Version of Record: https://www.sciencedirect.com/science/article/pii/S0034425717300019
Manuscript_9b9ac9324e65c9865ef4d20323742d94

Exploration of Scaling Effects on Coarse Resolution Land Surface Phenology

Xiaoyang Zhang[1,2], Jianmin Wang[1], Feng Gao[3], Yan Liu[4], Crystal Schaaf[4], Mark Friedl[5], Yunyue Yu[6], Senthilnath Jayavelu[1], Joshua Gray[5], Lingling Liu[1], Dong Yan[1], and Geoffrey M. Henebry[1,7]

[1] Geospatial Sciences Center of Excellence, South Dakota State University, Brookings, SD 57007, USA

[2] Department of Geography, South Dakota State University, Brookings, SD 57007, USA

[3]USDA, Agricultural Research Service, Hydrology and Remote Sensing Laboratory, 10300 Baltimore Avenue, Beltsville, MD 20705, United States

[4]School for the Environment, University of Massachusetts Boston, Boston, MA, 02125, USA

[5]Department of Earth and Environment, Boston University, Boston, MA, 02215, USA

[6]NOAA NESDIS Center for Satellite Applications and Research, College Park, MD, USA

[7]Department of Natural Resource Management, South Dakota State University, Brookings, SD 57007, USA

(A revised version submitted to ***Remote Sensing of Environment***)

ABSTRACT

Numerous land surface phenology (LSP) datasets have been produced from various coarse resolution satellite data and different detection algorithms from regional to global scales. In contrast to field-observed phenological events that are defined by clearly evident organismal changes with biophysical meaning, current approaches to detecting transitions in LSP only determine the timing of variations in remotely sensed observations of surface greenness. Since activities to bridge LSP and field observations are challenging and limited, our understanding of the biophysical characteristics of LSP transitions is poor. Therefore, we set out to explore the scaling effects on LSP transitions at the nominal start of growing season (SOS) by comparing detections from coarse resolution data with those from finer resolution imagery. Specifically, using a hybrid piecewise-logistic-model-based LSP detection algorithm, we detected SOS in the agricultural core of the United States—central Iowa—at two scales: first, at a finer scale (30m) using reflectance generated by fusing MODIS data with Landsat 8 OLI data (OLI SOS) and, second, at a coarser resolution of 500m using Visible Infrared Imaging Radiometer Suite (VIIRS) observations. The VIIRS SOS data were compared with OLI SOS that had been aggregated using a percentile approach at various degrees of heterogeneity. The results revealed the complexities of SOS detections and the scaling effects that are latent at the coarser resolution. Specifically, OLI SOS variation defined using standard deviation (SD) was as large as 40 days within a highly spatially heterogeneous VIIRS pixel; whereas, SD could be less than 10 days for a more homogeneous set of pixels. Furthermore, the VIIRS SOS detections equaled the OLI SOS (with an absolute difference less than one day) in more than 60% of OLI pixels within a homogeneous VIIRS pixel, but in less than 20% of OLI pixels within a spatially heterogeneous VIIRS pixel. Moreover, the SOS detections in a coarser resolution pixel reflected the timing at which vegetation greenup onset occurred in 30% of area, despite variation in SOS heterogeneities. This result suggests that (1) the SOS detections at coarser resolution is controlled more by the earlier

SOS pixels at the finer resolution rather than by the later SOS pixels, and (2) it should be possible to well simulate the coarser SOS value by selecting the timing at 30th percentile SOS at the finer resolution. Finally, it was demonstrated that in homogeneous areas the VIIRS SOS was comparable with OLI SOS with an overall difference of less than 5 days.

Key Words: Land Surface Phenology; Scaling Effects; Spatial Heterogeneity; VIIRS; OLI

1. INTRODUCTION

Remote sensing has been widely used to characterize seasonal vegetation dynamics at continental and global scales during the last three decades, because it can provide frequent and consistent measurements that are spatially exhaustive. Due to the coarse spatial resolution (>500m) of synoptic sensors, remote sensing monitors seasonal dynamics of the vegetated land surface that often includes multiple types of vegetation mixed with other scene objects, such as soil, water, and human structures. Land surface phenology (LSP) is the term used to distinguish the object of remote sensing from traditional notions of species-specific organismal phenology observed at ground level (de Beurs and Henebry 2004; Henebry and de Beurs 2013). The most commonly used satellite data for LSP characterization have been from the Advanced Very High Resolution Radiometer (AVHRR) instruments at a spatial resolution from 5km-8km (White et al., 2009; Zhang et al., 2007, 2014; de Jong et al., 2011; Julien and Sobrino, 2009; Zhou et al., 2001), because they boast the longest and densest time series available at a global coverage. With the availability of the

Moderate-resolution Imaging Spectroradiometer (MODIS) data since 2000, which substantially improved radiometric and geometric properties, atmospheric correction, and cloud screening of the time series, it has been possible to characterize a more reliable and consistent LSP at spatial resolutions from 250 m to 1000 m (Ganguly et al. 2010; Tan et al. 2011; Zhang et al. 2006). Recently, Landsat data at a spatial resolution of 30 m has also been applied to retrieve LSP (Fisher et al. 2006; Krehbiel et al. 2015; Melaas et al. 2013; Walker et al. 2012); however, Landsat's relatively long period for repeat observations (~16 days) have made it impractical to consistently produce annual time series at a regional scale for most parts of the planet.

A number of approaches have been developed to detect LSP, particularly, the start of growing season (SOS), based on the time series of satellite observations. Most approaches first smooth and gap-fill time series of vegetation indices using one or more of the methods that include asymmetric Gaussians (Jonsson and Eklundh 2002), piecewise logistic function (Zhang et al. 2003), Savitzky-Golay filter (Chen et al. 2004), best index slope extraction algorithm (BISE) (Viovy et al. 1992), moving average (Reed et al. 1994), moving median, iterative interpolation (Julien and Sobrino 2010), Fourier fitting (Moody and Johnson 2001; Wagenseil and Samimi 2006), polynomial curve fitting (Bradley et al. 2007), or the convex quadratic model based on thermal time (de Beurs and Henebry 2004; Henebry and de Beurs 2013). The timings of phenophase transitions during the vegetation growing season are then extracted based on either predefined absolute or relative thresholds of vegetation indices (Jonsson and Eklundh 2002; Lloyd 1990; Reed et al. 1994; White et al. 1997), or features of the fitted curves such as the inflection points (de Beurs and Henebry 2010; Tan et al. 2011; Zhang et al. 2003).

While a great number of LSP data have been produced from various satellite datasets and approaches, the biophysical meaning and scaling effects of these phenological data have rarely been investigated. Relative to the large number of LSP datasets produced, the validation activities have been surprisingly limited and simple. Validation efforts have been typically conducted in one or more of the following ways.

First, the extracted LSP transition or phenometrics have been indirectly compared with model outputs or other variables observed at ground level. For example, the LSP SOS calculated from 8 km 15-day composite AVHRR NDVI data was linked to phenological timings from empirical or bioclimatic models, such as the climate data-driven phenology (Schaber and Badeck 2003; Schwartz and Reed 1999), and associated with ground-based records from cryospheric and hydrological networks (White et al. 2009). These comparisons have generally shown poor correlations, such as no significant relationship between LSP SOS and the modeled phenology (Schwartz and Hanes 2010), or differences between AVHRR SOS and ground observations that could exceed two months (White et al. 2009).

Second, pixel-based LSP has also been compared with phenological observations of vegetation communities within field plots. For example, the MODIS SOS in a 1 km^2 footprint exhibited a root mean square error (RMSE) of 20.5 days and a bias of 17 days compared with *in-situ* observations of 36 trees in a 0.5ha (0.005km^2) plot in France (Soudani et al. 2008). Satellite derived green-up timing had a RMSE of about 15 days as compared with leaf-out dates of four woody species observed from the PlantWatch citizen science project across Canada (Delbart et al. 2015).

Third, LSP SOS dates have also been compared with landscape scale observations. By aggregating individual plants to population, community, and landscape scales within homogeneous regions consisting of deciduous and conifer plants, indices of landscape phenology —a concept distinct from land surface phenology (Liang and Schwartz 2009)— were derived and compared with MODIS SOS dates (Liang et al. 2011). The results indicated the LSP SOS dates matched well with full bud burst in deciduous forests, but not so well in conifer forests, which lagged LSP SOS dates by about 10 days.

Fourth, LSP SOS dates have recently been compared to PhenoCam observations. PhenoCam provides consistent and continuous monitoring of vegetation canopy conditions using tower-mounted webcams that collect images multiple times a day (Hufkens et al. 2012; Richardson et al. 2009; Richardson et al. 2007; Sonnentag et al. 2012). It has provided important information for validating and understanding satellite-derived LSP. However, PhenoCam analyses rely on vegetation indices derived from visible wavelengths, introducing some differences from satellite vegetation indices that are derived from both red and near infrared reflectance. Moreover, a mismatch of camera field of view angle and its large variation with the view angle of satellite pixel-coverage may cause major uncertainties (Elmore et al. 2012; Graham et al. 2010; Hufkens et al. 2012; Keenan et al. 2014).

Validation efforts have shown a discrepancy of more than 10 days between LSP and other phenological observations. This discrepancy arises in part from the arguably erroneous assumptions that (1) field observations are obtained from large homogeneous sites, and (2) LSP measurements should be consistently equivalent to the field observations despite the scaling differences. Homogeneous SOS values within a moderate or coarse satellite footprint are rarely observed because the timing of phenophase transitions vary greatly among different species and even within the same species due to ecotypic variation or local site conditions. Indeed, woody understory plants often leaf out more than three weeks earlier than the forest canopy (Augspurger et al. 2005). Budburst dates for coexisting tree species in temperate forests can vary by three weeks or more (Lechowicz 1984). Similarly, budburst dates among woody species within an area of locally homogeneous forests can even vary by roughly six weeks (Richardson and O’Keefe 2009). Even in relatively homogeneous deciduous forests (with similar composition, age, and structure), leaf out timing in a same species can vary more than two weeks spatially within a 500m area (Fisher et al. 2006).

These findings indicate that simple comparisons of LSP with field observations may only illuminate their differences rather than provide meaningful validation. This situation arises mainly because the scaling effects on the coarse resolution LSP are poorly understood. Field phenological measurements have sharply defined life cycle events, such as the appearance of first bloom, first leaf unfolding, first leaf coloration, etc. In contrast, “events” in LSP are not sharply defined, but rather are transitions within fitted curves of remotely sensed “greenness” that has equivocal biophysical meaning. This study, therefore, aims to explore the question: what kinds of SOS occurrences at the field scale translate into coarser resolution LSP “events”?

Our hypothesis is that SOS at coarser resolution becomes detectable once the vegetation starts to greenup in a certain proportion of finer resolution pixels. A corollary to this hypothesis is that coarser resolution SOS is driven by the portion of earlier SOS pixels at the finer resolution rather than the later SOS pixels. To explore this hypothesis, we made the assumptions that (1) vegetation phenology, environmental conditions, and microclimate within the 30 m scale are relatively homogenous, and (2) the SOS derived at the finer scale could well represent the start of surface vegetation leaf seasonality. Thus, we first detected LSP at finer scale (30 m) using the reflectance data from the fusion of MODIS data with Landsat 8 OLI observations, and then at the coarser resolution (500 m) using Visible Infrared Imaging Radiometer Suite (VIIRS) observations during 2013 and 2014. The scaling effect on SOS at coarser resolution was then investigated by linking to the SOS observations at the finer scale. Our study area is central Iowa in the United States (US), where agricultural lands dominate in the northern part of the State and forests and grasslands occur in the south. The timing of phenological events spans a wide range in central Iowa from low spatiotemporal heterogeneity within crop fields, to moderate spatiotemporal heterogeneity between different crop types, to high spatiotemporal heterogeneity in mixtures of croplands and natural vegetation.

2. Methodology

2.1 Datasets

The data used here include land cover classifications, Landsat-MODIS fused surface reflectance, and VIIRS surface reflectance in central Iowa in the Western Corn Belt.

2.1.1 Land cover data

We used land cover data from the USDA (United States Department of Agriculture) National Agricultural Statistics Service (NASS) Cropland Data Layer (CDL) in 2013 and 2014. The CDL is a crop-specific land cover data layer with a ground resolution of 30 m. The CDL products were generated using satellite imagery from the Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) on Landsat 8 and the Disaster Monitoring Constellation (DMC) DEIMOS-1 and UK2 sensors, which were collected during the crop growing season. Imperviousness and natural vegetation cover data were obtained from the USGS (United State Geological Survey) National Land Cover Database 2011 (Homer et al. 2015).

The overall classification accuracies for major crops (soybeans and corn) in NASS CDL were generally above 96%. The typical commodity crop rotation alternates between corn and soybeans. We simplified the CDL crop classes to corn, soybean, hay (aggregating these classes: alfalfa, other hay/non alfalfa), other crops (aggregating these classes: barley, wheat, other small grains, rye, oats, millet, and spelt), grass (grassland/pasture), forests, shrublands, non-vegetated areas (aggregating these classes: fallow/idle cropland, developed/open space, developed area, and barren), and open water/wetlands (aggregating these classes: open water, woody wetlands, and herbaceous wetlands) (Figure 1). The aggregated land cover data were then reprojected and resampled to match the Landsat scene (path 26 and row 31).

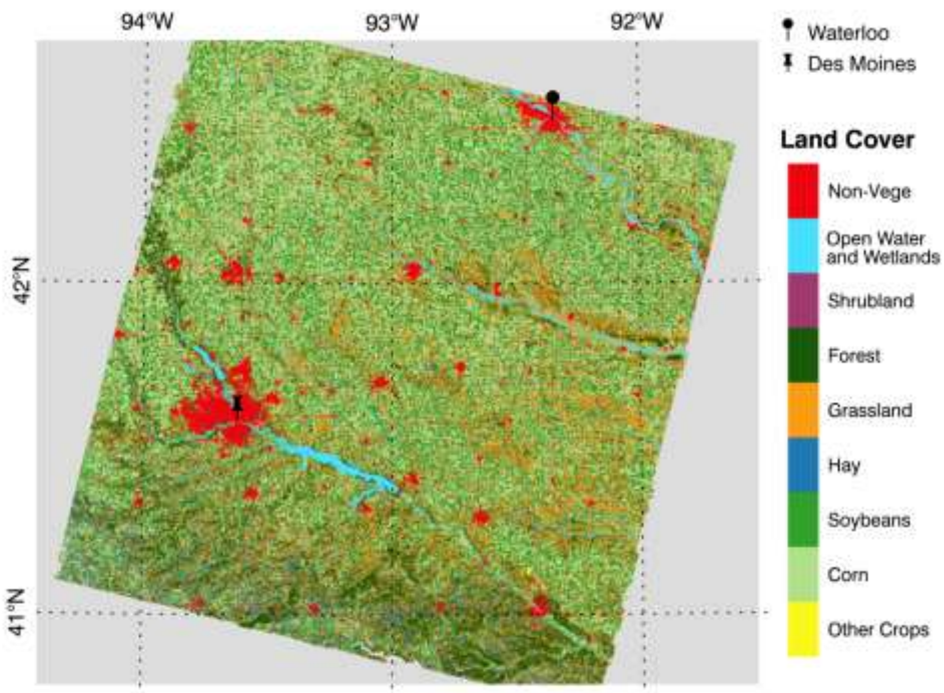

Figure 1. Spatial pattern of land cover types from NASS in 2014. Waterloo and Des Moines are the two largest cities in the area and are indicated by the black pushpins. Land cover types in 2013 were similar to those in 2014, with some spatial changes arising from crop rotation.

2.1.2 Daily Landsat-MODIS fused data

Satellite observations with high temporal frequency and high spatial resolution can be generated by fusing Landsat and MODIS data together (Gao et al. 2006). One commonly used data fusion methodology is the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM), which combines the higher spatial resolution of Landsat data with the high temporal MODIS observations to produce higher spatiotemporal resolution data (Gao et al. 2006; Hilker et al. 2009; Zhu et al. 2010). This approach compares one or more pairs of observed Landsat and MODIS datasets collected on the same day to predict maps at Landsat-scale on other MODIS observation dates (Gao et al., 2006). Recently, STARFM has been modified and extended for different applications, which includes the Spatial Temporal Adaptive Algorithm for mapping Reflectance Change (STAARCH) for the detection of reflectance changes associated with land cover change and disturbance (Hilker et al. 2009), and an enhanced STARFM (ESTARFM) approach for the fusion of very heterogeneous scenes without "pure" pixels (Zhu et al. 2010).

The STARFM approach was used here to produce Landsat-MODIS fused daily 30 m surface reflectance in 2013 and 2014 (Gao et al., 2006; 2017). Specifically, the MODIS daily directional surface reflectance (250 m MOD09GQ and 500 m MOD09GA) (Vermote et al. 2002) in tiles H10V04 and H11V04 were obtained and corrected to Nadir Bidirectional Reflectance Distribution Function (BRDF)-Adjusted Reflectance (NBAR) data using MODIS BRDF product (500 m MCD43A1) (Schaaf et al. 2002). The Landsat 8 OLI surface reflectance data (in path 26 and row 31) were downloaded from the USGS EROS (Earth Resources Observation and Science) Data Center, in which the Landsat digital number data were calibrated and atmospherically corrected using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (Masek et al. 2006). The OLI observations we used here were acquired at the following days of year (DOY) in 2013: 140, 188, 252, 268, 284, 300; and in 2014: 79, 95, 127, 143, 175, 191, 271, 303, 351. Note that Landsat 7 ETM+ imagery was not used because of the gaps resulting from the failure of the Scan Line Corrector (SLC). Finally, Landsat images on each MODIS date were then simulated with STARFM using co-temporal pairs of Landsat and MODIS imagery. The fused daily 30 m time series exhibited mean biases of ±0.01 for the red band and ±0.02 for the NIR band. Henceforth this data will simply be called “OLI”. In this dataset, an observation was defined as good quality if it was from either a cloud-free observation in Landsat or MODIS NBAR produced using a full BRDF inversion model (Schaaf et al. 2002), while the remaining observations were considered as other (poor) quality.

2.1.3 VIIRS NBAR data

The VIIRS instrument onboard the Suomi National Polar-orbiting Partnership (NPP) has a similar design to MODIS. VIIRS observes the surface at local time around 1:30pm. It acquired its first measurements on November 21, 2011. The spatial resolution is 375 m at nadir for the red (0.60-0.68μm) and near infrared (0.846-0.885 μm) bands. The VIIRS NBAR is produced utilizing a similar algorithm as the MODIS Collection V006 daily BRDF/Albedo/NBAR product (Schaaf et al. 2002). The NBAR product is ideal for land surface analysis since the view angle effects have been removed using BRDF estimates and the daily cloud and aerosol contaminations have been reduced or corrected in the surface reflectance product. Although the BRDF estimation is based on directional reflectance within a temporal window, the reflectance on the day of interest is emphasized to retain the phenological characteristics of that day. This product also provides quality assurance (QA) field indicating the quality of the surface reflectance, which includes snow flag, good quality, other (poor) quality, and fill values (Schaaf et al. 2002). In this study, the daily 500m NBAR data were produced for the tiles of H10V04 and H11V04 from January 1 2013 to December 31 2014.

2.2 Land surface phenology detection

First, the daily two-band enhanced vegetation index (EVI2) was generated from both the VIIRS NBAR and OLI datasets. The EVI2 is calculated from red and near infrared reflectance by removing the blue-reflectance influence on enhanced vegetation index (EVI) through an empirical relation between red and blue reflectance (Jiang et al. 2008; Rocha and Shaver 2009). Thus, EVI2 can also be derived from satellite sensors without blue reflectance, such as the AVHRR. EVI2 remains functionally equivalent to EVI (enhanced vegetation index) and has previously been used to monitor vegetation phenology (Jiang et al. 2008; Rocha and Shaver 2009; Zhang et al. 2014), but it is less sensitive to background reflectance, including bright soils and non-photosynthetically active vegetation (i.e., litter and woody tissues) than some other vegetation indices (Rocha et al. 2008).

Second, land surface phenological metrics were then retrieved using the hybrid piecewise-logistic-model-based LSP detection algorithm (HPLM-LSPD; Zhang 2015; Zhang et al. 2003). The HPLM-LSPD first reconstructed the EVI2 temporal trajectory in a pixel following previously described methods (cf., Zhang 2015). Briefly, spuriously large daily EVI2 values were removed if they were larger than 90% of the corresponding daily NDVI in the time series, which were likely subject to red band overcorrection in some observations that were contaminated by either residual snow or atmosphere (Justice et al. 2013; Zhang 2015). The daily EVI2 values were used to generate a 3-day composite dataset by applying the maximum value composite approach to the EVI2 data selected with best quality observation within the 3-day window. The EVI2 values contaminated by snow were identified using the VIIRS snow flag and were replaced using a background EVI2 value at each pixel. The background EVI2 value is referred to as the minimum EVI2 within the vegetation growing cycle that is not contaminated by snow and clouds or the maximum EVI2 during the phase of vegetation dormancy. It was determined by averaging five good observations (without cloud and snow contamination) during the winter period, which was identified using a MODIS LST (land surface temperature) climatology (LST<278K). Short gaps caused by clouds in the time series were replaced using a moving average of two neighboring good quality values starting from the point close to larger EVI2 values. If a gap was longer than one month, the corresponding EVI2 values were replaced using good quality observations in preceding or succeeding years, but the detected LSP was labeled as low confidence. The time series of EVI2 data at each pixel was further smoothed using a Savitzky-Golay filter and a running local median filter with a five 3-day window. The median filter could remove local sharp peaks or troughs in the time series. Finally, the hybrid piecewise logistic functions were applied to reconstruct the temporal EVI2 time series.

Phenological transition dates within each growth or senescence phase were detected using the rate of change in the curvature of the modeled curves. Specifically, transition dates correspond to the day of year on which the rate of change in curvature in the EVI2 time series data exhibits local minima or maxima (Zhang et al., 2003). Because phenological detections are significantly impacted by the number of good satellite observations during the period of phenological occurrences (Zhang et al. 2009), we further calculated the proportion of good quality (PGQ_{sos}) EVI2 observations during three 3-day periods before and after the start of growing season (SOS), respectively. This critical period was selected because the phenological metrics could be reasonably detected, if there was a good quality EVI2 observation within 8 days (Zhang et al. 2009).

2.3 Matchup of SOS detected from OLI and VIIRS data

OLI SOS and VIIRS SOS were matched spatially and qualitatively in order to compare these two datasets properly. Two VIIRS titles (H10V04 and H11V04) were first adjoined to cover the entire Landsat 8 OLI scene (path 26 and row 31). Both OLI and VIIRS data were then re-projected to the Universal Transverse Mercator *(*UTM*)* projection with a spatial resolution of 30 m and 450 m, respectively, resulting in one VIIRS pixel containing 225 OLI pixels. OLI SOS detections were also spatially matched with a grid of 3 by 3 VIIRS pixels (hereafter called the VIIRS grid) to reduce the spatial mismatch between these two datasets. The mismatch is caused by the following factors. First, the pixel size in VIIRS red and near infrared bands is 375 m at nadir while it is over 500 m at high scan angles. Second, the actual pixel size is 463.312 m (instead of 500 m) in the NASA 500 m VIIRS reflectance product, but the spectral reflectance data represent a median effective resolution of 565m x 595m (Campagnolo et al. 2016). Therefore, we also investigated SOS using the matched VIIRS grid that contains 9 VIIRS pixels or 2025 OLI pixels to ensure a better spatial match.

To qualitatively match the SOS, the SOS pixels with low PGQsos were t removed from both OLI and VIIRS detections. Based on sensitivity analysis (Zhang et al., 2009), we considered the SOS detection as high confidence if PGQ_{sos}>40%. The precision of SOS detection can be greatly reduced if there were very few or no good satellite observations during the period of SOS occurrence. Therefore, we only selected the pixels with PGQ_{sos}>40%, which are hereafter referred to as “high confidence SOS pixels”. The pairs of VIIRS and OLI SOS observations were also removed if the number of high confidence OLI SOS pixels was fewer than 200 (~90%) within a VIIRS footprint. Further, the VIIRS grids were excluded if the number of good VIIRS SOS detections was fewer than 7 out of 9 pixels.

2.4 Comparison of OLI SOS and VIIRS SOS

We compared VIIRS SOS with OLI SOS in order to characterize the biophysical context of the SOS derived from a coarser pixel to the SOS at finer scale. Therefore, the comparison was conducted across various levels of heterogeneity and with a set of aggregated OLI SOS values.

The SOS can vary greatly in heterogeneous areas, while it is relatively similar in homogeneous areas. To understand the impact of spatial heterogeneity on SOS detections at a coarser scale, we divided the entire study area into five levels of SOS heterogeneity. To do this, the standard deviation

(SD) of OLI SOS within a VIIRS pixel and grid was calculated, and its cumulative frequency distribution was established across the entire study area in 2013 and 2014, separately. The five levels of heterogeneity for VIIRS pixels were then determined using the proportion of OLI SOS SD frequency (PSD) at an interval of 20%: 0-20% PSD represents the most homogeneous level, whereas 80-100% PSD indicates the most heterogeneous level.

The OLI SOS was then aggregated to be comparable with VIIRS SOS. The aggregated OLI SOS is called "SOSag" hereafter. In previous studies, the SOS at the coarser scale was generally averaged from all SOS values or high frequency SOS values at a finer scale (Delbart et al. 2015; Ganguly et al. 2010). Biophysically, the SOS becomes detectable from satellite sensors after a certain amount of leaves within the pixel start to emerge. This means that the SOS value detected in a coarser pixel is associated with earlier SOS values (the plant leaves that emerge earlier) at finer pixels rather than later SOS pixels. To explore the correspondence of VIIRS SOS to OLI SOS, we aggregated a set of SOSag by selecting the timing at a specific percentile at an interval of 5% (starting from 0.5% which represents the earliest OLI SOS) from the cumulative OLI SOS frequency distribution within a VIIRS pixel or grid (Figure 2). We call this approach "percentile aggregation". In this way, we obtained 21 potential timings of SOSag in a VIIRS pixel or grid. From a biophysical perspective, SOSag from this percentile approach represents the date at which vegetation greenup has occurred in a certain percent of the OLI pixels, namely 0.5%, 5%, 10%, 15%, … 100%.

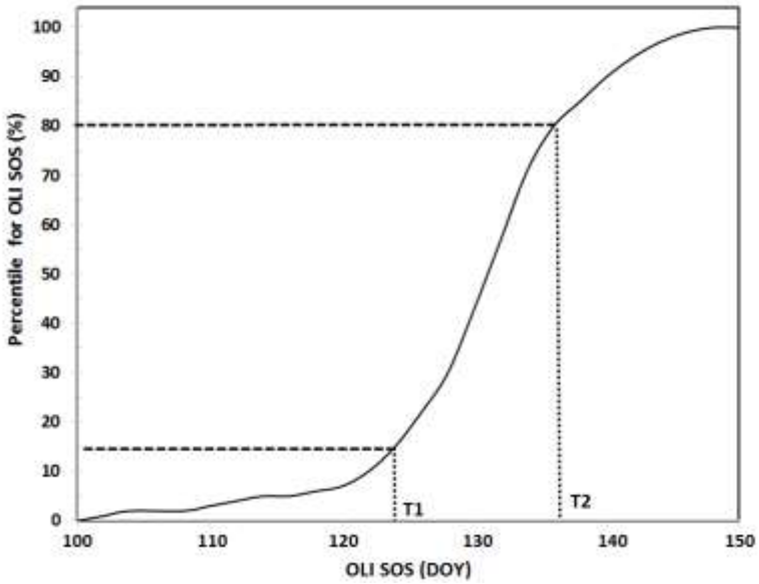

Figure 2. Schematic diagram of the percentile approach to aggregate SOS from finer scales (OLI) to coarser scales (VIIRS pixel or grid). The percentile represents OLI SOS distribution within a VIIRS pixel. The T1 and T2 are the examples of the SOSag aggregated using 15th and 80th percentile, separately.

VIIRS SOS was statistically compared with the SOSag using average absolute difference (AAD), mean difference (bias), root mean square difference (RMSD), and linear regression. AAD was a measure of statistical dispersion equal to the average absolute difference of two independent variables. RMSD was used to evaluate the average uncertainty between two observations. Note that root mean square error (RMSE) was not used here since both the OLI SOS and VIIRS SOS are remote sensing estimates without a clear reference a priori. Bias was used to evaluate the overestimation (positive bias) or underestimation (negative bias) of the two variables. Linear regression was used to examine the overall relation between the samples.

$$AAD = \frac{\sum_{i=1}^{N} \left| SOS_{OLI} - SOS_{VIIRS} \right|}{N} \qquad (1)$$

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N}\left(SOS_{OLI} - SOS_{VIIRS}\right)^2}{N}} \tag{2}$$

$$Bias = \frac{\sum_{i=1}^{N}\left(SOS_{OLI} - SOS_{VIIRS}\right)}{N} \tag{3}$$

The statistical comparison between VIIRS SOS and OLI SOSag was conducted for 21 different potential SOSag timings and five levels of heterogeneity in 2013 and 2014, separately, for a total of 210 comparisons. The analysis allowed us to determine the scale effects on the SOS at coarser resolutions. It was further used to reveal the most appropriate approach to aggregate SOS from finer resolution to coarser resolution.

3. Results

The SOS data derived from finer (OLI) and coarser (VIIRS) resolutions are investigated and compared. The spatial patterns of SOS and the corresponding confidence (including all different confidence levels) are presented in section 3.1, which provides the impacts of EVI2 data quality in OLI and VIIRS observations on SOS detections across the entire study area. Next, the scaling effects on SOS detections are illustrated and evaluated in sections 3.2-3.4 based on the OLI SOS and VIIRS SOS data that were of high confidence and matched spatially and qualitatively.

3.1 Spatial pattern of SOS detected from OLI data and VIIRS data

Figure 3 shows the spatial patterns of SOS (including all different confidence levels) derived from OLI data and VIIRS observations in 2013 and 2014. SOS dates were similar in 2013 and 2014, although dates were slightly earlier in 2013 than in 2014. However, PGQsos was relatively poorer in 2013 compared to 2014 in the southeastern region of the study area for both OLI and VIIRS data, and the northwestern region only for OLI. Overall, SOS was delayed moving northward: from DOY 100 to 160. Early SOS was mainly distributed in the southern portion of the study area, where forests dominate. Relatively later SOS was found in the northwestern region, where the croplands were the main land cover. In the central eastern portion of the study area, crops and natural vegetation were interspersed: SOS exhibited dates that were earlier for natural vegetation while later for croplands.

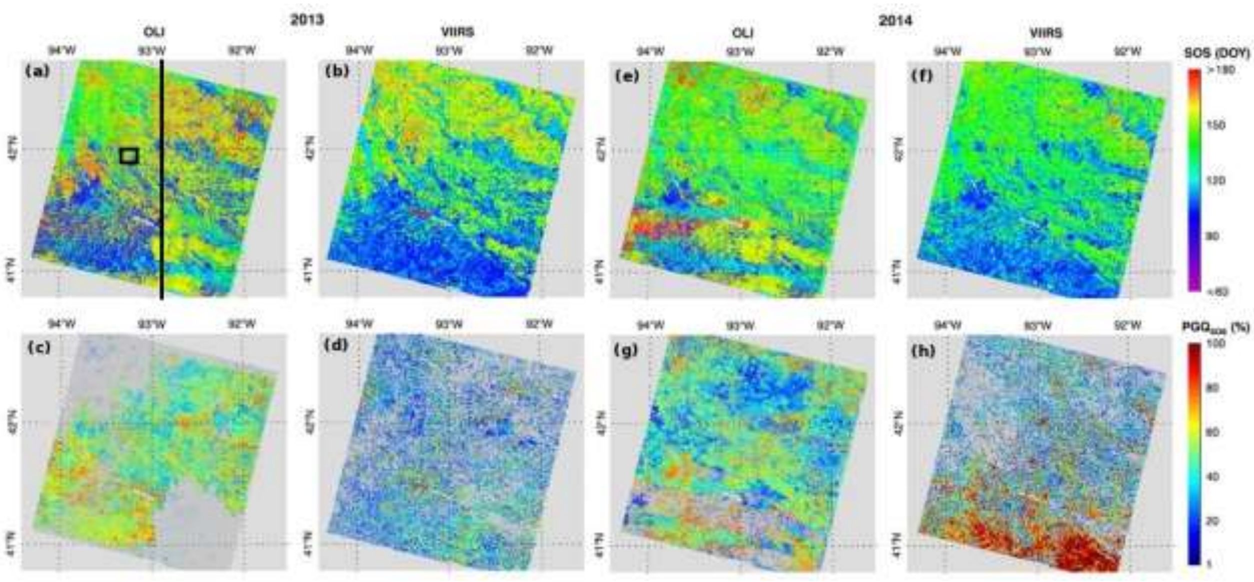

Figure 3. Spatial distributions of SOS and data quality in the time series of OLI and VIIRS in 2013 and 2014. The top row represents the SOS detected from OLI in 2013 (a) and 2014 (e) and from VIIRS in 2013 (b) and 2014 (f). The bottom row is PGQsos around SOS occurrence along the OLI time series in 2013 (c) and 2014 (g) and the VIIRS time series in 2013 (d) and 2014 (h). The black vertical line and box in (a) indicate the

locations for Figures 4 and 5, respectively. The dark gray color indicates no good EVI2 observations around SOS occurrence (PGQsos=0).

The north-south SOS gradient was more evident in the VIIRS data than in the OLI data. To the north, OLI SOS matched well with VIIRS SOS while larger differences were apparent in the south. This spatial inconsistency was apparently associated with the quality of the SOS detections, which was determined by the frequency and availability of OLI and VIIRS observations. The OLI PGQsos was generally higher than 20% from southwest to northeast in 2013, but it was very low in large areas across both the northwest and southeast corners because of lack of high quality OLI observations during the SOS periods. In 2014, the OLI PGQsos revealed no high quality temporal observations in large parts of the southern region. In contrast, the VIIRS PGQsos was relatively high in the southern region, particularly in 2014, although there were still various randomly distributed spots without high quality VIIRS observations during SOS periods.

The spatial transect of SOS dates exhibits patches of earlier and later occurrences, although there is a clear trend from earlier in the south to later in the north (Figure 4). This change mainly follows land cover types, but a latitudinal effect cannot be discounted as the transect spans nearly two degrees of latitude. The SOS could be more than one month earlier in forests and grasslands than in croplands. This pattern was evident from 41.5°N northward, where corn and soybean were abundant and forests and grasslands were sparse. In contrast, forest and grasslands were the main cover types in the southern region, so that SOS was conspicuously early with small proportion of late SOS dates over croplands in the VIIRS time series. However, there was also a spatial pattern of poor PGQsos in the south driven by a lack of sufficient high confidence OLI observations around the estimated timing of SOS.

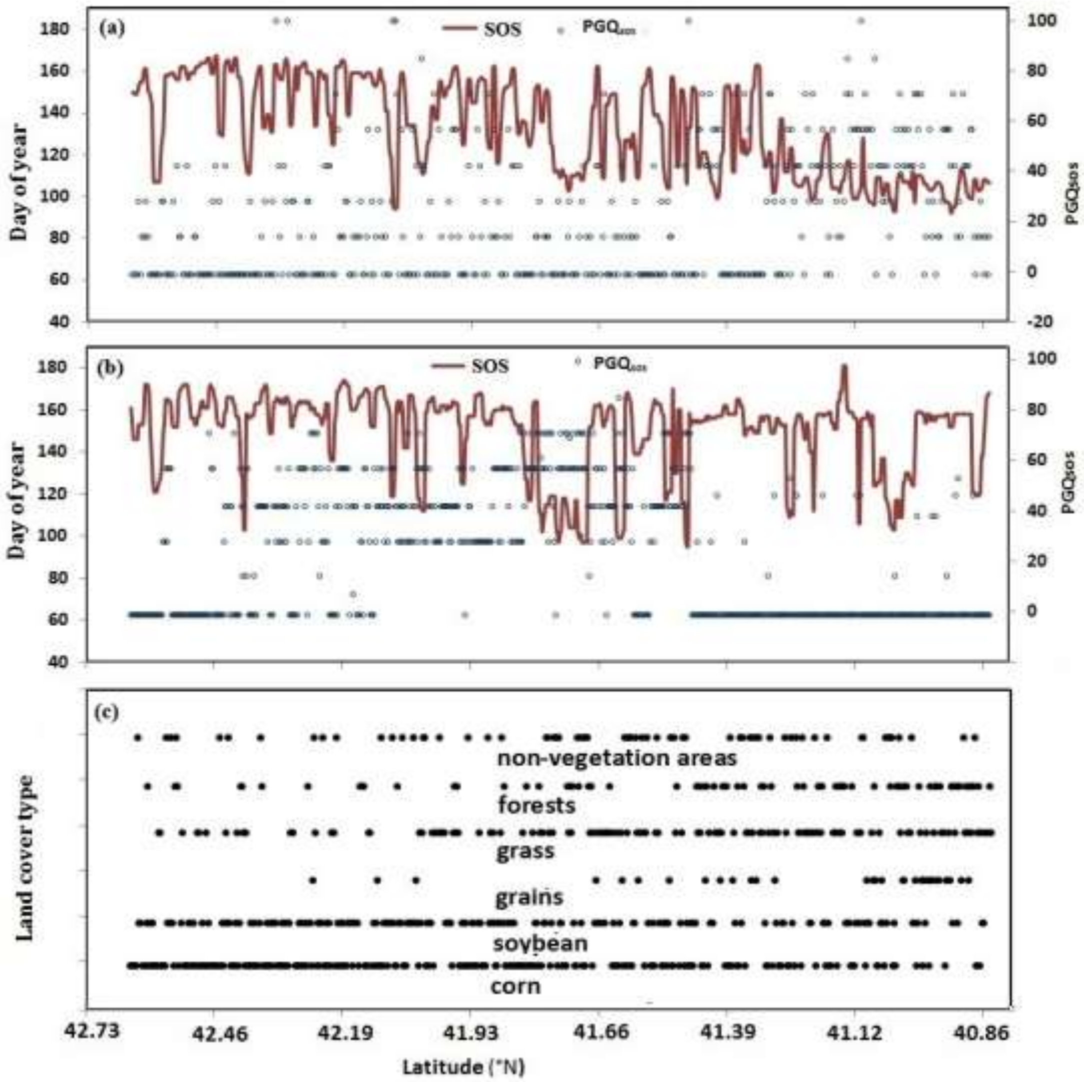

Figure 4. Variation in SOS, PGQsos, and land cover types along a north-south transect in 2013. (a) detections from VIIRS data, (b) detections from OLI data, and (c) land cover type. The transect location is identified in Figure 3.

Closer examination revealed that OLI SOS varied substantially even within a limited area (Figure 5). The OLI SOS displayed sharp boundaries with a difference as large as one month among neighboring crop fields and among different crop types; whereas the SOS was generally homogeneous within larger fields of the same crop type. Similarly, OLI SOS presented heterogeneous patterns between forests and croplands while it was relatively homogeneous within local forests and grasslands. The SOS heterogeneity among different crop fields and between croplands and forests or grasslands was verified using a high resolution image (Google Earth) from June 2012, which visually indicated growth conditions among different fields (although of course the crop types might be not exactly the same between 2012 and 2013). In contrast, the VIIRS SOS only captured the large spatial patterns of SOS rather than the details in and among individual fields, but the overall coarse-scale spatial pattern corresponded well with the OLI SOS.

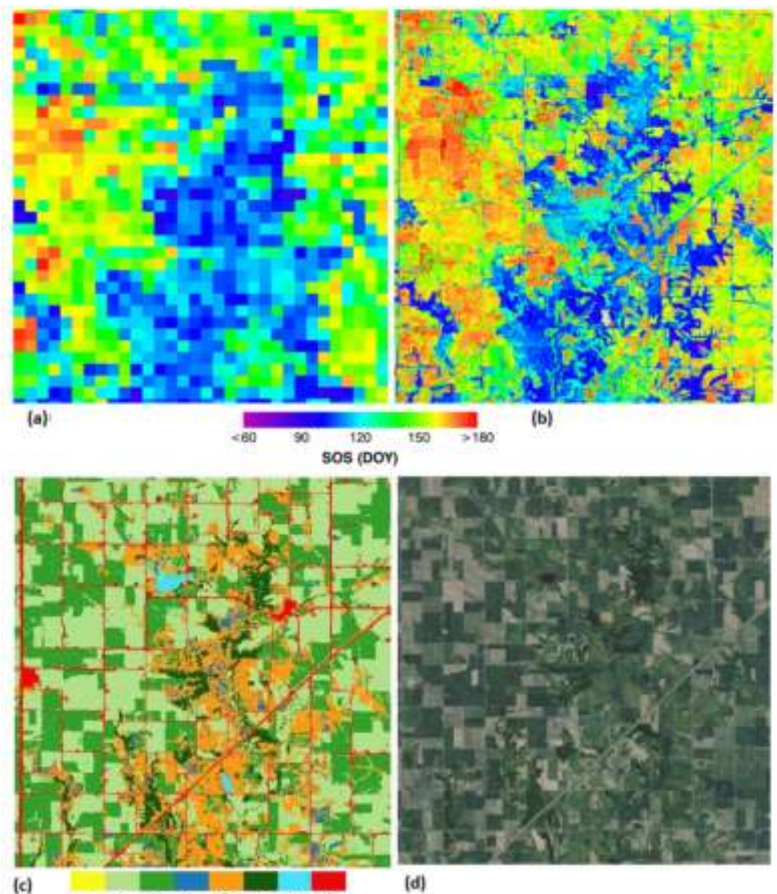

Figure 5. Local pattern (14280m x 14880m) of SOS from OLI and VIIRS across different land cover types in 2013. (a) VIIRS SOS, (b) OLI SOS, (c) land cover type (1- other crops, 2- corn, 3- soybean, 4-hay, 5-grasslands, 6-forests, 7-water and wetlands, 8- non-vegetated area), and (d) Google Earth imagery from June 2012. The location is defined in Figure 3.

3.2 Heterogeneity within VIIRS SOS pixels

Figure 6 presents the heterogeneity of the OLI SOS within a VIIRS pixel (225 OLI pixels) after the low confidence SOS pixel pairs were removed using the criterion that the percent of high confidence OLI SOS pixels within a high confidence VIIRS pixel (PGQsos>40%) was over 90% (~200 pixels). The result indicates that the heterogeneities were only distinguished in 12,583 VIIRS pixels in 2013 and in 20,707 pixels in 2014. The spatial patterns of the selected pixels between the two years were generally inconsistent (Figures 6a and 6c), because of differences in the ranges of OLI SOS SD (Figures 6b and 6d).

The frequency distribution of OLI SOS SD within VIIRS pixels varied between 2013 and 2014 (Figure 6b and 6d). A peak in both years appeared around 4 days of SD. However, SD frequency indicated that OLI was more heterogeneous in 2013 than in 2014. The SD frequency in 2014 was skewed right (positively skew), and the cumulative frequency was larger than 60% for SD<10 days. In contrast, the frequency in 2013 was relatively uniform in the SD range between 10 and 27 days, and the cumulative frequency was about 40% for SD<10 days. The largest SD was 40 days in 2013 and 30 days in 2014.

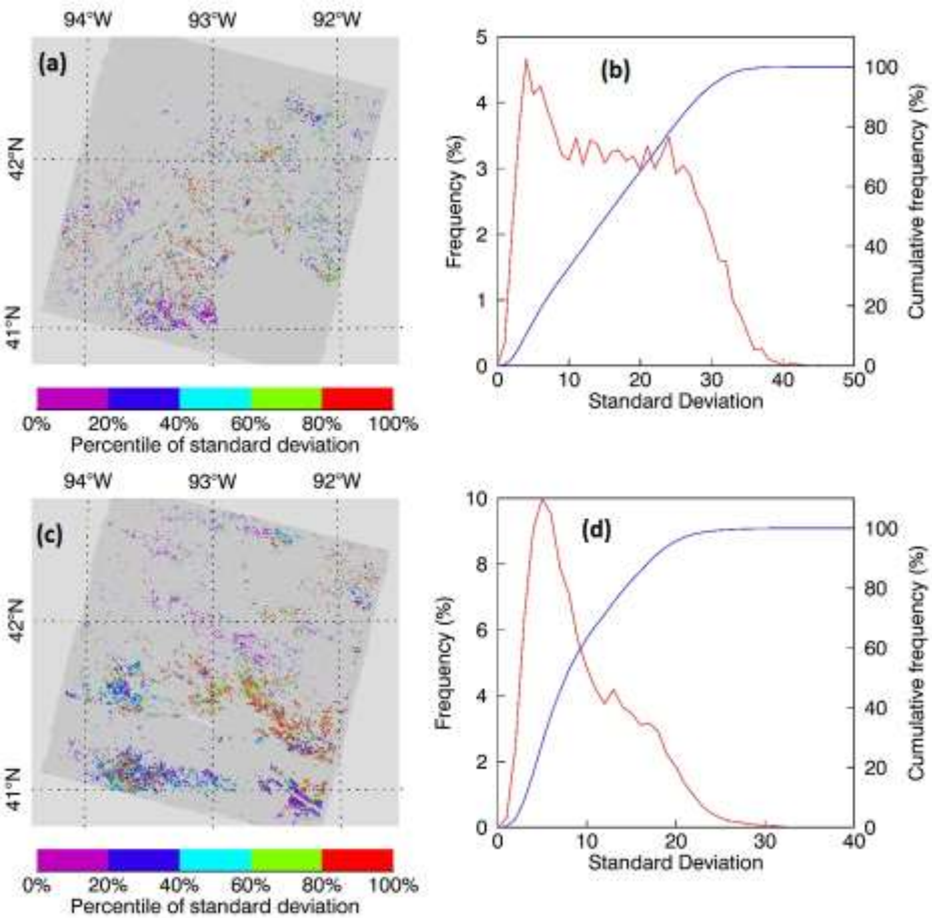

Figure 6. Heterogeneity of OLI SOS within a VIIRS pixel with high confidence VIIRS SOS detection. Heterogeneity levels were defined using percentile of OLI SOS standard deviation in 2013 (a) and 2014 (c), and frequency distributions represented VIIRS pixels varying with OLI SOS standard deviation in 2013 (b) and 2014 (d). Gray color in (a) and (c) represents the VIIRS pixels with either PGQsos<40% or the percent of high confidence OLI SOS pixels<90%.

Figure 7 displays the frequency distribution of high confidence OLI SOS dates within a VIIRS pixel at five levels of heterogeneity. These randomly selected VIIRS pixels represent several typical types of SOS heterogeneity across the study area. Within homogenous VIIRS pixels (PSD<20%), the OLI SOS frequency displayed a strong peak of more than 10% (>23 pixels) at the same SOS and most OLI SOS estimates were within 10 days of each other. Correspondingly, the cumulative frequency showed a pattern of sharp increase. Within heterogeneous VIIRS pixels (PSD>60%), by contrast, the OLI SOS dates varied within a wide range spanning more than three months, and the cumulative distribution exhibited a relatively flat pattern. The frequency at the same SOS date was less than 3% (<6 OLI pixels). In some cases, where a VIIRS pixel contained several different crop types and natural vegetation with divergent SOS values, the OLI SOS frequency displayed multiple distinct peaks.

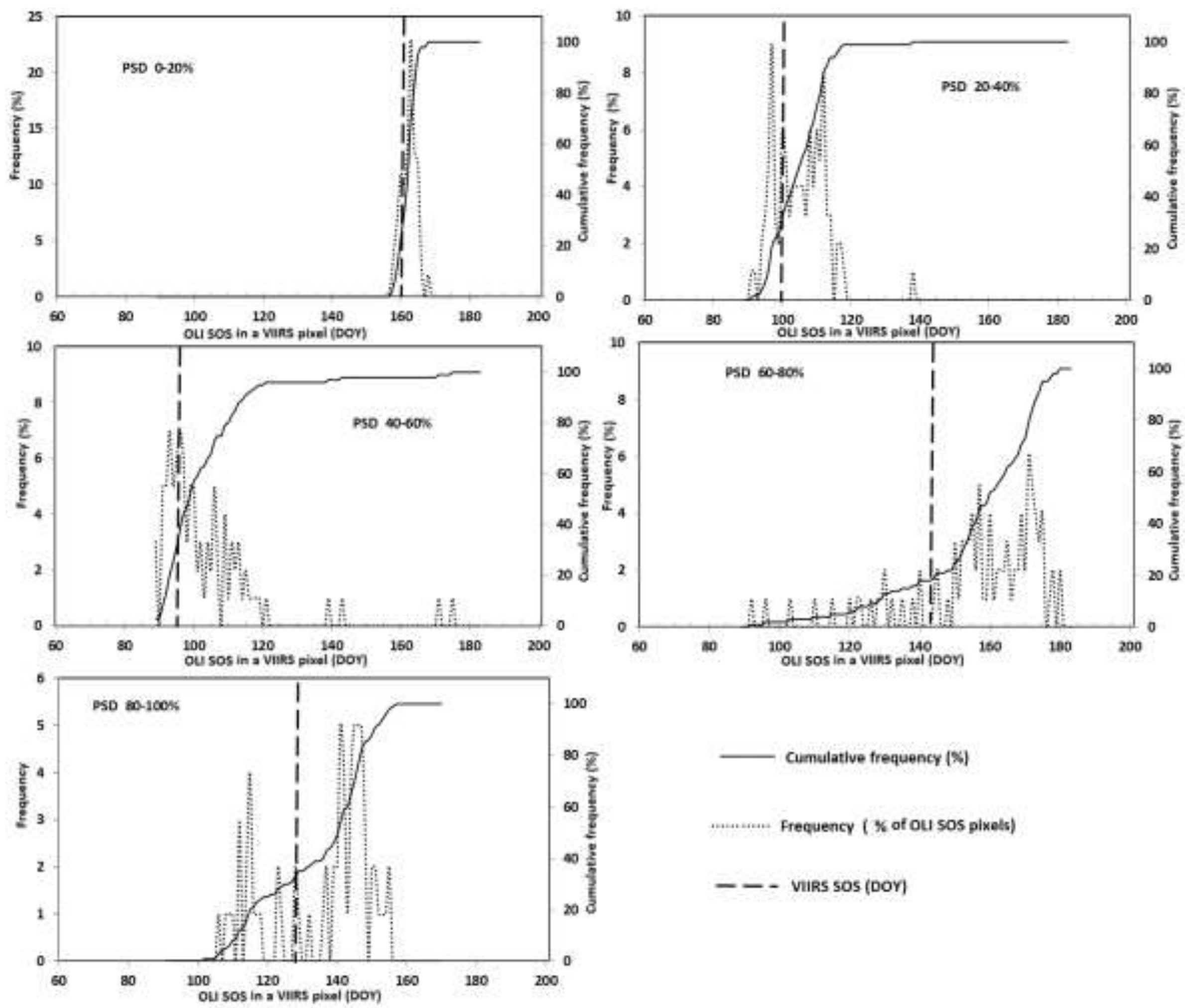

Figure 7. Frequency and cumulative frequency distributions of high confidence OLI SOS within a high confidence VIIRS pixel across different levels of heterogeneity and the corresponding VIIRS SOS in 2013.

3.3 Difference between VIIRS SOS and OLI SOSag

Figure 8 shows the average absolute difference between VIIRS SOS and a set of OLI SOSag aggregated by the percentile approach (as described in Figure 2) over the various levels of heterogeneity. For OLI SOSag aggregated using a specific percentile, AAD increased with increasing heterogeneity, resulting in AAD values for the most heterogeneous pixels (PSD=80-100%) that were more than twice as large as the most homogenous pixels (PSD=0-20%). AAD differences in relatively homogenous pixels (PSD<60%) were generally less than 10 days, but generally larger than 10 days in heterogeneous pixels (PSD>60%). If all pixels (PSD=0-100%) were considered, AAD was similar to the values from the middle heterogeneity level, i.e., PSD=40-60%.

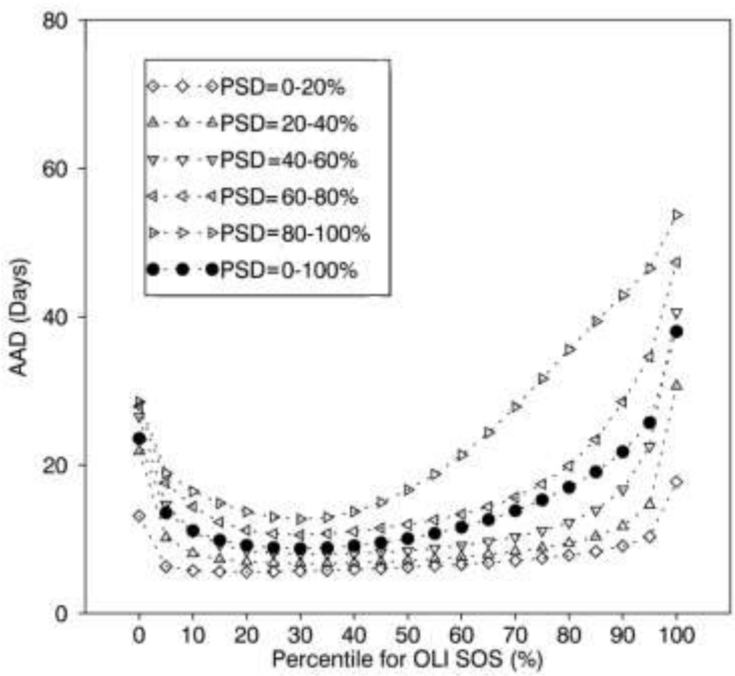

Figure 8. Average absolute difference between VIIRS SOS and OLI SOSag at different levels of heterogeneity based on data in both 2013 and 2014. OLI SOSag was aggregated within a VIIRS pixel using the percentile approach.

At a same level of heterogeneity, AAD varied more than threefold (Figure 8), if the OLI SOSag was aggregated using different OLI SOS percentiles. AAD was very large, if SOSag was aggregated from either <10^{th} or >90^{th} percentile of OLI SOS within a VIIRS pixel. For example, AAD was 57 days, if SOSag was obtained using 100^{th} percentile of OLI SOS; whereas it was 32 days, if SOSag was obtained using 0.5^{th} OLI SOS percentile in the most heterogeneous regions (PSD=80-100%) (Figure 8). However, AAD reached minimum (AADmin), if SOSag was aggregated from an optimal OLI SOS percentile. In homogenous regions (PSD=0-20%), the low AAD (< AADmin + 1 day) was reached, if OLI SOSag was selected from 5^{th} - 70^{th} percentiles. However, in the most heterogeneous region (PSD=80-100%), the low AAD (< AADmin + 1 day) was obtained, if OLI SOSag was selected from 20^{th} - 40^{th} OLI SOS percentile. The range of optimal percentiles with the low AAD varied from larger than 45% in homogeneous regions to less than 20% in heterogeneous regions. If the SOS values in the entire region were considered together (PSD=0-100%), then the AAD was distributed between those from homogenous and heterogeneous regions. Overall, AAD was smallest if OLI SOSag in a VIIRS pixel was aggregated as the date when SOS had occurred in 20%-40% of OLI pixels. In contrast, AAD was largest if OLI SOSag was considered as the date when SOS had appeared in more than 80% of the OLI pixels.

Figure 9 depicts the bias between VIIRS SOS and OLI SOSag aggregated from different OLI SOS percentiles across various levels of heterogeneity. Negative bias appeared if OLI SOSag was aggregated from the 0.5^{th} - 30^{th} percentiles of OLI SOS within a VIIRS pixel, while positive bias mainly occurred if OLI SOSag was aggregated from the 40^{th} - 100^{th} percentiles. This pattern was similar for all the levels of OLI SOS heterogeneity. Similar to AAD, the bias was smaller in homogeneous regions than in heterogeneous regions. Moreover, the negative bias could be as large as 30 days and the positive bias could be as large as 50 days. Overall, if OLI SOSag was aggregated using the timing around the 30th percentile, then the bias approached zero.

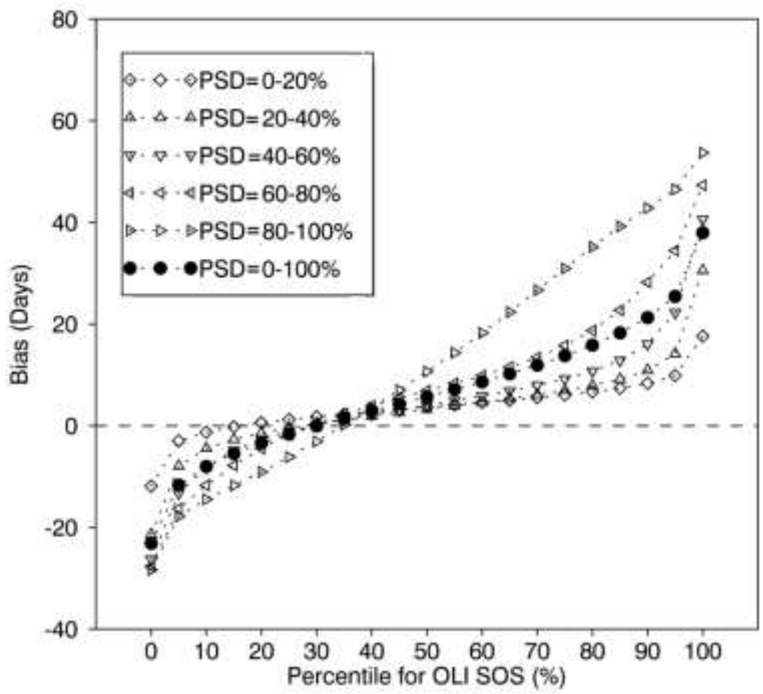

Figure 9. Bias between VIIRS SOS and OLI SOS aggregated using the percentile approach within a VIIRS pixel at different levels of heterogeneity based on data from both 2013 and 2014.

3.4 Evaluation of SOS aggregation at different scales

The relationship between VIIRS SOS and OLI SOS was evaluated by comparing VIIRS SOS with a set of OLI SOSag within a VIIRS grid (3 by 3 VIIRS pixels). We first generated VIIRS SOS in a VIIRS grid ($SOS_{VIIRS}ag$) using the approach of 30th percentile based on the result obtained at individual VIIRS pixels (see section 3.3). The $SOS_{VIIRS}ag$ was then compared with a set of OLI SOSag aggregated using the percentile approach within a VIIRS grid. AAD in a VIIRS grid displayed similar patterns as in a single VIIRS pixel, but with much smaller magnitudes (Figure 10). AAD was smallest in the most homogeneous regions and increased with heterogeneity. At the same level of heterogeneity, AAD was smallest when SOSag was aggregated by choosing the date when OLI SOS occurred in 30% of OLI pixels in the relatively homogeneous grids, and at 40% of OLI pixels in the more heterogeneous grids. If SOSag in a VIIRS grid was aggregated using the timing of 30^{th}-40^{th} percentile of OLI SOS, then the AAD was less than 5 days in homogeneous grids and 15 days in heterogeneous grids.

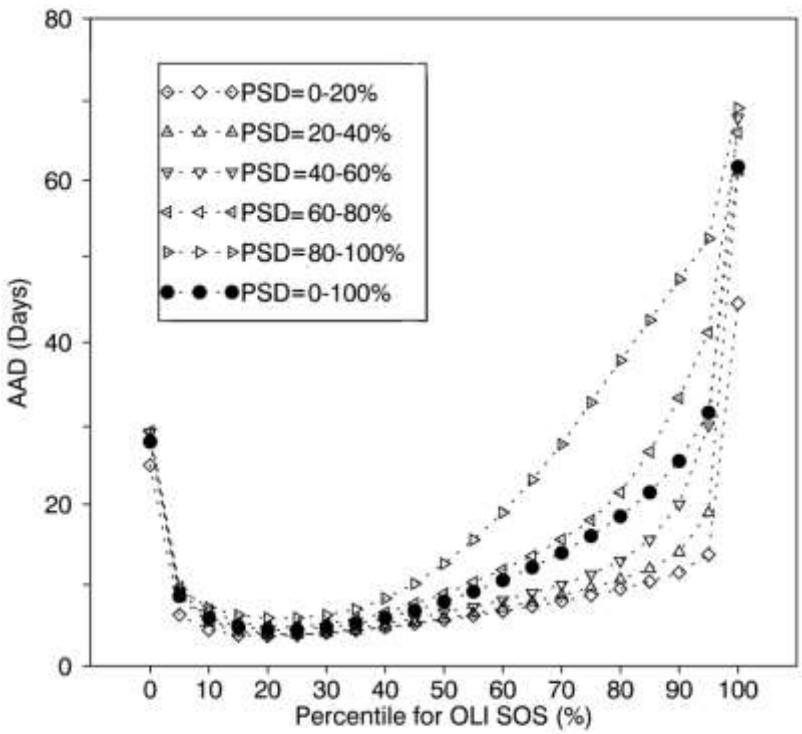

Figure 10. AAD between VIIRS SOS and SOSag within VIIRS grids aggregated using the percentile approach at different levels of heterogeneity based on data in both 2013 and 2014.

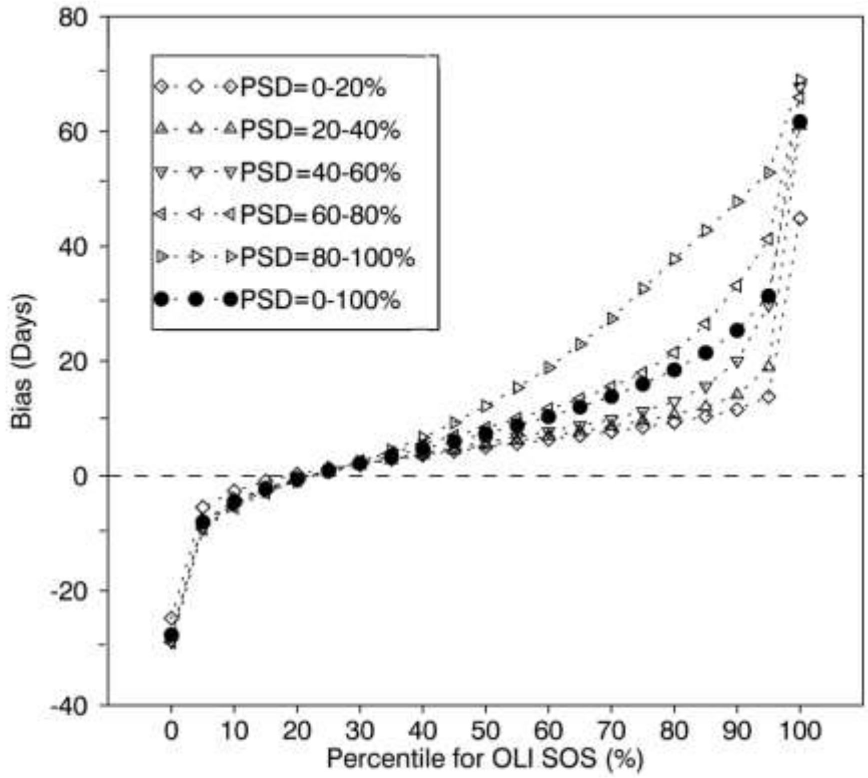

Figure 11. Bias between VIIRS SOS and SOSag within VIIRS grids aggregated using the percentile approach at different levels of heterogeneity based on data in both 2013 and 2014.

Figure 11 displays the bias between VIIRS SOS$_{VIIRS}$ag and SOSag in a VIIRS grid. Similar to the comparison at a VIIR pixel (Figure 9), the bias was negative if SOSag was aggregated using <20^{th} percentile of OLI SOS and the bias was positive if SOSag was aggregated from 35-100^{th} percentile. The bias approached zero if SOSag was selected from 20-30^{th} percentile of OLI SOS, identical to the pixel-based result.

Figures 12 presents the difference between VIIRS SOS$_{VIIRS}$ag and OLI SOSag aggregated using the timing at the SOS occurrence of 30% OLI pixels in a VIIRS grid. The samples were closely distributed along the 1:1 lines with slopes close to 1 in homogeneous regions. With the increase of heterogeneity, the intercept in the linear regression increased while the slope decreased, and the correlation coefficients were also reduced. AAD was less than 5 days and RMSD less than 6 days in the homogeneous region (PSD<60%). In the most heterogeneous region (PSD=80-100%), AAD was 6 days and RMSD was 8 days. If all the good SOS pixels across the regions were considered (0-100% PSD), then the AAD and RMSD were 5 days and 6 days, respectively.

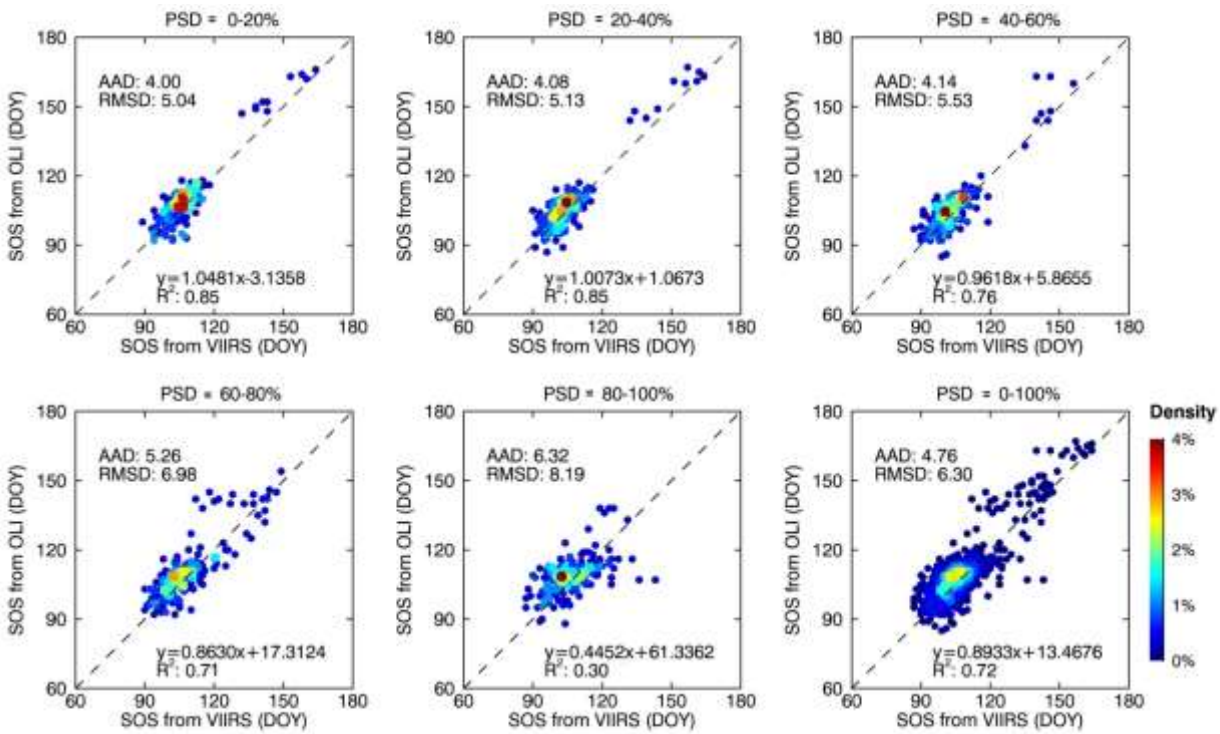

Figure 12. Scatterplots of VIIRS SOS and OLI SOSag at different levels of heterogeneity in 2013 and 2014. The color indicates the sample density, increasing from blue to red.

4. Discussion and Conclusion

This study investigated and compared SOS dates as estimated from remote sensing data at two common spatial resolutions: 500 m and 30 m. The SOS at different scales was retrieved from the fused OLI data and from VIIRS observations, instead of aggregation from the same finer resolution, which avoided the risk that SOS agreement arose due to the data source being identical at both scales. It should be noted, however, that the time series of 30 m data were fused from observations with high spatial resolution of Landsat and higher frequency of MODIS using the STARFM algorithm. Because the fusing algorithm relies on the existence of co-temporal pairs of Landsat and MODIS image to predict Landsat images on a MODIS date (Gao et al., 2006), the quality of the fused time series is dependent on the number of observations from Landsat. In 2013, there are only 6 Landsat OLI observations available to use and no data available before DOY 140, so that the fused time series likely contains large uncertainties during spring. In contrast, there are 9 Landsat OLI observations in 2014 and 5 observations from DOY 70-175 (spring to early summer), which is likely to produce more reliable fused time series and more accurate SOS dates. Moreover, fused time series were affected by off-nadir observations from Terra MODIS images with reduced spatial resolution (Campagnolo et al. 2016). We expect that these results will be further explored and verified once time series observations from Sentinel-2 satellite and Landsat 8 OLI are well-calibrated and combined.

The selected research area covers a wide range of SOS heterogeneity, which enabled us to explore the complexities of SOS variation in coarser resolution pixels. Within the same crop field, SOS patterns and growth conditions were relatively homogeneous because of the same management practice. In a 30 m pixel, SOS could well reflect the planting date and crop germination for specific crop varieties, because the mean size of crop fields that had a prominent and contiguous boundary with the same crop type was 0.193 km^2 (~214 Landsat pixels) across the central US (Yan and Roy 2016). However, changing agronomic practices resulted in dramatic changes of crop types and varieties among neighboring fields. As a result, the crop planting timing for various fields could vary sharply with a time difference of more than three months. A sharp difference in SOS was also evident between crops and natural vegetation across the study area, where SOS was more than one month earlier in natural vegetation than croplands. Among natural vegetation, SOS could also shift greatly due to microclimatic changes (Augspurger et al. 2005; Fisher et al. 2006; Richardson and O'Keefe 2009). Consequently, OLI SOS SD in a VIIRS pixel was observed to be as large as 40 days in 2013 and 30 days in 2014, although the peak frequency of OLI SD appeared at 4 SD days across the study area. The heterogeneous regions with SD larger than 10 days were 40% in 2014 and 60% in 2013. Although the heterogeneity levels were defined using the cumulative frequency distribution of OLI SD in a given year and the SD differed within the same heterogeneity level between years, the relative impacts of heterogeneity on SOS detections remained constant.

Comparisons between VIIRS SOS and OLI SOS by selecting only high confidence SOS retrievals (PGQsos>40%) ensured the reliability of the results. SOS detections were significantly affected by the quality of satellite observations used in the time series. Poor SOS detections were removed using PGQsos threshold during the period of SOS occurrences. PGQsos showed no consistent spatial and temporal patterns, because it was driven mainly by patterns of missing data that typically resulted from cloud cover. OLI PGQsos was very poor (insufficient high quality temporal observations) in large portions of the southern region in both years. In contrast, the VIIRS PGQsos was relatively high in the southern region and the poor VIIRS PGQsos retrievals were randomly distributed across large parts of the region. The difference between OLI PGQsos and VIIRS PGQsos was associated with the time lag of the satellite observations. OLI time series were fused using observations around 10:30am from Terra MODIS and Landsat 8, while VIIRS observations were obtained around 1:30pm. This time lag could have significantly impacted the level of cloud contamination, which was particularly evident in the southern region of the study area.

Comparisons between VIIRS SOS dates and OLI SOS dates across a wide range of heterogeneities improved our understanding of the scaling effect on land surface phenology (LSP) at coarse resolutions. This step is critical in evaluating LSP quality and bridging LSP across scales.

VIIRS SOS could be well represented using optimal OLI SOS values within a VIIRS pixel or grid. In homogeneous regions, OLI SOS values in more than 60% of pixels were equivalent to VIIRS SOS. However, about 5% of earliest OLI SOS and 20% of the latest OLI SOS within a VIIRS pixel relatively deviated from the VIIRS SOS dates. This level of deviation is reasonable because the VIIRS pixels or grids were more or less mixed covers with several vegetation types and completely homogeneous VIIRS pixels were rare. This result suggests that plot-based, *in-situ* observations in homogeneous regions can be generally effective for the validation of LSP (Roman et al. 2011). Unsurprisingly, the most homogeneous SOS is likely to be observed within a single crop field because of similarity of agronomic management practices. In comparison, SOS in a "homogeneous" forest area could still vary considerably due to forest species distribution and microclimate resulting in SOS dates that are larger than 10 days (Fisher et al, 2006; Richardson and O'Keefe, 2009; Liang et al., 2011). In contrast, the proportion of OLI pixels with SOS dates similar to VIIRS SOS dates greatly decreased with increasing heterogeneity. Within a heterogeneous VIIRS pixel containing various plant species, the range of OLI SOS could be as large as three months. In these situations, the OLI SOS values in less than 20% of pixels were comparable to VIIRS SOS dates.

Comparing VIIRS SOS with OLI SOSag further revealed that the AAD was smallest and bias approached zero, if the OLI SOSag data were aggregated by selecting the date when SOS transition had occurred in about 30% of OLI pixels. This result was consistent for individual VIIRS pixels and for the VIIRS grids (3×3 pixels) in both 2013 and 2014. This finding is also consistent with other remote sensing studies that have found MODIS SOS dates corresponds to the timing when budburst has occurred in 20%-33% of individual stems monitored from the ground in the Harvard Forest (Zhang et al., 2006; Ganguly et al., 2010). Thus, we can conclude that the SOS detected from satellite data represents the timing at which vegetation greenup onset occurred in 30% of area in an individual pixel despite the heterogeneity in SOS dates. The finding also supports our hypothesis that the SOS at a coarser resolution becomes detectable when vegetation starts to greenup in a certain proportion of finer resolution pixels, and that the coarser resolution SOS is associated with the earlier SOS pixels at the finer resolution rather than the later SOS pixels.

Understanding the scaling effect on LSP helps the process of validating coarser resolution SOS using finer resolution observations. Validation of satellite-based products is an important and challenging task in remote sensing; however, one of the main difficulties is how to scale plot level measurements up to the coarser resolution of spaceborne sensors (Buermann et al. 2002; de Beurs et al. 2009; Herold et al. 2008; Weiss et al. 2007). Coarser resolution LSP is commonly validated using the simple average of finer resolution data (Delbart et al. 2015; Roman et al. 2011); however, this study suggests that selecting the timing of the 30th percentile at the finer resolution is biophysically meaningful, particularly in very heterogeneous areas. Based on this criterion, we have demonstrated that the VIIRS SOS was well detected because its overall difference with the OLI SOSag was less than 5 days in homogeneous regions, although the difference was larger in heterogeneous regions.

Finally, it should be noted that the result of coarser resolution SOS equivalent to finer resolution value at 30th percentile was derived from OLI (30m) and VIIRS (~450m) and verified by comparing SOS aggregated in 3by3 VIIRS pixels with OLI SOS. Further studies are needed to explore how the SOS scales across various landscapes and ecosystems. To investigate SOS variations across scales is challenging, because it requires multiple sets of SOS data across spatial scales. Moreover, these multiple sets should be derived from vegetation index time series at various spatial resolutions independently rather than simply aggregated from the same finer resolution SOS dataset using aggregation approaches such as averaging, thinning, or majority filtering. Finally, further research is needed to verify if the 30th percentile is always the optimal percentile for the aggregation of LSP SOS values that are extracted by various methods.

ACKNOWLEDGEMENTS

This work was supported in part by NASA contracts NNX15AB96A, NNX14AJ32G, and NNX14AQ18A and by NOAA contract JPSS_PGRR2_14.

REFERENCES

Augspurger, C.K., Cheeseman, J.M., & Salk, C.F. (2005). Light gains and physiological capacity of understorey woody plants during phenological avoidance of canopy shade. *Functional Ecology*, *19*, 537-546

Bradley, B.A., Jacob, R.W., Hermance, J.F., & Mustard, J.F. (2007). A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. *Remote Sensing of Environment, 106*, 137-145

Buermann, W., Wang, Y.J., Dong, J.R., Zhou, L.M., Zeng, X.B., Dickinson, R.E., Potter, C.S., & Myneni, R.B. (2002). Analysis of a multiyear global vegetation leaf area index data set. *Journal of Geophysical Research-Atmospheres*, *107*

Campagnolo, M.L., Sun, Q.S., Liu, Y., Schaaf, C., Wang, Z.S., & Roman, M.O. (2016). Estimating the effective spatial resolution of the operational BRDF, albedo, and nadir reflectance products from MODIS and VIIRS. *Remote Sensing of Environment*, *175*, 52-64

Chen, J., Jonsson, P., Tamura, M., Gu, Z.H., Matsushita, B., & Eklundh, L. (2004). A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sensing of Environment*, *91*, 332-344

de Beurs, K.M., & Henebry, G.M. (2004). Land surface phenology, climatic variation, and institutional change: Analyzing agricultural land cover change in Kazakhstan. *Remote Sensing of Environment*, *89*, 497-509

de Beurs, K.M., Wright, C.K., & Henebry, G.M. (2009). Dual scale trend analysis for evaluating climatic and anthropogenic effects on the vegetated land surface in Russia and Kazakhstan. *Environmental Research Letters*, *4*

de Beurs, K.M., & Henebry, G.M. (2010). Spatio-temporal statistical methods for modeling land surface phenology. In: (I.L. Hudson & M.R. Keatley, eds.) *Phenological Research: Methods for Environmental and Climate Change Analysis*: Springer: New York. pp. 177-208.Delbart, N., Beaubien, E., Kergoat, L., & Toan, T.L. (2015). Comparing land surface phenology with leafing and flowering observations from the PlantWatch citizen network. *Remote Sensing of Environment*, *160*, 273-280

Elmore, A.J., Guinn, S.M., Minsley, B.J., & Richardson, A.D. (2012). Landscape controls on the timing of spring, autumn, and growing season length in mid-Atlantic forests. *Global Change Biology*, *18*, 656-674

Fisher, J.I., Mustard, J.F., & Vadeboncoeur, M.A. (2006). Green leaf phenology at Landsat resolution: Scaling from the field to the satellite. *Remote Sensing of Environment, 100*, 265-279

Ganguly, S., Friedl, M.A., Tan, B., Zhang, X.Y., & Verma, M. (2010). Land surface phenology from MODIS: Characterization of the Collection 5 global land cover dynamics product. *Remote Sensing of Environment*, *114*, 1805-1816

Gao, F., Anderson, M.C., Zhang, X., Yang, Z., Alfieri, J.G., Kustas, W.P., Mueller, R., Johnson, D.M., Prueger, J.H., 2017, Toward mapping crop progress at field scales using Landsat and MODIS imagery, *Remote Sensing of Environment*, *188, 9–25*

Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing, 44*, 2207-2218

Graham, E.A., Riordan, E.C., Yuen, E.M., Estrin, D., & Rundel, P.W. (2010). Public Internet-connected cameras used as a cross-continental ground-based plant phenology monitoring system. *Global Change Biology, 16*, 3014-3023

Henebry, G.M., & de Beurs, K.M. (2013). Remote Sensing of Land Surface Phenology: A Prospectus. In M.D. Schwartz (Ed.), *Phenology: An Integrative Environmental Science* (pp. 385-411): Springer

Herold, M., Woodcock, C.E., Loveland, T.R., Townshend, J., Brady, M., Steenmans, C., & Schmullius, C.C. (2008). Land-Cover Observations as Part of a Global Earth Observation System of Systems (GEOSS): Progress, Activities, and Prospects. *IEEE Systems Journal*, *2*, 414-423

Hilker, T., Wulder, M.A., Coops, N.C., Linke, J., McDermid, G., Masek, J.G., Gao, F., & White, J.C. (2009). A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sensing of Environment, 113*, 1613-1627

Homer, C., Dewitz, J., Yang, L.M., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N., Wickham, J., & Megown, K. (2015). Completion of the 2011 National Land Cover Database for the Conterminous United States - Representing a Decade of Land Cover Change Information. *Photogrammetric Engineering and Remote Sensing*, *81*, 345-354

Hufkens, K., Friedl, M., Sonnentag, O., Braswell, B.H., Milliman, T., & Richardson, A.D. (2012). Linking near-surface and satellite remote sensing measurements of deciduous broadleaf forest phenology. *Remote Sensing of Environment*, *117*, 307-321

Jiang, Z.Y., Huete, A.R., Didan, K., & Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, *112*, 3833-3845

Jonsson, P., & Eklundh, L. (2002). Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing, 40*, 1824-1832

Julien, Y., & Sobrino, J.A. (2010). Comparison of cloud-reconstruction methods for time series of composite NDVI data. *Remote Sensing of Environment, 114*, 618-625

Justice, C.O., Roman, M.O., Csiszar, I., Vermote, E.F., Wolfe, R.E., Hook, S.J., Friedl, M., Wang, Z.S., Schaaf, C.B., Miura, T., Tschudi, M., Riggs, G., Hall, D.K., Lyapustin, A.I., Devadiga, S., Davidson, C., & Masuoka, E.J. (2013). Land and cryosphere products from Suomi NPP VIIRS: Overview and status. *Journal of Geophysical Research-Atmospheres, 118*, 9753-9765

Keenan, T.F., Darby, B., Felts, E., Sonnentag, O., Friedl, M.A., Hufkens, K., O'Keefe, J., Klosterman, S., Munger, J.W., Toomey, M., & Richardson, A.D. (2014). Tracking forest phenology and seasonal physiology using digital repeat photography: a critical assessment. *Ecological Applications, 24*, 1478-1489

Krehbiel, C.P., Jackson, T., & Henebry, G.M. (2015). Web-Enabled Landsat Data time series for monitoring urban heat island impacts on land surface phenology. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*

Lechowicz, M.J. (1984). Why do temperate deciduous trees leaf out at different times –adaptations and ecology of forest communities. *American Naturalist*, *124*, 821-842

Liang, L., & Schwartz, M.D. (2009). Landscape phenology: an integrative approach to seasonal vegetation dynamics. *Landscape Ecology*, *24*, 465-472

Liang, L.A., Schwartz, M.D., & Fei, S.L. (2011). Validating satellite phenology through intensive ground observation and landscape scaling in a mixed seasonal forest. *Remote Sensing of Environment, 115*, 143-157

Lloyd, D. (1990). A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. *International Journal of Remote Sensing, 11*, 2269-2279

Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., Gao, F., Kutler, J., & Lim, T.K. (2006). A Landsat surface reflectance dataset for North America, 1990-2000. *IEEE Geoscience and Remote Sensing Letters, 3*, 68-72

Melaas, E.K., Friedl, M.A., & Zhu, Z. (2013). Detecting interannual variation in deciduous broadleaf forest phenology using Landsat TM/ETM plus data. *Remote Sensing of Environment, 132*, 176-185

Moody, A., & Johnson, D.M. (2001). Land-surface phenologies from AVHRR using the discrete fourier transform. *Remote Sensing of Environment*, *75*, 305-323

Reed, B.C., Brown, J.F., VanderZee, D., Loveland, T.R., Merchant, J.W., & Ohlen, D.O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, *5*, 703-714

Richardson, A.D., Braswell, B.H., Hollinger, D.Y., Jenkins, J.P., & Ollinger, S.V. (2009). Near-surface remote sensing of spatial and temporal variation in canopy phenology. *Ecological Applications, 19*, 1417-1428

Richardson, A.D., Jenkins, J.P., Braswell, B.H., Hollinger, D.Y., Ollinger, S.V., & Smith, M.L. (2007). Use of digital webcam images to track spring green-up in a deciduous broadleaf forest. *Oecologia, 152*, 323-334

Richardson, A.D., & O'Keefe, J. (2009). Phenological Differences Between Understory and Overstory: A Case Study Using the Long-Term Harvard Forest Records. In A. Noormets (Ed.), *Phenology of Ecosystem Processes*: Springer Science

Rocha, A.V., Potts, D.L., & Goulden, M.L. (2008). Standing litter as a driver of interannual CO(2) exchange variability in a freshwater marsh. *Journal of Geophysical Research-Biogeosciences, 113*

Rocha, A.V., & Shaver, G.R. (2009). Advantages of a two band EVI calculated from solar and photosynthetically active radiation fluxes. *Agricultural and Forest Meteorology, 149*, 1560-1563

Roman, M.O., Gatebe, C.K., Schaaf, C.B., Poudyal, R., Wang, Z.S., & King, M.D. (2011). Variability in surface BRDF at different spatial scales (30 m-500 m) over a mixed agricultural landscape as retrieved from airborne and satellite spectral measurements. *Remote Sensing of Environment, 115*, 2184-2203

Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X.W., Tsang, T., Strugnell, N.C., Zhang, X.Y., Jin, Y.F., Muller, J.P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R.P., Hu, B.X., Liang, S.L., Privette, J.L., & Roy, D. (2002). First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment*, *83*, 135-148

Schaber, J., & Badeck, F.W. (2003). Physiology-based phenology models for forest tree species in Germany. *International Journal of Biometeorology, 47*, 193-201

Schwartz, M.D., & Hanes, J.M. (2010). Intercomparing multiple measures of the onset of spring in eastern North America. *International Journal of Climatology, 30*, 1614-1626

Schwartz, M.D., & Reed, B.C. (1999). Surface phenology and satellite sensor-derived onset of greenness: an initial comparison. *International Journal of Remote Sensing*, *20*, 3451-3457

Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A.M., Friedl, M., Braswell, B.H., Milliman, T., O'Keefe, J., & Richardson, A.D. (2012). Digital repeat photography for phenological research in forest ecosystems. *Agricultural and Forest Meteorology, 152*, 159-177

Soudani, K., le Maire, G., Dufrene, E., Francois, C., Delpierre, N., Ulrich, E., & Cecchini, S. (2008). Evaluation of the onset of green-up in temperate deciduous broadleaf forests derived from Moderate Resolution Imaging Spectroradiometer (MODIS) data. *Remote Sensing of Environment, 112*, 2643-2655

Tan, B., Morisette, J.T., Wolfe, R.E., Gao, F., Ederer, G.A., Nightingale, J., & Pedelty, J.A. (2011). An Enhanced TIMESAT Algorithm for Estimating Vegetation Phenology Metrics From MODIS Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 4*, 361-371

Vermote, E.F., El Saleous, N.Z., & Justice, C.O. (2002). Atmospheric correction of MODIS data in the visible to middle infrared: first results. *Remote Sensing of Environment, 83,* 97-111

Viovy, N., Arino, O., & Belward, A.S. (1992). The best index slope extraction (BISE): A method for reducing noise in NDVI time series. *International Journal of Remote Sensing, 13*, 1585-1590

Wagenseil, H., & Samimi, C. (2006). Assessing spatio-temporal variations in plant phenology using Fourier analysis on NDVI time series: results from a dry savannah environment in Namibia. *International Journal of Remote Sensing, 27*, 3455-3471

Walker, J.J., de Beurs, K.M., Wynne, R.H., & Gao, F. (2012). Evaluation of Landsat and MODIS data fusion products for analysis of dryland forest phenology. *Remote Sensing of Environment, 117*, 381-393

Weiss, M., Baret, F., Garrigues, S., & Lacaze, R. (2007). LAI and fAPAR CYCLOPES global products derived from VEGETATION. Part 2: validation and comparison with MODIS collection 4 products. *Remote Sensing of Environment, 110*, 317-331

White, M.A., de Beurs, K.M., Didan, K., Inouye, D.W., Richardson, A.D., Jensen, O.P., O'Keefe, J., Zhang, G., Nemani, R.R., van Leeuwen, W.J.D., Brown, J.F., de Wit, A., Schaepman, M., Lin, X.M., Dettinger, M., Bailey, A.S., Kimball, J., Schwartz, M.D., Baldocchi, D.D., Lee, J.T., & Lauenroth, W.K. (2009). Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982-2006. *Global Change Biology, 15*, 2335-2359

White, M.A., Thornton, P.E., & Running, S.W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles, 11*, 217-234

Yan, L., & Roy, D.P. (2016). Conterminous United States crop field size quantification from multi-temporal Landsat data. *Remote Sensing of Environment, 172*, 67-86

Zhang, X., Friedl, M., & Schaaf, C. (2009). Sensitivity of vegetation phenology detection to the temporal resolution of satellite data. *International Journal of Remote Sensing, 30*, 2061-2074

Zhang, X.Y. (2015). Reconstruction of a complete global time series of daily vegetation index trajectory from long-term AVHRR data. *Remote Sensing of Environment*, *156*, 457-472

Zhang, X.Y., Friedl, M.A., & Schaaf, C.B. (2006). Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements. *Journal of Geophysical Research-Biogeosciences, 111*

Zhang, X.Y., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., Reed, B.C., & Huete, A. (2003). Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment, 84*, 471-475

Zhang, X.Y., Tan, B., & Yu, Y.Y. (2014). Interannual variations and trends in global land surface phenology derived from enhanced vegetation index during 1982-2010. *International Journal of Biometeorology, 58*, 547-564

Zhu, X.L., Chen, J., Gao, F., Chen, X.H., & Masek, J.G. (2010). An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of Environment*, *114*, 2610-2623