9-2017

# The shared and unique values of optical, fluorescence, thermal and microwave satellite data for estimating large-scale crop yields

Kaiyu Guan

Jin Wu

John S. Kimball
*University of Montana - Missoula*

Martha C. Anderson

Steve Frolking

*See next page for additional authors*

Authors

Kaiyu Guan, Jin Wu, John S. Kimball, Martha C. Anderson, Steve Frolking, Bo Li, Christopher R. Hain, and
David B. Lobell

CrossMark

# The shared and unique values of optical, fluorescence, thermal and microwave satellite data for estimating large-scale crop yields

Kaiyu Guan [a,*], Jin Wu [b], John S. Kimball [c,d], Martha C. Anderson [e], Steve Frolking [f], Bo Li [g], Christopher R. Hain [h], David B. Lobell [i]

[a] Department of Natural Resources and Environmental Sciences, and National Center for Supercomputing Applications, University of Illinois at Urbana Champaign, Urbana, IL 61801, USA
[b] Biological, Environmental & Climate Sciences Department, Brookhaven National Lab, Upton, New York, NY 11973, USA
[c] Numerical Terradynamic Simulation Group, College of Forestry & Conservation, The University of Montana, Missoula, MT 59812, USA
[d] Department of Ecosystem and Conservation Sciences, College of Forestry & Conservation, The University of Montana, Missoula, MT 59812, USA
[e] Agricultural Research Services, United States Department of Agriculture, Hydrology and Remote Sensing Laboratory, Beltsville, MD, USA
[f] Institute for the Study of Earth Oceans and Space, University of New Hampshire, Durham, NH 03824, USA
[g] Department of Statistics, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA
[h] NASA Marshall Space Flight Center, Huntsville, AL 35811, USA
[i] Department of Earth System Science and Center on Food Security and the Environment, Stanford University, Stanford, CA 94305, USA

## ARTICLE INFO

## ABSTRACT

Large-scale crop monitoring and yield estimation are important for both scientific research and practical applications. Satellite remote sensing provides an effective means for regional and global cropland monitoring, particularly in data-sparse regions that lack reliable ground observations and reporting. The conventional approach of using visible and near-infrared based vegetation index (VI) observations has prevailed for decades since the onset of the global satellite era. However, other satellite data encompass diverse spectral ranges that may contain complementary information on crop growth and yield, but have been largely understudied and underused. Here we conducted one of the first attempts at synergizing multiple satellite data spanning a diverse spectral range, including visible, near-infrared, thermal and microwave, into one framework to estimate crop yield for the U.S. Corn Belt, one of the world's most important food baskets. Specifically, we included MODIS Enhanced VI (EVI), estimated Gross Primary Production based on GOME-2 solar-induced fluorescence (SIF-GPP), thermal-based ALEXI Evapotranspiration (ET), QuikSCAT Ku-band radar backscatter, and AMSR-E X-band passive microwave Vegetation Optical Depth (VOD) in this study, benchmarked on USDA county-level crop yield statistics. We used Partial Least Square Regression (PLSR), an effective statistical model for dimension reduction, to distinguish commonly shared and unique individual information from the various satellite data and other ancillary climate information for crop yield estimation. In the PLSR model that includes all of the satellite data and climate variables from 2007 to 2009, we assessed the first two major PLSR components and found that the first component (an integrated proxy of crop aboveground biomass) explained 82% variability of modelled crop yield, and the second component (dominated by environmental stresses) explained 15% variability of modelled crop yield. We found that most of the satellite derived metrics (e.g. SIF-GPP, radar backscatter, EVI, VOD, ALEXI-ET) share common information related to aboveground crop biomass (i.e. the first component). For this shared information, the SIF-GPP and backscatter data contain almost the same amount of information as EVI at the county scale. When removing the above shared component from all of the satellite data, we found that EVI and SIF-GPP do not provide much extra information; instead, Ku-band backscatter, thermal-based ALEXI-ET, and X-band VOD provide unique information on environmental stresses that improves overall crop yield predictive skill. In particular, Ku-band backscatter and associated differences between morning and afternoon overpasses contribute unique information on crop growth and environmental stress. Overall, using satellite data from various spectral bands significantly improves regional crop yield predictions. The additional use of ancillary climate data (e.g. precipitation and temperature) further improves model skill, in part because the crop reproductive stage related to harvest index is highly sensitive to environmental stresses but they are not fully captured by the satellite data used in our study. We conclude that using satellite data across various spectral ranges can improve monitoring of large-scale crop growth and yield beyond what can be achieved from individual sensors. These results also inform the synergistic use and development of current and next generation satellite missions, including NASA ECOSTRESS, SMAP, and OCO-2, for agricultural applications.

© 2017 Elsevier Inc. All rights reserved.

## 1. Introduction

Accurately estimating large-scale crop productivity has critical value for scientific and societal benefits. For scientific research, knowing crop productivity at large scales can contribute to better assessments and understanding of differences, or gaps, between actual and potential yields (Lobell et al., 2009; Van Ittersum et al., 2013). Accurate large-scale yield estimation also facilitates better understanding of how crop growth responds to environmental stress (Guan et al., 2016; Sibley et al., 2014), and provides a valuable benchmark to improve and calibrate crop models (Basso et al., 2013; Thorp et al., 2012) for better near-term crop yield forecasts and long-term climate change projections. From the practical and societal perspective, accurate crop yield estimation is critical for economic planning and development (Carletto et al., 2015; Whitcraft et al., 2015), and facilitates the price forecast of commodity markets and international trades (Hoffman et al., 2015). It also helps in the design and assessment of crop insurance (Sherrick et al., 2014), food security monitoring and agriculturally-related humanitarian crises (Ross et al., 2009; Thornton et al., 1997), and can potentially provide better information for field-level management and decision-making for individual farmers (Mulla, 2013).

Large-scale crop yield estimation often relies on satellite remote sensing. Traditional approaches have primarily used optical and near-infrared (NIR) remote sensing, such as through the use of vegetation indices (VIs) that provide a general indicator of photosynthetic canopy cover or leaf area index (Sellers et al., 1992). The Normalized Difference Vegetation Index (NDVI) was one of the first VIs developed Tucker (1979) and has been widely used for operational crop monitoring (e.g. VegScape: https://nassgeodata.gmu.edu/VegScape/; GEOGLAM initiative, http://cropmonitor.org/). However, VIs derived from visible and near-infrared remote sensing data only utilize information from a small portion of the electromagnetic spectrum, while other available spectral bands have been comparatively less studied and may provide unique and/or complementary information for crop assessments.

Remote sensing observations from current global earth observation satellites encompass a diversity of spectral ranges, including visible, infrared, thermal and microwave wavelengths (Fig. 1). Besides the visible and NIR based VIs, an emerging satellite product called solar-induced fluorescence (SIF) is also derived from a narrow spectral window in the NIR (Frankenberg, Butz, and Toon, 2011; Guanter et al., 2007; Guanter et al., 2012, 2014; Joiner et al., 2011). SIF is the active emission from plant chlorophyll in its photo-machinery (Porcar-Castell et al., 2014; van der Tol et al., 2014), and SIF has been used as a proxy of plant photosynthesis (Guan et al., 2016; Guanter et al., 2014). Moving towards the thermal bands, land-surface temperature (LST) has been used to estimate evapotranspiration (ET) (Anderson et al., 1997; Anderson et al., 2007; Maes and Steppe, 2012), a critical variable that may be correlated with crop growth for closed canopies and can also be used to assess plant water stress (Anderson et al., 2013a). Longer wavelength bands from passive microwave sensors detect natural microwave emissions of the land surface, vegetation and atmosphere; the resulting brightness temperature ($T_b$) retrievals can be used with microwave radiative transfer models to derive vegetation optical depth (VOD) (Du et al., 2015; Jones and Kimball, 2010). VOD provides a frequency dependent metric of canopy water content and biomass changes (Guan et al., 2014; Jones et al., 2013; Liu et al., 2015; Liu et al., 2011), and has also been used to infer plant water potential (Konings and Gentine, 2016). Active microwave sensors (i.e. Radars) are sensitive to land surface dialectic properties, roughness, and vegetation properties (Ulaby et al., 1982). Microwave signals respond to vegetation in a manner that depends on sensor wavelength or frequency, with lower frequency microwave retrievals (e.g. C-band, either from passive microwave $T_b$ or radar backscatter) generally more sensitive to deeper canopy biomass layers than higher frequency (e.g. Ku-, X-band) retrievals (Ulaby, 1987). Very long wavelength (i.e. low frequency) microwave retrievals (e.g. L or P band) can penetrate low to moderate vegetation cover and detect surface soil moisture (Jackson, 1993; Njoku et al., 2003), whereas shorter wavelength (i.e. higher frequency) microwave retrievals are more suitable for detecting canopy biomass, since these frequencies are proportional to the size of vegetation scattering elements such as leaves and are primarily sensitive to the upper portion of the canopy (Frolking et al., 2005; Guan et al., 2012; Guan et al., 2013; van Emmerik et al., 2015). The above is a brief and inconclusive synopsis of various information that can be detected from different spectral bands.

Given the widespread availability of global satellite data and diversity of potentially valuable information for agricultural monitoring, there have been very limited studies so far devoted to clarifying and understanding the unique and additive value of different satellite data for estimating crop yield. Several guiding questions arise regarding the potential integration and use of these data, including: (1) How much common information is shared by various satellite and climate data for monitoring crop yield? What unique information is contributed by individual data? (2) Can a systematic approach be used to decouple and distinguish commonly shared and unique information from various satellite data? (3) And finally, how can we integrate different satellite
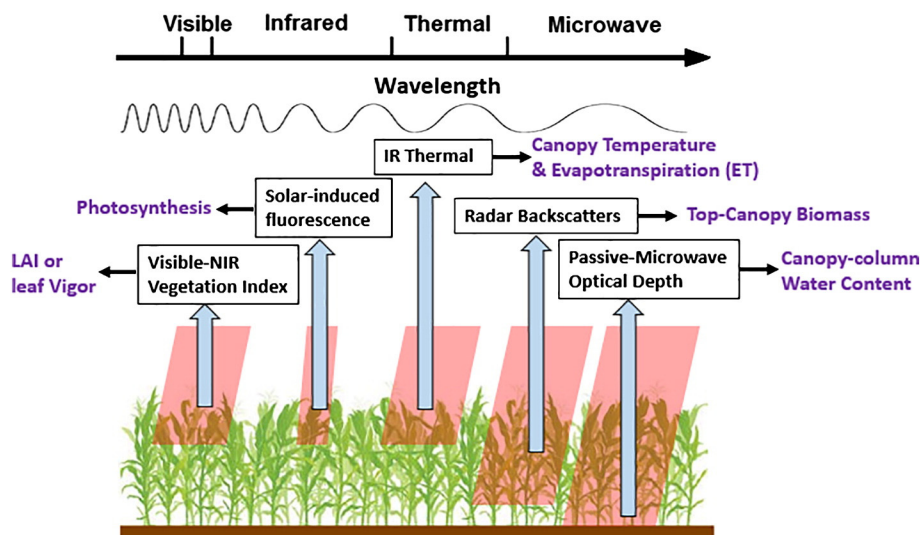


Fig. 1. Major satellite remote sensing data that are analyzed in the current study, which span a large range of the electromagnetic spectrum. The diagram shows a typical corn field during the peak-growing season in the US Corn Belt, which usually has a closed canopy.

data to optimize the estimation of crop productivity? In this study, we analyze a diverse set of satellite data introduced above to quantify their shared and unique contributions for estimating crop yields. Though other satellite data may also be available and useful for monitoring crop growth, e.g. Light Detection and Ranging (Lidar) or hyperspectral data, here we only consider data that have relatively long-term time series and provide consistent coverage over a broad region of the U.S. Corn Belt. We adopt a statistical inference approach called "Partial Least Square Regression" (PLSR) to disentangle commonly shared and unique individual information from a number of predictor variables that are highly correlated or collinear (see details in Section 2.3). Our study area encompasses the U.S. Corn Belt, which contributes about 50% of global maize and 40% of global soybean production, and is one of the most critical food production regions in the world.

## 2. Materials and methods

In this section, we first describe the benchmark data (Section 2.1) and different satellite and climate data used in our study (Section 2.2). We then describe the methodology (i.e. PLSR, Section 2.3) and experimental design (Section 2.4).

### 2.1. Benchmark data: Crop-yield based NPP (Y-NPP)

We use county-level crop yield and area statistics from the National Agricultural Statistics Service (NASS) of the U.S. Department of Agriculture (USDA) to derive county-level net primary production (NPP). We convert the NASS crop yields to the area-weighted growing season NPP at the county level using the following equation detailed in Lobell et al. (2002) and Guan et al. (2016):

$$\text{NPP} \left( {}^{g\,C}\!/_{m^2} \right) = \sum_{i=1}^{N} \frac{Y_i \cdot MRY_i \cdot (1-MC_i) \cdot 0.45 \frac{g\,C}{g} \cdot \text{Harv\_Area}_i}{HI_i \cdot fAG_i} \Big/ \sum_{i=1}^{N} \text{Harv\_Area}_i \tag{1}$$

where $Y_i$ is the reported yield of crop $i$, $MRY_i$ is the mass per unit of reported yield for crop $i$, $MC_i$ is the crop moisture content (i.e. the percentage of water in the total grain weight), $\text{Harv\_Area}_i$ is the harvested area, $HI_i$ is the Harvest Index, and $fAG_i$ is the fraction of NPP allocated aboveground (fAG). The crop-type-specific values for MRY, MC, HI and fAG are provided in Guan et al. (2016).

We use the above approach to unify various crop types into a consistent productivity metric (i.e. NPP). We only focus on counties where the total crop fraction is >40% of the total area and the associated corn fraction exceeded 20% of the area – this screening process narrowed our study domain to essentially the core part of the Corn Belt (Fig. 2a and b), which produced 73% of the total US corn production during the 2006–2010 period. Choosing only high crop fraction counties reduces potential uncertainty due to assumptions of spatially consistent yields (unit area crop productivity) within a given county, and also reduces potential errors introduced by other natural vegetation types. Soybean and corn account for >90% of all crops in our study area; therefore Y-NPP is only calculated for these two crop types. We assume a constant Harvest Index (HI) for each specific crop type in Eq. (1), which essentially means that Y-NPP is a linear scaling of raw crop yield. The validity of the constant HI assumption is further discussed in Section 4.

### 2.2. Satellite-based datasets and climate data

#### 2.2.1. MODIS EVI

The Enhanced Vegetation Index (EVI) is derived from visible-NIR remote sensing and is a widely-used VI based on the leaf red-edge spectral feature in the red and NIR spectral bands. The EVI is similar to the NDVI, but sensitive to higher canopy LAI and less affected by atmospheric aerosol impacts (Huete et al., 2002). The formulation of the EVI is

(Huete et al., 2002):

$$\text{EVI} = G \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + C_1 \times \rho_{red} + C_2 \times \rho_{blue} + L} \tag{2}$$

where $\rho$ is atmospherically corrected surface reflectance, L is the canopy background adjustment addressing nonlinear radiant transfer through a canopy for NIR and red spectra, and $C_1$ and $C_2$ are coefficients that adjust for aerosol influences in the red band using blue band information (Huete et al., 2002). EVI has a general linear relationship with fPAR at the biome-specific level (Myneni et al., 2002) and can be used as a proxy for unstressed canopy-level photosynthetic capacity (Sellers et al., 1992). Here we use EVI from the NASA Terra MODIS MOD13C1 (Collection 5) record, with 16-day repeat and global 0.05° resolution. The 16-day time series is interpolated to the daily step using a robust smooth algorithm (Garcia, 2010) and then aggregated to the monthly step for this study.

#### 2.2.2. GOME-2 SIF-based GPP

We use a newly developed SIF-based GPP product (Guan et al., 2016) derived from SIF retrievals near the $\lambda = 740$ nm spectral window from the Global Ozone Monitoring Experiment-2 (GOME-2) instrument onboard the Eumetsat MetOp-A satellite. The SIF retrieval algorithm (Joiner et al., 2013) disentangles three spectral components near the peak of the far-red chlorophyll fluorescence emission feature: atmospheric absorption (due to water vapor), surface reflectance, and fluorescence radiance. To derive GPP from the raw GOME-2 SIF (version 26) record, we first scale the raw SIF retrieval to the electron transport rate (ETR) of photosynthesis based on a derived relationship (Zhang et al., 2014) calibrated using the SCOPE model at five eddy-covariance flux tower sites in the U.S. Corn Belt (Guan et al., 2016); we then scale ETR to GPP by taking into account the different photosynthetic pathways ($C_3$ and $C_4$) for cereal and broadleaf crops, and the temperature effect and operating $CO_2$ concentration inside the leaf (Guan et al., 2016). The resulting SIF-GPP record has been found to have better performance in capturing spatial and temporal patterns of crop productivity compared with other GPP observational benchmarks, including MODIS (MOD17A2) and MPI-MTE products (Guan et al., 2016). The GOME-2 SIF-GPP gridded product extends from 2007 onward, with 0.5° spatial gridding and monthly compositing interval for the study domain.

#### 2.2.3. Thermal-based ALEXI ET

We use an advanced ET product derived using the Atmosphere-Land Exchange Inverse (ALEXI) model (Anderson et al., 1997, 2007). ALEXI uses the morning rise in surface radiometric temperature, remotely sensed using geostationary satellites, as the major inputs for estimating ET. It employs a two-source (soil + vegetation) model of the surface energy balance, with a simple model of atmospheric boundary layer (ABL) development (McNaughton and Spriggs, 1986) to provide energy closure. In other words, ALEXI follows the principle that wetter landscapes warm less rapidly during the morning hours. A previous study (Anderson et al., 1997) demonstrates that use of a time-differential temperature signal reduces the errors in absolute temperature retrievals, increasing the accuracy of ET estimation. Along with thermal IR retrievals of land surface temperature, ALEXI also uses MODIS LAI as input to guide soil/canopy partitioning. The reported errors in ALEXI-ET at daily time steps are 10–20% across a broad range of vegetation and climate conditions (Anderson et al., 2007; Cammalleri et al., 2014). The ALEXI ET product has 4-km resolution and monthly temporal fidelity from 2001 onward for the conterminous US.

#### 2.2.4. QuikSCAT Ku-band radar backscatter

QuikSCAT was a Ku-band (13.4 GHz, or 2.1 cm wavelength) satellite radar scatterometer operating from 1999 to 2009, with two rotating pencil beam antennas operating in H and V polarizations at incidence angles of 55.8° and 46.8° (Long and Early, 2001). The Ku-band
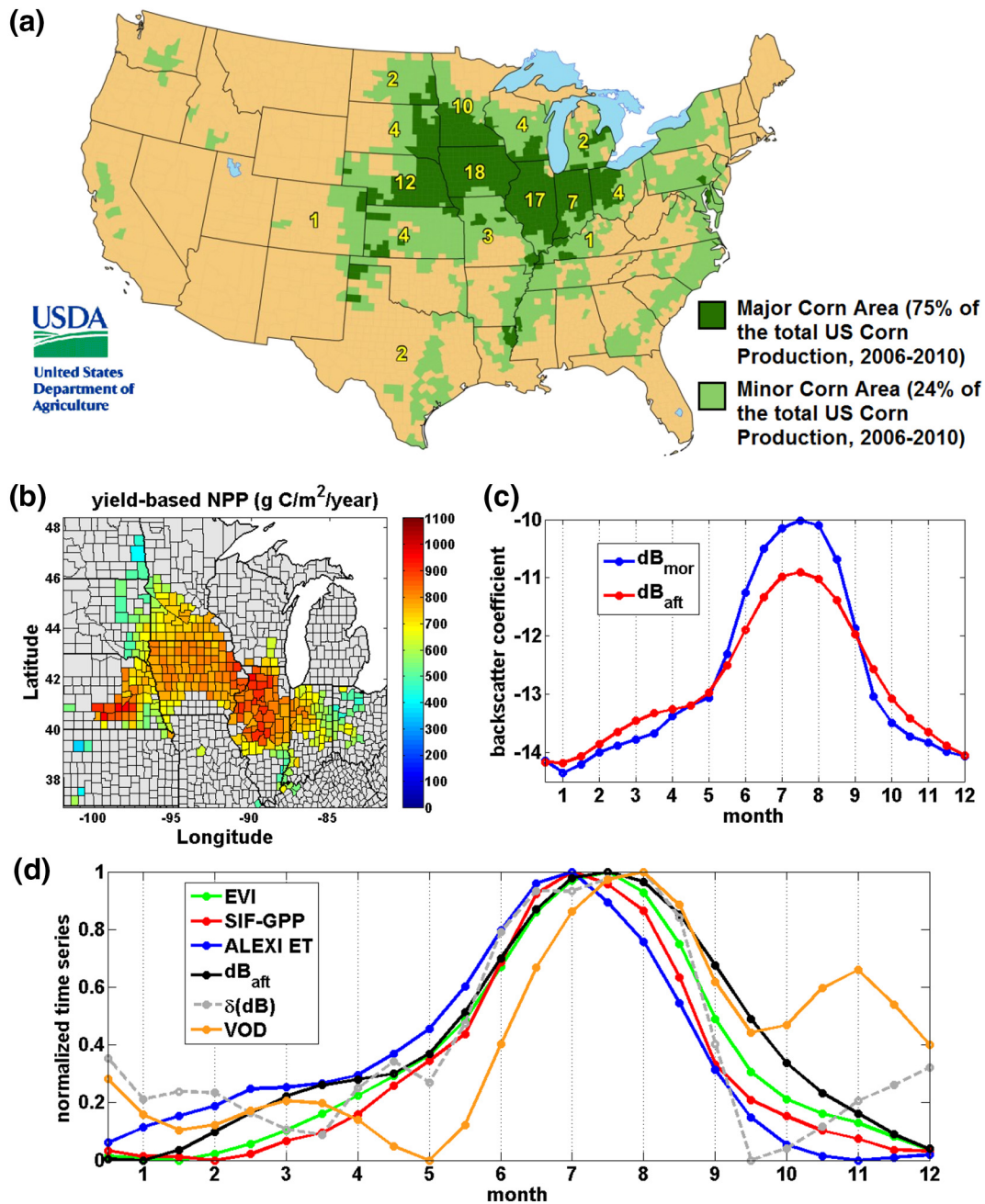
**(a)**



Major Corn Area (75% of
the total US Corn
Production, 2006-2010)

Minor Corn Area (24% of
the total US Corn
Production, 2006-2010)

**(b)** yield-based NPP (g C/m²/year)

**(c)**

**(d)**

Fig. 2. (a) The extent of the US corn production area, where the dark green shading refers to the "Major Corn Area", which includes 75% of the total US corn production based on 2006–2010 USDA statistics, and light green shading refers to the "Minor Corn Area", which includes 24% of US corn production (Copyright: CTG Publishing under Creative Commons Attribution 3.0 Unported License). (b) Study area used here (core counties of the US Corn Belt, defined as those with total crop area fraction >40% and corn area fraction >20%) shown in yield-based NPP. Based on (a), the current study includes the area that produced 73% of the total US corn production during 2006–2010. (c) Average seasonality of QuikSCAT Ku-band dB from morning and afternoon overpasses aggregated in absolute units (backscatter coefficient in dB). (d) Average seasonality of different satellite remote sensing data aggregated for the study, with all time-series normalized from 0 to 1 to match their minimum and maximum values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

backscatter product (hereafter denoted as "Ku-band dB") from QuikSCAT has ascending (morning, c.06:00 local equatorial crossing time) and descending (afternoon, c.18:00 local equatorial crossing time) overpasses. The radar-backscatter is primarily sensitive to the dielectric property and surface roughness of the landscape. Depending on the wavelength and canopy density, short-wavelength radars (e.g. Ku, X, or C-band) are relatively more sensitive to upper canopy water content and biomass structure (Ulaby et al., 1982; van Emmerik et al., 2015), whereas longer-wavelength radars (e.g. L-band) are sensitive to a larger canopy volume and surface soil moisture under low to moderate vegetation cover (Entekhabi et al., 2014; Njoku et al., 2003). It has

been found that morning and afternoon overpasses of the Ku-band radar backscatter data from QuikSCAT may sense diurnal variations in canopy water content; in the morning prior to active transpiration, canopy water content is at a daily maximum level due to overnight moisture replenishment from the soil rooting zone and xylem water storage, while canopy water is at a daily minimum in the mid-afternoon due to cumulative canopy water loss from daily transpiration (Frolking et al., 2012; Saatchi et al., 2013). Thus the difference between morning and afternoon radar backscatter may be related to canopy transpiration and water stress (van Emmerik et al., 2015). Some empirical studies support this hypothesis in the tropical forests (Frolking et al., 2012;

Madsen and Long, 2016; Saatchi et al., 2013), but no study has been done for croplands. Here we use QuikSCAT Ku-band dB data from local (6 AM/PM) morning and afternoon satellite overpasses (denoted as $dB_{mor}$ and $dB_{aft}$), as well as the difference between the two overpasses (denoted as $\delta(dB)$) to study their use for crop yield estimation. Since the H- and V-polarization data from QuikSCAT show little difference for our analysis, the V-Pol data are selected for the final investigation. We use QuikSCAT data processed by the NASA Scatterometer Climate Record Pathfinder (SCP) project located at Brigham Young University (http://www.scp.byu.edu/). The QuikSCAT Ku-band dB data record extends from 1999 to 2009, with global coverage at 0.25° resolution and 3-day temporal fidelity, while the monthly average backscatter for July is used in this study.

### 2.2.5. AMSR-E vegetation optical depth (VOD)

We use a recently developed satellite passive-microwave-based vegetation optical depth (VOD) product (Jones and Kimball, 2012) derived from daily brightness temperature retrievals ($T_b$) from the NASA Advanced Microwave Scanning Radiometer for EOS (AMSR-E). The VOD is a frequency-dependent measure of the vegetation canopy attenuation of land surface microwave emissions, that is sensitive to vegetation canopy biomass, structure and water content (Jones et al., 2013; Shi et al., 2008; Ulaby et al., 1982). The VOD has been used as an additional canopy phenology measure that is sensitive to both photosynthetic and nonphotosynthetic (e.g., woody) biomass (Guan et al., 2012; Guan et al., 2014; Jones et al., 2011; Liu et al., 2015; Liu et al., 2013) and plant-level water potential (Konings and Gentine, 2016). Here we use the near-daily 10.7 GHz frequency (X-band) VOD retrievals at the constant incidence angle of 55° from nadir from AMSR-E (2003 − 2011) (Du et al., 2014). The 0.25° VOD record at the 3-day frequency is used here, and a robust smoothing algorithm (Garcia, 2010) is applied to gap-fill missing values and smooth the VOD time series spanning the AMSR-E record.

### 2.2.6. Climate data

We use the Parameter-elevation Relationships on Independent Slopes Model (PRISM) dataset of temperature and precipitation across the conterminous U.S. (Daly et al., 2008). The PRISM data is derived using daily in situ observations from approximately 13,000 precipitation stations and 9800 air temperature stations, which are spatially interpolated using digital elevation data and empirical modeling to derive climate fields at 30 s spatial resolution. We use the daily, 4 km version of the PRISM data record aggregated to a monthly time step for the analysis.

### 2.2.7. Data processing

All of the gridded satellite and climate data (Table 1) are processed to extract county-level data based on a shapefile of U.S. county boundaries. The satellite and climate data used in this study cover various time periods, which is discussed in the Experimental Design Section 2.4 below.

### 2.3. Partial Least-Square regression (PLSR)

We use PLSR in this study (Geladi and Kowalski, 1986; Wold et al., 2001; Wu et al., 2016) to evaluate relationships between crop yield and the satellite derived metrics/climate variables. The yield-based productivity (Y-NPP) is the only response variable, and the various satellite data (and climate variables in some cases) are used as predictor variables. PLSR, as a dimension reduction approach, is suitable for distinguishing the shared and unique components among a large set of potential predictor variables that are mutually correlated. PLSR is similar to Principle Component Regression (PCR) – in that both approaches attempt to model a response variable from a large number of predictor variables that are correlated or collinear. Both approaches construct new predictor variables (known as "latent components" or "latent variables", hereafter termed "components") as linear combinations of the original predictor variables. However, they construct components in a different fashion. The components constructed in PCR are to maximize the variability in the predictor variables themselves, without taking into account the response variable. Essentially this is the principle components analysis for the predictor variables. However, the components in PLSR are constructed to optimize the explained power of each component in predicting the response variable(s); as a result, PLSR often leads to models with fewer components than PCR, and in many cases PLSR captures more variability in the response variable(s) than the PCR. PLSR is considered more suitable than PCR in our case, since a major goal of this study is to disentangle the unique contributions of the different satellite and climate inputs in predicting crop yield.

The following steps outline the PLSR implementation in our study. First, each variable is normalized to a standardized anomaly. This is accomplished by pooling all county-level data points over all years in the period of analysis, and computing the mean and standard deviation (std) over the full pool. The normalization is then done by calculating the following metric: (raw-mean)/std. In this way, the data are normalized over the full spatial and temporal domain encompassed by the study. This preprocessing step ensures that all variables have mean zero and unit standard deviation, so that all of the fitted coefficients in the PLSR are comparable. We then perform the following steps: (1) Outliers are identified and filtered following the Monte-Carlo sampling method for outlier detection (Xu and Liang, 2001), which removes ~5% of the original data. (2) The filtered dataset is then randomly split into training (70%) and testing (30%) datasets. (3) The number of components is then selected by running a 10-fold cross-validation 100 times with a different random split of the filtered dataset for each run. (4) The trained regression coefficients are then applied to the testing dataset, and the resulting model performance is assessed using the root mean squared error (RMSE) (mean Euclidean distance between the prediction and the observed crop yield), and $R^2$ (the proportion of variability in crop yield explained by model) metrics calculated with the test data. The major uncertainty in the current PLSR protocol arises from the randomness in splitting between training and testing data. However we find that the regression coefficients derived from the different data

**Table 1**
Specifics of the satellite products and climate data used in this study.

| Satellite/Climate data | Spectral range | Spatial resolution | Temporal resolution | Available record | Reference |
|---|---|---|---|---|---|
| MODIS EVI | Visible and NIR | 0.05°(MOD) | 16-day | 2000–present | Huete et al., 2002 |
| GOME-2 SIF-GPP | NIR (far-red chlorophyll fluorescence emission) | 0.5° (gridded) | Monthly (gridded) | 2007–present | Guan et al., 2016; Joiner et al., 2013 |
| ALEXI-ET | Thermal (Land Surface Temperature) VIs and NIR (LAI) | 4 km | daily | 2001–present | Anderson et al., 1997; 2013b |
| QuikSCAT Ku-band backscatter | Ku-band (13.4 GHz) | 0.25° | 3 days | 1999–2009 | Long et al., 2001; Guan et al., 2012 |
| AMSR-E VOD | 10.7 GHz | 0.25° | 3 days | 2002–2011 | Jones and Kimball, 2012 |
| PRISM precipitation and temperature | N/A | 4 km | daily | 1981–present | Daly et al., 2008 |

splitting schemes are consistent in terms of their general pattern and relative magnitude. The $R^2$ shows some sensitivity to the data splitting, but this uncertainty is largely eliminated by conducting the 10-fold cross-validations for 100 times.

## 2.4. Experiment design

In this study we use all satellite and climate data for the month of July, which represents the peak growing season for the U.S. Corn Belt as indicated from past studies (Lobell et al., 2013) and also clearly shown in the seasonal cycles of the different satellite data records (Fig. 2). In addition, July has been found to be the most critical period for climate to affect crop yield (Lobell et al., 2013). Finally, using July data also provides lead time for forecasting crop yield that is usually harvested in October in the U.S. Corn Belt. When designing the experiments for the PLSR analysis, we consider the following factors:

(FACTOR 1) Data availability: The satellite data used in this study cover different periods. We therefore select two major periods for the analysis: Period 1 (2007–2009) is relatively short in length but included all of the satellite data examined; Period 2 (2002–2009) had a longer duration which is more suitable for analyzing model predictive skill and inter-annual variations, but it excludes the SIF-GPP record. Note, that Period 1 (2007–2009) experienced relatively low inter-annual variability in drought conditions across the Corn Belt, in comparison with the longer timescale represented in Period 2 (2002–2009).

(FACTOR 2) Distinction between spatial and temporal patterns: We consider three options, as illustrated in Fig. 3. Option 1 ("Raw") uses all raw data and treats each county-year record as an independent sample, which essentially lumps both spatial and temporal variability. Option 2 ("Detrend") first removes linear trends of the raw data (both yield and predictor variables) at the county level (i.e. detrending) such that the average of the county data is zero, and then treats each of the county-year records as an independent sample. This option thus removes spatial variability at the county level, while preserving temporal variability in the data. Comparing results between Option 1 and 2 can reveal distinctions in explaining spatial and temporal variability. Option 3 ("Detrend + Mean") first calculates the average of the raw county data for each county, and then adds this county-specific average to the "Detrend" data of Option 2. Option 3 thus does not have the multi-year linear trend when compared with Option 1, but preserves the between-county spatial pattern when compared with Option 2.

(FACTOR 3) Evaluating impacts from the different combinations of satellite and climate inputs. We use the following four options to distinguish impacts of the different satellite and climate inputs on estimated yield: Option 1: only using satellite data ("Satellite Only"); Option 2: using both satellite data and climate data ("Satellite + Climate"); Option 3: using EVI data and climate data ("EVI + Climate"), and Option 4: only using climate data ("Climate Only"). Option 3 is based solely on EVI since most conventional studies only use VIs in yield estimation

(Sibley et al., 2014; Tucker and Holben, 1980), while comparisons between Options 2 and 3 can reveal the potential added predictive power of other satellite data in estimating crop yield. Our investigation primarily examines the climate record for July, though the May–July aggregated climate is also tested but does not produce noticeable qualitative differences from the July results.

The full implementation of the above options represents 24 ($2 \times 3 \times 4$) total experiments. We primarily focus on the Period 1 (2007–2009) results because this period includes all of the satellite data examined. The Period 1 results are also found to be largely consistent with those of Period 2 (2002–2009). The Period 2 temporal analysis results are also presented due to the longer record, even though it excludes the shorter SIF-GPP record. In the Results section, we use the following naming convention for the different experiments: "FACTOR1, FACTOR2, FACTOR3". For example, the experiment denoted as "2002–2009. Satellite Only.Raw" means that we are using only satellite data (FACTOR 3) from the 2002–2009 period (FACTOR 1), and the Raw county level data (FACTOR 2).

## 3. Results

### 3.1. Seasonal cycle of different satellite-based data

Fig. 2 shows the study area and spatially aggregated mean seasonal cycles of all satellite data examined for the study area. All of the satellite data shows a generally consistent pattern of a mid-summer peak during the crop growing season for the U.S. Corn Belt (May-Sep), though with some differences in temporal onset, rate of increase, peak timing and rate of decrease among the different metrics. SIF-GPP slightly lags behind EVI in the green-up stage and also drops earlier than EVI in the senescence stage, which is consistent with previous findings that crop greenness (revealed by EVI) generally leads photosynthesis (revealed by SIF-GPP) at the beginning of the growing season and lags photosynthesis in the senescence stage over North America (Joiner et al., 2014). We also find that ALEXI-ET increases earlier than the other metrics at the beginning of the growing season when bare soil evaporation is expected to be the major contributor of this signal. The ALEXI-ET drops earlier than the other metrics and closely tracks SIF-GPP, indicating that ET and photosynthesis are synchronized during the latter portion of the growing season. The microwave VOD and dB seasonal cycles both lag behind the other metrics at the end of the growing season, while VOD also lags at the beginning of the growing season. This lagged response of the microwave data (both VOD and dB) has also been reported for other ecosystems, including grassland and shrublands of North America (Jones et al., 2012), tropical savannas (Guan et al., 2014) and temperate forests (Jones et al., 2011). These studies found that VOD lags in spring to the rapid greening shown by VI, and this is because that above ground biomass and water contents reflected by VOD require time to accumulate and grow (Jones et al., 2012). The reported delayed VOD decrease during Fall (relative to VI) is largely due
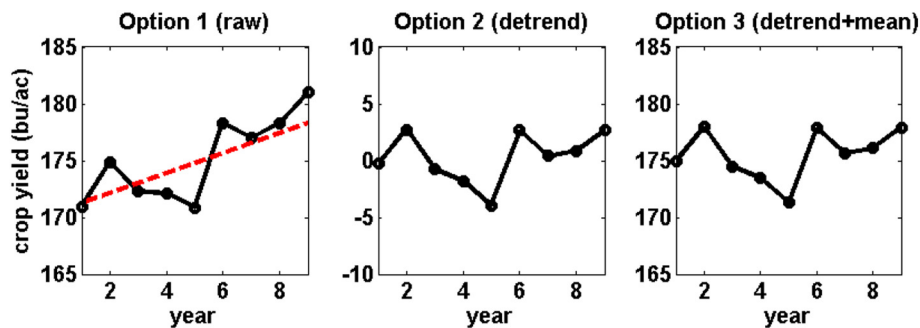


**Fig. 3.** A synthetic example to illustrate Options 1–3 for the "Distinction between spatial and temporal patterns" in Section 2.4 "Experiment design" for crop yield data in a specific county. Option 1 uses raw county-level data (the red dashed line shows its linear fit), Option 2 removes the linear trend (the red dashed line shown in the 1st panel) from the raw data for each county, and Option 3 adds the multi-year average from the raw county-level data to the Option 2 data.

to residual biomass remaining during canopy senescence (Guan et al., 2014; Jones et al., 2011; Saatchi et al., 2013). Both $dB_{mor}$ and $dB_{aft}$ show similar seasonal cycles (Fig. 2c), though $dB_{mor}$ shows a larger magnitude signal than $dB_{aft}$. Previous studies have hypothesized that the morning/afternoon difference in Ku-band dB indicates that plants replenish water overnight and lose water over the day through transpiration (Frolking et al., 2011; Saatchi et al., 2013), though this difference may also reflect diurnal differences in canopy structure (e.g. leaf wilt) and dewfall (Vereecken et al., 2012). The difference between $dB_{mor}$ and $dB_{aft}$ (i.e. $\delta(dB)$) also shows a clear seasonal cycle (Fig. 2c in gray), which is largely following EVI and SIF during the growing season. The regionally aggregated pattern shown in Fig. 2 is consistent with the spatial pattern of the peak month of the mean seasonal cycle for all the satellite data (Fig. S1), indicating that the aggregated pattern that we discussed above is representative for the study domain.

## 3.2. Exploratory data analysis

We first conduct the exploratory data analysis by looking at correlations among all of the satellite/climate metrics and yield-based NPP using the 2007–2009 raw data (Fig. 4). Below we summarize the findings using single asterisk (*) and double asterisks (**) to indicate a correlation coefficient (r) with statistical significance levels of $p$-value < 0.01 and $p$-value < 0.001, respectively. No asterisk means the significance level is above 0.01.

(1) SIF-GPP, EVI, $dB_{mor}$ and $dB_{aft}$ are highly correlated with each other, with all correlation coefficients exceeding 0.69**. These four metrics are also highly correlated with Y-NPP (with all

correlations exceeding 0.62**), while $dB_{aft}$ has the highest correlation with Y-NPP ($r = 0.66^{**}$) relative to the other metrics examined.

(2) VOD and ALEXI-ET are relatively less correlated with the data of the above group, but all of the correlation coefficients are still significant and above 0.47**. VOD and ALEXI-ET are also less correlated with Y-NPP ($r = 0.42^{**}$ for VOD, and $r = 0.47^{**}$ for ALEXI-ET). $\delta(dB)$ is even less correlated with other satellite data and with Y-NPP than all other satellite-based measurements.

(3) Mean July temperature (T), in general, has a significantly negative correlation with all of the satellite data and Y-NPP, except with ALEXI-ET. ALEXI-ET and T are positively correlated at a high significance level ($p < 0.001$).

(4) Mean July precipitation (Precip) generally has a significant positive correlation with EVI, ET, $dB_{mor}$, $dB_{aft}$, $\delta(dB)$ and Y-NPP. Precip and T do not have a significant correlation in the study region.

Fig. 5 shows the three-year-averaged (2007–2009) spatial pattern of the different metrics, which provides a qualitative visual justification of some of the above findings, for example, the similar spatial patterns among SIF-GPP, EVI, $dB_{mor}$, $dB_{aft}$, ALEXI-ET and Y-NPP.

## 3.3. PLSR results

### 3.3.1. Interpreting the PLSR results

Here we first demonstrate how we interpret the PLSR results. Using the "2007–2009. Satellite Only. Raw" (Fig. 6a and b) results as an example, Fig. 6a shows the performance of the PLSR models with different
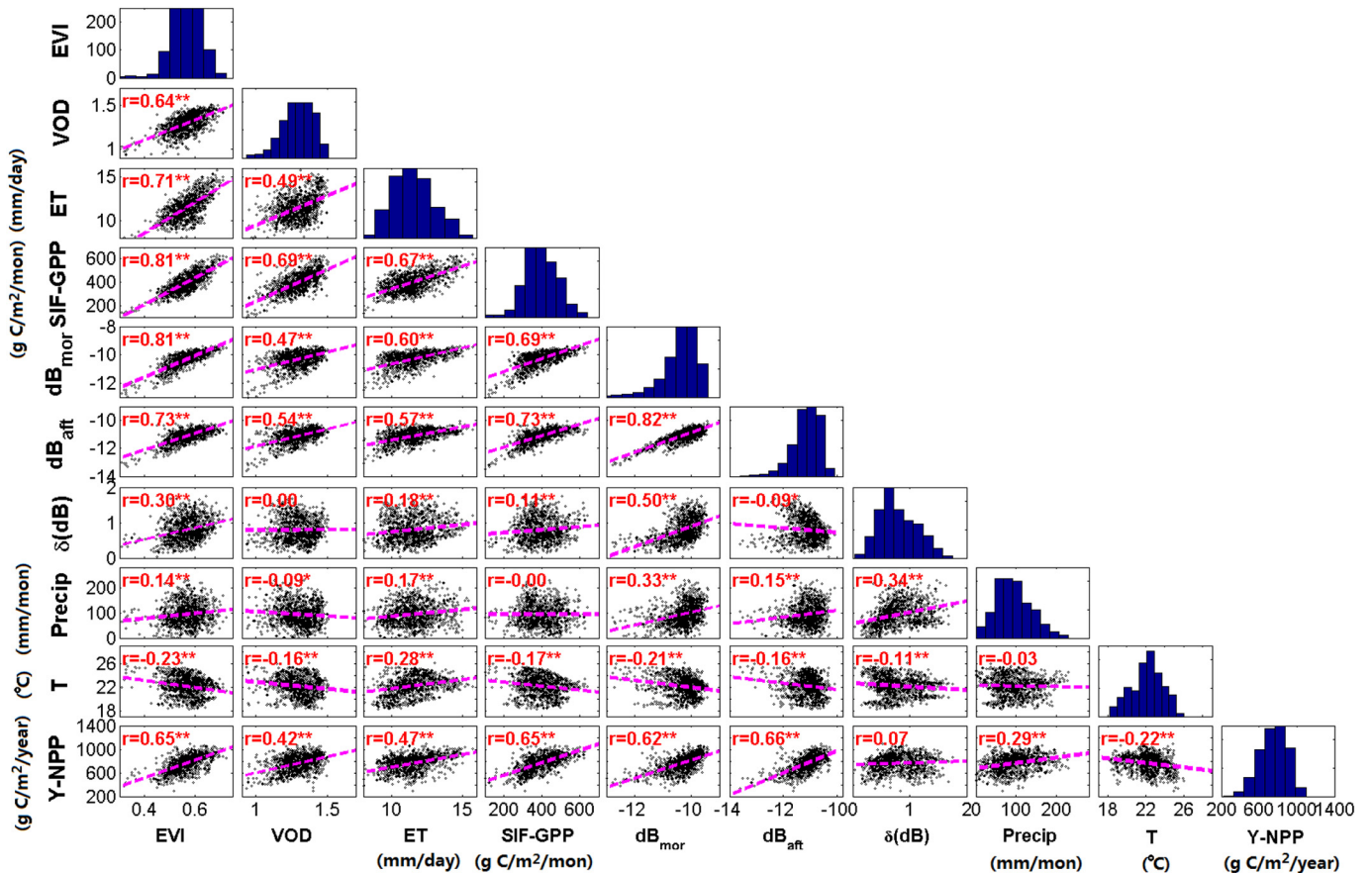


**Fig. 4.** Scatterplots between different variables including various satellite data, climate (precipitation and temperature for July), and crop yield-based July NPP for all counties examined for the 2007–2009 period. Each dot in the scatterplot represents a single county-year record (July only). Three-year-aggregated (2007–2009) spatial patterns after the normalization are shown in Fig. 5. Single asterisk (*) and double asterisks (**) denote statistical significance levels of $p$-value < 0.01 and $p$-value < 0.001, respectively; no asterisk indicates significance levels above 0.01.
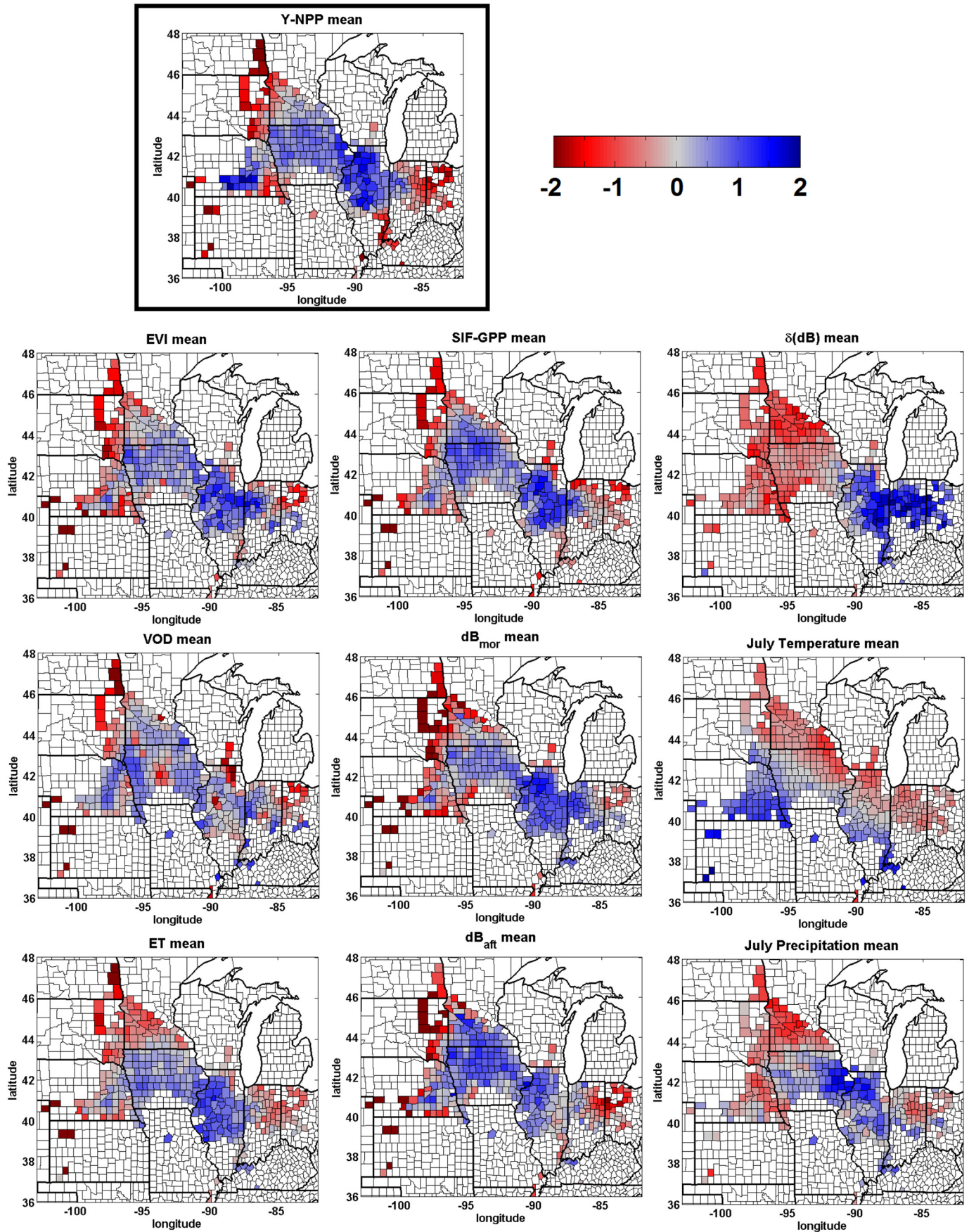
**Fig. 5.** Three-year-averaged (2007–2009, July) spatial patterns after normalization (all of the data have a mean of zero and standard deviation of one; see details in Section 2.3) of the satellite-based data, climate variables, and Y-NPP for all counties in the study region.
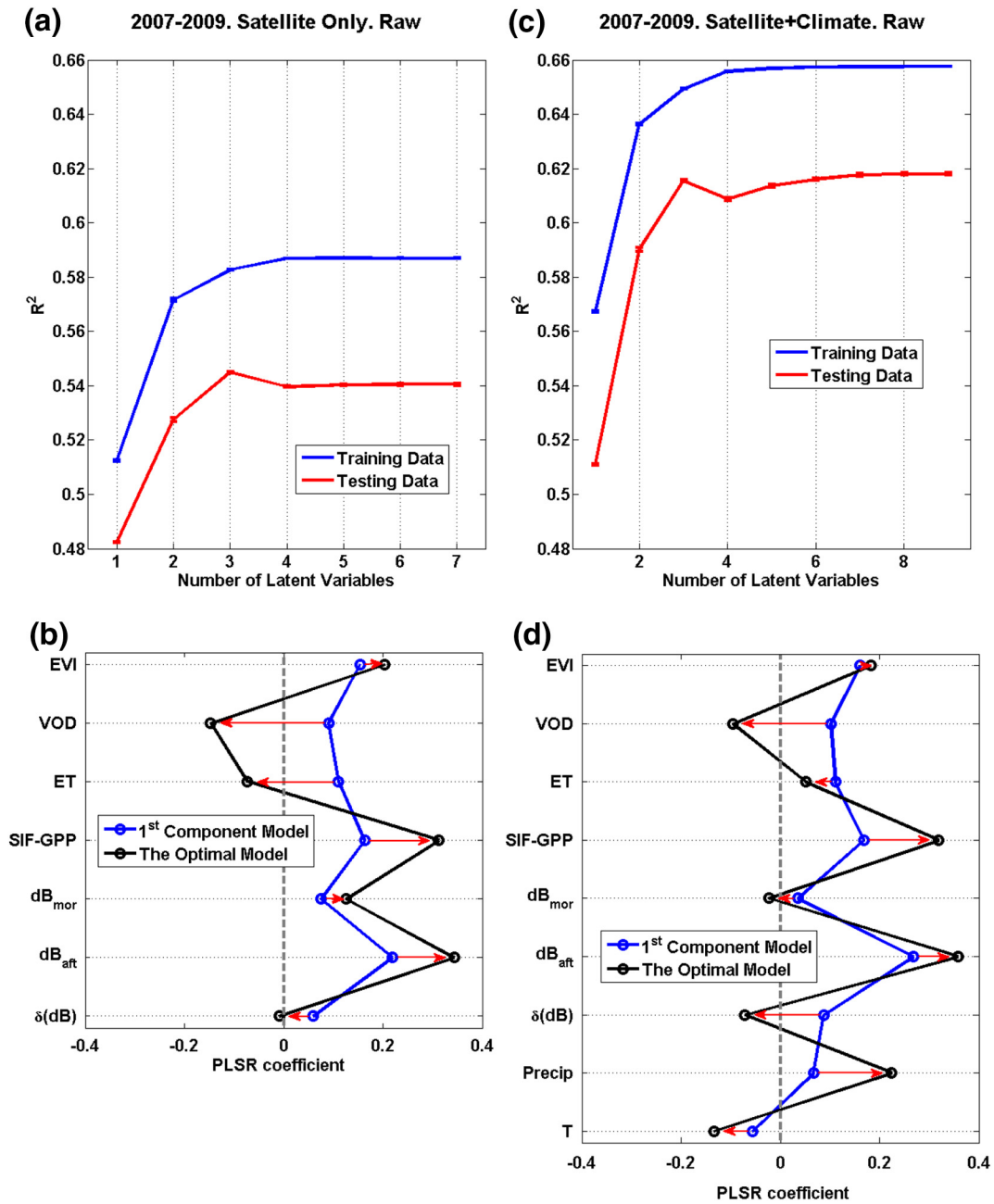
**Fig. 6.** The PLSR results for the two experiments, i.e. "2007–2009. Satellite Only. Raw" (6a, b) and "2007–2009. Satellite + Climate. Raw" (6c, d). The upper panels (6a, c) show the performance of the PLSR models with different numbers of components (i.e. $R^2$ between the model prediction and observation, with error bars showing the 95% confidence interval). The bottom panels (6b, d) show the regression coefficients of each predictor variable for the model with only the 1st component (in blue) and the optimal model (in black), respectively, with the red arrows showing the direction and magnitude of their difference for each predictor variable between the two models. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

numbers of components (i.e. $R^2$ between the model prediction and observation). Specifically, the 1st component is a vector derived as a linear combination of the different satellite data such that the 1st component can best explain the variance in crop yield (i.e. Y-NPP). The 2nd component is a perpendicular vector to the 1st component, and is the linear combination of the different satellite data that can best explain the remaining variance in Y-NPP after removing the part of Y-NPP explained by the 1st component. The same rule applies to the 3rd component, which needs to be perpendicular to both the 1st and 2nd components, and so on for subsequent components. Since each component is a linear combination of the different satellite data, the weight of each data input in the linear combination is the coefficient shown in Fig. 6b.

As expected, the $R^2$ for the training data increases with more components until saturation, and it is always higher than the counterpart $R^2$ for the testing data. The $R^2$ for the testing data increases from 0.48** for the model with only one component to 0.55** for the model with three components. The model performance slightly decreases ($R^2 = 0.54$**) using additional components, indicating model overfitting of the training data when the number of components is more than three. The $R^2$ for the testing data becomes stabilized when the number of components is more than four. Thus the optimal PLSR model in this case has three components.

Fig. 6b shows the regression coefficients of each predictor variable for the model with only one component (in blue) and the optimal model (in black), respectively. The red arrows in Fig. 6b show the

magnitude and direction of change in the coefficients between the two models. The regression coefficients for the model with only one component are all positive, which is explained in Section 3.3.3. However, the regression coefficients for the optimal model (in this case the PLSR model with three components) have clear differences from the model with only one component: a few predictor variable coefficients increase (EVI, SIF-GPP, $dB_{mor}$, and $dB_{aft}$), while others decrease and become negative (VOD, ET, and $\delta(dB)$).

Understanding the meaning of each PLSR component and how they can be reconstructed is essential to interpreting the PLSR results. The first PLSR component is a linear combination of all of the predictor variables weighted by the regression coefficients of the model with only one component. The second PLSR component can be calculated as the linear combination of all of the predictor variables but weighted by the coefficient difference between the two-component PLSR model and the one-component PLSR model. The same rule applies for other

PLSR components. Thus for Fig. 6b, the linear combination of all of the predictor variables weighted by the red vector arrows is equal to the sum of the 2nd and 3rd components, since the optimal model has three components.

### 3.3.2. "Satellite Only" vs. "Satellite + Climate"

Fig. 6 shows the comparison of results from the two experiments: "2007–2009. Satellite Only. Raw" (6a, b) vs. "2007–2009. Satellite + Climate. Raw" (6c, d). The only difference between these two experiments is whether or not the climate information is used. Both experiments attain the best performance using three components. Including climate information improves the $R^2$ performance to 0.61** (vs. 0.55** without climate information involved) for the optimal models in the testing data. The general trends of the regression coefficients for the one component model are similar in both experiments (blue lines in Fig. 6b, d). The regression coefficients for the optimal models also share similar
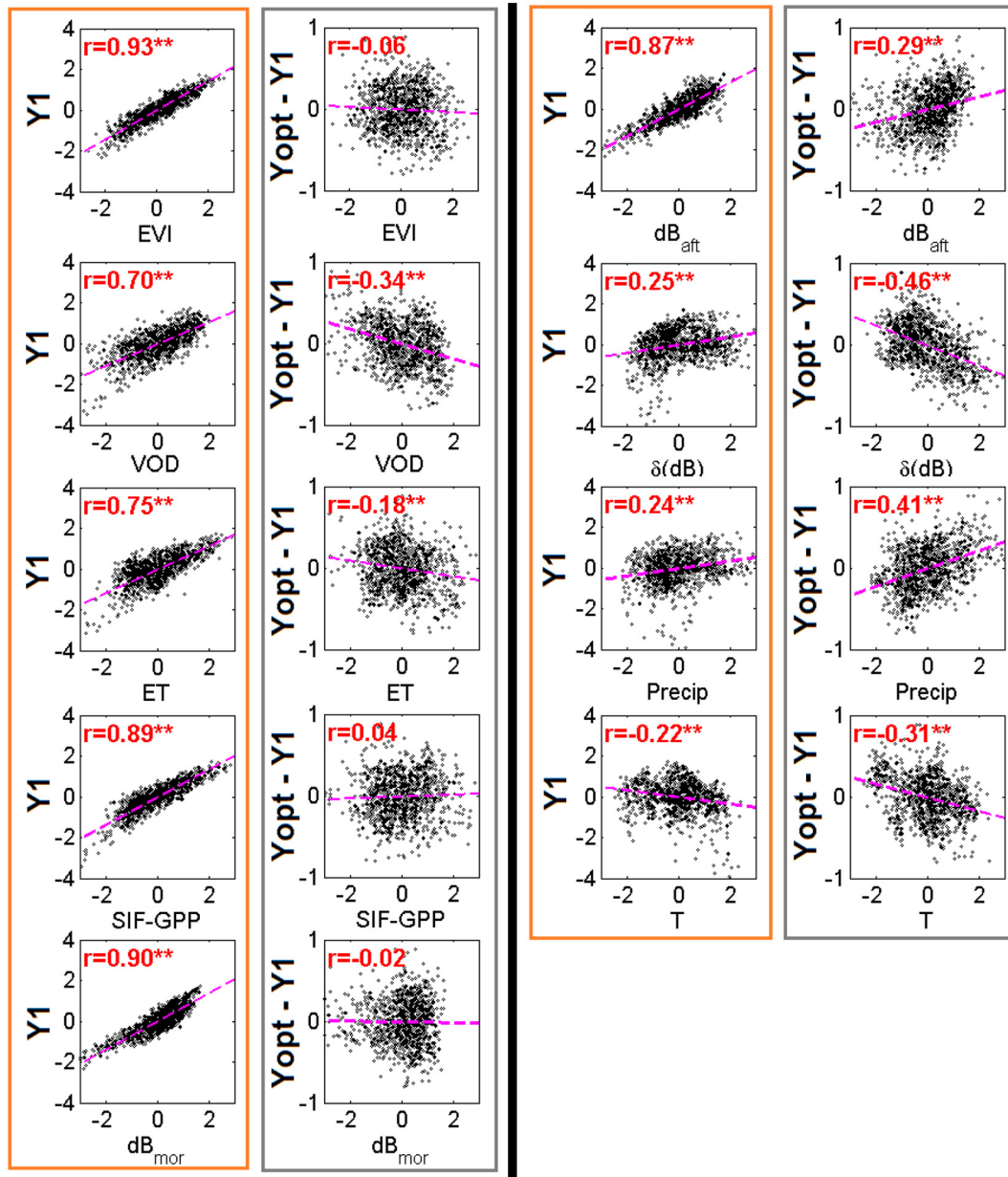


**Fig. 7.** Linear correlations between the first PLSR component (denoted as "$Y_1$") and the different predictor variables (1st column on both sides, in the orange box), and linear correlations between the added values of other PLSR components ("$Y_{opt}$ - $Y_1$") and the different predictor variables (2nd column on both sides, in the gray box); "$Y_{opt}$" refers to the predicted Y-NPP from the optimal PLSR model, and "$Y_{opt}$ - $Y_1$" refers to the extra information from other components when the contribution from the first component is removed. The results are from the experiment: "2007–2009. Satellite + Climate. Raw"; Single asterisk (*) and double asterisks (**) denote statistical significance levels of $p$-value < 0.01 and $p$-value < 0.001, respectively.

trends, except that the coefficients for ALEXI-ET in the optimal model have opposite signs in the two experiments, as well as for $dB_{mor}$. The co-efficient change for ALEXI-ET is largely attributed to the positive correlation between ALEXI-ET and T (Fig. 4); when T is incorporated in the model (Fig. 6d), T largely absorbs the variability previously carried by ET in the "Satellite Only" experiment (see more details in Section 3.3.3). The changes in coefficients from the one component model to the optimal model (red arrows in Fig. 6c, d) also show similar trends in direction, with the only exception for $dB_{mor}$, which shows opposite signs in the two models.

### 3.3.3. Shared and unique information from different predictor variables for crop yield estimation

We now use the results from the experiment "2007–2009. Satellite + Climate. Raw" (Fig. 6c, d) to distinguish both shared and unique information from the different predictor variables. Fig. 6c shows that the single component and optimal PLSR models explain 51% and 62% of the variability in the response variable (i.e. Y-NPP), respectively. In other words, the 1st component achieves 82% ($= 0.51/0.62$) of the optimal model performance, while additional components capture the remaining 18% of the performance.

We calculate the 1st component ("$Y_1$" in Fig. 7) and also the combined 2nd and 3rd model components ("$Y_{opt}$-$Y_1$" in Fig. 7, see their three-year averaged spatial patterns in Fig. 8), and then correlate these components with all of the predictor variables (Fig. 7). Our results show that the correlation coefficients between the 1st component and predictor variables closely correspond to the correlation coefficients between Y-NPP and the predictor variables (Fig. 4, $R^2 = 0.95$). This is expected as the 1st component from the PLSR model is derived to maximize the explained variability in the response variables (i.e. "Y-NPP" in this case). A ranking of the predictor variables from high to low correlation with the 1st component ranges from: EVI ($r = 0.93^{**}$), $dB_{mor}$ ($r = 0.90^{**}$), SIF-GPP ($r = 0.89^{**}$), $dB_{aft}$ ($r = 0.87^{**}$), ALEX-ET ($r = 0.75^{**}$), VOD ($r = 0.70^{**}$), and $\delta(dB)$ ($r = 0.25^{**}$). These results indicate that all of the satellite metrics contain the information carried by the 1st component, though some satellite metrics contain more of this information than others. The 1st component is also positively correlated with July precipitation ($r = 0.24^{**}$) and negatively correlated with July temperature ($r = -0.24^{**}$). The relatively higher correlation between the 1st component and EVI, dB and SIF-GPP metrics indicates that the 1st component is sensitive to cropland aboveground biomass production.

The combined 2nd and 3rd model component ("$Y_{opt}$-$Y_1$" in Fig. 7) has a small or insignificant correlation with EVI ($r = -0.05$), SIF-GPP ($r = 0.04$), $dB_{mor}$ ($r = -0.02$), but has stronger correspondence with VOD ($r = -0.34^{**}$), ET ($r = -0.18^{**}$), $dB_{aft}$ ($r = 0.29^{**}$), $\delta(dB)$ ($r = -0.46^{**}$), Precip ($r = 0.41^{**}$), and T ($r = -0.31^{**}$). In brief, the above results indicate that when the 1st component is removed, EVI, SIF and $dB_{mor}$ provide little extra information in terms of estimating crop

yield; however, the other satellite metrics and climate variables all provide non-trivial contributions to "$Y_{opt}$-$Y_1$". Specifically, we find that this relationship explains why the regression coefficients change from the single component model to the optimal model (Fig. 6d) – the change directions of the red arrows are consistent with the signs of the correlation coefficients between "$Y_{opt}$-$Y_1$" and the different predictor variables, while the magnitudes of the red arrows are also highly correlated with the magnitudes of the above regression coefficients ($R^2 = 0.83^{**}$). The interpretation of the combined 2nd and 3rd component is less intuitive than that of the 1st component, but some plausible explanations can be given based on existing knowledge and literature. In the U.S. Corn Belt during July (usually the key reproductive stage), higher precipitation usually leads to higher crop yields, while higher temperatures generally lead to yield declines (Guan et al., 2016; Lobell et al., 2013; Schlenker and Roberts, 2009); these characteristics are confirmed in the current study by the positive correlation between Precip and "$Y_{opt}$-$Y_1$", and negative correlation between T and "$Y_{opt}$-$Y_1$". Due to the relatively high correlation between ALEXI-ET and T ($r = 0.28^{**}$, Fig. 4) during this study period (2007–2009), ALEXI-ET tends to show similar information as T for "$Y_{opt}$-$Y_1$". The positive correlation between $dB_{aft}$ and "$Y_{opt}$-$Y_1$" may indicate that $dB_{aft}$ contains better information about water content in the aboveground biomass (van Emmerik et al., 2015) and associated plant water stress (Konings and Gentine, 2016). Thus, higher $dB_{aft}$ indicates higher plant water content and less water stress, which can translate to higher crop yields. Though VOD and $\delta(dB)$ both have a significant negative correlation with "$Y_{opt}$-$Y_1$", the specific processes behind these relationships remain elusive and require further research. In general, the combined 2nd and 3rd component contains information that is more related to the environmental stresses that crop experiences.

Fig. 9 further encapsulates the above interpretations by showing the PLSR loadings for the 1st and 2nd components of the PLSR model. PLSR loadings refer to the projection of each predictor variable to the PLSR components. Since the different PLSR components are perpendicular to each other, Fig. 9 only shows the first two components, which combined achieve 95% of the optimal model performance. The 2nd component shown here contains 85% of the information in "$Y_{opt}$-$Y_1$" (the combined 2nd and 3rd component) shown above, so the 2nd component can largely represent the information of "$Y_{opt}$-$Y_1$". We find that EVI, SIF-GPP, and $dB_{mor}$ all have similarly large loadings in the direction of the 1st component, but contribute little to the dimension of the 2nd component. $dB_{aft}$, VOD and ET all have a significant contribution to both "$Y_1$" and "$Y_{opt}$-$Y_1$". However, $\delta(dB)$, Precip and T primarily contribute to the dimension of the 2nd component (i.e. "$Y_{opt}$-$Y_1$"). Based on the previous analysis and interpretation, we argue that the 1st component primarily contains information about crop aboveground biomass, while the 2nd component primarily includes information about environmental stresses.
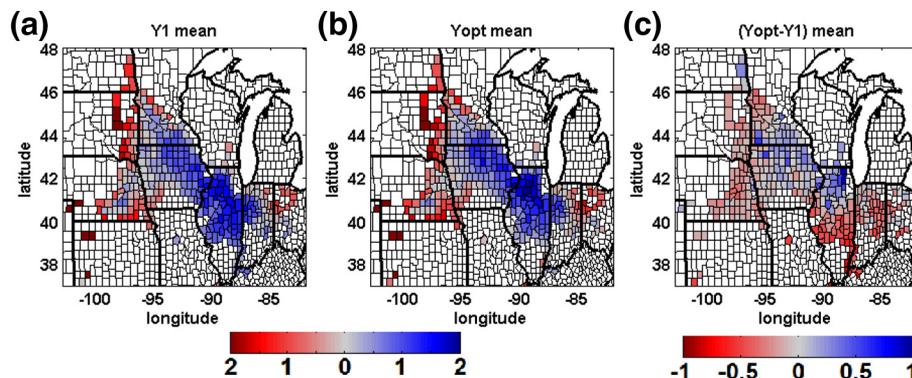


**Fig. 8.** (a) The three-year-averaged spatial pattern of the 1st component ("$Y_1$"), (b) the predicted response variable from the optimal model ("$Y_{opt}$"), and (c) the combined 2nd and 3rd component ("$Y_{opt}$-$Y_1$") for the experiment of "2007–2009. Satellite + Climate. Raw".
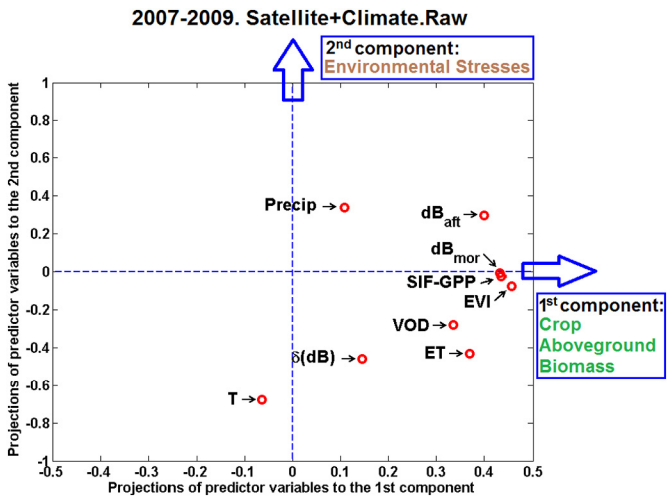
Fig. 9. PLSR loadings for different predictor variables for the 1st component (x-axis) and the 2nd component (y-axis) resulting from experiment: "2007–2009. Satellite + Climate. Raw". Based on the results, we interpret the 1st component as primarily related to "crop aboveground biomass", and the 2nd component primarily relating to "environmental stress".

### 3.3.4. "Raw" (spatial + temporal) vs. "Detrend" (temporal only) vs. "Detrend + Mean"

Next, we compare the performances of the PLSR models in order to explain spatial and temporal patterns of the observed crop yield. In accordance with the experimental design (Fig. 3), "Raw" data contains

both spatial and temporal variability, "Detrend" only contains temporal variability information, and "Detrend + Mean" removes the multi-year linear trend from the "Raw" data. Since data for the "2007–2009" period is too short for analyzing temporal characteristics and trends, we use the "2002–2009" data record here, which does not include the SIF-GPP data. However, as we have shown that SIF-GPP contains similar information as EVI for explaining county-level crop yields in the current study (e.g. Fig. 9), omission of the SIF-GPP may be largely mitigated by the use of EVI in the longer record.

The top row in Fig. 10 shows that the $R^2$ of the "Raw" experiment is higher than "Detrend", indicating that it is usually more difficult to correctly predict inter-annual yield variability than between-county yield variability. The $R^2$ of "Raw" is lower than "Detrend + Mean", meaning that the multi-year linear trend matters and that removing this trend leads to a higher $R^2$ of the predictive model. Fig. S2 shows the time trend of the satellite-based data used here, and we observe that $dB_{aft}$ and ALEXI-ET show a more negative trend in most counties, while EVI, $dB_{mor}$, VOD and $\delta(dB)$ do not show clear trend patterns. How these different trends affect the results may be worth further investigation, but is beyond the scope of the current study. However, the regression coefficients for "Raw" have a very similar trend as "Detrend + Mean", indicating that the multi-year linear trend may not affect the relative contributions of different predictor variables. The regression coefficients in the "Detrend" and "Raw" experiments have some similarities, e.g. EVI, T, $\delta(dB)$ and $dB_{aft}$ have large absolute magnitudes in regression coefficients for both experiments, indicating that these predictor variables are effective in explaining both temporal and spatial variability in crop yield. However, there are also clear differences in the regression coefficients between the "Raw" and "Detrend" results. Specifically, Precip has
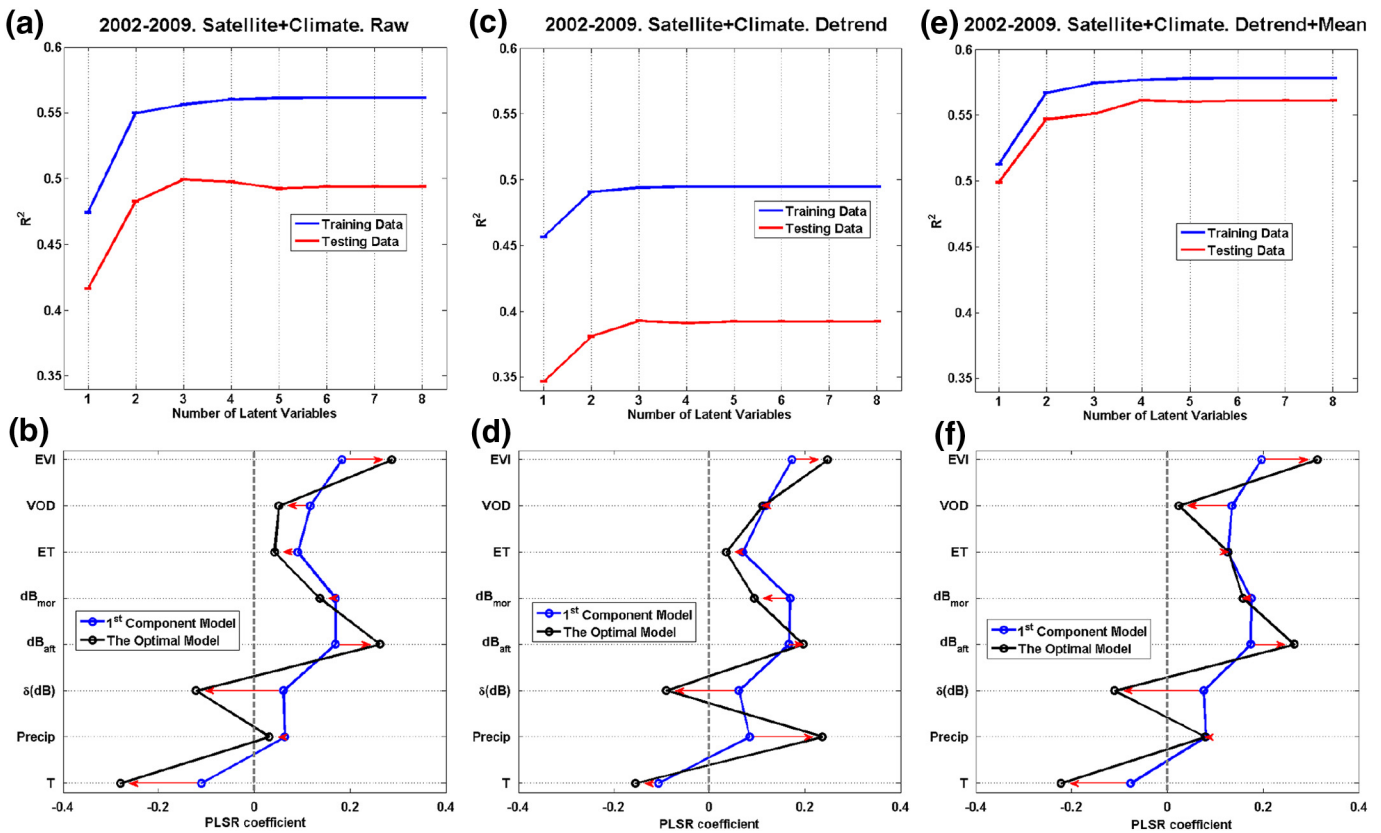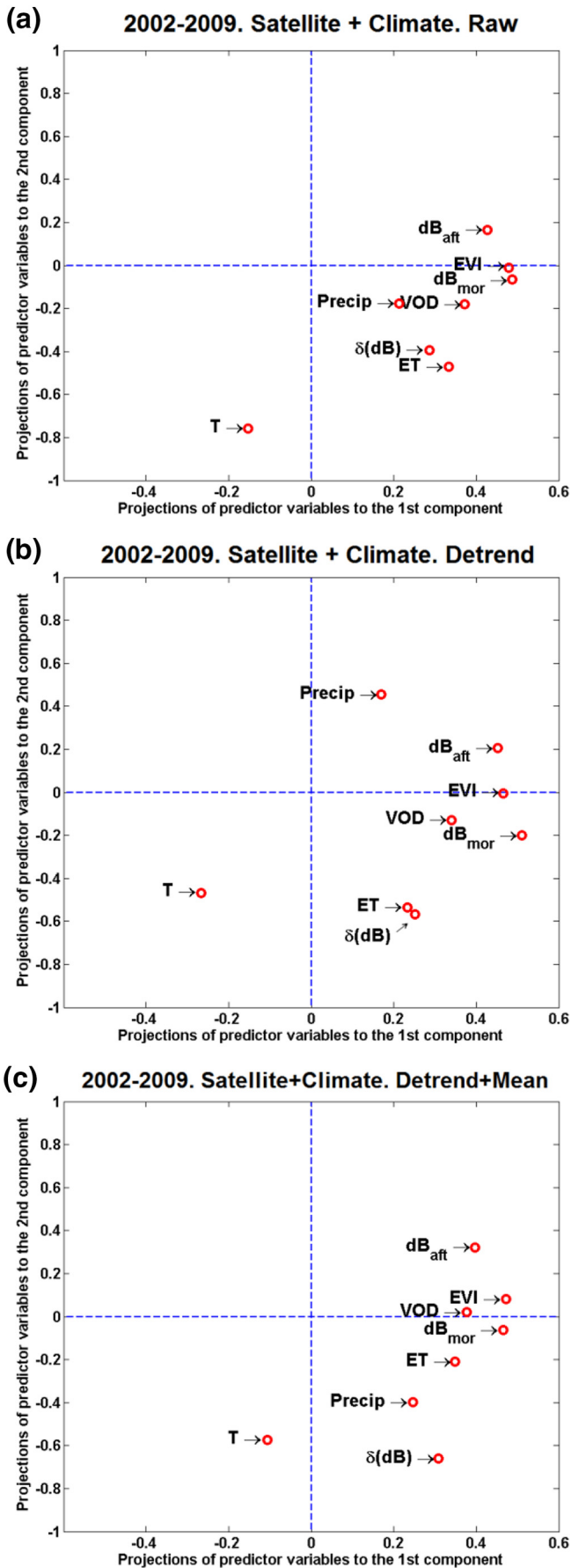


Fig. 10. The PLSR results for the three experiments using the data from: "2002–2009. Satellite + Climate". (6a, b) "Raw"; (6c, d) "Detrend"; and (6e, f) "Detrend + Mean". The upper panels (6a, c, e) show the $R^2$ of the PLSR models with different numbers of components, and errors bars show the 95% confidence interval. The bottom panels (6b, d, f) show the regression coefficients of each predictor variable for the model with only the 1st component (in blue) and for the optimal model (in black), respectively, with the red arrows showing the direction and magnitude of their difference for each predictor variable. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Fig. 11.** PLSR loadings for different predictor variables for the 1st component (x-axis) and the 2nd component (y-axis) of the three experiments using the data from: "2002–2009. Satellite + Climate". (a) "Raw"; (b) "Detrend"; and (c) "Detrend + Mean".

relatively more importance in the "Detrend" than in the "Raw" experiment, which confirms that precipitation is among the major factors causing inter-annual variability in crop yield (Urban et al., 2015).

Fig. 11 shows the PLSR loadings for the "Raw", "Detrend" and "Detrend + Mean" experiments using the "2002–2009" data period. We argue that the interpretation of the 1st and 2nd components in Fig. 9 still holds in Fig. 11, i.e. the 1st and 2nd components represent crop aboveground biomass and environmental stress factors, respectively. The PLSR loading patterns are generally consistent between Fig. 11a ("2002–2009.Satellite + Climate.Raw") and Fig. 9 ("2007–2009. Satellite + Climate.Raw"), though one major difference is that Precip changes from the upper-right quadrant in Fig. 9 to the lower-right quadrant in Fig. 11a. This means that though precipitation is still positively correlated with the 1st component (i.e. "aboveground biomass"), precipitation imposes different impacts on the 2nd component for the two data periods. Precipitation is positively correlated with the 2nd component for the "2007–2009" period, but negatively correlated for the longer "2002–2009" period. July 2003 represented an extreme rainfall month for the study area, with up to 6 standard deviations above the multi-year average in some counties for July precipitation. These extremely high precipitation events have been documented (http://mrcc.isws.illinois.edu/cliwatch/0307/030724.htm), and are known to lead to severe yield loss (Pantaleoni et al., 2007). Due to the nature of PLSR as a linear model, these extreme events heavily skew the results and lead to a negative correlation between precipitation and the 2nd component for the "2002–2009" record. This is further confirmed by excluding 2003 from the "2002–2009" data period, which produces positive correlation between precipitation and the 2nd component, similar to the results in Fig. 9. However, the final Precip regression coefficients of the optimal models for the three experiments ("Raw", "Detrend" and "Detrend + Mean") are all positive, indicating that the Precipitation contribution to the 1st component is more dominant than its contribution to the other components.

*3.3.5. Added value of other satellite data beyond EVI for crop yield estimation*

Fig. 12 shows the $R^2$ performance of the models derived using different combinations of predictor variables and the "2007–2009.Raw" data inputs. We find that only using July Climate information (i.e. "Climate Only") explains just ~5% of the total crop yield variability. Adding EVI with Climate (i.e. "EVI + Climate") achieves higher performance ($R^2$ = 0.4**) for the testing data, demonstrating the added value of EVI relative to the "Climate Only" inputs. Only using satellite data inputs (i.e. "Satellite Only") attains higher model performance ($R^2$ = 0.5**) for the testing data. Combining all of the satellite and climate data produces the best performance ($R^2 > = 0.6$**) among all of the models for both training and testing data. These results indicate that other satellite data beyond EVI provide added value in explaining crop yield variability, with significantly higher $R^2$ performance using "Satellite + Climate" than "EVI + Climate" inputs. Furthermore, comparing the "Satellite Only" and "Satellite + Climate" results indicates that the use of only satellite data could not explain all of the observed crop yield variability, while the additional use of climate information significantly improves model performance. Using the "2002–2009" data period for both "Raw" and "Detrend" cases generates similar results as Fig. 12 and are thus not shown here. It is worth noting that the current approach (i.e. PLSR) is primarily used to study the shared and unique contributions of the different data (including satellite and climate data), while the PLSR method has limitations for achieving higher model performance in yield estimation, as discussed in Section 4.
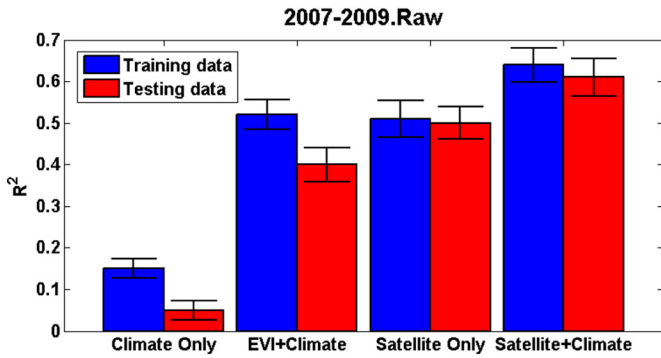
**Fig. 12.** Comparison of model performance derived using different data combinations (see FACTOR 3 in Section 2.4) using the July data from: "2007–2009. Raw". The error bars denote one standard deviation of the derived $R^2$ for the 100 PLSR simulations derived using different splitting of training and testing data.

## 4. Discussion

### 4.1. Benefits of a multi-sensor approach

Visible-NIR based VI metrics have been the dominant approach for operational crop monitoring since the start of the global satellite era (Kogan, 1994; Tucker, 1979), though more than a decade of other satellite observations using other spectral wavelengths have also been acquired, including thermal and microwave observations. One reason for the dominance of visible-NIR VIs is that these indices have a relatively long-term record, large signal-to-noise ratio and high spatial resolution. In contrast, other satellite data have been underutilized and understudied for monitoring crop growth and estimating yield. This study represents one of the first efforts to integrate available satellite data spanning visible, NIR, thermal and microwave spectral ranges for studying large-scale crop yield. Understanding and demonstrating the shared and unique information gained from the different satellite data also contributes to better planning of new satellite missions or the continuation of existing missions.

We find that rich information about crop growth can be revealed by other satellite data from spectral ranges extending beyond visible and NIR spectra. Most importantly, our results indicate that the different satellite data share similar information that the visible-NIR based EVI carries at large scales. In our case, SIF-GPP, Ku-band dB, X-band VOD and thermal-based ET all share information related to aboveground canopy biomass (Fig. 6); these data also share similar seasonality congruent with the crop growth cycle (Fig. 2). Both SIF-GPP and Ku-band dB (including both morning and afternoon overpass retrievals) are found to carry similar information as the EVI (Fig. 6). When excluding this shared information component from all of the satellite metrics, the EVI and SIF-GPP metrics are found to contain little additional information, while the Ku-band dB, VOD and thermal-based ET metrics reveal other unique information discussed below. We thus suggest to continue taking advantage of the visible-NIR VI record, while incorporating other satellite metrics that contain additional unique information.

The most surprising performance in our study comes from the Ku-band dB backscatter data. Being originally developed for monitoring ocean wind fields (Naderi et al., 1991), Ku-band dB from SeaWinds on QuikSCAT has only been used in terrestrial vegetation applications in limited cases [e.g. Frolking et al., 2011; Guan et al., 2012; Saatchi et al., 2013]. Our study suggests that the Ku-band dB from both morning and afternoon overpasses contains rich information on crop growth, similar to the primary information that the EVI carries, which is related to aboveground biomass. However, the Ku-band dB data also contains additional unique information that is associated with crop yield, which may be attributed to higher microwave sensitivity to canopy water content and less signal saturation at higher biomass levels than the visible-NIR based EVI (Ulaby et al., 1982). When combining the

shared and unique information from these metrics, $dB_{aft}$ has an even higher correlation with yield-based NPP than EVI at the county scale. The observed favorable performance of the radar backscatter data for crop monitoring may not be universal, and should be highly dependent on wavelength/frequency, as well as other sensor characteristics including overpass time and incidence angle. The Ku-band (13.4 GHz) frequency is expected to be most sensitive to the top-of-canopy for a fully developed corn crop, while lower microwave frequencies (e.g. C-band or L-band) may contain more information from deeper canopy layers and surface soil moisture (Ulaby et al., 1982), especially prior to canopy closure. We find that C-band dB data from the Advanced Scatterometer (ASCAT) on the European Space Agency's METOP-A satellite was less sensitive to crop growth and more sensitive to soil moisture in our study area (results not shown) than the Ku-band dB data from QuikSCAT, which is consistent with theoretical understanding (Ulaby et al., 1982; van Emmerik et al., 2015).

The Ku-band morning and afternoon overpass dB data (i.e. $dB_{mor}$ and $dB_{aft}$) share much overlapping information, but also have unique differences. The two metrics both show a seasonal cycle similar to EVI and SIF-GPP, while their difference (i.e. $\delta(dB) = dB_{mor} - dB_{aft}$) also shows a similar seasonal cycle (Fig. 2). Previously, Frolking et al. (2011) and Saatchi et al. (2013) hypothesized that the Ku-band morning and afternoon dB difference (i.e. $\delta(dB)$) is a potential indicator of water loss from the landscape: plants absorb water from deep soil overnight, and have a higher water content in the morning; and over the course of a day, plants transpire water and top soil also evaporates water, which leads to a lower landscape water content in the afternoon. Based on this reasoning, $\delta(dB)$ is expected to be correlated with ET, which is marginally confirmed by the relatively weak correlation between $\delta(dB)$ and thermal-based ALEXI-ET ($r = 0.18^{**}$, Fig. 4). ALEXI-ET, as one available ET product among many, may have its own uncertainty and bias. Meanwhile the nuances and rich information of $\delta(dB)$ for croplands has not been fully studied before. We find that $\delta(dB)$ contains little information related to the 1st component representing aboveground crop biomass, but it has a significant negative contribution to the secondary component representing environmental stress (Figs. 7 & 9).

Our work confirms the findings from previous studies (Anderson et al., 2011, 2007; Otkin et al., 2013) that thermal-based ET data contain useful information on crop yield. Over the relatively moisture stress-free conditions encountered in Period 1 (2007–2009), we find that ALEXI-ET conveys two primary levels of information: (1) vegetation aboveground biomass (Fig. 9); and (2) spatio-temporal information on temperature (Fig. 3). During this period, ET is driven primarily by atmospheric demand rather than soil moisture, and atmospheric demand is a function of air temperature. This relationship between ALEXI-ET and temperature leads to an apparent negative relationship between the secondary component of ALEXI-ET (after removing the aboveground biomass information) and crop yield. When temperature is not included in the PLSR for Period 1 (Fig. 6a, b), ALEXI-ET essentially contributes information related to temperature, and when temperature is also included in the PLSR (Fig. 6c, d), the contribution of ALEXI-ET is close to zero, as most ALEXI-ET information that is related to temperature has been represented directly by temperature. However, for Period 2 (2002–2009), which includes significant drought events in some parts of the Corn Belt, the loading for ET in the PLSR remains sizable even when air temperature is included (Fig. 10), indicating that it is conveying useful independent information regarding moisture stress conditions.

### 4.2. Implications and limitations of the methodology

The utility of different spectral wavelength remote sensing data appears to depend on whether one is attempting to track primarily spatial or temporal crop yield variability. We use three levels of data in our PLSR analysis: "raw" data contains information of spatial, inter-annual variations and time trend; "detrend" data removes the spatial pattern and time trend, and only contains inter-annual variations; "detrend

+ mean" data contains spatial and inter-annual variations, but removes the time trend. Little difference is found between the "raw" and "detrend + mean" results (based on the 2002–2009 study period), indicating that temporal trends over this period may not be a significant contributor of model variability. However, the "raw" and "detrend" data show some difference in the PLSR results. The better model performance using "raw" data than "detrend" data confirms findings from previous work (Guan et al., 2016; Lobell et al., 2015) that found predicting inter-annual variability is usually more challenging than predicting the spatial pattern of yield. It is encouraging to see that EVI, T, $\delta$(dB) and $dB_{aft}$ are useful predictors for capturing both spatial and temporal patterns.

The PLSR fundamentally is a linear model which does not account for potential non-linear relationships between predictor and response variables. This has been shown in the precipitation example (Fig. 10 and Section 3.3.4). We only use July data to predict the final crop yield in the current study, while using data from other periods may further improve model performance, which will be explored in future work. However, the current PLSR results provide an improved understanding of the relative value and impact of the different satellite data, which can serve as a guide for choosing the optimal combination of data for estimating crop yields. More sophisticated machine learning algorithms (e.g. neural network) can then be used to achieve the best model performance (You et al., 2017).

Our current study focuses on county-level crop yield estimation for two practical reasons: (1) the USDA NASS benchmark yield data is only available at county level; and (2) most of the satellite datasets we used are only available at relatively coarse spatial resolution (e.g. GOME-2 SIF has a 30–40 km, while QuikSCAT and AMSR-E data are gridded to 0.25° resolution), except for EVI and ALEXI-ET. The U.S. Corn Belt is an ideal study region due to a high spatial fraction of croplands dominated by only two crop types (corn and soybean). Our current study area covers the counties that produced 73% of the total US corn production, and thus nationally representative. However, we caution that applying the current methods to regions beyond this study area (e.g. going to the "Minor Corn Area" in Fig. 2a) faces greater challenges, as the assumption that the coarse-resolution satellite pixels (~>10 km) include mainly homogeneous croplands no longer holds. The utility of relatively coarse resolution satellite observations is expected to be degraded over more spatially heterogeneous croplands, such as small-holder farmlands in Africa and Asia, due to the inability of these sensors to resolve field-level biomass and yield characteristics. However, other satellite sensors may provide more effective information in these areas due to their finer spatial footprint observations, including satellite sensors in visible and NIR (e.g. Landsat, Sentinel-2), thermal (e.g. ECOSTRESS, GOES-R), and high frequency (C-band) synthetic aperture radar (e.g. Sentinel-1). Effectively integrating multi-sensor information for crop monitoring and yield estimation will be useful across spatial scales.

### 4.3. Role of climate information in crop yield modeling

We find that including climate information (i.e. precipitation and temperature) along with satellite data can further improve model predictive power for crop yield. Our results indicate that most of the satellite data selected for this study are highly sensitive to changes in aboveground biomass, but crop yield ultimately depends on grain weight. Current satellite technology is mostly unable to detect changes in grain biomass, which is largely obscured by the canopy during the development stage. Crop yield represents only a portion of the aboveground biomass, and this ratio (crop yield:aboveground biomass) is denoted as the "Harvest Index" (HI). The HI represents the translocation of carbohydrate, from either new photosynthesis or the existing crop biomass, to crop grain. The HI is determined by two main phases: the flowering period which determines grain number, and the grain-filling period which determines individual grain size. The rich agronomy

literature has shown that processes in these phases are highly sensitive to heat and drought stresses (Hay, 1995; Lobell et al., 2014). This also explains why July precipitation and temperature have significant respective positive and negative effects on crop yield in our results. The above reasoning means that although current satellite sensors are sensitive to aboveground biomass, additional HI information is needed to estimate crop yield. Our approach empirically demonstrates the necessity of including climate data for estimating HI and higher order crop yield information. Similar approaches have been implemented to model the HI either implicitly or explicitly in other studies (Leroux et al., 2015; Lobell et al., 2015).

### 4.4. Broader implications for current and future satellite missions and applications

Finally, the current study provides implications for effective utilization and development of existing and future satellite missions for cropland applications. The excellent performance of Ku-band backscatter data identified here motivates further use of satellite scatterometer data, such as RapidSCAT on the International Space Station (Madsen and Long, 2016; Rodriguez, 2013), which unfortunately ceased operating in August 2016. Further use of SIF data for crop monitoring may be catalyzed by the NASA OCO-2 satellite that was launched in July 2015 (Frankenberg et al., 2014) and the incoming FLEX mission by the European Space Agency planned for launch by 2022 (Rascher et al., 2015). The NASA AMSR-E sensor ceased functioning in September 2011, while the JAXA AMSR2 follow-on sensor provides a continuing VOD global data record extending from May 2012 to present (Du et al., 2014). The NASA SMAP radiometer is capable of generating an L-band (1.41 GHz) VOD retrieval for both morning and afternoon satellite overpasses (Konings et al., 2016), which may also be useful for crop monitoring. The combined use of VOD retrievals at different spectral frequencies, including L-band SMAP retrievals vs C– and X-band (6.93 GHz, 10.65 GHz) retrievals from AMSR-2 (Du et al., 2014) may provide enhanced delineation of vegetation phenology, biomass water content variations and water use within different canopy layers and the underlying soil. The NASA ECOSTRESS mission (Hook, 2015) has a planned launch in late 2017 and will capture the approximate diurnal temperature cycles over weekly periods covering most U.S. croplands; these observations are expected to provide unprecedented information on ET for monitoring crop growth. Other spectral indices, such as the Photochemical Reflectance Index (PRI) (Gamon et al., 1997; He et al., 2016) and Land Surface Water Index (LSWI) (Xiao et al., 2005), have been demonstrated to contain distinctive vegetation information and should be examined in context with other potentially synergistic satellite information in future studies.

### 5. Conclusion

Here we present one of the first attempts to synergize multiple satellite data spanning a diverse spectral range (from visible, near-infrared, thermal to microwave) into a coherent framework to estimate crop yield over the U.S. Corn Belt. Our PLSR methodology successfully distinguishes both commonly shared and unique information from the different satellite sensor data (see Fig. 9 for a summary of different data). Most of the satellite data examined (e.g. SIF-GPP, dB, EVI, VOD, ALEXI-ET) share common information related to aboveground biomass; for this shared information, SIF-GPP and $dB_{aft}$ data are found to contain similar information as the EVI. Once the above shared component from all of the satellite data is removed, we find that the EVI and SIF-GPP metrics provide little additional explanatory information at the county level; instead, the Ku-band dB and $\delta$(dB) data, thermal-based ALEXI-ET, and AMSR-E X-band VOD metrics all provide unique information that improve the overall model predictive skill for crop yield at the county level. Our findings support efforts for utilizing a wide spectral range of satellite data for cropland vegetation monitoring, while climate data

also has value for crop yield estimation, in part to capture variations in HI that are not observable from current satellite data. It is worth noting that our results have not yet optimized the predictive skills of the models for estimating crop yield. The modeling framework developed here provides a means to identify useful predictor variables for capturing the spatiotemporal variability of crop yield, which provides a path to fully integrating these data using more powerful machine learning approaches (e.g. deep learning techniques) for optimizing crop yield prediction.

## Acknowledgements

## Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.rse.2017.06.043.

## References

Anderson, M.C., Norman, J.M., Diak, G.R., Kustas, W.P., Mecikalski, J.R., 1997. A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. Remote Sens. Environ. 60 (2):195–216. http://dx.doi.org/10.1016/s0034-4257(96)00215-5.

Anderson, M.C., Norman, J.M., Mecikalski, J.R., Otkin, J.A., Kustas, W.P., 2007. A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 2. Surface moisture climatology. J. Geophys. Res.-Atmos. 112 (11):1–13. http://dx.doi.org/10.1029/2006JD007507.

Anderson, M.C., Hain, C., Wardlow, B., Pimstein, A., Mecikalski, J.R., Kustas, W.P., 2011. Evaluation of drought indices based on thermal remote sensing of evapotranspiration over the continental United States. J. Clim. 24 (8):2025–2044. http://dx.doi.org/10.1175/2010JCLI3812.1.

Anderson, M.C., Hain, C., Otkin, J., Zhan, X., Mo, K., Svoboda, M., ... Pimstein, A., 2013a. An intercomparison of drought indicators based on thermal remote sensing and NLDAS-2 simulations with U.S. drought monitor classifications. J. Hydrometeorol. 14 (4): 1035–1056. http://dx.doi.org/10.1175/JHM-D-12-0140.1.

Anderson, M., Kustas, W., Hain, C.R., 2013b. In Remote Sensing of Energy Fluxes and Soil Moisture Content. In: Petropoulos, G.P (Ed.). Taylor & Francis:pp. 207–232. [online] http://books.google.com/books?hl=en&lr=&id=Wn9FAQAAQBAJ&oi=fnd&pg=PA207&dq=Mapping+Surface+Fluxes+and+Moisture+Conditions+from+Field+to+Global+Scales+Using+ALEXI/DisALEXI&ots=hWFoyINnqt&sig=iDNUQhCyKoqmCORUsz0SmdG7JpI (Accessed November 28, 2014).

Basso, B., Cammarano, D., Carfagna, E., 2013. Review of crop yield forecasting methods and early warning systems. The First Meeting of the Scientific Advisory Committee of the Global Strategy to Improve Agricultural and Rural Statistics:pp. 1–56 http://dx.doi.org/10.1017/CBO9781107415324.004.

Cammalleri, C., Anderson, M.C., Gao, F., Hain, C.R., Kustas, W.P., 2014. Mapping daily evapotranspiration at field scales over rainfed and irrigated agricultural areas using remote sensing data fusion. Agric. For. Meteorol. 186:1–11. http://dx.doi.org/10.1016/j.agrformet.2013.11.001.

Carletto, C., Jolliffe, D., Banerjee, R., 2015. From tragedy to renaissance: improving agricultural data for better policies. J. Dev. Stud. 51 (2):133–148. http://dx.doi.org/10.1080/00220388.2014.968140.

Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Pasteris, P.P., 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. Int. J. Climatol. http://dx.doi.org/10.1002/joc.

Du, J., Kimball, J., Shi, J., Jones, L., Wu, S., Sun, R., Yang, H., 2014. Inter-calibration of satellite passive microwave land observations from AMSR-E and AMSR2 using overlapping FY3B-MWRI sensor measurements. Remote Sens. 6 (9):8594–8616. http://dx.doi.org/10.3390/rs6098594.

Du, J., Kimball, J.S., Jones, L.A., 2015. Passive microwave remote sensing of soil moisture based on dynamic vegetation scattering properties for AMSR-E. IEEE Trans. Geosci. Remote Sens. 1:597–608. http://dx.doi.org/10.1109/TGRS.2015.2462758 (in press).

van Emmerik, T., Steele-Dunne, S.C., Judge, J., van de Giesen, N., 2015. Impact of diurnal variation in vegetation water content on radar backscatter from maize during water stress. IEEE Trans. Geosci. Remote Sens.:1–15 http://dx.doi.org/10.1109/TGRS.2014.2386142.

Entekhabi, D., Yueh, S., O'Neill, P.E., Kellogg, K.H., 2014. SMAP Handbook. JPL 400–1567. https://smap.jpl.nasa.gov/files/smap2/SMAP_Handbook_FINAL_1_JULY_2014_Web.pdf (182).

Frankenberg, C., Butz, A., Toon, G.C., 2011. Disentangling chlorophyll fluorescence from atmospheric scattering effects in O2 A-band spectra of reflected sun-light. Geophys. Res. Lett. 38 (3):1–5. http://dx.doi.org/10.1029/2010GL045896.

Frankenberg, C., O'Dell, C., Berry, J., Guanter, L., Joiner, J., Köhler, P., ... Taylor, T.E., 2014. Prospects for chlorophyll fluorescence remote sensing from the orbiting carbon observatory-2. Remote Sens. Environ. 147:1–12. http://dx.doi.org/10.1016/j.rse.2014.02.007.

Frolking, S., Fahnestock, M., Milliman, T., McDonald, K., Kimball, J., 2005. Interannual variability in North American grassland biomass/productivity detected by SeaWinds scatterometer backscatter. Geophys. Res. Lett. 32 (21), L21409. http://dx.doi.org/10.1029/2005GL024230.

Frolking, S., Milliman, T., Palace, M., Wisser, D., Lammers, R., Fahnestock, M., 2011. Tropical forest backscatter anomaly evident in SeaWinds scatterometer morning overpass data during 2005 drought in Amazonia. Remote Sens. Environ. 115 (3):897–907. http://dx.doi.org/10.1016/j.rse.2010.11.017.

Frolking, S., Hagen, S., Milliman, T., Palace, M., Shimbo, J.Z., Fahnestock, M., 2012. Detection of large-scale forest canopy change in pan-tropical humid forests 2000–2009 with the seawinds Ku-band scatterometer. IEEE Trans. Geosci. Remote Sens. 50 (7 PART1):2603–2617. http://dx.doi.org/10.1109/TGRS.2011.2182516.

Gamon, J.A., Serrano, L., Surfus, J.S., 1997. The photochemical reflectance index: An optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. Oecologia 112, 492–501.

Garcia, D., 2010. Robust smoothing of gridded data in one and higher dimensions with missing values. Comput. Stat. Data Anal. 54 (4):1167–1178. http://dx.doi.org/10.1016/j.csda.2009.09.020.

Geladi, P., Kowalski, B.R., 1986. Partial least-squares regression: a tutorial. Anal. Chim. Acta 185 (C):1–17. http://dx.doi.org/10.1016/0003-2670(86)80028-9.

Guan, K., Wood, E.F., Caylor, K.K., 2012. Multi-sensor derivation of regional vegetation fractional cover in Africa. Remote Sens. Environ. 124:653–665. http://dx.doi.org/10.1016/j.rse.2012.06.005.

Guan, K., Wolf, A., Medvigy, D., Caylor, K., Pan, M., Wood, E.F., 2013. Seasonal coupling of canopy structure and function in African tropical forests and its environmental controls. Ecosphere 4 (March):1–21 Retrieved from. http://www.esajournals.org/doi/abs/10.1890/ES12-00232.1.

Guan, K., Wood, E.F., Medvigy, D., Kimball, J., Pan, M., Caylor, K.K., ... Jones, M.O., 2014. Terrestrial hydrological controls on land surface phenology of African savannas and woodlands. J. Geophys. Res. Biogeosci. 119 (8):1652–1669. http://dx.doi.org/10.1002/2013JG002572.

Guan, K., Berry, J.A., Zhang, Y., Joiner, J., Guanter, L., Badgley, G., Lobell, D.B., 2016. Improving the monitoring of crop productivity using spaceborne solar-induced fluorescence. Glob. Chang. Biol. 22 (2):716–726. http://dx.doi.org/10.1111/gcb.13136.

Guanter, L., Alonso, L., Gómez-Chova, L., Amorós-López, J., Vila, J., Moreno, J., 2007. Estimation of solar-induced vegetation fluorescence from space measurements. Geophys. Res. Lett. 34 (8), L08401. http://dx.doi.org/10.1029/2007GL029289.

Guanter, L., Frankenberg, C., Dudhia, A., Lewis, P.E., Gómez-Dans, J., Kuze, A., ... Grainger, R.G., 2012. Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements. Remote Sens. Environ. 121:236–251. http://dx.doi.org/10.1016/j.rse.2012.02.006.

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. a, ... Griffis, T. J., 2014. Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. Proc. Natl. Acad. Sci. U. S. A. 111 (14):E1327–E1333. http://dx.doi.org/10.1073/pnas.1320008111.

Hay, R., 1995. Harvest index: a review of its use in plant breeding and crop physiology. Ann. Appl. Biol. 126:197–216 Retrieved from. http://onlinelibrary.wiley.com/doi/10.1111/j.1744-7348.1995.tb05015.x/abstract.

He, M., Kimball, J.S., Running, S., Ballantyne, A., Guan, K., Huemmrich, F., 2016. Satellite detection of soil moisture related water stress impacts on ecosystem productivity using the MODIS-based photochemical reflectance index. Remote Sens. Environ. 186: 173–183. http://dx.doi.org/10.1016/j.rse.2016.08.019.

Hoffman, L.A., Etienne, X.L., Irwin, S.H., Colino, E.V., Toasa, J.I., 2015. Forecast performance of WASDE price projections for U.S. corn. Agric. Econ. (U.K.) 46:157–171. http://dx.doi.org/10.1111/agec.12204.

Hook, S.J., 2015. HyspIRI thermal & ECOSTRESS. HyspIRI Symposium Retrieved from. https://hyspiri.jpl.nasa.gov/downloads/2015_Symposium/day1/4_HyspIRIThermalECOSTRESS_Hook.pdf.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ. 83, 195–213.

Jackson, T.J., 1993. Microwave remote sensing. Hydrol. Process. 7 (May 1992), 139–152.

Joiner, J., Yoshida, Y., Vasilkov, A.P., Corp, L.A., Middleton, E.M., 2011. First observations of global and seasonal terrestrial chlorophyll fluorescence from space. Biogeosciences 8 (3):637–651. http://dx.doi.org/10.5194/bg-8-637-2011.

Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A.P., Middleton, E.M., ... Frankenberg, C., 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. http://dx.doi.org/10.5194/amtd-6-3883-2013.

Joiner, J., Yoshida, Y., Vasilkov, A.P., Schaefer, K., Jung, M., Guanter, L., ... Belelli Marchesini, L., 2014. The seasonal cycle of satellite chlorophyll fluorescence observations and its relationship to vegetation phenology and ecosystem atmosphere carbon exchange. Remote Sens. Environ. 152:375–391. http://dx.doi.org/10.1016/j.rse.2014.06.022.

Jones, L.A., Kimball, J.S., 2010. A global daily record of land parameter retrievals from AMSR - E. NASA ATBD Report.

Jones, L.A., Kimball, J.S., 2012. Daily Global Land Surface Parameters Derived from AMSR-E. Version 1. http://dx.doi.org/10.5067/JIKQZ6WO5C5M.

Jones, M.O., Jones, L.a., Kimball, J.S., McDonald, K.C., 2011. Satellite passive microwave remote sensing for monitoring global land surface phenology. Remote Sens. Environ. 115 (4):1102–1114. http://dx.doi.org/10.1016/j.rse.2010.12.015.

Jones, M.O., Kimball, J.S., Jones, L.a., McDonald, K.C., 2012. Satellite passive microwave detection of North America start of season. Remote Sens. Environ. 123:324–333. http://dx.doi.org/10.1016/j.rse.2012.03.025.

Jones, M.O., Kimball, J.S., Jones, L.a., 2013. Satellite microwave detection of boreal forest recovery from the extreme 2004 wildfires in Alaska and Canada. Glob. Chang. Biol. 19 (10):3111–3122. http://dx.doi.org/10.1111/gcb.12288.

Kogan, F.N., 1994. Global drought watch from space. Bull. Am. Meteorol. Soc. 621–636.

Konings, A.G., Gentine, P., 2016. Global variations in ecosystem-scale isohydricity. Glob. Chang. Biol. http://dx.doi.org/10.1111/gcb.13389 In press.

Konings, A.G., Piles, M., Rötzer, K., McColl, K.A., Chan, S.K., Entekhabi, D., 2016. Vegetation optical depth and scattering albedo retrieval using time series of dual-polarized L-band radiometer observations. Remote Sens. Environ. 172:178–189. http://dx.doi.org/10.1016/j.rse.2015.11.009.

Leroux, L., Baron, C., Zoungrana, B., Traoré, S.B., Lo Seen, D., Bégué, A., 2015. Crop monitoring using vegetation and thermal indices for yield estimates: case study of a rainfed cereal in semi-arid West Africa. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.: 1–16 http://dx.doi.org/10.1109/JSTARS.2015.2501343.

Liu, Y.Y., de Jeu, R.A., McCabe, M.F., Evans, J.P., van Dijk, A.I.J.M., 2011. Global long-term passive microwave satellite-based retrievals of vegetation optical depth. Geophys. Res. Lett. 38 (18). http://dx.doi.org/10.1029/2011GL048684 (n/a–n/a.).

Liu, Y.Y., van Dijk, A.I.J.M., McCabe, M.F., Evans, J.P., de Jeu, R. a. M., 2013. Global vegetation biomass change (1988–2008) and attribution to environmental and human drivers. Glob. Ecol. Biogeogr. 22 (6):692–705. http://dx.doi.org/10.1111/geb.12024.

Liu, Y.Y., van Dijk, A.I.J.M., de Jeu, R.A.M., Canadell, J.G., McCabe, M.F., Evans, J.P., Wang, G., 2015. Recent reversal in loss of global terrestrial biomass. Nat. Clim. Chang. (March): 1–5 http://dx.doi.org/10.1038/nclimate2581.

Lobell, D.B., Hicke, J.A., Asner, G.P., Field, C.B., Tucker, C.J., Los, S.O., 2002. Satellite estimates of productivity and light use efficiency in United States agriculture, 1982–98. Glob. Chang. Biol. 8 (8), 722–735.

Lobell, D.B., Cassman, K.G., Field, C.B., 2009. Crop yield gaps: their importance, magnitudes, and causes. Annu. Rev. Environ. Resour. 34 (1):179–204. http://dx.doi.org/10.1146/annurev.environ.041008.093740.

Lobell, D.B., Hammer, G.L., McLean, G., Messina, C., Roberts, M.J., Schlenker, W., 2013. The critical role of extreme heat for maize production in the United States. Nat. Clim. Chang. 3 (5):497–501. http://dx.doi.org/10.1038/nclimate1832.

Lobell, D.B., Roberts, M.J., Schlenker, W., Braun, N., Little, B.B., Rejesus, R.M., Hammer, G.L., 2014. Greater sensitivity to drought accompanies maize yield increase in the U.S. Midwest. Science New York, N.Y. 344 (6183). http://dx.doi.org/10.1126/science.1251423 (516–9).

Lobell, D.B., Thau, D., Seifert, C., Engle, E., Little, B., 2015. A scalable satellite-based crop yield mapper. Remote Sens. Environ. http://dx.doi.org/10.1016/j.rse.2015.04.021.

Long, D.G., Early, D.S., 2001. Image reconstruction and enhanced resolution imaging from irregular samples image reconstruction and enhanced resolution imaging from irregular samples. IEEE Trans. Geosci. Remote Sens. 39 (2), 291–302.

Madsen, N.M., Long, D.G., 2016. Calibration and validation of the RapidScat scatterometer using tropical rainforests. IEEE Trans. Geosci. Remote Sens. 54 (5):2846–2854. http://dx.doi.org/10.1109/TGRS.2015.2506463.

Maes, W.H., Steppe, K., 2012. Estimating evapotranspiration and drought stress with ground-based thermal remote sensing in agriculture: a review. J. Exp. Bot. 63 (13): 4671–4712. http://dx.doi.org/10.1093/jxb/ers165.

McNaughton, K.G., Spriggs, T.W., 1986. A mixed-layer model for regional evaporation. Bound.-Layer Meteorol. 34 (3):243–262. http://dx.doi.org/10.1007/BF00122381.

Mulla, D.J., 2013. Twenty five years of remote sensing in precision agriculture: key advances and remaining knowledge gaps. Biosyst. Eng. 114 (4):358–371. http://dx.doi.org/10.1016/j.biosystemseng.2012.08.009.

Myneni, R.B., Hoffman, S., Knyazikhin, Y., Privette, J.L., Glassy, J., Tian, Y., ... Running, S.W., 2002. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. Remote Sens. Environ. 83, 214–231.

Naderi, F.M., Freilich, M.H., Long, D.G., 1991. Spaceborne Radar Measurement of Wind Velocity Over the Ocean—An Overview of the NSCAT Scatterometer System. Proc. IEEE 79, 850–866.

Njoku, E.G., Jackson, T.J., Lakshmi, V., Chan, T.K., Nghiem, S.V., 2003. Soil moisture retrieval from AMSR-E. IEEE Trans. Geosci. Remote Sens. 41 (2 PART 1):215–228. http://dx.doi.org/10.1109/TGRS.2002.808243.

Otkin, J.a., Anderson, M.C., Hain, C., Mladenova, I.E., Basara, J.B., Svoboda, M., 2013. Examining rapid onset drought development using the thermal infrared–based evaporative stress index. J. Hydrometeorol. 14 (4):1057–1074. http://dx.doi.org/10.1175/JHM-D-12-0144.1.

Pantaleoni, E., Engel, B.A., Johannsen, C.J., 2007. Identifying agricultural flood damage using Landsat imagery. Precis. Agric. 8 (1–2):27–36. http://dx.doi.org/10.1007/s11119-006-9026-5.

Porcar-Castell, A., Tyystjärvi, E., Atherton, J., van der Tol, C., Flexas, J., Pfündel, E.E., ... Berry, J.a., 2014. Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges. J. Exp. Bot. 65 (15):4065–4095. http://dx.doi.org/10.1093/jxb/eru191.

Rascher, U., Alonso, L., Burkart, A., Cilia, C., Cogliati, S., Colombo, R., ... Zemek, F., 2015. Sun-induced fluorescence - a new probe of photosynthesis: first maps from the imaging spectrometer HyPlant. Glob. Chang. Biol. (0) http://dx.doi.org/10.1111/gcb.13017.

Rodriguez, E., 2013. The NASA ISS-RapidScat mission. American Geophysical Union, Fall Meeting 2013.

Ross, K.W., Brown, M.E., Verdin, J.P., Underwood, L.W., 2009. Review of FEWS NET biophysical monitoring requirements. Environ. Res. Lett. 4 (2), 024009. http://dx.doi.org/10.1088/1748-9326/4/2/024009.

Saatchi, S., Asefi-Najafabady, S., Malhi, Y., Aragão, L.E.O.C., Anderson, L.O., Myneni, R.B., Nemani, R., 2013. Persistent effects of a severe drought on Amazonian forest canopy. Proc. Natl. Acad. Sci. U. S. A. 110 (2):565–570. http://dx.doi.org/10.1073/pnas.1204651110.

Schlenker, W., Roberts, M.J., 2009. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. Proc. Natl. Acad. Sci. U. S. A. 106 (37): 15594–15598. http://dx.doi.org/10.1073/pnas.0906865106.

Sellers, P.J., Berry, J.A., Collatz, G.J., Field, C.B., Hall, F.G., 1992. Canopy reflectance, photosynthesis, and transpiration. III. A reanalysis using improved leaf models and a new canopy integration scheme. Remote Sens. Environ. 42, 187–216.

Sherrick, B.J., Lanoue, C.A., Woodard, J., Schnitkey, G.D., Paulson, N.D., 2014. Crop yield distributions: fit, efficiency, and performance. Agric. Finance Rev. 74 (3), 348–363.

Shi, J., Jackson, T., Tao, J., Du, J., Bindlish, R., Lu, L., Chen, K., 2008. Microwave vegetation indices for short vegetation covers from satellite passive microwave sensor AMSR-E. Remote Sens. Environ. 112 (12):4285–4300. http://dx.doi.org/10.1016/j.rse.2008.07.015.

Sibley, A.M., Grassini, P., Thomas, N.E., Cassman, K.G., Lobell, D.B., 2014. Testing remote sensing approaches for assessing yield variability among maize fields. Agron. J. 106 (1):24. http://dx.doi.org/10.2134/agronj2013.0314.

Thornton, P.K., Bowen, W.T., Ravelo, A.C., Wilkens, P.W., Farmer, G., Brock, J., Brink, J.E., 1997. Estimating millet production for famine early warning: an application of crop simulation modelling using satellite and ground-based data in Burkina Faso. Agric. For. Meteorol. 83 (1–2):95–112. http://dx.doi.org/10.1016/S0168-1923(96)02348-9.

Thorp, K.R., Wang, G., West, A.L., Moran, M.S., Bronson, K.F., White, J.W., Mon, J., 2012. Estimating crop biophysical properties from remote sensing data by inverting linked radiative transfer and ecophysiological models. Remote Sens. Environ. 124:224–233. http://dx.doi.org/10.1016/j.rse.2012.05.013.

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

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8, 127–150.

Tucker, C., Holben, B., 1980. Relationship of spectral data to grain yield variation. Photogramm. Eng. Remote. Sens. 46 (5):657–666 Retrieved from. http://www.cabdirect.org/abstracts/19810714985.html.

Ulaby, F.T., 1987. Dielectric Spectrum of Vegetation Part II: Dual-Dispersion Model. 5 pp. 550–557.

Ulaby, F., Moore, R., Fung, A., 1982. Microwave Remote Sensing: Vols. I, II, and III. Addison-Wesley, Reading, MA.

Urban, D.W., Roberts, M.J., Schlenker, W., Lobell, D.B., 2015. The effects of extremely wet planting conditions on maize and soybean yields. Clim. Chang. 130 (2):247–260. http://dx.doi.org/10.1007/s10584-015-1362-x.

Van Ittersum, M.K., Cassman, K.G., Grassini, P., Wolf, J., Tittonell, P., Hochman, Z., 2013. Yield gap analysis with local to global relevance-a review. Field Crop Res. 143:4–17. http://dx.doi.org/10.1016/j.fcr.2012.09.009.

Vereecken, H., Weihermüller, L., Jonard, F., Montzka, C., 2012. Characterization of crop canopies and water stress related phenomena using microwave remote sensing methods: a review. Vadose Zone J. 11 (2). http://dx.doi.org/10.2136/vzj2011.0138ra.

Whitcraft, A.K., Becker-Reshef, I., Justice, C.O., 2015. A framework for defining spatially explicit earth observation requirements for a global agricultural monitoring initiative (GEOGLAM). Remote Sens. 7 (2):1461–1481. http://dx.doi.org/10.3390/rs70201461.

Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemom. Intell. Lab. Syst. 58 (2):109–130. http://dx.doi.org/10.1016/S0169-7439(01)00155-1.

Wu, J., Chavana-Bryant, C., Prohaska, N., Serbin, S.P., Guan, K., Albert, L.P., ... Saleska, S.R., 2016. Convergence in relationships between leaf traits, spectra and age across diverse canopy environments and two contrasting tropical forests. New Phytol. http://dx.doi.org/10.1111/nph.14051.

Xiao, X., Zhang, Q., Saleska, S., Hutyra, L., De Camargo, P., Wofsy, S., ... Moore, B., 2005. Satellite-based modeling of gross primary production in a seasonally moist tropical evergreen forest. Remote Sens. Environ. 94 (1):105–122. http://dx.doi.org/10.1016/j.rse.2004.08.015.

Xu, Q.-S., Liang, Y.-Z., 2001. Monte Carlo cross validation. Chemom. Intell. Lab. Syst. 56 (1): 1–11. http://dx.doi.org/10.1016/S0169-7439(00)00122-2.

You, J., Li, X., Low, M., Lobell, D., Ermon, S., 2017. Deep gaussian process for crop yield prediction based on remote sensing data. 2017 Association for the Advancement of Articial Intelligence.

Zhang, Y., Guanter, L., Berry, J.A., Joiner, J., van der Tol, C., Huete, A., ... Köhler, P., 2014. Estimation of vegetation photosynthetic capacity from space-based measurements of chlorophyll fluorescence for terrestrial biosphere models. Glob. Chang. Biol. 20: 3727–3742. http://dx.doi.org/10.1111/gcb.12664.