ICPA
17th INTERNATIONAL CONFERENCE ON PRECISION AGRICULTURE
ConBAP 2026
CONGRESSO BRASILEIRO DE AGRICULTURA DE PRECISÃO E DIGITAL
13-16 July 2026 | Porto Alegre, Brazil
ISPA
AsBraAP

UAV-Based Multispectral Modelling of Biomass and Crude Protein Yield for Green Biorefinery Applications

Michele Canciani[1], Yuma Ikeda[2], Nikolaj Peder Hansen[3], Søren Krogh Jensen[3], Martin Riis Weisbjerg[3], Thomas Didion[3], Eusun Han[2], Uffe Jørgensen[2]

1) Department of Agricultural. Food, Environmental and Animal Sciences, University of Udine, Via delle Scienze, 206, 33100 Udine, Italy
2) Department of Agroecology, Aarhus University, Foulum, Blichers Allé 20, DK-8830, Tjele, Denmark
3) Department of Animal and Veterinary Sciences, Aarhus University, Foulum, Blichers Allé 20, DK-8830, Tjele, Denmark

Abstract.

The increasing demand for sustainable protein sources in animal production systems has intensified interest in Green Biorefinery (GBR) approaches based on perennial grass biomass. Accurate estimation of above-ground biomass (AGB) and crude protein (CP) yield is essential to optimize harvest timing and maximize protein extraction efficiency. This study aimed to (i) optimize feature extraction from a single unmanned aerial vehicle (UAV) flight integrating RGB and 10-band multispectral data, and (ii) improve estimation accuracy for support harvest timing decision in the production of biomass for GBR.

Field experiments were conducted over two growing seasons (2023–2024) in Denmark on five perennial grass varieties under three nitrogen fertilization regimes. A comprehensive dataset of 3,280 features was derived, including structural metrics, spectral bands, vegetation indices, and texture-based indicators. Two feature selection approaches, Competitive Adaptive Reweighted Sampling (CARS) and Recursive Feature Elimination with Random Forest (RFE-RF), were applied prior to model development.

Results showed that RFE-RF provided the highest predictive performance for both AGB ($R^2 = 0.885$; RMSESD = 0.240) and CP yield ($R^2 = 0.900$; RMSESD = 0.302), outperforming or matching CARS across datasets. Models trained on the combined multi-variety dataset consistently outperformed single-variety models. Key predictors included canopy volume for AGB estimation and red-edge texture-derived variables for CP yield estimation.

These findings demonstrate that integrating multi-variety datasets from a single flight with robust feature selection methods enhances model accuracy while reducing data dimensionality, supporting efficient UAV-based monitoring for GBR optimization.

Keywords.

UAV, Above-ground biomass, Crude Protein, Feature selection, Multispectral imagery, texture indices, Grassland, Green Biorefering

Introduction

Global population growth continues to increase the demand for protein in animal production systems, intensifying the need for alternative and sustainable protein sources. Green Biorefinery (GBR) systems represent a promising solution, as they enable efficient extraction of protein from plant biomass and may substantially reduce Denmark's reliance on imported plant-based protein (Hermansen et al., 2017). Previous studies have demonstrated that perennial grasses can produce high-quality biomass suitable for protein extraction (Larsen et al., 2024). Recent findings further indicate that the relationship between above-ground biomass (AGB) yield and crude protein (CP) yield varies throughout the growing season, reaching coefficients of determination (R^2) between 0.90 and 0.98 during the July and August harvests (Canciani et al., unpublished data). These results highlight the importance of accurately estimating biomass and CP yield across the season to determine the optimal harvest timing to optimize the CP extraction from GBR.

Several studies have developed predictive models for AGB, CP, and nitrogen (N) content in grasslands by integrating data from different unmanned aerial vehicle (UAV) sensors with statistical or machine-learning approaches. These models have achieved high predictive performance, with R^2 values exceeding 0.90 for AGB, 0.91 for CP, and 0.85 for N content (Bazzo et al., 2023; da Silva et al., 2026; Oliveira et al., 2024; Qi et al., 2025). However, the integration of multiple sensors or the use of computationally intensive deep-learning architectures can increase workflow complexity and substantially extend processing time.

The present study had two main objectives: (1) to optimize feature extraction from a single UAV flight combining RGB and 10-band multispectral sensors, and (2) to improve estimation accuracy by selecting the most informative variables from an expanded feature set.

Material and Methods

Field trials were conducted in 2023 and 2024 at the Foulumgaard Research Farm of Aarhus University (Tjele, Denmark; 8830). Five perennial grass varieties were sown in April 2023: two *Festulolium* cultivars, two perennial *Ryegrass* cultivars, and one *Tall Fescue* cultivar. Crops were grown under three nitrogen fertilization regimes (300, 450, and 600 kg N ha^{-1} $year^{-1}$). Each treatment was replicated four times, resulting in a total of 60 plots (6 m × 10 m). Plots were arranged in three rows of 20 plots each following a randomized layout. Across the two growing seasons, three harvests were performed in 2023 and five in 2024 using a Haldrup F-55 plot harvester (Haldrup GmbH, Germany). The first harvest of May 2024 was excluded due to low CP concentration not suitable for GBR. Biomass yield was recorded for all harvests, whereas CP and nitrogen content were measured only during the 2024 production season.

Dry matter (DM) yield was determined by collecting a subsample of fresh biomass from each plot. Subsamples were oven-dried at 60 °C for 48 h until constant weight, and DM content was used to estimate total DM yield per plot. Plot-level DM yield was subsequently scaled to a per-hectare basis. N and CP concentrations (g 100 g^{-1} DM) were quantified by DLF Seeds (Store Heddinge, Denmark) using a FOSS NIRS-DS2500 instrument and an in-house near-infrared calibration developed from chemically determined reference data. Chemical analyses for calibration development followed the ANKOM 2000 fiber analyzer protocol (ANKOM Technology). Crude protein yield was calculated as DM yield multiplied by CP concentration.

Prior to each harvest, UAV flights were conducted using a Matrice 300 RTK platform (DJI Technology Co., Shenzhen, China) at an altitude of 30 m. RGB imagery was acquired using a Zenmuse P1 camera, while multispectral data were collected using a 10-band Dual Camera System (MicaSense Inc., Seattle, WA, USA). Four ground control points (GCPs) were deployed for georeferencing. Raw imagery was processed using Pix4Dmapper (Pix4D, S.A., Lausanne, Switzerland). From RGB data, canopy height and canopy volume were derived using the generated mesh models (Valluvan et al., 2023). Multispectral reflectance orthomosaics were used to extract band-level reflectance and compute eight texture metrics (Contrast, Dissimilarity, Homogeneity, Angular Second Moment, Entropy, Mean, Variance, and Correlation) (Yamaguchi et al., 2024). The 10 reflectance bands and 80 texture metrics were used to calculate 44 normalized vegetation indices (NVIs) and 3,240 normalized difference texture indices (NDTIs). The final dataset comprised 3,280 features.

Due to high collinearity among the extracted variables, two feature selection methods were applied prior to model fitting: Competitive Adaptive Reweighted Sampling (CARS) (Li et al., 2009) and a new Recursive Feature Elimination with Random Forest (RFE-RF) model. For each ground-truth variable, separate estimation models were developed for each variety and for the combined dataset (all varieties). CARS parameters were set as follows: maximum latent variables = 20, iterations = 300, cross-validation = 10, fold-CV = 10, and training set proportion = 0.75. RFE-RF parameters included: random cross-validation strategy, 10 outer iterations, removal of 10 features per step, and a training set proportion of 0.75. The selection of the best parameters for the RF models were done by OPTUNA hyperparameter optimization (Akiba et al., 2019). Model performance was evaluated using the R^2 and $RMSE_{SD}$. Feature importance was assessed based on selection frequency and standardized coefficient matrices.

Results and Discussion

The performance of the two models, CARS and RFE-RF, was comparable across the different varieties and target variables. The most suitable model for AGB estimation was RFE-RF trained on the combined dataset including all varieties ($R^2 = 0.885$; $RMSE_{SD} = 0.240$), yielding results consistent with those reported by Bazzo et al. (2023). A total of 11 variables were consistently selected across iterations, with canopy volume emerging as the most influential predictor.

For CP content estimation, CP yield was the best-predicted variable among all targets. The RFE-RF model achieved the highest accuracy ($R^2 = 0.900$; $RMSE_{SD} = 0.302$), in agreement with findings by Qi et al. (2022). Fifteen variables were identified as most relevant, among which the ratio of $red_{edge740\ variance}$ to $red_{edge717\ homogeneity}$ showed the strongest influence.

The global dataset outperformed single-variety datasets, likely due to the limited number of samples within individual varieties. Among the latter, Hipast showed the best performance for biomass estimation (CARS: $R^2 = 0.845$; $RMSE_{SD} = 0.296$), while Perseus achieved the highest accuracy for CP yield estimation (CARS: $R^2 = 0.888$; $RMSE_{SD} = 0.233$).

Conclusion

Overall, the integration of multi-variety datasets showed great potential in enhancing model robustness and predictive performance in remote sensing-based trait estimation. The new feature selection method, RFE-RF, prove effective in identifying key spectral and structural variables driving model accuracy.

References

ANKOM Technology. (2017a). ADF Method 12: Acid detergent fiber, in feeds—Filter bag technique (for A2000 and A2000I). ANKOM Technology. https://www.ankom.com/sites/default/files/documentfiles/Method_12_ADF_A2000.pdf

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2623–2631.

Bazzo, C. O. G., Kamali, B., Hütt, C., Bareth, G., & Gaiser, T. (2023). A review of estimation methods for aboveground biomass in grasslands using UAV. *Remote Sensing*, *15*(3), 639.

da Silva, R. C., Tommaselli, A. M., Imai, N. N., Martins-Neto, R. P., da Silveira, D. S., & Moro, E. (2026). Estimation of grassland nitrogen content using UAV ultra-wide RGB images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *10*, 99–104.

Hermansen, J. E., Jørgensen, U., Lærke, P. E., Manevski, K., Boelt, B., Jensen, S. K., Weisbjerg, M. R., Dalsgaard, T. K., Danielsen, M., Asp, T., & others. (2017). *Green biomass-protein production through biorefining*. DCA-Nationalt Center for Fødevarer og Jordbrug.

Larsen, S. U., Manevski, K., Lærke, P. E., & Jørgensen, U. (2024). Biomass yield, crude protein yield and nitrogen use efficiency over nine years in annual and perennial cropping systems. *European Journal of Agronomy*, *161*, 127336.

Li, H., Liang, Y., Xu, Q., & Cao, D. (2009). Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. *Analytica Chimica Acta*, *648*(1), 77–84.

Oliveira, R. A., Näsi, R., Korhonen, P., Mustonen, A., Niemeläinen, O., Koivumäki, N., Hakala, T., Suomalainen, J., Kaivosoja, J., & Honkavaara, E. (2024). High-precision estimation of grass quality and quantity using UAS-based VNIR and SWIR hyperspectral cameras and machine learning. *Precision Agriculture*, *25*(1), 186–220.

Qi, H., Chen, A., Yang, X., & Xing, X. (2025). Estimation of crude protein content in natural pasture grass using unmanned aerial vehicle hyperspectral data. *Computers and Electronics in Agriculture*, *229*, 109714.

Valluvan, A. B., Raj, R., Pingale, R., & Jagarlapudi, A. (2023). Canopy height estimation using drone-based RGB images. *Smart Agricultural Technology*, *4*, 100145.

Yamaguchi, T., Sasano, K., & Katsura, K. (2024). Improving efficiency of ground-truth data collection for UAV-based rice growth estimation models: Investigating the effect of sampling size on model accuracy. *Plant Production Science*, *27*(1), 1–13.