Evaluation of satellite data for estimation of legume proportion in clover grass

Frandsen, T.S., Malskaer, L.M., Torp. M.T., *SEGES Innovation, Agro Food Park 15, 8200 Aarhus N, Denmark; tsf@seges.dk*

Temporary grasslands in Denmark are mostly used for cutting and typically consist of mixtures of grasses (perennial ryegrass, festulolium) and legumes (white and red clover) as the yield and forage quality are more stable compared to monocultures of species. However, the legume proportion differs within and between fields, which results in different optimal nitrogen requirements, as the nitrogen requirement is negatively correlated with the legume proportion. Determination of the legume proportion can be challenging and time consuming. The purpose of this study was to evaluate the use of sentinel 2 data to estimate the legume proportion in grass clover fields. The ground truth (validation data) was primarily NIR-data on forage choppers from 397 fields and data based on image analysis from 13 fields- in total a dataset of almost 300.000 observations. The legume proportion in the fields ranged from 0-95 percent with an average of 22 percent. The final model with satellite images as features showed a mean absolute error of 7.8 percentage point on field level. This study indicates that sentinel 2 data can be useful for differentiated nitrogen fertilization according to the legume proportion but requires further refinement.
Keywords: legume proportion, estimation, satellite data.

Introduction

Inclusion of legumes in grasslands generally increases yield stability and forage quality compared to fertilized grass only leys (Egan et al. 2018, Lüsher et al. 2014, Søegaard 2009 & Johansen et al. 2017). The legume proportion in grasslands has a major impact on the nitrogen requirement of the crop as the legume perform symbiotic fixation of atmospheric N. Eriksen et al., (2019) illustrated the beneficial effects of differentiated nitrogen fertilization according to the legume proportion. Estimation of the legume proportion can either be done by visual estimation with a high uncertainty or by botanically fractionation which is time consuming. Skovsen et al. (2017) and Hennesy et al. (2021) demonstrated a method for estimation of the legume proportion using analysis of RGB images of mixed swards – however good quality images typically require flash or intensive light to reduce effects of shadows. For commercial use such systems require further development before a wide implementation. More self-propelled forage choppers are equipped with NIRS sensors to measure different quality traits in the forage i.e., content of dry matter, fibre crude protein, but also the legume proportion. However, these NIRS sensors are few and requires calibration to deliver useful information. The purpose of this study was to evaluate whether a model based of satellite data could estimate the legume proportion in accordance with the measurements of NIRS sensors on forage choppers.

Materials and methods

A dataset based on measurements of the legume proportion in 1185 grass clover fields in Denmark in 2018, 2022 and 2023 was used as input.

A rather large proportion of the fields (775) had either negative or values for the legume proportion above 100% or had only a sparse coverage and were discharged from the analysis leaving a training set of 410 fields with reliable data. Of those 373 originates from 2023. For these 410 fields the number of registrations per field ranged from around 100 to well above 4000 with an average of 730.

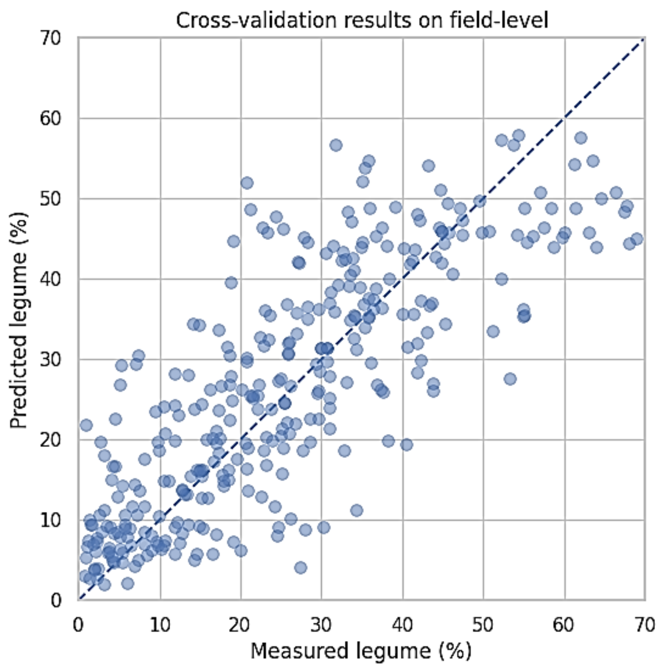
As the data source is primarily the NIRS sensors on forage choppers the data are recorded as the farmers are cutting the fields- typically 3-5 times a year. Consequently, the number of fields measured by the NIRS sensors differs in each month, resulting in a data set of 96, 86, 182 and 46 fields for May, June, September and October, respectively. For each field with a documented legume proportion, the polygon was extracted, followed by the collection of 13 bands from the Sentinel 2 satellite's cloud-free image nearest to the harvest day. Vegetation indices such as NDVI, gNDVI and GCI were calculated to capture information not contained in the raw bands. Additionally, the altitude of each field was incorporated into the model. The final model included these indices, all atmospheric-corrected (L2A) Sentinel 2 bands, field altitude, and a day-of-the-year indicator for the harvest date.

To evaluate the effectiveness of the model, a cross-validation technique was used. In this process, each farm was excluded one by one from the training data. The model was then trained using data from the remaining farms and applied to predict the outcomes for the excluded farm's fields. The model's overall performance was determined by calculating the average of the mean absolute error (MAE) from the predictions of each farm that was left out. This method provided a clear measure of how accurately the model could predict across different farms.

Results and discussion

The legume proportion varied from 0-69 percent at field level with an average of 21 percent. On monthly basis the average legume proportion was 20.8, 10.8, 25.3 and 24.1 for May, June, September and October respectively. Data from mid-June until September was excluded from the analysis as the estimation accuracy in these months was lower – perhaps due to the flowering of the legumes resulting in different satellite values.

The best model predicts the legume proportion with a mean absolute error of 7.8 percent. The scatter plot reveals a clustering of data points near the 45-degree line, which symbolizes the line of perfect prediction. This clustering indicates that the model's predictions are, on average, closely aligned with the actual values. Despite this, there is a notable dispersion of points around the line, reflecting considerable variability in prediction accuracy. This variability manifests as a broad spread of values, suggesting that while the model is generally reliable, there are instances where its predictive capacity deviates from the measured values.



Conclusion

The model needs further development and refinement but indicates that satellite data can be useful for a vast estimation of the legume proportion as a tool for the farmers and advisors to differentiate nitrogen fertilization according to the legume proportion. The outcome might be a higher net margin for the farmer, lower GHG emissions and lower potential nitrogen leaching from grasslands and the following crop.

Acknowledgements

We would like to thank “Promilleafgifsfonden for Landbrug”, who supported this study.

References

Egan M., Galvin N. and Hennessy D. (2018) Incorporating white clover (Trifolium repens L.) into perennial ryegrass (Lolium perenne L.) swards receiving varying levels of nitrogen fertiliser: effects on milk and herbage production. Journal of Dairy Science, 101:3412–3427.

Eriksen J., Frandsen T.S., Knudsen L., Skovsen S., Jørgensen R.N., Steen K.A., Green O. and Rasmussen J. (2019) Nitrogen fertilization of grass-clover leys, Grassland Science in Europe, Vol. 24, 103-109

Hennessy D., Saad M., Mac Namee B., O'Connor N.E., McGuinness K., Albert P., Narayanan B., Fitzpatrick E. and O’Connor A.H. (2021) Using image analysis and machine learning to estimate sward clover content. Grassland Science in Europe, Vol. 26, 132-134

Johansen M., Søegaard K., Lund P. and Weisbjerg M.R. (2017) Digestibility and clover proportion determine milk production when silages of different grass and clover species are fed to dairy cows. Journal of Dairy Science 100, 8861-8880

Lüscher A., Mueller-Harvey I. Soussana J.F., Rees R.M. and Peyraud J.L. (2014) Potential of legume-based grassland-livestock systems in Europe: a review. Grass and Forage Science 69, 206-228

Skovsen S., Dyrmann M., Mortensen A. K., Steen K. A., Green O., Eriksen J., Gislum R., Jørgensen R. N. and Karstoft H. (2017) Estimation of the botanical composition of clover-grass leys from RGB images using data simulation and fully convolutional neural networks. Sensors (Basel, Switzerland), 17(12). https://doi.org/10.3390/s17122930

Søegaard K. (2009) Nitrogen fertilization of grass/clover swards under cutting or grazing by dairy cows. Acta Agriculturae Scandinavica Section B-Soil and Plant Science 59, 139-150