPM data were expected as the RPM flattens the canopy during the measurement. The fact that this was not observed for the first three weeks of the second GP suggests that the sparse canopy (consequence of a severe drought period) and the open areas in which the ground was visible influenced the mean value calculated for the area covered by the polygon in QGIS, and that the compression effect by the RPM was reduced as the canopy was low. The compression by the plate was higher at later weeks of the regrowth period. A linear regression and a loess (non-parametric locally weighted) regression were carried out for the entire set of flights (Figure 5f) between CHM and RPM to study their relation. It was observed that the relation between both variables changed depending on the mean canopy height. Clearly, three zones could be determined with a different relation: from 5 to 13 cm, 13–25 cm and 25–40 cm.

3.3 | Correlation between variables

Correlation results are presented in a correlation map (Figure 6). Dry-matter yield displayed a good correlation with the RPM values (r = .79). Correlation between DMY and CHMv such as CH_min, CH_mean, CH_P50 and CH_P90 was also in the same range. The highest correlation was obtained between DMY and CH_P50 (r = .80). Low r-values were obtained between DMY and VIv (0.28–0.53). A negative correlation was found between GDD and DMY (r = −.42).

RPM values were highly correlated with CHM data, especially with CH_mean and CH_P50 (r = .88). RPM values were also significantly correlated with VIv, but the correlation values were lower compared with CHMv. The highest correlations were found for both ExG_stdev and ExGR_stdev (r = .67). The correlation is related to variation in greenness of the canopy. No clear preferred correlation could be found for ExG or ExGR-derived variables in relation to RPM. This was related to the high correlations between ExG and ExGR variables (r-values for mean, P50 and P90, ranging between .94 and .99).

3.4 | Dry-matter yield models

Different models were built to estimate DMY. Results are presented based on the validation data set (Table 2; see Materials and Methods above). The first model was a linear regression with the RPM variable, which served as reference for other models. For our trial, this linear regression resulted in an RMSE of 986 kg/ha and a NRMSE of 31.0%. The linear regression based on the best performing UAV-derived variable, CH_P50, performed better than the RPM, with RMSE equal to 876 kg/ha (NRMSE = 27.6%), which is an improvement of 11%.

Using multiple linear regressions (MLR), we evaluated whether a further improvement in the DMY prediction could be achieved. ANOVA was used to compare the different MLR model performances and their significance. In a first multiple linear regression model (MLR), only CHMv were considered in the model. The resulting RMSE was 825 kg/ha (NRMSE = 26.0%) and thus slightly lower than the linear model with CH_P50 and a 16% improvement compared to the linear RPM model. This result was paired with similar AIC and a slightly higher R² and achieved by including the CH_stdev. However, the observed improvement compared with the linear model incorporating CH_P50 was just not large enough to be significant (p = .056).

CHMv and VIv were combined, and the resulting model hardly improved the prediction and resulted in an RMSE of 822 kg/ha (NRMSE = 25.9%). Indeed, no significant improvement was found compared with the first MLR (p = .100). To investigate whether the RPM contained further complementary information compared with the CHMv and VIv, a third MLR was calculated including also RPM. A higher RMSE was achieved in this case: 857 kg/ha (NRMSE = 27%) compared with the first MLR (p = .031). To compare the different GP, an ANOVA was used to compare the different MLR model performances and their significance. In a fourth MLR model, GGDv were included in combination with CHMv and VIv. By incorporating GGDv, we added information on growth drivers and on the different GP. In this case, a significant (p < .001) improvement in the DMY estimation of 31% compared with the linear RPM model was obtained, resulting in an RMSE = 679 kg/ha (NRMSE = 21.3%).

In the fifth model, a principal component regression (PCR) was used for which five components were selected based on the eigenvalues, resulting in an RMSE of 714 kg/ha (NRMSE = 22.5%) or a 28% improvement compared with the linear RPM model. Finally, partial least square (PLS) and random forest (RF) methods were evaluated to predict DMY. After optimization, the PLS and RF resulted in RMSEs of 739 (NRMSE = 23.3%) and 769 kg/ha.
TABLE 3  Best MLR to predict DMY with all variables sets included in the model. For all variables mean, standard deviation (SD), effect size and standardized effect size are presented.

<table>
<thead>
<tr>
<th>Variable set</th>
<th>Variable</th>
<th>Mean ± SD</th>
<th>MLR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DM</td>
<td>3,233.38 ± 1,613.8</td>
<td>-959.3</td>
</tr>
<tr>
<td></td>
<td>Intercept (j0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CH_Mv</td>
<td>205.42 ± 89.9</td>
<td>10.14***</td>
</tr>
<tr>
<td></td>
<td>CH_stdev</td>
<td>20.15 ± 10.9</td>
<td>-31.34**</td>
</tr>
<tr>
<td></td>
<td>Vv</td>
<td>0.08 ± 0.02</td>
<td>14,057.15***</td>
</tr>
<tr>
<td></td>
<td>GDDv</td>
<td>2,331.10 ± 936.3</td>
<td>-0.63***</td>
</tr>
<tr>
<td></td>
<td>GDDdif</td>
<td>561.85 ± 145.9</td>
<td>5.46***</td>
</tr>
</tbody>
</table>

Significance levels:
***p < .001
**p < .01

(NRMSE = 24.2%) respectively. However, these models did not contribute any improvement compared with the best MLR model (MLR with CH_Mv + Vv + GDDv). In Table 3, the summary of the best MLRs is presented. The variable with the highest standard effect size in the model is CH_P50 followed by the GDDdif, which was calculated as the GDD accumulated after each cut.

4 | DISCUSSION

4.1 | Accuracy of UAV-derived data

Prior to the use of DEMs and orthomosaics, an evaluation whether these products do provide correctly georeferenced information is needed and is especially relevant in time series. Here, we used two methods: a set of validation ground control points (vGCPs) and artificial height references (AHR). Similar RMSE values were obtained for the eight vGCPs used. However, the flights just before a cut tended to show a slightly higher RMSE (from about 1.24 cm up to 1.89 cm). The higher RMSE was caused by the partial covering of cGCPs by the taller surrounding grass, resulting in fewer images where the cGCPs were visible and could thus be used to calculate the 3D model. This effect was accentuated when the cGCP differed more from the nadir position in the image (results not shown). In our case, the flight plan, overlap and number of images were sufficient and correct, as demonstrated by a mean calculated error in the Z-axis below 2 cm. The mean error calculated in the Z-axis was comparable or even better than shown in other studies in grasslands. For example, Näsi et al. (2018) or Viljanen et al. (2018) presented RMSE values in the Z-axis of 6.9 cm and 1.0–4.8 cm respectively. The size of the area “free of vegetation” needed around a GCP depends on the canopy height and the lens used. For a taller crop, more open area around the GCP is needed. Therefore, this is an important point of attention when planning the experiment. Furthermore, the stitching software needs sufficient overlap between images for the edges of the field trial to provide the same height accuracy as the central area of the DEM. Thus, the flight plan of the drone should be taken wide enough. In this study, every point of the field trial appeared in a minimum of 15 images, and the GCPs located in the central part of the field were present in 20 images. The number of images was higher than in other studies where the area of interest was captured in a minimum of nine images (Bendig et al., 2015; Viljanen et al., 2018).

Regarding the AHR, slightly higher variances were obtained for the set located closer to the border (zone 4). This result was expected as they appear in a lower number of images and was covered by a smaller angle of view to reconstruct them in the photogrammetry processing compared with the references located in the middle of the field. The lowest AHR (5 cm) showed higher variances in height. This outcome can be explained by the higher difficulty involved in the detection of less tall objects and their modelling due to (a) their closer distance to the soil that increases the difficulty for the SfM software and (b) the taller surrounding grass, especially at the end of the growth period that reduces its visibility in the extreme oblique images. To obtain the DTM needed to build the CHM, it is possible to use a flight without vegetation (Bendig et al., 2015), interpolation of points (manually or automatically) classified as ground using the borders of the plots or visual bare soil areas (Viljanen et al., 2018; Yue et al., 2017), to use a ultrasonic or laser scanner (Pittman et al., 2015) or interpolating X, Y and Z coordinates of the plot corners measured with a GPS (Chang, Jung, Maeda, & Landivar, 2017). In our case, very good results were obtained using interpolated coordinates measured with the RTK GPS. Using that information, it was possible to determine the slope of the field and to detect bumps or irregularities present. Consequently, a higher precision in AHR height was achieved.

4.2 | Canopy height determined using the RPM and CHM

The rising plate meter is often used as a reference for non-destructive assessment of biomass in pastures (e.g. Holshof et al., 2015; Mathieu & Fiorelli, 1985). In this study, it was also used as a reference even though we were aware that the RPM compresses the canopy, which makes the height measurement dependent on the canopy density, resulting in lower canopy heights in comparison with the CHM (information acquisition without physical contact). Furthermore, the RPM encounters problems when the...
sward is sparse, non-uniform, presents low growth rate or is too high. Therefore, it has been reported to be more reliable when the sward is dense and has reached a height of 20–30 cm (Viljanen et al., 2018). Having ruler-based height measurements as an undisturbed reference measurement was not an option as this measurement can only be done very locally, and questions would remain with respect to the height of the canopy to be measured (e.g. average canopy height, tallest leaves,...), the number of measurements needed for a good estimation or how to avoid human bias (Sanderson, Rotz, Fultz, & Rayburn, 1999).

The correlation obtained between the canopy height measured with the RPM, and the variables derived from the CHM for the Lp-CPs was 0.92 (Figure 5b). The correlation is similar to the values reported by Viljanen et al. (2018) for a timothy and meadow fescue mixture (r = .91–.94), similar to the values reported in a 3-year study by Bareth and Schellberg (2018) (r = .83–.91) and higher than the value of .55 reported by Poss och et al. (2016). It was observed that the CHM-CH was underestimated for the earliest measurement dates of a growth period and during a growth period with reduced growth due to abiotic stress (i.e. drought stress). This observation may be attributed to the period and season and during a growth period with reduced growth due to abiotic stress 

4.3 | Dry-matter yield prediction

In total, 19 variables related to the canopy height model (CHMv), the vegetation indices (VIs) and the environment (GDDv) were used to build models that predict DMY. We tested whether the combination of different types of data resulted in a better DMY prediction and investigated the influence of the methodology selected for the regression. As demonstrated by Bendig et al. (2015) and Brocks and Bareth (2018) for barley, Yue et al. (2017) for winter wheat, Näsi et al. (2018) for a timothy and meadow fescue mixture or for maize in Geipel et al. (2014), and as is shown in our regression analysis, the canopy height is (well) correlated to yield. Slightly lower correlation values (r = .79) were obtained in comparison with those reported by Viljanen et al. (2018) for a timothy and meadow fescue mixture (r ranges from .79 to .98). This can probably be explained by (a) the taller crop (up to 0.7 m) in their study, and (b) the much more diverse set of accessions in the present study, implying differences in canopy density which cannot be captured by CH_P50. Here, we used the model for DMY estimation based on RPM measurements as reference. Even if the RPM measured a weighted or averaged canopy height when the linear regression was built, it presented a higher RMSE than the model using the CH_P50, which achieved an 11% reduction in the RMSE.

To increase the performance of the prediction, the MLR model with more CHMv resulted in a slight improvement of 16% compared with the RPM model when CH_stev was included. Variables such as mean, median or 90th percentile described the vertical distribution of the canopy and the standard deviation and coefficient of variation, the vertical complexity and heterogeneity (variation). In a second MLR, spectral information was included in the form of vegetation indices (VIs). As VIs provide information on soil coverage and canopy density, they showed significant correlations with the RPM which is related to sward density. Nevertheless, they presented a lower correlation with DMY. Hence, it was difficult to achieve a relevant decrease in the RMSE compared with the first MLR model when ExG_P90, ExG_stdev and ExG_cv were included in the model. Indeed, this model did not show a significant improvement compared with the MLR with only CHMv. From this non-significant result, it can be concluded that the visual VIs are not crucial to achieve a better DMY estimation. The inclusion of other VIs such as NDVI or other wavelength ranges might add relevant information related to vegetation status and canopy cover in the earlier weeks after a cut. However, NDVI may reach saturation in dense canopies or when LAI values are about 2.5–3 (Tilly et al., 2014) and thus may not be relevant for DMY estimation, because 4 or 5 weeks after the cut, grasslands can reach LAI values of 4 (Lambert, Peeters, & Tousaint, 1999). As we focused on the use of a (low-cost) RGB camera with high resolution to obtain accurate height estimations, the performance/contribution of NDVI measurements could not be evaluated in our study. Finally, as growth is strongly driven by temperature (e.g. Voorend et al., 2014), two variables that incorporate environmental data were selected (GDD and GDDdif) and showed to improve the prediction. Indeed, a seasonal growth pattern of high DMY in the spring and to a lesser extent in the autumn has been reported for Lolium perenne (Barre et al., 2018). Furthermore, the growth of plants has been shown to be related to temperature sum (Yang, Logan, & Coffey, 1995) taking into account a species dependent base temperature (Tb for Lolium perenne = 0°C) (Lootens et al., 2016).
With each step of the incorporation of a new group of variables, the RMSE decreased without increasing AIC, reaching a DMY model prediction improvement of 31% compared with the reference model based on RPM. Including the RPM data in the MLR resulted in no complementary effects. As can be seen in Table 3, the standardized effect size (calculated as standard deviation multiplied by effect size) was estimated for the significant variables for the best MLR model to explain the variability of the DMY. As expected, the median canopy (CH_P50) explained the most variation followed by the temperature sum per growth period (GDDdiff).

As an alternative for MLR, a PCR was executed to safeguard the regression against the correlations between the explanatory variables. The five principal components included in the regression model captured 96% of the variance in the X data, but this model did not result in an improvement of the prediction of DMY compared with the best MLR model.

Because of the typically large data size, the use of machine learning techniques is increasing in remote sensing. Random forest (RF) is one of the most popular algorithms because it requires no feature selection and is less sensitive to overfitting than MLR. Viljanen et al. (2018) incorporated this method for DMY estimation, and they compared its performance with that of MLR. Similar to our study, one of their best results was obtained with RF, but it did not perform better than the best MLR that incorporated all the variable sets available. The use of PLS in our case did also not result in a further reduction of the RMSE.

To the best of our knowledge, no studies have been performed for perennial ryegrass plots related to DMY estimation with UAV imagery using an RGB camera to compare our results with. This might be due to the problems that could be encountered such as lower canopy or the high variability between the different accessions or cultivars. Compared with previous UAV imagery research in other crops (Geipel et al., 2014; Raymond et al., 2005; Li et al., 2016; Näsi et al., 2018; Yue et al., 2017), in this study a higher number of flights (and thus dates), samples and accessions were involved in the estimation of DMY though different regression techniques, from a simple linear regression to more advanced techniques such as partial least squares regression and random forest. We have shown that it is possible to estimate canopy height of perennial ryegrass plots using with a cheap, high-resolution RGB camera mounted on a drone and predict DMY in a better way (lower RMSE) than with a rising plate meter. Indeed, a high spatial resolution (1.5 cm) and accurate CHMs in the Z-axis (1–1.9 cm) were achieved. By performing frequent flights throughout the growing season, also a high temporal resolution was achieved, which allowed to monitor all plots in a large field trial in detail in a non-destructive way.

AKNOWLEDGMENTS

The research reported in this article was carried out at ILVO as part of the iSense project (iSense.farm and www.ilvo.vlaanderen.be) without any specific grant from funding agencies in the public, commercial or not-for-profit sectors. The authors wish to thank Thomas Vanderstocken, Filip De Brouwer and Ruben Van De Vijver for piloting UAV flights, the Grasslandscape team for the permission to fly over the trial, special thanks for Jean-Paul Sampoux and Miriam Levenson for English language corrections.

ORCID

Hilde Muylle https://orcid.org/0000-0001-7350-4179
Peter Lootens https://orcid.org/0000-0002-3275-3459

REFERENCES


