Assessing the sensitivity of multi-frequency passive microwave vegetation optical depth to vegetation properties

Luisa Schmidt[1], Matthias Forkel[1], Ruxandra-Maria Zotta[2], Samuel Scherrer[2], Wouter A. Dorigo[2], Alexander Kuhn-Régnier[3,4], Robin van der Schalie[5], Marta Yebra[6,7]

[1] Technische Universität Dresden, Institute of Photogrammetry and Remote Sensing, 01069 Dresden, Germany
[2] Technische Universität Wien, Department of Geodesy and Geoinformation, Vienna, Austria
[3] Leverhulme Centre for Wildfires, Environment, and Society, London, SW7 2AZ, UK
[4] Department of Physics, Imperial College London, London, SW7 2AZ, UK
[5] Planet, Wilhelminastraat 43A, 2011 VK Haarlem, The Netherlands
[6] Fenner School of Environment & Society, Australian National University, ACT 2601, Australia
[7] School of Engineering, Australian National University, ACT 2601, Australia

Correspondence to: Luisa Schmidt (luisa.schmidt1@tu-dresden.de)

Abstract. Vegetation attenuates the microwave emission from the land surface. The strength of this attenuation is quantified in models in terms of the parameter Vegetation Optical Depth (VOD), and is influenced by the vegetation mass, structure, water content, and observation wavelength. Earth observation ~~satellites~~satellite sensors operating in the microwave frequencies are used for global VOD retrievals, enabling the monitoring of vegetation ~~status~~ at large scales. VOD has been used to determine above-ground biomass, monitor phenology or estimate vegetation water status. VOD can be also used for constraining land surface models or modelling wildfires at large scale. Several VOD products exist differing by frequency/wavelength, sensor, and retrieval algorithm. Numerous studies present correlations or empirical functions between different VOD datasets and vegetation variables such as normalised difference vegetation index, leaf area index, gross primary production, biomass, vegetation height or vegetation water content. However, an assessment of the joint impact of land cover, vegetation biomass, leaf area, and moisture status on the VOD signal is challenging and has not yet been done.

This study aims to interpret the VOD signal as a multi-variate function of several descriptive vegetation variables. The results will help to select ~~certain~~ VOD ~~wavelengths~~at the most suitable wavelength for specific applications and can guide the development of appropriate observation operators to integrate VOD with large-scale land surface models. Here we use VOD from the Land Parameter Retrieval Model (LPRM) ~~of~~in Ku-, X- and C-bands ~~of~~from the harmonised VODCA dataset and ~~level 3~~ L-band VOD derived from SMOS and SMAP sensors. ~~Within a multivariable regression random forest model for simulating these VOD signals, leaf~~Leaf area index, live-fuel moisture content, above-ground biomass, and land cover are able to explain up to ~~0.~~93% and 95% of the variance (Nash-Sutcliffe model efficiency coefficient~~ of determination).~~) in 8-aily and monthly VOD within a multivariable random forest regression. Thereby, the ~~variance in~~ regression reproduces spatial patterns of L-band VOD ~~is reproduced spatially~~ and ~~for~~spatial and temporal patterns of Ku-, X- and C-band VOD~~ spatially as well as temporally.~~. Analyses of accumulated local effects demonstrate that Ku-, X- and C-band VOD ~~is~~are mostly sensitive to leaf area index and L-band VOD to above-ground biomass. However, for all VODs the global relationships with vegetation properties are non-monotonic and complex and differ with land cover type. This indicates that the use of simple global regressions to estimate single vegetation properties (e.g. above-ground biomass) from VOD is over-simplistic.

1 Introduction

Vegetation Optical Depth (VOD) describes the attenuation ~~by the vegetation layer~~ of microwave radiation ~~emitted by~~in the ~~Earth.~~vegetation layer. Quantifying this attenuation effect is important for an accurate retrieval of surface soil moisture from passive microwave satellite observations (Wang, 1985; Njoku and Entekhabi, 1996). In the radiative transfer equation for microwave emissions, the opacity of the vegetation layer (i.e. the VOD) is also commonly referred to as τ (Jackson et al., 1982). VOD can be retrieved e.g. from the passive microwave radiative transfer equation using measurements of passive microwaves (Jackson and Schmugge, 1991; Owe et al., 2008; Sawada et al., 2016). ~~However,~~However, VOD is a parameter in these microwave radiative transfer models for vegetation and hence it is not directly measurable and verifiable with in-situ measurements~~, which is why~~. Therefore, different authors have correlated VOD with different vegetation properties to understand ~~what VOD is sensitive to.~~the sensitivity of VOD to vegetation properties (Jones et al., 2011; Rodríguez-Fernández et al., 2018; Konings et al., 2019a). Generally, the opacity of passive microwaves in the vegetation layer increases with increasing vegetation water content but this relationship varies with vegetation structure including leaf and woody components and wavelength (Jackson and Schmugge, 1991; Wigneron et al., 1993; Njoku and Entekhabi, 1996). Based on radiometer measurements over various crops and a wide range of wavelengths (0.8 – 30 cm), Jackson and Schmugge (1991) report a clear linear relationship of VOD to vegetation water content (VWC):

$$VOD = b * VWC, \tag{1}$$

~~where the parameter b depends on vegetation type and wavelength. The authors find that b exponentially decreases with increasing wavelength, which implies that vegetation opacity (the VOD) is smaller for longer wavelengths (i.e. L-band) than for shorter wavelengths (i.e. Ku-, X- and C-bands). The vegetation water content can also be expressed as a product of above-ground biomass (AGB) and a relative water content parameter, often referred to as live-fuel moisture content (LFMC) (Konings et al., 2019):~~where the parameter b depends on vegetation type and wavelength. The authors find that b exponentially decreases with increasing wavelength, which implies that vegetation opacity (the VOD) is smaller for longer wavelengths (i.e. L-band) than for shorter wavelengths (i.e. Ku-, X- and C-bands). The parameter b is usually kept constant which might be insufficient due to its possible dependency on polarization. In addition, neglecting surface soil roughness can lead to an underestimation of VOD, especially when the vegetation does not completely cover the ground (Togliatti et al., 2022).

The vegetation water content can also be expressed as a product of above-ground biomass (AGB) and a relative water content parameter, often referred to as live-fuel moisture content (LFMC) (Konings et al., 2019b):

$$VOD = b * AGB * LFMC. \tag{2}$$

Based on ~~those~~these relationships, many studies use VOD to estimate AGB or other vegetation properties. For example, Liu et al. (2015) use Ku-band VOD to estimate long-term changes in global AGB, finding a gain of above-ground biomass carbon considering forest and non-forest vegetation for 1993-2012. ~~(Rodríguez-Fernández et al. (2018)~~Rodríguez-Fernández et al. (2018) correlate spatial patterns in AGB and yearly averaged values of L-band VOD from the Soil Moisture and Ocean Salinity (SMOS) mission with the INRA-CESBIO algorithm (SMOS-IC) for Africa with correlation coefficients up to 0.85. They find linear relationships between VOD and AGB within single land cover classes, but the relationship across land cover classes is shown to be nonlinear, with a weaker nonlinearity for L-band VOD compared to Ku-/X-/C-band VOD. Chaparro et al. (2018) use L-band from the Soil Moisture Active Passive mission (SMOS) derived with the Multi-Temporal Dual Channel Algorithm (MT-DCA) to determine crop biomass of the north-~~center of the~~central USA. Both Rodríguez-Fernández et al. (2018) and Chaparro et al. (2018) find better results for pixels with higher homogeneity~~, not just for~~ in land cover types ~~but~~or even~~ for~~ plant types, implying that relationships between VOD and vegetation properties change with land cover and plant types. Li et al. (2021b) find high correlation of L-band VOD and AGB leading to the conclusion that longwave VOD is more sensitive to woody parts of the vegetation than shortwave VOD. However, Konings et al. (2021) show that the relation between L-band VOD and AGB dominates in space but that short-term temporal dynamics in VOD are dominated by VWC. As a proxy for vegetation water status, VOD can be related to LFMC or VWC or both ~~(Fan et al., 2018; Konings et al., 2019; Frappart et al., 2020)~~(Fan et al., 2018; Konings et al., 2019b; Frappart et al., 2020) and can be used to estimate leaf water potential (Konings and Gentine, 2017; Momen et al., 2017; Zhang et al., 2019).

Furthermore, VOD is frequently compared with other vegetation properties such as canopy greenness, leaf area index (LAI), or plant productivity. For example, VOD shows similar temporal patterns to normalised difference vegetation index (NDVI) and LAI (Liu et al., 2011; Momen et al., 2017; Bousquet et al., 2021). In spatial comparisons, the vegetation indices and variables tend to saturate over densely vegetated areas. This saturation is less distinct for VOD (Rodríguez-Fernández et al., 2018) due to the ability of microwaves to penetrate deeper into the vegetation layer. Therefore, VOD ~~can therefore provide~~provides complementary information to the usually visible-infrared based metrics~~, e.g. for assessments of land surface phenology~~ (Jones et al., 2011). ~~Metrics~~For example, metrics sensitive to biomass or water content shifts can be derived from VOD (Jones et al., 2011, 2014). VOD can also be used for assessing land surface phenology (Jones et al. 2011). VOD and temporal changes in VOD are also correlated with gross primary production (GPP) (Teubner et al., 2018), which ~~allowed~~allows VOD to ~~develop~~be used as a ~~sink-driven~~predictor for GPP ~~estimation approach based on VOD~~ (Teubner et al., 2019, 2021; Wild et al., 2022).

Recently, several new VOD datasets became available for X-band from the Advanced Microwave Scanning Radiometer – Earth Observing System sensor (AMSR-E) and Advanced Microwave Scanning Radiometer 2 (AMSR2) sensors (Du et al., 2017; Wang et al., 2021), in L-band from SMOS (van der Schalie et al., 2016; Fernandez-Moran et al., 2017; Al Bitar et al., 2017; Wigneron et al., 2018, 2021) and SMAP (Konings et al., 2017)~~, or Ku-, X- and C-band~~. VOD was also retrieved jointly from several sensors (van der Schalie et al., 2017) ~~as well as~~and harmonized long-term multi-sensor datasets ~~(~~have been produced (e.g. Vegetation Optical Depth Climate Archive VODCA, Moesinger et al., 2020). A recent comparison study by Li et al. (2021) of different X-, C- and L-band VOD datasets and Moderate Resolution Imaging Spectroradiometer (MODIS) derived vegetation indices like NDVI and enhanced vegetation index (EVI) as well as tree height and AGB showed that X-band VOD is more suitable to detect temporal variations of the green vegetation parts, especially for less densely vegetated areas, than C- and L-band VOD. Additionally, Li et al. (2021) as well as Moesinger et al. (2022) found time lags between VOD and vegetation indices and climate variables, ~~showing~~which are not yet fully understood. This shows the need to include ~~a~~ further ecological parameters or vegetation ~~property to improve the~~variables which could account for a delayed response of VOD to temporal changes in the vegetation ~~status. To~~indices. Approaches with the ability to ~~take~~ consider ~~into account~~ VOD variations caused by vegetation water content have been developed, which are more complex than simple regression functions ~~have been developed~~ (e.g. Momen et al., 2017). Momen et al. (2017) were able to estimate VOD by using two predictors, LAI and leaf water potential. ~~Teubner et al. (2019) linked VOD and GPP by using generalized additive models and the differential equation between VOD and AGB by~~Among others, the studies by Momen et al. (2017) and ~~Liu et al. (2015). Among others, these two studies~~Teubner et al. (2019) show that the water content of the vegetation is not only influencing the relation between vegetation indices and VOD but also the relation between VOD and AGB.

The increasing availability of VOD data for vegetation studies also increases the possibilities to assimilate or integrate VOD with ecosystem or land surface models (LSM) (Scholze et al., 2019; Kumar et al., 2020). Therefore, observation operators are needed that link the modelled vegetation properties with the satellite-retrieved VOD. Scholze et al. (2019) use ~~a~~the sum of an empirical AGB function and a linear term for LAI to describe annual SMOS-IC L-band VOD within the Carbon Cycle Data Assimilation System (CCDAS) for estimating European carbon fluxes. Kumar et al. (2020) use CDF matching to convert VODCA X- and C-band VOD, and SMAP L-band VOD to LAI, which is then assimilated into the Noah-MP LSM. X- and L-band VOD showed partially complementary improvements of the modelled land surface variables. Both studies by Scholze et al. (2019) and Kumar et al. (2020) find an improvement of the model results by incorporating passive microwave data, demonstrating the benefits of the vegetation information contained in VOD. In another model-data-fusion approach, Liu et al., 2021 use VOD to derive plant hydraulic parameters for a soil-plant system model that accounts for the hydraulic state of the vegetation explicitly. However, as VOD reflects both dynamics in biomass and water content (Jackson and Schmugge, 1991; Konings et al., 2021), relations between VOD and AGB or LAI as observation operators are simplifications and demonstrate the need for a more detailed understanding of the effects of vegetation properties on VOD.

The increasing use of VOD for ecosystem studies (e.g. ~~Dorigo, 2021)~~Dorigo et al., 2021) and land surface modelling poses the question how different vegetation properties affect VOD in both time and space. Hence, a more detailed investigation of the relative effects of vegetation properties on VOD could improve the understanding of the VOD signal in terms of interpretation of the corresponding vegetation status. Such investigations will also help to identify a suitable VOD dataset for a specific ecological application~~.~~ in addition to the technical aspects of the datasets like the observation resolution depending on wavelength, errors and artefacts induced by the retrieval algorithm or the observation time depending on overpass times of the satellites. Furthermore, due to the high temporal resolution and temporal coverage of VOD datasets (partly since 1987), global analyses of vegetation properties and status as well as land cover change can be conducted for enhanced understanding of long-term environmental changes and to improve model predictions.

Here we aim to assess VOD in response to multiple vegetation properties at large (i.e. inter-continental) scales. ~~We apply a multi-variate framework~~Specifically, our objectives are to predict VOD from LFMC, LAI and AGB by using two machine learning ~~approaches (random forest,~~ regression approaches and to investigate the relationship between VOD and the predictors. This objective goes beyond previous empirical studies that compared VOD with vegetation properties based on bivariate correlations or regressions but not by estimating VOD within a multivariate framework.

We use random forests (RF) and generalized additive ~~model,~~ models (GAM) to ~~quantify sensitivities of VOD to and interactions with~~ predict VOD from LFMC, LAI, AGB, and land cover. ~~Generalized additive models and random forests are used to predict VOD from vegetation properties and accumulated~~Accumulated local effect (ALE) curves are used to assess the sensitivities of VOD to these properties. ~~Comparing~~While GAM is suitable to capture non-linear and non-monotonic relationships with additive effects of the predictors, RF can predict more complex interactions but is less suitable to capture a possible additive behavior. Therefore, comparing both machine learning algorithms gives insights into the structure of the relationship between VOD and vegetation properties and provides confidence in the findings. Additionally, we inspect how different temporal resolutions (i.e. 8-daily and monthly data) affect the relationships between VOD and vegetation properties for identifying the role of vegetation variables at quasi-weekly and seasonal time scales. The analyses are carried out for five VOD datasets, which differ in wavelength but were derived with the same algorithm (Land Parameter Retrieval Model, LPRM) (van der Schalie et al., 2016; van der Schalie et al., 2017) to exclude differences due to retrieval algorithms.

2 Data and methods

2.1 Datasets

2.1.1 VOD data

An overview of the datasets is given in Table 1 and Figure 1~~. All used VOD datasets are derived from passive sensors using the LPRM algorithm . .~~ . All used VOD datasets are derived from passive sensors using the LPRM algorithm (van der Schalie et al., 2016) to reduce the degrees of freedom of this analysis.

The VODCA dataset (Moesinger et al., 2020) provides harmonised long-term records of shortwave VOD for Ku-, X- and C-band (further named Ku-VOD, X-VOD and C-VOD, respectively), using data from the AMSR-E, AMSR-2, Special Sensor Microwave Imager (SSM/I), TRMM Microwave Imager (TMI), and Windsat sensors. Unfortunately, Ku-VOD is only available until 1st August 2017 due to a bias in the ~~eleven brightness temperatures of~~ AMSR-2 Ku-band VOD causing unexpected low values of the VOD retrievals after this date (Moesinger et al., 2020), which is not fixed in the version 01.0. Therefore, all datasets are analysed until 31st July 2017.

~~Two LPRM-derived L-band VOD datasets are used as longwave VOD, one sensed with SMAP, the other with SMOS (Schalie et al., 2016, further named as SMAP L-VOD and SMOS L-VOD, respectively). The SMAP satellite was launched in January 2015, and therefore SMAP L-VOD defines the start date of the analysis of all datasets.~~Two LPRM-derived L-band VOD datasets are used as longwave VOD, one sensed with SMAP, the other with SMOS (van der Schalie et al., 2016, further named as SMAP L-VOD and SMOS L-VOD, respectively). The SMAP satellite was launched in January 2015, and therefore SMAP L-VOD defines the start date of the analysis of all datasets.

All VOD datasets are provided as daily data with a spatial resolution of 0.25° on a global scale. As VOD generally decreases with increasing wavelength, the five VOD datasets have different dynamic ranges. As we are not interested in the absolute value but only the temporal dynamics and spatial patterns, the VOD datasets were globally normalised using minimum and maximum value to a range of 0 to 1 based on the available global data within the time span 2015-2017 to provide comparability. For normalising we use the scikit-learn function 'MinMaxScaler'. The normalised VOD data form the basis for the subplots d)-h) of Figure 1. These maps of temporal averaged VOD data show different patterns and scales even after the normalisation process. This illustrates that VOD data derived from different wavelengths and sensors are not related to the same vegetation properties inducing the need for this study.

2.1.2 Predictor data

Following the relationship between VOD, LFMC, and AGB as shown in ~~equation~~Equation 2, proxies related to biomass (AGB and LAI), water content (LFMC), and the structure parameter (plant types) are used as predictors for VOD.

As proxies for woody and non-woody biomass, we used a map of AGB and a time series of LAI. The ESA CCI AGB map (Santoro and Cartus, 2019) for the year 2017 with 100 m spatial resolution is used as a predictor for woody biomass~~.~~. This AGB map describes the oven-dry mass of woody parts of living trees per pixel. Thereby only above-ground mass is considered, i.e. stem and bark as well as twigs and branches, but not stumps and roots.

LAI is used as a proxy for canopy biomass. Specifically, we use the MOD15A2H version 6 dataset from MODIS, which is available at 500 m spatial and on a 8-daily temporal resolution on a global scale (Myneni et al., 2015). We excluded LAI retrievals under (partial) cloud cover, snow or high solar zenith angle.

For ~~live fuel moisture content,~~ LFMC, we used a product derived from MODIS MCD43A2 Collection 6 reflectance data for ~~three regions (California/~~western ~~US,~~USA (derived from the MODIS tiles h08v04, h08v05 and h09v04), South Africa and Australia (Figure 1 b) at a 500 m spatial and on a 4-daily temporal resolution using the approach described in Yebra et al. (2018). ~~Additionally~~Yebra et al. (2018) use three radiative transfer models (RTM) for the simulation of spectra corresponding to different LFMC values. More specifically, they use PROSPECT 1 coupled to SAILH 1 and GeoSail to simulate the spectra of grasslands/ shrublands and forest, respectively. Based on these simulations three different look-up tables (LUT) were generated. For a given location they use the MODIS land cover product (MCD12Q1 Collection 5) to select the LUT corresponding to the specific fuel type characterising that location. That fuel specific LUT is used to invert the RTM and retrieve LFMC from the MODIS spectra. The results were evaluated with LFMC field measurements and the model achieved an explained variance of 58% and a RMSE of 40% for Australia (Yebra et al., 2018). For Europe, we used ~~a European~~the LFMC product ~~computed~~produced by the European Union Joint Research Centre (JRC) ~~for~~and which is included in the European Forest Fire Information System (EFFIS). This product ~~uses also methods described in~~follows the same methodology as Yebra et al. (2018) but uses EFFIS's fuel type map to select the LUT and MODIS MCD43A2 Collection 5 data to invert the RTM before 2016. Therefore, ~~it is~~for those years, the LFMC estimates are produced with a temporal resolution of 8 days.

The LAI, LFMC, as well as the AGB datasets were resampled to 0.25° resolution to match the VOD spatial extent using a first order conservative remapping.

We used the land cover map by the European Space Agency (ESA) Climate Change Initiative (CCI, ~~ESA, 2017)~~ESA, 2017) and its continuation from the Copernicus Climate Change Service which provide yearly data for the period 1992-2018 at 300 m spatial resolution. The land cover classes were converted to fractions of plant functional types and aggregated to 0.25° spatial resolution using the cross-walking approach as described in Poulter et al. (2015). Specifically, we made use of the fractions per 0.25° grid cell of broad-leaved evergreen (treeBE), needle-leaved evergreen (treeNE), deciduous trees (treeD), shrublands (shrub), croplands (crop), and herbaceous vegetation (herb). Deciduous trees were not further segregated into broad- and needle-leaved trees as especially the latter would result in only a small sample when intersected with the VOD data. In another test, we also combined the fractional coverage of all tree PFTs (treeAll = treeBE + treeNE + treeD) and of short vegetation (short = shrub + herb + crop).

Table 1: Overview of the used datasets ~~used~~ and their original technical attributes.

Dataset	Variable and unit	Sensors	Temporal coverage / resolution	Spatial coverage / resolution	Reference
VODCA v01.0	Ku-VOD (-)	AMSR-2, SSMI/I, TMI and Windsat scaled to AMSR-E	1987-2017 / daily	Global / 0.25°	Moesinger et al. (2020)
	X-VOD (-)*		1997-2018 / daily		
	C-VOD (-)**		2002-2018 / daily		
SMAP L-VOD	L-VOD (-)	SMAP radiometer	2015-2019 / daily		van der Schalie et al. (2016)
SMOS L-VOD	L-VOD (-)	MIRAS	2010-2020 / daily		
ESA CCI AGB v1.0	AGB (Mg/ha)	PALSAR-2, Sentinel-1 (1A and 1B), Landsat	2017 / representative for one year	Global / 100 x 100 m	Santoro and Cartus (2019)
MOD15A2H v006	LAI (-)	MODIS sensors	2000-2020 / 8-daily	Global / 500 x 500 m	Myneni et al.(2015)
MODIS-LFMC	LFMC (%)	MODIS sensors	2000-2019 / 4-daily	Regional / 500 x 500 m ~~California~~western USA, South Africa, Australia	Yebra et al. (2018)
			8-daily	Europe	
ESA CCI Land cover v2.0.7	Plant functional types (PFT) derived from land cover classes	AVHRR, PROBA-V, Envisat MERIS, SPOT-VGT	1992-2018 / yearly	Global / 300 x 300 m	~~ESA (2017)~~ESA (2017)

* does not contain SSM/I ** does not contain SSM/I and TMI

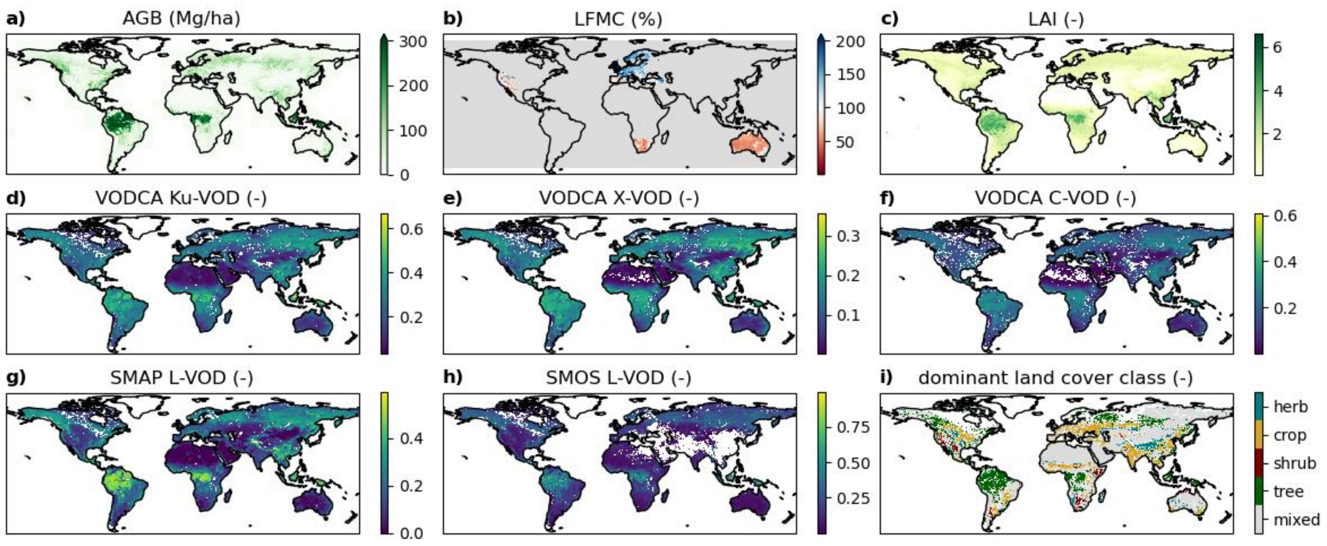


Figure 1: Overview of the datasets used a) Above-ground biomass (AGB) for 2017 based on the ESA CCI biomass dataset, b) Live Fuel Moisture Content (LFMC) derived from MODIS whereby grey indicates areas of non-available data, c) Leaf Area Index (LAI) derived from MODIS, d) Ku-band VOD (Ku-VOD) from VODCA, e) X-band VOD (X-VOD) from VODCA, f) C-band VOD (C-

VOD) from VODCA, g) L-band VOD from SMAP (SMAP L-VOD), h) L-band VOD from SMOS (SMOS L-VOD), and i) the dominant land cover class for 2016 based on the ESA CCI land cover map. LAI, LFMC, and VOD maps are temporal averages over the period January 2015-July 2017~~.~~, whereby the VOD maps are based on data scaled to 0-1 for the available range within the mentioned timespan. Note that LFMC is only available for ~~California, Southern~~western USA, South Africa, Europe, and Australia.

2.1.3 Data combination

All datasets were cropped to the extent of the LFMC data (Australia, Europe, ~~California~~western USA, South Africa) for further analyses. This implies that the 'global' models as stated in the following are indeed inter-continental models restricted to the spatial extent of the LFMC dataset. To provide comparability of the analyses of the different VOD datasets, only the overlapping timespan is used (January 2015-July 2017). The rather short time period does not impede the framework of this study, because instead of analysing coherent pixel time series this approach uses each time step of each pixel as an individual data point. The ESA CCI AGB map represents the year 2017, but we assume that the biomass does not dramatically change over two years. Therefore, the AGB values are kept constant for the whole time series. The PFT fractions are taken from the annual land cover maps for the respective years in 2015 to 2017 without any interpolation. During the analyses, models were trained and tested for 8-daily and monthly temporal resolutions of the LAI and LFMC time series. For the 8-daily resolution, only the VOD values matching the same timestamp of the MODIS LAI and LFMC products are used. For the monthly resolution, the mean VOD, LAI or LFMC within the regarding month were calculated.

As a final step, pixels were excluded when the fractional coverage of bare ground or water exceeds 5 % to avoid the interpretation of marginal effects of bare soils or water on VOD. Models were specifically trained for single land cover classes. A threshold of 55 % was used to discern when a land cover class was dominant compared to the other classes.

2.2 Regression methods

~~To assess the influence of the vegetation variables on VOD, we applied two methods: generalized additive models (GAM) and random forest regressor (RF). Both methods are compared to evaluate if the relationship between the features and the predictor variable is rather simple additive assembled (adequately captured by GAM) or more complex (requires RF). GAM can represent non-linear and non-monotonic relations with single predictors whereby all predictors have a joined additive effect.~~To assess the influence of the vegetation variables on VOD, we applied two methods: generalized additive models (GAM) and random forest regressor (RF). ~~RF can represent more complex relations and interactions between the single predictors, but are not well suited for capturing additive structures in the data (Hastie et al., 2009). Another reason to use GAM simultaneously to RF is that models that are designed for short vegetation use just two predictors (LAI and LFMC). AGB is only representative for woody biomass of trees and can therefore not be included for short vegetation. While GAMs can utilise a small number of predictors, the application of RF with only two predictors will likely result in overfitting as the random choice of a predictor variable during the development of decision trees is very limited. Both methods allow the qualitative and quantitative assessment of the sensitivities of VOD to the predictors via Accumulated Local Effects (ALE, see chapter 2.5).~~

The RF algorithm incorporates multiple independent decision trees, where the final prediction is the average prediction of the individual trees (Breiman, 2001; Hutengs and Vohland, 2016; Liang et al., 2018). Using the scikit-learn package ~~(~~version 24.1~~)~~, (Pedregosa et al., 2011) multiple hyper-parameters can be tuned~~. During~~, which will define the RF model structure. The optimization of the hyper-parameter combination is crucial to achieve a well performing model. The scikit-learn package provides a grid-search ~~using the scikit-learn function 'RandomizedSearchCV'~~function 'RandomizedSearchCV' which enables for an automatized search for an optimized parameter set by splitting the multi-variate space of the hyper-parameters into a grid of parameter combinations which are then used to train a RF. During this grid-search for an exemplary dataset (predicting monthly inter-continental Ku-VOD with LAI, LFMC, AGB, and land cover), the minimum number of samples within a leaf (1 and 4), number of estimators (100, 200-2000 with 200-steps), maximum features (functions: 'auto', 'sqrt', 'log2'), maximal depth (10-110 with 20-steps, None), and minimum samples split (2 and 10) were tested. For a detailed description of the available hyper-parameters and their effect on the result please refer to the documentation of the scikit-learn module 'sklearn.ensemble.RandomForestRegressor' (https://scikit-learn.org/0.24/modules/generated/sklearn.ensemble.RandomForestRegressor.html). The best combinations were again tested with monthly inter-continental predictions of X-, C-, SMOS and SMAP L-VOD. Some combinations led to partly improved results compared to the scikit-learn default hyper-parameters, but also partly degraded results. We finally selected the following hyper-parameters: minimum samples within a leaf=1, number of estimators=100, maximum features='auto', maximal depth=None, minimum samples split=2 and criterion=mean squared error. This setup provided the best results across all tested models ~~the best results.~~. The chosen maximum features parameter leads to the consideration of all features for all splits, thereby omitting one of the strengths of RF. This parameter may have been selected due to the low number of our chosen vegetation variables. However, RF is still able to capture complex relationships, which is our main focus.

~~GAMs~~GAM are a progression of standard linear regression models and generalized linear models (~~GLMs~~GLM) (Hastie and Tibshirani, 1987). In comparison to standard linear regression models, ~~GLMs~~GLM use a link function to connect the mean response of the target variable with the predictors, which can also represent other distributions of the target variable besides the Gaussian distribution, like binomial, gamma or Poisson distributions (Nelder and Wedderburn, 1972). In addition, ~~GAMs~~GAM incorporate smoothing functions for each predictor variable (Yee and Mitchell, 1991). This allows modeling non-linear and non-parametric relationships between the target and predictor variables. A general GAM equation can be written as:

$$g(\mu) = \cancel{b}b + \sum_{i=1}^{p} f_j(x_i) \,, \qquad (3)$$

with $g()$ as link function, μ as mean response of target variable, b as intercept term, $f()$ as smoothing functions, and x as predictor variables. Thereby, $g(\mu)$ represents the target variable, i.e. predicted VOD data, and $f(x_i)$ the predictors, i.e. the vegetation variables LAI, AGB, LFMC and land cover expressed as PFT data sets. Here the GAM is developed for a Gaussian distribution with an 'identity' link function and spline terms as smoothing functions using the Python package pyGAM ~~(~~version 0.8.0~~).~~ (Servén et al., 2018).

Both methods are compared to evaluate if the relationship between the features and the target variable is additive (adequately captured by GAM) or more complex (requires RF). GAM can represent non-linear and non-monotonic relations with single predictors whereby all predictors have a joint additive effect. RF can represent more complex relations and interactions between the single predictors, but are not well suited for capturing additive structures in the data (Hastie et al., 2009). Another reason to use GAM simultaneously to RF is that models that are designed for short vegetation use just two predictors (LAI and LFMC). The AGB dataset is only representative for woody biomass of trees and can therefore not be included for short vegetation. While GAM can utilise a small number of predictors, the application of RF with only two predictors will likely result in overfitting as the random choice of a predictor variable during the development of decision trees is very limited. Both methods allow the qualitative and quantitative assessment of the sensitivities of VOD to the predictors via Accumulated Local Effects (ALE, see chapter 2.5).

2.3 Model experiments

We applied two main classes of regression models to predict VOD. The first class are global models that use the PFTs from the land cover map in addition to the vegetation predictors LAI, LFMC, and AGB. This means that the individual maps of treeBE, treeBD, treeNE, treeND, shrub, crop, and herb are used as additional predictors. The second model class is comprised of land cover–specific models using LAI, LFMC, and AGB as inputs. These models are only applied to the spatial extent of one dominant land cover class. In models for short vegetation classes, AGB is not used as a predictor because this map is only representative of forest biomass. All model setups were trained both for GAM and RF, and using monthly as well as 8-daily values for each VOD dataset. Table 2 gives an overview of the models and the input data.

Table 2: List of tested models, with N = needleleaf, B = broadleaf, E = evergreen, D = deciduous, All = not differentiated, CCI PFT = ESA Climate Change Initiative Plant Functional Type; each model is run with GAM and RF as well as with datasets with 8-daily and monthly temporal resolution for each VOD dataset. The land cover-specific models are only trained and tested within a cross validation for pixels which are dominated by certain land cover (threshold PFT fraction > 0.55).

Land cover class/ Model name	Spatial domain (defined by dominant land cover)	Predictors
Land cover-specific models		
treeAll	CCI PFT treeAll > 55%	AGB + LFMC + LAI
treeNE	CCI PFT treeNE> 55%	AGB + LFMC + LAI
treeBE	CCI PFT treeBE > 55%	AGB + LFMC + LAI
treeB	CCI PFT (treeBE + treeBD) > 55%	AGB + LFMC + LAI
treeN	CCI PFT (treeNE + treeND) > 55%	AGB + LFMC + LAI
treeD	CCI PFT (treeBD + treeND) > 55%	AGB + LFMC + LAI
treeE	CCI PFT (treeBE + treeBD) > 55%	AGB + LFMC + LAI
shrub	CCI PFT shrub > 55%	LFMC + LAI
crop	CCI PFT crop > 55%	LFMC + LAI
herb	CCI PFT herb > 55%	LFMC + LAI
short vegetation	CCI PFT (shrub + crop + herb) > 55%	LFMC + LAI
Global model (including ~~land cover~~distinct CCI PFT data as ~~predictor~~additional predictors)		
global	inter-continental (all grid cells in ~~southern~~South Africa, ~~California~~western USA, Australia, and Europe)	AGB + LFMC + LAI + PFT treeNE + PFT treeND + PFT treeBE + PFT treeBD + PFT shrub + PFT crop + PFT herb

2.4 Model evaluation

For the evaluation of the models, 5-fold cross-validation is used. The same randomly computed folds are used for RF and GAM. The results are averages across all folds. The performance of the models is evaluated using the Nash-Sutcliffe model efficiency coefficient (NSE):

$$NSE = 1 - \frac{\sum_{i=1}^{n}(a_i - b_i)^2}{\sum_{i=1}^{n}(a_i - \bar{a})^2}, \quad (4)$$

with a as the true value, b as the predicted value and $\bar{a}$ as mean of ~~determination (R^2) and~~observed values, as well as the root mean squared error (RMSE) between the satellite-derived and the modelled VOD. ~~R^2~~NSE commonly ranges between 1 (perfect agreement) and 0, where the latter is the score for a model which solely predicts the mean of the reference data. Models that perform worse than this can also yield negative ~~R^2~~NSE values. In addition to the overall evaluation of the models, we evaluate the spatial distribution of ~~R^2~~NSE, i.e. ~~R^2~~NSE of the satellite and modelled VOD time series.

2.5 Partial relationships: Accumulated Local Effects (ALE)

The relationships ~~and sensitivities~~ of VOD to the predictors are examined via Accumulated Local Effects (ALE) plots (Apley and Zhu, 2020). ~~ALE~~ Like the commonly used PDP plots ~~are improvements over Partial Dependence Plots (PDP) (Friedman, 2001; Kuhn-Régnier et al., 2021) but can be interpreted similarly, i.e. as a partial relationship between a~~(Friedman, 2001), they show the marginal effect of a single predictor ~~and the target variable, taking~~on the model predictions. This marginal effect is reflected in the local gradient of the ALE plot; for example, a positive gradient indicates that an increase in the investigated predictor should lead to an increase in the predicted model outcome all other predictors being equal. While both techniques take into account all other predictors~~. Unlike PDPs,~~ to approximate the underlying relationship with the single investigated predictor, ALE does not combine each plotted predictor value with all possible combinations of the other predictors. ~~For ALE,~~Especially for correlated predictors, ALE plots are therefore more robust than PDPs (Kuhn-Régnier et al., 2021), as unlikely and unrealistic feature combinations are prevented. This is achieved by defining evenly spaced quantiles across the range of ~~an~~the examined ~~feature are defined. Every such~~predictor. Each quantile is then ~~combined~~used with only the closest~~,~~ existing ~~value combinations of all other features. This procedure prevents unlikely and unrealistic feature combinations, which increases robustness, especially when features are strongly correlated.~~ combinations of the other predictors to calculate the marginal effects. The~~-~~ ALE~~-~~ plots were generated from the final models, where all available data were used for training. To quantify the influence of the predictors on the target variable~~,~~ (sensitivities), we calculated the amplitude of the~~-~~ ALE~~-~~ curve (Δ_A).

3 Results

3.1 Performance of the models

The different regression models ~~used to predict satellite-derived VODs~~ showed large differences in model performance in predicting VOD (-0.04 ≤ ~~R^2~~NSE ≤ 0.97; 0.004 ≤ RMSE ≤ 0.~~2~~15) (Figure 2, ~~Figure 2,~~and Figure S1 ~~and S2~~ in supplement). ~~This difference was~~In summary, these differences were dominated by

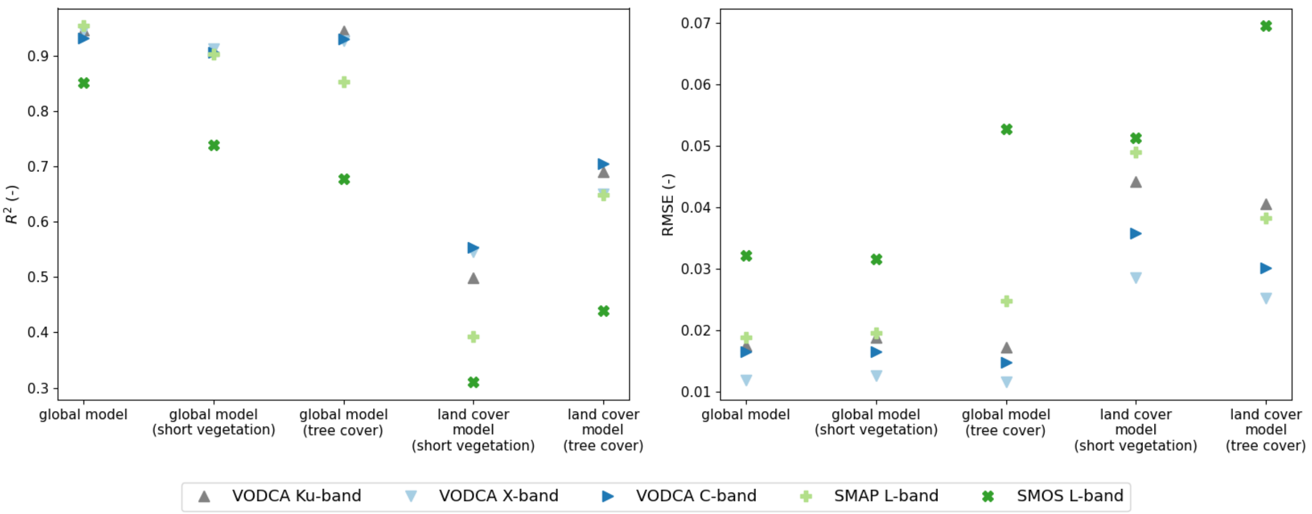
1) the type of regression model (RF or GAM~~),~~ , Figure 2 left subplots vs. right subplots, section 3.1.1);

2) by the use of 8-daily or monthly VOD data~~,~~ (symbols in Figure 2, section 3.1.2);

3) by the inclusion of land cover information as a predictor (~~land cover-specific vs. global models,~~ section 3.1.3);

4) by the wavelength of the predicted VOD (i.e. from Ku- to L-band~~),~~, section 3.1.4); and

5) by the vegetation type to which the model is applied to~~, i.e. spatial variability of global model performance, which will be discussed in more detail in the following.~~ (section 3.1.5).



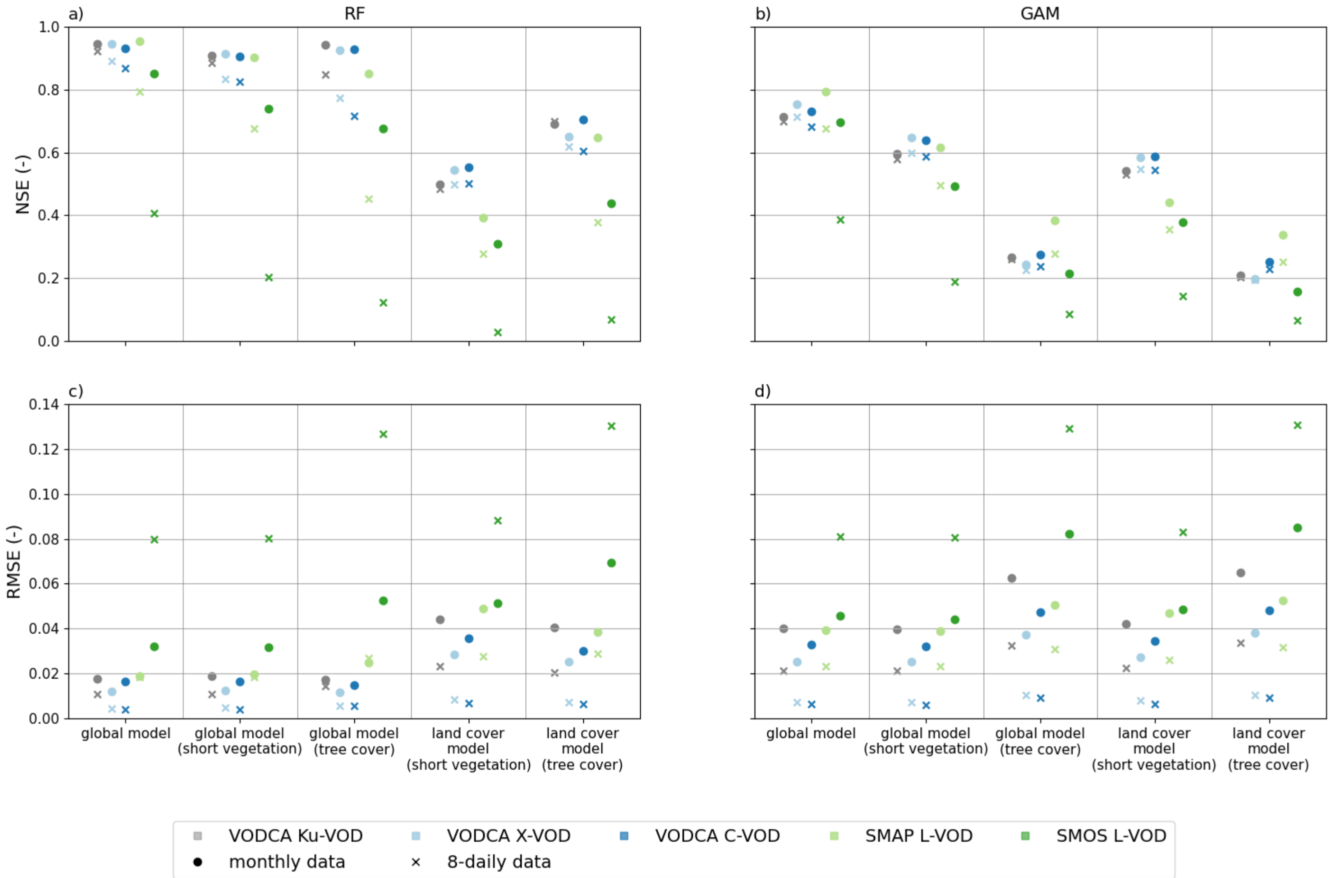


Figure 2: ~~Coefficient of determination (R^2, left) and~~ Nash-Sutcliffe model efficiency coefficient (NSE, top) and RMSE (~~right~~bottom) of random forest ~~(RF)~~ models (RF, left) and generalized additive models (GAM, right) using monthly (circle) or 8-daily (crosses) data. The global model uses PFTs as predictors, contrary to the land cover models, which were calibrated and applied only to the spatial extent of a certain dominant land cover class. Global model for short vegetation and tree cover ~~uses~~show results of the global model, but filtered by dominant land cover class.

3.1.1 Effect of the type of regression model used for calibrating the models (RF vs. GAM)

In general, RF performed better than GAM in predicting VOD, except for land cover-specific models for short vegetation classes where GAM reached ~~similar or~~ slightly higher ~~performance than~~ NSE (Figure 2 a vs. b) and similar RMSE compared to RF ~~(~~(Figure 2 c vs d). Another exception occurs for SMOS L-VOD where GAM performed better regarding the land cover-specific models for cropland and shrubland based on 8-daily data (see Figure S1~~, Figure S2 for corresponding RMSE results). In most cases, GAMs underestimated~~ for all models). While all models tended to underestimate high VOD values~~.~~, RF approximated them better than GAM. Based on these findings, in the following sections, we only refer to the results of RF models. If not stated otherwise, similar results were found for GAM.

3.1.2 Effect of the temporal aggregation of the ~~explanatory~~predictor variables (8-daily vs. monthly data)

Regression models based on monthly data ~~had~~usually exhibited higher ~~R^2~~NSE and lower RMSE than models based on 8-daily data (~~Figure S1~~comparison of circle and ~~S2).~~crosses in Figure 2 and S1). The superior performance of monthly over 8-daily models increased with increasing wavelength. For example, the difference was especially large for the prediction of SMOS L-VOD for which ~~R^2~~NSE doubled from 8-daily to monthly data ~~(Figure S1).~~(Figure 2 a). The performance in predicting Ku-, X- or C-VOD was more similar or monthly data presented slightly higher performance than 8-daily data. Given the higher performance of models based on monthly data, the following description of results is based on ~~these~~models with monthly data, unless mentioned otherwise. Section 3.2 examines than the differences of VOD sensitivities to the predictors based on the considered time scale.

3.1.3 Effect of including land cover information as a predictor (global vs. land cover-specific models)

Considering RF models based on monthly data, the global models (defined as models including fractional cover of PFTs as predictors, see Table 2) showed better model performances than the land cover-specific models ~~(~~that were trained and applied only to one specific land cover~~).~~. The global models performed with an ~~R^2~~NSE of 0.85 to 0.95 and an RMSE of 0.01 to 0.03 depending on VOD wavelength (Figure 2~~Figure 2).~~ a and c). We also compared the model performance of a specific land cover type within the global model with the related land cover-specific model. ~~Regarding RF models, the~~The land cover-specific RF models had ~~an R^2 of~~ a lower NSE (-0.09 to -0.59~~ lower~~) and a higher RMSE ~~of~~ (+0.006-0.03~~ higher~~) than the ~~related land cover~~global model within the ~~global models.~~same land cover. Considering ~~GAMs~~GAM, land cover-specific models performed better within a certain land cover type than the global model for the same land cover type. This applies especially for land cover types with simpler vegetation structure, e.g. shrubland, herbaceous vegetation or broad-leaved evergreen trees, and less for more complex land cover types like the ~~treeAll~~tree cover and short vegetation classes. These results indicate that the relationship between vegetation properties and VOD can be modelled with simpler relationships as represented by GAM only within a land cover type but that global relationships require more complex relationships as represented by RF.

3.1.4 Effect of ~~VOD~~ wavelength

In general, the ~~R^2~~NSE of predicting short-wavelength VOD was higher than for predicting L-VOD~~ predictions~~ and RMSE decreased from long to short wavelengths (Figure 2). All SMOS L-VOD models performed with a lower ~~R^2~~NSE and a higher RMSE than the other VOD models including SMAP L-VOD. For RF models based on 8-daily data, ~~R^2~~NSE was highest for Ku-VOD, followed by X-VOD and C-VOD~~ (Figure S1).~~. For monthly data and ~~GAMs~~GAM, the order in performance was slightly different between Ku-, X- and C-VOD for ~~R^2~~NSE and RMSE.

In the global model, the land cover-specific model performance ~~depends~~depended on the different VOD wavelengths. The prediction of monthly Ku-, X- and C-VOD using RF reached the highest performance for broad-leaved evergreen trees (0.95 $\leq$ ~~R^2~~NSE $\leq$ 0.97, 0.009 $\leq$ RMSE $\leq$ 0.013) and the lowest performance for croplands (0.82 $\leq$ ~~R^2~~NSE $\leq$ 0.85, 0.015 $\leq$ RMSE $\leq$ 0.023). Predicting monthly SMAP L-VOD using RF had the highest performance in herbaceous vegetation (~~R^2~~NSE = 0.93, RMSE = 0.016) and the lowest performance in deciduous trees (~~R^2~~NSE = 0.74, RMSE = 0.031). RF prediction of monthly SMOS L-VOD attained the highest performance in herbaceous vegetation (~~R^2~~NSE = 0.84, RMSE = 0.023) and the lowest performance in needle-leaved and deciduous trees and croplands (~~R^2~~NSE ~ 0.6, 0.032 $\leq$ RMSE $\leq$ 0.059).

3.1.5 Spatial variability in model performance

The performance in predicting VOD shows large spatial differences (Figure 3). Across all VOD datasets, the prediction of VOD was best in Australia; followed by South Africa, Europe, and western USA (Figure S2). As for the global model results (chapter 3.1.4), the best performance was achieved in predicting Ku-, X-, and C-VOD and the lowest performance for SMOS L-VOD. This is indicated by the dominant colour distribution in Figure 3 and by the corresponding histograms (Figure S2), whereby the more right-skewed and narrower the distribution the better the prediction of all pixel time series (e.g. Ku-VOD for Australia).

Several geographical patterns of high ~~predictability (high R^2, blue areas in Figure 3), like the cropland area in~~or low model performance appear for all VOD datasets. High model performance occur mainly in regions with croplands (e.g. south-western and south-eastern Australia~~, and~~), large shrublands (e.g. northern Australia and central South Africa) and grasslands (north-western and south-eastern South Africa and western Australia) (high NSE, blue areas in Figure 3). Regions in the south-western USA show a poor ~~predictability~~performance (low ~~R^2~~NSE, red areas in Figure 3~~), e.g. southern California, are similar for all VOD datasets.~~).

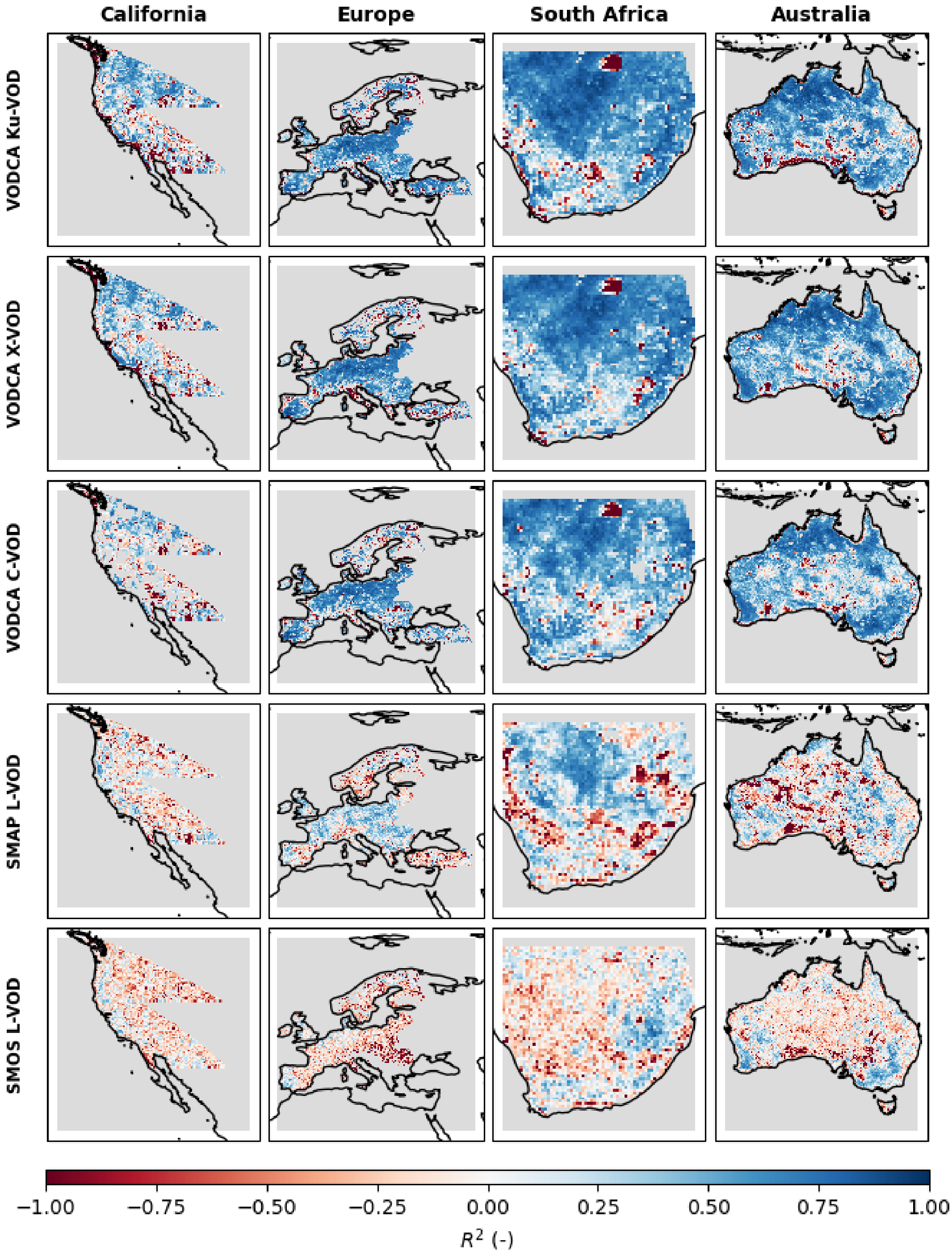
~~Across all VOD datasets, the prediction of VOD was best in Australia; followed by South Africa, Europe, and California (Figure S3). As for the global~~

Higher model ~~results (chapter 3.1.4), the best~~ performance ~~was achieved in predicting Ku-, X-, and C-VOD and the lowest performance for SMOS L-VOD. This is indicated by the dominant colour distribution in Figure 3 and by the corresponding histograms (Figure S3), whereby the more right-skewed and narrower the distribution the better the prediction of all pixel time series (e.g.~~ occur~~Ku-VOD for Australia).~~

~~Not only areas with a high crop fraction have high R^2, but~~ also ~~areas with large shrub fractions, e.g. northern Australia and central South Africa, and high herbaceous fractions, like in north-western and south-eastern South Africa and western Australia.~~

~~Pixels with large~~in regions with larger seasonality in LAI and LFMC (e.g. eastern Europe and northern part of ~~California), show higher R^2 results per pixel. Increasing pixel~~western USA) (Figure 4 c) and in pixels with homogenous land cover ~~homogeneity also contributes to improved results. This implies worse results for~~ than in pixels with a more heterogeneous ~~pixels and in regions with less pronounced seasonality in LAI due to lack of defoliation (needleleaf, evergreen), or in LFMC due to more or less stable weather conditions or to more drought resistance of less plant-water sensitivity~~ (Rao et al., 2022)~~, such as the central areas of California, northern Europe and central Australia.~~land cover distribution (Figure 4 a and b). With increasing wavelength, the VOD of ~~these~~ areas ~~is~~with less pronounced seasonality was getting more difficult to predict.

Additionally, regions with mean VOD values less than 0.1 and marginal changes over time tend to have low or even negative ~~R^2~~NSE. This is noticeable in central Australia and central South Africa. ~~The comparison of the high R^2~~Investigating the differences in the overall NSE based on all values (section ~~3.1, >~~3.1) with the grid cell based NSE in Figure 3 and S3 allows an insight if the RF models are able to represent not only spatial patterns but also time series. The comparison of the high overall NSE (>1.000 data samples) with the ~~R^2 in Figure 3 and S3~~NSE shown here (monthly time series January 2015 – July 2017 resulting in a maximum time series of 31 months i.e. < 32 data samples) indicates that ~~R^2~~NSE seems to be sensitive to the data size, leading to small ~~R^2~~NSE when few data points are available. The reference and modelled mean VOD and the variance of VOD are highly correlated in space (Spearman correlation coefficient > 0.75) which shows that the models ~~captures~~capture the variability and spatial patterns of VOD. With higher mean VOD the ~~R^2~~NSE increases, e.g. such as for the tree-covered areas dominated by deciduous broadleaf trees. Whereby this finding is based on the VOD range constrained by the proceeded data preparation, it might be not valid for very high VOD values, e.g. in rainforests, which are not considered here.



California
Europe
South Africa
Australia
VODCA Ku-VOD
VODCA X-VOD
VODCA C-VOD
SMAP L-VOD
SMOS L-VOD
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
R^2 (-)

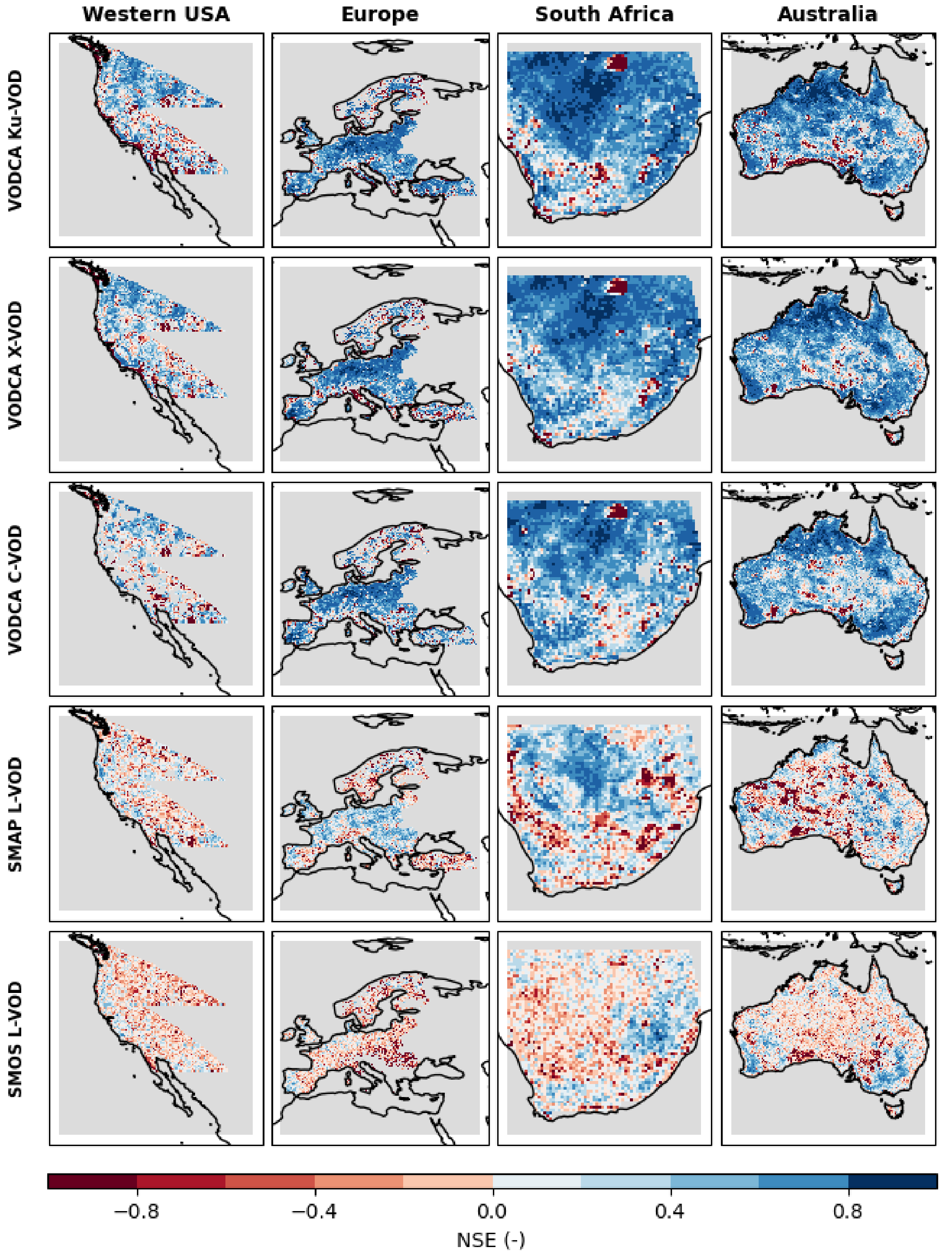


Figure 3: ~~Coefficient of determination (R^2~~Nash-Sutcliffe model efficiency coefficient (NSE)) per pixel for the global random forest model (PFTs included as predictor) based on monthly values. Rows indicating results for the different VOD datasets and columns the different regions as dictated by the availability of the LFMC dataset.

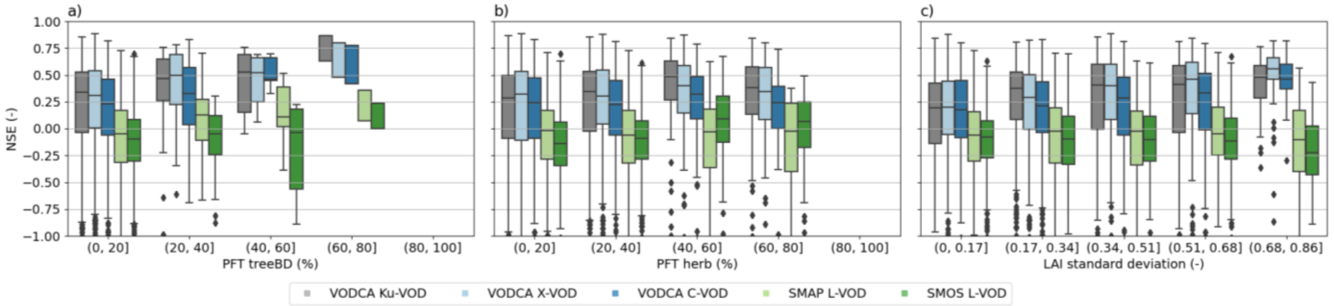


Figure 4: Spatial Nash-Sutcliffe model efficiency coefficient (NSE) based on the global random forest model computed with monthly data stratified by the land cover homogeneity of a pixel exemplary shown for a tree cover class a) plant functional type deciduous broadleaf trees (PFT treeBD) and b) for herbaceous vegetation (PFT herb). Note, that no data with 80-100% of these specific land-cover classes are available. c) shows NSE stratified by the seasonality of LAI expressed as the intra-annual standard deviation of LAI.

3.2 Relationships between VOD and vegetation properties

3.2.1 Global (inter-continental) relationships

~~The ALE plots in~~ The effects of vegetation properties on VOD for all wavelengths on a monthly or a 8-daily data basis are shown in the ALE plots in Figure 5~~Figure 5 (Figure S4 – S7 for all global predictors based on monthly and 8-daily RF and GAM models) demonstrate the effects of vegetation properties on VOD for all wavelengths. For Ku-VOD, LAI and herbaceous land cover have the highest influence followed by AGB and LFMC with an amplitude of Δ_A = 0.109, 0.054, 0.045 and 0.017, respectively. For X- and C-VOD the order of influence is LAI (Δ_A = 0.072 for X-VOD; and 0.099 for C-VOD), AGB (Δ_A = 0.032; and 0.046), herbaceous land cover (Δ_A = 0.029; and 0.031) and LFMC (Δ_A = 0.017; and 0.028).~~

(Figure S3 and S4 for all global predictors and GAM). The amplitudes Δ_A of the ALE curves can be used as a measure of the importance of a predictor for the estimation of VOD. The amplitudes Δ_A are usually higher for monthly data than for 8-daily data (Figure 6 a) except for the relationship between AGB and SMOS L-VOD (Figure 6 c). This result indicates that the used predictors are of higher importance for monthly data than for 8-daily data. However, the high Δ_A values in the global RF model based on 8-daily data for SMOS L-VOD and the relative low performance of this model (NSE=0.41) indicates that the influence of the used predictors might be overestimated. A predictor that could reproduce the main temporal dynamics in the 8-daily SMOS L-VOD signal is indeed missing in the analysis.

The order of Δ_A of the predictors within a certain model are generally similar for 8-daily and monthly models. The coverages of trees are for all models the main contributors to the VOD predictions. LAI is the second most important predictor for Ku-VOD and the most important for X- and C-VOD. For the L-VODs the importance of LAI is lower than for the short-wavelength VODs. The importance of AGB increases from low to middle importance for the shortwave VODs to the highest importance for the L-VODs. The coverages of short vegetation classes have middle to low influence on the VOD and decreases with increasing wavelength but as an exception the coverages of shrubs is the second- and third-most important predictor for monthly and 8-daily SMAP L-VOD, respectively. The Δ_A of LFMC are increasing with wavelength, with low influence on Ku- and X-VOD and higher influence on L-VOD. An exception here is the 8-daily SMOS L-VOD model, where LFMC has also a low impact on the predictions, but given the low performance of this model, the estimates importance of LFMC on SMOS L-VOD might be unreliable. Interestingly, the amplitude of the ALE plots varies between wavelengths, within monthly and 8-daily models although these results are based on normalised data. For LAI and land cover a clear decrease of the ALE amplitude with increasing wavelength is visible, which corresponds to the fact that the magnitude of VOD value range decreases with increasing wavelength. For AGB and LFMC, the ALE amplitude increases with increasing wavelength.

Given the similar shape of 8-daily and monthly based ALEs but with smaller amplitudes, we will focus on the examination of the monthly ALE curves. All VOD datasets show a positive relationship with LAI, but all curves saturate around an LAI value of 2.3, which corresponds approximately to the 95%-ile of LAI in our dataset~~.~~ (Figure 5 a). LAI has a much stronger effect on Ku-, X-, and C-VOD than on L-VOD. Interestingly, the relationship between LAI and SMAP L-VOD ~~(Δ_A 0.054)~~ is more similar to the relationship of LAI and shortwave ~~VOD~~VODs e.g. X-VOD than ~~for LAI and~~the relationship with SMOS L-VOD ~~(Δ_A 0.024).~~.

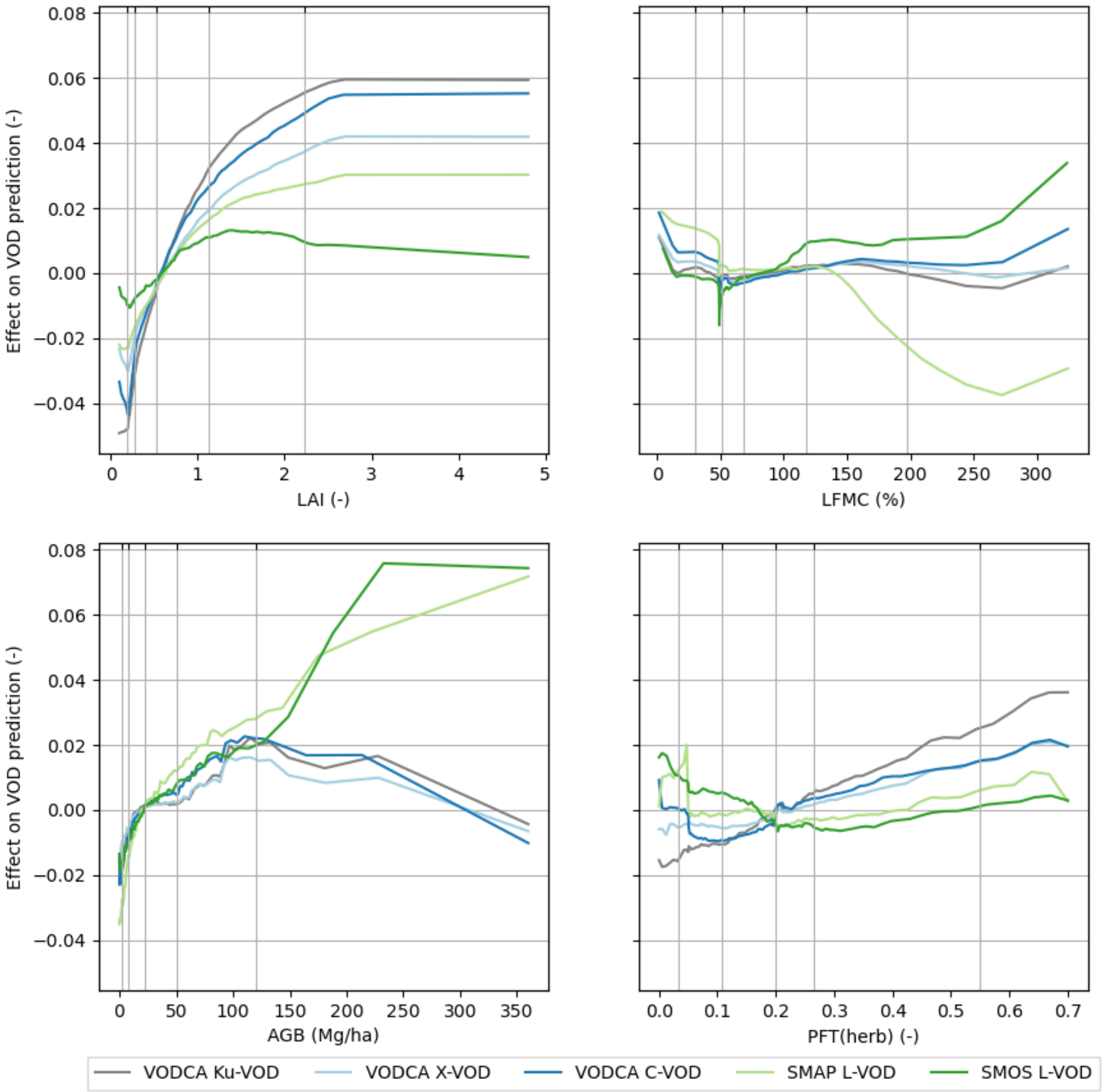
~~The relationship with LFMC is more complex for all VOD datasets. From 0 % to 50 % LFMC, the relationships are negative with a negative spike at 50 % LFMC. Afterwards, VOD increases with increasing LFMC, which is most pronounced for SMOS L-VOD (overall Δ_A = 0.05). However, SMAP L-VOD shows a strong negative relationship with LFMC after around 140 % LFMC (overall Δ_A = 0.056).~~The relationship with LFMC is more complex for all VOD datasets (Figure 5 b). From 0 % to 50 % LFMC, the relationships are negative with a negative spike at 50 % LFMC. Afterwards, VOD increases with increasing LFMC, which is most pronounced for SMOS L-VOD. However, SMAP L-VOD shows a strong negative relationship with LFMC after around 140 % LFMC. Generally, the relation within the last 95 %-percentile have to be interpreted with caution, because higher LFMC values also have a higher uncertainty (Yebra et al., 2018). In addition, the validation of the LFMC data set is impeded by uncertainties due to difficulties of comparison between measurements on the ground and what is detected by the satellite. Uncertainties in the used LFMC dataset arise from the temporal matching procedure of in-situ samples and MODIS data and from the canopy closure of the forest cover and the contribution of understory to the measured surface reflectance. However, these factors are difficult to quantify and can only be discussed in a qualitative manner, but they still might influence the results presented here.

All VOD datasets show a similar increase with AGB until 120~~-~~ Mg/ha (corresponding to the 95%-~~ile~~percentile) but the relationships differ at higher AGB values~~. While~~ (Figure 5 c). Ku-, X- and C-VOD show a ~~decreasing relationship~~decrease with increasing AGB~~,~~ above 120 Mg/ha but SMOS and SMAP L-VOD continue to increase ~~(overall Δ_A = 0.095 for SMOS L-VOD; and 0.107 for SMAP L-VOD).~~.

The relationships with land cover fractions are positive for most VOD datasets. As an example, we show here ~~show~~ the relationship with the fraction of ~~herbaceous~~shrubland cover~~. Ku- and X-VOD show an almost~~ (Figure 5 d). SMAP L-VOD shows a nearly monotonic increase with increasing ~~herbaceous~~shrubland cover. ~~On the other hand, C- and L-VOD (Δ_A 0.025 for SMAP; and Δ_A 0.024 for~~ The shortwave VODs and SMOS L-VOD~~)~~ show ~~a negative relationship~~no relation with ~~herbaceous~~shrubland cover ~~at very low~~below 10% coverage ~~(up to ca. 0.15, corresponding approx. to the 40%-ile of herbaceous cover)~~ but ~~increase afterwards.~~ show a positive relationship at higher coverage. SMOS L-VOD shows a non-monotonic relationship with shrubland cover.

Taken together, we find the following effects of vegetation properties on the different VOD datasets: SMOS L-VOD is most strongly affected by AGB (positive relationship), followed by tree cover and LFMC (positive relationship at LFMC > 50 %), short vegetation cover and LAI (positive relationship for LAI < 1.5~~), and herbaceous vegetation.~~). SMAP L-VOD is most strongly affected by AGB (positive relationship), followed by ~~LAI (positive relationship for < 2.5),~~ LFMC (negative relationship~~),~~) and shrubland cover, and ~~herbaceous vegetation.~~LAI (positive relationship for LAI < 2.5). Ku-, X-, and C-VOD show very similar relationships and are most strongly affected by LAI (positive relationship~~),~~) and tree cover, followed by AGB (positive relationship up to 120~~-~~ Mg/ha), ~~herbaceous~~short vegetation cover, and LFMC. ~~The relationships with LFMC and herbaceous cover differ mostly between Ku- and X-VOD on the one hand and C-VOD on the other.~~

Interestingly, the amplitude of the ALE plots varies between wavelengths. For LAI and land cover a clear decrease of the ALE amplitude with increasing wavelength is visible, which corresponds to the fact that the magnitude of VOD decreases with increasing wavelength. For AGB and LFMC, the ALE amplitude increases with increasing wavelength.



Effect on VOD prediction (-)
0.08
0.06
0.04
0.02
0.00
−0.02
−0.04
LAI (-)
LFMC (%)
AGB (Mg/ha)
PFT(herb) (-)
VODCA Ku-VOD
VODCA X-VOD
VODCA C-VOD
SMAP L-VOD
SMOS L-VOD

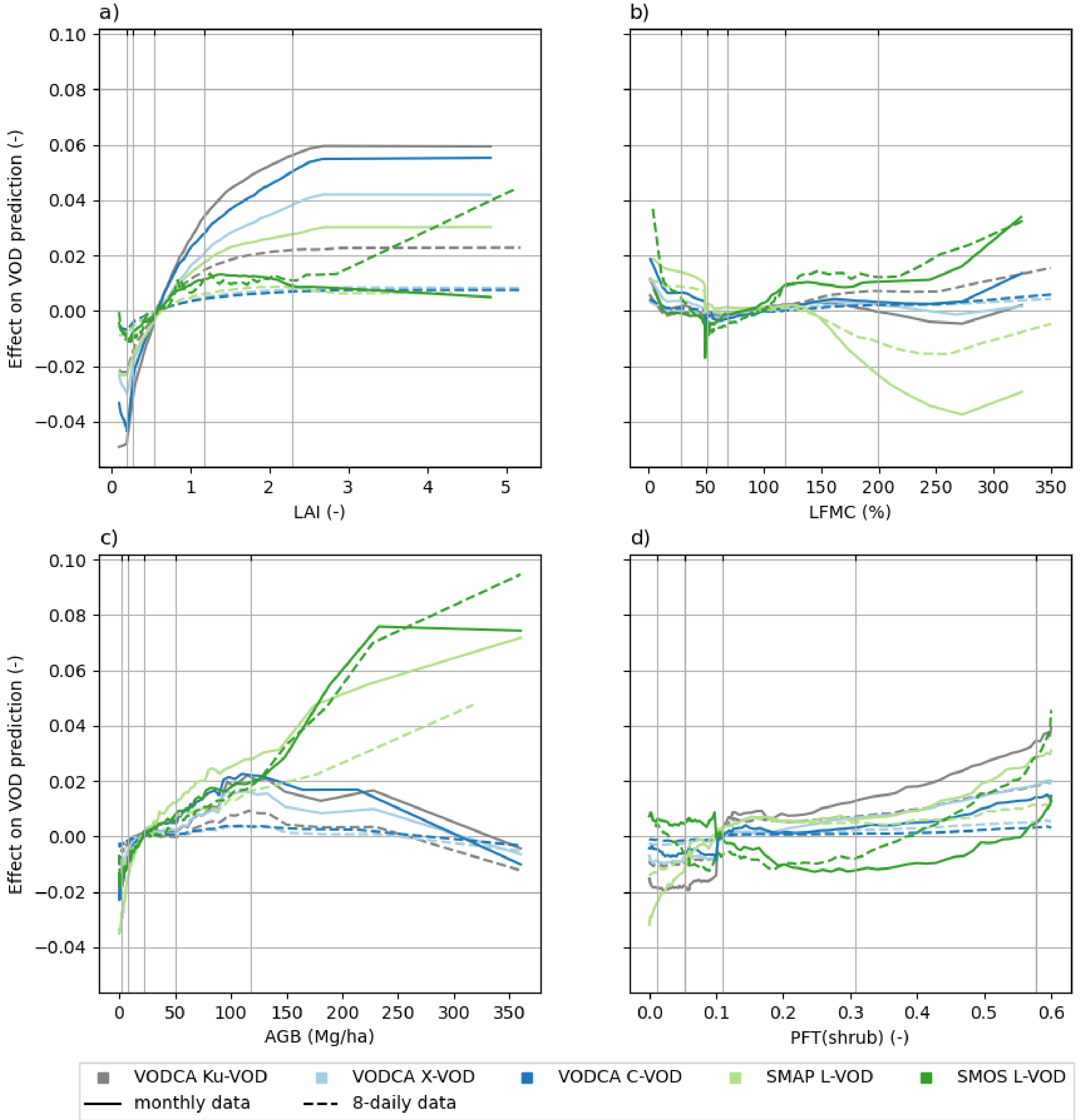


Figure 5: ALE plots of predicted normalised VOD with respect to ecosystem properties based on the global monthly or 8-daily RF model with plant functional type (PFT) of ~~herbaceous~~shrubland vegetation (~~herb~~shrub) as an example of the influence of land cover fractions on VOD. Vertical lines indicate the quantiles of the data sample size 0.05, 0.25, 0.5, 0.75, and 0.95, respectively.

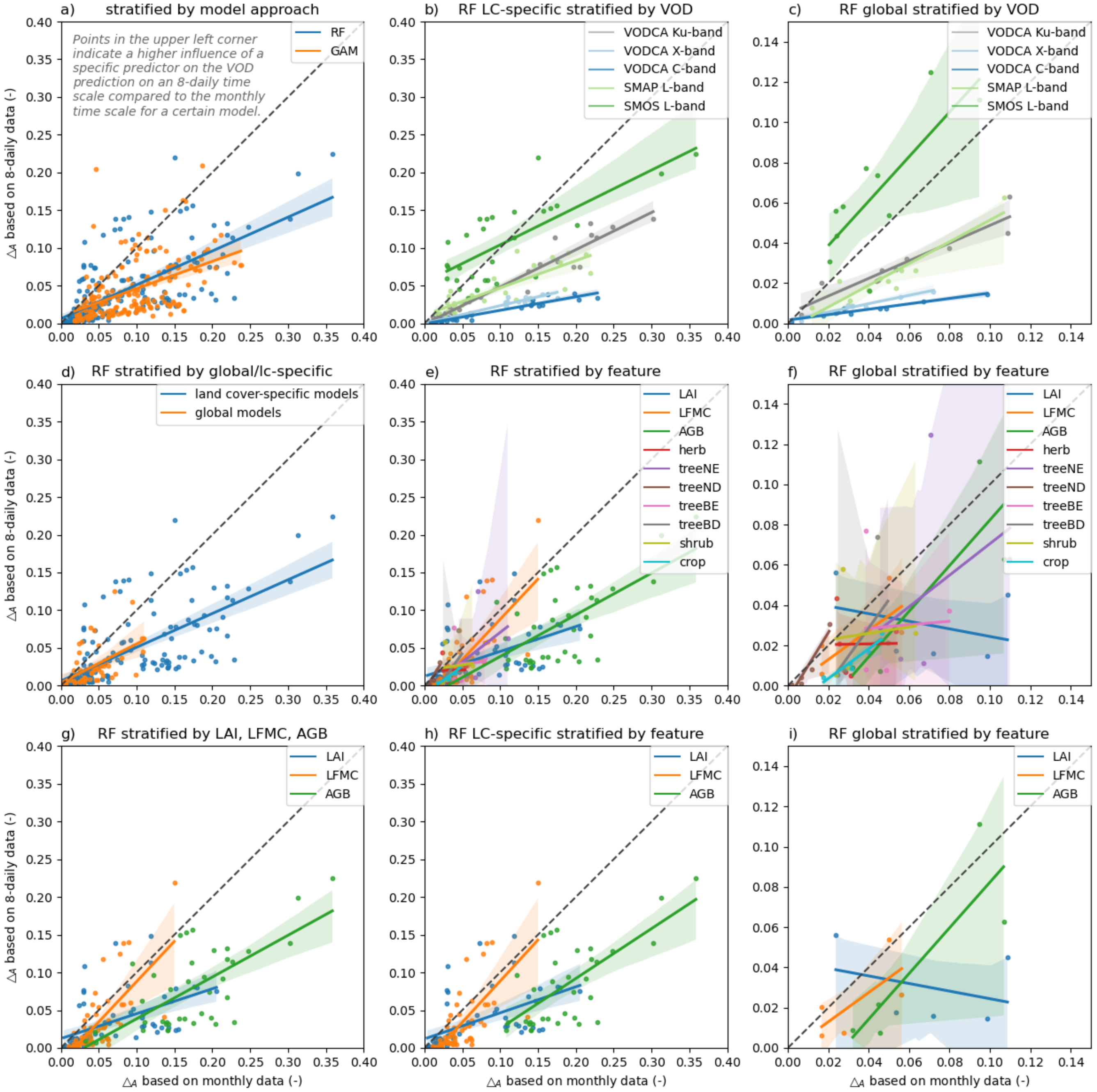


Figure 6: Regression plots of individual ALE amplitudes Δ_A on a monthly data basis versus on a 8-daily data basis. Panel a) shows the Δ_A of ALE curves from GAM and RF whereby panels b) to i) show only Δ_A of RF models but colourised by different factors: 'global' indicates the global models which use also PFT fractions as predictors and 'LC-specific' identifies the land cover-specific models which only use LAI, LFMC and AGB (for tree cover) as predictors and used data filtered for the specific land cover type. Note that c), f) and i) are zoomed in compared to the other subplots. Points located in the upper left corner indicate a higher influence of a specific predictor on the VOD prediction on an 8-daily time scale compared to the monthly time scale for a certain model. Points located on the 1:1 line indicate a constant influence on VOD regardless of the considered time scale. Points located in the lower right corner indicate a higher influence of a predictor on a monthly time scale.

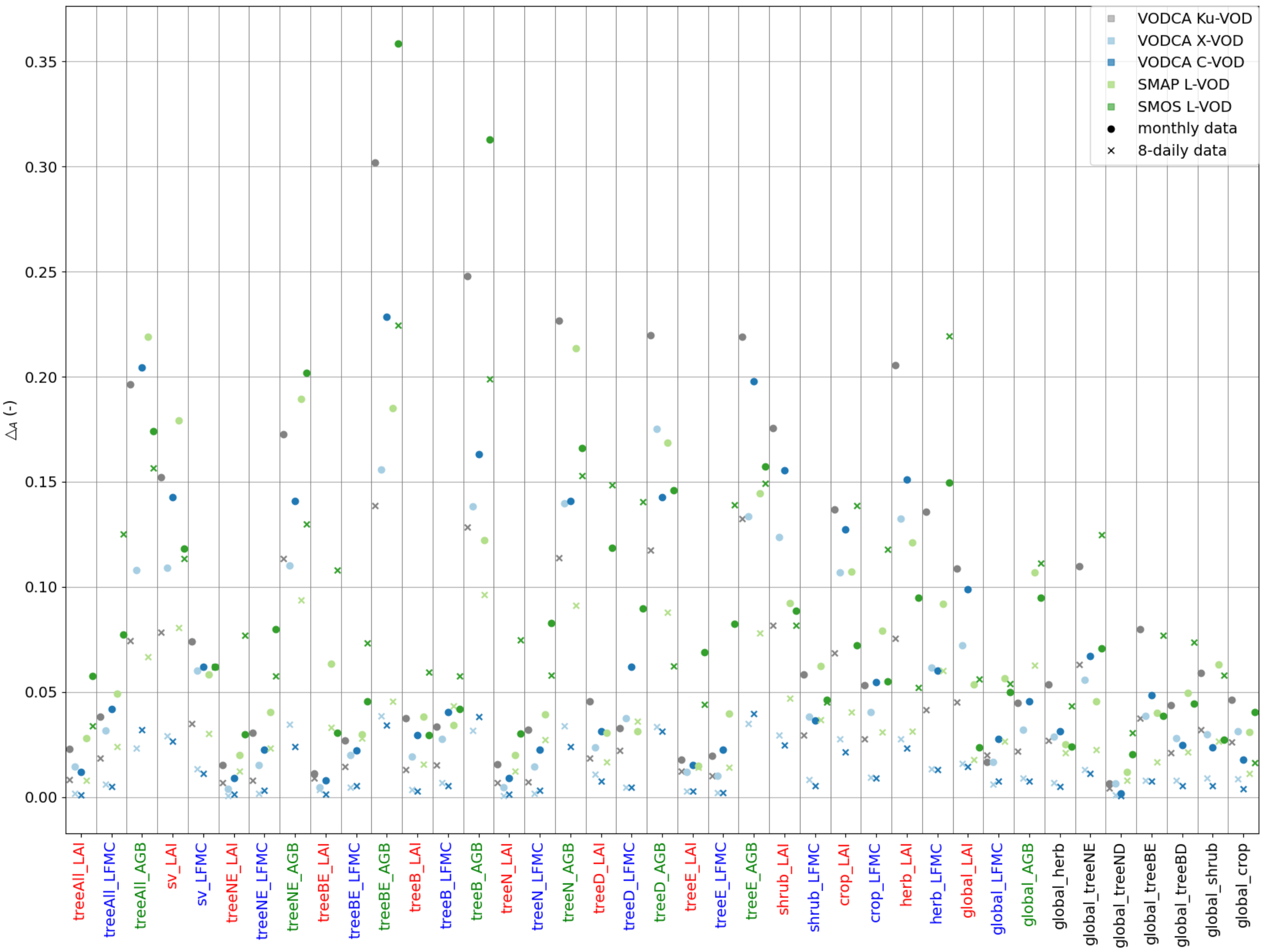


Figure 7: Amplitudes of RF ALEs for 8-daily and monthly models. The prefix defines the model, e.g. 'global' indicates the ALEs of the global model, all other prefixes indicate land cover-specific models. The suffix defines the predictor (red = LAI, blue = LFMC, green = AGB, black = PFTs – only used within global models).

3.2.2 Relationships ~~of~~within land cover~~-specific~~ types

In this chapter, we summarize the results of the RF models for relationships within a certain land cover type. The individual predictors in the land cover-specific models have a ~~partially~~ higher influence on the VOD prediction than in the global model because the land cover-related predictors are not used within the land cover specific models (~~Figure 5, Figure S4 – S15).~~Figure 6 d). ALE amplitudes Δ_A for monthly data are mostly larger than for 8-daily data with some exceptions for SMOS L-VOD (Figure 6 b). The order of the Δ_A for the different VODs is in the land cover-specific models like in the global model with the highest values for SMOS L-VOD, followed by Ku- and SMAP L-VOD and X- and C-VOD.

~~For~~In models for specific tree cover types, AGB has the largest Δ_A, followed by LFMC and LAI (Figure 7). The model for deciduous trees for 8-daily SMOS L-VOD data is an exception, in which LAI has the largest importance, followed by LFMC and AGB. Due to the poor performance of this model, this result might be questionable.

Models for short vegetation types, usually have LAI as the most important predictor, followed by LFMC (Figure 7). Exceptions are the models for the herbaceous vegetation with 8-daily SMAP L-VOD, and 8-daily and monthly SMOS L-VOD, where LFMC has the highest importance. In general, for the tree cover models AGB and for short vegetation cover LAI has a higher influence on the predictions than LFMC. Nevertheless, the Δ_A-LFMC regression line in Figure 6 h) indicates that LFMC has a similar effect on both time scales. This is contrary to AGB and LAI where the effect is higher for monthly than for 8-daily data. For short vegetation, the ALE plot between VOD and LFMC shows a similar form as in the global model with a drop around 50 % LFMC (Figure S6), which indicates that the global VOD-LFMC relationship is dominated by dynamics in short vegetation areas. Particularly, the drop is based on the herbaceous land cover type, which is also visible in the 8-daily based models and in the GAM (Figure S6 and S8). The importance of LAI in predicting VOD decreases for herbaceous and shrubland cover models with increasing wavelength. A similar dependence occurs for LFMC for shrublands and monthly data above 140 % LFMC. Globally, the positive relationship between VOD and LFMC in the range of 50 % and 140 % LFMC and the negative relationship at higher LFMC originates from croplands, because this decrease is only visible in the LFMC-ALE from the cropland model.

In tree-covered areas (treeAll model), the ~~LAI-~~ALE shows ~~a slight positive relationship~~that VOD increases with ~~VOD~~LAI up to ~~an~~ LAI ~~of ca.~~= 2 and is then ~~a~~ stable or slightly ~~decreasing~~decreases (Figure S5). The relation~~. The relation~~ of VOD with LFMC is positive for Ku-, X-, and C-VOD but unimodal for both L-VODs. AGB is the dominant predictor for all tree-covered models but the relationship with VOD is non-linear and non-~~monotonic.~~monotonic.

Comparing the ALEs of the treeAll ~~and~~model with the models for individual forest ~~type models~~types (i.e. treeB, treeN, treeD, treeE, Figure ~~S8),~~S5) shows that the influence of a specific forest type is partially recognizable within the treeAll ALEs. For example, the ~~highly non-linear~~ relationship between LFMC and VOD ~~and LAI until LAI~2.0 is based on the VOD-LAI relationship of deciduous trees. The~~ in the treeAll ~~LFMC-ALE~~model is highly influenced by the relationship for needle-leaved and evergreen trees. The ~~apparent~~decline of SMOS L-VOD ~~decrease~~ with LFMC is also pronounced within most tree types but not within deciduous trees. The relationships with AGB~~-ALE~~ for needle-leaved trees is ~~less non-~~more linear in comparison to the other tree cover models. Deciduous and broad-leaved trees exhibit a more complex relationship with AGB than evergreen and needle-leaved trees~~.~~

~~For short vegetation, LAI is the main influencing~~ for all VODs. The amplitudes of ALE curves with AGB are highest for X-VOD for deciduous trees (treeD Δ_A = 0.175) and for SMOS L-VOD for broadleaved trees (treeB Δ_A = 0.313). These results demonstrate that biomass is also an important predictor for short-wavelength VODs but that this importance varies with wavelength and ~~shows a positive relationship with all VOD bands. The ALE plot between VOD and LFMC shows a similar form as in the global model with a drop around 50 % LFMC, which indicates that the global VOD-LFMC relationship is dominated by dynamics in short vegetation areas. Particularly, the drop is based on the herbaceous land cover type, which is also visible in the 8-daily based models and in the GAMs (Figure S10, S11, S14 and S15). Contrary to the global model, the land cover-specific models do not exhibit the clear dependency of the ALE amplitude to the wavelengths. The dependency of the ALE amplitude on wavelength is still visible in the LAI-ALE of the herbaceous and shrubland cover models, especially for LAI greater than 1 and more pronounced in the 8-daily based models than in the monthly based models. The same is true for the LFMC-ALE of the monthly based shrubland models above 140 % LFMC. The positive relationship between 50 % and~~

~~140 % LFMC and the following decrease (especially for the L-VOD LFMC-ALE) for short vegetation is influenced by cropland cover, because the decrease is only visible in the LFMC-ALE from the cropland model~~forest type.

Contrary to the global model, the land cover–specific models do not exhibit a clear dependency of the ALE amplitude on the wavelengths.

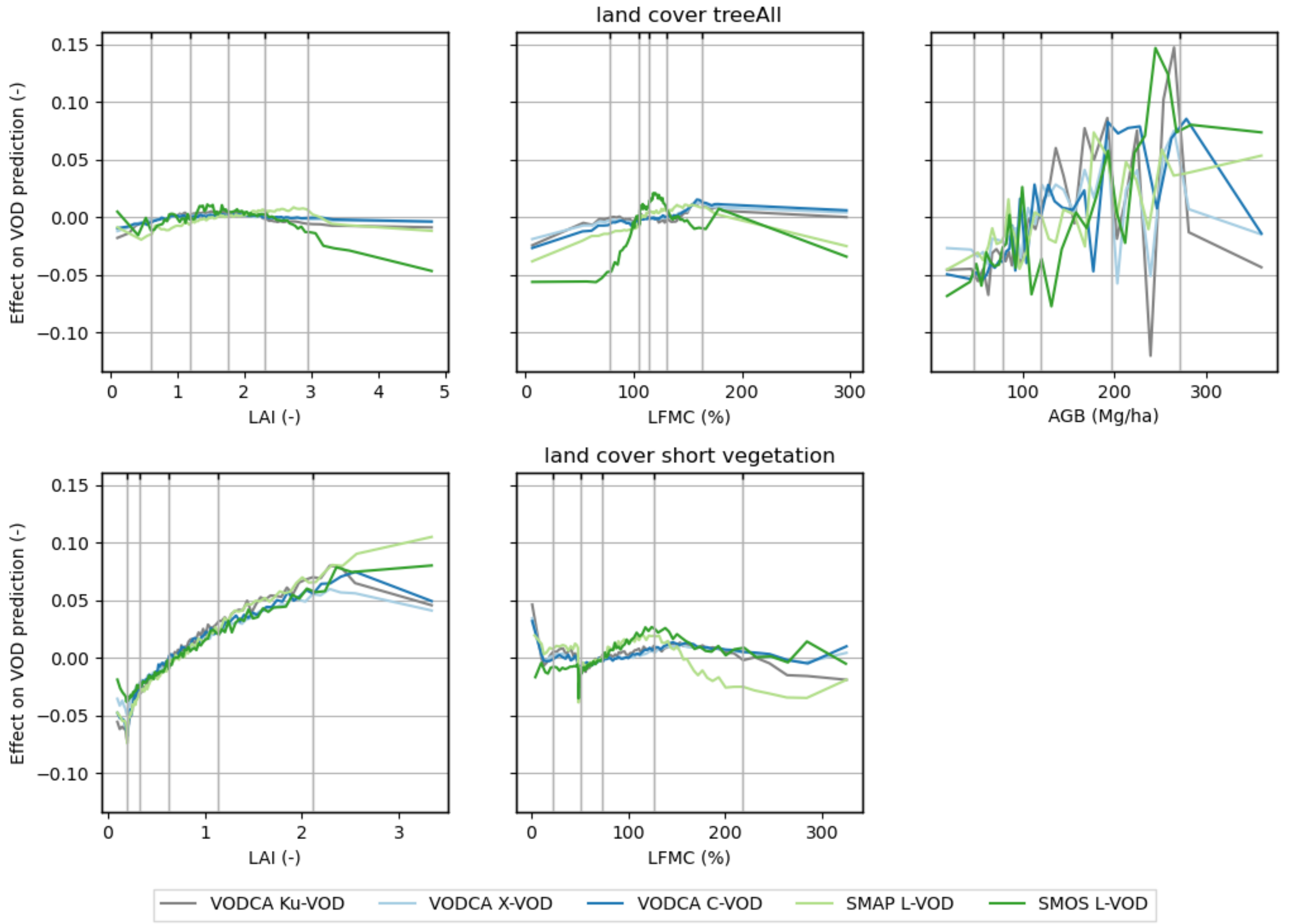


Figure 8: ALE plots of normalised VOD to ecosystem properties based on land cover-specific monthly RF models. Shown models are based on pixels with a PFT treeAll fraction more than 0.55 (top), and for short vegetation (bottom). Vertical lines indicate the quantiles of the data sample size 0.05, 0.25, 0.5, 0.75, and 0.95, respectively.

4 Discussion and conclusions

4.1 Predictors and predictability of VOD

The results demonstrate that for the global prediction of VOD, i.e. over different biomes, a ~~more~~ flexible modelling approach such as RF is better suited ~~as opposed to~~than an additive approach like GAM. The lower global performance of GAM suggests that local factors, e.g. intercepted or standing water or heterogeneous soil properties, and interactions between factors play a role in the dynamics of VOD~~, which were not considered and used as additional additive predictors.~~. In contrast, RF is partly able to account for this due to its ability of flexible modelling ~~leading to better~~which results~~.~~ in higher model performance. The simpler structure of GAM compared to RF is, in most cases, insufficient to predict VOD, but within single land cover types a simpler additive approach like GAM is sufficient. This indicates that the relationship between VOD and ecosystem properties cannot be easily captured with global linear, monotonic, and bivariate regressions but requires accounting for the non-linear interactions between various ecosystem properties. The results imply that the set of predictors allows the estimation of the dynamics of short wavelength VODs at high temporal resolution (8-daily and monthly) with very good performance, but the set of used predictors is insufficient to explain the dynamics in L-VOD due to ignoring local effects or possibly disregarded predictors.

This conclusion is supported by the performance difference between the four studied regions. For example, Europe has a much more fragmented landscape than most areas in Australia causing mixed effects on VOD within the coarse 0.25° grid cells leading to a lower predictability in Europe than Australia. Even if PFT fractions are used as predictors, the mismatch between the coarse resolution and land cover complexity cannot be resolved. This is especially pronounced in the longwave VOD, for which the footprint is often significantly larger than 0.25° (> 40 km). ~~A filtering of neighbouring grid cells would not then reduce the impact of the surrounding land cover.~~ Local complex effects on VOD are likely related to land cover changes, intercepted~~,~~ or standing water, or soil properties. For example, Saleh et al. (2006) showed for ~~the~~a grassland ~~test~~ site that ~~because of interception L-VOD~~ intercepted water could double ~~in value after a rainfall event.~~L-VOD after a rainfall event. Comparable to this finding, Wigneron et al. (1996) also reports a possible doubling in C-VOD due to interception at a wheat field. Although interception has reduced influence on the coarse resolution data (Baur et al., 2019; Wigneron et al., 2021) or might not impede temporal VOD analyses (Feldman et al., 2020), temporary flooding leads to an evident ~~decrease in VOD. This is not only valid~~change in VOD. For example, a decreased L-VOD signal at flooding was recognised for short vegetation areas using K-VOD derived from the microwave radiometer of the Chinese satellite FY-3B (Liu et al., 2019)~~, but also~~ as well as for forests ~~(Jones et al., 2011; Bousquet et al., 2021). These impacts of~~using AMSR-E K-VOD (Jones et al., 2011) or using SMOS-IC L-VOD (Bousquet et al., 2021). The effect of such local ~~effects~~events on VOD implies that large-scale spatial relations between VOD and e.g. AGB (Liu et al., 2015; Rodríguez-Fernández et al., 2018; Mialon et al., 2020) will likely wrongly associate changes in VOD to changes in AGB, which might result in unrealistic estimates of local AGB dynamics. This conclusion is supported by the findings of Konings et al. (2021), who show that regional temporal anomalies of X- and L-VOD are mostly uncorrelated with temporal anomalies of AGB but show a higher correlation with root-zone soil moisture, an indicator for water stress and availability.

The comparison of the global and the land cover-specific models highlights the complexity of the relation between VOD and vegetation properties. An interesting result is that the ALE amplitudes (i.e. sensitivity) increase with increasing wavelength in the global model but not in the land cover-specific model. The land cover-specific models only include pixels with a coverage > 55 % of the specific land cover type but do not use PFT fractions as predictors. This indicates that PFT fractions serve as a descriptor of vegetation structure and hence as a descriptor of land cover heterogeneity in the global model. This ~~causes~~results in a VOD-LAI relationship that varies by microwave wavelength~~, but~~. But this wavelength-dependency cannot be resolved within ~~a~~the land cover-specific ~~model~~models, because ~~it does not~~those models cannot account for the impact of sub-pixel land cover heterogeneity. Furthermore, the differences in the VOD-AGB relationship between the global and the land cover-specific models also highlights that a monotonic AGB-VOD relationship is only valid over a large spatial scale but does not hold within a vegetation type or at smaller scales. The high model performance in regions with high biomass areas were enabled using PFT maps as predictors, which compensate for the saturating effect at high AGB. Similar to the VOD-LAI relationship, the relative sensitivity of the LFMC-ALE increases with increasing wavelength for the global models and it also shows that LFMC has relative more influence on an 8-daily time scale compared to the monthly time scale for the global as well as in the land cover-specific models.

~~The high R^2 results for high biomass areas were enabled using PFT maps as feature, which compensate for the saturating effect of AGB. Both LFMC and LAI vary in time and space and are strongly correlated. The temporal and spatial variation of our model is dominated by LAI, leading to a decreased influence of LFMC on shortwave VOD, which has a higher short-term variation than L-VOD.~~ Both LFMC and LAI are strongly correlated. The temporal and spatial variation of our global models are dominated by LAI, leading to a lower influence of LFMC on shortwave VOD than of LAI. Although LFMC appears as the less important predictor for VOD than LAI in our models, the strong correlation of LAI and LFMC is nevertheless the reason why in-situ measured LFMC show medium to strong correlations with VOD and can be used to estimate LFMC from short-wavelength VOD (Fan et al., 2018; Forkel et al., 2022).

Globally, the L-band VOD is highly influenced by AGB, which is in agreement with the ability of longwave VOD to better penetrate dense vegetation and its higher sensitivity to the woody plant parts (Liu et al., 2011). However, the much lower predictability of L-VOD compared to Ku-, X-, and C-VOD indicates that L-VOD cannot be sufficiently explained by the combination of AGB, LAI, LFMC, and land cover. The performance in predicting L-VOD is much lower at pixel-level (Figure 3) than computed across the full spatial and temporal extent of the data. Hence, the low performance in predicting L-VOD is mostly related to the temporal dynamics at pixel-level because our model correctly explains the spatial patterns. The low performance in predicting SMOS L-VOD might be caused by a noisy signal of the SMOS sensor (van der Schalie et al., 2017). Especially the daily raw L-VOD data, as used for the 8-daily analyses, can be very noisy (Wigneron et al., 2021). Vittucci et al. (2016) found moderate seasonal differences (but within the standard variation) of the SMOS L-VOD signal over ~~boreal forest,~~forests located at higher latitudes than +20°, which are partly explainable due to the deciduous character of the forest but moreover because of random effects. The L-band signal, and also the C-band signal, is strongly disturbed by radio-frequency interference (RFI, ~~Liu et al., 2019)~~Liu et al., 2019). The spatial and temporal inconsistency of RFI complicates the RFI correction of the L-band (Wigneron et al., 2021). This indicates a noisy, or until now not fully understood, variation of the SMOS L-VOD, especially within the lower value range. Due to the uncertain proportion of noise and short-term changes of water content, Ebrahimi et al. (2018) averaged SMOS L-VOD over 15 days and ~~Rodríguez-Fernández et al.(2018)~~Rodríguez-Fernández et al. (2018) even over 2 years to reduce related uncertainties of the VOD signal. ~~Vaglio Laurin et al. (2020~~Vaglio Laurin et al. (2020) found a time lag of up to 6 months ~~for mostly tree-covered areas in South America and Africa~~ between SMOS L-VOD and ecosystem functional properties~~.~~ in tree-covered areas in South America and Africa. This time lag shows that the ~~relations~~relationships between SMOS L-VOD and vegetation properties need further investigation in densely-vegetated regions. In addition to the possible noisy signal of SMOS L-VOD, which might hamper the interpretation, errors within the L-VOD values can also be introduced by the retrieval algorithm itself. With the use of a tau-omega model, soil moisture and VOD are often retrieved simultaneously which can introduce errors in the VOD retrievals. Zwieback et al. (2019) found spurious correlations of soil moisture and VOD especially for sub-monthly time scales over forests. Besides that, the correctness of the retrieval product focuses on soil moisture at the cost of the VOD retrieval. The resulting error shifts from soil moisture to VOD are more prone to short-term changes and to higher VOD values (Feldman et al., 2021), which might contribute to the underestimation of high VOD values of our models and the reduced performance of the 8-daily models compared to the monthly models. A more robust L-VOD product might be achieved by analysing and adjusting the necessary degree of regularization for a VOD retrieval depending on time scale and land cover (Zwieback et al., 2019; Feldman et al., 2021).

An interesting finding is the higher sensitivity of L-VOD to LFMC than to LAI. This indicates that L-band indeed penetrates deeper in the canopy (low sensitivity to LAI) but is sensitive to the plant water status (i.e. LFMC). However, AGB and LFMC are insufficient predictors to reach high predictability of L-VOD. This might be caused by the fact that the AGB dataset used in this study does not contain any temporal information, and hence changes in AGB are not considered in our model. Using an alternative dataset (e.g. Xu et al., 2021), which provides a global time series of AGB could be a benefit for improving the understanding of temporal VOD variations. However, as we included annual land cover maps as predictors, our models do indeed account for land cover change such as deforestation which is strongly related to a change in AGB (Andela et al., 2013). The use of LFMC and LAI as ~~a predictor~~predictors might be insufficient for L-VOD. The used LFMC and LAI data ~~used in our study was~~were both derived from optical observation by MODIS, which ~~in, the case of closed forest canopies,~~ is only sensitive to the top of the canopy~~.~~ in closed forest canopies. Root-zone soil moisture was used as a proxy for water availability in other studies (e.g. Konings et al., 2021), however, it is not an ideal predictor for vegetation water content, as some plants can regulate their water potential or moisture content independent of soil moisture (Konings and Gentine, 2017; Hochberg et al., 2018). Therefore, it is necessary to further investigate the daily to seasonal temporal dynamics of L-VOD with respect to e.g. local and regional observations of water availability and plant water status.

4.2 Towards developing advanced approaches to link VOD with vegetation properties

~~The long time series, global coverage and multiple frequencies of VOD retrievals provide valuable information or can be used to derive vegetation properties at large scale or to evaluate and parametrize land surface models in data assimilation studies. Those applications of VOD require, however, a solid understanding of the biophysical controls on VOD. The relatively high effect of LAI on the short wavelength VODs indicates that data assimilation approaches that only use LAI for estimating the temporal dynamic of VOD (as they were used by Scholze et al., 2019 and Kumar et al., 2020) are valid approximations. This means that even models without an explicit representation of plant water status are suitable for VOD assimilation, but the observation operators need to take into account non-linear, non-additive, and local effects.~~The long time series, global coverage and multiple frequencies of VOD retrievals provide valuable information that can be used to derive vegetation properties at large scale or to evaluate and parametrize land surface models in data assimilation studies. However, those applications of VOD require a solid understanding of the biophysical controls on VOD. The relatively high effect of LAI on the short wavelength VODs indicates that data assimilation approaches that only use LAI for estimating the temporal dynamic of VOD (as they were used by Scholze et al., 2019 and Kumar et al., 2020) are valid approximations. This means that even models without an explicit representation of plant water status are suitable for VOD assimilation, but the observation operators need to take into account non-linear, non-additive, and local effects.

LFMC or similar measures for plant water status have only recently been introduced into land surface models commonly used for global-scale simulations (e.g. Kennedy et al., 2019; Niu et al., 2020; Eller et al., 2020; Li et al., 2021a). LFMC has therefore not been used in assimilation studies so far. The long time series of especially Ku-VOD could help to constrain model simulations of LFMC or ~~plant water~~support studies of plant water status but ~~require~~requires a good representation of LAI dynamics.

~~For observation operators for L-VOD, AGB should be the main predictor for spatial patterns. For example, Rodríguez-Fernández et al. (2018) modelled C-VOD from AMSR-E as an empirical function of AGB.~~For observation operators for L-VOD, AGB should be the main predictor for spatial patterns. Scholze et al. (2019) used the empirical function between VOD and AGB evaluated by Rodríguez-Fernández et al. (2018), to simulate L-VOD from AGB. Thereby, AGB was replaced with a function of net primary production and effective turnover time. However, temporal changes in L-VOD ~~might result in an overestimation in dynamics of biomass production, turnover or biomass loss if effects of~~that are caused by changes in plant water status are not considered~~might result in an overestimation in dynamics of biomass production, turnover or biomass loss~~ (Konings et al., 2021). Scholze et al. (2019) tried to avoid incorporating short-term changes ~~due to~~in VWC and therefore averaged the VOD simulations to yearly means. The temporal dynamics should include the effect of plant water status, but further investigations on the drivers of the temporal dynamics of L-VOD are necessary to make full use of the data.

Including a proxy for VWC and exploring the influence of short-term changes of vegetation properties on VOD, we assessed the temporal dynamics not only for L-VOD but also for Ku-, X-, and C-VOD, which will help to make explicit use of VOD temporal changes within modelling and assimilation studies~~.~~

Author contribution

LS and MF developed the research idea, objectives and methodology. LS, RvdS and MY prepared datasets. LS implemented and applied the analysis. AKR contributed to the implementation of ALE plots and the interpretation of related results. RMZ and SS helped with interpreting results and reviewed the application of VOD data. LS and MF wrote the initial draft of the manuscript. All authors reviewed and revised the initial draft of the manuscript.

Competing interests

MF is guest editor of the special issue “Microwave remote sensing for improved understanding of vegetation–water interactions”. The peer-review process was guided by an independent editor, and the authors have also no other competing interests to declare.

Disclaimer

Special issue statement

Acknowledgements

The computations were performed on an HPC system at the Center for Information Services and High-Performance Computing (ZIH) at TU Dresden. The authors thank Christopher Marrs for his detailed proofreading of this manuscript.

References

Andela, N., Liu, Y. Y., M. Van Dijk, A. I. J., De Jeu, R. A. M., and McVicar, T. R.: Global changes in dryland vegetation dynamics (1988-2008) assessed by satellite remote sensing: Comparing a new passive microwave vegetation density record with reflective greenness data, 10, 6657–6676, https://doi.org/10.5194/bg-10-6657-2013, 2013.

Apley, D. W. and Zhu, J.: Visualizing the effects of predictor variables in black box supervised learning models, J. R. Stat. Soc. Ser. B Stat. Methodol., 82, 1059–1086, https://doi.org/10.1111/RSSB.12377, 2020.

Baur, M. J., Jagdhuber, T., Feldman, A. F., Akbar, R., and Entekhabi, D.: Estimation of relative canopy absorption and scattering at L-, C- and X-bands, Remote Sens. Environ., 233, 111384, https://doi.org/10.1016/j.rse.2019.111384, 2019.

Al Bitar, A., Mialon, A., Kerr, Y. H., Cabot, F., Richaume, P., Jacquette, E., Quesney, A., Mahmoodi, A., Tarot, S., Parrens, M., Al-Yaari, A., Pellarin, T., Rodriguez-Fernandez, N., and Wigneron, J. P.: The global SMOS Level 3 daily soil moisture and brightness temperature maps, Earth Syst. Sci. Data, 9, 293–315, https://doi.org/10.5194/essd-9-293-2017, 2017.

Bousquet, E., Mialon, A., Rodriguez-Fernandez, N., Prigent, C., Wagner, F. H., and Kerr, Y. H.: Influence of surface water variations on VOD and biomass estimates from passive microwave sensors, Remote Sens. Environ., 257, 112345, https://doi.org/10.1016/j.rse.2021.112345, 2021.

Breiman, L.: Random Forests, Mach. Learn., 45, 5–32, https://doi.org/https://doi.org/10.1023/A:1010933404324, 2001.

Chaparro, D., Piles, M., Vall-llossera, M., Camps, A., Konings, A. G., and Entekhabi, D.: L-band vegetation optical depth seasonal metrics for crop yield assessment, Remote Sens. Environ., 212, 249–259, https://doi.org/10.1016/j.rse.2018.04.049, 2018.

Dorigo, W., Moesinger, L., van der Schalie, R., Zotta, R.-M., Scanlon, T., and Jeu, R. A. M.: Long-term monitoring of vegetation state through passive microwave satellites, in: State of the Climate in 2020, vol. 102 (8), Bulletin of the American Meteorological Society, 110–112, https://doi.org/10.1175/BAMS-D-21-0098.1, 2021.

Du, J., Kimball, J. S., Jones, L. A., Kim, Y., Glassy, J., and Watts, J. D.: A global satellite environmental data record derived from AMSR-E and AMSR2 microwave Earth observations, Earth Syst. Sci. Data, 9, 791–808, https://doi.org/10.5194/essd-9-791-2017, 2017.

Ebrahimi, M., Alavipanah, S. K., Hamzeh, S., Amiraslani, F., Neysani Samany, N., and Wigneron, J. P.: Exploiting the synergy between SMAP and SMOS to improve brightness temperature simulations and soil moisture retrievals in arid regions, J. Hydrol., 557, 740–752, https://doi.org/10.1016/j.jhydrol.2017.12.051, 2018.

Eller, C. B., Rowland, L., Mencuccini, M., Rosas, T., Williams, K., Harper, A., Medlyn, B. E., Wagner, Y., Klein, T., Teodoro, G. S., Oliveira, R. S., Matos, I. S., Rosado, B. H. P., Fuchs, K., Wohlfahrt, G., Montagnani, L., Meir, P., Sitch, S., and Cox, P. M.: Stomatal optimization based on xylem hydraulics (SOX) improves land surface model simulation of vegetation responses to climate, New Phytol., 226, 1622–1637, https://doi.org/10.1111/NPH.16419, 2020.

ESA: Land Cover CCI Product User Guide Version 2. Tech. Rep., 2017.

Fan, L., Wigneron, J. P., Xiao, Q., Al-Yaari, A., Wen, J., Martin-StPaul, N., Dupuy, J. L., Pimont, F., Al Bitar, A., Fernandez-Moran, R., and Kerr, Y. H.: Evaluation of microwave remote sensing for monitoring live fuel moisture content in the Mediterranean region, Remote Sens. Environ., 205, 210–223, https://doi.org/10.1016/J.RSE.2017.11.020, 2018.

Feldman, A., Chaparro, D., and Entekhabi, D.: Error Propagation in Microwave Soil Moisture and Vegetation Optical Depth Retrievals, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 14, 11311–11323, https://doi.org/10.1109/JSTARS.2021.3124857, 2021.

Feldman, A. F., Short Gianotti, D. J., Trigo, I. F., Salvucci, G. D., and Entekhabi, D.: Land-Atmosphere Drivers of Landscape-Scale Plant Water Content Loss, Geophys. Res. Lett., 47, https://doi.org/10.1029/2020GL090331, 2020.

Fernandez-Moran, R., Al-Yaari, A., Mialon, A., Mahmoodi, A., Al Bitar, A., De Lannoy, G., Rodriguez-Fernandez, N., Lopez-Baeza, E., Kerr, Y., and Wigneron, J. P.: SMOS-IC: An alternative SMOS soil moisture and vegetation optical depth product, Remote Sens., 9, 1–21, https://doi.org/10.3390/rs9050457, 2017.

Forkel, M., Schmidt, L., Zotta, R., Dorigo, W., and Yebra, M.: Estimating leaf moisture content at global scale from passive microwave satellite observations of vegetation optical depth, Hydrol. Earth Syst. Sci. Discuss. [preprint], in review, https://doi.org/https://doi.org/10.5194/hess-2022-121, 2022.

Frappart, F., Wigneron, J. P., Li, X., Liu, X., Al-Yaari, A., Fan, L., Wang, M., Moisy, C., Le Masson, E., Lafkih, Z. A., Vallé, C., Ygorra, B., and Baghdadi, N.: Global Monitoring of the Vegetation Dynamics from the Vegetation Optical Depth (VOD): A Review, Remote Sens. 2020, Vol. 12, Page 2915, 12, 2915, https://doi.org/10.3390/RS12182915, 2020.

Friedman, J. H.: Greedy function approximation: A gradient boosting machine., 29, 1189–1232, https://doi.org/10.1214/AOS/1013203451, 2001.

Hastie, T. and Tibshirani, R.: Generalized additive models: Some applications, J. Am. Stat. Assoc., 82, 371–386, https://doi.org/10.1080/01621459.1987.10478440, 1987.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning, https://doi.org/10.1007/978-0-387-84858-7, 2009.

Hochberg, U., Rockwell, F. E., Holbrook, N. M., and Cochard, H.: Iso/Anisohydry: A Plant–Environment Interaction Rather Than a Simple Hydraulic Trait, Trends Plant Sci., 23, 112–120, https://doi.org/10.1016/j.tplants.2017.11.002, 2018.

Hutengs, C. and Vohland, M.: Downscaling land surface temperatures at regional scales with random forest regression, Remote Sens. Environ., 178, 127–141, https://doi.org/10.1016/j.rse.2016.03.006, 2016.

Jackson, T. J. and Schmugge, T. J.: Vegetation effects on the microwave emission of soils, Remote Sens. Environ., 36, 203–212, https://doi.org/10.1016/0034-4257(91)90057-D, 1991.

Jackson, T. J., Schmugge, T. J., and Wang, J. R.: Passive microwave sensing of soil moisture under vegetation canopies, Water Resour. Res., 18, 1137–1142, https://doi.org/10.1029/WR018I004P01137, 1982.

Jones, M. O., Jones, L. A., Kimball, J. S., and McDonald, K. C.: Satellite passive microwave remote sensing for monitoring global land surface phenology, Remote Sens. Environ., 115, 1102–1114, https://doi.org/10.1016/j.rse.2010.12.015, 2011.

Jones, M. O., Kimball, J. S., Small, E. E., and Larson, K. M.: Comparing land surface phenology derived from satellite and GPS network microwave remote sensing, Int. J. Biometeorol., 58, 1305–1315, https://doi.org/10.1007/s00484-013-0726-z, 2014.

Kennedy, D., Swenson, S., Oleson, K. W., Lawrence, D. M., Fisher, R., Lola da Costa, A. C., and Gentine, P.: Implementing Plant Hydraulics in the Community Land Model, Version 5, J. Adv. Model. Earth Syst., 11, 485–513, https://doi.org/10.1029/2018MS001500, 2019.

Konings, A. G. and Gentine, P.: Global variations in ecosystem-scale isohydricity, Glob. Chang. Biol., 23, 891–905, https://doi.org/10.1111/GCB.13389, 2017.

Konings, A. G., Piles, M., Das, N., and Entekhabi, D.: L-band vegetation optical depth and effective scattering albedo estimation from SMAP, Remote Sens. Environ., 198, 460–470, https://doi.org/10.1016/j.rse.2017.06.037, 2017.

Konings, A. G., Rao, K., and Steele-Dunne, S. C.: Macro to micro: microwave remote sensing of plant water content for physiology and ecology, New Phytol., 223, 1166–1172, https://doi.org/10.1111/NPH.15808, 2019a.

Konings, A. G., Rao, K., and Steele-Dunne, S. C.: Macro to micro: microwave remote sensing of plant water content for physiology and ecology, New Phytol., 223, 1166–1172, https://doi.org/10.1111/nph.15808, 2019b.

Konings, A. G., Holtzman, N., Rao, K., Xu, L., and Saatchi, S. S.: Interannual Variations of Vegetation Optical Depth Are Due to Both Water Stress and Biomass Changes, Geophys. Res. Lett., 48, 1–9, https://doi.org/10.1029/2021gl095267, 2021.

Kuhn-Régnier, A., Voulgarakis, A., Nowack, P., Forkel, M., Prentice, I. C., and Harrison, S. P.: The importance of antecedent vegetation and drought conditions as global drivers of burnt area, 18, 3861–3879, https://doi.org/10.5194/bg-18-3861-2021, 2021.

Kumar, S. V., Holmes, T. R., Bindlish, R., De Jeu, R., and Peters-Lidard, C.: Assimilation of vegetation optical depth retrievals from passive microwave radiometry, Hydrol. Earth Syst. Sci., 24, 3431–3450, https://doi.org/10.5194/HESS-24-3431-2020, 2020.

Li, L., Yang, Z. L., Matheny, A. M., Zheng, H., Swenson, S. C., Lawrence, D. M., Barlage, M., Yan, B., McDowell, N. G., and Leung, L. R.: Representation of Plant Hydraulics in the Noah-MP Land Surface Model: Model Development and Multiscale Evaluation, J. Adv. Model. Earth Syst., 13, e2020MS002214, https://doi.org/10.1029/2020MS002214, 2021a.

Li, X., Wigneron, J.-P., Frappart, F., Fan, L., Ciais, P., Fensholt, R., Entekhabi, D., Brandt, M., Konings, A. G., Liu, X., Wang, M., Al-Yaari, A., and Moisy, C.: Global-scale assessment and inter-comparison of recently developed/reprocessed microwave satellite vegetation optical depth products, Remote Sens. Environ., 253, 112208, https://doi.org/10.1016/j.rse.2020.112208, 2021b.

Liang, L., Di, L., Huang, T., Wang, J., Lin, L., Wang, L., and Yang, M.: Estimation of leaf nitrogen content in wheat using new hyperspectral indices and a random forest regression algorithm, Remote Sens., 10, https://doi.org/10.3390/rs10121940, 2018.

Liu, R., Wen, J., Wang, X., Wang, Z., Li, Z., Xie, Y., Zhu, L., and Li, D.: Derivation of Vegetation Optical Depth and Water Content in the Source Region of the Yellow River using the FY-3B Microwave Data, Remote Sens., 11, 1536, https://doi.org/10.3390/rs11131536, 2019.

Liu, Y., Holtzman, N. M., and Konings, A. G.: Global ecosystem-scale plant hydraulic traits retrieved using model-data fusion, Hydrol. Earth Syst. Sci., 25, 2399–2417, https://doi.org/10.5194/HESS-25-2399-2021, 2021.

Liu, Y. Y., De Jeu, R. A. M., McCabe, M. F., Evans, J. P., and Van Dijk, A. I. J. M.: Global long-term passive microwave satellite-based retrievals of vegetation optical depth, Geophys. Res. Lett., 38, 1–6, https://doi.org/10.1029/2011GL048684, 2011.

Liu, Y. Y., Van Dijk, A. I. J. M., De Jeu, R. A. M., Canadell, J. G., McCabe, M. F., Evans, J. P., and Wang, G.: Recent reversal in loss of global terrestrial biomass, Nat. Clim. Chang., 5, 470–474, https://doi.org/10.1038/nclimate2581, 2015.

Mialon, A., Rodríguez-Fernández, N. J., Santoro, M., Saatchi, S., Mermoz, S., Bousquet, E., and Kerr, Y. H.: Evaluation of the sensitivity of SMOS L-VOD to forest above-ground biomass at global scale, Remote Sens., 12, 1–10, https://doi.org/10.3390/RS12091450, 2020.

Moesinger, L., Dorigo, W., De Jeu, R., Van Der Schalie, R., Scanlon, T., Teubner, I., and Forkel, M.: The global long-term microwave Vegetation Optical Depth Climate Archive (VODCA), Earth Syst. Sci. Data, 12, 177–196, https://doi.org/10.5194/ESSD-12-177-2020, 2020.

Moesinger, L., Zotta, R.-M., Schalie, R. van der, Scanlon, T., Jeu, R. de, and Dorigo, W. A.: Monitoring Vegetation Condition using Microwave Remote Sensing: The Standardized Vegetation Optical Depth Index SVODI, [preprint], in review, https://doi.org/https://doi.org/10.5194/bg-2021-360, 2022.

Momen, M., Wood, J. D., Novick, K. A., Pangle, R., Pockman, W. T., McDowell, N. G., and Konings, A. G.: Interacting Effects of Leaf Water Potential and Biomass on Vegetation Optical Depth, J. Geophys. Res. Biogeosciences, 122, 3031–3046, https://doi.org/10.1002/2017JG004145, 2017.

Myneni, R., Knyazikhin, Y., Park, T.: MOD15A2H MODIS Leaf Area Index/FPAR 8-Day L4 Global 500m SIN Grid V006, NASA EOSDIS L. Process. DAAC, https://doi.org/http://doi.org/10.5067/MODIS/MOD15A2H.006, 2015.

Nelder, J. A., & Wedderburn, R. W. M.: Generalized Linear Models Why Generalized Linear Models ?, J. R. Stat. Soc., 135, 370–384, https://doi.org/10.2307/2344614, 1972.

Niu, G. Y., Fang, Y. H., Chang, L. L., Jin, J., Yuan, H., and Zeng, X.: Enhancing the Noah-MP Ecosystem Response to Droughts With an Explicit Representation of Plant Water Storage Supplied by Dynamic Root Water Uptake, J. Adv. Model. Earth Syst., 12, e2020MS002062, https://doi.org/10.1029/2020MS002062, 2020.

Njoku, E. G. and Entekhabi, D.: Passive microwave remote sensing of soil moisture, J. Hydrol., 184, 101–129, https://doi.org/10.1016/0022-1694(95)02970-2, 1996.

Owe, M., de Jeu, R., and Holmes, T.: Multisensor historical climatology of satellite-derived global land surface moisture, J. Geophys. Res. Earth Surf., 113, 1–17, https://doi.org/10.1029/2007JF000769, 2008.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, J. Mach. Learn. Res., 12, 2825--2830, https://doi.org/https://doi.org/10.48550/arXiv.1201.0490, 2011.

Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., Bontemps, S., Boettcher, M., Brockmann, C., Defourny, P., Hagemann, S., Herold, M., Kirches, G., Lamarche, C., Lederer, D., Ottlé, C., Peters, M., and Peylin, P.: Plant functional type classification for earth system models: Results from the European Space Agency's Land Cover Climate Change Initiative, Geosci. Model Dev., 8, 2315–2328, https://doi.org/10.5194/gmd-8-2315-2015, 2015.

Rao, K., Williams, A. P., Diffenbaugh, N. S., Yebra, M., and Konings, A. G.: Plant-water sensitivity regulates wildfire vulnerability, Nat. Ecol. Evol., 6, 332–339, https://doi.org/10.1038/s41559-021-01654-2, 2022.

Rodríguez-Fernández, N. J., Mialon, A., Mermoz, S., Bouvet, A., Richaume, P., Al Bitar, A., Al-Yaari, A., Brandt, M., Kaminski, T., Le Toan, T., Kerr, Y. H., and Wigneron, J. P.: An evaluation of SMOS L-band vegetation optical depth (L-VOD) data sets: High sensitivity of L-VOD to above-ground biomass in Africa, 15, 4627–4645, https://doi.org/10.5194/bg-15-4627-2018, 2018.

Saleh, K., Wigneron, J. P., De Rosnay, P., Calvet, J. C., Escorihuela, M. J., Kerr, Y., and Waldteufel, P.: Impact of rain interception by vegetation and mulch on the L-band emission of natural grass, Remote Sens. Environ., 101, 127–139, https://doi.org/10.1016/j.rse.2005.12.004, 2006.

Santoro, M. and Cartus, O.: ESA Biomass Climate Change Initiative (Biomass_cci): Global datasets of forest above-ground biomass for the year 2017, Cent. Environ. Data Anal., v1, https://doi.org/10.5285/bedc59f37c9545c981a839eb552e4084, 2019.

Sawada, Y., Tsutsui, H., Koike, T., Rasmy, M., Seto, R., and Fujii, H.: A field verification of an algorithm for retrieving vegetation water content from passive microwave observations, IEEE Trans. Geosci. Remote Sens., 54, 2082–2095, https://doi.org/10.1109/TGRS.2015.2495365, 2016.

van der Schalie, R., Kerr, Y. H. H., Wigneron, J. P. P., Rodríguez-Fernández, N. J. J., Al-yaari, A., and Jeu, R. A. M. d.: Global SMOS Soil Moisture Retrievals from The Land Parameter Retrieval Model, Int. J. Appl. Earth Obs. Geoinf., 45, 125–134, https://doi.org/https://doi.org/10.1016/j.jag.2015.08.005, 2016.

van der Schalie, R., de Jeu, R. A. M., Kerr, Y. H., Wigneron, J. P., Rodríguez-Fernández, N. J., Al-Yaari, A., Parinussa, R. M., Mecklenburg, S., and Drusch, M.: The merging of radiative transfer based surface soil moisture data from SMOS and AMSR-E, Remote Sens. Environ., 189, 180–193, https://doi.org/10.1016/j.rse.2016.11.026, 2017.

Scholze, M., Kaminski, T., Knorr, W., Voßbeck, M., Wu, M., Ferrazzoli, P., Kerr, Y., Mialon, A., Richaume, P., Rodríguez-Fernández, N., Vittucci, C., Wigneron, J. P., Mecklenburg, S., and Drusch, M.: Mean European Carbon Sink Over 2010–2015 Estimated by Simultaneous Assimilation of Atmospheric CO2, Soil Moisture, and Vegetation Optical Depth, Geophys. Res. Lett., 46, 13796–13803, https://doi.org/10.1029/2019GL085725, 2019.

Servén, D., Brummitt, C., Abedi, H., and hlink: dswah/pyGAM: v0.8.0, https://doi.org/10.5281/ZENODO.1476122, 2018.

Teubner, I. E., Forkel, M., Jung, M., Liu, Y. Y., Miralles, D. G., Parinussa, R., van der Schalie, R., Vreugdenhil, M., Schwalm, C. R., Tramontana, G., Camps-Valls, G., and Dorigo, W. A.: Assessing the relationship between microwave vegetation optical depth and gross primary production, Int. J. Appl. Earth Obs. Geoinf., 65, 79–91, https://doi.org/10.1016/j.jag.2017.10.006, 2018.

Teubner, I. E., Forkel, M., Camps-Valls, G., Jung, M., Miralles, D. G., Tramontana, G., van der Schalie, R., Vreugdenhil, M., Mösinger, L., and Dorigo, W. A.: A carbon sink-driven approach to estimate gross primary production from microwave satellite observations, Remote Sens. Environ., 229, 100–113, https://doi.org/10.1016/j.rse.2019.04.022, 2019.

Teubner, I. E., Forkel, M., Wild, B., Mösinger, L., and Dorigo, W.: Impact of temperature and water availability on microwave-derived gross primary production, 18, 3285–3308, https://doi.org/10.5194/bg-18-3285-2021, 2021.

Togliatti, K., Lewis-Beck, C., Walker, V. A., Hartman, T., Vanloocke, A., Cosh, M. H., and Hornbuckle, B. K.: Quantitative Assessment of Satellite L-Band Vegetation Optical Depth in the U.S. Corn Belt, IEEE Geosci. Remote Sens. Lett., 19, https://doi.org/10.1109/LGRS.2020.3034174, 2022.

Vaglio Laurin, G., Vittucci, C., Tramontana, G., Ferrazzoli, P., Guerriero, L., and Papale, D.: Monitoring tropical forests under a functional perspective with satellite-based vegetation optical depth, Glob. Chang. Biol., 26, 3402–3416, https://doi.org/10.1111/gcb.15072, 2020.

Vittucci, C., Ferrazzoli, P., Kerr, Y., Richaume, P., Guerriero, L., Rahmoune, R., and Laurin, G. V.: SMOS retrieval over forests: Exploitation of optical depth and tests of soil moisture estimates, Remote Sens. Environ., 180, 115–127, https://doi.org/10.1016/j.rse.2016.03.004, 2016.

Wang, J. R.: Effect of vegetation on soil moisture sensing observed from orbiting microwave radiometers, Remote Sens. Environ., 17, 141–151, https://doi.org/10.1016/0034-4257(85)90070-7, 1985.

Wang, M., Wigneron, J.-P., Sun, R., Fan, L., Frappart, F., Tao, S., Chai, L., Li, X., Liu, X., Ma, H., Moisy, C., and Ciais, P.: A consistent record of vegetation optical depth retrieved from the AMSR-E and AMSR2 X-band observations, Int. J. Appl. Earth Obs. Geoinf., 105, 102609, https://doi.org/10.1016/J.JAG.2021.102609, 2021.

Wigneron, J. P., Calvet, J. C., Kerr, Y., Chanzy, A., and Lopes, A.: Microwave Emission of Vegetation: Sensitivity to Leaf Characteristics, IEEE Trans. Geosci. Remote Sens., 31, 716–726, https://doi.org/10.1109/36.225537, 1993.

Wigneron, J. P., Calvet, J. C., and Kerr, Y.: Monitoring water interception by crop fields from passive microwave observations, Agric. For. Meteorol., 80, 177–194, https://doi.org/10.1016/0168-1923(95)02296-1, 1996.

Wigneron, J. P., Mialon, A., De Lannoy, G., Fernandez-Moran, R., Al-Yaari, A., Ebrahimi, M., Rodriguez-Fernandez, N., Kerr, Y., Quets, J., Pellarin, T., Fan, L., Tian, F., Fensholt, R., and Brandt, M.: SMOS-IC: Current status and overview of soil moisture and VOD applications, Int. Geosci. Remote Sens. Symp., 2018-July, 1451–1454, https://doi.org/10.1109/IGARSS.2018.8519382, 2018.

Wigneron, J. P., Li, X., Frappart, F., Fan, L., Al-Yaari, A., De Lannoy, G., Liu, X., Wang, M., Le Masson, E., and Moisy, C.: SMOS-IC data record of soil moisture and L-VOD: Historical development, applications and perspectives, Remote Sens. Environ., 254, 112238, https://doi.org/10.1016/J.RSE.2020.112238, 2021.

Wild, B., Teubner, I., Moesinger, L., Zotta, R.-M., Forkel, M., van der Schalie, R., Sitch, S., and Dorigo, W.: VODCA2GPP – a new, global, long-term (1988–2020) gross primary production dataset from microwave remote sensing, Earth Syst. Sci. Data, 14, 1063–1085, https://doi.org/10.5194/ESSD-14-1063-2022, 2022.

Xu, L., Saatchi, S. S., Yang, Y., Yu, Y., Pongratz, J., Anthony Bloom, A., Bowman, K., Worden, J., Liu, J., Yin, Y., Domke, G., McRoberts, R. E., Woodall, C., Nabuurs, G. J., De-Miguel, S., Keller, M., Harris, N., Maxwell, S., and Schimel, D.: Changes in global terrestrial live biomass over the 21st century, Sci. Adv., 7, https://doi.org/http://doi.org/10.1126/sciadv.abe9829, 2021.

Yebra, M., Quan, X., Riaño, D., Rozas Larraondo, P., van Dijk, A. I. J. M., and Cary, G. J.: A fuel moisture content and flammability monitoring methodology for continental Australia based on optical remote sensing, Remote Sens. Environ., 212, 260–272, https://doi.org/10.1016/j.rse.2018.04.053, 2018.

Yee, T. W. and Mitchell, N. D.: Generalized additive models in plant ecology, J. Veg. Sci., 2, 587–602, https://doi.org/10.2307/3236170, 1991.

Zhang, Y., Zhou, S., Gentine, P., and Xiao, X.: Can vegetation optical depth reflect changes in leaf water potential during soil moisture dry-down events?, Remote Sens. Environ., 234, 111451, https://doi.org/10.1016/J.RSE.2019.111451, 2019.

Zwieback, S., Bosch, D. D., Cosh, M. H., Starks, P. J., and Berg, A.: Vegetation–soil moisture coupling metrics from dual-polarization microwave radiometry using regularization, Remote Sens. Environ., 231, 111257, https://doi.org/10.1016/j.rse.2019.111257, 2019.