Methods for spatially explicit estimation of NATURA 2000 grassland forage quality using satellites

Raab C.[1,4], Riesch, F.[1,4], Tonn, B.[1], Meißner, M.[2], Balkenhol, N.[3,4] and Isselstein, J.[1,4]
[1]Grassland Science, Department for Crop Sciences, University of Goettingen, Von-Siebold-Str. 8, 37075 Goettingen, Germany; [2]Institut für Wildbiologie Göttingen und Dresden e.V., Buesgenweg 3, 37077 Goettingen, Germany; [3]Faculty of Forest Sciences and Forest Ecology, Wildlife Sciences, University of Goettingen, Buesgenweg 3, 37077 Goettingen, Germany; [4]Centre of Biodiversity and sustainable Land Use (CBL), University of Goettingen, Grisebachstr. 6, 37077 Goettingen, Germany

Abstract

Spatially explicit mapping of grassland forage quality is of major interest for sustainable grazing management of NATURA 2000 areas, especially if those are large or have limited accessibility. Therefore, this study is concerned with the estimation of crude protein (CP) and organic acid detergent fiber (oADF) content at regional scale using Sentinel-2 and Landsat 8 remote sensing data. Field data were collected in the Grafenwoehr military training area in Bavaria, Germany. Different combinations of predictor variables were applied using cross-validated random forest regression, linear regression with lasso penalty and linear regression with ridge penalty models. The rede-dge band of Sentinel-2, centered at 705 nm, as well as the shortwave infrared bands of both sensors and related vegetation indices contributed the most to the respective models. Linear regression with lasso penalty and Sentinel-2 data performed consistently better, compared to the other models. The results (CP (10.1 - 23.1 %): max R^2 0.53, RMSE 1.78 %; oADF (22.7 - 39.5 %): max R^2 0.72, RMSE 2.3 %) demonstrate the potential of remote sensing as an information tool supporting the conservation management of grassland areas with limited access.

Keywords: remote sensing, regression, forage quality, wildlife grazing, NATURA 2000

Introduction

Regional mapping of grassland forage quality is of major interest for sustainable grazing management of NATURA 2000 areas, especially if they are large or have limited accessibility. The availability and quality of potential forage areas affects the spatial distribution and activities of herbivores (Merkle *et al.*, 2016) and thus their influence on the ecosystem by e.g. grazing, trampling and nutrient cycling (Fløjgaard *et al.*, 2017). Spatially continuous information about the quality of grassland can be seen as pivotal to understanding the spatial and temporal behavior, as well as the sustainable management and conservation of wildlife, such as red deer (*Cervus elaphus*). Therefore, this study is concerned with the prediction of crude protein (CP) and organic acid detergent fiber (exclusive of residual ash, oADF) by remote sensing and field data using different combinations of predictor variables and regression models. We hypothesise that (1) a relationship between CP, oADF and spectral information obtained by remote sensing exists; (2) a relevant subset of predictor variables can be identified to be sufficient for predicting grassland forage quality and (3) a higher spectral resolution is superior over a higher spatial resolution (3).

Materials and methods

The study was conducted on the military training area Grafenwoehr (GTA) in Bavaria, Germany (49° 40' 56" N, 11° 47' 20" E; about 230 km²). Roughly 85 % of the GTA are part of the NATURA 2000 network and contain numerous rare and highly protected habitat types (Warren & Büttner, 2008). To map the grass CP and oADF content at regional scale (about 200 ha) 72 grass samples were collected (16 - 18 May 2017). At each sampling location, grass was clipped within a 10 m radius (simulated grazing by hand plucking) and dried at 60 °C for 48 h. Subsequently we measured all samples for CP and oADF content using near infrared spectroscopy (NIRS). Access to the study sites was restricted due to military use and to avoid wildlife disturbance. Landsat 8 (path = 193, row = 25) and Sentinel-2 (T32UPA) surface reflectance data, acquired on 17 May 2017, were obtained from the EarthExplorer and Sentinel Data Hub, respectively. In addition to the multi-spectral bands (MS), a list of commonly used vegetation indices (VI) (Sentinel-2 20 m: n = 89; Landsat8 30 m: n = 40) and texture (TXT) measures (n = 28 for each) were included. We included Sentinel-2 data resampled to 30 m as a third data set. Different combinations of standardised predictor variables (MS+VIs, MS+TXT and MS+VIs+TXT) were applied using linear regression with lasso penalty (LRL), linear regression with ridge penalty (LRR) (lambda optimised using cross-validation) and random forest regression (RF) (ntree = 1000, mtry = p/3) models. All models were cross-validated (k = 5) with 100 repetitions and variable importance was calculated as mean decrease of RMSE (100 permutations) using the *sperrorest* package implemented in R (v3.4.1; R Core Team 2017).

Results and discussion

The range of NIRS results for CP and oADF was: 10.1 - 23.1 % (average 15.9 % and *sd* (standard deviation) 2.6 %), 22.7 - 39.5 % (average 29.7 % and *sd* 4.5 %), respectively. On average, the LRL (mean R^2: CP = 0.42, *sd* = 0.08; oADF = 0.64, *sd* = 0.064) models were superior over LRR (mean R^2: CP = 0.34, *sd* = 0.13; oADF = 0.57, *sd* = 0.12) and RF (mean R^2: CP = 0.26, *sd* = 0.08; oADF = 0.51, *sd* = 0.15). The moderately high R^2 values support hypothesis (1). As illustrated in Figure 1, the best model performance for CP was reached using the Sentinel-2 20 m data, combining MS and VI predictor variables (RMSE: 1.78, *sd* = 0.06). For oADF, the Sentinel-2 30 m data with the same variable combination performed best (RMSE: 2.3, *sd* = 0.06). In addition, the SR green divided by rede-dge (Sentinel-2 band 5, at 705 nm), the NMDI (Normalized Multi-band Drought Index) and the PSRI (Plant Senescence Reflectance Index) can be seen as important contributers to the model performance (Figure 2). Confirming hypothesis (2), the derived variable importances in Figure 2 indicate the strong relevance of the SWIR (shortwave infrared) bands for the prediction of CP and oADF. The presented results indicate the strength of LRL in selecting the most important variables by penalising, and consequently disregarding, the less relevant predictors. As reported by Verrelst *et al.* (2015), only few studies concerned with the prediction of bio-physical parameters using LRL exist, even though LRL seems to perform better than other regression methods in a high-dimensional data context (Zandler *et al.*, 2015). The Sentinel-2 sensor with a higher spectral resolution showed better performance results compared to Landsat 8, confirming hypothesis (3). This was supported by the derived variable importances (Figure 2), which underlined the importance of the rede-dge bands for mapping CP and oADF. The SWIR-associated variables can be seen as essential, as they can be related to the moisture content of soil and vegetation, which is associated with vegetation health and, accordingly, CP and oADF content.

Conclusions

Remote-sensing-based regression models have potential for mapping grassland forage quality, and can therefore support the conservation management and monitoring of grassland areas with limited access. The high spectral resolution of Sentinel-2, especially the red-edge and SWIR bands can be seen as crucial for good performance. Future research should investigate the potential of pan-sharpened Sentinel-2 data and LRL-based forage quality estimation.

References

Fløjgaard, C., De Barba M., Taberlet P. and Ejrnæs R. (2017) Body Condition, Diet and Ecosystem Function of Red Deer (*Cervus Elaphus*) in a Fenced Nature Reserve. *Global Ecology and Conservation* 11, 312-323.
Merkle J. A., Monteith K. L., Aikens E. O., Hayes M. M., Hersey K. R., Middleton D., Oates B. A., Sawyer H., Scurlock B. M. and Kauffman M. J. (2016) Large Herbivores Surf Waves of Green-up during Spring. In: *Proc. R. Soc. B* (Vol. 283, No. 1833, p. 20160456). The Royal Society.
Verrelst J., Camps-Valls G., Muñoz-Marí J., Rivera J. P., Veroustraete F., Clevers J. G.P.W. and Moreno J. (2015) Optical Remote Sensing and the Retrieval of Terrestrial Vegetation Bio-Geophysical Properties-A Review. *ISPRS Journal of Photogrammetry and Remote Sensing* 108, 273-290.
Warren S. D. and Büttner R. (2008) Active Military Training Areas as Refugia for Disturbance-Dependent Endangered Insects. *Journal of Insect Conservation* 12 (6), 671-76.
Zandler H., Brenning A. and Samimi C. (2015) Quantifying Dwarf Shrub Biomass in an Arid Environment: Comparing Empirical Methods in a High Dimensional Setting. *Remote Sensing of Environment* 158, 140-155.

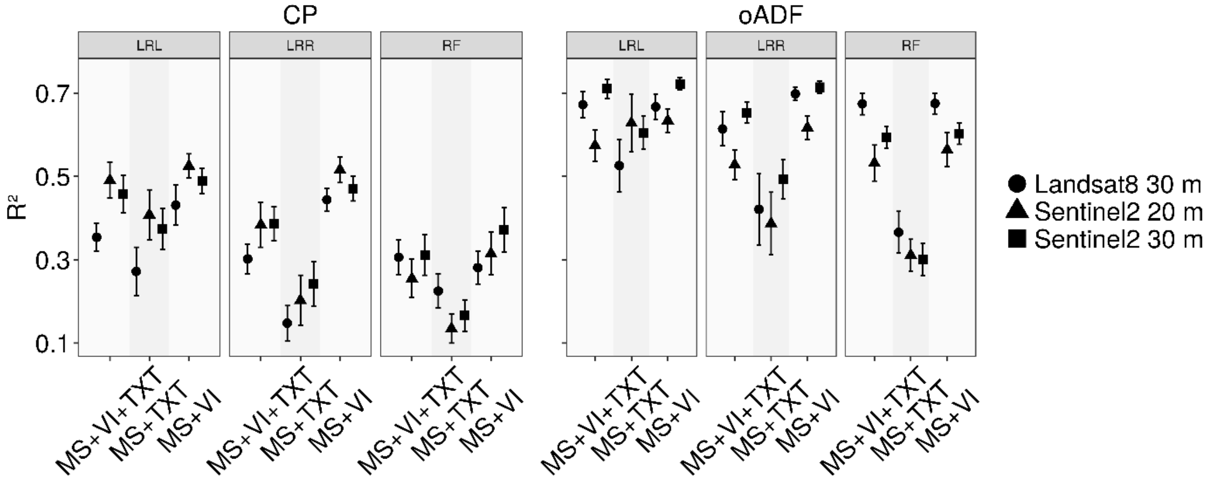

Figure 1. Cross-validated (100 repetitions) R^2 values for the estimation of CP and oADF using different predictor combinations (MS = multi-spectral, VI = Vegetation indices, TXT = texture) and regression models (LRL = linear regression with lasso penalty, LRR = linear regression with ridge penalty, RF = random forest regression).

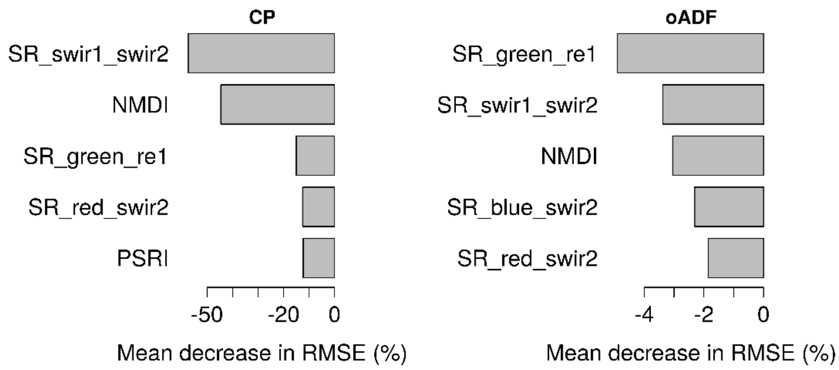

Figure 2: Permutation-based variable importance for CP (Sentinel-2 20m, MS+VI) and oADF (Sentinel-2 30m, MS+VI). SR = simple ratio, NMDI = Normalized Multi-band Drought Index, PSRI = Plant Senescence Reflectance Index, re1 = Sentinel-2 band 5, at 705 nm.