Mapping forest leaf dry matter content from hyperspectral data

Abebe Mohammed Ali [a, b], Andrew K. Skidmore[a], Roshanak Darvishzadeh[a], Iris van Duren[a], Stefanie Holzwarth[c] and Joerg Mueller[d]

[a]Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands–ali31165@itc.nl

[b]Department of Geography and environmental studies, Wollo University, P.O.Box 1145, Dessie, Ethiopia

[c]German Aerospace Center (DLR), German Remote Sensing Data Center (DFD) Oberpfaffenhofen, 82234 Wessling, Germany

[d]Bavarian Forest National Park, 94481 Grafenau, Germany

Abstract: Leaf dry matter content (LDMC) is a central vegetation property that plays an important role in assessments of ecosystem functions. In this study, LDMC was estimated from hyperspectral airborne image by inversion of the INFORM radiative transfer model using Continuous Wavelet Analysis (CWA). Stand parameters were collected for 33 sample plots during a field campaign in July 2013 in the Bavarian Forest National Park, Germany. The INFORM model was used to simulate the canopy reflectance of the study area and was then inverted by applying CWA in the shortwave infrared region. The results were evaluated using R^2 and RMSE of the estimated and measured LDMC. Our results revealed significant correlations of six wavelet features with LDMC. The wavelet feature at 1741 nm (scale 5) was the strongly correlated feature in the studied spectral region to LDMC variation. The combination of all the identified

wavelet features for LDMC gave the most accurate prediction (R^2= 0.59 and RMSE= 4.39%).

1. Introduction

Leaf dry matter content (LDMC) is one of the key plant functional traits, which is the ratio of leaf dry mass to leaf fresh mass. it is a proxy for relative growth rate and carbon assimilation, and is an important predictor of a plant's location on an axis of resource capture, usage and availability (Wilson *et al.* 1999). LDMC is also used to estimate other traits and ecological indicators, such as leaf thickness (Vile *et al.* 2005), leaf life span (Marenco *et al.* 2009), relative growth rate (Shipley 2006), and soil fertility (Hodgson *et al.* 2011). Generally, the quantitative information and spatial distribution of LDMC improves our understanding of and capacity to investigate community structure and ecosystem functioning (Lavorel *et al.* 2011). However, this trait is currently quantified through labor-intensive methods of field sampling.

Remote sensing, as a relatively fast and efficient approach for estimating LDMC across a wide range of spatial and temporal scales, has so far received little attention. Until recently, none of the remote sensing techniques has been tested for direct estimation of LDMC. A leaf scale study using PROSPECT model inversion reported the potential of remote sensing for quantifying LDMC (Ali *et al.* 2016). Because of redundancy and multi-collinearly in hyperspectral data (Blackburn 2007a), RTM inversions are often applied on selected bands sensitive to a given vegetation variable. Although a subset of spectral bands proved to be a stable and accurate predictor for vegetation parameters (Weiss *et al.* 2000), no general criteria have been formulated for the selection of bands (Banskota

et al. 2013). Wavelet transformation seems to be a promising alternative technique for selecting the most informative features from hyperspectral data.

Wavelet analysis enables spectral data to be transformed into a new representation by decomposing the original spectra into various scales (frequencies). Subsequently, correlation between the concentration of parameters and the wavelet scales can be ascertained, in order to identify the most sensitive spectral feature for predicting a given parameter. Previous studies have investigated the potential of wavelet analysis in estimating leaf parameters from leaf spectra measured in the laboratory and simulated data using RTMs (e.g., Jingcheng *et al.* 2011). However, the applicability of wavelet analysis at canopy level using canopy spectra obtained from airborne and spaceborne hyperspectral data has not yet been investigated in detail. Our study aimed at quantifying and mapping LDMC of a mountain forest from hyperspectral data by inverting the INFORM radiative transform model using an optimized predictive model constructed from wavelet coefficients.

2. Methodology

2.1. Study area and ground truth (field) data collection

The study area for this study was the mixed mountain forest of the Bavarian Forest National Park. The park is located in south-eastern Germany along the border with the Czech Republic (49^0 3' 19" N, 13^0 12' 9" E). Elevation of the study area varies from 600m to 1,473m above sea level. The climate of the region is temperate, with high annual precipitation (1,200 mm to 1,800 mm) and low average annual temperature (3^0 to 6^0 Celsius). The soils in the area are naturally acidic and low in nutrient content. The

natural forest ecosystems of the National Park vary with altitude: there are alluvial spruce forests in the valleys, mixed mountain forests on the hillsides and mountain spruce forests in the high areas. The dominant tree species include European beech (*Fagus sylvatica*), Norway spruce (*Picea abies*) and Fir (*Abies alba*).

Table 1

Summary statistics of the measured and estimated leaf and canopy parameters for the 33 sample plots in the study area: leaf mass per leaf area (C_m), leaf water content (C_w), leaf dry matter content (LDMC), leaf area index (LAI), stem density (SD), canopy closure (CC), crown diameter (CD) and stand height (SH).

Parameter	C_m (g/cm^2)	C_w (g/cm^2)	LDMC (g/g)	LAI (m^2/m^2)	SD (n/ha)	CC (%)	CD (m)	SH (m)
Minimum	0.0061	0.0071	0.3999	2.42	222	38	2.91	12.26
Maximum	0.0292	0.0309	0.5075	6.18	1722	91	10.55	27.36
Mean	0.0147	0.0178	0.4534	4.3	778.4	75.19	5.67	20.23
St. dev.	0.0059	0.0071	0.0254	0.81	405.5	5.33	1.56	4.52

A field campaign was conducted during summer 2013 to collect field measurements from 33 plots. Each plot was square, with sides 30m long. In all 33 sample plots, forest structural variables such as LAI, stem density (SD), canopy closure (CC), crown diameter (CD), and stand height (SH) were measured. Moreover in each plot, leaf samples were collected from mature sunlit leaves at the top of the canopy and their characteristics were measured (n=130). For details on the leaf samples' physical variable measurements, see (Ali *et al.* 2016). Leaf samples were oven dried at 65^0 C for

48 hours and then LDMC, leaf mass per leaf area (C_m) and leaf water content (C_w) were computed (Table 1).

2.2. Image acquisition and pre-processing

HySpex is a new airborne hyperspectral sensor developed by the Norwegian Norsk Elektro Optikk (NEO) company. It comprises two imaging spectrometers with spectral ranges of 400–1000 nm and 1000–2500 nm and up to 416 spectral channels. It records radiance data in contiguous bands at a spectral resolution of 3.7 nm for 400-992 nm spectral range (sensor 1) and 6 nm for 968-2498 spectral range (sensor 2). Its spatial resolution is 1.6 meters for sensor 1 and 3.4 meters for sensor 2. The instrument was flown over the study area onboard a Cessna 208B Grand Caravan at average altitudes of 3006.5 m above sea level on July 22 2013 between 9:00 and 11:15 local time. The HySpex image data were supplied by the DLR team after atmospheric correction performed with the ATCOR4 model, orthorectified and georeferenced using standard aircraft in-flight information. As only the spectral bands in the SWIR were utilized in this study, only the images from sensor 2 of HySpex were mosaicked and resampled. The average reflectance of the sample plots was extracted and used for evaluation. The noisy bands in the water absorption region (1345-1450nm and 1790–1980 nm) and bands from 2450–2498 nm were assigned zero values. This left a total of 203 bands with valid reflectance values.

2.3. RTM parameterization and continuous wavelet analysis

To simulate the spectral property of the study area we used the Invertible Forest Reflectance model “INFORM” (Atzberger 2000, Schlerf and Atzberger 2006). In

INFORM, LAI is represented by the leaf area indices of single trees. Hence, the ground truth values for LAIs were computed from LAI and CC.

$$\boldsymbol{LAI_s} = \frac{\boldsymbol{LAI}}{\boldsymbol{CC}} \quad (1)$$

And for every combination of input parameters, LDMC were indirectly calculated from C_m and C_w as:

$$\boldsymbol{LDMC} = \frac{\boldsymbol{C_m}}{\boldsymbol{C_m + C_w}} \quad (2)$$

where LAI_s is single tree leaf area index, CC is canopy closure; C_m is leaf dry mass and and C_w is leaf water content per leaf area. See Ali *et al.* (2016) for further details.

The INFORM model was run by generating the input parameters (C_m, C_w, N, LAI_s, SD, SH, CD and ALA) using a multivariate normal distribution function based on the mean and covariance matrix of their ground truth values (Table 1). Leaf chlorophyll content was fixed at an average value of 40µg/cm^2. A sensitivity study had previously reported insignificant effect of solar zenith and azimuth angles on INFORM simulated canopy reflectance (Ali *et al.* (2015). Therefore, other leaf, canopy, and external input parameters were fixed at average values based on the priori field knowledge and HySpex image specifications. The field spectra of understory vegetation and the forest floor elements were averaged and used as a fixed background reflectance during the simulation.

Wavelet features that significantly correlated with LDMC were determined in four steps as described in Cheng *et al.* (2011). A continuous wavelet transform was applied on the simulated and measured (HySpex) spectra in order to represent them in the wavelet domain where the wavelet power was a function of the wavelength and the scale. The

coefficient of determination (R^2)between the LDMC and simulated spectra wavelet power at each wavelength and scale location were calculated. The wavelet amplitude was correlated to log-transformed LDMC values. Wavelet features with large R^2 values imply high sensitivity to LDMC. The features of correlation scalogram were ranked and a threshold value of 1% was applied to delineate and define the feature regions most sensitive to LDMC. The features with a central wavelength within each region of the combined scalogram were selected as predictor variables The predictive performance of wavelet features was evaluated using the coefficient of determination (R^2) and root mean square error in percent (RMSE %). The land cover maps were used to mask out the non-forest areas in the HySpex image. The continuous wavelet transformation was applied to extract sensitive wavelet features for each pixel in the HySpex image and the predictive model developed earlier was applied for pixel-by-pixel estimation of LDMC and producing the final map of LDMC in Bavaria forest national park.

3. Results

3.1. Wavelet analysis and identification of wavelet features sensitive to LDMC

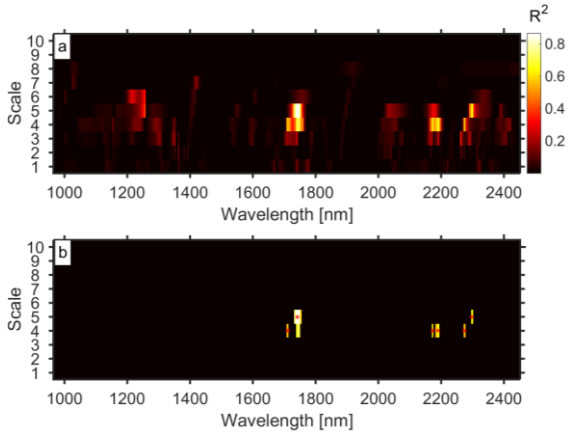

Figure 1. Correlation scalograms for the identification of wavelet features which significantly correlated with leaf dry matter content (LDMC) (a). Scalograms are derived from continuous wavelet analysis of simulated spectra. Brightness represents the coefficient of determination

(R^2) relating wavelet power to LDMC. Colored feature regions in scalograms (b) depict the wavelet features with the top 1% greatest R^2 values.

Figure 1 shows the sensitivity of wavelet features (which are transformed from the simulated spectra) plotted for LDMC. After a number of tests, the top 1 % strongly correlated wavelet features were found to be good predictors. There were six most sensitive wavelet features for LDMC. The wavelet features selected were at scales 4 and 5, (Figures 1b). Prediction and validation of LDMC from wavelet features

The stepwise linear model and quadratic regression were selected for their goodness of fit to our calibration dataset. The regression models were tested for each wavelet feature separately and for different combinations of wavelet features. Figure 2 depicts the prediction capacity of one of the many wavelet features selected and its corresponding wavelength reflectance for LDMC estimation. The overlaid measured spectra fall within the range of the simulated spectra. Directly correlating LDMC with simulated spectra without any wavelet transformation performed less satisfactorily than correlating it with wavelet power. The wavelet feature centered at 2191 nm yielded an R^2 = 0.64 of linear correlation with LDMC (Figure 2).

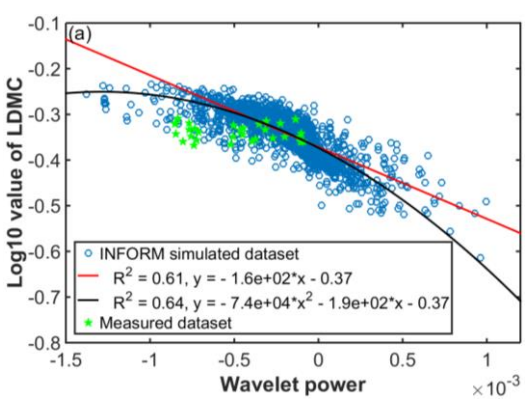

Figure 2. Relationships between the wavelet features and reflectance of the Calibration dataset for LDMC: the relationship between the logarithmic value of LDMC and the wavelet power at feature 2191 nm, scale 4, For comparison, the measured HySpex dataset (validation dataset) are shown as pentagrams.

Table 2

Coefficients of determination (R^2) and root mean square error (RMSE%) between the logarithmic values of LDMC and predictions made using wavelet features derived from the validation dataset (HySpex image).

Spectral (wavelet) feature	R^2	RMSE (%)
A. Combination of All	**0.59**	**4.39**
B. (1741 nm scale 5)	0.39	23.10
C. (2191 nm scale 4)	0.34	24.29
D. (2173 nm scale 4)	0.33	17.26
E. (2299 nm scale 5)	0.2	20.08
F. (1711 nm scale 4)	0.30	21.27
G. 2275 nm scale 4	0.28	20.82

Thus, predictions made using selected wavelet features and their combination revealed high R^2 and low RMSE (Table 2). The R^2 values range from 0.2 to 0.59 for LDMC. The single wavelet features that showed the highest correlation and the lowest NRMSE were 2191 nm at scale 4.

The comparison of predicted values obtained from the best combinations of wavelet features against the HySpex data after converting logarithmic values to normal values are presented in figure 3. In this figure predicted values were computed by applying models developed from the calibration dataset using six wavelet features combinations. As can be observed from the figure the prediction is precise (RMSE = 4.39 %). There is

a tendency for LDMC to be overestimated, especially for higher values of LDMC. Nevertheless, the predicted values were scattered closely around the 1:1 relationship line, which indicates the sensitivity of the selected wavelet features in capturing the variation in LDMC concentrations. Furthermore, no saturation problems were observed in the predicted values.

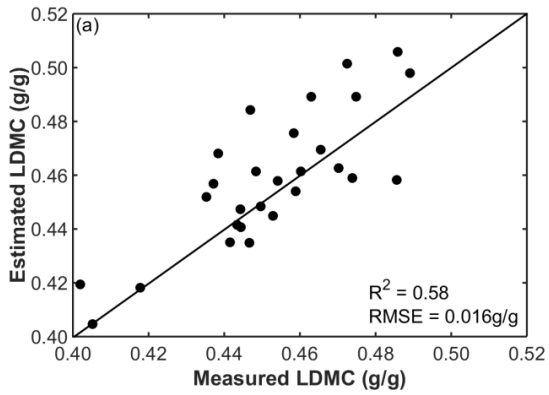

Figure.3. Scatter plots of measured and predicted LDMC All data points are the measured dataset extracted from HySpex airborne image. The solid line shows the 1:1 relationship between the predicted and measured data.

3.2. Mapping LDMC of the study area

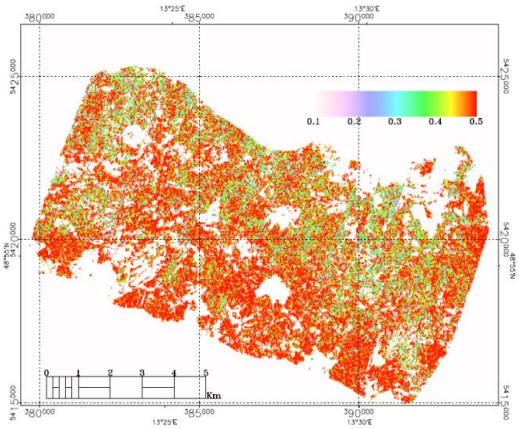

Figure.4. LDMC in g/g map derived from the HySpex imagery of July 22, 2013. The map is based on a predictive model developed on all six selected wavelet features.

The concentration of LDMC across the study area is presented in Figures 4. It shows perceptible variability across the study area. The mean obtained for all image pixels were 0.4235, which is close to the mean of the samples measured during the field

measurements shown in Table 1. A comparison with a forest type map of the study area revealed that LDMC values were higher for conifer and mixed stands than for deciduous forests (not shown)

4. Discussion and conclusion

In this study, the CWA approach was used for the retrieval of LDMC from HySpex hyperspectral reflectance data. One advantage of CWA over other approaches is that it has the potential to identify the most sensitive spectral features from large hyperspectral dataset. The second advantage of using CWA is transformation of the original spectra that resulted in better correlation of variables. Decomposition of INFORM simulated spectra and measured HySpex spectral reflectance using CWA (R2 = 0.59) provided higher correlations with LDMC than original reflectance.

After wavelet transformation, parameters which seemed to be uncorrelated to simulated canopy reflectance (without transformation), were observed to be correlated (Figures 3). This may be attributed to the effectiveness of CWA in decomposing the traits' absorption features into various scales of narrow and broad band absorption features and identifying those that correlate most with the variation in the traits' concentration. The performance of wavelet analysis compared with narrow-band indices and stepwise selection of narrow-band reflectance for retrieval of pigment concentrations in vegetation at leaf and canopy scales has also been reported by Blackburn (2007b). By comparison with narrow-band indices, wavelet analysis captures more information contained within the hyperspectral data and creates an opportunity to develop robust and extendible

methods for quantifying plant traits over extended areas (Blackburn 2007a). We determined six wavelet features for LDMC in a mixed mountain forest (Table 2).

As expected, the measured HySpex spectra overlapped exactly with the simulated spectra without transformation. However, despite the addition of 0.3% random Gaussian noise to the simulated spectra, systematic shifts to higher or lower values were observed when wavelet transformation was applied to the measured spectra (Figures 3). This in turn led to systematic overestimation and underestimation (Figure 4). Probable causes for this shift could be atmospheric effects and sensor noise on the measured spectra. These factors may cause variation in local absorption peaks on the measured spectra, and lead to higher or lower wavelet power values during the transformation.

References

Ali, A. M., R. Darvishzadeh, A. K. Skidmore and I. V. Duren, 2015. Effects of Canopy Structural Variables on Retrieval of Leaf Dry Matter Content and Specific Leaf Area From Remotely Sensed Data*,Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

Ali, A. M., R. Darvishzadeh, A. K. Skidmore, I. v. Duren, U. Heiden*, et al.*, 2016. Estimating leaf functional traits by inversion of PROSPECT: Assessing leaf dry matter content and specific leaf area in mixed mountainous forest*,International Journal of Applied Earth Observation and Geoinformation* 45, Part A: 66-76.

Atzberger, C., 2000. Development of an invertible forest reflectance model:The INFOR-Model*,In Buchroithner (Ed.), A decade of trans-European remote sensing*

cooperation: Proceedings of the 20th EARSeL Symposium Dresden, Germany, 14- 16. June 2000 (pp. 2039- 2044).

Banskota, A., R. H. Wynne, V. A. Thomas, S. P. Serbin, N. Kayastha*, et al.*, 2013. Investigating the Utility of Wavelet Transforms for Inverting a 3-D Radiative Transfer Model Using Hyperspectral Data to Retrieve Forest LAI*,Remote Sensing* 5(6): 2639-2659.

Blackburn, G. A., 2007a. Hyperspectral remote sensing of plant pigments*,Journal of Experimental Botany* 58(4): 855-867.

Blackburn, G. A., 2007b. Wavelet decomposition of hyperspectral data: A novel approach to quantifying pigment concentrations in vegetation*,International Journal of Remote Sensing* 28(12): 2831-2855.

Cheng, T., B. Rivard and A. Sánchez-Azofeifa, 2011. Spectroscopic determination of leaf water content using continuous wavelet analysis*,Remote Sensing of Environment* 115(2): 659-670.

Hodgson, J. G., G. Montserrat-Marti, M. Charles, G. Jones, P. Wilson*, et al.*, 2011. Is leaf dry matter content a better predictor of soil fertility than specific leaf area?*,Annals of Botany* 108(7): 1337-1345.

Jingcheng, Z., L. Juhua, H. Wenjiang and W. Jihua, 2011. Continuous wavelet analysis based spectral feature selection for winter wheat yellow rust detection*,Intelligent Automation and Soft Computing* 17(5): 531-540.

Lavorel, S., K. Grigulis, P. Lamarque, M. Colace, D. Garden*, et al.*, 2011. Using plant functional traits to understand the landscape distribution of multiple ecosystem services*,Journal of Ecology* 99(1): 135-147.

Marenco, R. A., S. A. Antezana-Vera and H. C. S. Nascimento, 2009. Relationship between specific leaf area, leaf thickness, leaf water content and SPAD-502 readings in six Amazonian tree species,*Photosynthetica* 47(2): 184-190.

Schlerf, M. and C. Atzberger, 2006. Inversion of a forest reflectance model to estimate structural canopy variables from hyperspectral remote sensing data,*Remote Sensing of Environment* 100(3): 281-294.

Shipley, B., 2006. Net assimilation rate, specific leaf area and leaf mass ratio: which is most closely correlated with relative growth rate? A meta-analysis,*Functional Ecology* 20(4): 565-574.

Vile, D., É. Garnier, B. Shipley, G. Laurent, M.-L. Navas, *et al.*, 2005. Specific Leaf Area and Dry Matter Content Estimate Thickness in Laminar Leaves,*Annals of Botany* 96(6): 1129-1136.

Weiss, M., F. Baret, R. B. Myneni, A. Pragnere and Y. Knyazikhin, 2000. Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data,*Agronomie* 20(1): 3-22.

Wilson, P. J., K. E. N. Thompson and J. G. Hodgson, 1999. Specific leaf area and leaf dry matter content as alternative predictors of plant strategies,*New Phytologist* 143(1): 155-162.