Technische Universität Wien
Fakultät für Mathematik und Geoinformation
Department für Geodäsie und Geoinformation

DIPLOMARBEIT

VODCA2GPP - A new global, long-term (1988-2020) GPP dataset from microwave remote sensing

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Masterstudiums

Geodäsie und Geoinformation

eingereicht von

Benjamin Wild, BSc
Matrikelnummer 01504060

ausgeführt am
Department für Geodäsie und Geoinformation
der **Fakultät für Mathematik und Geoinformation**
der **Technischen Universität Wien**

unter der Betreuung von
Univ.Prof. Dr.rer.nat. Wouter Arnoud Dorigo

Wien, am 10. Juli 2021

_____   _____
Unterschrift Verfasser   Unterschrift Betreuer

Declaration

I hereby declare that I have written this thesis by myself without any help or assistance of others. External literature used to clarify the content or provided data sources are fully cited. All mentioned information is in accordance with fact or truth up to my knowledge.

Vienna, on July 11, 2021 _____

Abstract

Long-term global monitoring of terrestrial Gross Primary Production (GPP) is crucial for assessing ecosystem response to global climate change. In recent years and decades, great advances in estimating GPP on a global level have been made and many global GPP datasets have been published. These global data records are either based on observations from optical remote sensing, are upscaled from in situ measurements, or rely on process-based models. The different estimation approaches are well established within the scientific community but also exhibit significant discrepancies among each other.

Here, the new VODCA2GPP dataset is introduced, which utilizes microwave remote sensing estimates of Vegetation Optical Depth (VOD) to estimate GPP on a global scale. VODCA2GPP is able to complement existing products with long-term GPP estimates covering the period 1988 - 2020. VODCA2GPP applies a previously developed carbon sink-driven approach (Teubner et al. [2019], [2021]) to estimate GPP from the Vegetation Optical Depth Climate Archive (Zotta et al. [in preperation]; Moesinger et al. [2020]), which merges VOD observations from multiple sensors into one long-running, coherent data record. VODCA2GPP was trained and evaluated against FLUXNET in situ observations of GPP and assessed against largely independent state-of-the art GPP datasets (MODIS GPP, FLUXCOM GPP, and GPP estimates from the TRENDY-v7 model ensemble).

These assessments show that VODCA2GPP exhibits very similar spatial patterns compared to existing GPP datasets across all biomes but with a consistent positive bias. In terms of temporal dynamics, a high agreement was found for regions outside the humid tropics, with median correlations around 0.75. Concerning anomalies from the long-term climatology, VODCA2GPP correlates well with MODIS and TRENDY-v7 GPP (Pearson's r: 0.53 and 0.61) but less with FLUXCOM GPP (Pearson's r: 0.29). A trend analysis for the period 1988-2019 did not exhibit a significant trend in VODCA2GPP on a global scale but rather suggests regionally differing long-term changes in GPP. Significant similar increases of global GPP that were found for VODCA2GPP, MODIS GPP, and the TRENDY-v7 ensemble for the shorter overlapping observation period (2003-2015) supports the theory of elevated CO_2 uptake potentially induced by increased

atmospheric CO2 concentrations and the associated rising temperatures.

Preface

This Master Thesis builds on a paper that was recently submitted to the Earth System Science Data Journal (ESSD, Wild et al. [2021a]). The submitted paper is lead-authored by me, goes under the same title as this thesis and has not been published yet. The co-authors of the submitted paper are: Irene Teubner, Leander Moesinger, Ruxandra-Maria Zotta, Matthias Forkel, Stephen Sitch, Robin van der Schalie, and Wouter Dorigo. I contributed to the paper by curating the dataset, conducting and assessing the analysis and drafting the manuscript.

For this Master Thesis I extended the paper by providing more theoretical background, especially on Gross Primary Production and the role of remote sensing in the quantification of Gross Primary Production. (Sub-)chapters that were directly adopted or contain adapted content from the paper are indicated with a footnote ([1]) at the beginning of the chapter. The abstract was also adopted from Wild et al. [2021a].

Following the open data philosophy that is important to all contributors of the dataset, VODCA2GPP is publicly and freely available (CC-BY-SA-NC 4.0) and can be accessed here: `https://doi.org/10.48436/1k7aj-bdz35` (Wild et al. [2021b]).

Contents

List of Figures

2

3

List of Tables

1 Introduction

[1]Gross Primary Productivity (GPP) describes vegetation's synthesis of atmospheric CO_2 to carbohydrates through photosynthesis and therefore plays a key role in the global carbon cycle (Bonan [2015]). Especially in the context of global climate change, GPP plays an important role as "filter" from the most critical greenhouse gas CO_2.

CO_2 concentrations in the atmosphere have steadily increased since the beginning of industrial revolution (Fig 1.1) which has triggered a significant temperature increase within the last decades (Allen et al. [2018]). Depending on the measures taken by governments and society to mitigate CO_2 emissions, CO_2 concentrations, and thus also temperature, are projected to increase at the same or at an even faster pace in the years and decades to come (Fig 1.1). As discussed in more detail in Chapter 2, GPP is strongly influenced by both, atmospheric CO_2 and temperature. Thus monitoring of GPP will become even more crucial for understanding the impact of climate change on the global carbon cycle (Haverd et al. [2020b]; Schimel et al. [2015]).
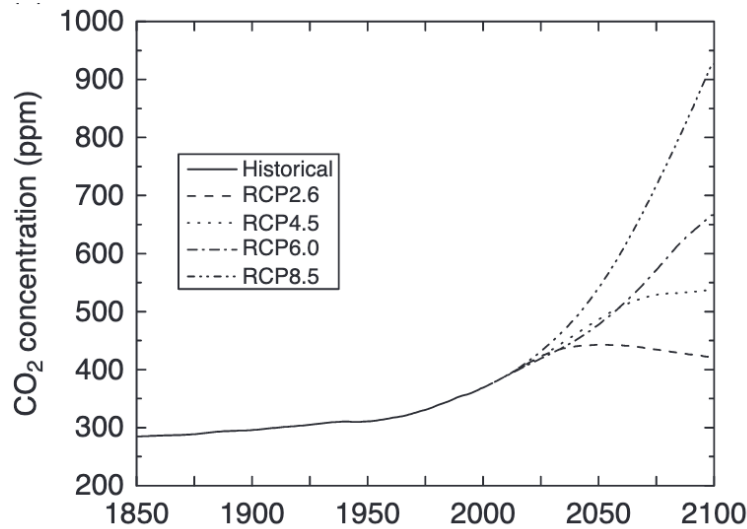
Figure 1.1: Evolution of historical and projection of future atmospheric CO_2 concentrations under various representative concentration pathways (RCPs) [Bonan, 2015]

However, direct measurements of GPP are difficult and costly which is why in-situ GPP measurement systems are mostly located in industrialized countries. This observation bias causes large uncertainties in many biomes of the world where hardly any direct GPP observations are available which massively hampers the analysis of global GPP variations.

On the global scale, GPP is commonly estimated using optical remote sensing data in combination with (semi-)empirical or machine learning models (e.g., O'Sullivan et al. [2020]; Alemohammad et al. [2017]; Gilabert et al. [2017]; Jung et al. [2020]; Tramontana et al. [2016]). However, optical remote sensing is strongly affected by cloud cover, leading to data gaps and high uncertainties in regions with frequent cloud cover and high production such as tropical forests.

In contrast to optical remote sensing observations, Vegetation Optical Depth (VOD) from microwave remote sensing is much less affected by weather conditions. VOD describes the vegetation's attenuation of radiation in the microwave domain, which is controlled by its water content, biomass, type and density (Jackson and Schmugge [1991]; Vreugdenhil et al. [2016]). Thus, VOD has been intensively used as proxy for aboveground biomass (Li et al. [2021]; Rodríguez-Fernández et al. [2018]; Tian et al. [2016]; Liu et al. [2011]) and is becoming increasingly important in the monitoring of vegetation dynamics (e.g., Frappart et al. [2020]; Piles et al. [2017]; Tian et al. [2016]).

Recent studies investigated how GPP can be estimated from VOD (Teubner et al. [2018], [2019], [2021]). GPP is significantly correlated with spatial patterns and temporal changes in VOD (Teubner et al. [2018]). Based on this relationship, Teubner et al. [2019] developed a theoretical framework and a machine-learning method using FLUXNET site observations to predict GPP using VOD. They showed that GPP can be adequately estimated for most regions of the world with an overall tendency for moderate overestimation and good temporal agreement with existing GPP products, especially for temperate regions. Recently, Teubner et al. [2021] improved this method by adding air-temperature in their model to account for temperature dependency of plant respiration, and found that this significantly improved the temporal agreement with reference GPP data.

However, long-term analysis of GPP from VOD was until recently hampered by the relatively short availability of passive microwave sensors. Moesinger et al. [2020] over-

came this issue by merging single-frequency VOD from various sensors into the long-term VOD Climate Archive (VODCA) which comprises VOD observations of more than 20 years for X-, and C-Band and more than 30 years for Ku-Band. A new version of VODCA (Zotta et al. [in preperation]) does not only combine single sensors from identical frequencies but also merges observations from different bands (X, C and Ku) into one long-running multi-frequency VOD climate archive.

1.1 Objective of this thesis

The objective of this thesis is to introduce a novel long-term GPP dataset (VODCA2GPP) which was generated by applying the method developed by Teubner et al. [2019], [2020] to the multi-frequency VODCA dataset. Furthermore, this work aims to provide a detailed evaluation the dataset, especially in regard to its complementarity to existing state-of-the-art GPP datasets.

This thesis is organised as follows: In Chapter 2 and 3 a general overview on the (gross) primary productivity is given (Chapter 2) with a focus on remote sensing techniques for assessing vegetation dynamics (Chapter 3). Chapter 4 describes the datasets that were used a) in the production process of VODCA2GPP and b) as references for the evaluation of VODCA2GPP. Chapter 5 gives a overview on the VOD2GPP-model which represents the theoretical basis for VODCA2GPP. Furthermore, the practical VOD2GPP implementation with Generalized Additive Models (GAMs) is outlined and other important steps in the processing chain of VODCA2GPP are explained. Lastly the utilized means of validation and comparison with the reference datasets are described. In chapter 6 the results of the validation process are shown, followed by a discussion of the results with potential sources of errors in Chapter 7.

2 Defining primary productivity

2.1 What is GPP

Primary productivity in ecosystems is defined as the amount of carbon that is fixed by primary producers (i.e. autotroph organisms like plants) by the means of photosynthesis (Bonan [2015]). Gross primary productivity can be split into

- plant respiration (R_a) and

- Net Primary Productivity (NPP).

Gross Primary Production is the sum of both, NPP and R_a, and the relationship between NPP, GPP and R_a can thus be formally written as (Bonan [2015]):

$$GPP = R_a + NPP \tag{2.1}$$

NPP describes vegetation's synthesis of atmospheric CO_2 to carbohydrates through photosynthesis but only includes the resulting amount of stored biomass which is used by plants to form tissues in order to adapt to environmental conditions (e.g., forming strong roots in arid regions). In contrast to that, R_a describes all respiration processes that are fueled by the remaining energy (Cain et al. [2014]). R_a can be further separated into

(a) growth respiration, which is mainly associated with the plant's ability to form organic compounds and amounts to aproximately 25% of GPP (Ryan [1991]) and

(b) maintanance respiration which is the process that maintains living cells by breaking down carbohydrates and in return reduces the amount of carbon that can be used for growth.

GPP is therefore defined as the total amount of carbon that is fixed by plants in a defined timespan over a certain area. Common units to quantify GPP are $[g\,C\,m^{-2}\,d^{-1}]$, $[kg\,C\,m^{-2}\,yr^{-1}]$ or, when quantifying global sums of GPP, $[Pg\,C\,m^{-2}\,yr^{-1}]$.

With recent estimates ranging from 120 (Beer et al. [2010]) up to 175 (Welp et al. [2011]) $Pg\,C\,m^{-2}\,yr^{-1}$, it is the largest CO_2 flux in flux in the carbon cycle (Beer et al. [2010]) and also considered to be the primary driver of the terrestrial carbon sink. GPP is responsible for the sequestration of approximately 30% of the global annual anthropogenic CO_2 emmisions (Friedlingstein et al. [2020]) and therefore not only plays a key role in the global carbon cycle but also significantly contributes to cleansing of the atmosphere from the most important greenhouse gas.

2.2 What drives GPP

The main source of energy that controls the primary production is photosynthetically active radiation (PAR; approximately in the wavelength range between 400 and 700 nm) which is emmited by the sun. This radiation is used by plants in the process of photosynthesis to convert CO_2 and H_2O (water) to CH_2O (i.e. carbohydrates such as sugar) (Bonan [2015]). Accordingly, the equation photosynthesis can be written as:

$$CO_2 + 2H_2O + light\ energy \rightarrow CH_2O + O_2 + H_2O \qquad (2.2)$$

where O_2 can be considered a byproduct.

Equation 2.2 shows the main drivers of GPP: energy from sunlight, water availability and the amount of available atmospheric CO_2. However, there are also other often co-varying environmental conditions that have a significant impact on GPP such as nutrient availability or temperature.
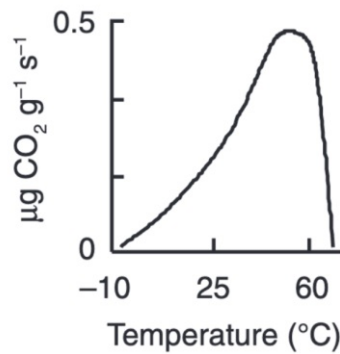
Figure 2.1: Temperature-dependence of maintenance respiration (Bonan [2015])

Varying temperature, for example, is known to strongly influence autotrophic respiration (Teubner et al. [2021]). Higher temperatures expedite maintenance respiration

exponentially up to 50°C (Fig 2.1) while growth respiration is hardly affected by temperature (Bonan [2015]).

The mentioned dependencies explain the high degree of (inter-)annual variability of GPP, which is particularly strong in regions with pronounced seasonality (e.g., temperate or boreal regions). Beside the distinct seasonal fluctuations of GPP also long-term changes are observable on a regional and global scale.

2.3 Recent trends in global GPP

In recent years, many studies have found positive trends in global GPP and linked those trends to changing meteorological conditions and atmospheric compositions. Campbell et al. [2017], for example, quantified the increase of GPP for the period 1900-2000 with ca. 31%. They hypothesised that increasing CO_2 concentrations are responsible for the strong increase. This so-called CO_2 fertilization effect has been investigated more closely by other studies (e.g. Schimel et al. [2015]; Haverd et al. [2020b]; Cox et al. [2000]; Canadell et al. [2007]) which all came to the conclusion that elevated GPP is indeed a response to increasing levels of atmospheric CO_2 (Fig 2.2).
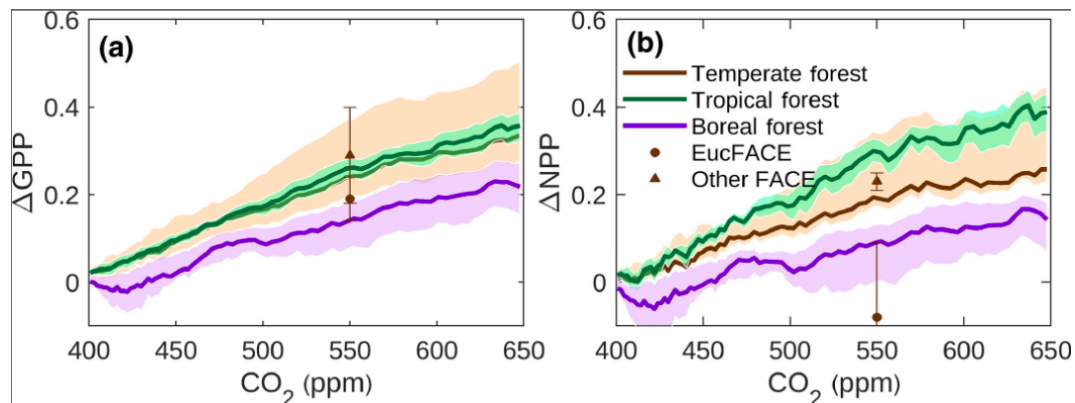
Figure 2.2: Relative regional changes in GPP (ΔGPP; left) and NPP (ΔNPP; right) as derived from the Community Atmosphere–Biosphere Land Exchange (CABLE) model plotted against absolute atmospheric CO_2 concentrations (Haverd et al. [2020b])

In addition to CO_2 fertilization effects, rising temperatures (Schimel et al. [2015]; Piao et al. [2013]; Cox et al. [2000]) as well as land-use change (Haverd et al. [2020b]) have been identified as an important driver for positive trends in GPP.

However, it is to be noted that most of these findings rely on GPP estimates that are derived from process-based models such as Dynamic Global Vegetation Models (DGVMs) or Land Surface Models (LSM) which are based on various theoretical model assumptions and thus exhibit a certain degree of uncertainty. The validation of such models is often difficult due to a general shortage of direct GPP observations. Its important role in the global carbon cycle, however, has lead to increased scientific effort in developing independent methods for observing GPP locally and globally.

2.4 Small-scale measurements of GPP

[1]Locally, GPP can be adequately determined at in-situ flux towers which measure carbon dioxide fluxes by the means of eddy-covariances. Eddy-covariance is a statisitcal method that allows the direct estimation of gas and energy fluxes at atmospheric boundary layers (e.g., between terrestrial ecosystems and the atmosphere) by the means of micro-meteorological observations (Liang et al. [2012]). These eddy-covariances can be further partitioned into GPP and ecosystem respiration (Baldocchi [2003]).

There exist numerous eddy-covariance stations that are continously measuring CO_2 fluxes. Most of these stations are associated to regional networks that incorporate from only a few to several hundred (e.g., AmeriFlux) individual eddy-covariance towers. FLUXNET is the global network of regional flux tower networks to provide the scientific community with harmonized and well-documented flux observations (Pastorello et al. [2020]). However, FLUXNET in-situ stations are sparsely and unevenly distributed (cf 8.1) across the planet with most stations located in industrialized countries which complicates the derivations of global ecosystem dynamics.

3 Remote sensing in monitoring global GPP

Remote sensing allows to overcome the strong spatial limitations that are associated with the in-situ based monitoring of GPP by providing information on biogeophysical variables, that serve as proxies for GPP on a global scale. Basically, remote sensing of vegetation can be separated into two categories:

- optical remote sensing and

- microwave remote sensing

The difference between the two categories lies within their operation frequencies which allows them to gain largely independent information about similar bio-geophysical properties. Both methods are being extensively used for the remote sensing of vegetation but only optical methods have so far been applied for the long-term global quantification of GPP (O'Sullivan et al. [2020]) because of their close relationship to PAR.

3.1 Optical remote sensing of vegetation

Optical remote sensing uses sensors that operate with wavelengths between approximately 0.4 μm and 15 μm (Kerekes [2009]). Remote sensing of vegetation most commonly makes use of the visible (VIS 0.4 μm - 0.7 μm) and the near-infrared (NIR; 0.7 μm - 1.1 μm) region of the electromagnetic (EM) spectrum because of the distinct vegetation's reflection mechanisms that come into play in these spectral regions.

The more photosythentically active plants are, the higher is also the degree of absorption of photosynthetically active radiation (PAR; corresponds approximately to the visible spectrum) which they use as source of energy for photosynthesis. While the degree of reflection of PAR is highly dependent on the photosynthetic capacity (i.e. the vitality) of vegetation, the absorption mechanisms in the NIR region are mostly controlled by the cellular structure of the leaves which reflect most of the incoming NIR radiation. Thus, vital vegetation exhibits a distinct drop in absorption between those two spectral regions while this drop becomes less distinct for stressed vegetation or non-existent for other objects (Tucker and Sellers [1986]).

13

This behaviour can be expressed by vegetation indices such as the Normalized Difference Vegetation Index (NDVI), which serves as a proxy for vegetation coverage and vegetation health (Carlson and Ripley [1997]). The NDVI, for example, is obtained by calculating the (normalized) difference between the (normalized) red (RED) and NIR reflectances:

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{3.1}$$

As such, the NDVI can take on values between -1 and 1. Values close to 1 indicate vegetation with high photosynthetic activity while positive values close to 0 either indicate stressed vegetation or non-vegetated surfaces. In rare cases also negative values occur which are often caused by atmospheric disturbances (Tucker and Sellers [1986]).

Besides the NDVI, there are other optical remote sensing vegetation indicators that rely on similar principles. Other vegetation health variables are for example:

- fraction of Absorbed Photosynthetically Active Radiation (fAPAR)

- Enhanced Vegetation Index (EVI)

- Solar Induced Fluorescence (SIF)

3.1.1 Optical vegetation variables as proxie for GPP

[1]So far, GPP has been commonly estimated on the basis of these optical remote sensing vegetation indices in combination with (semi-)empirical or machine learning models (e.g. Tramontana et al. [2016]; Gilabert et al. [2017] or Alemohammad et al. [2017]). Specifically, these models are based on light use efficiency (LUE) theory and/or statistical models that are applied to derive GPP based on optical remote sensing variables as discussed above.

The usage of optical remote sensing data makes these models closely linked to photosynthetic activity in plants and hence also the terrestrial carbon cycle. Furthermore, these optical remote sensing based GPP datasets have the advantage of being available virtually globally with high spatial (usually in the order of 10 m to 1 km) and decent temporal granularity (approximate revisit times of one to several weeks, depending on the spatial resolution and satellite orbits) which is also why optical remote sensing based GPP products such as FLUXCOM GPP (Jung et al. [2019]) or MODIS GPP (Running et al. [1999]) can be considered state-of-the-art in the monitoring of GPP on a global scale.

3.1.2 Issues related to optical remote sensing

Despite being able to provide detailed information on the photosynthetic activity in plants there are also several shortcoming related to the remote sensing in the optical domain. For example, optical remote sensing tends to underestimate GPP in very productive regions (Chen et al. [2006]; Asner et al. [2003]) presumably due to saturation effects of mentioned vegetation health indicators. Also, optical remote sensing is strongly affected by cloud cover which results in additional uncertainties especially in highly productive regions (e.g., tropical forests). Furthermore, observations from space are highly dependent on the illumination conditions on the ground which necessitates the usage of sun-synchronous orbits to obtain observations with a fixed sun-angle (Carver et al. [1985]).

3.2 Microwave remote sensing of vegetation

Due to the differences in operation frequencies, microwave remote sensing can be viewed as independent from optical earth observation techniques. This is why within the last years and decades increased effort was put into the development of innovative microwave sensor systems that are today also extensively used in the monitoring of global ecosystems and vegetation dynamics.

A concrete definition for the wavelength boundaries of the microwave spectrum does not exist but microwaves are commonly described as EM waves in the wavelength region between 1 mm to 1 m (Ulaby et al. [1981]). Microwave frequencies are classified in frequency regions (i.e. bands). For the remote sensing of vegetation the most commonly used bands are listed in Table 3.1.

Table 3.1: Band definitions of microwave frequencies according to the Radio Society of Great Britain (URL: `https://rsgb.org/`)

Band	Frequency range [GHz]	Wavelength range [mm]
L	1-2	150-300
C	4-8	37.5-75
X	8-12	25-37.5
Ku	12-18	16.7-25

Microwave remote sensing sensors can be split into two sub-categories (Konings et al. [2019]):

- active sensors (radars) and

- passive sensors (radiometers)

3.2.1 Active systems

Active systems are equipped with antennas that emit EM waves which are reflected, absorbed or scattered on the earth's surface. The main quantity that is measured is the amount of radiation that returns to the sensor (i.e. the backscatter) Konings et al. [2019]).

Active sensor systems have the potential to deliver observations at very high spatial resolution (up to 10 m for Synthethic Aperture Radar (SAR)). However, vegetation structures and geometry play a large role in the active case (Fig 3.1) which often complicates the seperation of the vegetation signal (i.e. the VOD) from the other contributors. Therefore, also the parametrization of vegetation can become very complex (Konings et al. [2019]). Nevertheless, there have been advances to retrieve global VOD also from the actively operating Advanced Scatterometer (ASCAT) (Vreugdenhil et al. [2016]). Their approach mainly relies on the the water-cloud model introduced by Attema and Ulaby [1978].

Until today, however, most VOD datasets are still based on observation from passive microwave sensors (e.g., TMI, AMSR-E, AMSR2, etc.) which is why a more detailed introduction of active microwave remote sensing based VOD retrievals is ommited.

3.2.2 Passive systems

In contrast to radar systems, radiometers do not emit EM radiation themselves but sense the amount of radiation that is naturally emitted by the earth's surface which is commonly expressed as brightness temperatures T_b (Konings et al. [2019]). Radiometers typically have coarse resolutions ($> 10\,km$) but as it can be seen in Fig 3.1 there are only three main (emission) contributors that come into play which facilitates the isolation of the vegetation signal.

According to Ulaby et al. [1981], the measured signal (i.e. T_B) can be seperated into contributions from soil (T_{Bs}), vegetation (T_{Bv}) and the interaction between soil and vegetation (T_{Bvs}):

$$T_B = T_{Bv} + T_{Bs} + T_{Bvs} \tag{3.2}$$

The three components are further defined as:

$$T_{Bv} = T_v(1 - \alpha)(1 - \gamma_v) \tag{3.3}$$

$$T_{Bs} = \epsilon_s T_s \gamma_v \tag{3.4}$$

$$T_{Bvs} = T_v R(1 - \alpha)\gamma_v(1 - \gamma_v) \tag{3.5}$$

where T_v and T_s denote the temperatures of the vegetation layer and the soil-surface. ϵ_s denotes soils emmisivity and R its reflectivity. α corresponds to the single-scattering albedo. The most important quantity in context of vegetation is γ_v which denotes the one-way transmissivity of the vegetation cover. In other words, γ_v describes the attenuation of the signal while travelling through the vegetation cover. γ_v is given by

$$\gamma_v = e^{\frac{-VOD}{cos(\theta)}} \tag{3.6}$$

The degree of attenuation is thus dependent on the incidence angle θ and VOD. Therefore, high VOD indicates a high degree of attenuation by the vegetation cover and vice versa (Ulaby et al. [1981]).

A common approximation for VOD is the linear relationship with Vegetation Water Content (VWC) Jackson and Schmugge [1991]:

$$VOD = bVWC \tag{3.7}$$

It is to be noted that the equations 3.2-3.7[1] are only valid under the assumption of a completely smooth and plane soil surface and an idealized, homogenous vegetation layer (i.e. "water cloud").

In reality, VOD is not only dependent on VWC but also influenced by the vegetation's structure. The slope b is mainly dependent on the landcover type and the frequency, which is why the used frequency needs to be considered when analysing VOD retrievals (Konings et al. [2019]). In general, high frequency microwaves (Ku-/X-/C-Band) are more sensitive to smaller vegetation structures like leaves and branches while low frequencies (L-Band) are sensitive to larger components like stems.

---

[1]The equations 3.2-3.6 are taken from Ulaby et al. [1981] and Konings et al. [2019]
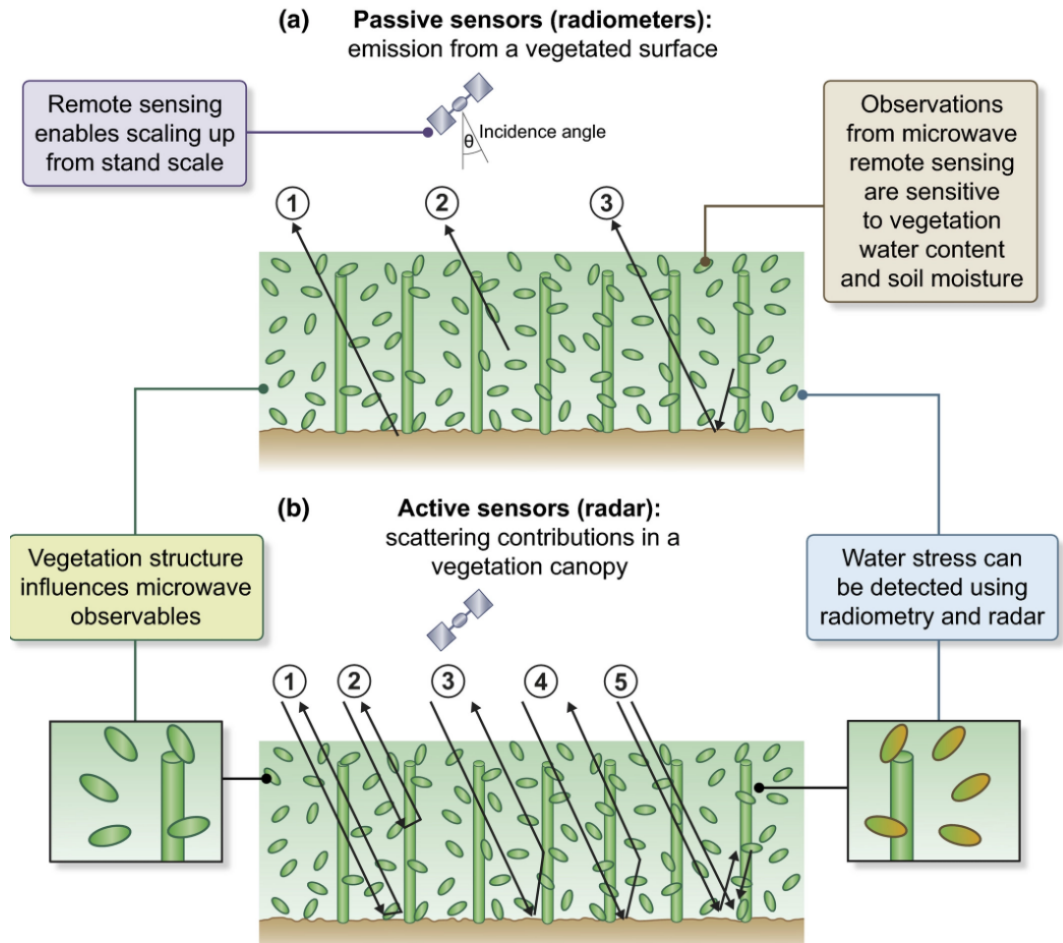
Figure 3.1: Illustration of an idealized vegetated surface and its main emission contrib-
utors (a): 1) bare soil, 2) the vegetation itself and 3) two-path emmision
that is first reflected by the soil and then attenuated by the vegetation. (b)
illustrates the backscatter contributors (in the active case): 1) backscatter
from bare soil (direct), 2) backscatter from plants (direct), 3) and 4) two-way
backscatter from plants-ground and 5) three-way backscatter ground-plants-
ground (Konings et al. [2019])

3.2.3 Advantages of microwave remote sensing over optical

[1]In contrast to optical remote sensing, microwave sensors are hardly affected by weather
conditions due to their much stronger capability to penetrate the atmosphere also at
cloudy conditions. This allows the observation of continuous time-series throughout the
year which is especially relevant in regions with significant cloud cover such as the Trop-
ics (Ulaby et al. [1981]). Apart from being virtually weather independent, microwave
remote sensing also has the advantage of being less dependent on illumination condi-

tions on the earth's surface (Carver et al. [1985]). However, the effect of surface heating during daytime needs to be considered which can be realised by only taking nightime obersvations into account (Owe et al. [2008]). Furthermore, unlike optical vegetation indices, VOD is less prone to saturation effects and theoretically can take on values between $[0,\infty[$.

3.3 Relationship between VOD and GPP

[1]Various studies showed that AGB and changes in AGB are correlated with R_a (Lavigne and Ryan [1997]) and NPP (Girardin et al. [2010]; Luyssaert et al. [2007]; Clark et al. [2001]) which are the both constituents of GPP (Eq. 2.1). Following the theoretical relationship between VOD, AGB and GPP it has been demonstrated that VOD as well as temporal changes in VOD are significantly correlated with GPP (Teubner et al. [2018]). According to the mentioned relationship between AGB, VOD and GPP, theoretically highest correlation with GPP would be expected for L-Band VOD which corresponds best to AGB (Rodríguez-Fernández et al. [2018]). However, best correspondence between GPP and VOD was found for X-Band observations. Teubner et al. [2021] explained the better agreement of GPP with X-Band VOD with the higher sensivitiy of X-Band to parts of the vegetation that exhibit the metabolic activity (e.g. leaves).

Based on these findings Teubner et al. [2019] derived the VOD2GPP-model which is the basis for the VODCA2GPP dataset.

4 Data

4.1 Model Input data

4.1.1 VODCA

[1]The Vegetation Optical Depth Climate Archive (VODCA v1; Moesinger et al. [2020]) was first published in 2020 and consisted of three single-frequency VOD products (Ku-, X- and C-band), covering the period from 1987-2017 (Ku-Band), 1997-2019 (X-Band), and 2002-2019 (C-band), respectively. For VODCA2GPP, an updated VODCA version (VODCA v2 CXKu; Zotta et al. [in preperation]) that merges all bands in a single dataset to obtain increased spatial and temporal coverage and reduced random errors compared to VODCA v1 was used. VODCA v2 CXKu utilizes observations from the same sensors and frequencies as VODCA v1 (Table 3.6) to generate a single long-running VOD multi-frequency time series. It merges 15 passive night-time VOD datasets retrieved from seven different sensors via the Land Parameter Retrieval Model (LPRM; van der Schalie et al. [2017]). LPRM is based on radiative transfer theory introduced by Mo et al. [1982] and uses forward modelling to simulate the top of atmosphere brightness temperatures under a wide range of conditions and compares this to the actual satellite observation. Although primarily developed for soil moisture, it simultaneously solves for the VOD using an analytical solution by Meesters et al. [2005], utilizing the ratio between H and V polarized observations (van der Schalie et al. [2017]). LPRM requires that the soil and vegetation temperatures are equal, which may not be the case during the day due to uneven heating from solar radiation. VODCA v2 therefore uses only night-time observations which are assumed to have thermal equilibrium (Owe et al. [2008]). Scaling of the single-sensor VOD observations is done by means of cumulative distribution function (CDF) matching (Moesinger et al. [2020]).

The preprocessing of LPRM level 2 VOD data used in VODCA v2 follows the steps described in detail in Moesinger et al. [2020]. These include projecting the data into a $0.25° \times 0.25°$ grid, using nearest neighbour resampling, selecting the closest night-time value in a window of $\pm$ 12 hours for every 0:00 UTC. The data are masked for radio-frequency interference (de Nijs et al. [2015]), and negative VOD values and temperature. Different from Moesinger et al. [2020], masking for low land surface temperature (LST

$< 0°C$), when the dielectric properties of water change drastically, is not based on Ka-band retrievals because these have high uncertainties over frozen land (Holmes et al. [2009]). Instead, VODCA v2 uses the ERA-5 Land (Muñoz-Sabater et al. [2021]) soil temperature level 1 (stl1) data. To ensure that all observations taken under frozen conditions are masked, all observations with an associated surface soil temperature (stl1) below $3°$ C are masked (Zotta et al. [in preperation]).

VODCA v2 CXKu is obtained by first scaling VODCA v2 observations from C- and Ku-band to X-band to remove systematic biases and then computing a weighted average in order to fuse overlapping observations. VODCA v2 CXKu provides a single, long-term vegetation metric covering over 30 years of observations (1988-2020) and thus exceeding the temporal length of the individual sensor products (VODCA v2 C-, X and Ku). Due to the novel weighted averaging technique used in computing this dataset, the random error levels are lower, enabling the detection of long-term trends and dynamics at global scales (Zotta et al. [in preperation]).

4.1.2 ERA5-Land

[1]For representing the autotrophic respiration's temperature dependency 2 m air temperature (T2m) was derived from the ERA5-Land dataset. T2m is the most commonly used parameter for the describing the relationship between autotrophic respiration and temperature (Teubner et al. [2021]; Drake et al. [2016]; Ryan et al. [1997]). ERA5-Land is a reanalysis dataset of meteorological variables which is provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) (Muñoz-Sabater et al. [2021]). ERA5-Land is produced at a spatial resolution of 8 km ($0.08°$) and is available hourly.

4.1.3 FLUXNET2015 in-situ GPP

[1]In-situ GPP data from Tier1 v1 FLUXNET2015 (Pastorello et al. [2020]) was used to train and evaluate the VODCA2GPP product. FLUXNET GPP estimates are available for night-time and day-time flux partitioning, which were averaged as suggested by (Pastorello et al. [2020]). FLUXNET data is available daily from 1991 until 2014 with a mean observation timespan of 7.27 ± 4.89 years for the used stations indicating significant variability in station data availability. An overview of the used FLUXNET2015 stations can be found in Table 8.2 and Fig A8.1.

Table 4.1: Input data for the used merged-band VODCAv2 with the main sensor specifications: Temporal coverage, local ascending equatorial crossing times (AECT) and available frequencies. Table information is taken from Moesinger et al. [2020] and adapted for VODCAv2 XCKu

Sensor	Time period used	AECT	C-Band [GHz]	X-Band [GHz]	Ku-Band [GHz]	reference
AMSR-E	Jun 2002 - Oct 2011	13:30	6.93	10.65	18.70	van der Schalie et al. [2017]
AMSR2	Jul 2012 - Dec2020 Jul 2012-Jan 2019 (Ku-Band)	13:30	6.93, 7.30	10.65	18.70	van der Schalie et al. [2017]
SSM/I F08	Jul 1987 - Dec 1991	18:15			19.35	Owe et al. [2008]
SSM/I F011	Dec 1991 – May 1995	17:00-18:15			19.35	Owe et al. [2008]
SSM/I F13	May 1995 – Apr 2009	17:45 – 18:40			19.35	Owe et al. [2008]
TMI	Dec 1997 – Apr 2015	Asynchronous	0.483	10.65	19.35	Owe et al. [2008]
WindSat	Feb 2003 - Jul 2012	18:00	0.680	10.70	18.70	Owe et al. [2008]

22

4.2 Reference datasets

4.2.1 MODIS GPP

[1]GPP estimates derived from Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data are based on light-use efficiency concept introduced by Monteith [1972] which relates the amount of absorbed solar radiation to vegetation productivity. The MODIS algorithm uses fAPAR as proxy for the absorbed solar energy. For this study the MOD17A2H v006 GPP product was used (Running et al. [2015]; Zhao et al. [2005]). It is available at 8-daily temporal resolution and 500 m sampling and was resampled to 0.25° to match the resolution of VODCA2GPP.

4.2.2 FLUXCOM GPP

[1]FLUXCOM GPP (Tramontana et al. [2016]; Jung et al. [2020]) is produced by upscaling GPP estimates from in-situ eddy covariances by means of various machine learning algorithms. Two FLUXCOM GPP setups exist: FLUXCOM RS uses high resolution land surface properties from MODIS observations as machine learning model input while FLUXCOM RS + METEO uses the mean seasonal cycle of land surface variables and additionally incorporates meteorological data (Jung et al. [2020]). For validating VODCA2GPP, FLUXCOM RS was used because it includes temporal properties of land surface variables at finer spatial and temporal resolution than FLUXCOM RS+METEO. FLUXCOM RS GPP has 10 km sampling and is available every, 8 days, in accordance with the MODIS input data. The data were aggregated to 0.25° to match VODCA2GPP's resolution.

4.2.3 TRENDY GPP

[1]Additionally to remote sensing-based datasets, GPP estimates from the reanalysis-driven TRENDY-v7 ensemble of 16 dynamic global vegetation models (DGVMs) were used as independent reference dataset (Le Quéré et al. [2018]; Sitch et al. [2015]). TRENDY-v7 simulations consider forcing effects of climate, land use and changes in atmospheric CO2 concentrations on ecosystem's productivity over the period 1950-2017. The TRENDY-v7 ensemble consists of the following global vegetation models: CABLE-POP, CLASS-CTEM, CLM5.0, DLEM, ISAM, JSBACH, JULES, LPJ, LPJ-GUESS, LPX, OCN, ORCHIDEE, ORCHIDEE-CNP, SDGVM, SURFEX and VISIT. All models simulate monthly GPP and have a 1°x1° grid. For the comparison with VODCA2GPP, all models were regridded to 0.25° using nearest neighbour resampling and the ensemble

mean GPP of all available TRENDY-v7 models was computed.

5 Methods

5.1 The VOD2GPP-model

[1]VODCA2GPP is based on the carbon-sink driven GPP estimation approach (the VOD2GPP-model) introduced by Teubner et al. [2019],[2021]. The VOD2GPP-model describes the theoretical relationship between GPP and VOD. The biogeochemical basis of this model is the relationship between GPP, the plant's net uptake of carbon (NPP) and autotrophic respiration (Ra) (Bonan [2015]):

$$GPP = R_a + NPP \tag{5.1}$$

where Ra can be again split into two terms: maintenance respiration and growth respiration (Bonan [2015]). The VOD2GPP-model makes use of several VOD variables to represent the sum of NPP and Ra: the original VOD time series (VOD) which relates to maintenance respiration, temporal changes in VOD ($\Delta(VOD)$) which relate to both growth respiration and NPP and the temporal median of VOD ($mdn(VOD)$) derived from the complete time series which serves as a proxy for the landcover The incorporation of $mdn(VOD)$ helps the model to adapt for different vegetation types without explicitly using additional ancillary data (Teubner et al. [2019]). NPP is mostly represented by $\Delta(VOD)$ while Ra is represented by the original VOD signal and $\Delta(VOD)$. Thus, the VOD-based-only VOD2GPP-model can be formulated as follows (Teubner et al. [2019]):

$$GPP(VOD) = s(VOD) + s(\Delta VOD) + s(mdnVOD) \tag{5.2}$$

where $s()$ denotes the mapping function that maps the input variables to GPP.

Eq. 5.2 represents a simplified model formulation that connects VOD to GPP but does not explicitly take into account the strong temperature dependency of aboveground vegetation's autotrophic respiration (Teubner et al. [2021]; Wythers et al. [2013]; Atkin et al. [2005]; Tjoelker et al. [2001]) which is mainly attributed to its maintenance part (Bonan [2015]; Ryan et al. [1997]) . Therefore, an enhanced formulation of the model was developed by taking into account the temperature dependency of autotrophic main-

tenance respiration through a term representing the interaction between temperature (T2m) and VOD (Teubner et al. [2021]):

$$GPP(VOD, T2M) = te(VOD, T2M) + s(\Delta VOD) + s(mdnVOD) \qquad (5.3)$$

The mapping and interaction functions were implemented using generalized additive models (GAMs).

5.2 Generalized Additive Models

[1]Generalized Additive Models (GAMs; Hastie and Tibshirani [1990]) are semi-parametric generalizations of linear models and combine properties of Generalized Linear Models (GLM) and additive models (Guisan et al., 2002). Link functions $f()$ are trained and summed up for each predictor in order to relate the expected value of a response variable $E(Y)$ to the explanatory variables x_i (Hastie and Tibshirani [1990]). The model can be written as:

$$E(Y) = \beta + \sum_{i=1}^{n} f_i(x_i) \qquad (5.4)$$

where β denotes a constant offset and n corresponds to the number of input predictor variables x_i. The link functions $s_i()$ are implemented as smooth spline functions and allow the representation of non-monotonic and non-linear relationships which give them a high degree of flexibility (Hastie and Tibshirani [1990]). Hence, the relationship between target and predictor variables does not require explicit a-priori knowledge but can be estimated from the data itself, which makes GAMs appropriate for the VOD2GPP-model for which the exact relationship between VOD, air temperature, and GPP is unknown (Teubner et al. [2019]).

5.3 Preprocessing

[1]The model-input data (FLUXNET GPP (response variable), VODCA v2 CXKu, ERA5-Land (predictor variables)) was resampled from daily to 8-daily resolution using the 8-day means over the respective time period in order to reduce noise and computation times. This means that also the final VODCA2GPP represents the mean of daily GPP for an 8-day period with an estimate every 8 days. Since VODCA v2 CXKu incorporates already extensive quality flagging (e.g., for temperature) no additional data cleansing was necessary.

For the computation of $\Delta(VOD)$ the resampled VOD observations were smoothed in order to increase the robustness of the derivation (Teubner et al. [2019]). The smoothing was performed using a Savitzky-Golay filter with a window size of 11 data points as suggested by Teubner et al. (2021). $\Delta(VOD)$ was then obtained by subtracting the VOD observations at two consecutive timesteps. Median VOD ($mdn(VOD)$) was derived by computing the temporal median VOD of the entire time series for each available pixel.

5.4 Model input and training

[1]For each valid FLUXNET2015 in-situ observation, the corresponding overlapping pixel values of VOD, $\Delta(VOD)$, $mdn(VOD)$ and T2m were gathered and used to set up the GAM. Data from days with one or more invalid or missing observations were not considered for model training. For training of the final VODCA2GPP model no data was retained as test data. The GAM-based implementation of the VOD2GPP-model is consistent with Teubner et al. [2021] and utilizes algorithms from the pygam python package (Servén et al. [2018]).

The trained VODCA2GPP model was applied on each pixel where all input variables (VOD, $\Delta(VOD)$, $mdn(VOD)$ and T2m) were available. The result of this upscaling process is VODCA2GPP which covers the period between January 1988 and July 2020. Its spatial resolution is 0.25° and its temporal resolution is 8 days.

It is to be noted that with this approach it is in extreme cases also possible to obtain negative GPP values. Since negative GPP is not possible per definition, negative GPP estimates were set to zero in VODCA2GPP.

5.5 Uncertainty analysis

[1]The robustness of the model was evaluated based on a uncertainty analysis during which the influence of the in-situ GPP station selection on the model was investigated. Specifically, the uncertainty was analysed by training 10 VODCA2GPP models with the proposed GAM approach while for each model run 10% of the station data was (pseudo-)randomly retained (Teubner et al. [2019]). This means that each of the 10 models was trained with 90% of the available FLUXNET stations. Every station was excluded exactly once which is why this approach is classified pseudo-randomly.

5.6 Product evaluation

[1]For model evaluation, the 10 different model outputs from the uncertainty analysis were investigated by computing the minimum/maximum range as well as the standard deviation of the resulting 10 mean annual accumulated GPP estimates for each pixel. The standard deviation is also incorporated as uncertainty map in the available dataset (layer name: 'Uncertainties') to support users with an indicator for known uncertainties in VODCA2GPP.

Furthermore, annual GPP from VODCA2GPP, MODIS GPP and FLUXCOM GPP were evaluated against annual GPP from FLUXNET. The used error metrics were Root Mean Square Errors (RMSE) and Pearson's r. Global spatial GPP patterns were compared between the products by computing the mean annual GPP and the differences in mean annual GPP over the common observation period. Temporal agreements were tested by means of a Pearson correlation analysis for 8-daily GPP.

A correlation analysis of monthly GPP anomalies was conducted for all available datasets. Anomalies were derived by subtracting the long-term mean of the overlapping observation periods from 8-daily GPP estimates for each product.

Additionally, a trend analysis was conducted for all available GPP products in order to compare long-term changes in GPP. Trends in yearly median GPP were quantified using the Theil-Sen estimator (Theil [1950]; Sen [1968]) which calculates the slopes for each line between two point pairs. The median of all computed slopes is then used for line-fitting making it very insensitive to outliers and far more robust than simple linear regression (Wilcox [2010]). Slopes were considered as significant when then signs of the lower and upper 90%-confidence intervals were equal. For the trend analysis yearly median GPP was used.

6 Results

6.1 Spatial patterns in global annual GPP

[1]The average annual GPP of VODCA2GPP exhibits similar spatial patterns as the remote sensing-based references MODIS and FLUXCOM (Fig 6.1). The agreement in annual GPP is high in Northern latitudes (e.g., Europe, Russia, Canada) while there are relatively large differences in the southern hemisphere, especially in tropical regions (Fig 6.1d). The highest positive bias is observed in the subtropics. Very arid regions (e.g., Australian deserts, Kalahari Desert, etc.) have low mean yearly productivity in all three datasets (Fig 6.1a) but tend to be higher in VODCA2GPP compared to MODIS and FLUXCOM (Fig 6.1b, c). Comparison of the latitudinal distribution of FLUXNET station shows that lowest differences in yearly GPP are generally found in regions with high density of FLUXNET in-situ stations while largest bias is observed in regions with little or no FLUXNET coverage.

Similarly to the bias, uncertainties have a tendency to be smaller in latitudes with a high density of FLUXNET station (Fig 6.1d). The uncertainty map (Fig 6.2) shows that arid regions (e.g., Sahara, Australian deserts, Arabian Peninsula) as well as various mountainous regions (e.g., Carpathians, Alps, Rocky Mountains, Andes) have the highest model uncertainties (Fig 6.2). Moderate to high model instabilities are also exhibited for the tropics. Furthermore, significant uncertainties in VODCA2GPP are found in parts of Eastern and Western Siberia's boreal forests as well as in parts of Southern China.
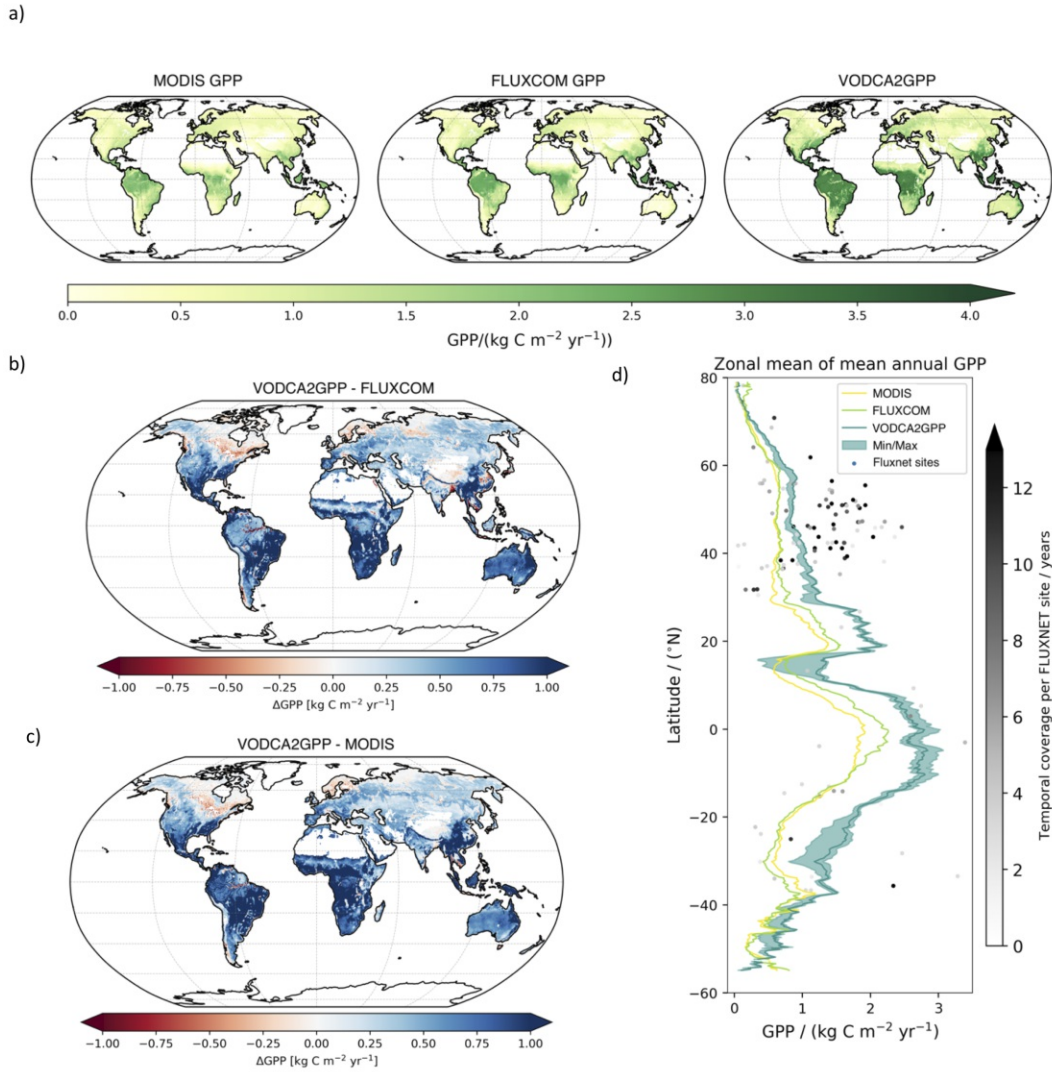
Figure 6.1: Figure 1: a) Mean yearly aggregated GPP for the common observation period of the three products (2002- 2016); b) and c) Difference in mean annual GPP between VODCA2GPP and the reference datasets; d) Latitude plot of zonal means of mean annual accumulated GPP. The means were computed based on 8-daily, 0.25 degree sampling. The Min/Max area denotes the minimum/maximum latitudinal mean for the ten model runs which were obtained during the uncertainty analysis. The dots represent the latitudinal location of the FLUXNET sites and the corresponding mean annual GPP. The brightness of the dots indicates the data availability for the respective FLUXNET station. Only data that is available in all three datasets was used for these plots.
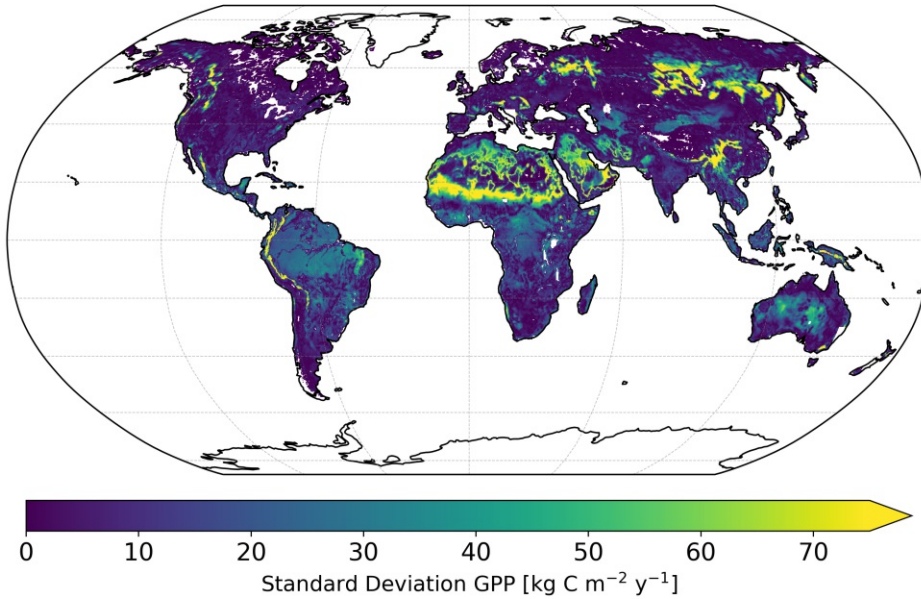
Figure 6.2: Standard deviation of mean yearly annual GPP (1988-2019) as obtained by
the uncertainty analysis

VODCA2GPP's tendency towards a positive bias with respect to MODIS and FLUX-COM GPP products is not mirrored in the comparison against FLUXNET GPP. The in-situ comparison for MODIS and FLUXCOM GPP suggests a slight but systematic underestimation of GPP across all biomes (Fig 6.3). VODCA2GPP has a higher RMSE and lower Pearson's r than the optical remote sensing based products indicating an overall slightly weaker performance. A landcover-based analysis (Fig A8.2) shows that uncertainties in annual VODCA2GPP are mostly occurring in (semi-)arid environments (e.g., savannas, open shrublands, grasslands). VODCA2GPP performs best in temperate environments (e.g., wetlands, evergreen broadleaf forest, croplands). Wetlands and evergreen broadleaf forests generally exhibit the best performance for all products while all three datasets underperform in open shrublands and deciduous broadleaf forest.
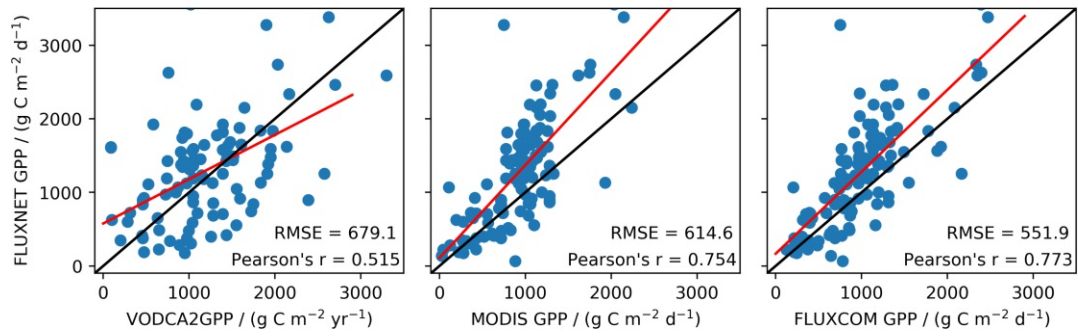
Figure 6.3: Mean annual in-situ GPP (FLUXNET) plotted against mean annual GPP from VODCA2GPP, FLUXCOM and MODIS for the respective grid cells. Mean annual GPP was computed from all available overlapping years the and thus each station is represented by one dot. Red lines indicate the best linear fits determined by ordinary linear regression and the black lines represent the 1:1 lines.

6.2 Comparison of temporal dynamics

[1]VODCA2GPP shows good temporal agreement with MODIS and FLUXCOM GPP. Pearson's r is highest in regions with distinct interannual variability such as sub-arctic, temperate, and semi-arid regions and lowest for dense tropical forests where partly even negative correlations occur (Fig 6.4). Median Pearson's r reaches 0.77 for MODIS GPP and 0.75 for FLUXCOM GPP.
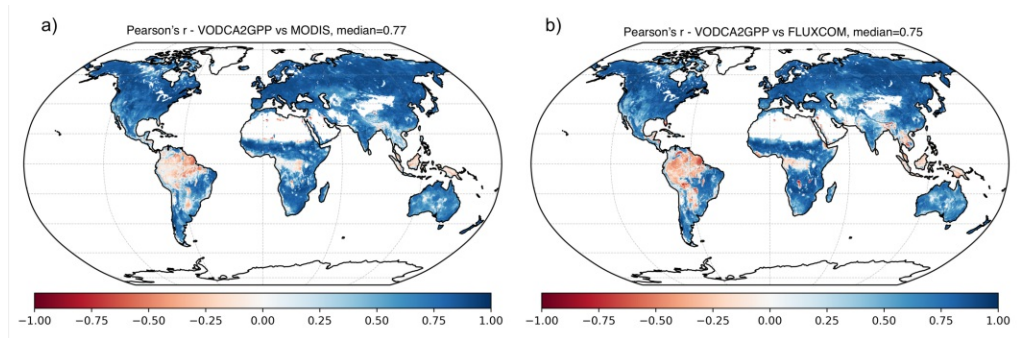
Figure 6.4: Pearson's r between VODCA2GPP and MODIS GPP (a) and VODCA2GPP and FLUXCOM GPP (b). The correlations are based on the common observation periods between 2002 and 2016 with 0.25°C spatial and 8-daily temporal resolution.

6.3 Anomaly patterns in space and time

[1]For a more detailed comparison of the products, monthly GPP anomalies were compared against each other. VODCA2GPP shows good correlation with MODIS and TRENDY-v7 GPP and weaker correlation with FLUXCOM GPP (Table 6.1, Fig 6.5). TRENDY-v7 correlates similarly well with VODCA2GPP and MODIS GPP and also shows worse correspondence with FLUXCOM GPP. The highest correlation is found between the two optical remote sensing based products MODIS and FLUXCOM GPP.

Table 6.1: Pearson's r correlation matrix for mean global monthly GPP anomalies between 2003 and 2016.

	VODCA2GPP	MODIS	FLUXCOM	TRENDYv7
VODCA2GPP	1			
MODIS	0.53	1		
FLUXCOM	0.29	0.69	1	
TRENDYv7	0.61	0.60	0.26	1
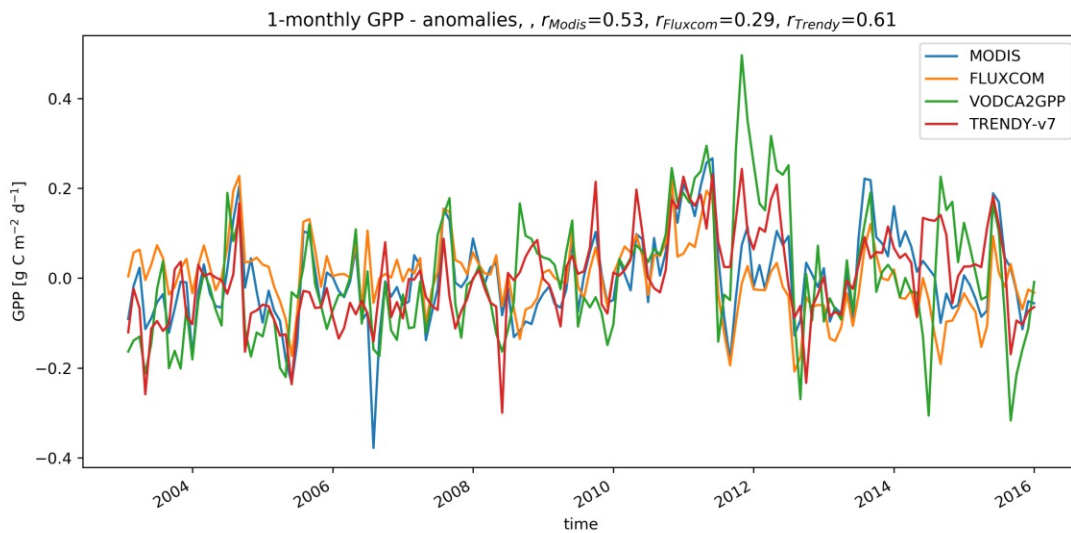
Figure 6.5: Time-series of mean global monthly GPP anomalies.

Analysis of the anomalies' temporal evolution and spatial distribution exhibits various similar patterns (Fig A8.3). Several extreme events are captured in VODCA2GPP and in at least one of the other GPP datasets. An example of such GPP extremes are the strong positive anomalies between 2010 and 2011 at around 25°CS which were mainly caused by record-breaking rainfalls in Australia (Wardle et al. [2013]). These positive

anomalies are clearly visible in all examined GPP products apart from FLUXCOM which does not exhibit similarly striking patterns (Fig A8.3). Other GPP extremes that are noticeable in all products apart form FLUXCOM is extremely low GPP in 2002/2003 and early 2005 around 20°CS (Fig A8.3). Both anomalies can be explained by extreme drought events that occurred in these years (Bureau of Meteorology, [2002],[2003],[2005]; Horridge et al. [2005]) which can be associated with El Niño events (Taschetto and England [2009]). Also, the distinct drop in GPP in 2015/2016 in similar latitudes might be linked to El Niño related drought events (Zhai et al. [2016]). Generally, extreme events in VODCA2GPP are more pronounced than in the other datasets.

6.4 Global GPP trends

[1]Trends in global annual median GPP between 2002 and 2016 are similar for VODCA2GPP, TRENDY-v7 GPP and MODIS GPP. All three products detect a significant increase of GPP during the observation period (Table 6.2, Fig 6.6). FLUXCOM GPP does not exhibit a significant trend but rather suggests a slight decrease in global annual median GPP.

The spatial distribution of GPP trends for the period 2003-2015 (Fig 6.7) exhibits many similarities between all analysed products. Large patterns of strong positive trends are, for example, found in Eastern parts of Siberia and China as well as in India and North America. Patterns of negative trends are found north of the Caspian Sea for all datasets. The remote sensing based products exhibit distinct patterns of declining GPP in central Siberia and significantly elevating GPP in Western Russia. Generally, the trends of VODCA2GPP match better with MODIS GPP and Trendy-v7 than with FLUXCOM. While there are many similar patterns on the Northern Hemisphere, trends on the Southern Hemisphere do not match well or are even contradictory. Especially in the Tropics hardly any similarities are visible.

For the full-time period (1988-2019) VODCA2GPP increases slightly on a global scale (Table 6.2), but this cannot be classified as significant due to contradictory upper and lower confidence intervals. The same is true for the slightly shorter period between 1988 and 2016 during which Trendy-v7 does detect a subtle significant trend on a global scale.

The spatial distribution of long-term (1988-2019; 6.8) trends in VODCA2GPP is similar to the shorter period (2003-2015) but in general, long-term VODCA2GPP trends are

less pronounced . The comparison of the fully overlapping period between VODCA2GPP and TRENDY-v7 (1988-2016, Fig A8.4) shows that TRENDY-v7 GPP exhibits weak but consistent positive trends for practically all biomes while VODCA2GPP trends are spatially strongly differing and for some regions even contradicting TRENDY-v7 trends.

A comprehensive comparison with in-situ GPP trends is hampered because most FLUXNET time series are too short to derive reliable trends. However, trends that were derived from available long-term time series (Fig A8.5) also suggest increasing GPP. Since this analysis is only based on very few stations this cannot be seen as evidence for positive GPP trends on a global scale.
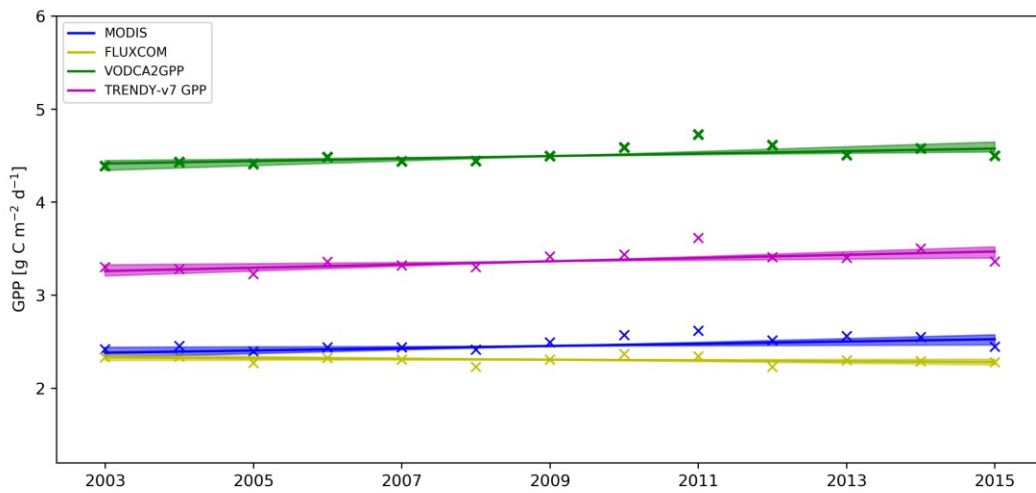
Figure 6.6: Time-series of yearly median GPP with the regression lines as obtained by the Theil-Sen estimator. Areas around the regression lines indicate the 90%-confidence intervals.

Table 6.2: Theil-Sen trends in global yearly median GPP. Same signs of the upper/lower 90%-confidence interval indicate significant trends. The analysed periods are 2003-2015 which corresponds to the fully overlapping period for all datasets, 1989-2016 which corresponds to the fully overlapping period of VODCA2GPP and TRENDY-v7 and 1988-2019 which corresponds to all fully available years of VODCA2GPP.

Product	2003-2015		2003-2015		1988-2019	
	Theil-Sen slope	lower/upper confidence interval	Theil-Sen slope	lower/upper confidence interval	Theil-Sen slope	lower/upper confidence interval
VODCA2GPP	0.013	+0.008 / +0.025	0.002	-0.001 / +0.006	0.002	-0.001 / +0.005
TRENDY-v7 GPP	0.017	+0.006 / +0.026	0.004	+0.000 / +0.008	-	-
MODIS GPP	0.012	+0.002 / +0.020	-	-	-	-
FLUXCOM GPP	-0.004	-0.009 / +0.001	-	-	-	-
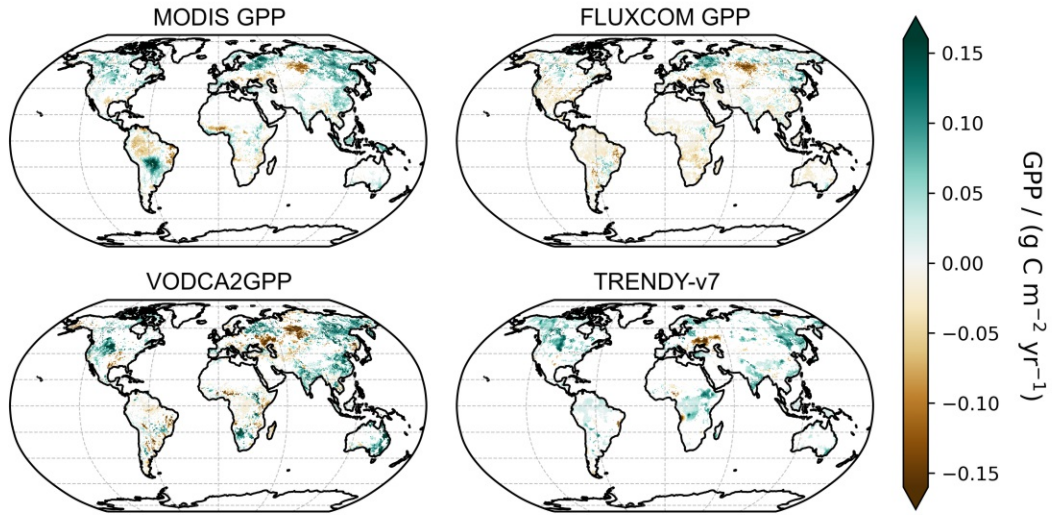
Figure 6.7: Global map of trends in yearly median GPP for the period 2003-2015 for all analysed datasets. White indicates non-significant trends.
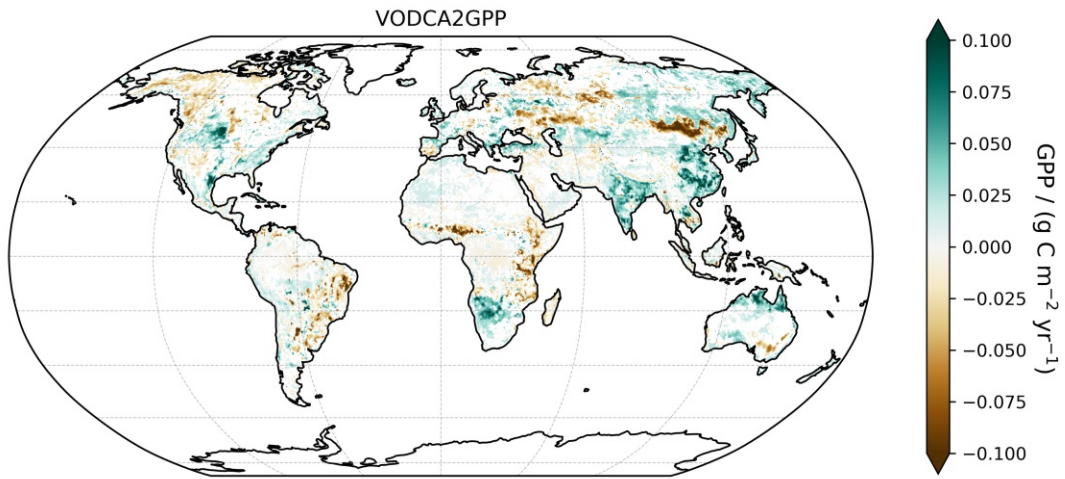
Figure 6.8: Global map of yearly median GPP trends for the period 1988-2019 for VODCA2GPP. White indicates non-significant trends.

7 Discussion

7.1 Uncertainties in the VODCA2GPP model

[1]The results from the uncertainty analysis and the comparison with in-situ GPP show that VODCA2GPP estimates can be viewed as very reliable across most biomes. However, substantial uncertainties were exhibited through the uncertainty analysis in some areas with extreme climatic or topographic conditions (e.g.; deserts and mountain ranges). Also, parts of Eastern and Western Siberia and parts of Southern China show a relatively large spreads in the predictions. The observed uncertainty patterns in Siberia might be associated with topography, generally lower data availability (due masking of frozen conditions in VODCA) and a lack of FLUXNET stations. Missing in-situ information might also be a driver for uncertainties in Southern China but the exact cause for the relatively high uncertainty there is unclear. The uncertainty analysis suggests that VODCA2GPP estimates have a tendency to be too high in these regions and thus should be interpreted with caution. Furthermore, moderate uncertainties were also found for the tropics which is likely due to the extremely low in-situ data availability and higher absolute GPP than in mid-latitudes.

The comparison with in-situ GPP shows clear differences in performance of the VODCA2GPP model across different biomes. High performance is achieved in densely vegetated biomes while the performance decreases in arid and less-vegetated regions. A reason for the weaker performance in areas with less water availability might be adapted water regulation strategies of plants. Plants in drought-prone regions often reduce transpiration by limiting stomatal conductance in order to maintain a constant water potential even in times of extreme water scarcity (Sade et al. [2012]). Since VOD is largely driven by the vegetation water content, this isohydric behaviour of vegetation could at least partly explain relatively high VOD and consequently also overestimated GPP in these regions (Teubner et al. [2021]).

Also, the observation bias which is introduced by unevenly distributed FLUXNET sites decreases the model's robustness. GPP is measured in situ only for few locations where in-situ carbon fluxes are measured and these stations are mostly located in temperate

regions (e.g., Europe and North America) while semi-arid and tropical forest regions are underrepresented in the training data.

7.2 Independence of reference datasets

[1]The lack of in-situ GPP is not only problematic in model training but also hampers a fair evaluation and validation on a global scale. Thus, validation in regions with low FLUXNET station density could only be done via global reference datasets. Remote sensing-based references (MODIS and FLUXCOM), however, are also calibrated or trained using in-situ GPP observations (Jung et al. [2020]; Running et al. [1999]) and can therefore not be viewed as fully independent from VODCA2GPP (Teubner et al. [2021]). In contrast to observation-based GPP products, estimates from the TRENDY ensemble can be considered largely independent from VODCA2GPP which makes it a very valuable additional reference for VODCA2GPP.

7.3 Limitations in VODCA and their impact on VODCA2GPP

[1]As outlined in Moesinger et al. [2020] there are certain limitations in the VODCA v1 product which are partly also evident in VODCA v2 (Zotta et al. [in preperation]) and thus also propagate to VODCA2GPP. A known issue of VODCA v2 is caused by an observation gap between October 2011 and July 2012 for AMSR-E and AMSR2 (Table 4.1), which prevents a direct bias removal between the sensors. However, scaling between the sensors is achieved by using TMI observations North/South of 35°N/35S for X and Ku-band. Beyond these latitudes for X- and Ku-band and generally globally for C-band, AMSR-E data from 2010-2012 was matched against AMSR2 data from 2012-2014 under the assumption that trends between 2010-2014 are negligible (Moesinger et al. [2020]). The result is that AMSR2 observations exhibit a slight positive bias in parts of North America which is also evident in a spatial break in VODCA v1 X- and Ku-Band trends (Moesinger et al. [2020]). Although the impact of this procedure on VODCA2GPP trends is small and spatially limited, users are advised to keep the potential bias in mind when analysing VODCA2GPP data after 2012 for latitudes North/South of 35°N/35°S. Other limitations in VODCA concern the mixing of observations that were retrieved at different geometries (e.g., incidence angles) or observation times (Moesinger et al. [2020]) and the data loss in certain regions, mostly in the Himalayas, which is caused by failure of the CDF-matching method due to insufficient input data (Moesinger et al. [2020]). These issues, however, only have a small or spatially very limited influence on the final VODCA2GPP product.

7.4 Observed bias between VODCA2GPP and other remote sensing based GPP data

[1]VODCA2GPP compares well to the analysed state-of-the-art remote sensing-based products FLUXCOM and MODIS in terms of spatio-temporal patterns. However, in comparison to the GPP datasets driven by optical remote sensing data, VODCA2GPP exhibits higher GPP across most biomes. This positive bias can be partly explained by a reported and observed tendency of FLUXCOM and MODIS to underestimate GPP which is especially pronounced in tropical regions (Turner et al. [2006]; Wang et al. [2017]; Fig 6.1; Fig A8.2). Furthermore, VODCA2GPP is in better alignment with the average station data (Fig 6.1d). Discrepancies in absolute GPP among the products outside the tropics might also be caused by assumed overestimation of VODCA2GPP in winter months (i.e., in times with very little or no primary productivity). This overestimation is explicable with the water content in vegetation that is also present in these dormant periods. The sensitivity of microwaves to this water content results in non-zero VOD and, consequently, non-zero GPP (Teubner et al. [2021]).

Another potential explanation for overestimation in VODCA2GPP is the presence of surface water and its impact on VOD retrievals. The presence of surface water is known to decrease the brightness temperature of the earth's surface and thus significantly decreases VOD retrievals (Bousquet et al. [2021]). The impact of surface water contamination is evident in VODCA pixels that partly contain water bodies (e.g., lakes, rivers). These pixels exhibit systematically lower values than neighbouring pixels without water share. This is on the one hand problematic for the output of the VODCA2GPP model where GPP is presumably systematically underestimated in pixels containing surface water. On the other hand, it also has a non-negligible effect on model training. This effect is caused by FLUXNET stations located close to water bodies, which hardly impact in-situ GPP retrievals but do cause erroneous VOD-retrievals. As a result, underestimated VOD is trained against unaffected in-situ GPP which causes a slight but systematic global overestimation. A potential solution would be the masking of water-contaminated VOD. However, due to the constraints with temperature in the interaction term (eq. 5.3) this is a non-trivial task and would require a reformulation of the VODCA2GPP model.

7.5 Potential factors for observed trends in VODCA2GPP

[1]There are several potential drivers for long-term increases in GPP, the most important ones being global warming, land-use changes and elevating CO_2 concentrations in the atmosphere (Piao et al. [2013]). VODCA2GPP does not show a clear global increase trend in microwave-derived GPP for the period 1988-2019 but does so for the period 2003-2015. MODIS GPP also exhibits a slight increase in median global GPP for 2003-2015 but at a slower rate. The strongest increase in global GPP was found for the TRENDY-v7 ensemble for both analysed periods. The observed long-term trends in GPP across the different products supports the theory of elevated atmospheric CO_2 leading to an increased uptake of CO_2 (Haverd et al. [2020a]; Walker et al. [2020]; Campbell et al. [2017]; Schimel et al. [2015]). The absence of trends in FLUXCOM is not contradicting this theory as FLUXCOM does not account for CO_2 fertilization effects (Jung et al. [2020]). However, an in-situ based trend analysis did not provide enough evidence to prove this hypothesis. Therefore, the global influence of atmospheric CO_2 on vegetation productivity remains uncertain but VODCA2GPP advocates a positive effect of rising CO_2 on GPP.

7.6 Complementarity of VODCA2GPP to existing products

[1]The comparison of spatio-temporal patterns showed that there are many similarities between VODCA2GPP and the reference datasets. The analysis of monthly anomalies exhibited various extreme events in VODCA2GPP that also are found in one or several existing products indicating high plausibility of VODCA2GPP derived anomalies. Furthermore, trends derived from VODCA2GPP contain several plausible patterns that match those derived from the TRENDY-v7 simulations but are not visible in MODIS and FLUXCOM and vice versa. This suggests that VODCA2GPP has the potential to bridge the gaps between existing observation-based and process-based state-of-the art products and can be used complementary to them.

8 Conclusions and outlook

[1]In this thesis VODCA2GPP, a long-term GPP data record which uses multi-frequency microwave VODCA and temperature data from ERA5-Land for the upscaling of in-situ GPP from FLUXNET2015 was introduced. The comparison of VODCA2GPP with FLUXNET in-situ GPP and global state-of-the art GPP datasets showed good correspondence between the products in both the spatial and temporal domain, but with varying performance differences across the biomes. In tropical and arid regions VODCA2GPP has significantly higher values than the reference datasets. Arid and mountainous areas were found to have the largest uncertainties. The observed anomalies and trends in VODCA2GPP overall match with anomalies and trends derived from MODIS and Trendy-v7 which is another argument for the validity of VODCA2GPP as well as for increased $CO2$ uptake by plants induced by increasing atmospheric $CO2$ concentrations. Overall, the analysis show that the novel microwave remote sensing based VODCA2GPP product is able to complement existing GPP data records.

In this thesis several potential improvements of the VODCA2GPP dataset have been identified and might be included in future versions of VODCA2GPP. An important step towards bias reduction of the VOD2GPP-model would be the accounting for water contamination in VOD pixels. Other enhancements concern updates of the underlying VODCA data which will likely include new missions and potentially further refined merging methods. Potentially updated versions of VODCA2GPP will be made available to the public as soon as possible. Therefore, interested users are asked to follow the developments at `https://doi.org/10.48436/1k7aj-bdz35`.

Appendix

Table 8.1: Overview of used FLUXNET2015 sites

ID	Name	Lon [°E]	Lat [°N]	Years used
AR-SLu	San Luis	-66.46	-33.46	2009-2011
AR-Vir	Virasoro	-56.19	-28.24	2010-2012
AT-Neu	Neustift	11.32	47.12	2002-2012
AU-ASM	Alice Springs	133.25	-22.28	2010-2013
AU-Ade	Adelaide River	131.12	-13.08	2007-2009
AU-Cpr	Calperum	140.59	-34	2010-2013
AU-Cum	Cumberland Plains	150.72	-33.61	2012-2013
AU-DaP	Daly River Savanna	131.32	-14.06	2008-2013
AU-DaS	Daly River Cleared	131.39	-14.16	2008-2013
AU-Dry	Dry River	132.37	-15.26	2008-2013
AU-Emr	Emerald, Queensland, Australia	148.47	-23.86	2011-2013
AU-Fog	Fogg Dam	131.31	-12.55	2006-2008
AU-GWW	Great Western Woodlands, Wester Australia, Australia	120.65	-30.19	2013-2014
AU-RDF	Red Dirt Melon Farm, Northern Territory	132.48	-14.56	2011-2013
AU-Rig	Riggs Creek	145.58	-36.65	2011-2013
AU-Rob	Robson Creek, Queensland, Australia	145.63	-17.12	2014
AU-Tum	Tumbarumba	148.15	-35.66	2001-2013
AU-Whr	Whroo	145.03	-36.67	2011-2013
BE-Bra	Brasschaat	4.52	51.31	2004-2013
BE-Lon	Lonzee	4.75	50.55	2004-2014
BE-Vie	Vielsalm	6	50.31	1996-2014
BR-Sa3	Santarem-Km83-Logged Forest	-54.97	-3.02	2000-2004
CA-NS1	UCI-1850 burn site	-98.48	55.88	2001-2005
CA-NS3	UCI-1964 burn site	-98.38	55.91	2001-2005
CA-NS4	UCI-1964 burn site wet	-98.38	55.91	2002-2005
CA-NS5	UCI-1981 burn site	-98.49	55.86	2002-2005
CA-NS6	UCI-1989 burn site	-98.96	55.92	2001-2005

43

CA-NS7	UCI-1998 burn site	-99.95	56.64	2002-2005
CA-Qfo	Quebec – Eastern Boreal, Mature Black Spruce	-74.34	49.69	2003-2010
CA-SF1	Saskatchewan – Western Boreal, forest burned in 1977	-105.82	54.49	2003-2006
CA-SF2	Saskatchewan – Western Boreal, forest burned in 1989	-105.88	54.25	2001-2005
CA-SF3	Saskatchewan – Western Boreal, forest burned in 1998	-106.01	54.09	2001-2006
CH-Cha	Chamau	8.41	47.21	2006-2012
CH-Fru	Frübüel	8.54	47.12	2006-2012
CH-Oe1	Oensingen grassland	7.73	47.29	2002-2008
CN-Cha	Changbaishan	128.1	42.4	2004-2005
CN-Cng	Changling	123.51	44.59	2007-2010
CN-Dan	Dangxiong	91.07	30.5	2004-2005
CN-Din	Dinghushan	112.54	23.17	2003-2005
CN-Du2	Duolun-grassland (D01)	116.28	42.05	2006-2008
CN-Ha2	Haibei Shrubland	101.33	37.61	2003-2005
CN-HaM	Haibei Alpine Tibet site	101.18	37.37	2002-2004
CN-Qia	Qianyanzhou	115.06	26.74	2003-2005
CN-Sw2	Siziwang Grazed (SZWG)	111.9	41.79	2010-2012
CZ-BK1	Bily Kriz forest	18.54	49.5	2004-2008
CZ-BK2	Bily Kriz grassland	18.54	49.49	2004-2006
DE-Akn	Anklam	13.68	53.87	2009-2014
DE-Gri	Grillenburg	13.51	50.95	2004-2014
DE-Hai	Hainich	10.45	51.08	2000-2012
DE-Kli	Klingenberg	13.52	50.89	2004-2014
DE-Lkb	Lackenberg	13.3	49.1	2009-2013
DE-Obe	Oberbärenburg	13.72	50.78	2008-2014
DE-RuS	Selhausen Juelich	6.45	50.87	2011-2014
DE-Spw	Spreewald	14.03	51.89	2010-2014
DE-Tha	Tharandt	13.57	50-96	1996-2014
DK-NuF	Nuuk Fen	-51.39	64.13	2008-2014
DK-Sor	Soroe	11.64	55.49	1996-2012
ES-LgS	Laguna Seca	-2.97	37.1	2007-2009
ES-Ln2	Lanjaron-Salvage logging	-3.48	36.97	2009

FI-Hyy	Hyytiala	24.3	61.85	1996-2014
FI-Jok	Jokioinen	23.51	60.9	2000-2003
FR-Gri	Grignon	1.95	48.84	2004-2013
FR-Pue	Puechabon	3.6	43.74	2000-2013
GF-Guy	Guyaflux (French Guiana)	-52.92	5.28	2004-2012
IT-CA1	Castel d'Asso 1	12.03	42.38	2011-2013
IT-CA2	Castel d'Asso 2	12.03	42.38	2011-2013
IT-CA3	Castel d'Asso 3	12.02	42.38	2011-2013
IT-Cp2	Castelporziano 2	12.36	41.7	2012-2013
IT-Isp	Ispra ABC-IS	8.63	45.81	2013-2014
IT-Lav	Lavarone	11.28	45.96	2003-2012
IT-Noe	Arca di Noé – Le Prigionette	8.15	40.61	2004-2012
IT-PT1	Parco Ticino forest	9.06	45.2	2002-2004
IT-Ren	Renon	11.43	46.59	1998-2013
IT-Ro1	Roccarespampani 1	11.93	42.41	2000-2008
IT-Ro2	Roccarespampani 2	11.92	42.39	2003-2012
IT-SR2	San Rossore 2	10.29	43.73	2013-2014
IT-SRo	San Rossore	10.28	43.73	1999-2012
IT-Tor	Torgnon	7.58	45.84	2008-2013
JP-MBF	Moshiri Birch Forest Site	142.32	44.39	2003-2005
JP-SMF	Seto Mixed Forest Site	137.08	35.26	2002-2006
NL-Hor	Horstermeer	5.07	52.24	2004-2011
NL-Loo	Loobos	5.74	52.17	1996-2013
NO-Adv	Adventdalen	15.92	78.19	2012-2014
RU-Che	Cherski	161.34	68.61	2002-2005
RU-Cok	Chokurdakh	147.49	70.83	2003-2013
RU-Fyo	Fyodorovskoye	32.92	56.46	1998-2013
RU-Ha1	Hakasia steppe	90	54.73	2002-2004
SD-Dem	Demokeya	30.48	13.28	2005-2009
US-AR1	ARM USDA UNL OSU Woodward Switch-grass 1	-99.42	36.43	2009-2012
US-AR2	ARM USDA UNL OSU Woodward Switch-grass 2	-99.6	36.64	2009-2012
US-ARM	ARM Southern Great Plains site - Lamont	-97.49	36.61	2003-2012
US-Blo	Blodgett Forest	-120.63	38.9	1997-2007
US-Ha1	Harvard Forest EMS Tower (HFR 1)	-72.17	42.54	1991-2012

US-Los	Lost Creek	-89.98	46.08	2000-2014
US-MMS	Morgan Monroe State Forest	-86.41	39.32	1999-2014
US-Me6	Metolius Young Pine Burn	-121.61	44.32	2010-2012
US-Myb	Mayberry Wetland	-121.77	38.05	2011-2014
US-Ne1	Mead – irrigated continuous maize site	-96.48	41.17	2001-2013
US-Ne2	Mead – irrigated maize-soybean rotation site	-96.47	41.16	2001-2013
US-Ne3	Mead – rainfed maize-soybean roatation site	-96.44	41.18	2001-2013
US-SRM	Santa Rita Mesquite	-110.87	31.82	2004-2014
US-Syv	Sylvania Wilderness Area	-89.35	46.24	2001-2014
US-Ton	Tonzi Ranch	-120.97	38.43	2001-2014
US-Tw3	Twitchell Alfalfa	-121.65	38.12	2013-2014
US-UMd	UMBS Disturbance	-84.7	45.56	2007-2014
US-Var	Vaira Ranch-Ione	-120.95	38.41	2000-2014
US-WCr	Willow Creek	-90.08	45.81	1999-2014
US-Whs	Walnut Gulch Lucky Hills Shrub	-110.05	31.74	2007-2014
US-Wkg	Walnut Gulch Kendall Grasslands	-109.94	31.74	2004-2014
ZM-Mon	Mongu	23.25	-15.44	2007-2009
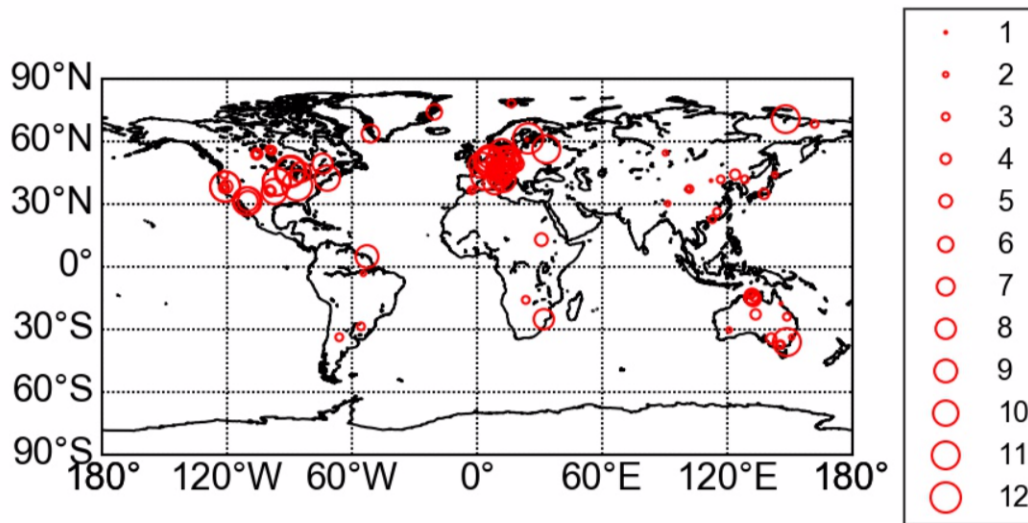
Figure 8.1: Spatial distribution of the FLUXNET2015 Tier v1 stations. The circle size indicates the observation length for each station (Teubner et al. [2021])
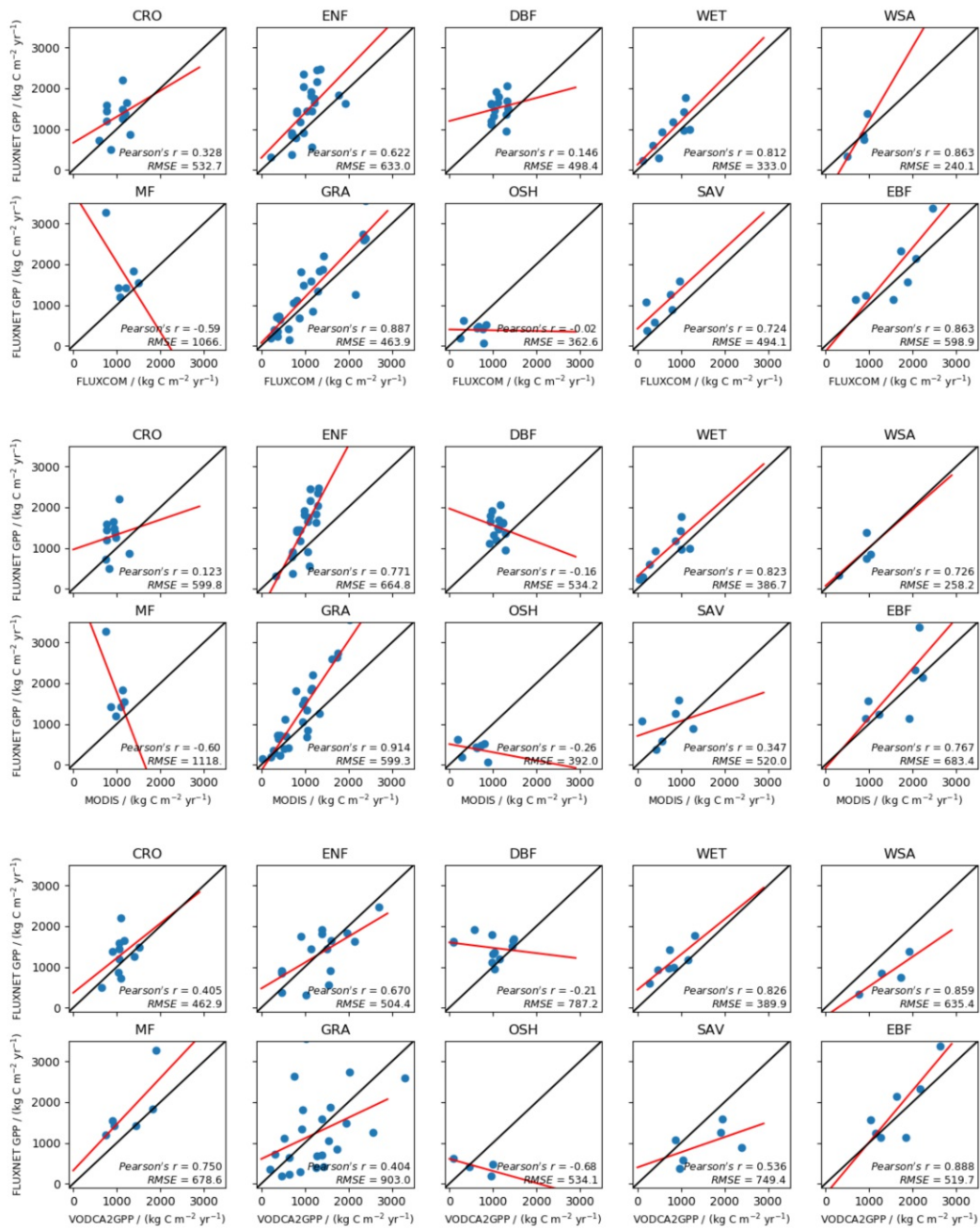
Figure 8.2: Scatterplots of mean annual GPP for the period 2002-2016 per vegetation type. Vegetation types indicate the pre-dominant IGBP-vegetation type at the respective FLUXNET station. Abbreviations: CRO: Croplands; ENF: Evergreen Needleleaf Forests; DBF: Deciduous Broadleaf Forests; WET: Permanent Wetlands; WSA: Woody Savannas; MF: Mixed Forests; GRA: Grasslands; OSH: Open Shrublands; SAV: Savannas; EBF: Evergreen Broadleaf Forests

47
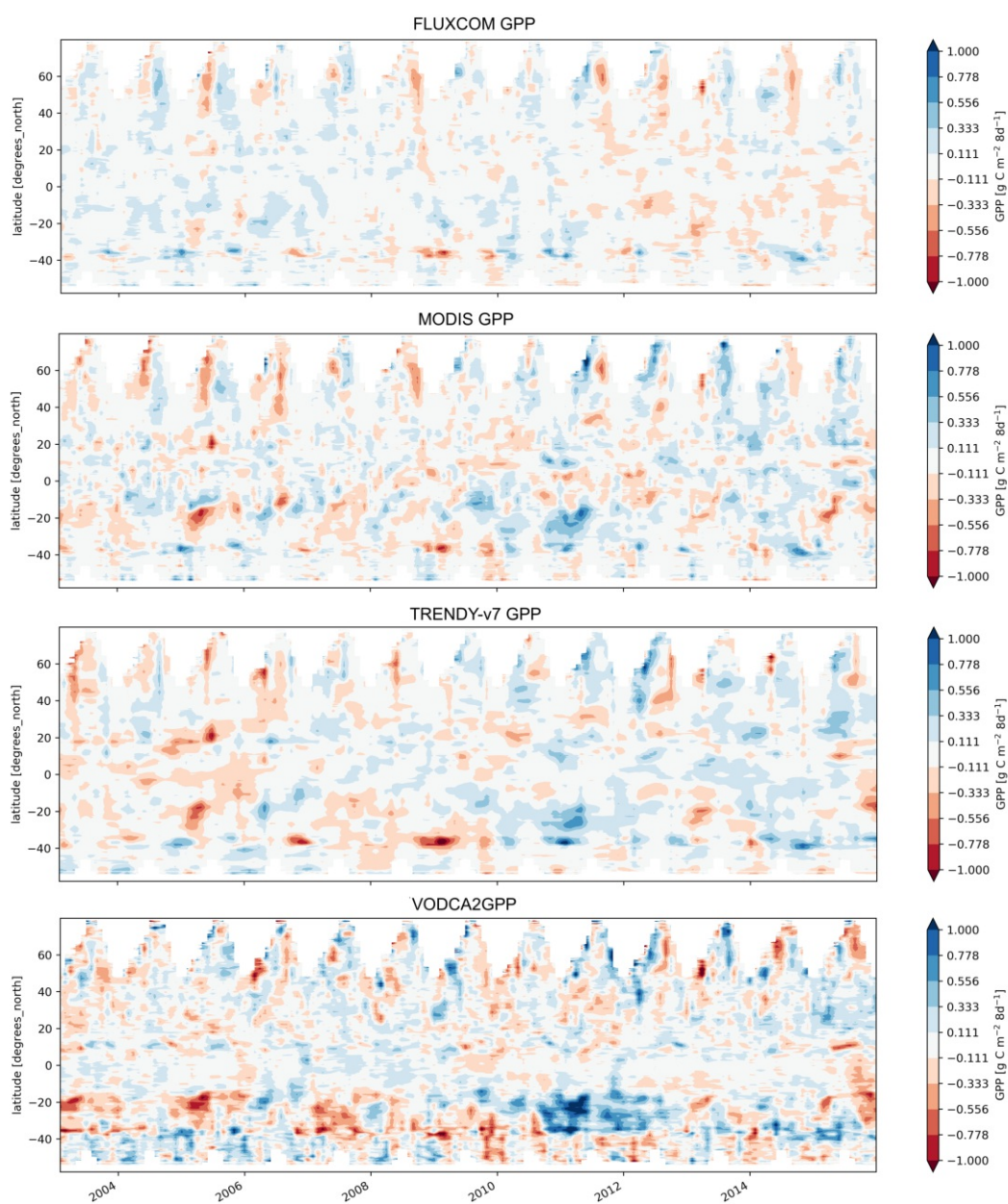
Figure 8.3: Hovmoeller diagrams of monthly GPP-anomalies for each dataset.
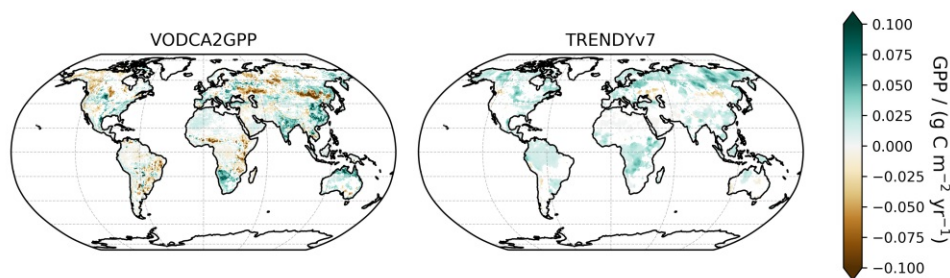
Figure 8.4: Global maps of yearly median GPP trends for the period 1988-2016 for VODCA2GPP and TRENDY-v7 GPP. White indicates non-significant trends.
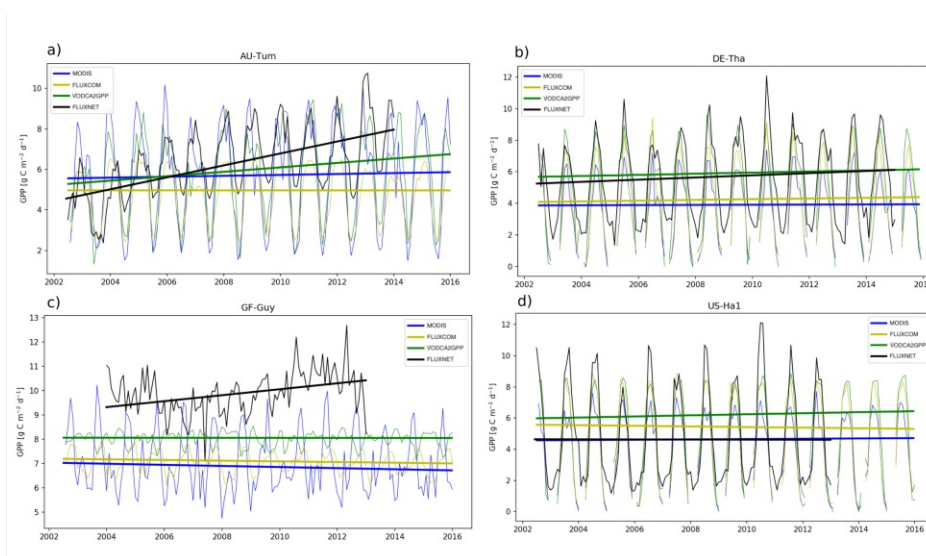
Figure 8.5: Exemplary collection of time series of in-situ FLUXNET GPP together with extracted time series from MODIS, FLUXCOM and VODCA2GPP. The stations were selected because of their high data availability for the respective landcover. The lines indicate the regression lines as obtained from the Theil-Sen slope estimation. The trends are computed for the common observation period with FLUXNET. The depicted stations are: a) AU-Tum: Tumbarumba, Australia; Lat: -35.65 °N, Lon: 148.15 °E; Landcover: EBF b) DE-Tha: Tharandt, Germany; Lat: 50.96 °N, Lon: 13.57 °E; Landcover: ENF c) GF-Guy: Guyaflux, French Guiana; Lat: 5.28 °N, Lon: -52.93 °E; Landcover: EBF d)US-Ha1: Harvard Forest EMS Tower, United States; Lat: 42.54 °N, Lon: -72.17 °E; Landcover: DBF

Bibliography

Alemohammad, S. H., Fang, B., Konings, A. G., Aires, F., Green, J. K., Kolassa, J., Miralles, D., Prigent, C., and Gentine, P. Water, energy, and carbon with artificial neural networks (wecann): a statistically based estimate of global surface turbulent fluxes and gross primary productivity using solar-induced fluorescence. *Biogeosciences*, 14(18):4101–4124, 2017.

Allen, M. R., de Coninck, H., Dube, O. P., Hoegh-Guldberg, O., Jacob, D., Jiang, K., Revi, A., Rogelj, J., Roy, J., Shindell, D., et al. Technical summary. In *Global warming of 1.5° C: An IPCC Special Report on the impacts of global warming of 1.5° C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty*, pages 27–46. Intergovernmental Panel on Climate Change, 2018.

Asner, G. P., Scurlock, J. M., and A. Hicke, J. Global synthesis of leaf area index observations: implications for ecological and remote sensing studies. *Global Ecology and Biogeography*, 12(3):191–205, 2003.

Atkin, O. K., Bruhn, D., Hurry, V. M., and Tjoelker, M. G. Evans review no. 2: The hot and the cold: unravelling the variable response of plant respiration to temperature. *Functional Plant Biology*, 32(2):87–105, 2005.

Attema, E. and Ulaby, F. T. Vegetation modeled as a water cloud. *Radio science*, 13 (2):357–364, 1978.

Baldocchi, D. D. Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. *Global change biology*, 9(4): 479–492, 2003.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., et al. Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science*, 329(5993):834–838, 2010.

Bonan, G. *Ecological climatology: concepts and applications.* Cambridge University Press, 2015.

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 Sensing of Environment*, 257:112345, 2021.

Cain, M. L., Bowman, W. D., Hacker, S. D., and Bowman, W. D. *Ecology (3rd ed.).* Sunderland, Mass.: Sinauer Associates, 2014.

Campbell, J., Berry, J., Seibt, U., Smith, S. J., Montzka, S., Launois, T., Belviso, S., Bopp, L., and Laine, M. Large historical growth in global terrestrial gross primary production. *Nature*, 544(7648):84–87, 2017.

Canadell, J. G., Le Quéré, C., Raupach, M. R., Field, C. B., Buitenhuis, E. T., Ciais, P., Conway, T. J., Gillett, N. P., Houghton, R., and Marland, G. Contributions to accelerating atmospheric co2 growth from economic activity, carbon intensity, and efficiency of natural sinks. *Proceedings of the national academy of sciences*, 104(47): 18866–18870, 2007.

Carlson, T. N. and Ripley, D. A. On the relation between ndvi, fractional vegetation cover, and leaf area index. *Remote sensing of Environment*, 62(3):241–252, 1997.

Carver, K. R., Elachi, C., and Ulaby, F. T. Microwave remote sensing from space. *Proceedings of the IEEE*, 73(6):970–996, 1985.

Chen, P.-Y., Fedosejevs, G., Tiscareno-Lopez, M., and Arnold, J. G. Assessment of modis-evi, modis-ndvi and vegetation-ndvi composite data using agricultural measurements: An example at corn fields in western mexico. *Environmental monitoring and assessment*, 119(1):69–82, 2006.

Clark, D. A., Brown, S., Kicklighter, D. W., Chambers, J. Q., Thomlinson, J. R., and Ni, J. Measuring net primary production in forests: concepts and field methods. *Ecological applications*, 11(2):356–370, 2001.

Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A., and Totterdell, I. J. Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, 408(6809):184–187, 2000.

de Nijs, A. H., Parinussa, R. M., de Jeu, R. A., Schellekens, J., and Holmes, T. R. A methodology to determine radio-frequency interference in amsr2 observations. *IEEE Transactions on Geoscience and Remote Sensing*, 53(9):5148–5159, 2015.

Drake, J. E., Tjoelker, M. G., Aspinwall, M. J., Reich, P. B., Barton, C. V., Medlyn, B. E., and Duursma, R. A. Does physiological acclimation to climate warming stabilize the ratio of canopy respiration to photosynthesis? *New Phytologist*, 211(3):850–863, 2016.

Frappart, F., Wigneron, J.-P., Li, X., Liu, X., Al-Yaari, A., Fan, L., Wang, M., Moisy, C., Le Masson, E., Lafkih, Z. A., et al. Global monitoring of the vegetation dynamics from the vegetation optical depth (vod): A review. *Remote Sensing*, 12(18):2915, 2020.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., et al. Global carbon budget 2020. *Earth System Science Data*, 12(4):3269–3340, 2020.

Gilabert, M. A., Sánchez-Ruiz, S., and Moreno, Á. Annual gross primary production from vegetation indices: A theoretically sound approach. *Remote Sensing*, 9(3):193, 2017.

Girardin, C. A. J., Malhi, Y., Aragao, L., Mamani, M., Huaraca Huasco, W., Durand, L., Feeley, K., Rapp, J., Silva-Espejo, J., Silman, M., et al. Net primary productivity allocation and cycling of carbon along a tropical forest elevational transect in the peruvian andes. *Global Change Biology*, 16(12):3176–3192, 2010.

Hastie, T. J. and Tibshirani, R. J. *Generalized additive models*, volume 43. CRC press, 1990.

Haverd, V., Smith, B., Canadell, J. G., Cuntz, M., Mikaloff-Fletcher, S., Farquhar, G., Woodgate, W., Briggs, P. R., and Trudinger, C. M. Higher than expected co2 fertilization inferred from leaf to global observations. *Global change biology*, 26(4): 2390–2402, 2020a.

Haverd, V., Smith, B., Canadell, J. G., Cuntz, M., Mikaloff-Fletcher, S., Farquhar, G., Woodgate, W., Briggs, P. R., and Trudinger, C. M. Higher than expected co2 fertilization inferred from leaf to global observations. *Global Change Biology*, 26(4):2390–2402, 2020b. doi: https://doi.org/10.1111/gcb.14950. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.14950.

Holmes, T., De Jeu, R., Owe, M., and Dolman, A. Land surface temperature from ka band (37 ghz) passive microwave observations. *Journal of Geophysical Research: Atmospheres*, 114(D4), 2009.

Horridge, M., Madden, J., and Wittwer, G. The impact of the 2002–2003 drought on australia. *Journal of Policy Modeling*, 27(3):285–308, 2005.

Jackson, T. and Schmugge, T. Vegetation effects on the microwave emission of soils. *Remote Sensing of Environment*, 36(3):203–212, 1991.

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., and Reichstein, M. The fluxcom ensemble of global land-atmosphere energy fluxes. *Scientific data*, 6(1):1–14, 2019.

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., Carvalhais, N., et al. Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the fluxcom approach. *Biogeosciences*, 17(5):1343–1365, 2020.

Kerekes, J. P. *Optical sensor technology*. SAGE Publications, 2009.

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 Phytologist*, 223(3):1166–1172, 2019. doi: https://doi.org/10.1111/nph.15808. URL `https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/nph.15808`.

Lavigne, M. and Ryan, M. Growth and maintenance respiration rates of aspen, black spruce and jack pine stems at northern and southern boreas sites. *Tree Physiology*, 17 (8-9):543–551, 1997.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., et al. Global carbon budget 2018. *Earth System Science Data*, 10(4):2141–2194, 2018.

Li, X., Wigneron, J.-P., Frappart, F., Fan, L., Ciais, P., Fensholt, R., Entekhabi, D., Brandt, M., Konings, A. G., Liu, X., et al. Global-scale assessment and intercomparison of recently developed/reprocessed microwave satellite vegetation optical depth products. *Remote Sensing of Environment*, 253:112208, 2021.

Liang, S., Li, X., and Wang, J. Chapter 16 - vegetation production in terrestrial ecosystems. In *Advanced Remote Sensing*, pages 501–531. Academic Press, Boston, 2012. ISBN 978-0-12-385954-9. doi: https://doi.org/10.1016/B978-0-12-385954-9.00016-2. URL `https://www.sciencedirect.com/science/article/pii/B9780123859549000162`.

Liu, Y. Y., De Jeu, R. A., McCabe, M. F., Evans, J. P., and Van Dijk, A. I. Global long-term passive microwave satellite-based retrievals of vegetation optical depth. *Geophysical Research Letters*, 38(18), 2011.

Luyssaert, S., Inglima, I., Jung, M., Richardson, A. D., Reichstein, M., Papale, D., Piao, S., Schulze, E.-D., Wingate, L., Matteucci, G., et al. Co2 balance of boreal, temperate, and tropical forests derived from a global database. *Global change biology*, 13(12):2509–2537, 2007.

Meesters, A. G., De Jeu, R. A., and Owe, M. Analytical derivation of the vegetation optical depth from the microwave polarization difference index. *IEEE Geoscience and Remote Sensing Letters*, 2(2):121–123, 2005.

Mo, T., Choudhury, B., Schmugge, T., Wang, J., and Jackson, T. A model for microwave emission from vegetation-covered fields. *Journal of Geophysical Research: Oceans*, 87 (C13):11229–11237, 1982.

Moesinger, L., Dorigo, W., Jeu, R. d., Schalie, R. v. d., Scanlon, T., Teubner, I., and Forkel, M. The global long-term microwave vegetation optical depth climate archive (vodca). *Earth System Science Data*, 12(1):177–196, 2020.

Monteith, J. Solar radiation and productivity in tropical ecosystems. *Journal of applied ecology*, 9(3):747–766, 1972.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., et al. Era5-land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data Discussions*, pages 1–50, 2021.

O'Sullivan, M., Smith, W. K., Sitch, S., Friedlingstein, P., Arora, V. K., Haverd, V., Jain, A. K., Kato, E., Kautz, M., Lombardozzi, D., Nabel, J. E. M. S., Tian, H., Vuichard, N., Wiltshire, A., Zhu, D., and Buermann, W. Climate-driven variability and trends in plant productivity over recent decades based on three global products. *Global Biogeochemical Cycles*, 34(12):e2020GB006613, 2020. doi: https://doi.org/10. 1029/2020GB006613. URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/ 10.1029/2020GB006613`. e2020GB006613 2020GB006613.

Owe, M., de Jeu, R., and Holmes, T. Multisensor historical climatology of satellite-derived global land surface moisture. *Journal of Geophysical Research: Earth Surface*, 113(F1), 2008.

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., et al. The fluxnet2015 dataset and the oneflux processing pipeline for eddy covariance data. *Scientific data*, 7(1):1–27, 2020.

Piao, S., Sitch, S., Ciais, P., Friedlingstein, P., Peylin, P., Wang, X., Ahlström, A., Anav, A., Canadell, J. G., Cong, N., Huntingford, C., Jung, M., Levis, S., Levy, P. E., Li, J., Lin, X., Lomas, M. R., Lu, M., Luo, Y., Ma, Y., Myneni, R. B., Poulter, B., Sun, Z., Wang, T., Viovy, N., Zaehle, S., and Zeng, N. Evaluation of terrestrial carbon cycle models for their response to climate variability and to co2 trends. *Global Change Biology*, 19(7):2117–2132, 2013. doi: https://doi.org/10.1111/gcb.12187. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.12187`.

Piles, M., Camps-Valls, G., Chaparro, D., Entekhabi, D., Konings, A. G., and Jagdhuber, T. Remote sensing of vegetation dynamics in agro-ecosystems using smap vegetation optical depth and optical vegetation indices. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 4346–4349. IEEE, 2017.

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., et al. An evaluation of smos l-band vegetation optical depth (l-vod) data sets: high sensitivity of l-vod to aboveground biomass in africa. *Biogeosciences*, 15(14):4627–4645, 2018.

Running, S., Mu, Q., and Zhao, M. Mod17a2h modis/terra gross primary productivity 8-day l4 global 500m sin grid v006. *NASA EOSDIS Land Processes DAAC.) Available at http://doi. org/10.5067/MODIS/MOD17A2H*, 6, 2015.

Running, S. W., Nemani, R., Glassy, J. M., and Thornton, P. E. Modis daily photosynthesis (psn) and annual net primary production (npp) product (mod17) algorithm theoretical basis document. *University of Montana, SCF At-Launch Algorithm ATBD Documents (available online at: www. ntsg. umt. edu/modis/ATBD/ATBD_MOD17_v21. pdf)*, 490, 1999.

Ryan, M. G. A simple method for estimating gross carbon budgets for vegetation in forest ecosystems. *Tree physiology*, 9(1-2):255–266, 1991.

Ryan, M. G., Lavigne, M. B., and Gower, S. T. Annual carbon cost of autotrophic respiration in boreal forest ecosystems in relation to species and climate. *Journal of Geophysical Research: Atmospheres*, 102(D24):28871–28883, 1997.

Sade, N., Gebremedhin, A., and Moshelion, M. Risk-taking plants: anisohydric behavior as a stress-resistance trait. *Plant signaling & behavior*, 7(7):767–770, 2012.

Schimel, D., Stephens, B. B., and Fisher, J. B. Effect of increasing co2 on the terrestrial carbon cycle. *Proceedings of the National Academy of Sciences*, 112(2):436–441, 2015.

Sen, P. K. Estimates of the regression coefficient based on kendall's tau. *Journal of the American statistical association*, 63(324):1379–1389, 1968.

Servén, D., Brummitt, C., and Abedi, H. pygam: Generalized additive models in python. zenodo, 2018.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., et al. Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences*, 12(3):653–679, 2015.

Taschetto, A. S. and England, M. H. El niño modoki impacts on australian rainfall. *Journal of Climate*, 22(11):3167–3174, 2009.

Teubner, I. E., Forkel, M., Wild, B., Mösinger, L., and Dorigo, W. Impact of temperature and water availability on microwave-derived gross primary production. *Biogeosciences*, 18(11):3285–3308, 2021. doi: 10.5194/bg-18-3285-2021. URL `https://bg.copernicus.org/articles/18/3285/2021/`.

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., et al. Assessing the relationship between microwave vegetation optical depth and gross primary production. *International journal of applied earth observation and geoinformation*, 65:79–91, 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 Sensing of Environment*, 229:100–113, 2019.

Theil, H. A rank-invariant method of linear and polynomial regression analysis. *Indagationes Mathematicae*, 12(85):173, 1950.

Tian, F., Brandt, M., Liu, Y. Y., Verger, A., Tagesson, T., Diouf, A. A., Rasmussen, K., Mbow, C., Wang, Y., and Fensholt, R. Remote sensing of vegetation dynamics

in drylands: Evaluating vegetation optical depth (vod) using avhrr ndvi and in situ green biomass data over west african sahel. *Remote Sensing of Environment*, 177: 265–276, 2016.

Tjoelker, M. G., Oleksyn, J., and Reich, P. B. Modelling respiration of vegetation: evidence for a general temperature-dependent q10. *Global change biology*, 7(2):223–230, 2001.

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M. A., Cescatti, A., Kiely, G., et al. Predicting carbon dioxide and energy fluxes across global fluxnet sites with regression algorithms. *Biogeosciences*, 13(14):4291–4313, 2016.

Tucker, C. J. and Sellers, P. Satellite remote sensing of primary production. *International journal of remote sensing*, 7(11):1395–1416, 1986.

Turner, D. P., Ritts, W. D., Cohen, W. B., Gower, S. T., Running, S. W., Zhao, M., Costa, M. H., Kirschbaum, A. A., Ham, J. M., Saleska, S. R., et al. Evaluation of modis npp and gpp products across multiple biomes. *Remote sensing of environment*, 102(3-4):282–292, 2006.

Ulaby, F. T., Moore, R. K., and Fung, A. K. Microwave remote sensing: Active and passive. volume 1-microwave remote sensing fundamentals and radiometry. 1981.

van der Schalie, R., de Jeu, R. A., Kerr, Y., 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 Sensing of Environment*, 189:180–193, 2017.

Vreugdenhil, M., Dorigo, W. A., Wagner, W., De Jeu, R. A., Hahn, S., and Van Marle, M. J. Analyzing the vegetation parameterization in the tu-wien ascat soil moisture retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 54(6):3513–3531, 2016.

Walker, A. P., De Kauwe, M. G., Bastos, A., Belmecheri, S., Georgiou, K., Keeling, R., McMahon, S. M., Medlyn, B. E., Moore, D. J., Norby, R. J., et al. Integrating the evidence for a terrestrial carbon sink caused by increasing atmospheric co2. *New Phytologist*, 2020.

Wang, L., Zhu, H., Lin, A., Zou, L., Qin, W., and Du, Q. Evaluation of the latest modis gpp products across multiple biomes using global eddy covariance flux data. *Remote Sensing*, 9(5):418, 2017.

Wardle, G. M., Pavey, C. R., and Dickman, C. R. Greening of arid australia: New insights from extreme years. *Austral Ecology*, 38(7):731–740, 2013.

Welp, L. R., Keeling, R. F., Meijer, H. A., Bollenbacher, A. F., Piper, S. C., Yoshimura, K., Francey, R. J., Allison, C. E., and Wahlen, M. Interannual variability in the oxygen isotopes of atmospheric co 2 driven by el niño. *Nature*, 477(7366):579–582, 2011.

Wilcox, R. R. *Fundamentals of modern statistical methods: Substantially improving power and accuracy.* Springer, 2010.

Wild, B., Teubner, I., Moesinger, L., Zotta, R.-M., Forkel, M., van der Schalie, R., Sitch, S., and Dorigo, W. A. Vodca2gpp–a new global, long-term (1988–2020) gpp dataset from microwave remote sensing. *Earth System Science Data Discussions*, pages 1–37, 2021a.

Wild, B., Teubner, I., Moesinger, L., Zotta, R.-M., Forkel, M., van der Schalie, R., Sitch, S., and Dorigo, W. A. Vodca2gpp [dataset]. *TU Data*, 2021b. doi: https://doi.org/10.48436/1k7aj-bdz35.

Wythers, K. R., Reich, P. B., and Bradford, J. B. Incorporating temperature-sensitive q10 and foliar respiration acclimation algorithms modifies modeled ecosystem responses to global change. *Journal of Geophysical Research: Biogeosciences*, 118(1): 77–90, 2013.

Zhai, P., Yu, R., Guo, Y., Li, Q., Ren, X., Wang, Y., Xu, W., Liu, Y., and Ding, Y. The strong el niño of 2015/16 and its dominant impacts on global and china's climate. *Journal of Meteorological Research*, 30(3):283–297, 2016.

Zhao, M., Heinsch, F. A., Nemani, R. R., and Running, S. W. Improvements of the modis terrestrial gross and net primary production global data set. *Remote sensing of Environment*, 95(2):164–176, 2005.

Zotta, R.-M., Dorigo, W. A., Moesinger, L., et al. Vodca cxku. in preperation.