# Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence

Luis Guanter[a,1,2], Yongguang Zhang[a,1], Martin Jung[b], Joanna Joiner[c], Maximilian Voigt[a], Joseph A. Berry[d], Christian Frankenberg[e], Alfredo R. Huete[f], Pablo Zarco-Tejada[g], Jung-Eun Lee[h], M. Susan Moran[i], Guillermo Ponce-Campos[i], Christian Beer[j], Gustavo Camps-Valls[k], Nina Buchmann[l], Damiano Gianelle[m], Katja Klumpp[n], Alessandro Cescatti[o], John M. Baker[p], and Timothy J. Griffis[q]

[a]Institute for Space Sciences, Freie Universität Berlin, 12165 Berlin, Germany; [b]Department for Biogeochemical Systems, Max Planck Institute for Biogeochemistry, 07745 Jena, Germany; [c]Laboratory for Atmospheric Chemistry and Dynamics (Code 614) National Aeronautics and Space Administration Goddard Space Flight Center, Greenbelt, MD 20771; [d]Department of Global Ecology, Carnegie Institution for Science, Stanford, CA 94305; [e]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109; [f]Plant Functional Biology and Climate Change Cluster, University of Technology Sydney, Sydney, 2007, Australia; [g]Instituto de Agricultura Sostenible, Consejo Superior de Investigaciones Científicas, 14004 Córdoba, Spain; [h]Geological Sciences, Brown University, Providence, RI 02912; [i]Southwest Watershed Research, Agricultural Research Service, US Department of Agriculture, Tucson, AZ 85719; [j]Department of Applied Environmental Science and Bolin Centre for Climate Research, Stockholm University, 10691 Stockholm, Sweden; [k]Image Processing Laboratory, Universitat de València, 46980 València, Spain; [l]Agricultural Sciences, Eidgenössische Technische Hochschule Zurich, 8092 Zurich, Switzerland; [m]Sustainable Agro-ecosystems and Bioresources Department, Research and Innovation Centre, Fondazione Edmund Mach, 38010 San Michele all'Adige, Italy; [n]Grassland Ecosystem Research Unit, Institut National de la Recherche Agronomique, Clermont-Ferrand, France 63122; [o]Institute for Environment and Sustainability, Joint Research Centre, European Commission, 20127 Ispra, Italy; [p]Soil and Water Management Research, Agricultural Research Service, US Department of Agriculture, St. Paul, MN 55108; and [q]Department of Soil, Water, and Climate, University of Minnesota, St. Paul, MN 55108

Edited by Gregory P. Asner, Carnegie Institution for Science, Stanford, CA, and approved February 25, 2014 (received for review October 24, 2013)

**Photosynthesis is the process by which plants harvest sunlight to produce sugars from carbon dioxide and water. It is the primary source of energy for all life on Earth; hence it is important to understand how this process responds to climate change and human impact. However, model-based estimates of gross primary production (GPP, output from photosynthesis) are highly uncertain, in particular over heavily managed agricultural areas. Recent advances in spectroscopy enable the space-based monitoring of sun-induced chlorophyll fluorescence (SIF) from terrestrial plants. Here we demonstrate that spaceborne SIF retrievals provide a direct measure of the GPP of cropland and grassland ecosystems. Such a strong link with crop photosynthesis is not evident for traditional remotely sensed vegetation indices, nor for more complex carbon cycle models. We use SIF observations to provide a global perspective on agricultural productivity. Our SIF-based crop GPP estimates are 50–75% higher than results from state-of-the-art carbon cycle models over, for example, the US Corn Belt and the Indo-Gangetic Plain, implying that current models severely underestimate the role of management. Our results indicate that SIF data can help us improve our global models for more accurate projections of agricultural productivity and climate impact on crop yields. Extension of our approach to other ecosystems, along with increased observational capabilities for SIF in the near future, holds the prospect of reducing uncertainties in the modeling of the current and future carbon cycle.**

crop productivity | carbon fluxes | Earth observation | carbon modeling | spaceborne spectroscopy

The rapidly growing demand for food and biofuels constitutes one of the greatest challenges for humanity in coming decades (1). It is estimated that we must double world food production by 2050 to meet increasing demand (2), but the once rapid growth seen in the "green revolution" has stalled, and even past advances are threatened by climate change (3–5). Much of past yield improvement has focused on increases in the harvest index and resistance to pests. However, all else being equal, the quantity of photosynthesis places an upper limit on the supply of food and fuels from our agricultural systems.

Ironically, we currently have very limited ability to assess photosynthesis of the breadbaskets of the world. Agricultural production inventories provide important information about crop productivity and yields (6–8), but these are difficult to compare between regions and lag actual production. Carbon cycle models, based on either process-oriented biogeochemistry or semiempirical data-driven approaches, have been used to understand the controls and variations of global gross primary production (GPP, equivalent to ecosystem gross photosynthesis) (9) and to investigate the climate impact on crop yields (10). However, uncertainty associated with inaccurate input data and much simplified process descriptions based on the plant functional type concept severely challenge the application of these models to agricultural systems. Recent model intercomparisons conducted as part of the North American Carbon Project found that GPP estimates for crop areas varied by a factor of 2 (11). The best available estimates of GPP of crop systems are from direct measurement of carbon dioxide exchange by so-called flux towers over agricultural fields (12). However, these generally sample small areas (<1 $km^2$) and are concentrated in North America and Europe.

Remote sensing of reflectance-based vegetation parameters has been used in the last decades to monitor agricultural

## Significance

Global food and biofuel production and their vulnerability in a changing climate are of paramount societal importance. However, current global model predictions of crop photosynthesis are highly uncertain. Here we demonstrate that new space-based observations of chlorophyll fluorescence, an emission intrinsically linked to plant biochemistry, enable an accurate, global, and time-resolved measurement of crop photosynthesis, which is not possible from any other remote vegetation measurement. Our results show that chlorophyll fluorescence data can be used as a unique benchmark to improve our global models, thus providing more reliable projections of agricultural productivity and climate impact on crop yields. The enormous increase of the observational capabilities for fluorescence in the very near future strengthens the relevance of this study.

Author contributions: L.G., Y.Z., M.J., and J.A.B. designed research; L.G., Y.Z., M.V., A.R.H., P.Z.-T., J.-E.L., M.S.M., and G.P.-C. performed research; L.G., Y.Z., M.J., J.J., C.B., G.C.-V., N.B., D.G., K.K., A.C., J.M.B., and T.J.G. contributed new reagents/analytic tools; L.G., Y.Z., J.J., M.V., and C.F. analyzed data; and L.G., Y.Z., J.A.B., C.F., and A.R.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]L.G. and Y.Z. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: luis.guanter@wew.fu-berlin.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1320008111/-/DCSupplemental.

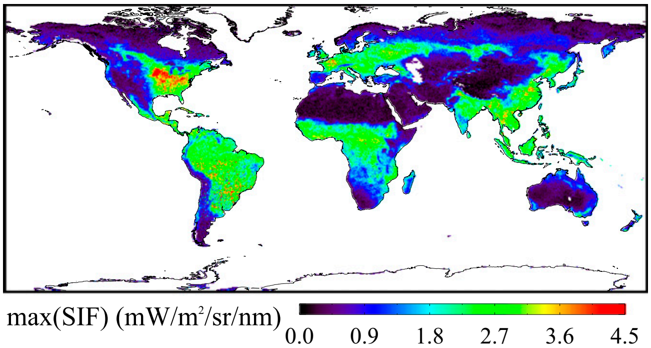



**Fig. 1.** Global map of maximum monthly sun-induced chlorophyll fluorescence (SIF) per 0.5° grid box for 2009. SIF retrievals are performed in a spectral window centered at 740 nm (see *Materials and Methods* and *SI Appendix, SIF Retrievals*). This maps illustrates the outstanding SIF signal detected at the US CB, which shows the highest SIF return of all terrestrial ecosystems. The maximum SIF over the largest part of the US CB region is detected in July.

resources (e.g., refs. 13, 14). The signal of the so-called spectral vegetation indices convolves leaf chlorophyll content, biomass, canopy structure, and cover (15, 16), such that estimating actual productivity from vegetation indices requires additional data and modeling steps, both associated with considerable uncertainty. Complementing reflectance-based indices, global space-based estimates of sun-induced chlorophyll fluorescence (SIF) became available recently. SIF is an electromagnetic signal emitted in the 650- to 850-nm spectral window as a by-product of photosynthesis (e.g., refs. 17–19). The first global maps of SIF were derived using data from the Greenhouse Gases Observing Satellite (GOSAT) (20–23). Despite the complicated photosynthesis-SIF relationships and the convolution of the signal with canopy structure (16), SIF retrievals showed high correlations with data-driven GPP estimates at global and annual scales (21, 22), as well as intriguing patterns of seasonal drought response in Amazonia (24, 25). Recently, a global SIF data set with better spatial and temporal sampling than that from GOSAT was produced using spectra from the Global Ozone Monitoring Experiment-2 (GOME-2) instrument onboard the MetOp-A platform (26) (see *SI Appendix, SIF Retrievals*).

Our attention is drawn to the remarkably high SIF returns from the US Corn Belt (CB) region (Fig. 1). This highly productive area (Fig. 2*D*) accounts for >40% of world soybean and corn production (30). We hypothesize that the high SIF indicates very high GPP for this area and report here on studies that compare SIF retrievals to GPP models and flux tower data with the aim of gaining a unique global perspective on crop photosynthesis.

## Results and Discussion

Looking at the spatial patterns of the maximum monthly gross carbon uptake from model results in the north temperate region (Fig. 2), we find a generally good agreement between the data-driven approach (27), that relies on data from a global network of micrometeorological tower sites (FLUXNET) (12), and the median of 10 state-of-the-art global dynamic vegetation models from the Trendy ("Trends in net land-atmosphere carbon exchange over the period 1980–2010") project (28, 29), the former showing somewhat larger values in a small region of the US CB (Fig. 2 *A* and *B*) (see *SI Appendix, Model-Based GPP Data*). It must be stated that the Trendy models do not include explicit crop modules, so the results from our comparisons with process-based models are intended to illustrate the potential impact of such crop-specific modules on simulations over agricultural regions. The SIF measurements, on the other hand, show large differences between the US CB and the cropland and grassland areas in Western Europe, with much enhanced SIF in the US CB (Fig. 2*C*). This pattern is roughly consistent with the distribution of C4 crops in the area, predominantly corn fields (Fig. 2*D*). Is the photosynthesis signal in the SIF retrievals disturbed by other factors, or is the US CB indeed much more productive than any area in Western Europe, which is not captured by the carbon models?

We compare year-round monthly means of flux tower-based GPP estimates at cropland and grassland sites in the United States and Europe with SIF retrievals, GPP estimates from carbon models, and spectral reflectance indices (Figs. 3 and 4 and *SI Appendix, Comparison of Flux Tower-Based GPP with Model GPP, SIF and Vegetation Indices*). Data-driven model GPP data are from the statistical model developed at the Max Planck Institute for Biogeochemistry (MPI-BGC) (27) (Fig. 3*B*) and the semiempirical moderate resolution imaging spectroradiometer (MODIS) MOD17 GPP model (31) (*SI Appendix,* Fig. S4). The same ensemble of 10 land surface models (28, 29) is used to evaluate the performance of process-based models (Fig. 3*C*). We present the comparisons in Fig. 3 without including the European cropland sites, as we want to illustrate the strong differences

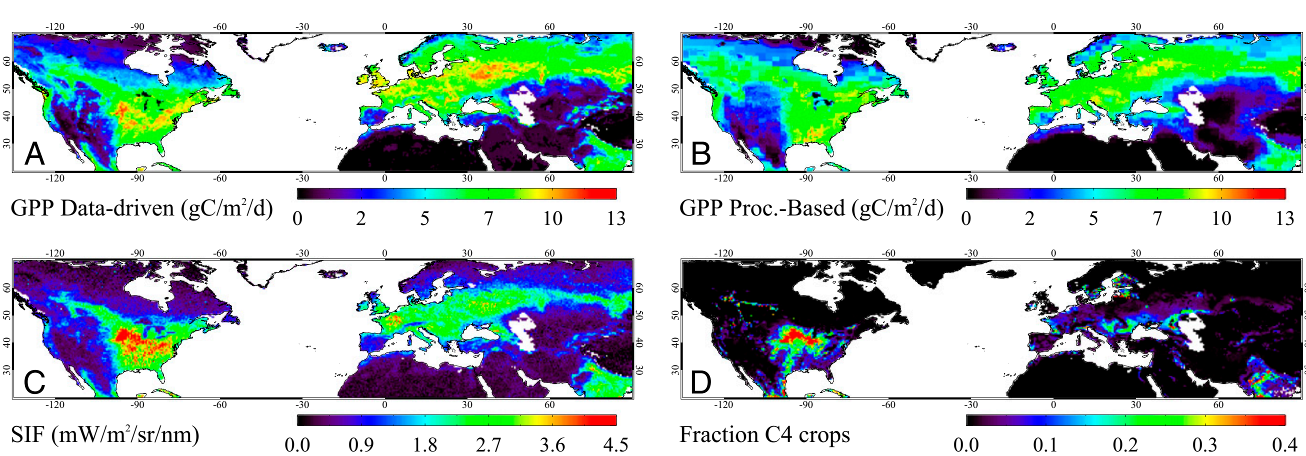



**Fig. 2.** Spatial patterns of maximum monthly gross primary production (GPP) per 0.5° grid box for 2009 from data-driven (*A*) and process-based (*B*) models together with maximum monthly SIF at 740 nm (*C*). The fraction of C4 crop area (mostly corn in this region) depicts the approximate area of the US Corn Belt (*D*). The data-driven GPP data correspond to the MPI-BGC model (27), the process-based GPP corresponds to the median of an ensemble of 10 global dynamic vegetation models from the Trendy ("Trends in net land-atmosphere carbon exchange over the period 1980–2010") project (28, 29), and SIF was retrieved from GOME-2 satellite measurements (26). The fraction of C4 crop data are described in Ramankutty et al. (6).

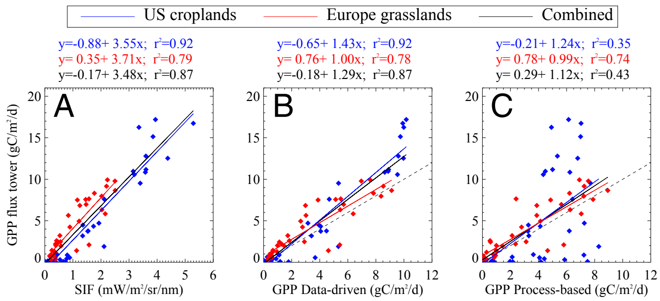



**Fig. 3.** Comparison of monthly mean GPP estimates at cropland flux tower sites in the US Corn Belt and grassland sites in Western Europe. Flux tower GPP estimates are compared with sun-induced fluorescence (SIF) observations at 740 nm (*A*) and with GPP estimates from the MPI-BGC data-driven model (27) (*B*) and from process-based models [median of an ensemble of 10 dynamic global vegetation models (28, 29)] (*C*). Each symbol depicts a monthly average for a 0.5° grid box and those months in the 2007–2011 period for which flux tower data were available (see *SI Appendix*, Table S1). The *P* value is <0.01 in all of the comparisons. The dashed line in *B* and *C* represents the 1:1 line. Similar comparisons but including also Western Europe cropland sites are provided in *SI Appendix*, Fig. S4.

between cropland and grassland GPP over the most homogeneous ecosystems (the European cropland sites are highly fragmented, which may not be properly sampled by the 0.5° resolution at which we can grid the GOME-2 SIF retrievals; see *SI Appendix, SIF Retrievals*). The comparison including all types of cropland and grassland sites is provided in *SI Appendix*, Fig. S4.

We find that the peak monthly mean GPP derived from the flux tower data in some of the US CB sites is very high (>15 $gC{\cdot}m^{-2}{\cdot}d^{-1}$), whereas for the grassland sites, monthly mean GPP never exceeds 10 $gC{\cdot}m^{-2}{\cdot}d^{-1}$ (Fig. 3). Process-based GPP estimates compare well with the tower-based estimates over the grassland sites but show a poor correlation over the US CB (Fig. 3*C*). Concerning the data-driven models, there is a clear nonlinear relation between flux tower and model GPP, showing that models strongly underestimate GPP at cropland sites with high fluxes. A piece-wise linear approximation reveals that deviations from the linear relation appear at GPP > 10 $gC{\cdot}m^{-2}{\cdot}d^{-1}$ for the MPI-BGC estimates (Fig. 3*B*) and at GPP > 8 $gC{\cdot}m^{-2}{\cdot}d^{-1}$ for the MODIS MOD17 (*SI Appendix*, Fig. S4). We observe that data-driven models produce similar peak GPP values for both grasslands and croplands, and that grasslands have even a higher GPP than croplands in results from the process-based models, which is not reflected by tower-based estimates. We find that SIF values exhibit a much stronger linear relationship with tower GPP at these cropland and grassland sites (Fig. 3*A*), and that a single linear model is able to link SIF with GPP for both croplands and grasslands. On the other hand, the good agreement between the model- and tower-based GPP estimates at grassland sites, including similar peak values, suggests that the direct comparison of flux tower data (typical footprint of <1 $km^2$) with SIF retrievals and model data at 0.5° is acceptable for these sites.

Hence, the comparisons in Fig. 3 support the following claims: (*i*) SIF captures high photosynthetic signals that are observed from flux towers in the US CB, and (*ii*) the models underestimate crop GPP, in particular for the highly productive crop sites at the US CB. The low correlation between the crop GPP estimates by the process-based models at the US CB sites may be explained by the lack of specific crop modules in the Trendy model ensemble. Concerning the underestimation of crop GPP by data-driven models, it can be argued that these cannot capture the complex dynamics required to link stable and structurally driven vegetation indices derived from remote sensing data with a highly variable physiological measure such as crop photosynthesis. On the other hand, those reflectance-based indices usually underestimate "greenness" for very dense crop canopies with high green biomass levels, such as cultivars with high fertilizer levels. This can lead to the underestimation of GPP by the data-driven models constrained by those vegetation indices.

The same flux tower-based GPP data set is compared with SIF retrievals and the enhanced vegetation index (EVI) extracted from the MODIS MOD13C2 product (15) in Fig. 4. This comparison illustrates that spectral reflectance indices, similar to the GPP models, do not scale linearly with GPP for these biomes despite the good representation of the temporal patterns: The highest EVI values for grassland sites are close to the values for some of the cropland sites, whereas GPP is very different. On the other hand, it is difficult to find a global baseline value for EVI to indicate the total absence of green vegetation activity. The minimum EVI value depends on the soil nature and especially on the presence of snow (32), which can be observed in the relatively high variability of EVI in the months in which no photosynthetic activity is observed (Fig. 4 *C* and *D*). This poses a problem for the identification of start- and end-of-season times in phenological studies based on reflectance-based remote sensing data (32). The SIF observations, in turn, drop to zero following photosynthesis, which provides an unambiguous signal of photosynthetical activity.

The linear relationship between SIF data and flux tower GPP observed in Fig. 3*A* may be rationalized by considering that

$$\mathrm{GPP} = \mathrm{PAR} \times \mathrm{fPAR} \times \mathrm{LUE_P}, \tag{1}$$

where PAR is the flux of photosynthetically active radiation received, fPAR is the fractional absorptance of that radiation, and $\mathrm{LUE_P}$ is the efficiency with which the absorbed PAR is used in photosynthesis (33). SIF may be similarly conceptualized as

$$\mathrm{SIF}(\lambda) = \mathrm{PAR} \times \mathrm{fPAR} \times \mathrm{LUE_F}(\lambda) \times f_{esc}(\lambda), \tag{2}$$

where λ is the spectral wavelength (~740 nm in our GOME-2 retrievals; see *Materials and Methods* and *SI Appendix, SIF Retrievals*), $\mathrm{LUE_F}$ is a light-use efficiency for SIF (i.e., the fraction of absorbed PAR photons that are re-emitted from the canopy as SIF photons at wavelength λ), and $f_{esc}(\lambda)$ is a term accounting for the fraction of SIF photons escaping from the canopy to space. These equations can be combined making the dependence on light implicit,

$$\mathrm{GPP} \approx \mathrm{SIF}(\lambda) \times \frac{\mathrm{LUE_P}}{\mathrm{LUE_F}(\lambda)}, \tag{3}$$

where we assume $f_{esc}(\lambda) \approx 1$ because of the low absorptance of leaves in the near-infrared wavelengths at which we perform the

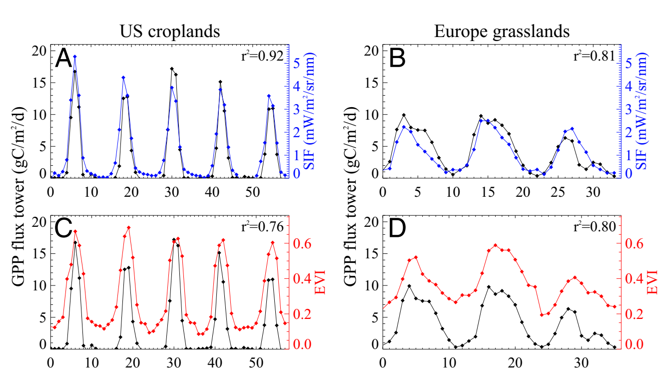



**Fig. 4.** Time series of flux tower-based GPP compared with SIF retrievals (*A* and *B*) and the MODIS MOD13C2 EVI (*C* and *D*) for the same cropland and grassland sites and spatiotemporal averages as in Fig. 3 (monthly averages in 0.5° grid boxes and the 2007–2011 period). SIF and EVI are plotted with the same vertical scale for cropland and grassland sites.

SIF retrievals and the relatively simple plant structure and high leaf area index of grasses and crops (34).

Empirical studies at the leaf and canopy scale indicate that the two light-use efficiency terms tend to covary under the conditions of the satellite measurement (35–37). Hence, the SIF data should provide information on both the light absorbed and the efficiency with which it is being used for photosynthesis. Vegetation indices derived from reflectance measurements from spaceborne instruments such as MODIS (15) and knowledge of the solar angle and atmospheric condition can be used to estimate PAR × fPAR (Eq. **1**), but $LUE_P$ is a free parameter. These data from the CB are consistent with $LUE_P$ being much higher for intensively managed crops than for native grasslands or less managed crops.

Based on the linear relationship obtained from the comparison of SIF with tower-based GPP at all of the US and Western Europe cropland and grassland flux tower sites [GPP(SIF) = −0.10 + 3.72 × SIF; see *SI Appendix, Comparison of Flux Tower-Based GPP with Model GPP, SIF and Vegetation Indices* and *Derivation of Spatially-Explicit Crop GPP Estimates*], we have produced unique global estimates of annual crop GPP. Even though tower data outside the US CB and Western Europe were not available for the derivation of the empirical GPP−SIF relationship, we assume it to hold for all of the ecosystems in which GPP is driven by canopy chlorophyll content such as croplands and grasslands (14). We have compared our SIF-based crop GPP estimates with the GPP predicted by ensembles of state-of-the-art data-driven (9) and process-based (28, 29) biogeochemistry models (see *SI Appendix, Model-Based GPP Data*). We evaluate the consistency of the different GPP estimates with the agricultural yield statistics from the National Agriculture Statistics Service of the US Department of Agriculture (USDA NASS) (38) (only North America, years 2006–2008) and the data set by Monfreda et al. (7) (global coverage, year 2000). These inventories provide large-scale cropland net primary production (NPP, biomass production by plants) estimates by combining national, state, and county-level census statistics with maps of cropland areas (see *SI Appendix, NPP Data from Agricultural Inventories*).

The comparison between our annual crop GPP estimates and the NPP from the USDA NASS inventory at the US CB shows that SIF-based GPP estimates are, similar to the flux tower comparisons, more linearly related to the inventory-based NPP than the model GPP (Fig. 5). Again, data-driven GPP estimates show a strongly nonlinear relationship with the inventory-based NPP, whereas the comparison with the process-based GPP estimates presents more scatter compared with the SIF-based and the data-driven estimates. The same conclusions hold for the comparison of the different GPP estimates over the US CB and Western Europe with the NPP data set from Monfreda et al. (7) (see *SI Appendix, NPP Data from Agricultural Inventories*). Assuming that annual GPP and NPP covary linearly across the entire US CB area, this result confirms our initial statement that GPP models substantially underestimate the photosynthetic uptake of highly productive crops. However, it is challenging to relate GPP and yield-based NPP estimates in a quantitative way, as it is difficult to account for heterogeneous land cover given the coarse resolution of current SIF retrievals. For example, much of Northern Europe is a mosaic of forests (which have low SIF) and agricultural fields. This may partly explain the apparently lower productivity of European agricultural regions.

Continuing the comparison of model estimates to SIF-based crop GPP over the globe (Figs. 6 and 7 and *SI Appendix, Derivation of Spatially-Explicit Crop GPP Estimates*), spatial patterns of SIF-based crop GPP estimates differ from data-driven models by 40–60% in the US CB area and by 50–75% in some regions of the Indo-Gangetic Plain, the North China Plain, and the Sahel belt in Africa. Smaller differences within 0–10% are found in Europe. In terms of area-integrated annual GPP estimates (*SI Appendix*, Table S2), the largest differences are found in the US CB region (+43% for the data-driven models and +18% for the process-based models) and the Indo-Gangetic Plain (+55% and +39%, respectively). A remarkable difference of −38% is also obtained

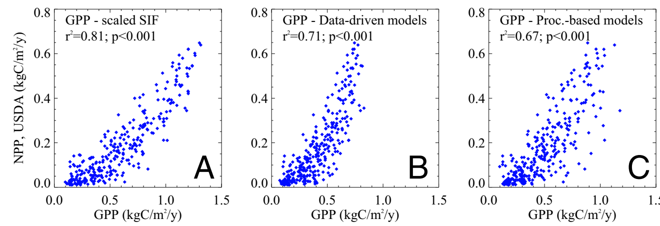



**Fig. 5.** Comparison of net primary production (NPP) estimates over the US Corn Belt (35°N–50°N, 80°W–105°W) from the USDA agricultural inventory (8) with crop GPP estimates from SIF retrievals (*A*) and data-driven and process-based model ensembles (*B* and *C*). Points correspond to 1° grid boxes with fraction of cropland area higher than 20%. GPP and NPP values are given in per-total-area units (see *SI Appendix, NPP Data from Agricultural Inventories*). The squared Pearson's correlation coefficient $r^2$ and the *P* value of the comparisons are shown. An analogous comparison with the inventory-based NPP from Monfreda et al. (7), which also includes Western Europe, can be found in *SI Appendix, NPP Data from Agricultural Inventories*.

between the SIF- and the process-based model estimates in the cropland areas between Brazil and Argentina. This area is often specified in biogeochemistry models as C4 grasslands, which have higher productivity than the C3 grasslands. Despite the relatively important local differences, the global cropland GPP estimated from SIF is in excellent agreement with the data-driven models (17.04 ± 0.19 $PgC \cdot y^{-1}$ and 17 ± 4 $PgC \cdot y^{-1}$, respectively), whereas a difference about −12% is found with the process-based models (global cropland GPP of 20 ± 9 $PgC \cdot y^{-1}$). These annual GPP numbers must be compared with the 14.8 $PgC \cdot y^{-1}$ given by Beer et al. (9) for croplands, and 123 $PgC \cdot y^{-1}$ for the total of all biomes.

Time series of SIF- and model-based crop GPP over some selected agricultural regions give insight into the differences observed in the annual GPP estimates (Fig. 7). The variation range of the monthly GPP estimates from SIF observations agrees well with the estimates from data-driven models in all of the selected cropland regions, which supports the consistency of our approach of scaling SIF to GPP using direct comparisons between GOME-2 SIF data and flux tower-based GPP. Also, the seasonal variations of data-driven and SIF-based GPP estimates are in general very consistent in all regions, and especially in Western Europe and China (Fig. 7 *B*–*D*). Estimates over the US CB and the Indo-Gangetic Plain also show the same phenological trends, but the SIF-based GPP estimates over the US CB are systematically higher than data-driven estimates by about 20% throughout the year (Fig. 7*A*). Over India, both GPP estimates coincide for the so-called *Rabi* crops sown in winter and harvested in the spring, but SIF-based GPP is about 40% higher than data-driven GPP for the *Kharif* or monsoon crops sown around June and harvested in autumn (Fig. 7*C*). This large difference in the estimated crop GPP over India in autumn explains the time shift of the global SIF-based crop GPP with respect to the data-driven models (Fig. 7*F*). On the other hand, the tested process-based models from the Trendy ensemble compare very well with data-driven models and SIF over the Western Europe region despite the lack of crop-specific modules in the Trendy models. We hypothesize that this is due to the fact that West European crops mostly follow the seasonality of grasslands, by which crops are often represented in the models. However, these models fail to describe crop phenology at the other regions and, more significantly, the multiple cropping in China and India. A time shift of the peak GPP estimates at the US CB with respect to SIF-based and data-driven GPP can be explained by modeling uncertainties associated to irrigation and also by the fact that sowing and harvesting time in the US CB is different from the lifetime of natural grassland (peak in June), as opposed to Western Europe. Also, process-based models substantially underestimate the peak GPP values for the US CB, India, and China regions, and tend to overestimate GPP in South America, which explains the

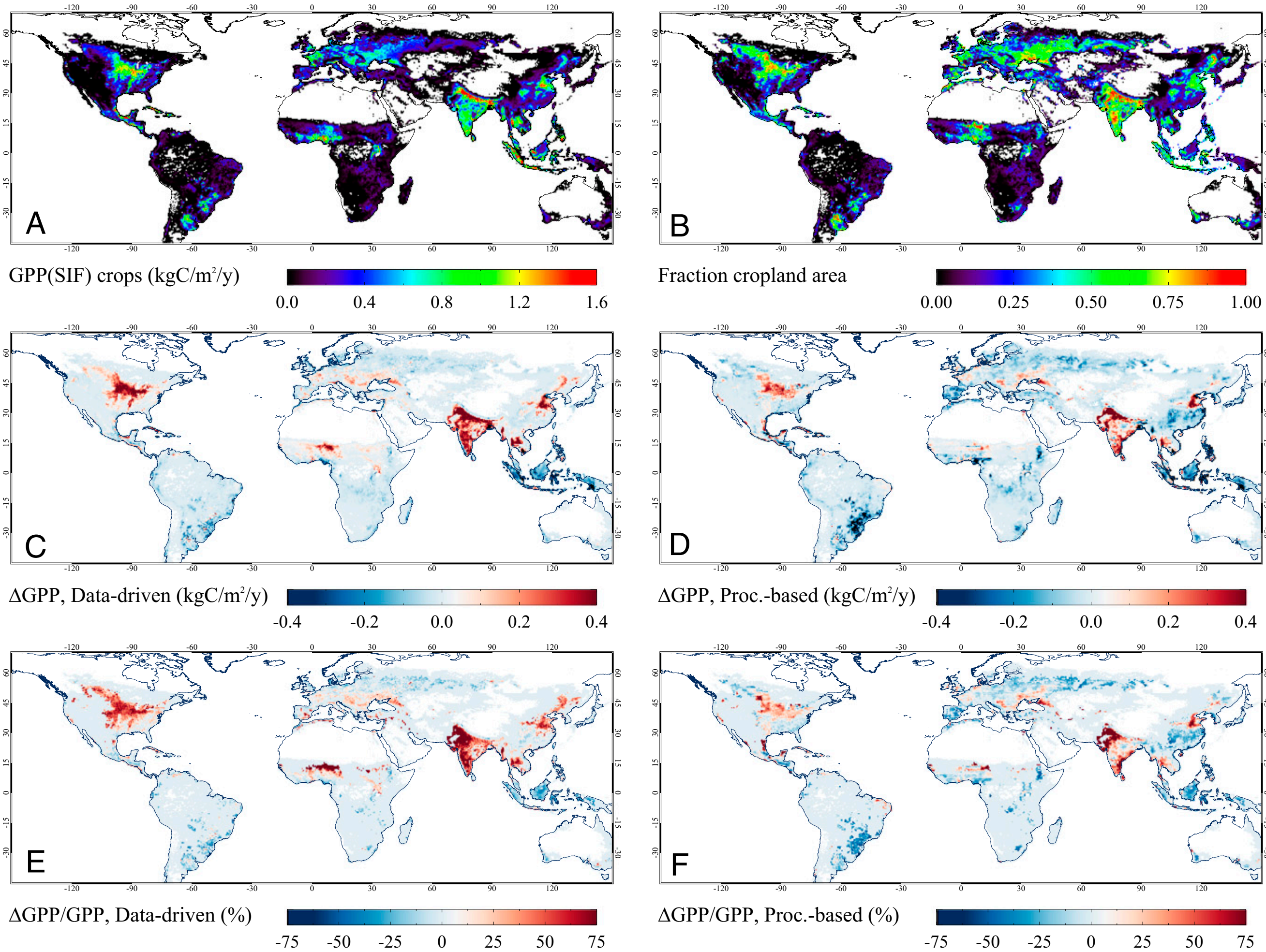



**Fig. 6.** Spatial details of the annual SIF-based crop GPP estimates over cropland areas (*A*), fraction of cropland area per grid box (*B*), and absolute and relative differences between annual SIF-based crop GPP estimates and the output of data-driven models (*C* and *E*) and process-based models (*D* and *F*). Spatially explicit GPP is derived through the scaling of SIF retrievals with the relationship $\mathrm{GPP(SIF)} = -0.10 + 3.72 \times \mathrm{SIF}$ (see *SI Appendix, Derivation of Spatially-Explicit Crop GPP Estimates*). Cropland GPP is given in per-total-area units. The absolute difference ΔGPP is calculated as GPP(SIF) − GPP(model), and the relative difference is calculated as ΔGPP over GPP(model).

spatial patterns observed in the annual GPP comparisons in Fig. 6. These results illustrate the need for specific crop modules in global dynamic vegetation models.

Considering the growing pressure on agricultural systems to provide for an increasing food and biofuel demand in the world, a global, time-resolved, and accurate analysis of crop productivity is critically required. Crop-specific models or improved process-based biogeochemistry models including explicit crop modules could provide projections of agricultural productivity and climate impact on crop yields (e.g., refs. 39–41). However, local information such as meteorology, planting dates and cultivar choices, irrigation, and fertilizer application are needed. In this

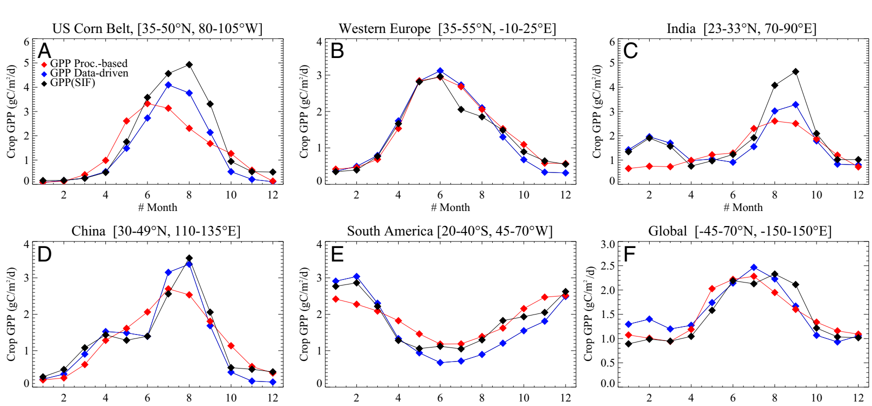



**Fig. 7.** (*A*–*F*) Time series of monthly crop GPP derived from SIF retrievals, process-based models, and data-driven models over different cropland regions in 2009. GPP area averages are weighted by the fraction of cropland area per grid box. Data-driven GPP corresponds to the MPI-BGC data-driven model (27). Process-based GPP estimates are calculated as the median of the monthly GPP estimates from the Trendy process-based model ensemble (28, 29) (see also *SI Appendix*, Table S2).

work, we have demonstrated that spaceborne SIF retrievals can provide realistic estimates of photosynthetic uptake rates over the largest crop belts worldwide without need of any additional information. This finding indicates that SIF data can help us improve our current models of the global carbon cycle, which we have shown to substantially underestimate GPP in some large agricultural regions such as the US CB and the Indo-Gangetic Plain. The launch of the Orbiting Carbon Observatory-2 and the Sentinel 5-Precursor satellite missions in 2014 or 2015 will enormously improve the observational potential for SIF, up to a 100-fold increase in spatiotemporal resolution (42, 43). This will especially benefit measurements over the typically fragmented agricultural areas, which suggests that SIF-based estimates of crop photosynthesis will soon become a unique data set for both an unbiased monitoring of agricultural productivity and the benchmarking of carbon cycle models.

## Materials and Methods

We have used monthly averages of SIF retrievals (26) from the GOME-2 instrument onboard the MetOp-A platform to produce unique estimates of global cropland GPP. GOME-2 SIF retrievals are performed in the 715- to 758-nm spectral window. Single retrievals are quality-filtered and aggregated in a 0.5° grid. The GOME-2 SIF data set used in this study covers the 2007–2011 time period (see *SI Appendix, SIF Retrievals*).

Ensembles of process-based and data-driven biogeochemistry models have been analyzed to assess the ability of global models to represent crop GPP (see *SI Appendix, Model-Based GPP Data*). The process-based model ensemble comprises the 10 global dynamic vegetation models (CLM4C, CLM4CN, HYLAND, LPJ, LPJ-GUESS, OCN, Orichidee, SDGVM, TRIFFID, and VEGAS) included in the Trends in net land carbon exchange over the period 1980–2010 (Trendy) project (28, 29). It must be noted that these models do not include explicit crop modules. The data-driven model ensemble consists of the MTE1, MTE2, ANN, KGB, and LUE models used by Beer et al. (9). In addition, monthly GPP estimates from the MPI-BGC data-driven model (27), which corresponds to the MTE1 in the data-driven model ensemble, and the MODIS GPP product (MOD17) (31) have been compared with monthly flux tower-based GPP over croplands and grasslands to evaluate the ability of data-driven models to reproduce GPP at those biomes. Cropland GPP is calculated from the SIF observations and the model ensembles as the product of the total GPP in each 0.5° grid box by the fraction of cropland area given by Ramankutty et al. (6) (see *SI Appendix, Derivation of Spatially-Explicit Crop GPP Estimates*). EVI data in Fig. 4 and *SI Appendix, Comparison of Flux Tower-Based GPP with Model GPP, SIF and Vegetation Indices*, have been extracted from the MODIS MOD13C2 product (15).

Flux tower-based GPP estimates covering the 2007–2011 period were extracted from 14 sites in Midwest United States and Western Europe. Sites correspond to the Ameriflux and the European Fluxes Database networks. Only the most spatially homogeneous sites have been selected to enable direct comparisons with the SIF observations and the GPP model outputs available in 0.5° grid cells. The relationship GPP = −0.1 + 3.72 × GPP derived from the comparison of GOME-2 monthly SIF composites with flux tower GPP data has been used to scale SIF to GPP (see *SI Appendix, Comparison of Flux Tower-Based GPP with Model GPP, SIF and Vegetation Indices*).

Large-scale NPP estimates have been derived from the USDA-NASS (38) and Monfreda et al. (7) agricultural inventory data sets. The USDA inventory covers North America and the 2006–2008 period. It is based on a statistical method to upscale county-level crop NPP data from the USDA National Agricultural Statistics Service (8, 38). The inventory by Monfreda et al. (7) is for 2000. It is based on the aggregation of 175 crop classes in a 5 min by 5 min grid. Inventory-based NPP is converted from per-harvested-area to per-total-area units through scaling by the fraction of harvested area, following Monfreda et al. (7) (see *SI Appendix, NPP Data from Agricultural Inventories*).

**ACKNOWLEDGMENTS.** We thank T. Meyers (National Oceanic and Atmospheric Administration Air Resources Laboratory), D. Cook and R. Matamala (Argonne National Laboratory), A. Suyker (University of Nebraska), C. Bernhofer (Technische Universität Dresden), Z. Nagy (Szent István University), M. Aubinet (Université de Liège), W. Kutsch (Johann Heinrich von Thuenen Institut), and K. Schneider (University of Cologne) for kindly providing eddy covariance data. We acknowledge C. Monfreda (Arizona State University), P. H. Verburg (Vrije Universiteit University Amsterdam), and N. Ramankutty (McGill University) for the crop fraction and NPP data sets and/or advice on their use, Eumetsat for the GOME-2 data, the Trendy project for the process-based model runs, and the USDA NASS for their agricultural inventory data. We also thank the two anonymous reviewers and Dr. Asner for their valuable suggestions and comments. MODIS MOD17 GPP data were downloaded from the server of the Numerical Terradynamic Simulation Group at the University of Montana, MODIS MOD13 data were obtained from the MODIS Land Processes Distributed Active Archive Center archive, and MERIS-MTCI from the Infoterra Ltd server. This work used eddy covariance data acquired by AmeriFlux and GHG-Europe. The work by L.G., Y.Z., and M.V. has been funded by the Emmy Noether Programme (GlobFluo project) of the German Research Foundation. J.J. is supported by the National Aeronautics and Space Administration (NASA) Carbon Cycle Science program (NNH10DA001N) and G.P.-C. is supported by NASA Soil Moisture Active Passive Science Definition Team (08-SMAPSDT08-0042). We also thank the W. M. Keck Foundation for funding the New Methods to Measure Photosynthesis from Space workshop held at the Caltech Keck Institute for Space Studies.

1. Foley JA, et al. (2011) Solutions for a cultivated planet. *Nature* 478(7369):337–342.
2. Tilman D, Balzer C, Hill J, Befort BL (2011) Global food demand and the sustainable intensification of agriculture. *Proc Natl Acad Sci USA* 108(50):20260–20264.
3. Wheeler T, von Braun J (2013) Climate change impacts on global food security. *Science* 341(6145):508–513.
4. Battisti DS, Naylor RL (2009) Historical warnings of future food insecurity with unprecedented seasonal heat. *Science* 323(5911):240–244.
5. Lobell DB, et al. (2013) The critical role of extreme heat for maize production in the United States. *Nat Clim Change* 3(5):497–501.
6. Ramankutty N, et al. (2008) Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochem Cycles* 22:GB1003, 10.1029/2007GB002952.
7. Monfreda C, et al. (2008) Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochem Cycles* 22:GB1022, 10.1029/2007GB002947.
8. Hicke JA, et al. (2004) Cropland area and net primary production computed from 30 years of USDA agricultural harvest data. *Earth Interact* 8(10):1–20.
9. Beer C, et al. (2010) Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate. *Science* 329(5993):834–838.
10. Ciais P, et al. (2005) Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature* 437(7058):529–533.
11. Schaefer K, et al. (2012) A model-data comparison of gross primary productivity: Results from the North American Carbon Program site synthesis. *J Geophys Res* 117:G03010, 10.1029/2012JG001960.
12. Baldocchi D, et al. (2001) FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bull Am Meteorol Soc* 82(11):2415–2434.
13. Moran MS, et al. (1997) Opportunities and limitations for image-based remote sensing in precision crop management. *Remote Sens Environ* 61(3):319–346.
14. Gitelson AA, et al. (2006) Relationship between gross primary production and chlorophyll content in crops: Implications for the synoptic monitoring of vegetation productivity. *J Geophys Res* 111:D08S11, 10.1029/2005JD006017.
15. Huete A, et al. (2002) Overview of the radiometric and biophysical performance of the modis vegetation indices. *Remote Sens Environ* 83(1-2):195–213.
16. Knyazikhin Y, et al. (2013) Hyperspectral remote sensing of foliar nitrogen content. *Proc Natl Acad Sci USA* 110(3):E185–E192.
17. Baker NR (2008) Chlorophyll fluorescence: A probe of photosynthesis in vivo. *Annu Rev Plant Biol* 59:89–113.
18. Meroni M, et al. (2009) Remote sensing of solar induced chlorophyll fluorescence: Review of methods and applications. *Remote Sens Environ* 113(10):2037–2051.
19. Rascher U, et al. (2009) CEFLES2: The remote sensing component to quantify photosynthetic efficiency from the leaf to the region by measuring sun-induced fluorescence in the oxygen absorption bands. *Biogeosciences* 6(7):1181–1198.
20. Joiner J, et al. (2011) First observations of global and seasonal terrestrial chlorophyll fluorescence from space. *Biogeosciences* 8(3):637–651.
21. 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* 38:L17706, 10.1029/2011GL048738.
22. Guanter L, et al. (2012) Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements. *Remote Sens Environ* 121:236–251.
23. Joiner J, et al. (2012) Filling-in of near-infrared solar lines by terrestrial fluorescence and other geophysical effects: Simulations and space-based observations from SCIAMACHY and GOSAT. *Atmos Meas Tech* 5(4):809–829.
24. Lee J-E, et al. (2013) Forest productivity and water stress in Amazonia: Observations from GOSAT chlorophyll fluorescence. *Proc Biol Sci* 280(1761):20130171.
25. Parazoo NC, et al. (2013) Interpreting seasonal changes in the carbon balance of southern amazonia using measurements of $XCO_2$ and chlorophyll fluorescence from GOSAT. *Geophys Res Lett* 40(11):2829–2833.
26. Joiner J, et al. (2013) Global monitoring of terrestrial chlorophyll fluorescence from moderate spectral resolution near-infrared satellite measurements: Methodology, simulations, and application to GOME-2. *Atmos Meas Tech Discuss* 6(2):3883–3930.
27. Jung M, et al. (2011) Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *J Geophys Res* 116:G00J07, 10.1029/2010JG001566.
28. Sitch S, et al. (2013) Trends and drivers of regional sources and sinks of carbon dioxide over the past two decades. *Biogeosci Discuss* 10(12):20113–20177.

29. Piao S, et al. (2013) Evaluation of terrestrial carbon cycle models for their response to climate variability and to $CO_2$ trends. *Global Change Biol* 19(7):2117–2132.
30. USDA Foreign Agricultural Service (August 2013) World Agricultural Production (US Department of Agriculture, Washington, DC), Circular Ser WAP 8-13, 26 pp.
31. Running SW, et al. (2004) A continuous satellite-derived measure of global terrestrial primary production. *Bioscience* 54(6):547–560.
32. Zhang X, et al. (2006) Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements. *J Geophys Res* 111:G04017, 10.1029/2006JG000217.
33. Monteith JL (1972) Solar-radiation and productivity in tropical ecosystems. *J Appl Ecol* 9(3):747–766.
34. Fournier A, et al. (2012) Effect of canopy structure on sun-induced chlorophyll fluorescence. *ISPRS J Photogramm Remote Sens* 68:112–120.
35. Flexas J, et al. (2002) Steady-state chlorophyll fluorescence (Fs) measurements as a tool to follow variations of net $CO_2$ assimilation and stomatal conductance during water-stress in C3 plants. *Physiol Plant* 114(2):231–240.
36. Zarco-Tejada PJ, et al. (2013) Spatio-temporal patterns of chlorophyll fluorescence and physiological and structural indices acquired from hyperspectral imagery as compared with carbon fluxes measured with eddy covariance. *Remote Sens Environ* 133:102–115.
37. Damm A, et al. (2010) Remote sensing of sun-induced fluorescence to improve modeling of diurnal courses of gross primary production (GPP). *Global Change Biol* 16(1):171–186.
38. US Department of Agriculture National Agricultural Statistics Service. Cropland data layer. Available at http://gcmd.nasa.gov/records/GCMD_USDA_NASS_CROPLAND.html. Accessed March 15, 2013.
39. Yang HS, et al. (2004) Hybrid-Maize—A maize simulation model that combines two crop modeling approaches. *Field Crops Res* 87(23):131–154.
40. Keating BA, et al. (2003) An overview of APSIM, a model designed for farming systems simulation. *Eur J Agron* 18(34):267–288.
41. Bondeau A, et al. (2007) Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Global Change Biol* 13(3):679–706.
42. Frankenberg C, et al. (2014) Prospects for chlorophyll fluorescence remote sensing from the Orbiting Carbon Observatory-2. *Remote Sens Environ*, 10.1016/j.rse.2014.02.007.
43. Veefkind JP, et al. (2012) TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sens Environ* 120:70–83.

# Supporting Information

## Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence

Luis Guanter[1], Yongguang Zhang[1], Martin Jung[2], Joanna Joiner[3], Maximilian Voigt[1], Joseph A. Berry[4], Christian Frankenberg[5], Alfredo Huete[6], Pablo Zarco-Tejada[7], Jung-Eun Lee[8], M. Susan Moran[9], Guillermo Ponce-Campos[9], Christian Beer[10], Gustavo Camps-Valls[11], Nina Buchmann[12], Damiano Gianelle[13], Katja Klumpp[14], Alessandro Cescatti[15], John M. Baker[16], and Timothy J. Griffis[17]

[1]Institute for Space Sciences, Freie Universität Berlin, Germany
[2]Department for Biogeochemical Systems, Max Planck Institute for Biogeochemistry, Jena, Germany
[3]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[4]Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA
[5]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[6]Plant Functional Biology and Climate Change Cluster, University of Technology Sydney, Australia
[7]Instituto de Agricultura Sostenible (IAS), CSIC, Córdoba, Spain
[8]Geological Sciences, Brown University, Providence, RI, USA
[9]USDA ARS Southwest Watershed Research Center, Tucson, AZ, USA
[10]Department of Applied Environmental Science (ITM) and Bert Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden
[11]Image Processing Laboratory, Universitat de València, Spain
[12]ETH Zurich, Agricultural Sciences, Zurich, Switzerland
[13]Sustainable Agro-ecosystems and Bioresources Dept., Research and Innovation Centre, Fondazione E. Mach, Italy
[14]Grassland Ecosystem Research Unit, INRA, Clermont-Ferrand, France
[15]European Commission, JRC, Institute for Environment and Sustainability, Ispra, Italy
[16]USDA ARS Soil and Water Management Research, St Paul, MN, USA
[17]Department of Soil, Water, and Climate, University of Minnesota, St Paul, MN, USA

## Contents


1 SIF retrievals 2

2 Model-based GPP data 3

3 Comparison of flux tower-based GPP with model GPP, SIF and vegetation indices 5

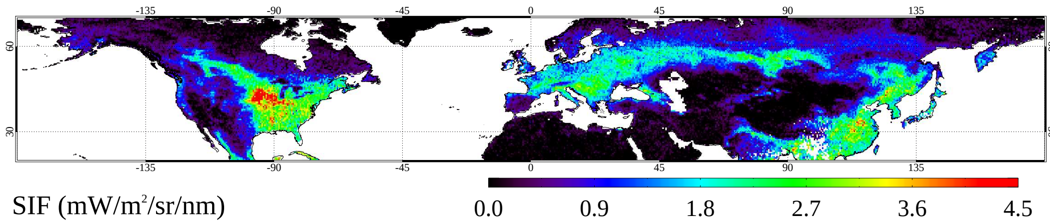
4 Derivation of spatially-explicit crop GPP estimates 11

5 NPP data from agricultural inventories 11

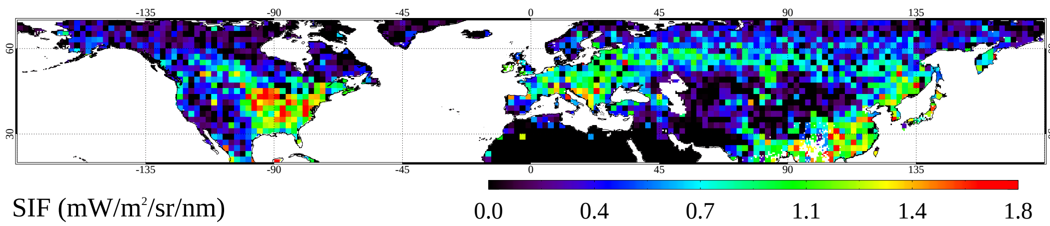

# 1 SIF retrievals

We use SIF data derived from spectral radiance measurements by the GOME-2 instrument onboard the Eumetsat's MetOp-A platform launched in October 2006. Details can be found in [1]. GOME-2 measures in the 240–790 nm spectral range with relatively high spectral resolution (~0.2–0.4 nm), signal-to-noise ratio (~1000–2000), and a footprint size of 40×80 km$^2$. SIF retrievals are performed in the 715–758 nm spectral window overlapping the second peak of the SIF emission. The retrieval method disentangles SIF from the spectral signals of atmospheric absorption and scattering and of surface reflectance which affect the measured top-of-atmosphere radiance. The retrievals are quality-filtered and binned in a 0.5° lat-lon grid. GOME-2 data between 2007 and 2011 have been used in this work.

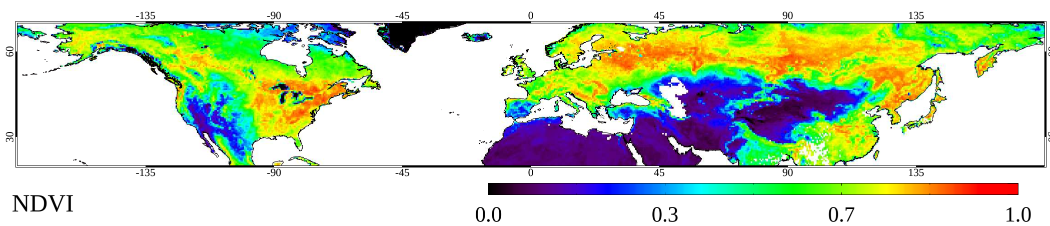
Fig. S1 presents SIF retrievals from GOME-2 and GOSAT's Fourier Transform Spectrometer (FTS) data over the northern temperate region. NDVI from the MODIS MOD13C2 product is also shown for reference. The retrieval approach applied to the GOSAT data is described in Guanter et al. [2]. The retrieval of SIF from GOSAT data is much simpler than that for GOME-2 thanks to the very high spectral resolution of the GOSAT's FTS (~0.025 nm), which allows to use narrow fitting windows (hence simpler modeling of the background surface reflectance) and to resolve individual solar Fraunhofer lines (i.e. free from contamination by atmospheric absorption, mostly $O_2$ in this spectral range). GOSAT/FTS measurements consist of round field-of-views of about 10 km diameter separated by hundreds of kilometers. The random component of the single-retrieval error is high, in the range of 50–100%, due to the narrow fitting window used for the retrieval and the relatively low signal-to-noise ratio (~100–300) of the FTS. Global composites of monthly SIF from GOSAT retrievals are typically produced by averaging in 2° gridboxes. Despite the noise and the low spatial resolution of the GOSAT SIF composites, we consider them to be highly accurate (free from systematic errors) due to the simplicity of the retrieval approach based on narrow fitting windows and solely Fraunhofer lines. Therefore,




(a) SIF GOME-2, July 2009, 0.5° grid, 715–758 nm, ∼9:30h local time overpass




(b) SIF GOSAT, July 2009, 2° grid, 757 nm, ∼13h local time overpass




(c) NDVI, July 2009, 0.5° grid (MODIS MOD13C2)

**Fig. S 1:** Monthly composites (July 2009) of SIF retrievals from GOSAT/FTS and MetOp-A/GOME-2 measurements. NDVI from the MODIS MOD13C2 product is also shown for reference. GOME-2 retrievals are for a spectral fitting window centered around 740 nm (715–758 nm) and are gridded in 0.5° cells, whereas GOSAT retrievals are for a narrow window at 757 nm and are gridded in 2° cells.

the good comparison between the spatial patterns in the GOSAT and the GOME-2 SIF supports the consistency of the GOME-2 SIF data used in this work, and in particular of the outstanding SIF levels observed at the Midwest US in the GOME-2 data (Fig. 1–2 of the main text). Slight differences in the spatial patterns of GOSAT and GOME-2 SIF can be explained by the lower precision of the GOSAT retrievals, which leads to noisier SIF composites, and the different overpass times (morning for MetOp-A, noon for GOSAT) which makes the latitudinal differences in the solar flux received in the north and the south to be greater for GOSAT than for GOME-2. The absolute SIF values differ for GOME-2 and GOSAT-FTS because of the different retrieval wavelengths and instantaneous illumination fluxes associated to the overpass time of each satellite.

## 2 Model-based GPP data

We have used global GPP estimates from ensembles of data-driven and process-based models as follows:

- **Data-driven models** are based on the calculation of GPP with empirical and semi-empirical relationships between GPP and a series of diagnostic variables (e.g. vegetation parameters such as the fraction of absorbed photosynthetically active radiation and meteorological variables such as short-wave radiation or vapor pressure deficit). As representative of state-of-the-art data-driven methods, we have used annual GPP estimates from 5 of the data-driven models described in Beer et al. [3], namely MTE1, MTE2, ANN, KGB and LUE. These models differ with each other in how the relationship between the diagnostic variables and GPP is expressed.

  In addition, monthly GPP estimates from the MTE1 model, referred to as Max Planck Institute for Biogeochemistry (MPI-BGC) model [4] in the main text, and from the MODIS GPP model (MOD17) [5] are used in the comparison with flux tower GPP in Fig. 2 of the main text and Fig. S4, respectively. The MPI-BGC GPP data set is produced through the global upscaling of site measurements of carbon dioxide fluxes. This is based on a Model Tree Ensemble approach for a statistical formulation of the relationship between GPP and vegetation parameters derived from remote sensing data and meteorological variables from re-analysis products. MOD17 GPP is derived from a production-efficiency approach consisting in the formulation of GPP as the product of absorbed photosynthetically-active radiation derived from satellite and meteorological data and tabulated light use efficiency.

- **Process-based models** or dynamic global vegetation models (DGVMs), are based on mathematical representations of physiological and ecological mechanisms driving productivity among other vegetation responses. The DGVMs in our ensemble of process-based models are part of the Trendy activity[1] intended to intercompare *Trends in net land - atmosphere carbon exchange* over the period 1980–2010. We have use the CLM4C, CLM4CN, HYLAND, LPJ, LPJ-GUESS, OCN, Orichidee, SDGVM, TRIFFID, and VEGAS models. Model outputs were available at different spatial resolutions. The data from the LPJ, LPJ-GUESS, Orchidee and VEGAS models were simulated at $0.5°\times0.5°$ resolution, CLM4C and CLM4CN at $2.5°\times1.875°$, and OCN, TRIFFID and HYLAND other at $3.75°\times 2.5°$. All 10 models have been resampled to the 0.5° grid used for the SIF measurements, the data-driven model ensemble and the NPP inventories.

Fig. S2 shows the median and the standard deviation of the annual GPP from the 5 data-driven models from Beer et al. [3] and the 10 process-based Trendy models from Piao et al. [6], Sitch et al. [7] that we have used in this study. The median of the annual GPP from the two model ensembles shows similar absolute values, although there are some spatial differences, especially in North America. The spread of GPP estimates is significantly smaller for the data-driven models than for the process-based models.

[1] `http://dgvm.ceh.ac.uk/node/9`

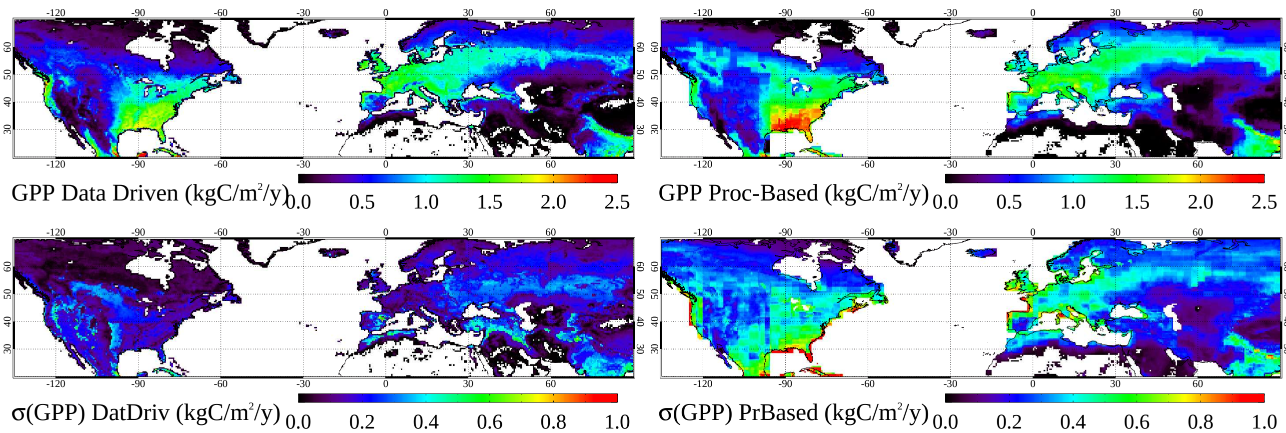



**Fig. S 2:** Median (top row) and mean absolute deviation (bottom row) of annual GPP estimates in North America and Western Europe from the data-driven and process-based model ensembles used in this work. Details about each model ensemble can be found in Beer et al. [3] and Piao et al. [6], Sitch et al. [7], respectively.

# 3 Comparison of flux tower-based GPP with model GPP, SIF and vegetation indices

We used fourteen eddy flux sites from the FLUXNET network [8] (Table S1). Six of these sites are located in crop fields in the US Corn Belt. The remaining eight stations include five crop sites and three grassland sites located across Europe. Sites have been selected on the basis of landscape homogeneity in the GOME-2 grid and on data availability in the period of interest (2007–2011). To determine landscape homogeneity, we used land cover type data from the MODIS Collection 5 MCD12C1 product (Friedl et al. [9]) and EVI data from the MODIS MOD13C2 product (Huete et al. [10]), both with spatial resolution of 0.05°. For a site to be selected for the study, the dominant vegetation cover type at the flux site (either cropland or grassland) must represent more than 60% of the GOME-2 pixel area, and the standard deviation of the EVI must be less than 0.10 (see Table S1). We used the Level 4 data product for the six US crop sites from the AmeriFlux website[2], and from the GHG-Europe database[3] for the eight Europe sites. Monthly GPP values were used in our investigation. GPP is estimated by partitioning the observed net flux into GPP and ecosystem respiration as discussed in Reichstein et al. [11] and Papale et al. [12].

For each site, SIF was extracted based on the coordinates of the flux tower, and averaged to monthly means when at least 5 SIF retrievals were available. Three US crop sites (US-IB1, Ne2-3, Ro1) are very close to big cities. To avoid signal contamination from urban areas, we extracted SIF from a nearby pixel fulfilling the homogeneity criteria. Given that flux measurements are usually representative of a large area in homogeneous landscapes (i.e., US-IB1 is representative of central Illinois), we assumed that SIF (or EVI and NDVI) from nearby grid boxes can represent the footprint of the flux towers. Monthly SIF

[2]http://ameriflux.ornl.gov/

[3]http://www.europe-fluxdata.eu/

and GPP were averaged over the 2007–2011 observation period for each month to minimize uncertainties due to the different spatial scales of the SIF retrievals and the flux tower data. This uncertainties occur because both corn and soybean fields exist in the GOME-2 footprint for the US flux sites. A mixed signal of corn and soybean is therefore sampled by the GOME-2 footprint, while the eddy covariance tower measured flux either from corn or soybean for each year. Multi-year averaging may help reduce this mismatch.

**Table S 1:** Details of the flux tower sites used in this study. LC stands for Land Cover class, max(LC) stands for the percent of dominant vegetation cover within the GOME-2 pixel, EVI is the MODIS Enhanced Vegetation Index, and $\sigma$(EVI) represents the standard deviation of EVI within the GOME-2 pixel.

| Site ID | Country | Lat. (°) | Lon. (°) | IGBP class | Study period | max(LC) (%) | mean EVI | $\sigma$ EVI | Vegetation type or crop rotations | Reference |
|---|---|---|---|---|---|---|---|---|---|---|
| US-Bo1 | USA | 40.00 | -88.29 | CRO | 2007 | 0.98 | 0.55 | 0.04 | Corn | Ryu et al. [13] |
| US-IB1 | USA | 41.85 | -88.22 | CRO | 2007–2009 | 0.98 | 0.44 | 0.08 | Soybean/Corn/Soyb. | Allison et al. [14] |
| US-Ne2 | USA | 41.16 | -96.47 | CRO | 2007–2010 | 0.94 | 0.56 | 0.07 | Corn/Soybean/Corn/Corn | Suyket et al. [15] |
| US-Ne3 | USA | 41.17 | -96.43 | CRO | 2007–2010 | 0.95 | 0.57 | 0.07 | Corn/Soybean/Corn/Soyb. | Suyker et al. [15] |
| US-Ro1 | USA | 44.71 | -93.09 | CRO | 2007–2010 | 1.00 | 0.49 | 0.10 | Corn/Soybean/Corn/Soyb. | Griffis et al. [16] |
| US-SFP | USA | 43.24 | -96.90 | CRO | 2007–2009 | 1.00 | 0.55 | 0.03 | Continuous corn | – |
| DE-Gri | Germany | 50.94 | 13.51 | GRA | 2007–2010 | 0.58 | 0.44 | 0.04 | Permanent grassland | Hussain et al. [17] |
| FR-Lq1 | France | 45.64 | 2.73 | GRA | 2007–2010 | 0.79 | 0.57 | 0.04 | Permanent grassland | Klumpp et al. [18] |
| HU-Bug | Hungary | 46.69 | 19.60 | GRA | 2007–2008 | 0.94 | 0.35 | 0.03 | Permanent grassland | Naggy et al. [19] |
| BE-Lon | Belgium | 50.55 | 4.74 | CRO | 2007–2010 | 0.71 | 0.49 | 0.07 | Winter wheat/sugar beet/ /winter wheat/seed potato | Aubinet et al. [20] |
| CH-Oe2 | Switzerland | 47.28 | 7.73 | CRO | 2007–2009 | 0.71 | 0.50 | 0.05 | Winter wheat/rapeseed/ /winter wheat | Dietiker et al. [21] |
| DE-Geb | Germany | 51.10 | 10.91 | CRO | 2007–2010 | 0.97 | 0.46 | 0.08 | Winter wheat/rapeseed/ /barley/sugar beet | Kutsch et al. [22] |
| DE-Seh | Germany | 50.87 | 6.44 | CRO | 2007–2010 | 0.60 | 0.45 | 0.07 | Winter wheat/winter wheat/ /sugar beet/winter wheat | Schmidt et al. [23] |
| IT-Cas | Italy | 45.06 | 8.66 | CRO | 2007–2010 | 0.97 | 0.43 | 0.09 | Continuous paddy rice | Skiba et al. [24] |

Reflectance-based vegetation indices derived from satellite observations [e.g. 10, 25] provide information about vegetation *greenness* (i.e. a combination of biomass, chlorophyll content and structural effects) and have also been reported to be good indicators of gross primary production [e.g. 26]. The data-driven GPP models combine these reflectance-based proxies for green biomass and canopy light interception with meteorological inputs modulating photosynthesis at the ecosystem scale.

To complete the comparison of model GPP with fluorescence and tower-based GPP discussed in the main text, we have also analyzed the relationship between flux tower GPP and the normalized difference vegetation index (NDVI) [27], the enhanced vegetation index (EVI) [10], both extracted from the MOD13C2 product, and the MERIS terrestrial chlorophyll index (MTCI) [28]. The NDVI is the most widely used vegetation index in the last decades. The EVI is a modification of the NDVI intended to improve the response of the NDVI for high green biomass levels and to reduce the sensitivity to atmospheric effects. The MTCI is designed to provide a high sensitivity to chlorophyll content through the sampling of the so-called red-edge window between the red and the near-infrared spectral regions.

Fig. S3 displays maps of the EVI, NDVI and MTCI for July 2009 and the same area as the GPP and SIF maps shown in Fig. 2 of the main text (please, note that maximum monthly values instead of July values are plotted in Fig. 2 of the main text, so this comparison is only approximate). The data-driven GPP from the MODIS MOD17 product is also shown. The NDVI appears to be close to saturation in the most densely vegetated areas of North America and Europe. This is not happening for the EVI, which shows a somewhat higher signal in the midwest and the east coast of the US than in Europe, in line with the spatial patterns of SIF and GPP MPI-BGC (Fig. 2 of the main text). No significant differences between Europe and the US are observed in the MOD17 GPP data. On the other hand, the spatial patterns of the MTCI at the US Corn Belt are the most similar ones to those of SIF. This could be due to the fact that both SIF and the MTCI are most sensitive to canopy chlorophyll content for the high levels of leaf-area index found at the peak of the growing season for the corn and soybean crops in the US Corn Belt.

The same three indices have been compared with flux tower-based GPP estimates as we have done with MPI-BGC GPP, process-based GPP from the Trendy models and SIF in Fig. 3 of the main text. Results are shown in Fig. S4, in this case also including the European crop sites not included in Fig. 3 of the main text. Points to be noted are (i) the relatively bad comparison between GPP and both EVI and NDVI for the US crops, (ii) the good correlation between EVI and GPP when the comparison is performed for all three biomes, (iii) the lower values of EVI and MTCI at the grasslands sites, which agrees with SIF and the tower-based GPP, but not with the data-driven GPP estimates, and (iv) the good performance of the MTCI to track GPP in the US crops. These results, together with the conclusions extracted from Fig. 3 of the main text, support our approach of selecting SIF as the best input to upscale cropland GPP from the tower footprint to the regional scale. The relationship GPP(SIF)$=-0.10+3.72\times$SIF) is used for this upscaling.

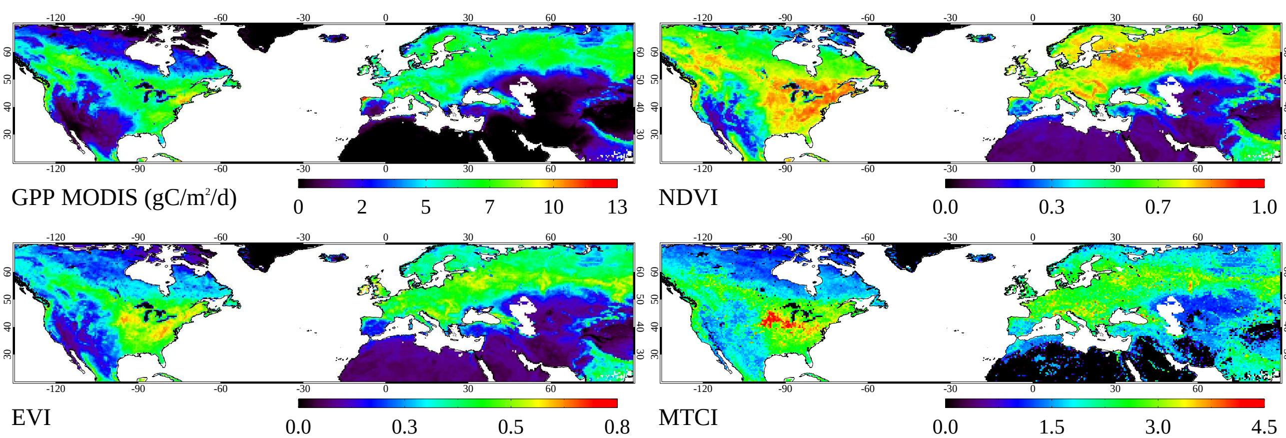



**Fig. S 3:** Maps of GPP from the MODIS MOD17 product, NDVI and EVI from the MODIS MOD13C2 product and the MERIS MTCI for July 2009 and the same region of the GPP and fluorescence maps displayed in Fig. 2 of the main text. Please, note that maximum monthly values instead of July values are plotted in Fig. 2 of the main text, so the comparison is only approximate.

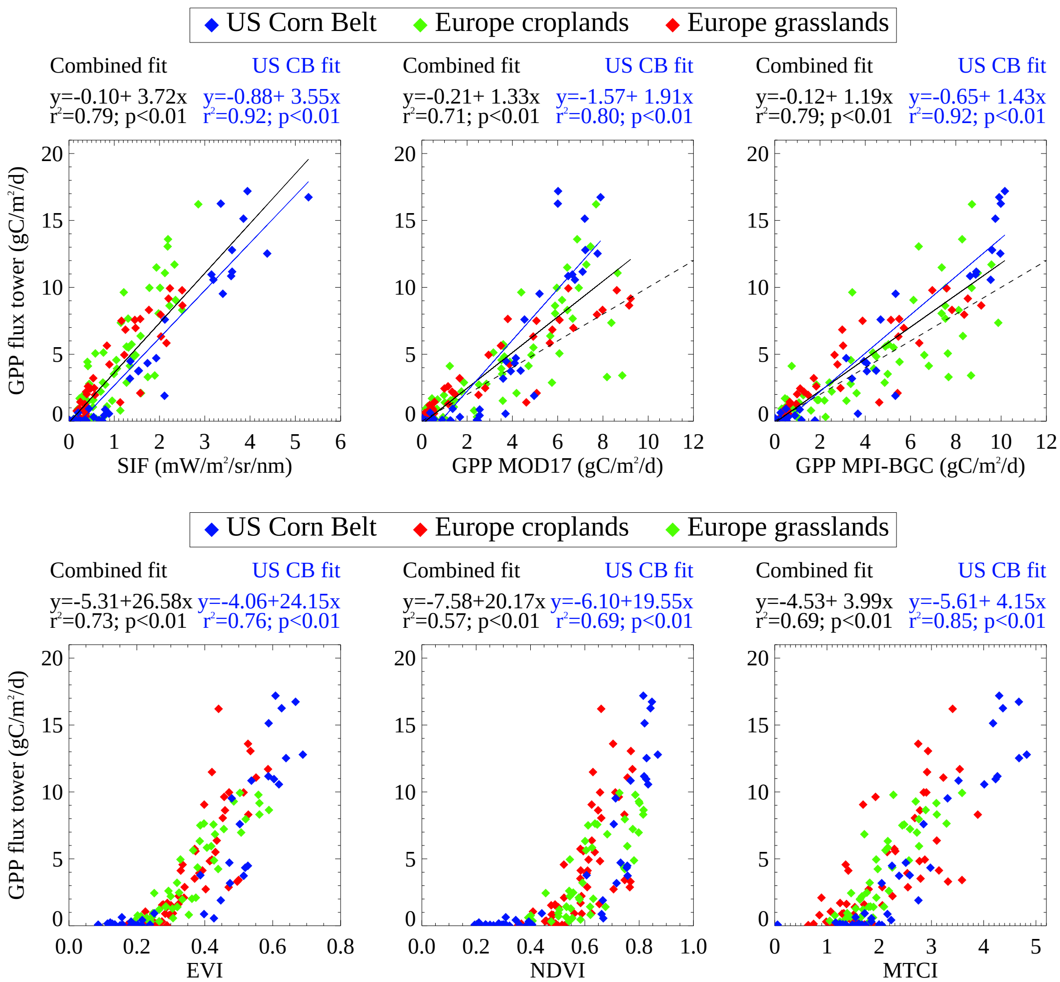



**Fig. S 4:** Similar to Fig. 3 of the main text but including the European cropland sites. Tower-based GPP is compared with SIF, GPP MPI-BGC and GPP MOD17 (top) and with EVI, NDVI and MTCI data (bottom).

# 4 Derivation of spatially-explicit crop GPP estimates

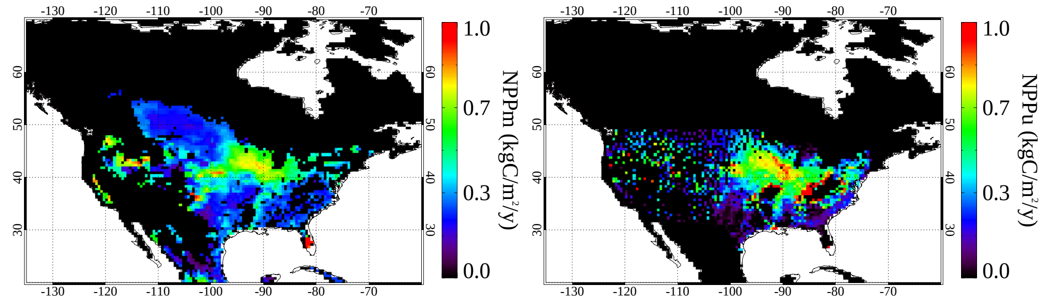
The monthly composites of SIF at 0.5° are scaled to GPP with the linear relationship derived from the comparison of SIF with flux tower-based GPP shown in Fig. S4a (GPP(SIF)=$-0.10+3.72\times$SIF). Model-based GPP maps are generated as the median GPP per grid cell from the data-driven and process-based model ensembles described before. We have estimated crop GPP from the total GPP in the grid box by multiplying the total GPP by the fraction of cropland area in the gridbox described in Ramankutty et al. [29] and downloadable from `http://www.geog.mcgill.ca/~nramankutty/Datasets/Datasets.html`. As a result, we obtain the cropland GPP per unit total area, as shown in Fig. 6a of the main text. Comparison of annual, area-integrated crop GPP estimated from SIF and the data-driven and process-based models are provided in Table S2.

**Table S 2:** Annual, area-integrated GPP estimates over the US Corn Belt (35–50°N, -105–80°E), Western Europe (35–55°N, -10–25°E), India (23–33°N, 70–90°E), China (30–49°N, 110–135°E), South America (−40–−20°N, −45–−70°E), and the globe from the median of the data-driven and process-based biogeochemistry model ensembles and the scaled SIF. These regions match those used to produce Fig. 7 of the main text. Relative ΔGPP is calculated as SIF-based GPP minus model GPP over model GPP. Uncertainties are derived from the standard deviation of the ensembles in the case of the GPP models and from the errors in the slope and intercept in the linear regression in Fig. S4a for the scaled SIF.

| | **Crop GPP (PgC y$^{-1}$)** | | | | | |
|---|---|---|---|---|---|---|
| | **US CB** | **WestEur** | **India** | **China** | **SouthAm** | **Global** |
| GPP(Data-Driven) | 1.1±0.2 | 1.3±0.3 | 0.8±0.3 | 0.73±0.16 | 0.95±0.15 | 17±4 |
| GPP(Proc.-based) | 1.3±0.5 | 1.5±0.6 | 0.9±0.4 | 0.9±0.3 | 1.2±0.4 | 20±9 |
| GPP(SIF) | 1.54±0.06 | 1.30±0.05 | 1.23±0.06 | 0.90±0.05 | 0.81±0.04 | 17.0±0.2 |
| ΔGPP(Data-Driven) | 43% | 0% | 55% | 24% | −14% | 3% |
| ΔGPP(Proc.-based) | 18% | −14% | 39% | −1% | −38% | −12% |
| Crop area ($10^6$ km$^2$) | 1.2 | 1.3 | 1.0 | 0.9 | 0.7 | 16.5 |

# 5 NPP data from agricultural inventories

The SIF- and model-based crop GPP estimates have been compared with crop net primary productivity (NPP) estimates derived from agricultural inventories to produce Fig. 5 of the main text. Large-scale NPP estimates have been provided by the agricultural inventory data sets described in USDA-NASS [30] and Monfreda et al. [31]. The USDA NPP inventory was estimated using a statistical method that includes factors for dry weight, harvest indices, and root:shoot ratios multiplied by yield data from the National Agricultural Statistics Service (NASS). This method has been documented and published by Hicke and Lobell [32], Hicke et al. [33], Prince et al. [34]. U.S. county-level estimates of croplands production (P, in units of MgC$y^{-1}$) dataset is available in `http://cdiac.ornl.gov/carbonmanagement/cropcarbon/`. Data from the three most recent years (2006–2008) was used for




(a) Monfreda et al.

(b) USDA

**Fig. S 5:** Crop NPP per harvested area in North America from the global inventory by Monfreda et al. for 2000 (a) and the USDA inventory (2006 and 2008) [33].

comparison. To derive the spatial distribution of cropland GPP, county-level NPP (kgC$m^{-2}y^{-1}$) was collocated in ArcGIS to a layer of the cultivated area of the US during 2008–2012. To compute NPP, we divide P by the total crop area of each county. The cultivated layer data is available from USDA NASS database at `http://www.nass.usda.gov/research/Cropland/Release/index.htm.`. Regarding the global inventory by Monfreda et al., it is based on the aggregation of 175 crop classes in a 5 min by 5 min grid following a method similar to the one proposed by Prince et al. [34] for the US. Monfreda et al. data corresponds to the year 2000.

Both USDA-NASS and Monfreda et al. NPP data sets are derived from the crop yields, and have units of per-harvested-areas (Fig. S5). NPP is converted from per-harvested-area to per-total-area units through the multiplication by the fraction of harvested area as described in Monfreda et al. (Fig. S6). The fraction of harvested area is calculated by summing the fraction of harvested area for each of the 175 crop classes considered by Monfreda et al. (data available from `http://www.geog.mcgill.ca/~nramankutty/Datasets/Datasets.html`).

The comparison of NPP from the USDA inventory with GPP from the SIF retrievals and the data-driven and process-based models for the US Western Corn Belt is shown in Fig. 5 of the main text. The same comparison for the NPP from Monfreda et al. for both the US and Western Europe is displayed in Fig S7.

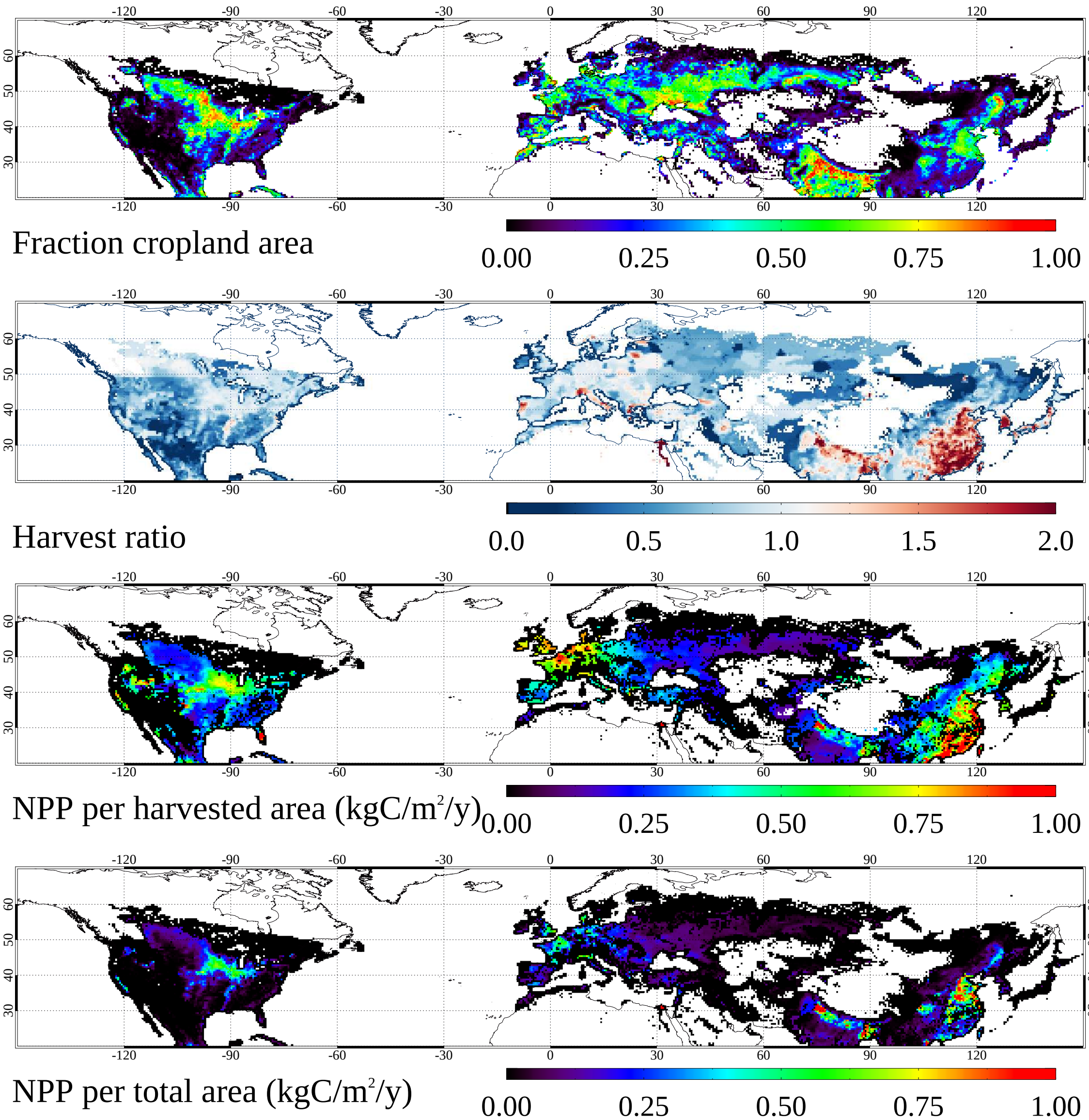



**Fig. S 6:** Cropland area and net primary production data sets from Ramankutty et al. [29] and Monfreda et al. [31] The fraction of cropland area expresses the ratio of cropland to total area in each 0.5° grid cell. The harvest ratio is the ratio of harvested-to-cropland area. The fraction of harvested area has been calculated from single fractions of harvested area provided by Monfreda et al. [31] for a total of 175 crop classes. The NPP per total area is calculated as the product of the original per-harvested-area NPP data from Monfreda et al. by the fraction of harvested area.

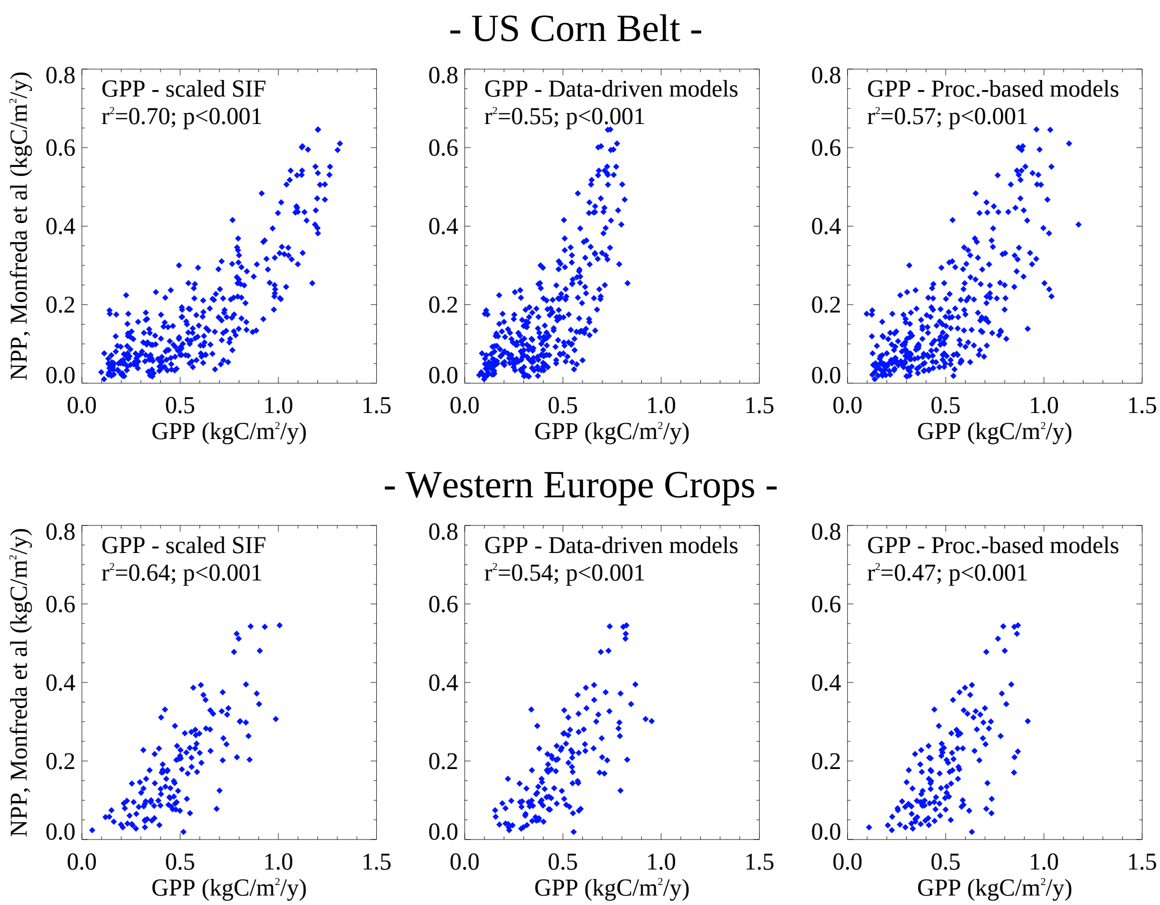



**Fig. S 7:** Same as Fig. 5 of the main text but for the NPP data set from the agricultural inventory by Monfreda et al. and showing results also for the Western Europe area (40–55°N, -5–15°E).

## References


[1] J. Joiner, et al. Global monitoring of terrestrial chlorophyll fluorescence from moderate spectral resolution near-infrared satellite measurements: methodology, simulations, and application to GOME-2. *Atmospheric Measurement Techniques Discussions*, 6(2):3883–3930, 2013. doi: 10.5194/amtd-6-3883-2013.

[2] L. Guanter, et al. Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements. *Remote Sensing of Environment*, 121:236–251, 2012.

[3] C. Beer, et al. Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate. *Science*, 329(5993):834–838, 2010. doi: 10.1126/science.1184984.

[4] M. Jung, et al. Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research: Biogeosciences*, 116(G3):n/a–n/a, 2011. ISSN 2156-2202. doi: 10.1029/2010JG001566.

[5] S. W Running, et al. A Continuous Satellite-Derived Measure of Global Terrestrial Primary Production. *BioScience*, 54(6):547–560, 2004.

[6] S. Piao, et al. Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO2 trends. *Global Change Biology*, 19(7):2117–2132, 2013.

[7] S. Sitch, et al. Trends and drivers of regional sources and sinks of carbon dioxide over the past two decades. *Biogeosciences Discussions*, 10(12):20113–20177, 2013. doi: 10.5194/bgd-10-20113-2013.

[8] D Baldocchi, et al. FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, 82(11):2415–2434, NOV 2001. ISSN 0003-0007.

[9] M. A. Friedl, et al. {MODIS} collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114(1):168 – 182, 2010. ISSN 0034-4257. doi: http://dx.doi.org/10.1016/j.rse.2009.08.016.

[10] A Huete, et al. Overview of the radiometric and biophysical performance of the modis vegetation indices. *Remote Sensing of Environment*, 83(1-2):195 – 213, 2002. ISSN 0034-4257. doi: 10.1016/S0034-4257(02)00096-2.

[11] M. Reichstein, et al. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Global Change Biology*, 11(9):1424–1439, 2005. ISSN 1365-2486. doi: 10.1111/j.1365-2486.2005.001002.x.

[12] D. Papale, et al. Towards a standardized processing of net ecosystem exchange measured with eddy covariance technique: algorithms and uncertainty estimation. *Biogeosciences*, 3(4):571–583, 2006. doi: 10.5194/bg-3-571-2006.

[13] Y. Ryu, et al. On the temporal upscaling of evapotranspiration from instantaneous remote sensing measurements to 8-day mean daily-sums. *Agricultural and Forest Meteorology*, 152(0):212 – 222, 2012. ISSN 0168-1923. doi: http://dx.doi.org/10.1016/j.agrformet.2011.09.010.

[14] V. J. Allison, et al. Changes in soil microbial community structure in a tallgrass prairie chronosequence. *Soil Science Society of America Journal*, 69(5):1412–1421, SEP-OCT 2005. ISSN 0361-5995. doi: {10.2136/sssaj2004.0252}.

[15] A. E. Suyker, et al. Gross primary production and ecosystem respiration of irrigated maize and irrigated soybean during a growing season. *Agricultural and Forest Meteorology*, 131(34):180 – 190, 2005. ISSN 0168-1923. doi: http://dx.doi.org/10.1016/j.agrformet.2005.05.007.

[16] T. J. Griffis, et al. Direct measurement of biosphere-atmosphere isotopic co2 exchange using the eddy covariance technique. *Journal of Geophysical Research: Atmospheres*, 113(D8):n/a–n/a, 2008. ISSN 2156-2202. doi: 10.1029/2007JD009297.

[17] M. Z. Hussain, et al. Summer drought influence on CO2 and water fluxes of extensively managed grassland in Germany. *Agriculture Ecosystems & Environment*, 141(1-2):67–76, APR 2011. ISSN 0167-8809. doi: {10.1016/j.agee.2011.02.013}.

[18] K. Klumpp, et al. Long-term impacts of agricultural practices and climatic variability on carbon storage in a permanent pasture. *Global Change Biology*, 17(12):3534–3545, 2011. ISSN 1365-2486. doi: 10.1111/j.1365-2486.2011.02490.x.

[19] Z. Nagy, et al. Some preliminary results of the hungarian grassland ecological research: Carbon cycling and greenhouse gas balances under changing. *Cereal Research Communications*, 33(1): 279–281, 2005.

[20] M. Aubinet, et al. Carbon sequestration by a crop over a 4-year sugar beet/winter wheat/seed potato/winter wheat rotation cycle. *Agricultural and Forest Meteorology*, 149(34):407 – 418, 2009. ISSN 0168-1923. doi: http://dx.doi.org/10.1016/j.agrformet.2008.09.003.

[21] D. Dietiker, et al. Testing the ability of the {DNDC} model to predict {CO2} and water vapour fluxes of a swiss cropland site. *Agriculture, Ecosystems & Environment*, 139(3):396 – 401, 2010. ISSN 0167-8809. doi: http://dx.doi.org/10.1016/j.agee.2010.09.002.

[22] W.L. Kutsch, et al. The net biome production of full crop rotations in europe. *Agriculture, Ecosystems & Environment*, 139(3):336 – 345, 2010. ISSN 0167-8809. doi: http://dx.doi.org/10.1016/j.agee.2010.07.016.

[23] M. Schmidt, et al. The carbon budget of a winter wheat field: An eddy covariance analysis of seasonal and inter-annual variability. *Agricultural and Forest Meteorology*, 165(0):114 – 126, 2012. ISSN 0168-1923. doi: http://dx.doi.org/10.1016/j.agrformet.2012.05.012.

[24] U. Skiba, et al. Biosphereatmosphere exchange of reactive nitrogen and greenhouse gases at the nitroeurope core flux measurement sites: Measurement strategy and first data sets. *Agriculture, Ecosystems & Environment*, 133(34):139 – 149, 2009. ISSN 0167-8809. doi: http://dx.doi.org/10.1016/j.agee.2009.05.018.

[25] R.B. Myneni, et al. The interpretation of spectral vegetation indexes. *Geoscience and Remote Sensing, IEEE Transactions on*, 33(2):481–486, 1995. ISSN 0196-2892. doi: 10.1109/36.377948.

[26] D. A. Sims, et al. On the use of modis evi to assess gross primary productivity of north american ecosystems. *Journal of Geophysical Research: Biogeosciences*, 111(G4):n/a–n/a, 2006. ISSN 2156-2202. doi: 10.1029/2006JG000162. URL `http://dx.doi.org/10.1029/2006JG000162`.

[27] C. J. Tucker. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8:127–150, 1979.

[28] J. Dash et al. The MERIS terrestrial chlorophyll index. *International Journal of Remote Sensing*, 25:5003–5013, 2004.

[29] N. Ramankutty, et al. Farming the planet: 1. geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles*, 22(1), 2008. doi: 10.1029/2007GB002952.

[30] USDA-NASS. USDA National Agricultural Statistics Service Cropland Data Layer. Published crop-specific data layer [Online]. Technical report.

[31] C. Monfreda, et al. Farming the planet: 2. geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, 22(1), 2008. doi: 10.1029/2007GB002947.

[32] J. A. Hicke et al. Spatiotemporal patterns of cropland area and net primary production in the central united states estimated from usda agricultural information. *Geophysical Research Letters*, 31(20):n/a–n/a, 2004. ISSN 1944-8007. doi: 10.1029/2004GL020927.

[33] J. A. Hicke, et al. Cropland Area and Net Primary Production Computed from 30 Years of USDA Agricultural Harvest Data. *Earth Interactions*, 8, 2004. ISSN 1087-3562.

[34] S. D. Prince, et al. Net primary production of US Midwest croplands from agricultural harvest yield data. *Ecological Applications*, 11(4):1194–1205, AUG 2001. ISSN 1051-0761. doi: {10.2307/3061021}.