# The Global Long-term Microwave Vegetation Optical Depth Climate Archive VODCA

Leander Moesinger[1], Wouter Dorigo[1], Richard de Jeu[2], Robin van der Schalie[2], Tracy Scanlon[1], Irene Teubner[1], and Matthias Forkel[1]

[1]Vienna University of Technology, Department of Geodesy and Geoinformation, Gußhausstraße 27-29, 1040 Vienna, Austria
[2]VanderSat, Wilhelminastraat 43A, 2011 VK Haarlem, The Netherlands

**Correspondence:** Leander Moesinger (Leander.Moesinger@geo.tuwien.ac.at, vodca@geo.tuwien.ac.at)

**Abstract.** Since the late 1970s, spaceborne microwave sensors have been providing measurements of radiation emitted by the Earth's surface. From these measurements it is possible to derive vegetation optical depth (VOD), a model-based indicator related to vegetation density and its relative water content. Because of its high temporal resolution and long availability, VOD can be used to monitor short- to long-term changes in vegetation. However, studying long-term VOD dynamics is generally

5    hampered by the relatively short time span covered by the individual microwave sensors. This can potentially be overcome by merging multiple VOD products into a single climate data record. But, combining multiple sensors into a single product is challenging as systematic differences between input products, e.g. biases, different temporal and spatial resolutions and coverage, need to be overcome.

10    Here, we present a new series of long-term VOD products, which combine multiple VOD data sets derived from several sensors (SSM/I, TMI, AMSR-E, Windsat, and AMSR-2) using the Land Parameter Retrieval Model. We produce separate VOD products for microwave observations in different spectral bands, namely Ku-band (period 1987-2017), X-band (1997-2018), and C-band (2002-2018). In this way, our multi-band VOD products preserve the unique characteristics of each frequency with respect to the structural elements of the canopy. Our approach to merge the single-sensor VOD products is similar to the one of

15    the ESA CCI Soil Moisture products (Liu et al., 2012; Dorigo et al., 2017): First, the data sets are co-calibrated via cumulative distribution function matching using AMSR-E as scaling reference. We apply a new matching technique that scales outliers more robustly than ordinary piece-wise linear interpolation. Second, we aggregate the data sets by taking the arithmetic mean between temporally overlapping observations of the scaled data, generating a VOD Climate Archive (VODCA)

20    The characteristics of VODCA are assessed for self-consistency and against other products: spatio-temporal patterns and anomalies of the merged products show consistency between frequencies and both with observations of Leaf Area Index derived from the MODIS instrument as well as Vegetation Continuous Fields from AVHRR instruments. Trend analysis shows that since 1987 there has been a decline in VOD in the tropics and in large parts parts of east-central and north Asia along with a strong increase in India, large parts of Australia, south Africa, southeastern China and central north America. Using an

25    autocorrelation analysis, we show that the merging of the multiple data sets successfully reduces the random error compared

to the input data sets. In summary, VODCA shows vast potential for monitoring spatio-temporal ecosystem behaviour complementary to existing long-term vegetation products from optical remote sensing.

The VODCA products (Moesinger et al., 2019) are open access and available under Attribution 4.0 International at https://doi.org/10.5281/zenodo.2575599

# 1   Introduction

Vegetation attenuates microwave radiation that is emitted or reflected by the Earth surface. The degree of attenuation can be derived from microwave satellite observations and is commonly referred to as Vegetation optical depth (VOD) (Jackson and Schmugge, 1991). The amount of attenuation depends on various factors, e.g. the density and type of vegetation, and the wavelength of observation (Jackson and Schmugge, 1991; Owe et al., 2008). Short wavelengths experience a higher attenuation by vegetation (and hence relate to higher VOD values) than longer ones (Liu et al., 2009; Owe et al., 2008). As a consequence, VOD estimates from long wavelengths contain more information on deeper vegetation layers (e.g. stems) than VOD estimates from short wavelengths (Chaparro et al., 2018). VOD increases with the Vegetation Water Content (VWC) (Jackson and Schmugge, 1991) and therefore by extension is related to the Above-Ground dry Biomass (AGB) (Liu et al., 2015) and its Relative Water Content (RWC) (Momen et al., 2017).

Satellite-derived VOD has a wide range of potential applications, including biomass monitoring (Liu et al., 2015; Brandt et al., 2018), drought monitoring (Liu et al., 2018), phenology analyzes (Jones et al., 2011) and fire risk management (Fan et al., 2018). VOD also correlates with various optical remote sensing indicators of plant productivity, e.g., Gross Primary Production (GPP) (Teubner et al., 2018), Leaf Area Index (LAI) (Vreugdenhil et al., 2017), Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index, and Normalized Difference Water Index (Grant et al., 2016). VOD has some distinct advantages over optical vegetation indexes, such as a slower saturation and the resulting higher sensitivity to high biomass areas such as rain-forests (Liu et al., 2015), or the ability to be retrieved despite of cloud cover (Liu et al., 2011a).

VOD products have been derived from multiple spaceborne microwave sensors that have been in orbit since the late 1970s (Owe et al., 2008). These sensors have varying lifetimes and characteristics, resulting e.g. from differences in microwave frequency used, measurement incidence angles, orbit characteristics, radiometric quality and spatial footprints. This complicates their joint use in studying long-term VOD dynamics. To overcome this issue, Liu et al. (2011a) proposed a long-term (1987-2008) harmonised multi-sensor VOD data set by merging VOD products derived from the Special Sensor Microwave/Imager (SSM/I), the Microwave Imager onboard the Tropical Rainfall Measuring Mission (TMI), and the Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E) through the Land Parameter Retrieval Method (LPRM; Owe et al. (2008). Their methodology was inherited from the methodology used to produce the first long-term satellite-based climate

data record of soil moisture within the the Climate Change Initiative of the European Space Agency (ESA CCI Soil moisture; (Dorigo et al., 2017, 2012; Liu et al., 2011b, 2012; Gruber et al., 2019)). In their methodology, all available observations were harmonised with respect to C-Band (6.9 GHz) VOD observations from AMSR-E, which was assumed to provide the highest quality observations (Liu et al., 2012). Only in periods where AMSR-E C-band observations were not available, other products

5 were used instead. This approach ignores the fact that in a statistical sense a high quality product can be fused with a low quality product to create a product with a higher quality than either of the original products. This was systematically demonstrated for the merging of two level 2 soil moisture products (Gruber et al., 2017). Since the release of the multi-satellite VOD product by Liu et al. (2011a), significant progress has been made towards a better understanding of the VOD signal, and it was shown that also the individual bands carry valuable information for different applications (Teubner et al., 2018; Chaparro et al., 2018),

10 which prompts the generation of frequency-specific VOD data sets. In addition, new sensors have been launched, allowing the observational VOD records to be extended to the running present.

In this paper, we present a new series of long-term, harmonised VOD climate data records, derived from multiple single-sensor level 2 products. Our data sets use a similar core methodology as in Liu et al. (2011a) and in ESA CCI Soil Moisture

15 but significantly progresses the current state-of-the-art, based on the insights on VOD and climate data record production gathered during the last few years, and by introducing recent satellite missions. We combine VOD observations from SSM/I, TMI, AMSR-E, WindSat, AMSR2 into global, harmonised long-term VOD products at a 0.25° spatial sampling and covering the period 1987-2018. Similar to the approach of Liu et al. (2011a), we use Cumulative Distribtion Function (CDF) matching techniques to scale all VOD data sets to the distribution of AMSR-E. However, since VOD is a function of the frequency and,

20 consequently, different canopy information is stored in different bands, we do not amalgamate all bands but instead produce three individual products: one for Ku-band, one for X-band, and one for C-band. Even though VOD products also exist in L-band derived from the sensors SMAP and SMOS, preliminary analysis concluded that their short overlapping time span and low temporal correlation between them do not warrant a product for it. Also, we discuss the susceptibleness of piece-wise linear CDF-matching to extreme values and as solution propose a new hybrid CDF-matching technique that scales very high

25 and very low values more robustly. Besides, in contrast to Liu et al. (2011a) we merge the scaled observations by taking the arithmetic mean whenever more than one observation is available. From a statistical perspective this should lead to data sets with reduced random error compared to the input products.
First, we describe the input VOD data sets, followed by an overview of the fusion methodology. We then describe the main characteristics of the merged data sets in terms of spatial and temporal coverage and patterns, and their random error character-

30 istics. The spatio-temporal characteristics are checked for plausibility by comparing them to those of related satellited-derived biogeophysical products, e.g. Vegetation Continuous Fields (VCF) and LAI. We complement the data set assessment by a trend analysis. The assessment will be concluded with the discussion on current limitations and ways forward.

## 2 Input data

### 2.1 VOD data sets

#### 2.1.1 The land-parameter retrieval model (LPRM)

LPRM (van der Schalie et al., 2017; Owe et al., 2008; Meesters et al., 2005) retrieves soil moisture and VOD at the same time

5 from vertical and horizontal polarized microwave data and is based on a radiative transfer model (Mo et al., 1982). The model

assumes that the earth emits microwave radiation depending on its surface temperature $T_s$ and emissivity $e$ which is a function

of its dielectric constant $k$, which in turn is dependent on the surface soil moisture. Part of this radiation is then absorbed or

scattered by water in the vegetation depending on its transmissivity $\Gamma$ and single scattering albedo $w$ while the vegetation itself

also emits radiation depending on its temperature $T_v$. The resulting brightness temperature $T_b$ measured at the sensor can then

10 be modeled as

$$T_{bp} = T_s e_p \Gamma + (1 - \Gamma) T_v (1 - w) + (1 - e_p)(1 - w) T_v (1 - \Gamma) \Gamma \tag{1}$$

where the subscript $p$ denotes either a vertical or horizontal polarization. Further, VOD ($\tau$) is related to $\Gamma$ and the incidence

angle $u$ by:

$$\Gamma = \exp\left(\frac{-\tau}{\cos(u)}\right) \tag{2}$$

15 Since observations from the sensors used in this study are available in both horizontal and vertical polarization, eq. 1 is used

to open a system of linear equations. While the the absolute measured $T_{bH}$ is lower than $T_{bV}$, it is more sensitive to changes

in soil moisture while $T_{bV}$ is more sensitive to vegetation and surface soil temperature. This relationship in combination with

the application of a separate retrieval algorithm to determine the temperature from 37-GHz vertical polarization measurements

(Holmes et al., 2009) allows to solve the system analytically as described in Meesters et al. (2005).

20 The actual temperatures are difficult to estimate during daytime due to surface heating, while during nighttime, soil and vege-

tation are in a near thermal equilibrium. This implies that nighttime retrievals are expected to have a lower temperature-related

error than daytime retrievals (Owe et al., 2008). Therefore, to minimize error sources, only nighttime retrievals are used in

VODCA.

#### 2.1.2 Sensor specifications

25 The used VOD data sets were derived from brightness temperature measurements of various spaceborne sensors active since

1987 (Fig. 1).

The Advanced Microwave Scanning Radiometer (AMSR-E) onboard AQUA retrieved microwave observations from 2002 to

2011 in six bands, of which we only consider the C-, X-, and Ku-band. Their spatial footprint is $75 \times 43$ km, $51 \times 29$ km and

$27 \times 16$ km respectively. AQUA is on an sun-synchronous circular orbit, passing the equator at 1:30 PM ascending and 1:30

30 AM descending mode (Knowles et al., 2006; Kawanishi et al., 2003).

The Advanced Microwave Scanning Radiometer 2 (AMSR2) is an improved version of AMSR-E onboard GCOM-W1 continuing AMSR-E's measurements since 2012 with similar bands, orbit and overpass times but with a slightly higher spatial resolution: $62 \times 35$ km, $42 \times 24$ km and $22 \times 14$ km, for C-, X-, and Ku-band respectively. In addition, AMSR2 also contains a second C-band (7.3 GHz) that can be used to cover areas where RFI is present in the primary C-band channel (6.9 GHz) (Meier

5   et al., 2018). During preliminary analysis, we discovered that the AMSR2 Ku-band VOD retrievals have an apparent break in late 2017. Since then, the values observed are globally systematically lower than before, indicating possibly a calibration error in Ku-band brightness temperatures. While the exact reasons are unknown to us, until the matter is resolved we do not include Ku-band data after 2017-08-01 into VODCA. This shortens the Ku-band VOD product by 16 months. VOD retrievals from X- and C-band AMSR2 seem unaffected and are used until the end of 2018.

10   The Special Sensor Microwave Imager (SSM/I) is onboard a series of DMSP satellites. We use the VOD data retrieved from F-8, F-11 and F-13. From the 7 available bands of SSM/I we use only VODfrom Ku-band which has a resolution of $69 \times 43$ km. The equatorial crossing time varies between the DMSP satellites, but all are on sun-synchronous orbits (Wentz, 1997).

Among other sensors the Tropical Rainfall Measuring Mission (TRMM) carried the TRMM Microwave Imager (TMI). TRMM is the only satellite used which has a non-near-polar orbit with an inclination of 35 degrees. Up to 2001 it had an altitude of 350

15   km, which then got boosted to 400 km leading to a slight decrease in spatial resolution. TMI was active from 1997 to 2015. Of the 9 channels we only use its X- and Ku-band, which has a spatial resolution of $63 \times 37$ km / $72 \times 43$ km and $30 \times 18$ km / $35 \times 21$ km pre/post boost, respectively (Kummerow et al., 1998).

WindSat onboard Coriolis was launched in 2003 on a sun-synchronous orbit providing radiometric measurements in five bands, of which the C-, X- and Ku-band were used to derive VOD. The spatial resolution is $39 \times 71$ km, $25 \times 38$ km and $16 \times 27$

20   km. Due to some periods of non-operation, WindSat contains temporal data gaps (Gaiser et al., 2004). Unfortunately we were unable to gain access to data past July 2012, even though WindSat is still operational.
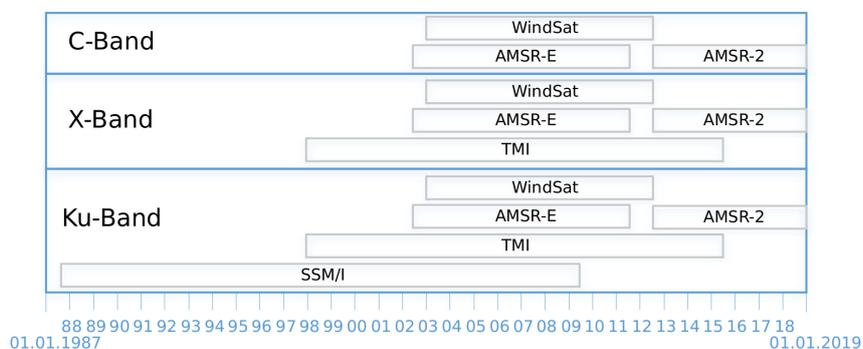
**Figure 1.** Time periods of the sensors used for each band.

**Table 1.** The input VOD data sets used with their temporal coverage and used frequencies for each product.

| Sensor | Time period used | C-Band | X-Band | Ku-Band | Reference |
|---|---|---|---|---|---|
| AMSR-E | Jun 2002 - Oct 2011 | 6.93 | 10.65 | 18.7 | van der Schalie et al. (2017) |
| AMSR2 | Jul 2012 - Jan 2019 | 6.93 & 7.3 | 10.65 | 18.7 | van der Schalie et al. (2017) |
| SSM/I F08 | Jul 1987 - Dec 1991 | | | 19.35 | Owe et al. (2008) |
| SSM/I F11 | Dec 1991 - May 1995 | | | 19.35 | Owe et al. (2008) |
| SSM/I F13 | May 1995 - Apr 2009 | | | 19.35 | Owe et al. (2008) |
| TMI | Dec 1997 - Apr 2015 | | 10.65 | 19.35 | Owe et al. (2008) |
| WindSat | Feb 2003 - Jul 2012 | 6.8 | 10.7 | 18.7 | Owe et al. (2008) |

## 2.2 Evaluation data

### 2.2.1 MODIS leaf area index

To vertify the plausibility of VODCA we compare it to MODIS leaf area index (LAI), MOD15A2H version 6 (Myneni et al., 2015). LAI is the one-sided ratio of leaf area to ground area and is estimated from the red and NIR MODIS data using a 5 look-up-table based approach with a back-up algorithm that uses empirical relationships between NDVI, LAI and fraction of photosynthetically active radiation (FPAR). The data is available globally since 2002 with an 8-day temporal resolution, and is for comparison purposes spatially downsampled from its native resolution of 500 metres to a quarter degree grid

### 2.2.2 AVHRR vegetation continuous fields

We us the vegetation continuous fields (VCF) version 1 by Hansen and Song (2018); Song et al. (2018) which show the 10 fractional cover of bare ground, short vegetation and tree canopy, where trees are defined as all vegetation taller than 5 metres in height and short vegetation all shorter vegetation. They produced yearly files from 1982 to 2016 indicating the fractional coverage during the local annual peak of growing season. The VCF are derived from data of Advanced Very High Resolution Radiometer (AVHRR) instruments distributed by the Land Long Term Data Record (LTDR) project.

Given the relation of vegetation height (Lefsky et al., 2005) and VOD to biomass, it seems sensible to assume that VOD would 15 be highest / medium / lowest in areas with high tree canopy / short vegetation / bare ground coverage, respectively. We use the VCF data for two purposes. First (sec. 4.1), we calculate the mean VCF from 2002 to 2016 and compare it to the mean of the VODCA products from 2002-2017. Second (sec. 4.4.2), we calculate the VCF trends from 1987 to 2016 and compare it to the merged Ku-band VOD of the same time period to spot differences in trends. Song et al. (2018) also calculated and distributed VCF trends by first determining whether there is a significant trends with a Mann-Kendall test and then calculating the slope 20 with a Theil-Sen estimator. Both are non-parametric tests that are robust to outliers, but using different methods to mask for significance and estimate the slope can lead to significant slopes that are still very small. To alleviate the issue, we too calculate

the slope using a Theil-Sen estimator, but we use the Theil-Sen estimator also to determine a 95% confidence interval for the slope and remove any slopes where the zero-slope is within the confidence interval.

## 3  Methods

The almost exactly the same methodology is used for the product of each product is computed without any influence of the others, with exceptions being described at the end of the respective section. The main difference between the three products is the time period spanned by each product, resulting from the varying availability of input data (Fig. 1). The fusion process involves three main processing steps. First, preprocessing involves masking for spurious observations and spatial and temporal collocation of the data sets. Second, bias between the different sensors is corrected for by scaling them to the VOD climatology of AMSR-E C-, X-, and Ku-band, respectively. Ultimately, the collocated and bias-corrected observations of all data sets are merged in time and space.

### 3.1  Preprocessing

Level 2 VOD data in swath geometry were first projected onto a common regular $0.25° \times 0.25°$ latitude-longitude grid using nearest neighbour resampling. The different sensors visit the same spot on the Earth surface at different times of the day. To facilitate further processing, all observations are temporally resampled to UTC midnight. This is done as in ESA CCI soil moisture (Dorigo et al., 2017) by taking for every 0:00 UTC the closest observation in a window of $\pm12$ hours if any is available.

Basic masking operations were applied to remove potentially spurious observations. Specifically we mask for radio frequency interference (RFI), low land surface temperatures (LST), and VOD values $\leq 0$ as follows:

- RFI: Artificial microwave emitters on the Earth's surface distort the signal received by the satellite, causing the resulting VOD values at those locations to be unreliable. RFI is typically frequency-specific. RFI flags were already provided with the level 2 VOD data and were based on de Nijs et al. (2015). Any observations affected by RFI are removed.

- LST: Due to the different dielectric properties of ice and water reliable retrievals can only be made if the ground is not frozen. Therefore, we remove observations where the LST is below $0°$ C. Masking for LST was based on the temperature retrievals of from Ka-band (Holmes et al., 2009), which is found on all the multi-channel instruments used in VODCA, and were provided with the level 2 VOD data.

- Negative VOD values: VOD retrievals $<0$ are physically impossible and are therefore removed from the data sets. We also remove VOD values of 0.0 (floating point zero). The reasoning is twofold: First, it is physically only possible to get floating point zero VOD if there is virtually no vegetation, making it very unlikely for most parts of the globe. Second, we observed that this case occurs surprisingly often, also in non-desert regions, and that these values never fitted well with the other observations. This indicates that most VOD values of zero are artifacts that have to be removed.

The above masking is applied independently to all sensors and bands. A special case is AMSR2, which has two channels in C-band, i.e. at 6.9 and 7.3 GHz. If possible, the observations from the 6.9 GHz band are used, but if the observation in this channel is masked, the 7.3 GHz observation is used instead (if unmasked) to fill gaps. A flag indicating the channel ultimately used in the merged data set for each observation is provided in the metadata.

## 3.2 Cumulative distribution function (CDF) matching

We use a new implementation of the CDF-matching technique, based on a combination of piece-wise linear interpolation and linear least squares regression. CDF-matching is used to correct for systematic differences between the VOD values of each sensor, which may result, e.g., from the individual sensor designs, incidence angles, spatial footprints and the slight differences

10 in the frequencies used. The goal of CDF-matching is to scale a source data set such that its empirical CDF becomes similar to the empirical CDF of the reference data set. CDF-matching is applied on a per-pixel basis and has been successfully used for similar tasks that require the correction of higher order differences between data sets (Liu et al., 2009, 2011a, 2012; Dorigo et al., 2017).

### 3.2.1 Ordinary piece-wise linear CDF-matching

15 Piece-wise linear CDF-matching (Liu et al., 2009, 2011a; Dorigo et al., 2017) predicts for each [0, 5, 10, 20, 30, ... , 80, 90, 95, 100] percentile of the source data the same percentile of the reference data set. Values between the $n^{th}$ and $n^{th}$+1 percentile are then scaled using linear interpolation. While the scaling parameters are determined only from temporally overlapping observations, during prediction there can be values outside the training range. These values are scaled by extrapolating the first

20 or last percentile interval. This method preserves the ranks of the source and computes rather fast. However, the first and last percentiles are defined by the lowest and highest observations, respectively, in both source and reference time series. Hence, a single outlier can greatly affect the parameters of these percentiles, making them unreliable.

### 3.2.2 Improvements

25 We improved the original method by fitting a linear model using the sorted observations smaller than the second percentile with an intercept through the second percentile. This gives more reliable scaling parameters for low values since all the data between the lowest and second-lowest percentiles are used instead of just the lowest value. In case a different number of observations exists in the source and reference, the data with less observations is padded by linear interpolation during training. In a similar fashion, a model is fitted for observations above the penultimate percentile.

30 We further increase the robustness of the CDF-matching parameters by dynamically increasing the step size of the percentiles if only few observations are available. The number of observations varies greatly from grid point to grid point and from sensor

to sensor. If too few observations exist between two subsequent matching-percentiles (a "bin"), the CDF-matching may overfit, leading to unreliable parameters. To counteract this, we dynamically reduce the number of bins and increase the size of the bins based on the number of observations.

### 3.2.3 Stability of parameters

5 To evaluate whether the new matching technique is more robust to outliers than the original piecewise linear cdf matching method, we simulate the variances of the derived parameters of each bin for a varying number of training observations using artificial values. The use of artificial values allows us to test the method without being influenced by the artifacts inherent to real data. To achieve this, we sample a set of source and reference values from a standard normal distribution, and then determine the resulting CDF-matching parameters. For each evenly spaced percentile bin, we determine the slope in radians. This is
10 repeated a few thousand times for various numbers of values (representing time series with a varying number of observations), each time drawing new values. If a CDF-matching method is robust, the determined slopes should have low variance due to the values always being drawn from the same distribution.

We run this both with piece-wise linear CDF-matching and our new method. However, for this simulation we do not dynam-
15 ically decrease the number of bins, as we are solely interested in the performance of the linear regression scaling the first and last percentile. Both methods are tested with the same randomly drawn data.

The resulting variances in the slope, for each percentile bin, for both methods, depend on the number of observations used for the parameter determination. This is shown in Fig. 2. The results in the middle bins are exactly the same, as the same
20 methodology in used for these bins. However, in the case of linear piece-wise interpolation, the slope parameters of the first and last bins have a much higher variance than the middle bins as they are affected by outliers. In contrast, the slopes determined by the least-squares method have a much lower variance. In both cases we can also see that the more observations we have, the lower the variance of the slope parameter is, showcasing the reasoning behind reducing the number of bins dynamically if too few observations are available.

### 25  3.2.4 Practical implementation and exceptions

While there is no "true" reference to scale to, AMSR-E has almost global coverage and has a long temporal overlap with all other sensors but AMSR2. Hence, the empirical CDFs of WindSat, TMI and SSM/I are directly scaled to the one of AMSR-E. To preserve any potential trends in both source and reference data, only dates when both have a valid observation are used. If at a certain location less than 20 temporal overlapping observations exist, no reliable scaling parameters can be determined and
30 the source time series is dropped.

AMSR2 does not share any temporal overlap with AMSR-E and therefore cannot be directly scaled based on overlapping observations. Instead, for X- and Ku-band, scaled observations of TMI can potentially be used to bridge this gap. This is
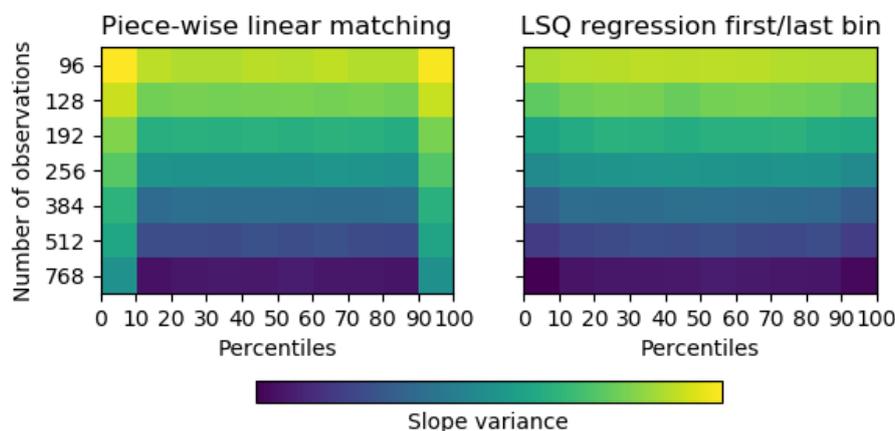
**Figure 2.** Variance of the derived slope, depending on the number of observations and the percentile bin for both piece-wise linear CDF-matching. The color is log normalized.

done according to the following logic (Fig. 3): If possible, AMSR2 is scaled to the rescaled TMI. Should there not be enough overlapping observations, the scaling parameters are determined from all observations of the first two years of AMSR2 and the last two years of AMSR-E. While this removes any potential trends in the first two years of the AMSR2 period, these trends are still assumed to be smaller than the removed bias. Last, if there are also not enough AMSR2 or AMSR-E obser-
5   vations available in those years, the whole AMSR2 time series is dropped. For C-band, which is not covered by TMI, the AMSR2 data are always matched directly to AMSR-E following the approach above. The published data sets contain a flag indicating the matching method, allowing the user to remove the AMSR2 observations matched directly to AMSR-E if desired.

Since the scaling parameters are determined using only a subset of all observations, during prediction there can be values
10  outside the training range. The regression is not forced to go through the origin, therefore if the predicted values can potentially be smaller than 0. These values are deemed unreliable and removed.

### 3.3   Merging

For all bands, the CDF-matched time series of all individual sensors are merged into a single long continuous time series. For
15  a certain pixel at a certain time step, three possible scenarios can occur:

1. If on a certain date no sensor has an observation, a data gap will result in the final product;

2. If only one sensor has an observation, the CDF-matched value will be directly integrated in the final product;

3. If multiple sensors have an observation on a certain date, their arithmetic mean is taken.

```
for each grid point:
    if at least 20 overlapping TMI observations available
        match AMSR2 to TMI using overlapping observations
    else if at least 20 AMSR-E and 20 AMSR2 observations are available:
        match first two years of AMSR2 to last two years of AMSR-E
    else:
        drop AMSR2 time series
```

**Figure 3.** Pseudo-code showing the CDF-matching logic of AMSR2

This means that the number of sensors contributing to each observation within a time series can vary greatly. For each observation in the final product there is a flag indicating which sensors have contributed to it. Although more sophisticated weighted merging methods based on least squares have been proposed to merge multiple satellite observations (Gruber et al., 2017, 2019), estimating these weights, i.e. indicators of the relative quality of the individual data sets, is a non-trivial task. This particularly applies to VOD, for which no appropriate independent reference data exist. However, in most cases, the arithmetic mean appears to be a robust approximation of optimal merging (Liu et al., 2012).

## 4 Properties of the long-term VOD data sets

### 4.1 Spatial patterns and temporal dynamics

Figure 4 shows an example of X-band VOD time series in Austria at different stages of merging procedure together with MODIS LAI. The original VOD time series have visible systematic differences between each sensor. The CDF-matched VOD time series have been scaled to AMSR-E and visually do not show systematic differences between sensors. The statistical distributions of VOD from the sensors are similar after matching (Fig. 4 b). This example grid point is north of 38N and thus outside the spatial coverage of TMI, therefore AMSR2 has been scaled to AMSR-E directly using non-temporally overlapping observations. The merged VOD time series shows comparable seasonal dynamics like LAI.

The global spatial patterns of average VOD between June 2002 and June 2017 is shown for each band in Figure 5 (a-c). This period was selected because all bands have global coverage in this time period. All bands show similar spatial patterns, matching the ones of the VCF land covers (Fig. 5 (e), with high VOD in tropical and northern forests and lower VOD in grassland and desert regions. The same pattern is also visible in canopy height (Simard et al., 2011) and MODIS LAI (Fig. 5 (d), even though the LAI in the tropical forests is much higher than in the boreal forests, while VOD is similarly high in both regions. Based on the principle that the penetration of microwaves increases with wavelength, the maximum VOD is highest at shorter
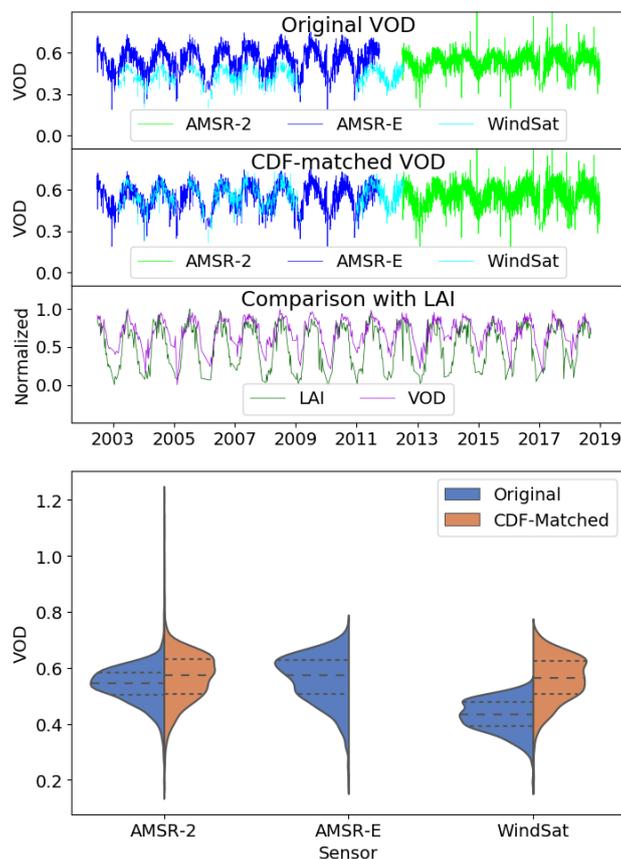
**Figure 4.** Example X-band time series (lon / lat: 15.125 / 48.125) at different processing steps (a) and violin plot showing the effect of CDF-matching on its distribution (b). Time series of the original VOD data of all available sensors for that band at that location are shown in the top panel, the CDF-matched series in the middle panel, and the final merged VOD (VODCA) is shown together with MODIS LAI in the bottom panel. In the bottom panel VOD and LAI are both normalized, VOD is downsampled by moving average to match the temporal 8-day resolution of LAI.

wavelengths (Ku-band) and smallest at longer wavelengths (C-band). This can also be seen in Figure 5 (f) which shows the average VOD of each band for locations dominated by TC, SV or BG, respectively. This figure also shows that as expected, on average VOD is highest in forests and lowest over bare ground.

5    The temporal dynamics of VOD across different latitudes shows plausible seasonal patterns of vegetation (Fig. 6). In general, summer months have the highest VOD: in the Tropics and Subtropics due to increased precipitation during that time, while in northern/southern regions due to the increased temperature and consequent vegetation growth and (leaf) biomass gain. The VOD patterns strongly correlate with LAI, quantified by a Spearman coefficient of 0.67, 0.66 and 0.58 between LAI and C-,

(a)  C-band

(b)  X-band

(c)  Ku-band

(d)  MODIS LAI

(e)  Tree cover (TC), short vegetation (SV) and bare ground (BG) coverage

(f)  Mean VOD of each band depending on VCF class
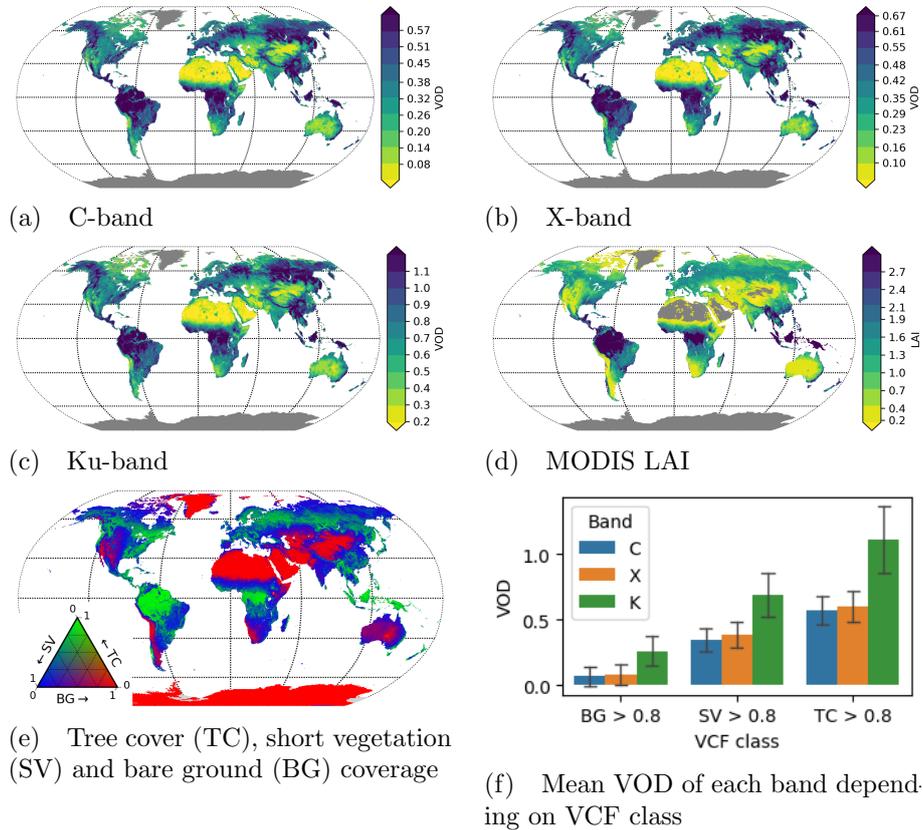
**Figure 5.** Global spatial patterns of average multi-sensor VOD from each band (2002 to 2017), average MODIS LAI (2002 to 2017) and average VCF (2002 - 2016 and distribution of VOD for locations with TC, SV and BG greater than 0.8. The error bars indicate the standard deviation within each group.

X- and Ku-band respectively

The VOD time series do not show any visible artificial breaks, indicating that the biases have overall been successfully removed from the individual sensors before merging. To make potential artificial breaks more visible, we investigated the
5  seasonal anomalies per latitude (Fig. 7). The anomalies, using the period from 2002 to 2017 as reference, are calculated by collecting all the observations of a latitude, calculating the monthly mean, subtracting the multi-year monthly average and removing any potential linear trends using ordinary least squares regression. Hence the anomalies should either represent natural variability or artifacts due to shifts in available sensors. In the latter case, one would expect global anomalies to be visible either due to bias or differing spatial extent.
10  Most anomalies are limited in both space and time and their start or end does not coincide with a change in sensors, indicating
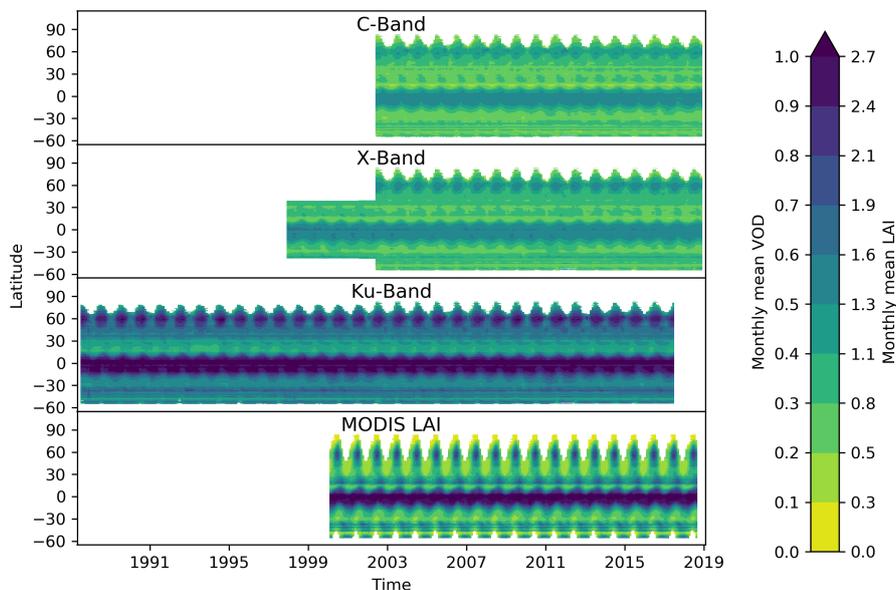
**Figure 6.** Hovmoeller diagrams showing the monthly mean VOD per latitude for each Band of VODCA and for LAI

that they are due to natural causes. MODIS LAI shows similar anomalies as the VOD products, the Spearman correlation coefficient is 0.29, 0.29 and 0.26 between LAI and C-, X- and Ku-band anomalies respectively.
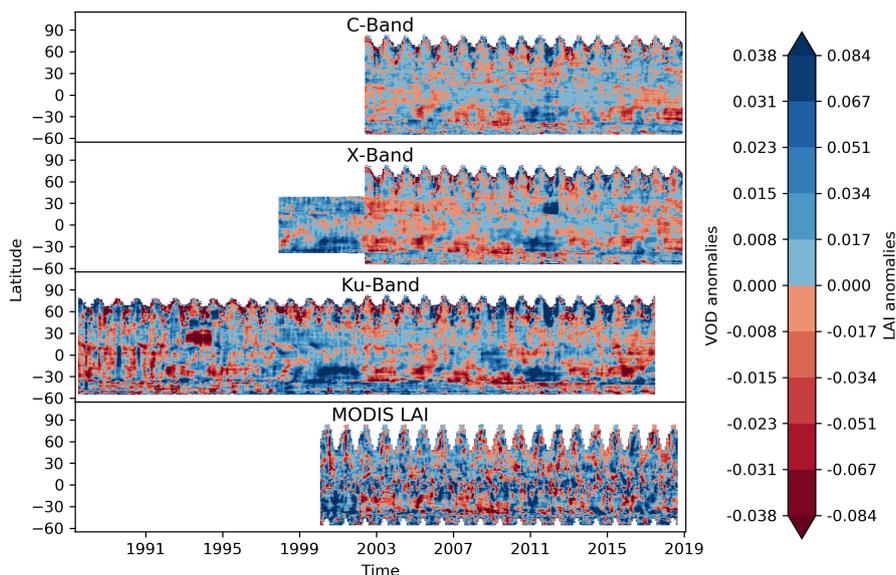
**Figure 7.** Hovmoeller diagrams showing anomalies of the monthly means per latitude for each band of VODCA and for LAI

## 4.2 Spatio-temporal coverage

The temporal and spatial coverage of the merged VOD time series for each band is shown in Figure 8. The coverage of the merged products is defined by the spatial and temporal coverage of sensors (Fig. 1). For any band in any time span with at least one sensor, most parts of the globe have for at least 40% of all days an observation, while in any time period with at least two

5    sensors about 70% of all days have a valid observation. TMI is the only sensor with a non-polar orbit of 35°N/S, leading to an increased coverage in that region in Ku- and X-band from 1997-2015. The latitude affects the coverage in multiple ways: Northern regions are generally more often covered by the polar-orbiting satellites but on the other hand, frozen grounds and snow cover inhibit the retrieval of VOD in Winter. The low coverage band near 23°N is the result of LPRM not converging on a valid solution in very arid regions due to the extreme soil dielectric constants in these regions (de Jeu et al., 2014).
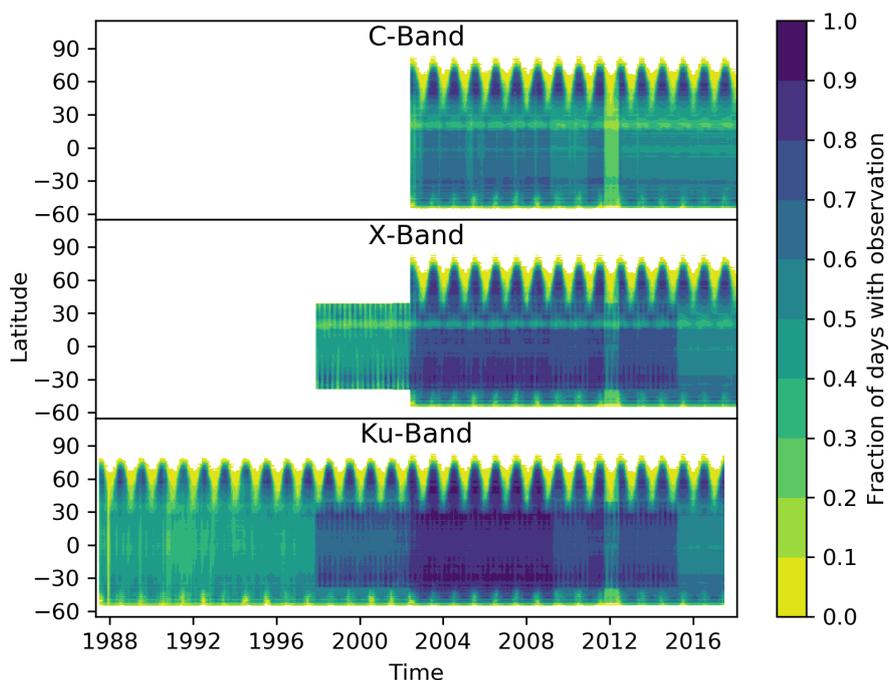
**Figure 8.** Hovmoeller diagrams showing the fraction of in days each month per latitude that have an observation. For this figure, the number of observations for each latitude and month are counted and then divided by the number of land grid points and days per month.

In some locations the merged VOD products have fewer observations than in the original products. This data loss can be caused by a failure of the merging procedure. The matching can fail due to insufficient AMSR-E data and hence the data loss occurs in similar regions for all sensors of one band. The lack of AMSR-E data is in most cases due to either RFI or low temperatures in mountainous regions. As an example Figure 9 shows, for all bands, where the CDF-matching failed for WindSat

15    data. Ku-band is the least affected (Fig. 9 (c)), where only about 2% of the grid points are lost, mostly in the Himalayas. In

X-band the matching fails for about 5% of the grid points, mostly in large parts of the Sahara (Fig. 9 (b)). C-band is most affected by data loss (10%), mostly in some parts of the USA where additional RFI prevents accurate retrievals (Fig. 9 (a)) (Njoku et al., 2005).
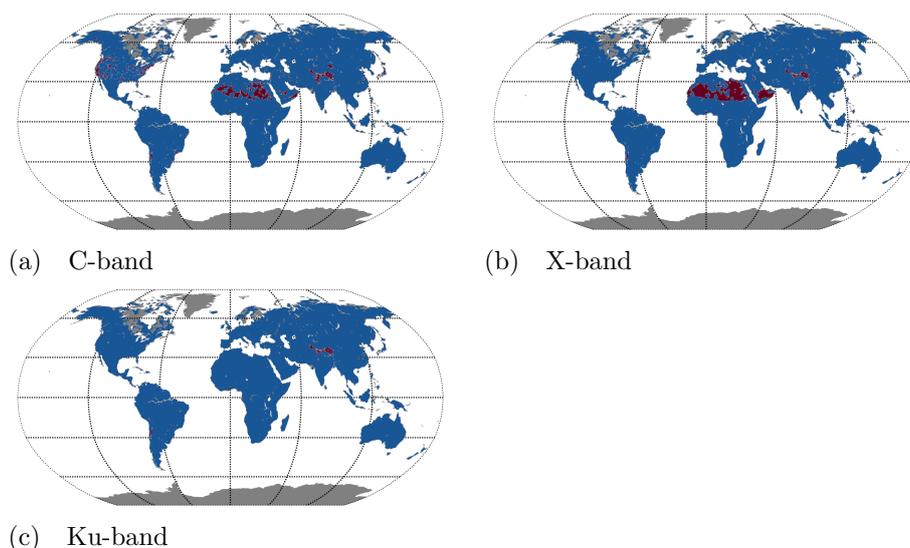
**Figure 9.** Data loss of during CDF-matching of different WindSat bands. CDF-matching failed for the red grid points and therefore the data of WindSat at that location is dropped. Very similar looking maps exist for the other sensors in the supplement Fig. 4

### 4.3   Random error characteristics

5   To validate the performance of our merging approach we evaluate the change in autocorrelation as an indicator for precision. Merging overlapping observations from multiple sensors is supposed to result in data that has a higher precision than the data of any of the individual sensors. But without a higher-quality external reference data set, assessing the change in precision is non-trivial. However, we can assume that there is supposed to be a high degree of temporal autocorrelation between subsequent observations because VOD is related to gradual changes in plant water content and biomass (Momen et al., 2017; Konings et al.,

10   2016). Therefore we calculated the difference between the first-order temporal autocorrelation before and after merging. The autocorrelation coefficient is strongly dependent on the temporal resolution. As seen in sec. 4.2, the temporal resolution of VODCA increases if multiple sensors are available. Therefore directly comparing the autocorrelation coefficients between the individual sensors and the merged products would lead to an increase in autocorrelation that is related to the temporal resolution rather than to the precision. Therefore the temporal resolution is kept unchanged by using only observation dates

15   existing both in the pre-merge and post-merge data set.

The autocorrelation differences for X-band are shown in Figure 10. The other bands show similar results and are available in the supplement Figs. 1-3. The autocorrelation of the merged time series is on average higher than the autocorrelation of the

input series, indicating an overall decrease in noise. However, sometimes the gain in autocorrelation of one sensor mirrors the loss of the autocorrelation of the other, likely due to the former sensor being more noisy than the latter, e.g. in Alaska or east Russia in X-band of AMSR-E vs. WindSat. This means that locally, sometimes a single sensor has a higher precision than VODCA. But there are also regions where the merged VOD autocorrelation is higher than any of the input time series, e.g in

5  Europe or central north America. This is likely to occur when all sensors have a similar precision, meaning that none of them is dragging the precision of the others down.

A noteworthy case is TMI where the autocorrelation of the merged time series is almost always higher. This could mean that the TMI data is very noisy and is dragging the overall quality of the merged data down. We investigated this possibility by experimentally not including TMI in VODCA. This resulted in average in a lower gain in autocorrelation for the other data

10  sets, indicating that the TMI data is still positively contribute to the precision of the merged products by reducing the noise of the end product (result not shown).
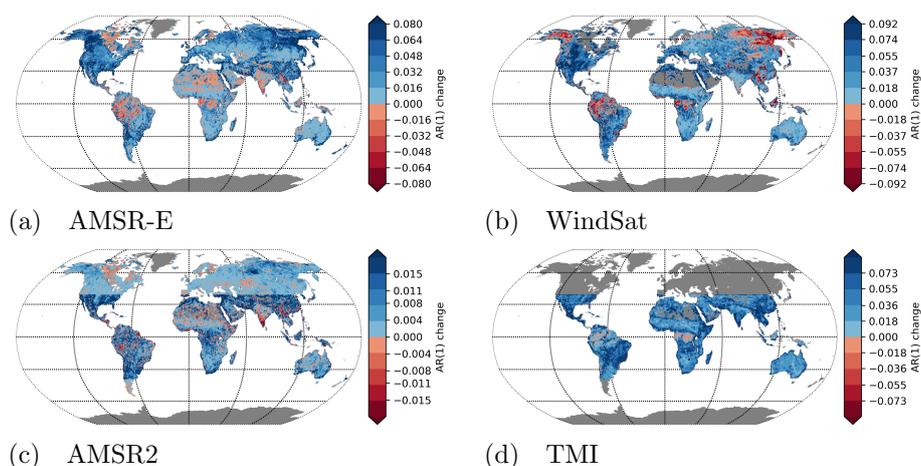
**Figure 10.** First-order auto-correlation change due to merging of X-band data for each sensor.

## 4.4 Comparison of VOD with LAI and Vegetation Continuous Fields

### 4.4.1 Correlation between VOD and LAI

A direct validation of VODCA is not possible because of the lack of appropriate in situ measurements. Hence it is only possible

15  to assess dynamics in VOD with dynamics in related variables such as LAI or land cover. Globally, LAI and VOD time series and their seasonal anomalies are positively correlated over large areas (Fig. 11). For all bands, the highest correlations with LAI can be found in grassland-dominated regions such as in African Savannahs, Australia and in parts of South America. Correlations are usually lower in forested regions and even slightly negative in parts of tropical forests such as in the Amazon. The negative correlations in tropical forests could be caused by drought periods where vegetation water content and hence

20  VOD should decline but LAI possibly increases (Myneni et al., 2007; Saleska et al., 2007), although a green-up of the Amazon

under drought is highly debated (Samanta et al., 2010, 2012; Morton et al., 2014). However, this comparison of VOD and LAI demonstrates that the merged VOD products reflect plausible seasonal and short-term changes in vegetation and will likely provide additional information on vegetation dynamics on top of LAI and other related optical biophysical vegetation products from optical remote sensing.
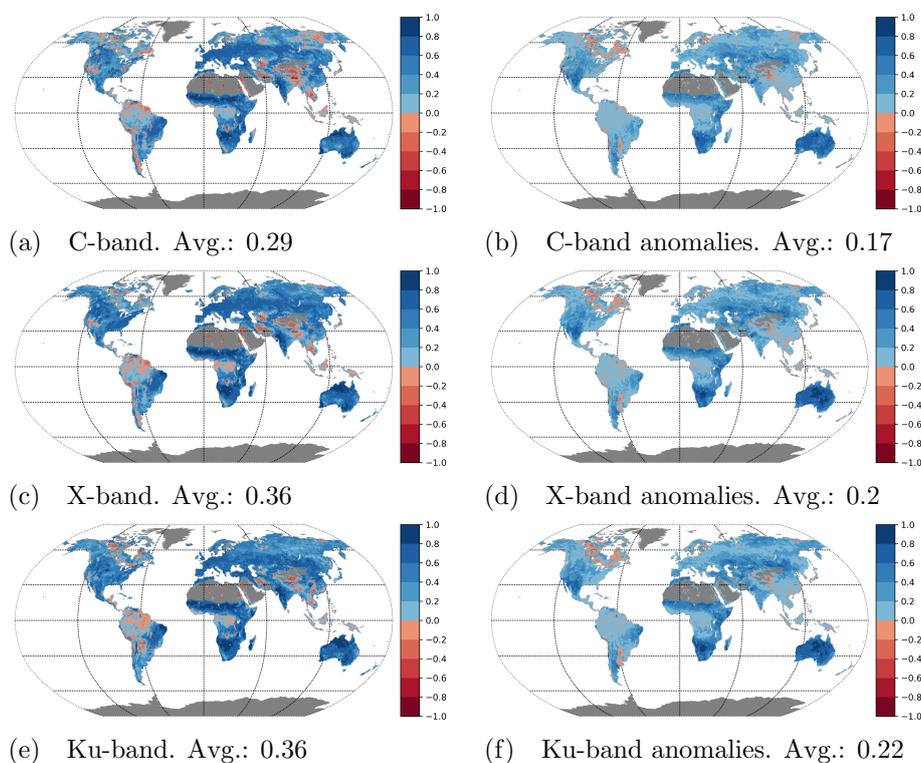
**Figure 11.** Spearman correlation coefficient between merged VOD and MODIS LAI for each band. In the left column are the coefficients between the raw values, in the right between the anomalies.

### 4.4.2 Trend—analysis of VOD, LAI and Vegetation Continuous Fields

To evaluate the relationship between C-, X-, Ku-band VOD, MODIS LAI and VCF changes and to gain a first insight into the long-term changes in VOD we assess linear trends in the data sets. Yearly averages are used to determine the trends and their confidence intervals via the Theil–Sen estimator. Trends whose upper and lower confidence interval do not have the same sign or either of them is zero are regarded as non-significant and are not displayed. Fig. 11 (a-c) show the C-, X- and Ku-band trends from 2002-06-19 to 2017-06-19 during which all bands have global coverage. The trends are visually very similar in all bands, confirmed by the spatial Spearman correlation coefficients of 0.88 between the C- and X-band trends, 0.89 between C- and Ku-band and 0.91 between X- and Ku-band, calculated using only locations where both bands have a significant trend. This further reinforces that all bands react very similarly to vegetation changes. The spatial overlap of trends is shown in Fig. 11

(d), where each location is classified based on the sum of positive and the sum of negative trends. Locations with no significant trend in any band are not displayed. The three classes with contradicting trends (1|1, 2|1, 1|2) are rare as together they make up only $4.2\%$ of the displayed points. Conversely, $48\%$ of the land points are covered by the four classes with at least two agreeing trend directions without any contradicting trend (2|0, 3|0, 0|2, 0|3). The agreement in trends between frequencies indicates that

5 the longer Ku-band series can be used as indicator of the shorter X- and C-band series in trend analyses. Further, the LAI trends of the same time period (Fig. 12 (e)) match the VOD trends very well overall, even though in detail the strength and location of the trends vary.

Taking advantage of the much longer length of the Ku-band, another trend analysis is done for this band using the data from 1987 - 2017 (Fig. 12 (g)) to give a first impression of the changes within the last thirty years. Overall we see a decline in

10 VOD in the tropics, likely due to deforestation, and in large parts of Mongolia, attributed to variations in rainfall and surface temperatures as well as increased life stock farming and wild fires (Liu et al., 2013). VOD increased strongly in India and large parts of China, mostly due to an increase in croplands in the former case and due to both an increase in forest and croplands in the latter (Chen et al., 2019). VOD also increased in northern parts of Australia, matching trends in FPAR and precipitation seen in Donohue et al. (2009). Other regions with increasing VOD are south Africa and central north America.

15 Of a questionable nature is the wide spread positive trend in the Sahara given LPRMs struggle to retrieve VOD here. Most of the changes observed for VOD are mirrored in the VCF changes from 1987 to 2016 (Fig. 12 (f), see sec. 2.2.2 for details). The large bare ground losses in India, China and the north African shrubland manifest as positive VOD trends. Likewise, the deforestation in south America and land degradation with hotspots in Mongolia, Afghanistan or southwestern USA coincide with a loss in VOD. Also the patterns of tree cover gain in eastern Europe and European Russia coincide with increased VOD.

20 While there do not seem to be any areas where VOD and VCF contradict each other clearly, some trends are only visible in one of the data sets. For example the strong increase in VOD in southern Africa cannot be observed in VCF.

## 5  Current limitations and possible improvements

### 5.1  AMSR2 scaling to TMI

Upon closer inspection of the trends in Figure 12, we can see in north America a spatial break in X- and Ku-band trends at

25 35°N. AMSR2 data of 2012-2014 have been matched to the AMSR-E data of 2010-2012 north of this latitude and in the south to temporally overlapping observations of scaled TMI. Unusual low VOD values can be observed in time series of that region in the years 2012 to 2015 in X-and Ku-band. This indicates that the CDF-matching does not correct the bias between the sensors but artificially removes the difference that are due to surface processes. Consequently, the matched AMSR2 data has a slight positive bias north of 35°N in large parts of north America. For users we advise to be careful when using X- and Ku-band

30 observations after July 2012 north/south of 35$^o$N/S as well as C-band observations after July 2012 globally as the AMSR2 data might induce a bias.
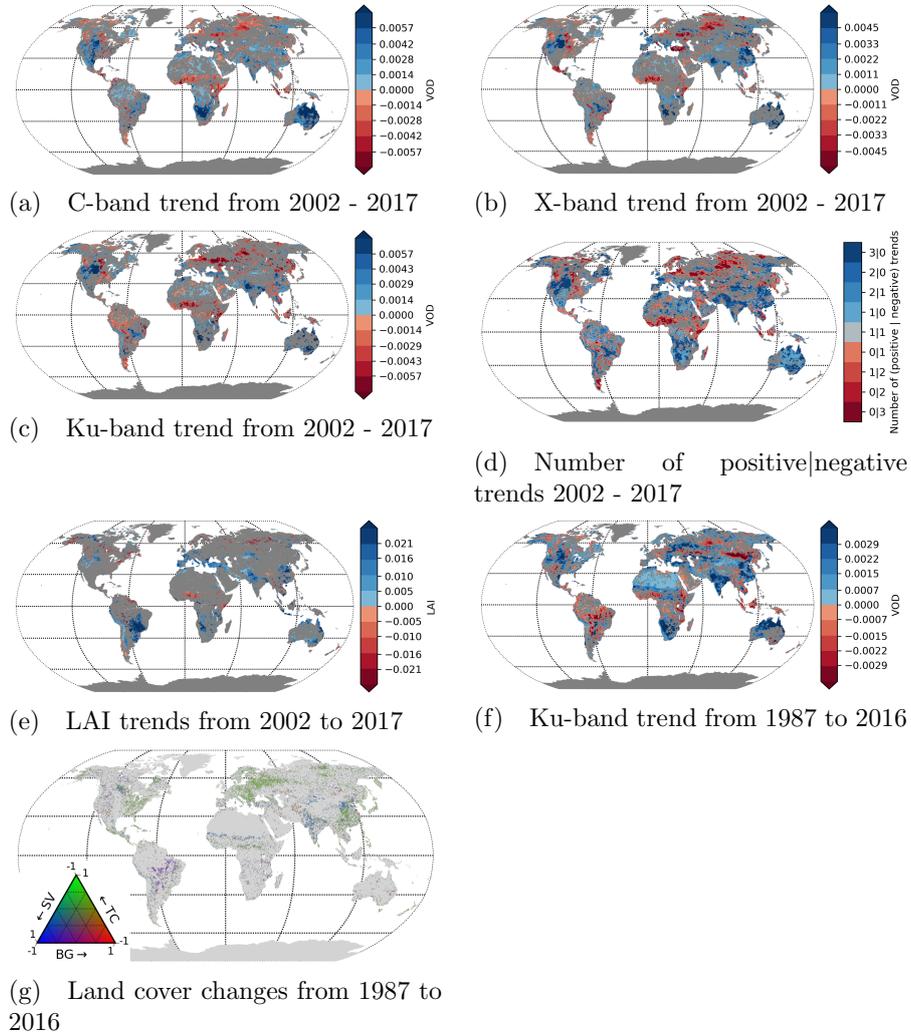
**Figure 12.** Trend-analyses for various bands and time spans of VOD (a-c, f), LAI (e) and tree canopy, short vegetation and bare ground (g). Non-significant trends are not displayed, the trends are calculated by linear regression using yearly mean values. Figure (d) shows trend classes based on the number of VOD bands showing a positive|negative trend. Their order and color are indicative of the likelihood of the trend.

## 5.2 Data loss while CDF-matching

As described earlier, CDF matching failed because of missing AMSR-E data in some regions, mostly in the Himalayas (Fig. 9). One possible solution to avoid this data loss would be to substitute the CDF-matching parameters of these locations with the parameters from locations with similar dynamics in VOD. This could be done by clustering the time series and using the parameters of another location within the same cluster. Taking this one step further, one could also investigate the possibility

of using all the data in one cluster to derive a single set of CDF-matching parameters and use these to scale all the source time series within it. Not only would this allow to scale all the data without loss, but the increased number of observations available for each parameter determination would also lead to more robust CDF-parameters. However, generating meaningful clusters from hundreds of thousands long time series containing missing observations while keeping the computational cost at bay is

5    anything but trivial (e.g. Mikalsen et al. (2018)).

## 5.3 Data gaps in the input data sets leading to increased noise

Averaging multiple temporally overlapping observations reduces noise (sec. 4.3). However, this can be only done if overlapping observations exist. While the maximum number of observations is limited to the number of available sensors, usually less

10    observations are available due to gaps in the individual time series of each sensor. Hence a potential short-time gap-filling of the original time series of each sensor could increase the precision of VODCA.

Since VOD changes slowly over time (Konings et al., 2016), it is intuitively clear that even if a sensor has no valid observation on a certain date, the value is expected to be similar to the value of the dates before and after. Therefore one could fill short gaps with a model that at least implicit uses autocorrelation for its predictions, such as gaussian processes as in Camps-valls

15    et al. (2017).

## 6    Conclusions

We present to the scientific community VODCA, three long term VOD data sets spanning the last three decades to be used in studies of the biosphere. For the most part we were able to remove the biases between the different input sensors by co-calibrating them to AMSR-E. The merging leads to observations with less noise than the input data sets. The trends of the

20    different VOD bands (C-, X-, Ku-band) correlate very strongly with each other and show similar spatial distributions and temporal dynamics as LAI and VCF, with the added benefit of having observations on a daily basis unaffected by cloud cover. A major ongoing issue is the potential bias in AMSR2 due to no temporally overlapping observations with other sensors. This and other problems still have to be resolved and as such we plan to maintain and improve VODCA with up-to-date data and continued development for the foreseeable future.

25    *Data availability.* The VODCA products (Moesinger et al., 2019) are open access (Attribution 4.0 International) and available at Zenodo https://doi.org/10.5281/zenodo.2575599

*Author contributions.* Wouter Dorigo, Leander Moesinger, and Matthias Forkel designed the study. Leander Moesinger performed the analyses and wrote the manuscript together with Matthias Forkel and Wouter Dorigo. All authors contributed to discussions about the methods and results and provided feedback on the manuscript.

*Competing interests.* The authors declare that they have no conflict of interest

5   *Acknowledgements.* The MODIS LAI data (MOD15A2H, v006) by Myneni et al. (2015) were retrieved from the online data pool, courtesy of the NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://lpdaac.usgs.gov/data_access/data_pool .

The VCF annual data (VCF5KYR, v001) by Hansen and Song (2018) were retrieved from the online NASA Earthdata Search, courtesy
10  of the NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://search.earthdata.nasa.gov .

The authors acknowledge the TU Wien University Library for financial support through its Open Access Funding Program.

## References

Brandt, M., Yue, Y., Wigneron, J. P., Tong, X., Tian, F., Jepsen, M. R., Xiao, X., Verger, A., Mialon, A., Al-Yaari, A., Wang, K., and Fensholt, R.: Satellite-observed Major Greening and Biomass Increase in South China Karst During Recent Decade, Earth's Future, https://doi.org/10.1029/2018EF000890, http://doi.wiley.com/10.1029/2018EF000890, 2018.

5 Camps-valls, G., Svendsen, D. H., Martino, L., and Campos-t, M.: Physics-Aware Gaussian Processes for Earth Observation, 10270, https://doi.org/10.1007/978-3-319-59129-2, http://link.springer.com/10.1007/978-3-319-59129-2, 2017.

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 Sensing of Environment, 212, 249–259, https://doi.org/10.1016/J.RSE.2018.04.049, https://www.sciencedirect.com/science/article/pii/S0034425718302062{#}bb0240, 2018.

10 Chen, C., Park, T., Wang, X., Piao, S., Xu, B., Chaturvedi, R. K., Fuchs, R., Brovkin, V., Ciais, P., Fensholt, R., Tømmervik, H., Bala, G., Zhu, Z., Nemani, R. R., and Myneni, R. B.: China and India lead in greening of the world through land-use management, Nature Sustainability, 2, 122–129, https://doi.org/10.1038/s41893-019-0220-7, http://www.nature.com/articles/s41893-019-0220-7, 2019.

de Jeu, R. A., Holmes, T. R., Parinussa, R. M., and Owe, M.: A spatially coherent global soil moisture product with improved temporal resolution, Journal of Hydrology, 516, 284–296, https://doi.org/10.1016/J.JHYDROL.2014.02.015, https://www.sciencedirect.com/science/article/pii/S0022169414001139, 2014.

15 de Nijs, A. H. A., Parinussa, R. M., de Jeu, R. A. M., Schellekens, J., and Holmes, T. R. H.: A Methodology to Determine Radio-Frequency Interference in AMSR2 Observations, IEEE Transactions on Geoscience and Remote Sensing, 53, 5148–5159, https://doi.org/10.1109/TGRS.2015.2417653, http://ieeexplore.ieee.org/document/7116527/, 2015.

Donohue, R. J., Mcvicar, T. R., and Roderick, M. L.: Climate-related trends in Australian vegetation cover as inferred from satellite observations, 1981-2006, Global Change Biology, 15, 1025–1039, https://doi.org/10.1111/j.1365-2486.2008.01746.x, http://doi.wiley.com/10.1111/j.1365-2486.2008.01746.x, 2009.

20 Dorigo, W., De Jeu, R., Chung, D., Parinussa, R., Liu, Y., Wagner, W., and Fernández-Prieto, D.: Evaluating global trends (1988-2010) in harmonized multi-satellite surface soil moisture, Geophysical Research Letters, 39, 3–9, https://doi.org/10.1029/2012GL052988, 2012.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.: ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions, Remote Sensing of Environment, 203, 185–215, https://doi.org/10.1016/j.rse.2017.07.001, http://linkinghub.elsevier.com/retrieve/pii/S0034425717303061, 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.: Evaluation of microwave remote sensing for monitoring live fuel moisture content in the Mediterranean region, Remote Sensing of Environment, 205, 210–223, https://doi.org/10.1016/J.RSE.2017.11.020, https://www.sciencedirect.com/science/article/pii/S0034425717305692, 2018.

Gaiser, P. W., St. Germain, K. M., Twarog, E. M., Poe, G. A., Purdy, W., Richardson, D., Grossman, W., Jones, W. L., Spencer, D., Golba, G., Cleveland, J., Choy, L., Bevilacqua, R. M., and Chang, P. S.: The windSat spaceborne polarimetric microwave radiometer: Sensor description and early orbit performance, IEEE Transactions on Geoscience and Remote Sensing, 42, 2347–2361, https://doi.org/10.1109/TGRS.2004.836867, 2004.

Grant, J. P., Wigneron, J.-P., De Jeu, R. A. M., Lawrence, H., Mialon, A., Richaume, P., Al Bitar, A., Drusch, M., Van Marle, M. J. E., and Kerr, Y.: Comparison of SMOS and AMSR-E vegetation optical depth to four MODIS-based vegetation indices, Remote Sensing of Environment, 172, 87–100, https://doi.org/10.1016/j.rse.2015.10.021, https://ac.els-cdn.com/S0034425715301735/1-s2.0-S0034425715301735-main.pdf?{_}tid=fd05dbf8-2a49-438c-8f7f-dce794bb90a6{&}acdnat=
5  1524568420{_}3f2adbe057045116fcd23d7de2f827ab, 2016.

Gruber, A., Dorigo, W. A., Crow, W., and Wagner, W.: Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals, IEEE Transactions on Geoscience and Remote Sensing, 55, 6780–6792, https://doi.org/10.1109/TGRS.2017.2734070, 2017.

Gruber, A., Scanlon, T., Van der Schalie, R., Wagner, W., and Dorigo, W. A.: Evolution of the ESA CCI Soil Moisture Climate Data Records and their underlying merging methodology (Manuscript in Preparation), 2019.

10  Hansen, M. and Song, X.: Vegetation Continuous Fields (VCF) Yearly Global 0.05 Deg [Data set], https://doi.org/10.5067/MEaSUREs/VCF/VCF5KYR.001, 2018.

Holmes, T. R. H., De Jeu, R. A. M., Owe, M., and Dolman, A. J.: Land surface temperature from Ka band (37 GHz) passive microwave observations, Journal of Geophysical Research, 114, D04 113, https://doi.org/10.1029/2008JD010257, http://doi.wiley.com/10.1029/2008JD010257, 2009.

15  Jackson, T. and Schmugge, T.: Vegetation effects on the microwave emission of soils, Remote Sensing of Environment, 36, 203–212, https://doi.org/10.1016/0034-4257(91)90057-D, https://www.sciencedirect.com/science/article/pii/003442579190057D, 1991.

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 Sensing of Environment, 115, 1102–1114, https://doi.org/10.1016/J.RSE.2010.12.015, https://www.sciencedirect.com/science/article/pii/S0034425710003615, 2011.

20  Kawanishi, T., Sezai, T., Ito, Y., Imaoka, K., Takeshima, T., Ishido, Y., Shibata, A., Miura, M., Inahata, H., and Spencer, R.: The advanced microwave scanning radiometer for the earth observing system (AMSR-E), NASDA's contribution to the EOS for global energy and water cycle studies, IEEE Transactions on Geoscience and Remote Sensing, 41, 184–194, https://doi.org/10.1109/TGRS.2002.808331, http://ieeexplore.ieee.org/document/1196037/, 2003.

Knowles, K., Savoie, M., Armstrong, R., and Brodzik, M. J.: AMSR-E/Aqua Daily EASE-Grid Brightness Temperatures,
25  https://doi.org/https://doi.org/10.5067/XIMNXRTQVMOX, 2006.

Konings, A. G., Piles, M., Rötzer, K., McColl, K. A., Chan, S. K., and Entekhabi, D.: Vegetation optical depth and scattering albedo retrieval using time series of dual-polarized L-band radiometer observations, Remote Sensing of Environment, 172, 178–189, https://doi.org/10.1016/J.RSE.2015.11.009, https://www.sciencedirect.com/science/article/pii/S003442571530198X, 2016.

Kummerow, C., Barnes, W., Kozu, T., Shiue, J., Simpson, J., Kummerow, C., Barnes, W., Kozu, T., Shiue, J., and Simp-
30  son, J.: The Tropical Rainfall Measuring Mission (TRMM) Sensor Package, Journal of Atmospheric and Oceanic Technology, 15, 809–817, https://doi.org/10.1175/1520-0426(1998)015<0809:TTRMMT>2.0.CO;2, http://journals.ametsoc.org/doi/abs/10.1175/1520-0426{%}281998{%}29015{%}3C0809{%}3ATTRMMT{%}3E2.0.CO{%}3B2, 1998.

Lefsky, M. A., Harding, D. J., Keller, M., Cohen, W. B., Carabajal, C. C., Del Bom Espirito-Santo, F., Hunter, M. O., and de Oliveira, R.: Estimates of forest canopy height and aboveground biomass using ICESat, Geophysical Research Letters, 32, n/a–n/a,
35  https://doi.org/10.1029/2005GL023971, http://doi.wiley.com/10.1029/2005GL023971, 2005.

Liu, Y., Dorigo, W., Parinussa, R., de Jeu, R., Wagner, W., McCabe, M., Evans, J., and van Dijk, A.: Trend-preserving blending of passive and active microwave soil moisture retrievals, Remote Sensing of Environment, 123, 280–297, https://doi.org/10.1016/J.RSE.2012.03.014, https://www.sciencedirect.com/science/article/pii/S0034425712001332, 2012.

Liu, Y. Y., van Dijk, A. I. J. M., de Jeu, R. A. M., and Holmes, T. R. H.: An analysis of spatiotemporal variations of soil and vegetation moisture from a 29-year satellite-derived data set over mainland Australia, Water Resources Research, 45, https://doi.org/10.1029/2008WR007187, http://doi.wiley.com/10.1029/2008WR007187, 2009.

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
5    retrievals of vegetation optical depth, Geophysical Research Letters, 38, 1–6, https://doi.org/10.1029/2011GL048684, 2011a.

Liu, Y. Y., Parinussa, R. M., Dorigo, W. A., De Jeu, R. A. M., Wagner, W., Van Dijk, A. I. J. M., Mccabe, M. F., and Evans, J. P.: Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, Hydrol. Earth Syst. Sci, 15, 425–436, https://doi.org/10.5194/hess-15-425-2011, www.hydrol-earth-syst-sci.net/15/425/2011/, 2011b.

Liu, Y. Y., Evans, J. P., McCabe, M. F., de Jeu, R. A. M., van Dijk, A. I. J. M., Dolman, A. J., and Saizen, I.: Changing Climate and
10    Overgrazing Are Decimating Mongolian Steppes, PLoS ONE, 8, e57599, https://doi.org/10.1371/journal.pone.0057599, https://dx.plos.org/10.1371/journal.pone.0057599, 2013.

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

Liu, Y. Y., van Dijk, A. I., Miralles, D. G., McCabe, M. F., Evans, J. P., de Jeu, R. A., Gentine, P., Huete, A., Parinussa, R. M., Wang,
15    L., Guan, K., Berry, J., and Restrepo-Coupe, N.: Enhanced canopy growth precedes senescence in 2005 and 2010 Amazonian droughts, Remote Sensing of Environment, 211, 26–37, https://doi.org/10.1016/J.RSE.2018.03.035, https://www.sciencedirect.com/science/article/pii/S0034425718301366, 2018.

Meesters, A., DeJeu, R., and Owe, M.: Analytical Derivation of the Vegetation Optical Depth From the Microwave Polarization Difference Index, IEEE Geoscience and Remote Sensing Letters, 2, 121–123, https://doi.org/10.1109/LGRS.2005.843983, http://ieeexplore.ieee.org/
20    document/1420287/, 2005.

Meier, W., JesefinoC, C., and Thorsten, M.: NRT AMSR2 Unified L3 Daily 25 km Brightness Temperature & Sea Ice Concentration Polar Grids V1, https://doi.org/undefined, http://undefined, 2018.

Mikalsen, K. Ø., Bianchi, F. M., Soguero-Ruiz, C., and Jenssen, R.: Time series cluster kernel for learning similarities between multivariate time series with missing data, Pattern Recognition, 76, 569–581, https://doi.org/10.1016/J.PATCOG.2017.11.030, https://www.
25    sciencedirect.com/science/article/pii/S0031320317304843, 2018.

Mo, T., Choudhury, B. J., Schmugge, T. J., Wang, J. R., and Jackson, T. J.: A model for microwave emission from vegetation-covered fields, Journal of Geophysical Research, 87, 11 229, https://doi.org/10.1029/JC087iC13p11229, http://doi.wiley.com/10.1029/JC087iC13p11229, 1982.

Moesinger, L., Dorigo, W., De Jeu, R., Van der Schalie, R., Scanlon, T., Teubner, I., and Forkel, M.: The Global Long-term Microwave
30    Vegetation Optical Depth Climate Archive VODCA [Data set], https://doi.org/10.5281/ZENODO.2575599, https://zenodo.org/record/2575599{#}.XIjXc4Uo82A, 2019.

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, Journal of Geophysical Research: Biogeosciences, 122, 3031–3046, https://doi.org/10.1002/2017JG004145, http://doi.wiley.com/10.1002/2017JG004145, 2017.

35  Morton, D. C., Nagol, J., Carabajal, C. C., Rosette, J., Palace, M., Cook, B. D., Vermote, E. F., Harding, D. J., and North, P. R. J.: Amazon forests maintain consistent canopy structure and greenness during the dry season, Nature, 506, 221–224, https://doi.org/10.1038/nature13006, http://www.nature.com/articles/nature13006, 2014.

Myneni, R., Knyazikhin, Y., and Park, T.: MCD15A2H MODIS/Terra+Aqua Leaf Area Index/FPAR 8-day L4 Global 500m SIN Grid V006 [Data set], https://doi.org/doi: 10.5067/MODIS/MCD15A2H.006, 2015.

Myneni, R. B., Yang, W., Nemani, R. R., Huete, A. R., Dickinson, R. E., Knyazikhin, Y., Didan, K., Fu, R., Negrón Juárez, R. I., Saatchi, S. S., Hashimoto, H., Ichii, K., Shabanov, N. V., Tan, B., Ratana, P., Privette, J. L., Morisette, J. T., Vermote, E. F., Roy, D. P., Wolfe, R. E., Friedl, M. A., Running, S. W., Votava, P., El-Saleous, N., Devadiga, S., Su, Y., and Salomonson, V. V.: Large seasonal swings in leaf area of Amazon rainforests., Proceedings of the National Academy of Sciences of the United States of America, 104, 4820–3, https://doi.org/10.1073/pnas.0611338104, http://www.ncbi.nlm.nih.gov/pubmed/17360360http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1820882, 2007.

Njoku, E., Ashcroft, P., Chan, T., and Li Li: Global survey and statistics of radio-frequency interference in AMSR-E land observations, IEEE Transactions on Geoscience and Remote Sensing, 43, 938–947, https://doi.org/10.1109/TGRS.2004.837507, http://ieeexplore.ieee.org/document/1424270/, 2005.

Owe, M., de Jeu, R., and Holmes, T.: Multisensor historical climatology of satellite-derived global land surface moisture, Journal of Geophysical Research, 113, F01 002, https://doi.org/10.1029/2007JF000769, http://doi.wiley.com/10.1029/2007JF000769, 2008.

Saleska, S. R., Didan, K., Huete, A. R., and da Rocha, H. R.: Amazon forests green-up during 2005 drought., Science (New York, N.Y.), 318, 612, https://doi.org/10.1126/science.1146663, http://www.ncbi.nlm.nih.gov/pubmed/17885095, 2007.

Samanta, A., Ganguly, S., Hashimoto, H., Devadiga, S., Vermote, E., Knyazikhin, Y., Nemani, R. R., and Myneni, R. B.: Amazon forests did not green-up during the 2005 drought, Geophysical Research Letters, 37, n/a–n/a, https://doi.org/10.1029/2009GL042154, http://doi.wiley.com/10.1029/2009GL042154, 2010.

Samanta, A., Ganguly, S., Vermote, E., Nemani, R. R., Myneni, R. B., Samanta, A., Ganguly, S., Vermote, E., Nemani, R. R., and Myneni, R. B.: Why Is Remote Sensing of Amazon Forest Greenness So Challenging?, Earth Interactions, 16, 1–14, https://doi.org/10.1175/2012EI440.1, http://journals.ametsoc.org/doi/abs/10.1175/2012EI440.1, 2012.

Simard, M., Pinto, N., Fisher, J. B., and Baccini, A.: Mapping forest canopy height globally with spaceborne lidar, Journal of Geophysical Research, 116, G04 021, https://doi.org/10.1029/2011JG001708, http://doi.wiley.com/10.1029/2011JG001708, 2011.

Song, X.-P., Hansen, M. C., Stehman, S. V., Potapov, P. V., Tyukavina, A., Vermote, E. F., and Townshend, J. R.: Global land change from 1982 to 2016, Nature, 560, 639–643, https://doi.org/10.1038/s41586-018-0411-9, http://www.nature.com/articles/s41586-018-0411-9, 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, International Journal of Applied Earth Observation and Geoinformation, 65, 79–91, https://doi.org/10.1016/j.jag.2017.10.006, http://dx.doi.org/10.1016/j.jag.2017.10.006, 2018.

van der Schalie, R., de Jeu, R., Kerr, Y., Wigneron, J., Rodríguez-Fernández, N., Al-Yaari, A., Parinussa, R., Mecklenburg, S., and Drusch, M.: The merging of radiative transfer based surface soil moisture data from SMOS and AMSR-E, Remote Sensing of Environment, 189, 180–193, https://doi.org/10.1016/J.RSE.2016.11.026, https://www.sciencedirect.com/science/article/pii/S0034425716304734, 2017.

Vreugdenhil, M., Hahn, S., Melzer, T., BauerMarschallinger, B., Reimer, C., Dorigo, W. A., and Wagner, W.: Assessing Vegetation Dynamics Over Mainland Australia With Metop ASCAT, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10, 2240–2248, https://doi.org/10.1109/JSTARS.2016.2618838, http://ieeexplore.ieee.org/document/7762756/, 2017.

Wentz, F. J.: A well-calibrated ocean algorithm for special sensor microwave / imager, Journal of Geophysical Research: Oceans, 102, 8703–8718, https://doi.org/10.1029/96JC01751, http://doi.wiley.com/10.1029/96JC01751, 1997.