THESIS

USING REMOTELY SENSED FLUORESCENCE AND SOIL MOISTURE TO BETTER UNDERSTAND THE SEASONAL CYCLE OF TROPICAL GRASSLANDS

Submitted by

Dakota Carlysle Smith

Department of Atmospheric Science

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2017

Master's Committee:

Advisor: A. Scott Denning

Melinda Smith
Christopher O'Dell
Christian Kummerow

ABSTRACT

USING REMOTELY SENSED FLUORESCENCE AND SOIL MOISTURE TO BETTER UNDERSTAND THE SEASONAL CYCLE OF TROPICAL GRASSLANDS

Seasonal grasslands account for a large area of Earth's land cover. Annual and seasonal changes in these grasslands have profound impacts on Earth's carbon, energy, and water cycles. In tropical grasslands, growth is commonly water-limited and the landscape oscillates between highly productive and unproductive. As the monsoon begins, soils moisten providing dry grasses the water necessary to photosynthesize. However, along with the rain come clouds that obscure satellite products that are commonly used to study productivity in these areas. To navigate this issue, we used solar induced fluorescence (SIF) products from OCO-2 along with soil moisture products from the Soil Moisture Active Passive satellite (SMAP) to "see through" the clouds to monitor grassland productivity. To get a broader understanding of the vegetation dynamics, we used the Simple Biosphere Model (SiB4) to simulate the seasonal cycles of vegetation. In conjunction with SiB4, the remotely sensed SIF and soil moisture observations were utilized to paint a clearer picture of seasonal productivity in tropical grasslands. The remotely sensed data is not available for every place at one time or at every time for one place. Thus, the study was focused on a large area from 15° E to 35° W and from 8°S to 20°N in the African Sahel. Instead of studying productivity relative to time, we studied it relative to soil moisture. Through this investigation we found soil moisture thresholds for the emergence of grassland growth, near linear grassland growth, and maturity of grassland growth. We also found that SiB4 overestimates SIF by about a factor of two for nearly every value of soil moisture. On the whole,

SiB4 does a surprisingly good job of predicting the response of seasonal growth in tropical grasslands to soil moisture. Future work will continue to integrate remotely sensed SIF & soil moisture with SiB4 to add to our growing knowledge of carbon, water, and energy cycling in tropical grasslands.

ACKNOWLEDGEMENTS

First, I'd like to thank my advisor, Dr. Scott Denning, for his scientific guidance and thoughtful mentorship. I've learned an enormous amount from his teaching and insight, both in and out of the classroom. Additionally, I'm grateful to my committee, Dr. Melinda Smith, Dr. Christopher O'Dell, and Dr. Christian Kummerow for their expertise.

I'm in debt to the Biocycle Group that provided significant aide to my research. Thank you to Dr. Katherine Haynes for the insight in modeling and phenology. Thank you to Dr. Ian Baker for being the jack of all trades and being a helping hand. Thank you to Nick Guyer for his work with OCO-2 SIF retrievals. Finally, thank you to Ammon Redman and Matt Bishop for their constant technological support.

The fluorescence data were produced by the OCO-2 project at the Jet Propulsion Laboratory, California Institute of Technology, and obtained from the OCO-2 data archive maintained at the National Aeronautics and Space Administration (NASA) Goddard Earth Science Data and Information Services Center. The soil moisture data were gathered by NASA's Soil Moisture Active Passive mission and made available by the National Snow and Ice Data Center Distributed Active Archive Center. The MODIS LAI subset data were acquired from NASA's Oak Ridge National Laboratory Distributed Active Archive Center for Biogeochemical Dynamics.

This research was supported by the National Science Foundation Science and Technology Center for Multi-Scale Modeling of Atmospheric Processes, managed by Colorado State University under cooperative agreement No.ATM-0425247.

TABLE OF CONTENTS

LIST OF FIGURES

Introduction

Tropical grasslands and savannas play a critical role in the carbon cycle and climate. The strongest land-atmosphere interactions in the climate system occur over grasslands and savannas [Koster et al., 2002]. Changes in tropical savannas and grasslands may be an important feedback to anthropogenic climate change [Staver et al., 2011]. Tropical grasslands are located in areas where growth is seasonally limited by moisture. Over the course of a year, the vegetation within these landscapes oscillates between highly productive and dormant. An example of this phenomenon exists in West Africa where seasonal rainfall associated with the inter-tropical convergence zone (ITCZ) creates a strong, north-to-south precipitation gradient. The timing and intensity of this rainfall determines the seasonal "green-up" and "brown-down" (phenology) of grasslands. Accurately predicting the timing and intensity of this seasonal cycle is crucial to estimating the seasonal carbon fluxes of this ecosystem and of Earth as a whole. Over many seasons, this prediction becomes more important as it can have a sizeable impact on global inter-annual fluxes of carbon.

Simulation of phenology is especially challenging in tropial grasslands. Unlike their temperate counterparts, phenology of these systems does not depend on strong seasonal changes in temperature and day length. In the model intercomparison presented by Schwalm et al. (2010), skill decreases in savannas and grasslands as compared to forested biomes. Model skill also decreased during fall and spring as models poorly handled initial leaf growth and senescence (browning). This was the case particularly for biomes with a seasonal cycle of leaf area index (LAI, the ratio of leaf area to the area of the underlying surface). One significant difference in how models handle this seasonal cycle is whether or not LAI is prescribed via satellite or is

prognosed. Overall, models that prescribe LAI have better skill compared to models that prognose LAI [Schwalm et al., 2010].

The use of satellite imagery to prescribe phenology of tropical grasslands is problematic because the beginning of the growing season is controlled by the onset of seasonal rains (the monsoon). This critical transition between the unproductive dry season and peak growing conditions coincides with heavy cloud cover associated with monsoon rains. This problem is highlighted in Figure 1 showing the contrast in cloud cover and vegetation over West Africa in January and July. Unfortunately, satellite imagery at the beginning of the growing season in tropical grasslands is very often missing because it is precisely the onset of rainfall that triggers plant growth. Rather than a well-observed and gradual transition from brown to green at the onset of the growing season, satellite data in visible wavelengths in this region instead shows a sudden transition from brown to white. Only when the clouds clear later in the season are highly productive grasslands revealed.

Vegetation properties derived from visible channels measured by the Moderate Resolution Imaging Spectroradiometer (MODIS) include LAI, the fraction of photosynthetically active radiation absorbed by vegetation (fPAR), and the Normalized Difference Vegetation Index (NDVI). All of these products are impossible to retrieve through heavy cloud cover, and produce unrealistically low values if the optical path is obscured by even thin or sub-pixel clouds. To compensate for cloud cover, all of these products are composited over eight days by assigning a value to each pixel that represents the maximum observed over the compositing period. Compositing smears out the brown-to-green transition associated with the first emergence of grasses, and may also alias the first appearance of green vegetation toward the end of the compositing period. Figure 2 shows a timeseries of LAI retrieved by MODIS along with the

number of valid pixels for 2015 across a 1°x1° (latitude x longitude) grid cell centered at 10° N and 11° E in Nigeria. Although valid LAI retrievals are available for over 90% of pixels in the dry season (days 300 to 75), the frequency of valid data drops dramatically with the onset of monsoon rains and there are very few valid data at all during the period of maximum growth (days 180 to 250). Worse, retrieved LAI shows strong dropouts during periods of heavy cloud cover when data are also sparse. These dropouts are unrealistic: actual grass may grow more or less quickly during the greenup period, but does not shed leaves until senescence much later in the year [Tan et al, 2011]. Models that use satellite data to prescribe LAI may substantially underpredict carbon when they are forced to repeatedly shed and regrow their canopies because of unrealistic early-season dropouts resulting from subgrid-scale cloud cover [Kevin Schaefer, personal communication].

As an alternative to prescribing seasonal and spatial variations of vegetation from satellite data, mechanistic models have been developed to predict phenology from local environmental conditions and the known behavior of plant physiology [e.g., Jolly and Running, 2004; Krinner et al, 2005; Corbin et al, 2016]. Advantages of this approach include the ability to conduct simulations during periods of heavy cloud cover, near-real time simulations when fresh imagery is not available, and simulations of both the distant past or the future for which there are no satellite data. The main disadvantage is that both the timing and magnitude of vegetation growth may be incorrectly simulated by prognostic phenology models. It is therefore very important to evaluate quantitative predictions of vegetation growth by prognostic models against measurements. Detailed multi-year datasets of the seasonal timing of grassland phenology are extremely sparse over sub-Saharan Africa.. Many studies [e.g. Prince, 1991; Fensholt et al., 2013; Dardel et al., 2013;] have used MODIS LAI and NDVI to describe and model Sahel

grasslands due to the absence of in situ measurements. While these studies are important for our understanding of this region, most focus on the interannual variability rather than seasonal variability. The studies that do focus on seasonality [Pierre et al., 2011; Bobee et al., 2012] provide only limited information about the transition from brown to green for the reasons outlined above.

Here we evaluate a new model of grassland phenology [Corbin et al, 2016], focusing on the early season greenup over North Africa which has previously been poorly observed. We take advantage of new data from NASA satellites which are less affected by cloud cover than visible imagery from MODIS. The Orbiting Carbon Observatory 2 (OCO-2) was used to retrieve solar-induced fluorescence (SIF) of chlorophyll as a proxy for photosynthesis. The Soil Moisture Active Passive (SMAP) satellite was used to retrieve surface (0-5cm) soil moisture. Chlorophyll fluorescence is retrieved in the near-infrared by the in-filling of solar Fraunhofer lines. Due to the nature of this signal, SIF retrievals are only affected by thick clouds that attenuate more than 90% of visible light [Frankenberg et al., 2012]. Soil moisture retrievals from SMAP are based on the dielectric properties of the soil, collected at low microwave frequencies and are almost completely insensitive to clouds. The nearly simultaneous retrieval of both surface soil moisture and chlorophyll fluorescence in tropical grasslands were used to quantify phenological triggers, time lags, and rates of growth as a function of soil moisture for early season grassland development over North Africa.

Retrievals of SIF in the near infrared (from OCO-2) and soil moisture in the microwave (from SMAP) are much more robust than retrieval of LAI from visible radiance by MODIS. Nevertheless, coverage of SIF on a any given day is quite sparse due to the narrow swath and orbital geometry of OCO-2 and heavy clouds still obscure some retrievals. Even using OCO-2

and SMAP, these limitations mean it is still impossible to document early season phenology at a given location. We therefore focus on analyzing changes in SIF with changes in soil moisture across many locations, rather than attempting to reconstruct seasonal timeseries of SIF and soil moisture at individual pixels. This approach allows us to quantify soil moisture triggers for growing season onset, the rate of early season growth, and the approach to maturity of North African grasslands as a function of soil moisture, and to compare these mechanistic relationships directly to those simulated by SiB4.

Hypotheses

1) The growing season onset in Sahelian grasslands will be defined by a soil moisture threshold. Plant growth will not begin until fraction plant available water in the top 18 centimeters reaches .12.
2) After reaching this threshold, growth begins as carbon in the roots is allocated to plant stem and leaves. As leaves begin to grow, carbon is then fixed from the atmosphere via photosynthesis. The rate of photosynthesis is governed by enzyme kinetics as described in the methods. During this period, the relationship between SIF and soil moisture will be linear.
3) Observed SIF will be similar in timing and magnitude as described by SiB4.

Methods

The area examined in both the model and observations was a box outlined from 8°N to 20°N and 15°W to 35°E. This area includes a variety of ecosystems from highly productive tropical forest to highly unproductive desert. The focus of this study is tropical (C4) grasslands that grow and senesce seasonally. Observations were only considered if the location was identified as a grassland or barren landcover type in OCO-2 retrievals using the International

Geosphere–Biosphere Programme (IGBP) classification based on 1 km MODIS imagery. Model output was also only considered for subgrid-scale areas that were identified as C4 grass from the same underlying MODIS classifications. These grasslands were analyzed during the growing season from May 1st to August 30th of 2015. Specifically, they were examined during the onset of growth which we defined separately at each location as the period before the seven day average of soil moisture retrieved by SMAP reaches its maximum.

Observations

The SMAP satellite was utilized for its passive, level two soil moisture data (SPL2SMP). As described in O'Neill et al., (2015), this data is gathered via the passive sensor on SMAP that measures the intensity of microwaves emitted by the soil. This microwave intensity is dependent on the dielectric properties of the soil, which are directly related to volumetric water content (vwc, the ratio of water volume to soil volume, in cm^3 cm^{-3}). This measurement can quantify the volumetric water content in the first five centimeters of soil with an accuracy of ± 0.04 cm^3/cm^3. At 6:00 am local overpass time, SMAP collects soil moisture retrievals for 36-km pixels in a swath approximately 1,000 km wide. SMAP provides global coverage after three days and repeats its track after sixteen days. These data were obtained through the NASA National Snow and Ice Data Center Distributed Active Archive Center at www.nsidc.org/data/smap.

SIF represents one of three outcomes for the energy of photons absorbed by a chlorophyll-*a* molecule: photosynthetic carbon assimiliation; fluorescent emission; and dissipation as heat (also known as non-photosynthetic quenching). As sunlight is absorbed, a chlorophyll molecule is raised to an excited state. Fluorescence occurs when radiation, between 600 and 800 nm, is re-radiated from the chlorophyll molecule as it returns to its ground state [Misra et al., 2012; Krause and Weis, 1984]. SIF is known to be an indicator of plant

photosynthesis [Krause & Weis, 1984; Campbell et al., 2007; Baker, 2008] and recent studies [Zhang et al., 2014; Guanter et al., 2012; Frankenberg et al., 2011b] have shown that remotely sensed SIF can be used as a proxy for Gross Primary Production (photosynthesis, GPP) in surface vegetation. SIF has been evaluated in a variety of biomes including croplands [Cheng et al., 2014; Gaunter et al., 2014], forests [Yang et al., 2015; Lee et al., 2013], and grasslands [Damm et al., 2015]. While the relationship between remotely sense SIF and GPP has been found to vary across ecosystems, SIF can still be used a proxy for plant productivity and has been found to be related to GPP more consistently than previously used remotely measured products, such as LAI, because it is emitted as a direct consequence of chlorophyll photochemistry [Damm et al., 2015].

We used SIF retrievals acquired from the OCO-2 satellite. OCO-2 takes advantage of Fraunhofer lines at wavelengths 758.8 nm and 770.1 nm to retrieve fluorescence [Frankenberg et al., 2011a, Joiner et al., 2011]. Fraunhofer lines are produced when particular wavelengths of radiation are less intense as they are emitted from a star due to gases in the star's atmosphere that are partially absorptive in those particular wavelengths. Without this effect, it would be impossible to identify signals of SIF against the much brighter background of scattered solar radiation at these wavelengths [Frankenberg et al., 2011a]. The fluorescence retrieval algorithm is outlined in Frankenberg et al. (2011a) and Joiner et al. (2011) with specifics to OCO-2 in Frankenberg et al. (2014). It is important to note that OCO-2 passes over at 1:35 pm local time, different from SMAP. The maximum OCO-2 swath width is 10.3 km [Frankenberg et al., 2014] with measurement resolution of 1.3 x 2.25 km^2. OCO-2 data were obtained through the NASA Goddard Earth Science Data and Information Services Center at www.co2.jpl.nasa.gov.

Using remotely measured SIF and soil moisture to diagnose productivity in Sahelian grasslands offers several coverage advantages, but spatial and temporal limitations still exist in both datasets. OCO-2 has near global data coverage after sixteen days and SMAP has global coverage after three days. SIF retrievals from OCO-2 are available along 10 km wide swaths, for successive orbits 24° apart in longitude. Precession of the OCO-2 orbit fills in the gaps between these narrow swaths over 16 days, but growing season onset is nearly certain to fall between successive retrievals at any given location. Retrievals of near-surface soil moisture by SMAP are less sparse because of the wider swath, but still average only once every three days for successive retrievals at the same location. Neither soil moisture nor GPP can therefore be estimated for one day at every location nor at one location for every day. Instead of evaluating ecosystem productivity at each pixel over time, we analyzed SIF relative to soil moisture. To accurately compare the observations to the model, SiB4 was evaluated in the same fashion.

The difference in time and location of SMAP and OCO-2 retrieval swaths was challenging. There were many days when the OCO-2 and SMAP swaths were completely out of sync, producing zero co-located points (Figure 3**)**. Near-surface soil moisture responds to the time integral of precipitation minus losses from evapotranspiration, runoff, and percolation to deeper layers. It therefore has a memory of several days in the study region. We addressed the space-time mismatch between OCO-2 and SMAP swaths by assuming that a given pixel retains its soil moisture in the top 5 cm for two days after it's measured, or until it is measured again, whichever comes first. In this way, we produced a fully populated 36-km grid of SMAP retrievals for every day during the study period. The SMAP data were composited into a three day grid (0.4° by 0.3°) (Figure 4**)**. Pairs of SIF and soil moisture retrievals were constructed by sampling the three-day composite soil moisture grid at the time and location of every valid SIF

retrieval. Co-locating of SIF and SMAP retrievals was done for every day from May 1st, 2015 to August 30th, 2015 across the study area. This co-location was done for each day by locating any SIF retrievals within the 0.4° by 0.3° SMAP 3-day composite grid cell. The OCO-2 data has a finer resolution than the SMAP data so each soil moisture retrieval was co-located with many SIF retrievals. Figure 4 shows the result of co-locating one day's worth of SIF retrievals within the 3-day soil moisture composite.

Another challenge was the isolation of data that represented early season growth only, because we wanted to focus on the relationships among soil moisture and emergence, growth rate, and approach to maturity rather than changes in grassland productivity during senescence and dormancy. To better identify this period, a weekly running average of soil moisture was created for each grid cell in the SMAP 3-day composite. Figure 5 shows the weekly running average of soil moisture at a given location over the course of a year. As shown, the shape of the soil moisture seasonality is dramatically different depending on the location of the co-located retrievals. The length and timing of the growing season varies across the study area. To hone in on the onset of growth, co-located retrievals were only considered if they occurred before the maximum weekly averaged soil moisture. The dark shaded area in Figure 5 shows which data were used at these locations. Using this method, the data are representative of the green-up period only, neglecting retrievals representative of the brown-down. This process produced over 100,000 co-located points.

The fjeld-of-view (FOV) of a SMAP observation is much larger than the OCO-2 FOV which means there can be many SIF retrievals inside one SMAP composite grid cell. All SIF retrievals were averaged for each soil moisture retrieval to produce one SIF value for each SMAP grid cell. Daily mean SIF averages for each SMAP cell were binned according to their

co-located soil moisture value. Soil moisture bins were created for every 0.025 cm^3/cm^3 from 0.0 cm^3/cm^3 to 0.30 cm^3/cm^3. The average and standard deviation of SIF were determined for each bin.

Model

The fourth version of the Simple Biosphere Model (SiB4) simulates heterogeneous land-atmosphere fluxes, environmentally responsive prognostic phenology, dynamic carbon allocation, and cascading carbon pools from live biomass to surface litter to soil organic matter. Rather than relying on satellite data for the vegetation state, SiB4 fully simulates the terrestrial carbon cycle by using the carbon fluxes to determine the above and belowground biomass, which in turn feeds back to impact carbon assimilation and respiration. For every time-step, SiB4 computes albedo, radiation, temperature, and soil moisture, as well as the resulting energy exchanges, moisture fluxes, and carbon fluxes. Photosynthesis depends directly on environmental factors (humidity, moisture, and temperature) and the fraction of absorbed photosynthetically active radiation (FPAR, calculated from aboveground biomass); and carbon uptake was determined using enzyme kinetics [*Farquhar et al.*, 1980] and stomatal physiology [*Collatz et al.,* 1991 and 1992]. Carbon release occurs from both autotrophic and heterotrophic respiration. Both biomass growth and maintenance contribute to autotrophic respiration, and heterotrophic respiration depends on moisture, temperature and the amount of dead plant material in the surface and soil carbon pools.

To calculate the carbon pools, carbon fluxes are summed daily, and the net carbon is allocated to the live pools. For grasslands, the carbon fixed during photosynthesis is allocated to four live carbon pools (leaf, fine root, stem, and seed). The allocation of carbon to the separate pools is determined from the current phenology stage, which depends on both the benefit of

growing new aboveground biomass and the plant stress from day length, soil moisture, and temperature. The phenology stage, combined with temperature and moisture environmental adjustments, dictates the fraction of photosynthesized carbon allocated to each live carbon pool. Once the carbon is allocated, the pools and related land surface properties are updated and used for sub-hourly photosynthetic assimilation as well as sub-hourly autotrophic and heterotrophic respiration. Carbon from the live carbon pools is transferred daily to six dead carbon pools (metabolic litter, structural litter, standing dead, soil litter, soil slow, and soil passive). This sequence completes the carbon cycle, providing self-consistent predicted vegetation state, carbon pools, and land-atmosphere exchanges.

Since grasslands are extremely responsive to precipitation, simulations take advantage of the stress-based phenological scheme in SiB4. To start a growing season, growth was stimulated once the grassland has sufficient light, temperature and moisture. Early in the growing season, the assimilated carbon is allocated primarily to the stems and leaves to simulate the rapid green-up. As the aboveground live biomass increases, the net benefit of adding new leaves decreases and the grassland reaches the mature phenological stage where the allocation was more evenly distributed amongst the pools. Once stresses begin increasing due to hot conditions and/or soil moisture stress, carbon allocation is shifted to the seed and root pools.

While phenology controls the addition of new carbon into the appropriate pools, SiB4 simulates senescence by altering the biomass transfer rates. The rate of transfer for the roots depends on moisture and temperature, and while it can change seasonally, it occurs year-round. In contrast, the aboveground live carbon pool transfer is highly varying. For grasslands, to represent both the gradual browning in temperate climates as well as the rapid brown-down in stressed desert environments, the transfer of the aboveground live biomass depends on the

assimilation rate, daylength, and temperature (both low and high): as these stresses increase, the transfer rate increases. The transfer of the leaf, stem, and seed biomass to the standing dead and litter pools turns the grassland from green to brown and eventually ends the growing season once the live pools are depleted. In the case of desert environments as seen across Africa, the high temperatures during the dry season rapidly increase the rate of aboveground biomass transfer, causing the rapid browning and subsequent end of the growing season.

An upgrade to SiB4, as compared to SiB3, was its ability to model SIF as part of photosynthesis. The three fates of photons absorbed by leaves include SIF, breaking chemical bonds in carbon dioxide and water to create sugars in photosystems 1 and 2 (PSI, PSII), or being released as heat via nonphotochemical quenching (NPQ) or thermal dissipation (D). The relative yields ($\mathbf{\Phi}_x$) of these processes accounts for all photons absorbed by chlorophyll [van der Tol et al., *2014*], so that:

$$\phi_p + \phi_D + \phi_f + \phi_{NPQ} = 1 \quad (1)$$

Fluorescence yield ($\mathbf{\Phi}_f$) has been shown to correlate with photosynthetic yield ($\mathbf{\Phi}_p$) under quiescent conditions [Genty et al., *1988*] as well as stressed conditions like drought or extreme temperature [Flexas et al., 2002; Daumard et al., 2010]. These stresses are immediately felt as changes GPP and/or SIF while canopy-scale physiology (i.e. LAI, fPAR) take longer to react. Since $\mathbf{\Phi}_p$ is calculated by SiB4, SIF can be estimated using the relationship between $\mathbf{\Phi}_p$ and $\mathbf{\Phi}_f$.

In van der Tol et al., (2014), a detailed canopy-level model is described. While useful, this model is cumbersome to simulate globally. Instead, SiB4 contains a streamlined set of equations that exploits the enzyme-kinetics features in SiB4 to simulate SIF in an efficient manner.

The yields of the photon pathways are determined by nondimensional rate constants, given by:

$$\phi_x = \frac{k_x}{\sum k} \tag{2}$$

A maximum quantum yield of photosynthesis, or $\mathbf{\Phi}_{pmax}$, in PSII can be expressed as a function of rate constants for fluorescence (k_f), thermal deactivation (k_D), photosynthesis (k_p), as:

$$\phi_{pmax} = \frac{k_{p0}}{(k_f + k_D + k_p)} \tag{3}$$

Within SiB4 the values of k_f and k_{p0} are constant (0.05 and 4.0, respectively) and the coefficient for thermal dissipation (k_D) is 0.95 when canopy temperature is below 300K. Otherwise, k_D is calculated by:

$$K_D = 0.95 \times (T_{canopy} - 300) \times 0.0236 \tag{4}$$

The photochemical yield below the optimum rate can be calculated as:

$$\phi_p = \phi_{pmax} \left[\frac{GPP\left(\frac{C_c + 2.0\Gamma^*}{C_c - \Gamma^*}\right)}{PAR}\right] \tag{5}$$

where C_c is chloroplast CO_2 concentration, $\mathbf{\Gamma}^*$ is CO_2 compensation point, and PAR is Photosynthetically Active Radiation. An x-factor, relating actual and optimal quantum yield, is obtained by:

$$x = 1.0 - \frac{\phi_p}{\phi_{pmax}} \tag{6}$$

The relationship between NPQ and photosynthetic yield is not completely understood, therefore empirical relationships outlined in van der Tol et al. (2014) will be used. It is shown that there is relationship between the rate constant for NPQ (k_{NPQ}) and x is different under conditions of low and high stress. At low stress, k_{NPQ} is defined as:

$$k_{NPQ} = k_{NPQ0} \frac{(1.0 - a2) \times x^{a1}}{(a2 + x^{a1})} \tag{7}$$

where k_{NPQ} = 2.48, a1 = 2.83, and a2 = 0.114. Under high vegetative stress the equation is the same but with coefficients k_{NPQ} = 5.01, a1 = 1.93, and a2 = 10.0. The exact definition of what constitutes vegetative stress is uncertain. For this study, we use a value of 0.5 for the *waterstress* parameter as defined by Baker et al. (2008) to decipher between non-stressed and stressed conditions. With these terms we can calculate a maximum fluorescence yield ($\mathbf{\Phi}_{fmax}$) as:

$$\phi_{fmax} = \frac{k_f}{(k_f + k_D + k_{NPQ})} \tag{8}$$

Actual fluorescence yield ($\mathbf{\Phi}_f$) is calculated as:

$$\phi_f = \phi_{fmax} (1.0 - \phi_p) \tag{9}$$

To approximate SIF at the satellite level, atmospheric scattering is accounted for by multiplying canopy-level SIF by the ratio of downwelling shortwave at the surface to top-of-atmosphere radiation.

A regional SiB4 simulation was performed over the study region with hourly output for every plant functional type within each .5° by .5° grid box producing over 5,000 grid points. Subgrid-scale areas covered by grass, shrubs, forest and bare soil were simulated separately using hourly surface weather derived from the NASA Modern Era Reanalysis (Rienecker, et al., 2011) on a 0.5 x 0.67 grid. The exact same method of co-locating, averaging, and binning as described for the observations, was done for model output. To best mirror the observations, soil moisture was collected at 6:00 am local time and SIF output was collected at 1:00pm local time. A cloud mask was used to exclude output for which less than half of the solar radiation at the top of the atmosphere reached the surface.

Results

We compiled a comprehensive data set of co-located contemporaneous SIF and soil moisture retrievals during the early growing season across the African Sahel. For a view of the spatial coverage of the data set, Figure 6 shows the number of co-location observations in each .5° by .5° grid box. Figure 7 shows the number of simulated co-locations from SiB4 output. The co-located retrievals are different for the model and observations because the model has ouput for every grid cell for every time step whereas the observations are only co-located when the satellite swaths overlap. Co-located observations and simulations are most common east of 20°E and north of 15°N. This distribution is a product of both the frequency of clear skies and of landcover type. Areas with the highest density of co-located observations are mainly grassland and barren while areas with the lowest density of co-located observations are mainly forest and savanna which were excluded from this study.

First, the two datasets were examined on different axes, to better compare the overall patterns. In Figure 8, average observed SIF is plotted against binned soil moisture with the number of corresponding retrievals for each soil moisture bin plotted as a histogram. Each soil moisture bin has at least three thousand co-located retrievals with over five thousand in the majority of them. While each SIF versus soil moisture point has a large variance, a clear pattern emerges in the data. When the top five centimeters of soil contain less than 0.1 cm^3/cm^3 of water, grasslands are dormant. From 0.0 to 0.1 cm^3/cm^3, SIF is low but fairly constant. Interestingly, SIF is still greater than zero in the driest soils. While only barren and grassland landcover pixels were choosen, some shrubs or trees may exist on a sub-pixel scale resulting in nonzero SIF. Productivty emerges around soil moisture levels of 0.1 cm^3/cm^3 with near linear growth to 0.2

cm^3/cm^3. From 0.1 to 0.2 cm^3/cm^3, SIF increases linearly with increasing soil moisture. Beyond soil moisture levels of 0.2 cm^3/cm^3, grasslands appear to mature with constant productivity.

Next, modeled SIF versus binned soil moisture will be examined in Figure 9. More co-located SIF and soil moisture retrievals were collected for the model as compared to the observations. This is due to the requirement that OCO-2 provided in only sampling a small portion of the study area every day while SiB4 has fluorescence simulations for every location, every day. The variance for each point is noticeably less as compared to the observations. A major reason for this is that grasslands are modeled by the same equations and act the same at every location. Modeled SIF has a much larger range from minimum to maximum as compared to the observations. Most markedly, the model overestimates SIF for nearly every soil moisture bin. Several potential reasons for this include: SiB4 overestimating LAI or V_{max}, not enough in-canopy absorption, issues with atmospheric radiative transfer, or excessive fluorescent yield.

Modeled SIF is dormant up to 0.0525 cm^3/cm^3 but begins to grow after 0.055 cm^3/cm^3, slightly earlier than the observations. SiB4 does not have a required volumetric soil moisture requirement before growth but rather a required fraction of total available water in the top three soil layers (.12) along with a required number of sustained days (4) with this requirement. Similar to the observations, SIF is likely nonzero in SiB4 due to a minimum LAI required by SiB4 regardless of soil moisture. As grasslands contiue to grow, SIF responds linearaly to increasing soil moisture. SiB4 seems to model the rate of this increase fairly well. Still, SIF is overestimated for each soil moisture value. Mature growth occurs around 0.225 cm^3/cm^3 but not before dipping around 0.200 cm^3/cm^3.

Overall, SiB4 models the response in grassland productivty to soil moisture relatively well in Sahelian grasslands. While there are some descrepancies between the model output and

observations, they have similar shape and are in good agreement. The biggest descrepancy is SiB4's systematic over estimation of SIF. This is due to either an overestimation of LAI with accurate SIF, an overestimation of SIF with accurate LAI, or something in between.

Figure 9 shows a similar pattern in modeled SIF versus soil moisture compared to the observations. Grasslands remain unproductive and dormant with soil moisture values less than 0.05 m^3/m^3. Growth begins near 0.05m^3/m^3 and becomes nearly linear until 0.225 m^3/m^3. Grasslands become mature with near constant fluorscence after 0.225 m^3/m^3. Finally, to compare the overall shape of both datasets, they will be compared, side-by-side, on the same scale in Figure 10. This plot shows that SiB4 does a good job modeling fluorescence versus soil moisture, particularly the overall pattern. As previously noted, the magnitude of fluorescence is still drastically higher than the remotely-sensed fluorescence retrievals.

Evaluation of Hypotheses

1) The onset of plant growth is characterized by a soil moisture threshold however it was predicted too early, occuring around .075 cm^3/cm^3 in the model and near .10 cm^3/cm^3 in the observations.
2) After the onset of growth, SIF and soil moisture were directed related following a linear relationship. The magnitude of SIF was not accurately predicted.
3) The methods and equations described by SiB4 largely overestimate SIF. The way SIF is scaled from leaf to canopy in SiB4 is a likely cause.

Conclusions

This study presents a successful analysis of early season phenology in Sahelian grasslands. By doing so, this study shines a light on relatively undersampled and less understood region. More than one hundred thousand of paired observations of near-surface soil moisture and

grassland productiviy obtained via satellites creates a dataset that has never beeen utilized before. This data set provides a more robust overview of this region, particularly compared to products like LAI retrieved by MODIS.

This new data set shows three regimes in the response to grassland productivity to soil moisture. When soils are driest, grasslands are relatively dorminant with nonzero SIF from shrubs and trees. As soil moisture increases, grassland productivity emerges and increases linearly. Grasslands then reach a soil moisture threshold for which productivity remains relatively constant. SiB4 models does a fairly good modeling these processs with a few exceptions. SiB4 begins grassland growth slightly earlier than observations. This is due to SiB4's moisture and day requirement of fraction of total available water. SiB4 also systematically overestimates grassland SIF for nearly every soil moisture. Several potential reasons for this are: 1) an overestimation of LAI; 2) poorly modeled radiative transfer at the canopy level; 3) excessive fluorescent yield.

To build this body of work, these methods can be utilized in other tropical grasslands across the world over multiple seasons. The growing seasons of 2016 and 2017 can now be analyzed as SMAP and OCO-2 data continue to be collected. The last piece to this body of work is to adjust SiB4 based on newfound relationships between fluorescence and soil moisture. Extending this to other models, fluorescence could be tested as a replacement for satellite derived LAI to prescribe initial plant growth.

This study is just one small stepping stone in the pursuit of accurately modeling and understanding the global carbon cycle. More analysis will be required, particularly on understudied and poorly understood regions like Sahelian grasslands, to get closer and closer to that goal.

Figures

Figure 1. Visible Satellite and NDVI in Two Seasons in African Sahel

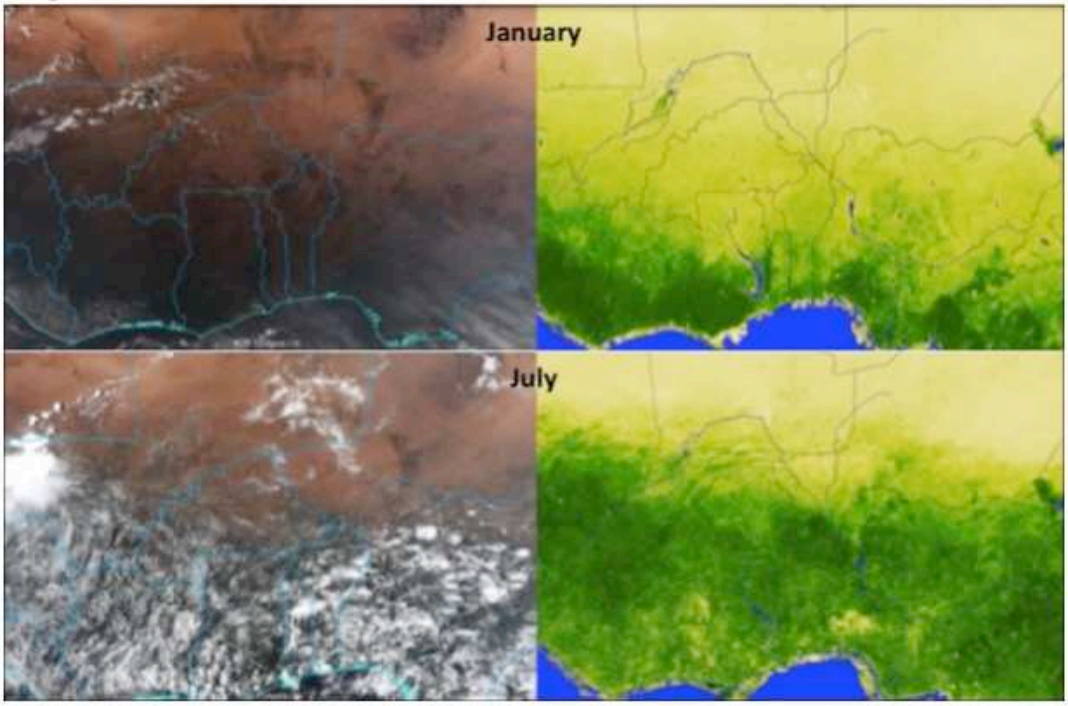

Figure 1: In January, skies are clear and plants are relatively unproductive in the Sahel. As rain and clouds approach in July, the Normalized Difference Vegetation Index (NDVI, retrieved by the National Oceanic and Atmospheric Administration's Advanced Very High Resolution Radiometer (AVHRR), responds as plant growth ensues. Due to satellite restraints, NDVI can mainly only be seen to respond after clouds have left the region.

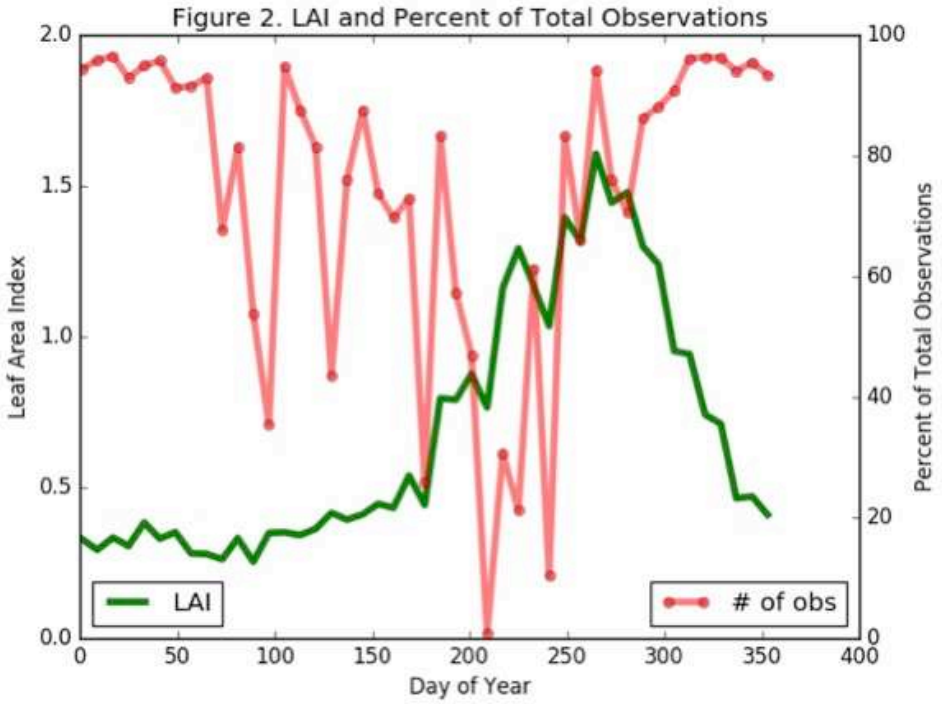

Figure 2: These observations were retrieved by MODIS-Terra in 2015 for a 101km by 101km grid cell centered at 10 degrees north and 11 degrees east. Each point is an 8-day composite of each 1km by 1km pixel inside the 101km by 101km grid cell.

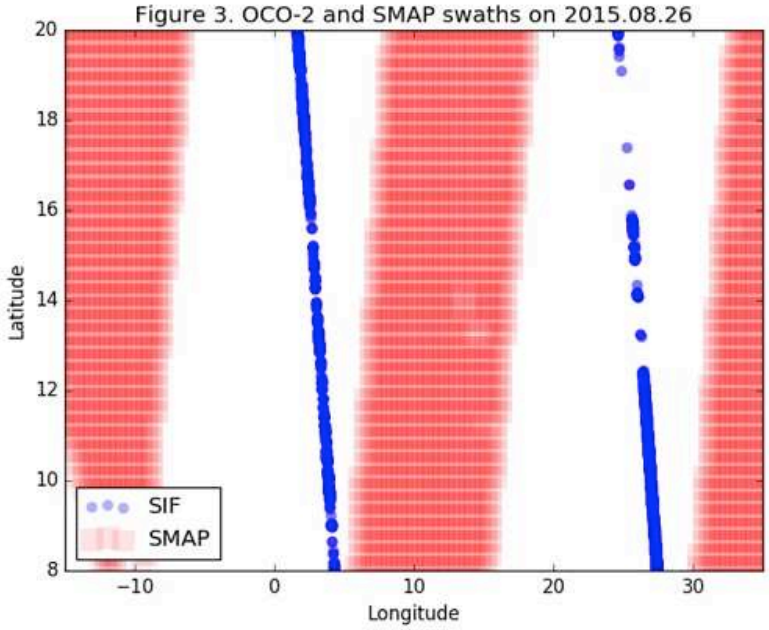

Figure 3: Each blue circle represents a SIF retrieval by OCO-2 and each red square represents a soil moisture retrieval by SMAP. For this day, the swaths are out of sync and zero co-located points are collected.

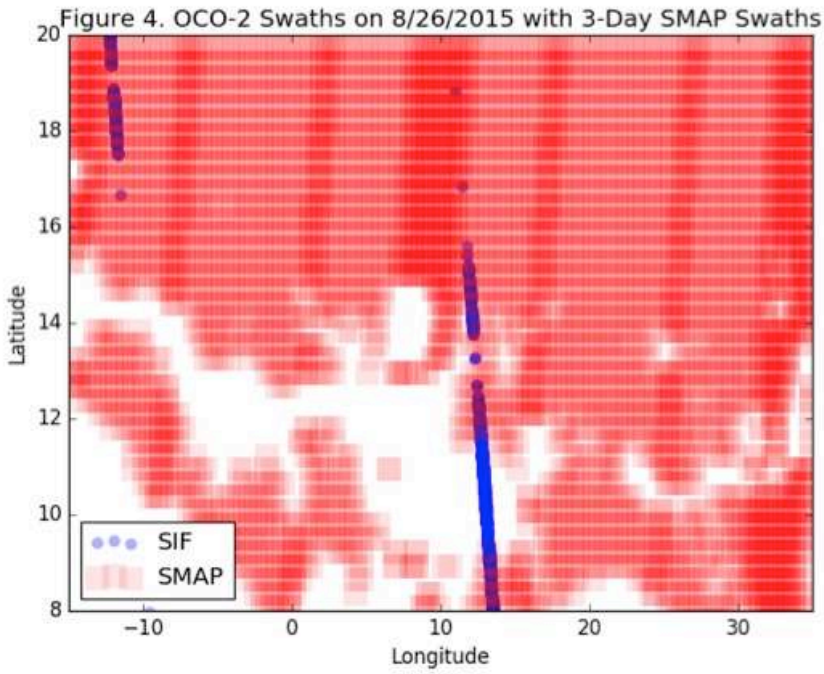

Figure 4: Each blue circle represents a SIF retrieval by OCO-2 and each red square represents a soil moisture retrieval by SMAP. With the SMAP 3-day composite grid, hundreds of new co-located points are collected. This new grid also filters out unwanted landcover types.

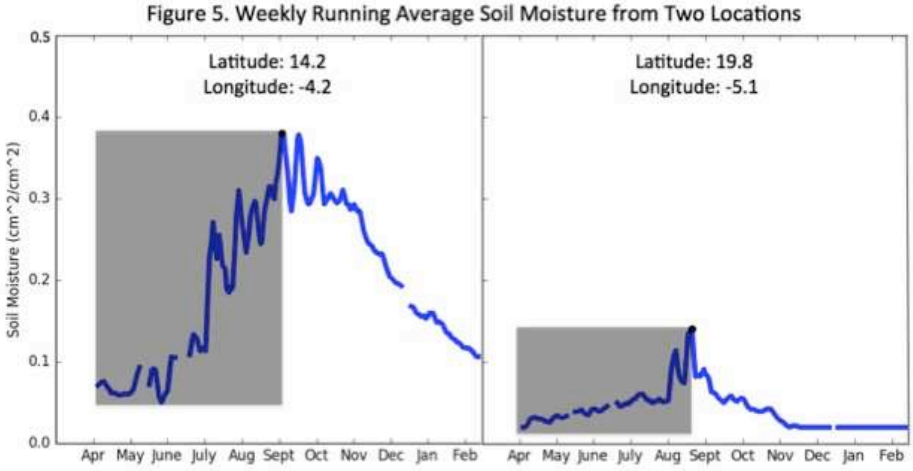

Figure 5: Soil moisture (cm^2/cm^2) retrieved by SMAP for two grid cells was averaged every seven days and plotted over the year of 2015. The black dot represents the maximum averaged seven day soil moisture and the black box represents the time prior to this maximum. Depending on the location, the shape of this graph varies.

Figure 6. Density Map of Co-located Observed SIF & Soil Moisture

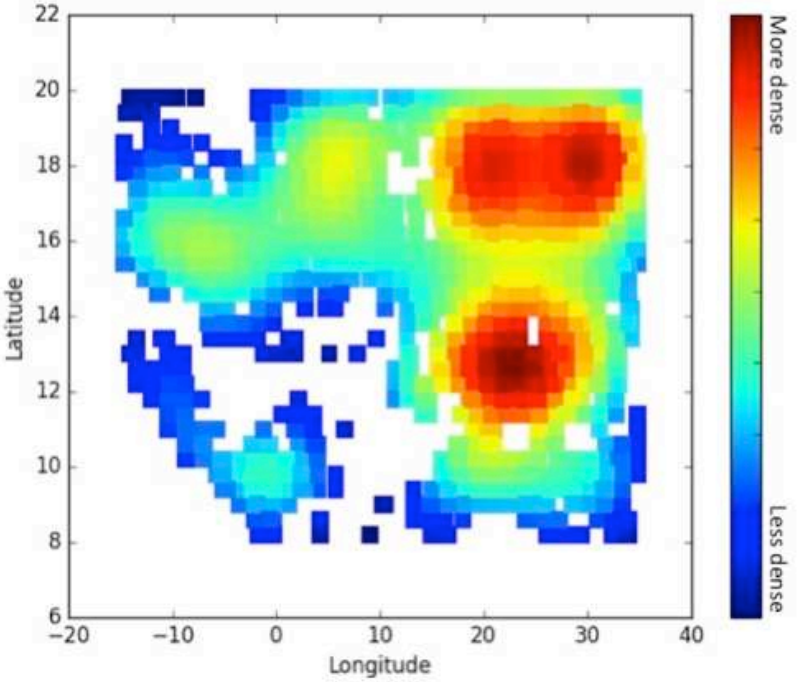

Figure 6: Each colored grid cell represents a location for which co-located observed SIF and observed soil moisture retrievals were found. Several hot spots exist in the eastern and northern half of the study area. Areas with few co-located retrievals exist due to landcover, satellite swath, and cloud restraints.

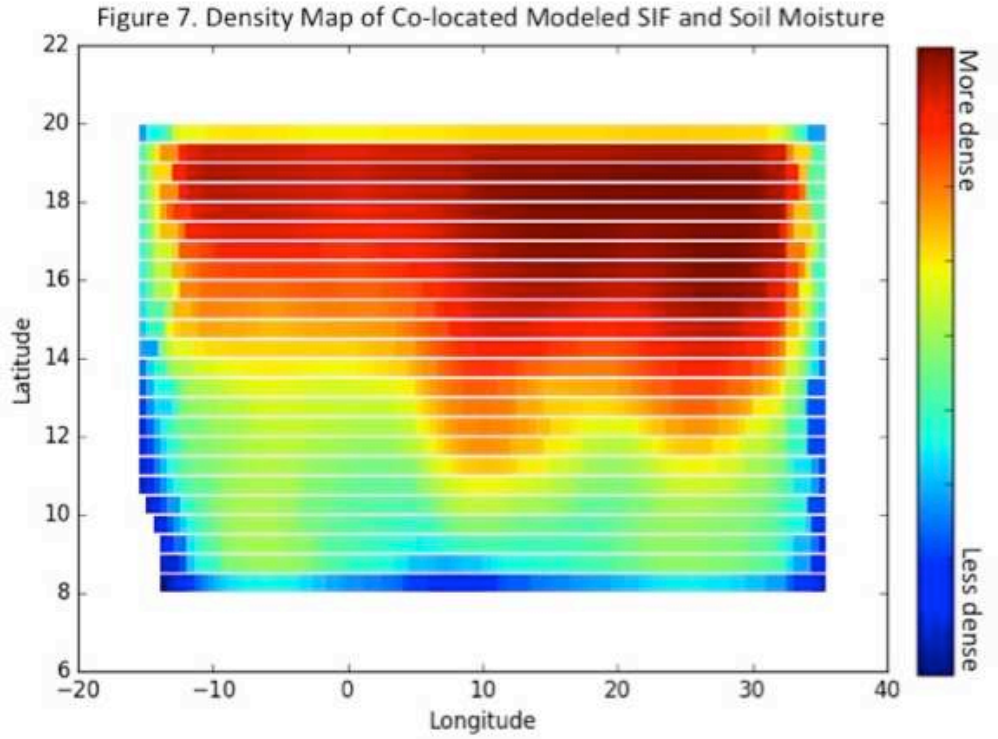

Figure 7: Each colored grid cell represents a location for which co-located modeled SIF and modeled soil moisture retrievals were found. Most retrievals were found in the northern half of the study area. Areas with less co-located retrievals exist due to landcover and cloud restraints.

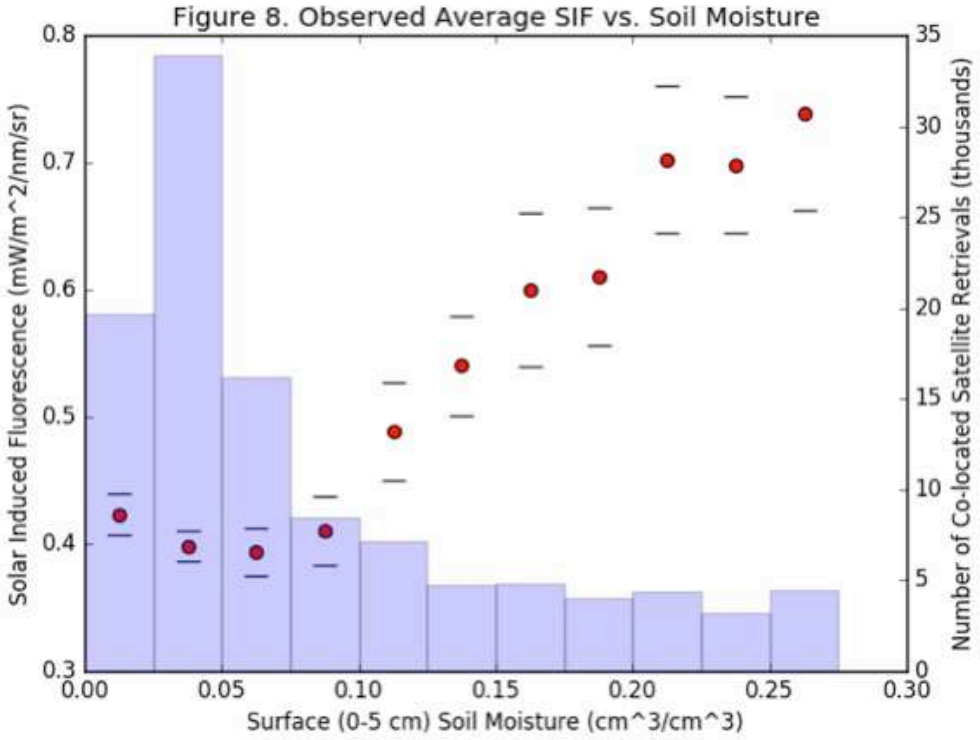

Figure 8: Each red dot represents SIF (mW/m^2/nm/sr) averaged for each surface (0-5 cm) soil moisture (cm^3/cm^3) bin. Bins range from 0 to 0.275 (cm^3/cm^3), every 0.025 (cm^3/cm^3). The number of co-located satellite retrievals is plotted a bar graph in light blue.

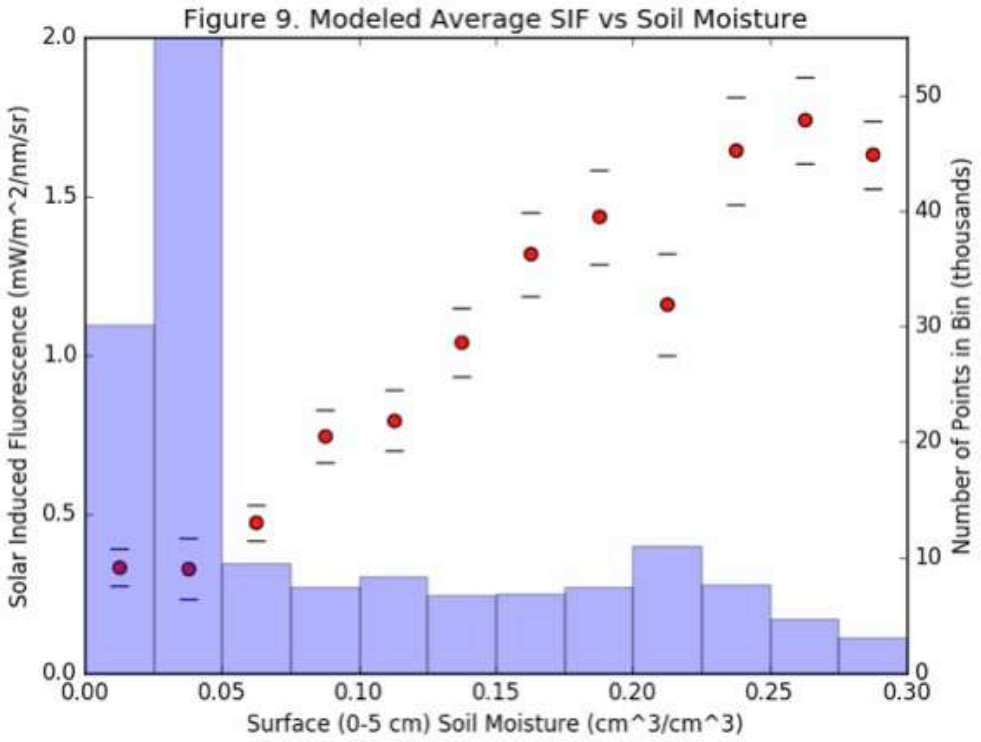

Figure 9: Each red dot represents SIF (mW/m^2/nm/sr) averaged for each surface (0-5 cm) soil moisture (cm^3/cm^3) bin. Bins range from 0 to 0.275 (cm^3/cm^3), every 0.025 (cm^3/cm^3). The number of co-located model output is plotted a bar graph in light blue.

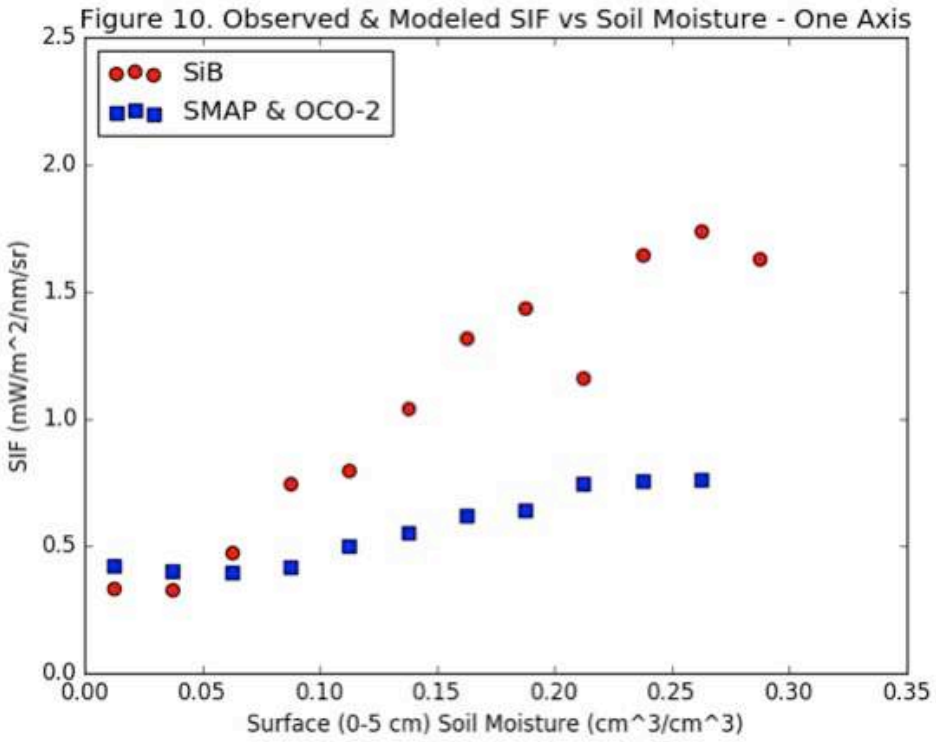

Figure 10: Each red dot and blue square represents SIF (mW/m^2/nm/sr) averaged for each surface (0-5 cm) soil moisture (cm^3/cm^3) bin in SiB4 and satellite retrievals, respectively. This demonstrates that SiB4 systematically overestimates SIF as compared to the observations.

REFERENCES

Baker, I. T., L. Prihodko, A. S. Denning, M. Goulden, S. Miller, and H. R. da Rocha (2008), Seasonal drought stress in the Amazon: Reconciling models and observations, J. Geophys. Res., 113, G00B01, doi:10.1029/2007JG000644.

Bobée, C., Ottlé, C., Maignan, F., Noblet-Ducoudré, N. D., Maugis, P., Lézine, A., & Ndiaye, M. (2012). Analysis of vegetation seasonality in Sahelian environments using MODIS LAI, in association with land cover and rainfall. Journal of Arid Environments, 84, 38-50. doi:10.1016/j.jaridenv.2012.03.005

Campbell, P. E., Middleton, E., Corp, L., & Kim, M. (2008). Contribution of chlorophyll fluorescence to the apparent vegetation reflectance. Science of The Total Environment, 404(2-3), 433-439. doi:10.1016/j.scitotenv.2007.11.004

Cheng, Y., Middleton, E., Zhang, Q., Huemmrich, K., Campbell, P., Corp, L., . . . Daughtry, C. (2013). Integrating Solar Induced Fluorescence and the Photochemical Reflectance Index for Estimating Gross Primary Production in a Cornfield. Remote Sensing, 5(12), 6857-6879. doi:10.3390/rs5126857

Chevallier, F., and C. W. O'Dell (2013), Error statistics of Bayesian CO2 flux inversion schemes as seen from GOSAT, Geophys. Res. Lett., 40(6), 1252–1256, doi:10.1002/grl.50228.

Damm, A., Paul-Limoges, E., Ammann, C., & Schaepman, M. (2015). Far-red sun-induced chlorophyll fluorescence shows ecosystem-specific relationships to gross primary production: An assessment based on observational and modeling approaches. *Remote Sensing of Environment, 166*, 91-105. doi:10.1016/j.rse.2015.06.004

Daumard, Fabrice & Champagne, Sébastien & Fournier, Antoine & Goulas, Yves & Ounis, Abderrahmane & Hanocq, J.F. & Moya, Ismaël. (2010). A Field Platform for Continuous Measurement of Canopy Fluorescence. Geoscience and Remote Sensing, IEEE Transactions on. 48. 3358 - 3368. 10.1109/TGRS.2010.2046420.

Dardel, C., Kergoat, L., Hiernaux, P., Mougin, E., Grippa, M., & Tucker, C. (2014). Re-greening Sahel: 30years of remote sensing data and field observations (Mali, Niger). Remote Sensing of Environment, 140, 350-364. doi:10.1016/j.rse.2013.09.011

Denning, S., D. A. Randall, G. J. Collatz, and P. J. Sellers (1996), Simulations of terrestrial carbon metabolism and atmospheric CO2 in a general circulation model, Tellus B, 48(4), 543–567.

Fensholt, R., Rasmussen, K., Kaspersen, P. S., Huber, S., Horion, S. M. A. F., & Swinnen, E. (2013). Assessing land degradation/recovery in the African Sahel from long-term earth

observation based primary productivity and precipitation relationships. Remote Sensing, 5(2), 664-686. DOI: 10.3390/rs5020664

J. FLEXAS, H. MEDRANO; Drought-inhibition of Photosynthesis in C_3 Plants: Stomatal and Non-stomatal Limitations Revisited, *Annals of Botany*, Volume 89, Issue 2, 1 February 2002, Pages 183–189, https://doi.org/10.1093/aob/mcf027

Frankenberg C, O'Dell C, Berry JA et al. (2014) Prospects for chlorophyll fluorescence remote sensing from the Orbiting Carbon Observatory-2. Remote Sensing of Envi- ronment, 147, 1–12.

Frankenberg, C., et al. (2011), New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity, Geophys. Res. Lett., 3 8, L17706, doi:10.1029/ 2011GL048738.

Frankenberg, C., C. O'Dell, L. Guanter, and J. McDuffie (2012), Remote sensing of near-infrared chlorophyll fluorescence from space in scattering atmospheres: implications for its retrieval and interferences with atmospheric CO2 retrievals, Atmos. Meas. Tech., 5(8), 2081–2094, doi:10.5194/amt-5-2081-2012.

Genty, B., Briantais, J., & Baker, N. R. (1989). The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. Biochimica et Biophysica Acta (BBA) - General Subjects, 990(1), 87-92. doi:10.1016/s0304-4165(89)80016-9

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. A., . . . Griffis, T. J. (2014). Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. Proceedings of the National Academy of Sciences, 111(14). doi:10.1073/pnas.1320008111

Gurney, K. R., R. M. Law, S. Denning, P. J. Rayner, D. F. Baker, P. Bousquet, L. Bruhwiler, Y. H. Chen, P. Ciais, and S. Fan (2002), Towards robust regional estimates of CO2 sources and sinks using atmospheric transport models, Nature, 415(6872), 626–630.

Joiner et al., 2011 - Joiner, J., Yoshida, Y., Vasilkov, A. P., Yoshida, Y., Corp, L. A., and Middleton, E. M.: First observations of global and seasonal terrestrial chlorophyll fluorescence from space, Biogeosciences, 8, 637-651, https://doi.org/10.5194/bg-8-637-2011, 2011.

Jolly, W. M., and S. W. Running (2004), Effects of precipitation and soil water potential on drought deciduous phenology in the Kalahari, Global Change Biology, 10(3), 303–308, doi:10.1046/j.1365-2486.2003.00701.x.

Koster, R. D., P. A. Dirmeyer, and A. N. Hahmann (2002), Comparing the degree of land-atmosphere interaction in four atmospheric general circulation models, J. Hydrometeor, 3(3), 363–375, doi:10.1175/1525-7541(2002)0030363:2.0.co;2.

Krause, G. H., & Weis, E. (1984). Chlorophyll fluorescence as a tool in plant physiology. Photosynthesis Research, 5(2), 139-157. doi:10.1007/bf00028527

Krinner, G., N. Viovy, N. de Noblet-Ducoudré, J. Ogée, J. Polcher, P. Friedlingstein, P. Ciais, S. Sitch, and I. C. Prentice (2005), A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochem. Cycles, 19(1), 1–44, doi:10.1029/2003GB002199.

Lee, J., Frankenberg, C., Tol, C. V., Berry, J. A., Guanter, L., Boyce, C. K., . . . Saatchi, S. (2013). Forest productivity and water stress in Amazonia: observations from GOSAT chlorophyll fluorescence. Proceedings of the Royal Society B: Biological Sciences, 280(1761), 20130171-20130171. doi:10.1098/rspb.2013.0171

Liu, J. et al. (2014), Carbon monitoring system flux estimation and attribution: impact of ACOS-GOSAT X CO 2sampling on the inference of terrestrial biospheric sources and sinks, Tellus B, 66(0), doi:10.3402/tellusb.v66.22486.

Peters, W., A. R. Jacobson, C. Sweeney, A. E. Andrews, T. J. Conway, K. A. Masarie, J. B. Miller, L. Bruhwiler, G. Petron, and A. Hirsch (2007), An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, Proceedings of the National Academy of Sciences, 104(48), 18925.

Pierre, C., Bergametti, G., Marticorena, B., Mougin, E., Lebel, T., & Ali, A. (2011). Pluriannual comparisons of satellite-based rainfall products over the Sahelian belt for seasonal vegetation modeling. Journal of Geophysical Research, 116(D18). doi:10.1029/2011jd016115

Prince, S. D. "Satellite remote sensing of primary production: comparison of results for Sahelian grasslands 1981-1988." International Journal of Remote Sensing, vol. 12, no. 6, 1991, pp. 1301–1311., doi:10.1080/01431169108929727.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., . . . Woollen, J. (2011). MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. Journal of Climate, 24(14), 3624-3648. doi:10.1175/jcli-d-11-00015.1

Schuh, A. E. et al. (2013), Evaluating atmospheric CO 2inversions at multiple scales over a highly inventoried agricultural landscape, Global Change Biology, 19(5), 1424– 1439, doi:10.1111/gcb.12141.

Staver, A. C., S. Archibald, and S. A. Levin, 2011. The global extent and determinants of savanna and forest as alternative biome states. Science, 334, 230-232.

Tan, B., J. T. Morisette, R. E. Wolfe, F. Gao, G. A. Ederer, J. Nightingale, and J. A. Pedelty (2017), An Enhanced TIMESAT Algorithm for Estimating Vegetation Phenology Metrics From MODIS Data, IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing, 4(2), 361–371, doi:10.1109/JSTARS.2010.2075916.

van der Tol, C., J. A. Berry, P. K. E. Campbell, and U. Rascher (2014), Models of fluorescence and photosynthesis for interpreting measurements of solar-induced chlorophyll fluorescence, J. Geophys. Res. Biogeosci., 119, 2312–2327, doi:10.1002/2014JG002713.

Yang, X., J. Tang, J. F. Mustard, J.-E. Lee, M. Rossini, J. Joiner, J. W. Munger, A. Kornfeld, and A. D. Richardson (2015), Solar-induced chlorophyll fluorescence that correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest, Geophys. Res. Lett., 42, 2977–2987, doi:10.1002/ 2015GL063201.

Zhang, Y., Guanter, L., Berry, J. A., Joiner, J., Tol, C. V., Huete, A., . . . Köhler, P. (2014). Estimation of vegetation photosynthetic capacity from space-based measurements of chlorophyll fluorescence for terrestrial biosphere models. Global Change Biology, 20(12), 3727-3742. doi:10.1111/gcb.12664