Strong relationship between canopy structure and maximum light use efficiency across biomes

# Summary

* Dechant et al. previously reported a linear relationship between temporal LUE and fesc at the site level.
* Our study builds on this finding by examining the relationship between LUEmax (the maximum annual LUE) and fesc across a much broader scale - spanning 271 sites from the AmeriFlux and FluxNet datasets, encompassing diverse biomes.
* By shifting the focus to LUEmax and analyzing the patterns at the biome level, we observed a stronger and more informative relationship, with an R2 of 0.85 between LUEmax and fesc.
* A clear pattern emerged, where biomes with more open canopies, such as shrublands and savannas, exhibit lower values of both fesc and LUEmax. In contrast, biomes with denser canopies, like forests, tend to have higher values of both fesc and LUEmax.
* This relationship can be implicitly explained by the link between canopy structure, light scattering, and plant physiological mechanisms. Specifically, the increased multiple scattering in open canopies leads to more light escaping the canopy (lower fesc), while also resulting in less efficient light utilization (lower LUEmax). Denser canopies facilitate more complete light absorption and utilization, leading to higher fesc and LUEmax.
* However, to explicitly explain the observed fesc-LUEmax relationship, further mechanistic modeling work integrating canopy structure, radiative transfer, and plant ecophysiology is needed. Such an approach can provide deeper insights into the drivers of this relationship across different biomes.

# Introduction

Light use efficiency (LUE) (Monteith 1972) is the efficiency with which plants convert absorbed radiation into fixed carbon. At the leaf scale, LUE can be defined as the initial slope (quantum yield) of the light response curve (Björkman 1981). At the canopy scale, LUE is defined as the ratio of canopy gross primary production (GPP) divided by absorbed photosynthetically active radiation (APAR) (Gitelson and Gamon 2015). Maximum LUE (εMax) is an important parameter that shows the potential of a plant to convert light to carbon. At the leaf scale, εMax is a function of leaf physiology and potentially has a small variation among different types of plants. For example, Bolton and Hall (1991) showed that εMax for green plants with two photosystems under ideal conditions and bright sunlight is 13%. However, at the canopy scale and within and between biomes, the theoretical εMax is influenced by environmental factors such as temperature, water, plant physiology and canopy strucrture (Medlyn 1998; Gan et al. 2021; Pei et al. 2022) leading to an apparent LUEMax. It is expected that LUEMax varies between different biomes and types of plants.

The variation in LUEMax is of interest as it provides critical information about the ecosystem's capacity for light-to-carbon conversion. It also serves as the foundation for LUE-based models used in estimating GPP (Pei et al., 2022). The variability in LUEMax has been attributed to a range of processes, from ecological evolution (Field, 1991) to factors such as water use efficiency, nitrogen availability, light quality, pigment distribution, vegetation type, and atmospheric CO2 levels (for a review, see Gitelson and Gamon, 2015; Gan et al., 2021).

One aspect highlighted as a primary driver of changes in LUE is canopy structure (Medlyn 1998; Dechant et al. 2020; de Mattos et al. 2020). In a recent study, Dechant et al. (2020) reported a moderate (R2=0.4 to 0.6) positive relationship between LUE and canopy structure estimated through remote sensing observations. Specifically, they correlated the half-hourly and daily LUE values with the fraction of photons escaping the canopy after multiple scattering within the canopy (fesc). The variation in fesc is largely driven by the variation in canopy structure (Knyazikhin et al. 2013; Dechant et al. 2020).

While the findings of Dechant et al. provide valuable insights, there are a few challenges that limit the generalization of their results. First, the LUE is not equivalent to the LUEmax. There is evidence that LUE can vary significantly throughout the day due to changes in factors such as photosynthetically active radiation (PAR), as well as seasonally due to factors like canopy structure, stress, and environmental conditions (Gitelson, Arkebauer, and Suyker 2018; Gamon and Berry 2012). As a result, the relationship between fesc and LUE in Dechant et al. likely confounds the impact of canopy structure with other dynamic factors. In contrast, LUEmax is a more stable metric (Medlyn 1998) and better represents the inherent capacity of plants or biomes to convert light into carbon. Furthermore, the LUE-fesc relationship reported in the study is based on a limited dataset of just three crop sites. A more comprehensive, global study on the relationship between LUEmax and fesc would provide deeper and more generalizable insights.

The main objective of this study is to address the challenges mentioned and examine the relationship between fesc and LUEmax across larger spatial scales. Dechant et al. (2020) demonstrated that up to 80% of the variation in gross primary production (GPP) can be explained by:

Here, APAR represents absorbed photosynthetic radiation, which can be calculated as:

(2)

Where PAR is the photosynthetic absorbed radiation, and fPAR is the fraction of absorbed photosynthetic radiation. The variable fesc can be approximated by (Yelu et al.):

(3)

Where NIRV can be acquired by multiplying the normalized difference vegetation index (NDVI = (NIR - RED) / (NIR + RED)) by the reflectance values in the near-infrared region of the spectrum (i.e., NIRV = NDVI \* NIR). On the other hand, from the classical approach (Monteith 1972), we have:

(4)

# By comparing equations 1 and 4, we hypothesize that there is a strong relationship between fesc and LUEmax. This is because changes in canopy structure, and hence fesc, are more dominant spatially over larger domains rather than temporally over a specific site. To test this hypothesis, we utilized data from more than 270 eddy covariance sites along with remote sensing observations and investigated the relationship between the two variables across various biomes and within individual biomes.

# Materials and Methods

The eddy covariance (EC) data used in this study comes from two global EC networks, the FLUXNET2015 Dataset and AmeriFlux. Table S1 (in the supplementary material) provides more information on all the sites used in this study. Both datasets use the same pipeline to process the data (Pastorello et al. 2020). If a site was present in both datasets, we selected its data from AmeriFlux since they usually include the most recent updates compared to its counterpart in FLUXNET.

To calculate fesc, we required reflectance and fPAR data. In this study, we used MODIS products. The reflectance data was obtained from the daily MCD43A4 Version 6.1 Nadir Bidirectional Reflectance Distribution Function (BRDF)-Adjusted Reflectance product, and the fPAR data was obtained from the MCD15A3H Version 6.1 Fraction of Photosynthetically Active Radiation product. Both products have a spatial resolution of 500 meters, which we assume roughly matches the footprint of the EC towers. The fPAR product has a 4-day temporal resolution. To make it consistent with the daily reflectance and EC data, we resampled the fPAR data to daily using linear interpolation.

From the EC data, we only used GPP and PAR data that met the following criteria: 1) 20% or less of the data was gap-filled, 2) the net ecosystem exchange (NEE) uncertainties were less than 3 g C m-2 d-1, and 3) the GPP partitioned using day- and night-time methods differed by less than 3 g C m-2 d-1. We then took the mean of the day and night GPP estimates as the daily GPP (Joiner et al. 2018; Tramontana et al. 2016; Joiner and Yoshida 2020). For the reflectance data, we used the red and near-infrared (NIR) bands. We filtered the data to include only good quality observations where the main BRDF algorithm was used to process the data. Similarly, for the fPAR data, we only included observations where the main radiative transfer algorithm was used to derive fPAR, and there were no significant cloud cover issues in the scene. After the data pre-processing and quality filtering, we were left with 276 sites which include 212,552 days of data across 13 different biomes (see Table 1 for details).

Table 1. The number of site and observations for each biome.

|  |  |  |  |
| --- | --- | --- | --- |
| Biome | Abbreviation | Number of sites | Number of observations |
| Grasslands | GRA | 48 | 36969 |
| Mixed Forests | MF | 9 | 7272 |
| Evergreen Needleleaf Forests | ENF | 67 | 48402 |
| Open Shrublands | OSH | 25 | 18744 |
| Cropland/Natural Vegetation Mosaics | CVM | 1 | 270 |
| Permanent Wetlands | WET | 29 | 17134 |
| Woody Savannas | WSA | 4 | 12756 |
| Deciduous Broadleaf Forests | DBF | 38 | 31851 |
| Croplands | CRO | 38 | 24540 |
| Evergreen Broadleaf Forests | EBF | 6 | 4337 |
| Closed Shrublands | CSH | 5 | 4285 |
| Savannas | SAV | 5 | 5771 |
| Snow and Ice | SNO | 1 | 221 |

To calculate the maximum LUEmax for each site, we first calculated the daily LUE as: . We then identified the annual maximum LUE value for each site by looking for the day within a year where the daily LUE was highest. If a site had data for multiple years, we took the median of the annual maximum LUE values as the LUEmax for that site. We used Equation 3 to calculate the fraction of absorbed radiation that escapes the canopy (fesc) and chose the fesc values corresponding to the days when the LUE was at its maximum. Five sites showed fesc values exceeding 1, which is theoretically impossible. We therefore dropped these sites from the analysis. Figure 1 shows the summary distribution of these LUEMax and fesc across the 271 remaining sites.

**A graph of a graph with numbers and lines

Description automatically generated with medium confidence**

Figure 1. The distribution of LUEMax and fesc for each biome.

After the data pre-processing based on quality flags and calculating all the variables, we removed any outliers by calculating the standard deviation of each variable for each site across all the days, and then only kept the data points that were within 2 standard deviations of the mean. To examine the relationship between fesc and LUEmax across different biomes, we first grouped the LUEmax and fesc data based on their respective biomes and calculated the median value for each biome group. We also investigated the relationship between fesc and LUEmax for individual sites within each biome. Additionally, we estimated the standard error for the LUEmax values by dividing the standard deviation of the LUEmax values by the square root of the number of observations presented in each biome or site. This provided a measure of the uncertainty associated with the LUEmax estimates.

# Results

Figure 2 shows a strong linear association between the biome-level median values of LUEmax and fesc (R2 = 0.86). As fesc increases, the corresponding LUEmax also tends to increase. A clear pattern emerges, where biomes with more open canopies, such as open shrublands and savannas, generally exhibit lower values of both fesc and LUEmax. In contrast, biomes with denser canopies, such as forested ecosystems, tend to have higher values of both fesc and LUEmax. Compared to the findings of Dechant et al., the relationship observed in this study is much stronger. Indeed, when we correlated the daily LUE and fesc for all sites (except for evergreen forests and closed shrublands), the R2 values are generally less than 0.5, with an average of around 0.3 (Figure sXXX). This suggests that the temporal correlation between LUE and fesc is weak across the majority of sites.

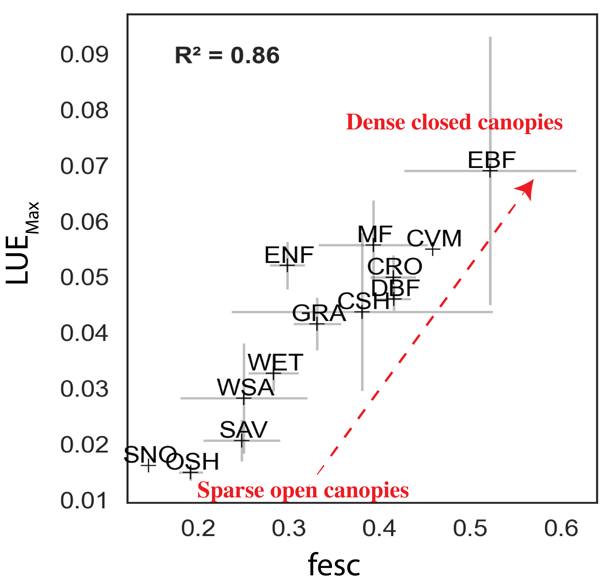


Figure 2. Relationship between fesc and LUEMax across various biomes. Refer to Table 1 for the definition of each biome abbreviation.

Figure 3 reveals a predominantly positive association between LUEmax and fesc when examined at the individual site level within each biome. In this analysis, we excluded the CVM and SNO biomes due to their single-site representation. This trend is evident across most biomes. For example, closed shrublands and woody savannas exhibit a remarkably strong linear relationship, reflected by R² values exceeding 0.96. This signifies a robust positive correlation between these variables within these specific biome types. Within the forested biomes, evergreen broadleaf sites display the strongest positive correlation. It's important to note, however, that the number of sites represented within each biome is uneven. Consequently, directly comparing the strength of these relationships across biomes might be misleading due to variations in sample sizes. Rather, the positive relationship should be paid attention to within each biome, as it suggests a consistent association between canopy structure (as captured by fesc) and LUEmax at the site level.

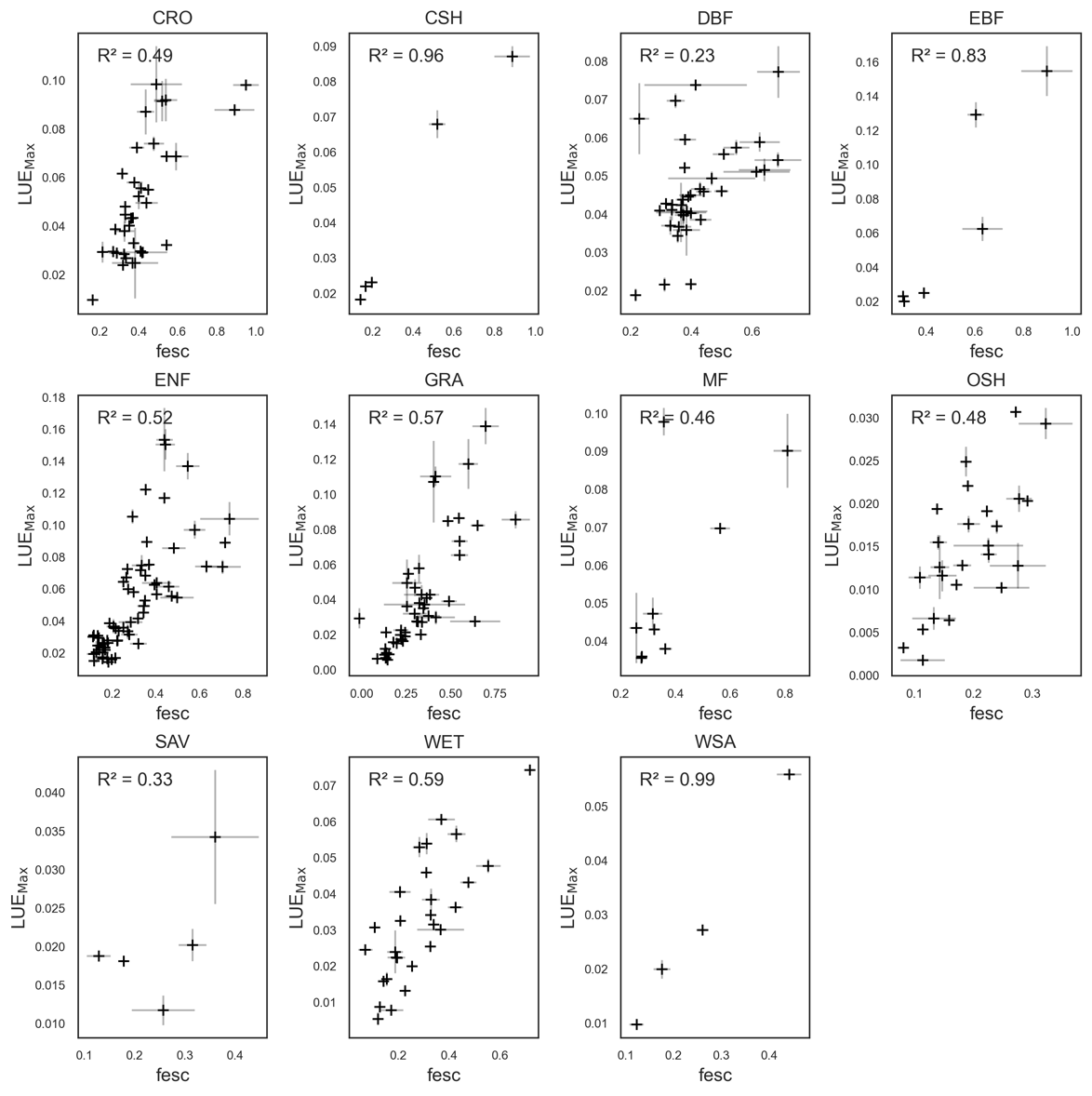


Figure 3. The relationship between fesc and LUEMax across different sites within each biome.

# Discussion

Our study builds upon prior research that found a positive link between temporal LUE and fesc (Dechant et al.). However, we demonstrate a significantly stronger and more informative relationship by focusing on two key aspects: 1) using LUEmax, the annual peak in light use efficiency, as a more robust and stable metric, and 2) analyzing the connection across both different biomes and individual sites within each biome. By comparing the relationship between biomes (Figure 2) and within biomes (Figure 3), we can see that the relationship is generally stronger at the biome level. This comparison, coupled with the clear trend from open canopies to closed canopies, highlights the fact that the relationship between fesc and LUEmax is best captured at larger, biome-level scales, confirming our hypothesis. To our knowledge, this is the first study to directly link fesc with LUEmax across a vast number of sites (>270) spanning diverse biomes. The observed strong linear association, particularly at the biome level, offers valuable insights into how variations in canopy structure influence the maximum light-use efficiency of different ecosystems. These findings underscore the importance of considering differences in biome when investigating the links between canopy structure and light use efficiency.

The pattern observed in the relationship between LUEmax and fesc can be explained from both remote sensing and plant physiological perspectives. In the lower left quadrant of Figure 2, we observe biomes with open canopies. These open canopy systems are generally not limited by light availability, resulting in a lower overall LUEmax. Additionally, the increased light penetration through open canopies potentially leads to more frequent multiple scattering of photons within the canopy, causing more photons to escape the canopy in various directions and reducing the number of photons that ultimately reach the sensor, as captured by the lower fesc values.

In contrast, denser canopies such as evergreen broadleaf forests represent closed canopy systems. In these environments, a significant portion of incoming photons is reflected from the top layers of the canopy toward the sensor without much multiple scattering in the lower canopy layers, leading to a higher fesc value. Additionally, leaves in the lower layers of such dense canopies often experience light limitation, indicating a greater capacity for converting the absorbed light into energy, resulting in higher LUEmax. The findings of this study underscore the significant link between canopy structure and optics, and canopy physiology, emphasizing the need for more explicit investigations that integrate both remote sensing observations and process-based models. Such studies could leverage 3D models of radiative transfer model with intergraded with plant physiology models, to further elucidate the mechanisms behind the observed fesc-LUEmax relationship across different biomes.

It's important to emphasize that our study is observational in nature. While we report a strong positive relationship between LUEmax and fesc, this correlation does not necessarily imply direct causation. We cannot definitively say that changes in fesc directly cause changes in LUEmax, or vice versa, without further elucidating the underlying mechanisms driving this relationship. However, the observed strong association between these variables across various biomes can still be leveraged to improve the estimation of LUEmax for various applications. For example, many LUE-based algorithms to estimate GPP, such as the MODIS GPP algorithm, often rely on fixed values of LUEmax for different plant functional types based on land cover data. Our results suggest this approach may be problematic, as LUEmax can vary significantly between sites within the same biome due to changes in canopy structure. Instead, canopy structure, and the remotely sensed fesc in particular, may be a more suitable variable for constraining estimates of LUEmax across different ecosystems.

There are several challenges and limitations associated with this study that present opportunities for future research. Firstly, while fesc is commonly used as an indicator of canopy structure, its approximation using the ratio of NIRv to fPAR has certain inherent constraints. Ideally, fesc should always be less than one, as it represents the fraction of photons escaping the canopy which cannot exceed the total incoming photons. However, during our analysis, we encountered a few sites where fesc values exceeded one. This discrepancy could stem from various factors, ranging from data quality issues to underlying assumptions in the fesc estimation. For instance, it's conceivable that a site with a closed canopy might experience some level of stress, reducing its chlorophyll content and affecting fPAR, while simultaneously exhibiting high NIRv values, leading to fesc values greater than one. Furthermore, in calculating fesc, we assume that soil does not contribute, and that red reflectance values of plants in the scene can be approximated using NDVI\*NIR (Yang and van der Tol 2018; Zeng et al. 2019). However, these assumptions may not hold true in scenarios such as sparse canopy cover, where soil can impact both NDVI and NIR. Additionally, it is important to note that in the hotspot angular region, the reflectance can potentially exceed 1 due to coherent backscattering, specular reflections, and multiple scattering effects (Hapke 1993; Verhoef and Bach 2007). This could lead to fesc values greater than 1 in the hotspot direction, even if the overall canopy reflectance is less than 1. More theoretical work is needed to fully understand the impact of various canopy elements and structural characteristics on the fesc metric.

Similarly, the definition and estimation of maximum LUE (LUEmax) poses its own challenges. In this study, we simply chose the median of annual maximum LUE values for each site as the representative LUEmax. However, this approach may overlook year-to-year variations in LUEmax due to factors such as environmental stress. Nonetheless, we assume that given the large number of sites included in our analysis, the impact of such inter-annual variations would be minimal.

Despite these limitations, we believe our results open new windows toward bridging our understanding of plant physiology and structure with their optical properties as observed from remote sensing. Further refinements in the estimation of key variables like fesc and LUEmax, as well as the integration of process-based models, can help elucidate the mechanisms underlying the observed relationships.

# References

Dechant, Benjamin, Youngryel Ryu, Grayson Badgley, Yelu Zeng, Joseph A. Berry, Yongguang Zhang, Yves Goulas, et al. 2020. “Canopy Structure Explains the Relationship between Photosynthesis and Sun-Induced Chlorophyll Fluorescence in Crops.” *Remote Sensing of Environment* 241 (May): 111733. https://doi.org/10.1016/j.rse.2020.111733.

Gamon, J.A., and J.A. Berry. 2012. “Facultative and Constitutive Pigment Effects on the Photochemical Reflectance Index (PRI) in Sun and Shade Conifer Needles.” *Israel Journal of Plant Sciences* 60 (1–2): 85–95. https://doi.org/10.1560/IJPS.60.1-2.85.

Gitelson, Anatoly A., Timothy J. Arkebauer, and Andrew E. Suyker. 2018. “Convergence of Daily Light Use Efficiency in Irrigated and Rainfed C3 and C4 Crops.” *Remote Sensing of Environment* 217 (November): 30–37. https://doi.org/10.1016/j.rse.2018.08.007.

Gitelson, Anatoly A., and John A. Gamon. 2015. “The Need for a Common Basis for Defining Light-Use Efficiency: Implications for Productivity Estimation.” *Remote Sensing of Environment* 156 (January): 196–201. https://doi.org/10.1016/j.rse.2014.09.017.

Joiner, Joanna, and Yasuko Yoshida. 2020. “Satellite-Based Reflectances Capture Large Fraction of Variability in Global Gross Primary Production (GPP) at Weekly Time Scales.” *Agricultural and Forest Meteorology* 291 (September): 108092. https://doi.org/10.1016/j.agrformet.2020.108092.

Joiner, Joanna, Yasuko Yoshida, Yao Zhang, Gregory Duveiller, Martin Jung, Alexei Lyapustin, Yujie Wang, and Compton J. Tucker. 2018. “Estimation of Terrestrial Global Gross Primary Production (GPP) with Satellite Data-Driven Models and Eddy Covariance Flux Data.” *Remote Sensing* 10 (9): 1346. https://doi.org/10.3390/rs10091346.

Knyazikhin, Yuri, Mitchell A. Schull, Pauline Stenberg, Matti Mõttus, Miina Rautiainen, Yan Yang, Alexander Marshak, et al. 2013. “Hyperspectral Remote Sensing of Foliar Nitrogen Content.” *Proceedings of the National Academy of Sciences* 110 (3): E185–92. https://doi.org/10.1073/pnas.1210196109.

Mattos, Eduardo M. de, Dan Binkley, Otavio C. Campoe, Clayton A. Alvares, and Jose L. Stape. 2020. “Variation in Canopy Structure, Leaf Area, Light Interception and Light Use Efficiency among Eucalyptus Clones.” *Forest Ecology and Management* 463 (May): 118038. https://doi.org/10.1016/j.foreco.2020.118038.

Medlyn, Belinda E. 1998. “Physiological Basis of the Light Use Efficiency Model.” *Tree Physiology* 18 (3): 167–76. https://doi.org/10.1093/treephys/18.3.167.

Monteith, J. L. 1972. “Solar Radiation and Productivity in Tropical Ecosystems.” *Journal of Applied Ecology* 9 (3): 747–66. https://doi.org/10.2307/2401901.

Tramontana, G., M. Jung, C.R. Schwalm, K. Ichii, G. Camps-Valls, B. Ráduly, M. Reichstein, et al. 2016. “Predicting Carbon Dioxide and Energy Fluxes across Global FLUXNET Sites with Regression Algorithms.” *Biogeosciences* 13 (14): 4291–4313. https://doi.org/10.5194/bg-13-4291-2016.

Yang, Peiqi, and Christiaan van der Tol. 2018. “Linking Canopy Scattering of Far-Red Sun-Induced Chlorophyll Fluorescence with Reflectance.” *Remote Sensing of Environment* 209 (May): 456–67. https://doi.org/10.1016/j.rse.2018.02.029.

Zeng, Yelu, Grayson Badgley, Benjamin Dechant, Youngryel Ryu, Min Chen, and J. A. Berry. 2019. “A Practical Approach for Estimating the Escape Ratio of Near-Infrared Solar-Induced Chlorophyll Fluorescence.” *Remote Sensing of Environment* 232 (October): 111209. https://doi.org/10.1016/j.rse.2019.05.028.