Predicting forage provision of grasslands across climate zones by hyperspectral measurements

Männer F.A.[1,2,3]; Muro, J.[4,5]; Ferner, J.[4]; Schmidtlein, S.[6]; Linstädter, A.[3]
[1] Fraunhofer Institute for Computer Graphics Research IGD, Competence Center Smart Farming, Rostock, Germany
[2] University of Bonn, Institute of Crop Science and Resource Conservation, Bonn, Germany
[3] University of Potsdam, Biodiversity Research/Systematic Botany, Potsdam, Germany
[4] University of Bonn, Centre for Remote Sensing of Land Surfaces ZFL, Bonn, Germany
[5] University of Bonn, Remote Sensing Research Group, Bonn, Germany
[6] Karlsruhe Institute of Technology, Institute for Geography and Geoecology, Karlsruhe, Germany

Keywords: Metabolizable energy; forage biomass; remote sensing; machine learning; neural networks

Abstract

The potential of grasslands' fodder production is a crucial management measure, while its quantification is still laborious and costly. Remote sensing technologies, such as hyperspectral field measurements, enable fast and non-destructive estimation. However, such methods are still limited in transferability to other locations or climatic conditions. With this study, we aim to predict forage nutritive value, quantity, and energy yield from hyperspectral canopy reflections of grasslands across three climate zones. We took hyperspectral measurements with a field spectrometer from grassland canopies in temperate, tropical and semi-arid grasslands, and analyzed corresponding biomass samples for their quantity (BM), metabolizable energy content (ME) and metabolizable energy yield (MEY). Three machine learning algorithms were used to establish prediction models for single and across climate regions. The normalized root mean squared error (nRMSE) for ME, BM and MEY varied between 0.12 – 0.19, 0.14 – 0.21, and 0.15 – 0.21, respectively. The ME trans-climatic model showed the best accuracy compared to the local models. Trans-climatic model predictions of climate-specific data, decrease in accuracy to 0.16 – 0.21, 0.17 – 0.24, and 0.19 – 0.28 for ME, BM and MEY compared to predictions with climate-specific models. Trans-climatic models with feed-forward neural networks showed similar performance for ME but higher accuracies for BM and MEY predictions. The trans-climatic models generally showed good performance for forage nutritive value and forage provision. Our results suggest that models based on hyperspectral measurements offer great potential to assess or even map the forage nutritive value of grasslands across climate zones.

Introduction

The world's grassland ecosystems provide a wide range of ecosystem services, with the provision of forage - both in terms of nutritive value and quantity - being among the most important. In particular, the potential of biomass productivity, metabolizable energy content, and yield are critical for sustainable and profitable grassland management. However, this potential is often unknown, and quantifying forage nutritive value and quantity remains expensive and time-consuming, typically requiring laboratory analysis of biomass samples. This is where remote sensing technologies, such as hyperspectral sensors, are becoming increasingly important as they provide rapid and non-destructive measurements (Ferner et al., 2015). Hyperspectral modelling approaches have been used to estimate, for example, metabolizable energy content and biomass quantity for West African savanna grasslands (Ferner et al., 2018; Ferner et al., 2021), nutritional value for South African savanna grasslands (Singh et al., 2017), and temperate ryegrass canopies (Smith et al., 2020). Nevertheless, these forage supply predictions have limited applicability to other sites or even climatic conditions. With this study, we aim to fill this gap by predicting biomass production (BM), metabolizable energy (ME) content, and metabolizable energy yield (MEY) from hyperspectral, remotely sensed canopy reflectance of grassland communities in three different climatic zones on the African and European continents. Here we aim to (1) investigate the accuracy of hyperspectral-based prediction models for ME, BM and MEY for a temperate, tropical and semi-arid climate, and a combined trans-climatic model, (2) test different machine learning algorithms (partial least squares, random forest and neural networks), and (3) validate the trans-climatic models against the three local models using the same validation dataset to investigate the transferability of a trans-climatic model for ME, BM, and MEY.

Methods

Our study sites included grasslands in temperate, subtropical, and tropical climates. Sites were located in (a) subtropical to tropical grasslands in the Sudanese savannas of West Africa (Ferner et al., 2015; Guuroh et al., 2018), (b) subtropical grasslands in the semi-arid thornbush savannas of Namibia, and (c) temperate central European meadows and pastures within the three Biodiversity Exploratory sites in northeastern, central, and southwestern Germany (Fischer et al., 2010). We made 456 hyperspectral measurements using full-range field spectrometers (177 from Germany, 105 from West Africa, and 174 from Namibia). Samples of aboveground biomass were collected from all measured plots at the time of spectrometer measurements by cutting the herbaceous vegetation of a 60 x 60 cm quare at stubble height (3 cm). The amount of dry matter per m² (BM) was determined by weighing after 48 hours of oven drying at 55°C. The dried samples were also analyzed for metabolizable energy (ME) content as a proxy for forage nutritive value using the procedure of Menke and Steingass (1988). From the values of BM and ME, the amount of metabolizable energy per m² was calculated as the so-called metabolizable energy yield (MEY). The spectral signatures of the hyperspectral measurements were smoothened and corrected for atmospheric dynamic artefacts. For each spectrum, we took the raw spectra and determined the first derivative, various vegetation indices (VIs), and absorption features from the entire spectrum as potential predictors of the ME, BM, and MEY forage provision measures. Partial least squares (PLS), random forest (RF) and a feed-forward neural network (NN) were used to build predictive models for each climatic zone (climate-specific), and for the three zones combined (trans-climatic) with a repeated k-fold cross-validation. Repeated backward selection according to the importance of the predictors was used to reduce the dimension of the 644 predictors with RF, while PLS reduced the dimension of the predictors by creating latent vectors from the entire predictor set and the NN used the entire set. We also split the local datasets into a training and a validation dataset at 80% and 20%, respectively, to calibrate and validate the trans-climatic models against the local models by using the same validation datasets for both the local and trans-climatic models. The different model accuracies were assessed and compared by calculating the normalized root mean square error (nRMSE) of the validation datasets.

Results and Discussion

Random Forest and Partial Least Squares Regression

Interestingly, the trans-climatic models achieved generally similar or better accuracies (nRMSE ~0.12) than the climate-specific models for ME, BM, and MEY, which could be due to greater variation in the calibration data (Fig. 1).

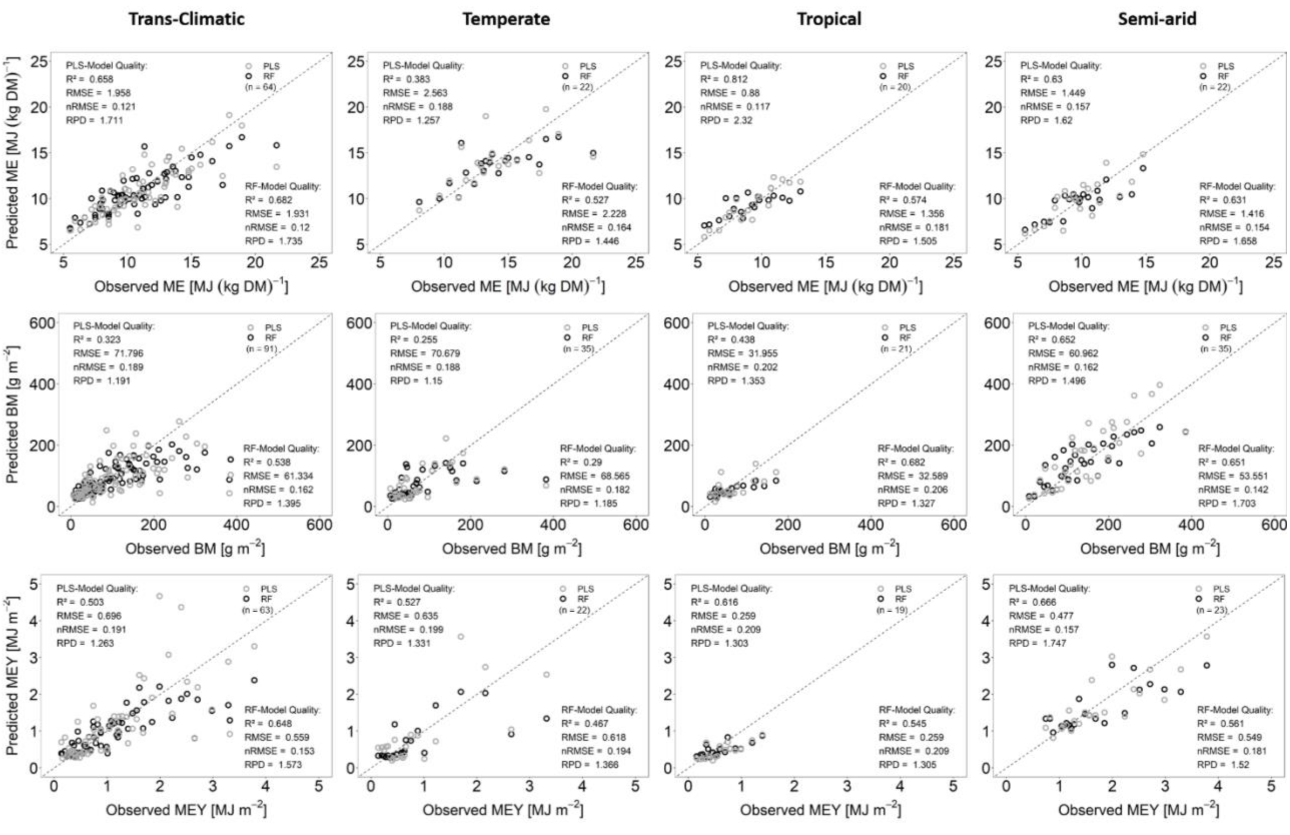

Fig. 1: Predicted vs observed metabolizable energy (ME), biomass (BM) and metabolizable energy yield (MEY) for climate-specific (temperate, tropical and semi-arid) and trans-climatic models partial least squares (PLS) and random forest (RF) regressions. Model accuracies are given by the coefficient of determination (R²), root mean squares error (RMSE), normalized RMSE (nRMSE), and the ratio of performance to deviation (RPD).

The first derivative of the canopy reflectance signature proved to be the most appropriate predictor for ME and BM prediction, while MEY was best predicted by a combination of the first derivative, absorption features, and vegetation indices. We did not find a clear advantage of PLS or RF, as both strategies lead to similar accuracies with minor differences. In general, we found a slight underestimation of high ME, BM, and MEY values, which could be an indication of overfitting (ME) or saturation of the models due to a closed canopy (BM). This saturation effect at a given vegetation density is well-known and remains a limitation of optical sensors (Wachendorf et al., 2018). Predicting standing biomass from hyperspectral canopy measurements proved difficult, as models appear to saturate at about 200 g m^{-2} for temperate, tropical, and transclimate models. This may be explained by the effect of nearly closed canopies in temperate and tropical grasslands at about 200 g dry mass per m^2, which is not the case for the more sparse vegetation in semi-arid savannas. Comparing the trans-climatic models to the local models by predicting the same data with both models, the accuracy of the trans-climatic models is slightly lower than that of the local models with an nRMSE of 0.16-0.21 compared to 0.12-0.19 for ME, 0.17-0.24 compared to 0.14-0.21 for BM, and 0.19-0.28 compared to 0.16-0.21 for MEY. Consequently, using the trans-climatic models to predict data from only one climate is still possible with reasonable, but somewhat lower, accuracy.

Neural Networks
Initially, NN models for the trans-climatic data showed similar accuracies for the ME, BM, and MEY predictions, with an nRMSE of 0.13 (Fig. 2). However, we also found a slight underestimation of higher BM and MEY values, which can be explained by model saturation due to a closed canopy at higher grassland vegetation biomass levels as described above.

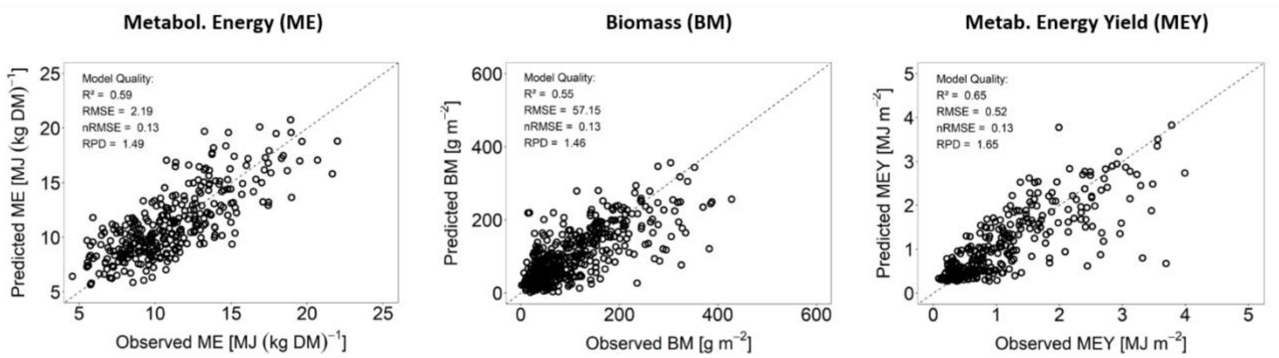

Fig. 2: Predicted vs observed metabolizable energy (ME, n = 318), biomass (BM, n = 456) and metabolizable energy yield (MEY, n = 316) prediction models for trans-climatic datasets using feed-forward neural networks (NN). Model accuracies are given by the coefficient of determination (R^2), root mean squares error (RMSE), normalized RMSE (nRMSE), and the ratio of performance to deviation (RPD).

The nRMSE of the local models for temperate, tropical and semi-arid areas varied between 0.18 and 0.22 for ME, 0.14 and 0.15 for BM, and 0.17 and 0.18 for MEY predictions. Similar to the RF and PLS models, the trans-climatic prediction is more accurate than the local models, which can be an effect of both broader variation, or a larger number of training observations, as neural networks are known to require more data than RF or PLS models do. The trans-climatic NN models, however, show overall better accuracy (nRMSE = 0.13) than the trans-climatic RF or PLS models (0.12 < nRMSE < 0.19).

Conclusions and Implications
Hyperspectral forage supply models performed well in both local and trans-climatic applications in temperate, tropical, and semiarid climates. Limitations can occur with saturation effects due to closed grassland canopies, especially for temperate or tropical grassland vegetation. As the size and variation of the calibration dataset increase, the models – especially deep-learning strategies – may even improve or be transferable to other parts of the world. Hyperspectral models thus offer great potential for agricultural or ecological applications to assess or even map forage nutritive value and quantity of grasslands worldwide, also in light of the recently launched EnMAP satellite. This will contribute to better-informed management of grasslands and rangelands and maintain or improve their agricultural and ecological value.

Acknowledgements

This work was partly funded by the DFG Priority Program 1374 within the framework of the SeBAS project (LI 1842/4-1 & DU 1596/1-1) and by the German Ministry of Education and Research (BMBF) through the NamTip project (01LC1821A), and through WASCAL (01LG1202A).

References

Ferner, J., Linstädter, A., Rogass, C., Südekum, K.-H. and Schmidtlein, S. 2021. *Towards Forage Resource Monitoring in subtropical Savanna Grasslands: going multispectral or hyperspectral?* European Journal of Remote Sensing, 54(1), 364–384.

Ferner, J., Linstädter, A., Südekum, K.-H. and Schmidtlein, S. 2015. *Spectral indicators of forage quality in West Africa's tropical savannas*. International Journal of Applied Earth Observation and Geoinformation, 41(September), 99–106.

Ferner, J., Schmidtlein, S., Guuroh, R.T., Lopatin, J. and Linstädter, A. 2018. *Disentangling effects of climate and land-use change on West African drylands' forage supply*. Global Environmental Change, 53, 24–38.

Fischer, M., Bossdorf, O., Gockel, S., Hänsel, F., Hemp, A. and Hessenmöller, D. 2010. *Implementing large-scale and long-term functional biodiversity research: The Biodiversity Exploratories*. Basic and Applied Ecology, 11(6), 473–485.

Guuroh, R.T., Ruppert, J.C., Ferner, J., Čanak, K., Schmidtlein, S. and Linstädter, A. 2018. *Drivers of forage provision and erosion control in West African savannas - A macroecological perspective.* Agriculture, Ecosystems & Environment, 251, 257–267.

Menke, K.H. and Steingass, H. 1988. *Estimation of the energetic feed value obtained from chemical analysis and in vitro gas production using rumen fluid.* Animal Research and Development, 28, 7–55.

Singh, L., Mutanga, O., Mafongoya, P. and Peerbhay, K. 2017. *Remote sensing of key grassland nutrients using hyperspectral techniques in KwaZulu-Natal, South Africa*. Journal of Applied Remote Sensing, 11(3), 36005.

Smith, C., Karunaratne, S., Badenhorst, P., Cogan, N., Spangenberg, G. and Smith, K. 2020. *Machine Learning Algorithms to Predict Forage Nutritive Value of In Situ Perennial Ryegrass Plants Using Hyperspectral Canopy Reflectance Data.* Remote Sensing, 12(6), 928.

Wachendorf, M., Fricke, T. and Möckel, T. 2018. *Remote sensing as a tool to assess botanical composition, structure, quantity and quality of temperate grasslands.* Grass and Forage Science, 73(1), pp. 1–14.