Agricultural and Forest Meteorology 173 (2013) 74–84

Contents lists available at SciVerse ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics

Douglas K. Bolton*, Mark A. Friedl

Department of Earth and Environment, Boston University, 675 Commonwealth Avenue, Boston, MA 02215, United States

ARTICLE INFO

Article history:
Received 14 September 2012
Received in revised form
30 December 2012
Accepted 22 January 2013

Keywords:
Remote sensing
Crop yield forecasting
Crop phenology
MODIS
Central United States

ABSTRACT

We used data from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) in association with county-level data from the United States Department of Agriculture (USDA) to develop empirical models predicting maize and soybean yield in the Central United States. As part of our analysis we also tested the ability of MODIS to capture inter-annual variability in yields. Our results show that the MODIS two-band Enhanced Vegetation Index (EVI2) provides a better basis for predicting maize yields relative to the widely used Normalized Difference Vegetation Index (NDVI). Inclusion of information related to crop phenology derived from MODIS significantly improved model performance within and across years. Surprisingly, using moderate spatial resolution data from the MODIS Land Cover Type product to identify agricultural areas did not degrade model results relative to using higher-spatial resolution crop-type maps developed by the USDA. Correlations between vegetation indices and yield were highest 65–75 days after greenup for maize and 80 days after greenup for soybeans. EVI2 was the best index for predicting maize yield in non-semi-arid counties ($R^2 = 0.67$), but the Normalized Difference Water Index (NDWI) performed better in semi-arid counties ($R^2 = 0.69$), probably because the NDWI is sensitive to irrigation in semi-arid areas with low-density agriculture. NDVI and EVI2 performed equally well predicting soybean yield ($R^2 = 0.69$ and 0.70, respectively). In addition, EVI2 was best able to capture large negative anomalies in maize yield in 2005 ($R^2 = 0.73$). Overall, our results show that using crop phenology and a combination of EVI2 and NDWI have significant benefit for remote sensing-based maize and soybean yield models.

1. Introduction

Remote sensing has been used for agricultural applications for over three decades (Knipling, 1970; Tucker et al., 1981; Moran et al., 1994; Wardlow and Egbert, 2008). In particular, remotely sensed vegetation indices (VIs) such as the Normalized Difference Vegetation Index (NDVI) have been widely utilized for agricultural mapping and monitoring (Maselli et al., 1992; Rasmussen, 1992; Benedetti and Rossini, 1993; Funk and Budde, 2009; Mkhabela et al., 2011). Data from the National Oceanic and Atmospheric Administration's (NOAA) Advanced Very High Resolution Radiometer (AVHRR) have been used since the early 1980's for large scale crop monitoring and yield forecasting (Tucker et al., 1985; Malingreau, 1986). In recent years, the focus of remote sensing-based yield forecasting research has shifted to the National Aeronautics and Space Administration's (NASA) Moderate Resolution Imaging Spectroradiometer (MODIS), which provides improved radiometric calibration and superior spectral and spatial resolution relative to the AVHRR (Doraiswamy et al., 2004, 2005; Ren et al., 2008; Funk and Budde, 2009; Becker-Reshef et al., 2010).

Both AVHRR and MODIS provide high frequency observations. However, their spatial resolution is quite coarse. MODIS data are available at 250-m, 500-m, and 1000-m spatial resolution depending on the specific product (Justice et al., 1998). AVHRR data, on the other hand, are available at spatial resolutions of 1.1 km for local area coverage and 4 km globally (Kidwell, 1998). Higher spatial resolution data from sensors such as the Landsat Thematic Mapper (30 m) have also been used in agricultural applications (Rudorff and Batista, 1991; Thenkabail et al., 1994). However, the repeat period for Landsat is relatively infrequent (16 days), which presents challenges for agricultural applications that rely on high frequency sampling during critical phases of the crop growth cycle. Recent studies have fused higher spatial resolution map products derived from Landsat with higher frequency MODIS or AVHRR data (e.g., Genovese et al., 2001; Becker-Reshef et al., 2010; Mkhabela et al., 2011). However, medium resolution map products derived from Landsat data such as the Cropland Data Layer (CDL; Boryan et al., 2011), which is produced by the United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS), are

* Corresponding author. Current address: Faculty of Forestry, University of British Columbia, Forest Sciences Centre, 2424 Main Mall, Vancouver, BC, Canada V6 T 1Z4. Tel.: +1 604 822 6592.
E-mail addresses: doug.k.bolton@alumni.ubc.ca (D.K. Bolton), friedl@bu.edu (M.A. Friedl).

0168-1923/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.agrformet.2013.01.007

generally only produced at regional scales and are therefore not available for much of the world. It is therefore important to explore and assess the utility of methods that rely on widely available datasets such as MODIS.

In this paper we explore methods for forecasting county-level maize and soybean yield in the United States using MODIS, ecoregion, and crop yield data for 2004 through 2009. Our analysis focuses on agricultural states in the Central United States because high quality yield data are available for this region. However, the more general motivation is to explore methods that could ultimately be used in other parts of the world, especially in developing countries where data on crop yields are less available and food security is an important societal issue. Within this framework, our analysis examines three main questions:

1. *What spectral information from MODIS provides the best basis for predicting crop yields?* In addition to the NDVI we examine two spectral indices that have not been previously explored in detail for crop yield forecasting applications: (1) a two-band variant of the Enhanced Vegetation Index (EVI2; Jiang et al., 2008), and (2) the Normalized Difference Water Index (NDWI; Gao, 1996). The potential advantages of these indices are that relative to the NDVI, EVI does not saturate as quickly at higher crop leaf area (Huete et al., 2002), and the NDWI is designed to be sensitive to water stress in plants and may therefore better capture the effect of drought on crop yields (Gu et al., 2007, 2008).
2. *What ancillary information is required to distinguish agriculture from non-agricultural land use and land cover?* A critical step in forecasting crop yields using remote sensing is to identify agricultural areas and minimize contamination from natural vegetation (Maselli et al., 2000; Maselli and Rembold, 2001; Funk and Budde, 2009). Thus, understanding how predicted crop yields depend on the resolution and quality of underlying crop extent maps is critical for assessing trades-offs between model accuracy and model transferability.
3. *Can remotely sensed information related to crop phenology be used to improve predictions of yield?* Several previous studies have examined how the relationship between remotely sensed data and crop yields vary as a function of time during the growing season (Wall et al., 2008; Mkhabela et al., 2011). Results from these studies have demonstrated that the correlation between VIs and yield varies through the crop cycle (Basnyat et al., 2004; Salazar et al., 2007; Ren et al., 2008; Wall et al., 2008; Mkhabela et al., 2011). However, crop planting dates and phenology vary by location and change from one year to the next. Thus, using a fixed calendar date to estimate remote sensing-based yield prediction models (e.g., Basnyat et al., 2004; Mkhabela et al., 2011) is not optimal. Several studies have accounted for variations in planting dates and phenology using the onset of the rainy season (Rojas, 2007; Funk and Budde, 2009) or the seasonal NDVI curve derived from AVHRR (Reynolds et al., 2000). Here we explore an approach that uses information related to crop phenology derived from MODIS to improve remotely sensed estimates of both intra- and inter-annual variability in crop yields.

Our analysis includes four main elements. First, we compare the effectiveness of different MODIS spectral indices for predicting maize and soybean yields. Second, we investigate the value added information provided by remotely sensed crop phenology metrics for yield prediction. Third, we compare and assess the utility of the NASS CDL and the MODIS Land Cover Type product for isolating the signature of crop yields in MODIS data. Fourth, we use our crop yield models to assess how well remote-sensing based approaches are able to capture county-level geographic and inter-annual variability in crop yields in the Central United States.

2. Data and methods

2.1. Study region

Our study region focuses on agricultural counties in the Central United States with large areas of maize and soybean, and includes data collected from 2004 through 2009 (Fig. 1). The NASS CDL, a collection of state-level crop type classifications, was used to identify these counties. The CDL is produced annually at 30- or 56-m spatial resolution using reference data collected through cropland census and imagery from the Landsat Thematic Mapper and the Advanced Wide Field Sensor (AWIFS) (Boryan et al., 2011; http://www.nass.usda.gov/research/Cropland/SARS1a.htm). Using these data, we estimated the total acreage for maize and soybeans in each county. Counties with more than 10,000 ha of maize or soybeans in each of the six study years were selected for inclusion in our analysis. The final dataset included 333 counties for maize and 325 counties for soybeans across ten states (Fig. 1). To account for first-order geographic variation in climate, topography and edaphic conditions over the study region, we used the Level II ecoregions defined by the Commission for Environmental Cooperation (1997) (available at www.epa.gov/wed/pages/ecoregions.htm).

2.2. Crop yield and MODIS data

Crop yield data, in units of bushels acre^{-1}, were obtained from datasets compiled by the NASS based on surveys collected from a random sample of farms (available at http://quickstats.nass.usda.gov/). Surveys are conducted monthly by NASS to provide state-level data on acreage planted, acreage harvested and yield for a variety of crops. At the end of each year, NASS derives estimates of crop production at the county-level from these survey data. We extracted annual maize and soybean yields at the county-level from these data for the counties shown in Fig. 1 from 2004 to 2009. Yield estimates were converted from bushels acre^{-1} to kg ha^{-1} using published conversion tables (Murphy, 1993).

NASA provides MODIS data for a suite of land products representing different land surface properties at several different spatial and temporal resolutions. MODIS "tiles" used in our analysis cover the major maize and soybean growing states in the Central United States (MODIS tile numbers h10v04, h10v05, h11v04 and h11v05). For each of these tiles, spectral indices were calculated using MODIS collection 5 nadir bidirectional reflectance distribution function adjusted surface reflectance (NBAR) data, which are available at 500-m spatial resolution and eight-day intervals (Schaaf et al., 2002). Using these data we computed three dimensionless VIs: the NDVI (Tucker, 1979), the EVI2 (Jiang et al., 2008) and the NDWI (Gao, 1996). The NDVI and EVI2 are designed to capture spatio-temporal variation in photosynthetically active vegetation (Tucker, 1979; Jiang et al., 2008), and the NDWI is designed to capture spatio-temporal patterns in surface moisture status (Gao, 1996). We used cubic spline interpolation (de Boor, 1978) to derive daily values for each index based on the eight-day NBAR data. All daily values falling within a 16-day window of missing NBAR data were marked as missing.

We tested two different data sources to distinguish pixels dominated by cropland from those dominated by natural vegetation in the MODIS NBAR data: the NASS CDL data set and the MODIS Land Cover Type product. The state-level CDL classifications were converted into two binary maps for each study year: one identifying maize pixels and the other identifying soybean pixels. These maps were reprojected into the MODIS sinusoidal grid, and each 500-m MODIS pixel was assigned the percentage of the pixel classified as either maize or soybean based on the CDL maps in each year.

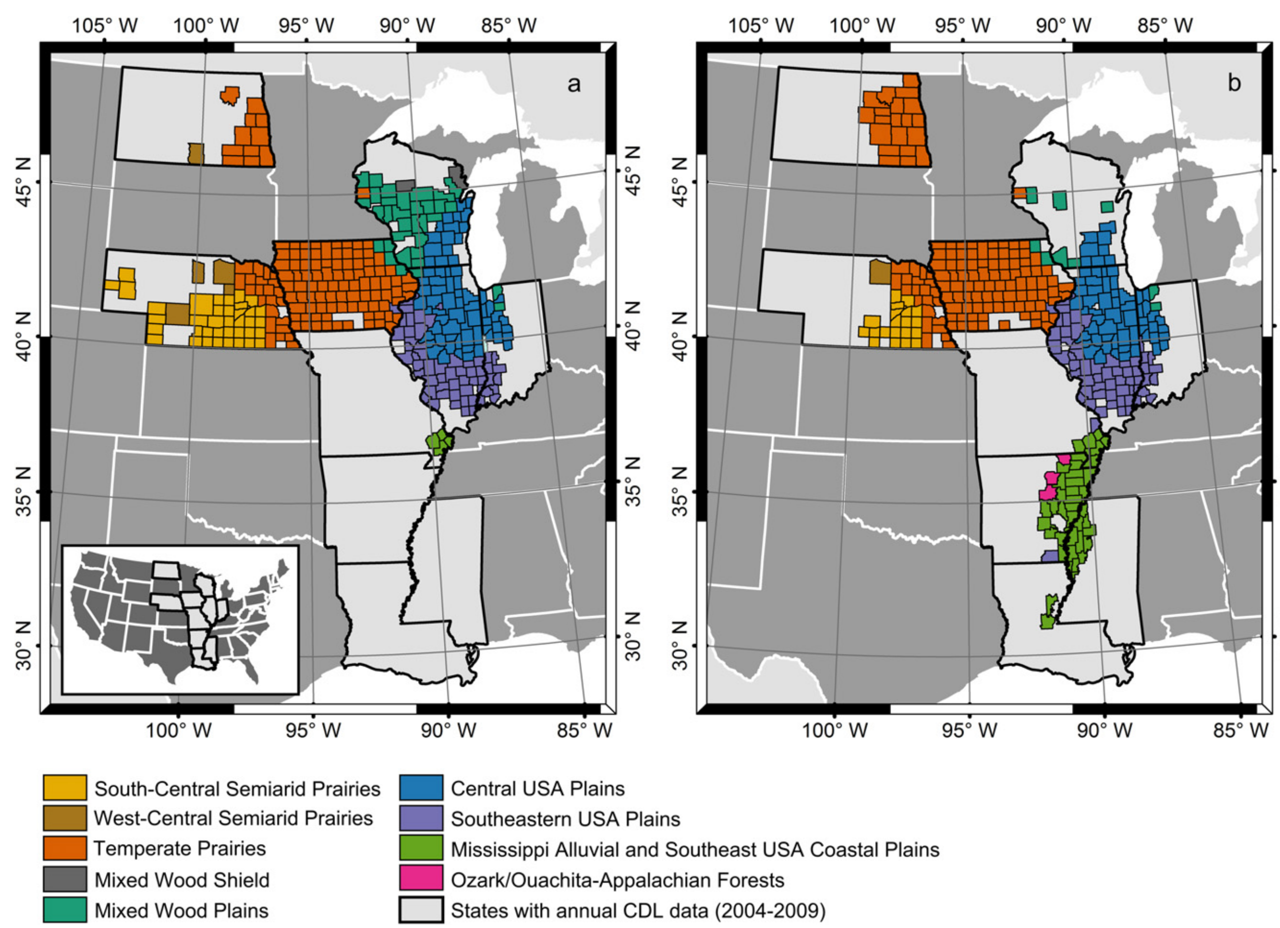

Fig. 1. Selected counties for (a) maize and (b) soybeans, shaded by the Level II ecoregions defined by the Commission for Environmental Cooperation (1997).

The second source of cropland extent that we tested was produced from MODIS data. Specifically, the MODIS Land Cover Type Product provides global maps of land cover at 500-m spatial resolution on an annual basis using the International Geosphere-Biosphere Programme (IGBP) classification system (Friedl et al., 2010). This system includes 17 land cover types including agriculture and agricultural mosaic classes, which are defined as pixels with >60% or 40–60% agriculture by area, respectively. Unlike the CDL, the MODIS Land Cover Type Product does not differentiate crop types. We included only pixels classified as agriculture in our analysis. Because agriculture is both intensive and extensive in the Central United States, exclusion of agricultural mosaic pixels should not affect our results. However, it is worth noting that the mosaic class is likely to be important in other parts of the world, especially in developing countries where low intensity agriculture is common.

Data from the MODIS Collection 5 Land Cover Dynamics product provided information related to the phenology of crops (Ganguly et al., 2010). This product is produced globally at 500-m spatial resolution using eight-day EVI values computed from NBAR data, and provides estimates for the day-of-year (DOY) corresponding to key phenophase transition dates at each pixel. To estimate these transition dates the algorithm used to generate this product fits piecewise logistic functions to EVI time-series at each pixel, and uses maxima in the rate of change in curvature to identify the DOY corresponding to four cardinal transition dates in each growth cycle: (1) the onset of EVI increase (i.e. spring onset or greenup), (2) the onset of EVI maximum (maximum canopy development), (3) the onset of EVI decrease (senescence), and (4) the onset of EVI minimum (dormancy). Complete details are provided in Zhang et al. (2003, 2006). For this work, we used the date of onset of EVI increase at each pixel to identify the start of the photosynthetically active growing season.

2.3. Yield modeling approach

Our modeling approach is similar to the methods described by Funk and Budde (2009) and Mkhabela et al. (2011), but includes several important refinements. The core of our strategy uses linear regression models to predict yield based on remotely sensed indices obtained during the crop growing period. Previous studies have used the NDVI for this purpose. Here we test the EVI2 and the NDWI, in addition to the NDVI. More importantly, the key innovation of our approach is that we use MODIS derived information related to the timing of crop phenology at each pixel to optimize the period used to estimate the model and predict yield, while Funk and Budde (2009) used coarse resolution information on the onset of rains and Mkhabela et al. (2011) used no adjustment. Our general approach includes three main steps (Fig. 2):

(i) *Calculate "phenologically adjusted" MODIS composites:* Using the MODIS Land Cover Dynamics product, we identify the date of onset of EVI increase at each pixel (nominally, the date at which crops emerge, are detectable from MODIS, and begin to grow). Hereafter we refer to this as the "greenup date." Using this information, we extract time series values of EVI2, NDVI, and NDWI interpolated to five-day intervals, stopping 120 days after the greenup date (Fig. 3). Thus, at each pixel we create a time series consisting of 25 "phenologically adjusted" values

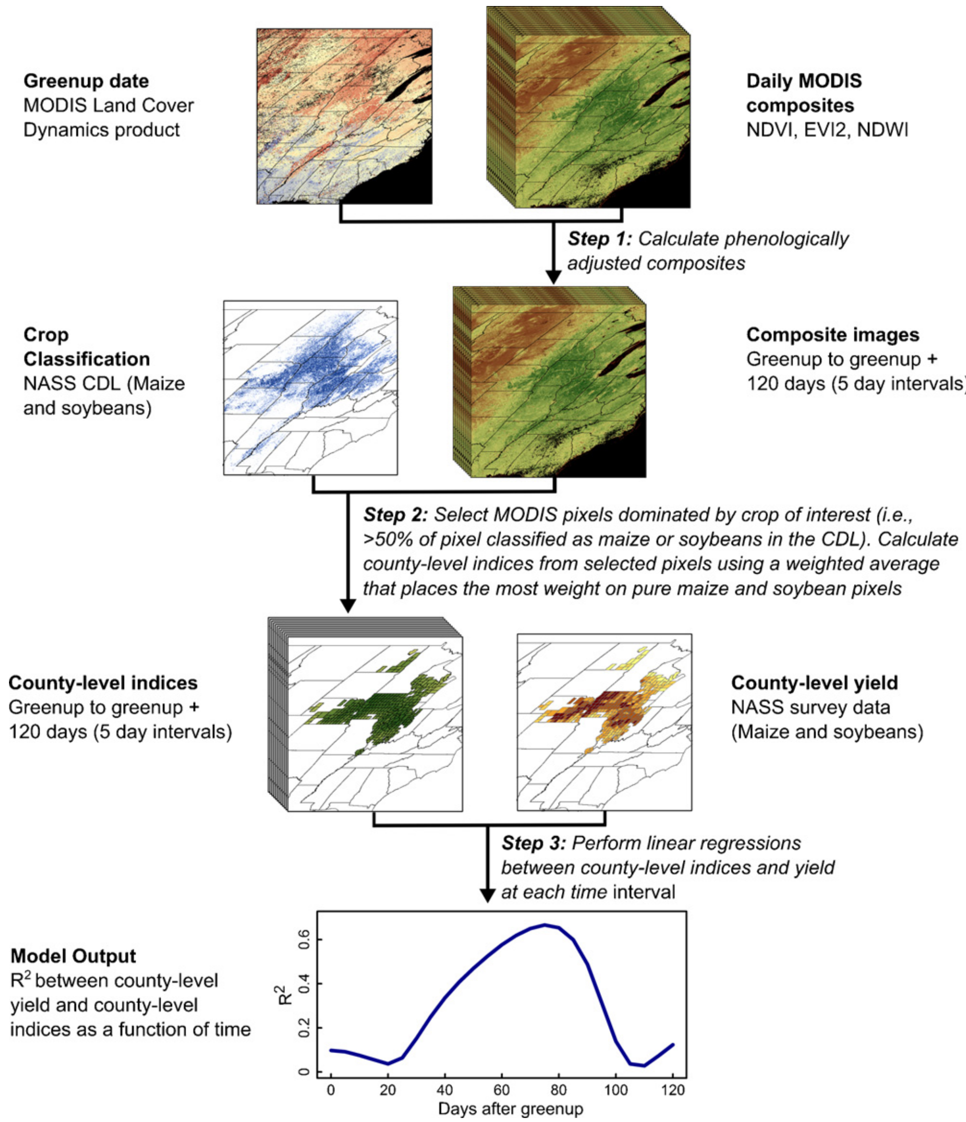

Fig. 2. A flow diagram of the datasets and processing steps used in the base model.

for each spectral index, where the start date for the time series at each pixel varies depending on the timing of greenup.

(ii) *Compute county-average spectral indices.* To estimate our models using county-level yield data, it was necessary to aggregate MODIS data to the same scale. To do this, we selected all MODIS pixels that were dominated by the crop of interest (i.e. had a sub-pixel proportion of >50% maize or soybeans according to the CDL), and computed county-level weighted averages for each index, where the contribution of each MODIS pixel was weighted by the proportion of the pixel classified as maize or soybean in the CDL. This technique eliminates highly mixed pixels while placing the most weight on pure maize and soybean pixels. Pixels with greenup dates before DOY 100 or after DOY 200 were assumed to either reflect errors in the MODIS Land Cover Dynamics product, or included significant proportions of crops with phase-shifted phenology (e.g., winter wheat). These pixels were therefore removed from the analysis.

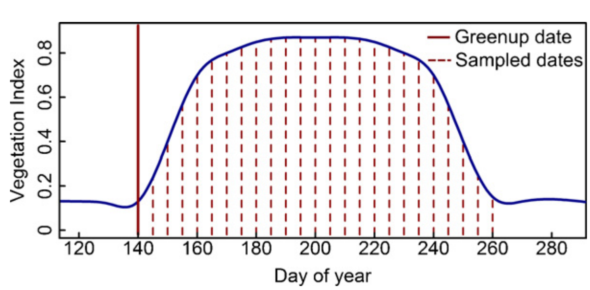

Fig. 3. A schematic showing how "phenologically adjusted" values were extracted from daily VI values. Values were extracted at five-day intervals from the greenup date to 120 days following the greenup date for each pixel.

(iii) *Estimate linear models between county-level indices and yield.* Using the phenologically adjusted and county-averaged spectral indices, we estimated linear regression models in R (R Development Core Team, 2009) to predict county-level maize and soybean yields across all of the counties in our study:

$$\hat{Y} = \beta_0 + \beta_1 VI_t$$

where $\hat{Y}$ is the predicted county-level yield, and VI_t is the crop-weighted county average of the MODIS VI at time step t after greenup. To determine when the correlation between VIs and crop yield is strongest, we estimated coefficients of determination (R^2) using data from each five-day time step between 0 and 120 days after greenup. We then selected the time step that produced the highest coefficients of determination to develop linear regression models based on all six years of data. In addition, we calculated cross-validated coefficients of determination by training the models on data from 2004 to 2006 and testing them on data from 2007 to 2009.

As a final test, we also developed regression models using a leave-one-year-out-approach (Schut et al., 2009; Mkhabela et al., 2011), in which the model is iteratively trained on five years of data and used to predict yield in the held-out year. The root mean square error (RMSE) was calculated between actual and predicted yield for the held-out year as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{Y}_i - Y_i\right)^2}$$

where n is the number of counties, $\hat{Y}_i$ is the predicted county-level yield and Y_i is the actual county-level yield. Because aridity has been shown to affect the relationship between VIs and crop yield (Mkhabela et al., 2011), we grouped the counties as semi-arid or non-semi-arid according to their ecoregion, and re-estimated the model for each group independently.

2.4. Additional analyses

We performed four additional analyses designed to improve understanding of results from the regression models described above. First, to assess the degree to which model results depend on information related to the location and extent of agriculture and crop types within counties, we re-estimated the models described above using the MODIS Land Cover Type product to identify pixels classified as agriculture (i.e., we used land cover from MODIS in place of the CDL). We did this to assess the importance of using high spatial resolution maps to delineate agricultural land use, which are not available for many parts of the world. However, because the MODIS Land Cover Type product does not provide information related to crop types, it was not possible to calculate weighted county averages for individual crop types following the procedure we used based on the CDL. The revised models were therefore based on average spectral indices for all agricultural pixels within each county, independent of crop type.

Second, to assess the importance of using “phenologically adjusted” spectral indices, we re-estimated the models using the same approach that we describe above, but using a fixed date for the start of season for all pixels in the study region. To do this, we developed models using MODIS data from DOY 140 to 260 at five-day intervals. Thus, the models developed through this exercise were based on spectral indices for a fixed DOY, with no allowance for variation in crop phenology.

Third, we evaluated the ability of phenologically adjusted NDVI, EVI2 and NDWI data to capture inter-annual variation in maize and soybean yields. To evaluate this, anomalies in both county-level crop yields and phenologically adjusted spectral indices were calculated as the deviation in annual values from six-year county averages for each quantity. The sensitivity of the phenologically adjusted indices to inter-annual variation in yield was evaluated by examining the covariance between inter-annual anomalies in yield and inter-annual anomalies in each spectral index.

Fourth, we assessed the degree to which maize and soybeans can be distinguished from each other in the Central United States based on remotely sensed crop phenology information. In the Central United States, soybeans are not usually planted until maize planting is complete, creating a 10–20 day lag between the average planting date of these two crops (United States Department of Agriculture, 2010). It should therefore be possible to differentiate areas dominated by maize from those dominated by soybeans in our study area based on time series remote sensing observations (Chang et al., 2007; Wardlow and Egbert, 2008; Ozdogan, 2010).

At the relatively coarse spatial resolution of MODIS, we expect that sub-pixel heterogeneity in land cover and crop type will affect our ability to differentiate maize and soybean phenology. To evaluate this, greenup dates from the MODIS Land Cover Dynamics product were compared at the county-level for maize and soybeans pixels. Before county-level greenup dates were calculated, pixels were stratified into five classes of maize and soybeans based on the sub-pixel proportion classified as each crop in the CDL (i.e., 10–30%, 30–50%, 50–70%, 70–90%, >90%; pixels with <10% of either crop were excluded from the analysis). County-level greenup dates were then calculated for each class independently to explore how the phenological information in the MODIS Land Cover Dynamics product is affected by sub-pixel heterogeneity. In this context, we expect that remote sensing observations should also be sensitive to inter-annual variability in the timing of crop emergence and growth. To investigate covariance between MODIS greenup and crop emergence dates across years, greenup dates for pure (>90%) maize and soybean pixels were averaged to the state-level and compared against crop emergence data from NASS. We defined the date of crop emergence in each state as the date for which 50% of the crop of interest (maize or soybeans) had emerged according to the Crop Progress report, which is produced weekly by the NASS from survey data and provides information on the progress of crop development within each state. Only counties containing >10,000 ha of both maize and soybeans in each year were considered in the analysis of greenup dates and results are only presented for the six states containing more than ten such counties.

3. Results and discussion

3.1. Yield forecast models

Fig. 4 presents multi-year average coefficients of determination derived from linear regression models using EVI2, NDVI and NDWI to predict crop yield at five-day intervals through the growing season for 2004–2009. For the base model, which uses the CDL to identify cropland, correlation between the remotely sensed indices and maize yield peaked 65–75 days after greenup, with NDWI showing the strongest correlation (R^2 = 0.65), followed by EVI2 (R^2 = 0.58) and NDVI (R^2 = 0.53). Peak correlation for soybeans occurred 80 days after greenup for all three indices, with EVI2 showing the strongest correlation (R^2 = 0.73), followed by NDVI (R^2 = 0.70) and NDWI (R^2 = 0.67).

Similar results were observed for models that used the MODIS Land Cover Type Product to identify croplands, suggesting that this product provides a viable alternative in regions where higher resolution products are not available. As for the base model, NDWI had the highest correlation with maize yield (R^2 = 0.65), followed by EVI2 (R^2 = 0.61) and NDVI (R^2 = 0.57). Maximum correlation with maize yield occurred 70–75 days after greenup. For soybean, EVI2 and NDVI were most highly correlated with yield (R^2 = 0.67 and 0.68, respectively), followed by NDWI (R^2 = 0.64), with maximum correlation 80–85 days after greenup.

Coefficients of determination between remotely sensed indices and both maize and soybean yields were systematically lower by roughly 10% when the indices were not adjusted for phenology. NDWI showed the highest correlation with maize yield (R^2 = 0.63) at DOY 200, followed by EVI2 (R^2 = 0.56) and NDVI (R^2 = 0.46) five days later. Maximum correlation between the three indices and soybean yield occurred between DOY 205–210, with EVI2 showing

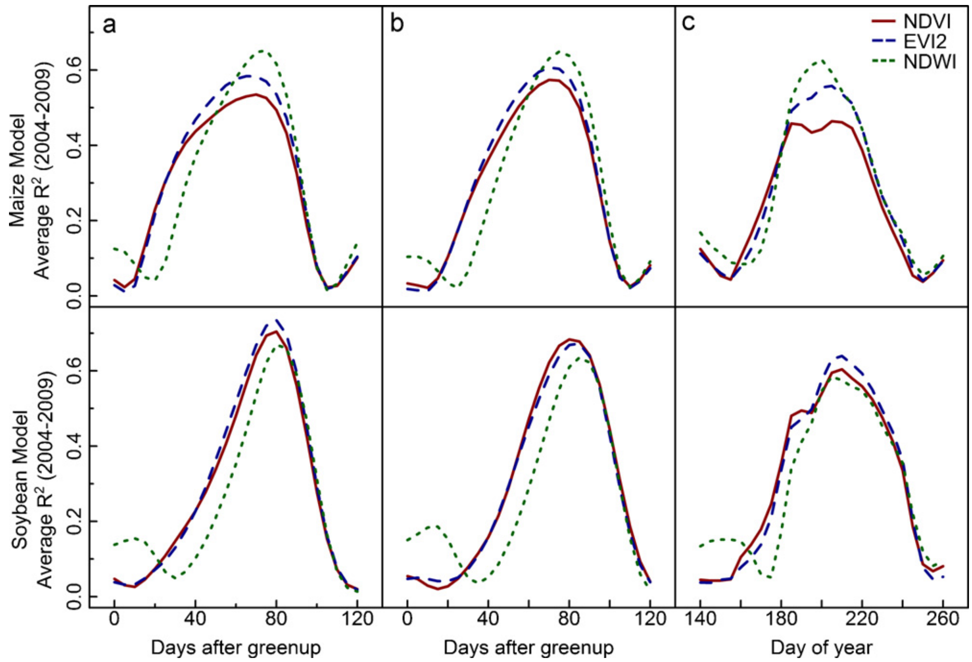

Fig. 4. Multi-year average coefficients of determination between VIs and yield for the (a) base model, (b) model using the MODIS Land Cover Type product to identify agriculture pixels, and (c) model based on fixed calendar dates.

the strongest correlation ($R^2 = 0.64$), followed by NDVI ($R^2 = 0.60$) and NDWI ($R^2 = 0.58$). The timing of peak correlation for maize and soybeans (i.e., mid- to late-July) coincides with the start of the reproductive stage of maize and soybeans in the Central United States (Sacks and Kucharik, 2011) and agrees with previous results showing this to be the best time to predict yield (Ma et al., 1996; Basnyat et al., 2004; Mkhabela et al., 2011).

Fig. 5 shows how the seasonal pattern in correlation between soybean yield and remotely sensed indices vary from year to year. Specifically, this figure shows that correlation between yield and EVI2 consistently peaks 75–85 days after greenup, while the timing of peak correlation (i.e., based on DOY) is much more variable, peaking between DOY 195 and 225. This strongly suggests that geographic and inter-annual variability in phenology affect the correlation between remotely sensed indices and yield, and that using models based on phenologically adjusted VIs provides significant benefit for yield forecast models.

Fig. 6 displays the relationship between each phenologically adjusted VI and yield at the time of peak correlation (i.e., 70 days after greenup for maize and 80 days after greenup for soybeans) for all study years, with counties identified as semi-arid or non-semi-arid according to their ecoregion. The estimated linear regression models are presented in Table 1 along with cross-validated model results, in which each model was trained on data from 2004 to 2006 and tested on data from 2007 to 2009. The results displayed in Fig. 6, Table 1 and from this point forward were produced using the base model (i.e., the CDL was used to identify croplands and VIs were phenologically adjusted). Fig. 6 seems to confirm the linear relationship between the VIs and crop yield and illustrates how the relationship can vary as a function of ecoregion. In particular, NDVI and EVI2 values for a handful of semi-arid maize counties were low compared to non-semi-arid counties with similar maize yield, suggesting that yield would be consistently underestimated in these areas if all counties were modeled together. Most of the

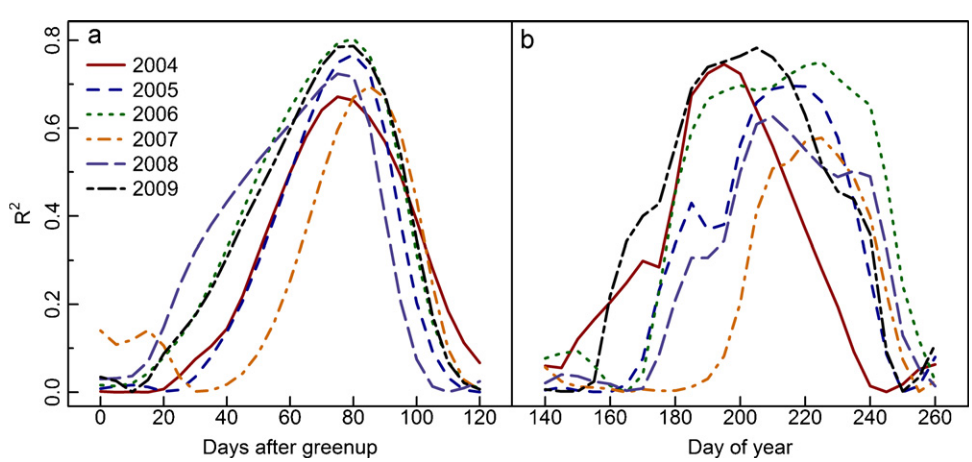

Fig. 5. Annual coefficients of determination between EVI2 and soybean yield for the (a) base model and (b) model based on fixed calendar dates.

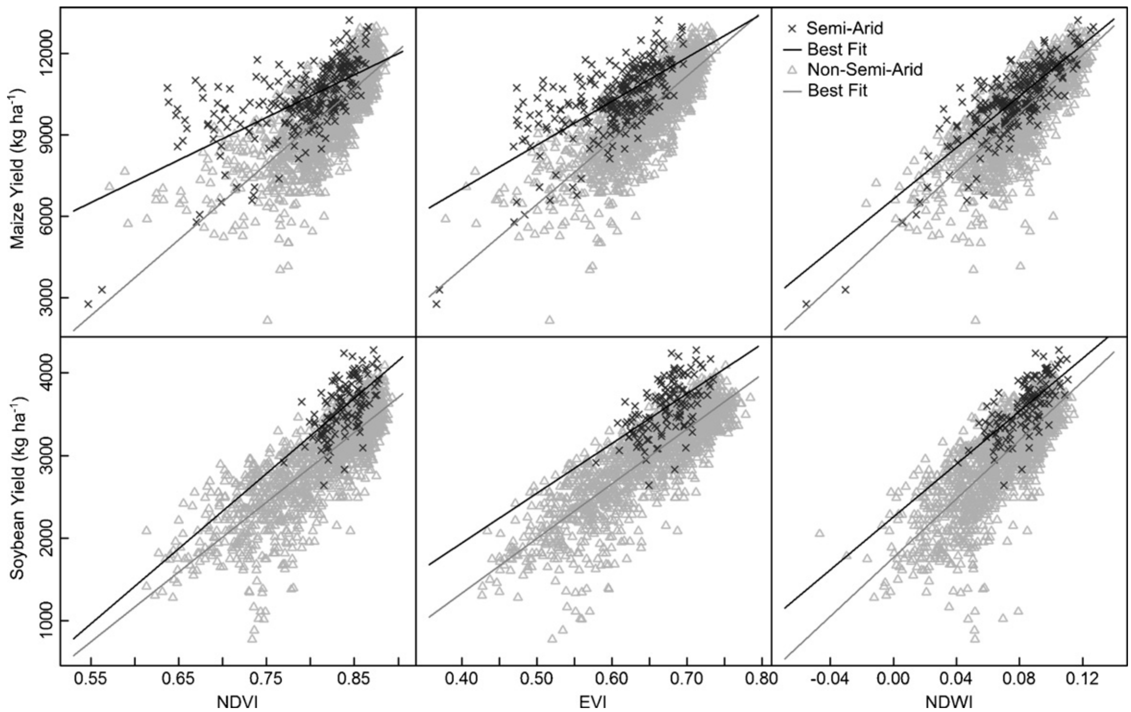

Fig. 6. Scatterplots between crop yield and phenologically adjusted VIs for all study years (2004–2009). VIs from 70 days after greenup are displayed against maize yield, while VIs from 80 days after greenup are displayed against soybean yield.

semi-arid counties with low NDVI and EVI2 values were located in western Nebraska and consisted of a mosaic of irrigated agriculture and grassland. Agriculture in western Nebraska is less dense and includes more fallow areas and winter wheat relative to the eastern portion of the state. The presence of these agricultural subtypes within and surrounding maize pixels lowers NDVI and EVI2 values in these counties relative to counties with more uniform agricultural land use. As a result, NDVI and EVI2 were not able to accurately predict maize yield in semi-arid counties ($R^2 = 0.43$ and 0.46, respectively). In contrast, the relationship between NDWI and maize yield was strong in semi-arid counties ($R^2 = 0.69$) and stronger than NDVI and EVI2 when all counties were modeled together ($R^2 = 0.62$), probably because the NDWI is sensitive to moisture content and therefore able to detect irrigation in the western Nebraska counties. This suggests that the NDWI may have an advantage relative to the EVI2 and NDVI in more arid areas that consist of mixed land cover that include irrigation. Note that soybean is uncommon in western Nebraska (i.e., the soybean models do not

Table 1
Linear regression models for predicting maize and soybean yield with phenologically adjusted VIs. Models were developed for all counties together and for semi-arid and non-semi-arid counties separately using all years of data (2004–2009). Cross-validated coefficients of determination were derived by developing the models on data from 2004 to 2006 and tested them on data from 2007 to 2009.

Crop	Index	Model	*n*	All years		Cross validated (2004–2006 against 2007–2009)	
				R^2	Equation	R^2	Equation
Maize	NDVI	All counties	1998	0.49	$\hat{Y} = 23{,}181.5VI_{70} - 8961.3$	0.44	$\hat{Y} = 21{,}774.3VI_{70} - 8001.3$
		Semi-arid	258	0.43	$\hat{Y} = 15{,}683.1VI_{70} - 2110.2$	0.40	$\hat{Y} = 14{,}770.9VI_{70} - 1588.0$
		Non-semi-arid	1740	0.59	$\hat{Y} = 27{,}918.5VI_{70} - 13{,}007.8$	0.54	$\hat{Y} = 26{,}235.6VI_{70} - 11{,}821.6$
	EVI2	All counties	1998	0.55	$\hat{Y} = 20{,}155.5VI_{70} - 2934.0$	0.48	$\hat{Y} = 19{,}972.7VI_{70} - 3031.4$
		Semi-arid	258	0.46	$\hat{Y} = 16{,}103.7VI_{70} + 583.0$	0.39	$\hat{Y} = 16{,}469.2VI_{70} + 131.5$
		Non-semi-arid	1740	0.67	$\hat{Y} = 23{,}689.1VI_{70} - 5395.5$	0.59	$\hat{Y} = 23{,}262.4VI_{70} - 5336.7$
	NDWI	All counties	1998	0.62	$\hat{Y} = 51{,}816.4VI_{70} + 5752.4$	0.59	$\hat{Y} = 56{,}323.3VI_{70} + 5275.9$
		Semi-arid	258	0.69	$\hat{Y} = 47{,}397.6VI_{70} + 6633.7$	0.62	$\hat{Y} = 53{,}170.2VI_{70} + 6091.3$
		Non-semi-arid	1740	0.64	$\hat{Y} = 53{,}625.8VI_{70} + 5520.0$	0.60	$\hat{Y} = 58{,}571.1VI_{70} + 5003.4$
Soybeans	NDVI	All counties	1950	0.69	$\hat{Y} = 8596.3VI_{80} - 3996.9$	0.69	$\hat{Y} = 8319.5VI_{80} - 3784.8$
		Semi-arid	126	0.34	$\hat{Y} = 9097.0VI_{80} - 4039.0$	0.32	$\hat{Y} = 8289.9VI_{80} - 3398.4$
		Non-semi-arid	1824	0.71	$\hat{Y} = 8438.3VI_{80} - 3893.3$	0.71	$\hat{Y} = 8253.5VI_{80} - 3755.3$
	EVI2	All counties	1950	0.70	$\hat{Y} = 6718.2VI_{80} - 1346.4$	0.70	$\hat{Y} = 6528.6VI_{80} - 1257.0$
		Semi-arid	126	0.33	$\hat{Y} = 6019.4VI_{80} - 461.4$	0.36	$\hat{Y} = 5250.3VI_{80} + 18.0$
		Non-semi-arid	1824	0.73	$\hat{Y} = 6632.3VI_{80} - 1318.7$	0.72	$\hat{Y} = 6537.0VI_{80} - 1291.7$
	NDWI	All counties	1950	0.64	$\hat{Y} = 18{,}138.6VI_{80} + 1759.5$	0.60	$\hat{Y} = 19{,}142.5VI_{80} + 1629.4$
		Semi-arid	126	0.49	$\hat{Y} = 15{,}976.8VI_{80} + 2256.6$	0.50	$\hat{Y} = 16{,}460.9VI_{80} + 2177.3$
		Non-semi-arid	1824	0.65	$\hat{Y} = 17{,}761.0VI_{80} + 1764.6$	0.59	$\hat{Y} = 18{,}978.7VI_{80} + 1619.0$

VI_{70} and VI_{80} are phenologically adjusted VIs 70 and 80 days after greenup, respectively.

Table 2
Root mean squared error between predicted and actual yield for all models. Models were trained on five years of data to predict yield in the removed year.

Crop	Model	Year	RMSE (kg ha^{-1})[a]		
			NDVI	EVI	NDWI
Maize	All ecoregions	2004	1092.43 (11)	941.31 (9)	1026.47 (10)
		2005	1221.81 (13)	1154.03 (12)	992.21 (10)
		2006	1150.94 (12)	1040.53 (11)	869.16 (9)
		2007	1106.15 (11)	1076.92 (11)	960.15 (10)
		2008	964.45 (9)	981.50 (10)	974.25 (10)
		2009	1038.60 (10)	919.63 (9)	809.74 (8)
	Semi-arid	2004	907.47 (9)	898.21 (9)	675.56 (7)
		2005	1169.56 (12)	1154.62 (12)	835.38 (8)
		2006	1397.27 (14)	1306.93 (13)	967.60 (10)
		2007	906.75 (9)	949.33 (9)	857.88 (8)
		2008	732.14 (7)	758.66 (7)	709.95 (7)
		2009	1036.26 (9)	953.07 (9)	597.18 (5)
	Non-semi-arid	2004	1173.56 (11)	910.42 (9)	1083.79 (10)
		2005	1077.26 (11)	1002.93 (11)	962.05 (10)
		2006	963.21 (10)	846.40 (9)	801.06 (8)
		2007	975.75 (10)	935.44 (9)	936.64 (9)
		2008	861.34 (8)	871.84 (9)	979.81 (10)
		2009	1058.31 (10)	863.16 (8)	855.90 (8)
Soybeans	All ecoregions	2004	413.90 (14)	430.36 (14)	456.48 (15)
		2005	264.41 (8)	272.42 (9)	277.94 (9)
		2006	257.70 (8)	254.08 (8)	261.41 (8)
		2007	328.94 (11)	338.43 (11)	345.01 (11)
		2008	285.36 (10)	266.71 (9)	333.75 (11)
		2009	279.40 (9)	261.25 (8)	310.57 (10)
	Semi-arid	2004	423.03 (13)	469.86 (14)	411.86 (12)
		2005	242.96 (7)	265.76 (7)	196.56 (5)
		2006	256.99 (7)	237.01 (7)	231.58 (6)
		2007	234.82 (7)	244.55 (7)	203.16 (6)
		2008	203.45 (6)	193.04 (5)	200.03 (6)
		2009	290.62 (8)	276.88 (7)	239.16 (6)
	Non-semi-arid	2004	412.94 (14)	420.09 (14)	456.89 (15)
		2005	239.40 (8)	241.15 (8)	265.09 (9)
		2006	226.90 (7)	215.66 (7)	240.16 (8)
		2007	324.40 (11)	338.09 (11)	352.07 (12)
		2008	265.70 (9)	242.16 (8)	324.64 (11)
		2009	252.60 (8)	225.66 (7)	299.59 (10)

[a] Percent RMSE is displayed in parentheses.

include data from these counties), which explains why NDWI does not outperform NDVI or EVI2 for soybeans when all counties are modeled together.

The results presented in Table 1 suggest that NDWI is the most suitable index for predicting maize yield in semi-arid counties ($R^2 = 0.69$, cross-validated $R^2 = 0.62$) while EVI2 is most suitable for non-semi-arid counties ($R^2 = 0.67$, cross-validated $R^2 = 0.59$). NDVI and EVI2 perform equally well for predicting soybean yield in non-semi-arid counties ($R^2 = 0.71$ and 0.73, cross-validated $R^2 = 0.71$ and 0.72, respectively), while the relatively small sample of semi-arid soybean counties (21) resulted in models with low R^2 in semi-arid counties. Therefore, we recommend using a single soybean model for the counties in this study area ($R^2 = 0.69$ and 0.70, cross-validated $R^2 = 0.69$ and 0.70 for NDVI and EVI2, respectively).

Table 2 presents cross-validated root mean square errors (RMSE) for predicted yields based on the leave-one-year-out method. The RMSE (expressed in units of percentage relative to the mean) ranged from 5 to 15% and averaged 9.4% across all maize and soybean models. The percent RMSE averaged 7.5% for maize yield predicted using the NDWI in semi-arid counties, and averaged 9.2% for EVI2 in non-semi-arid counties. The percent RMSE averaged 10% and 9.8% when predicting soybean yield for all counties using the NDVI and EVI2, respectively.

Finally, in 2005, many counties in the central and southeastern plains experienced low maize yield relative to the six-year average, which provided an opportunity to evaluate the sensitivity of our yield models to climate-induced yield anomalies. To explore this, Fig. 7 shows scatterplots of maize yield anomalies against anomalies for each of our phenologically adjusted VIs in 2005, where the index and yield anomalies were computed as deviations from the six-year average of each. These scatterplots clearly suggest that the remote sensing indices are sensitive to yield anomalies, with EVI2 showing the strongest explanatory power ($R^2 = 0.73$), followed by the NDVI ($R^2 = 0.62$), and the NDWI ($R^2 = 0.61$).

3.2. Distinguishing maize and soybean crop phenology with MODIS

Fig. 8 displays boxplots for county-level average greenup dates for maize and soybeans in six states, and illustrates how greenup dates from the MODIS Land Cover Dynamics product depends on the sub-pixel composition of crops. According to NASS, the date corresponding to 50% crop emergence is consistently 10–20 days later for soybeans than for maize. For pixels where no single crop covered >50% of MODIS pixels, differences between soybean and maize greenup dates from MODIS were not statistically significant ($p < 0.05$) in most states, suggesting that sub-pixel mixtures of crops affect the signature of crop phenology in MODIS. In addition, for pixels that included substantial proportions of natural vegetation, estimated greenup dates for crops from MODIS were typically 10–20 days earlier than NASS emergence dates and up to ∼30 days earlier in Wisconsin. This result suggests that crop phenology

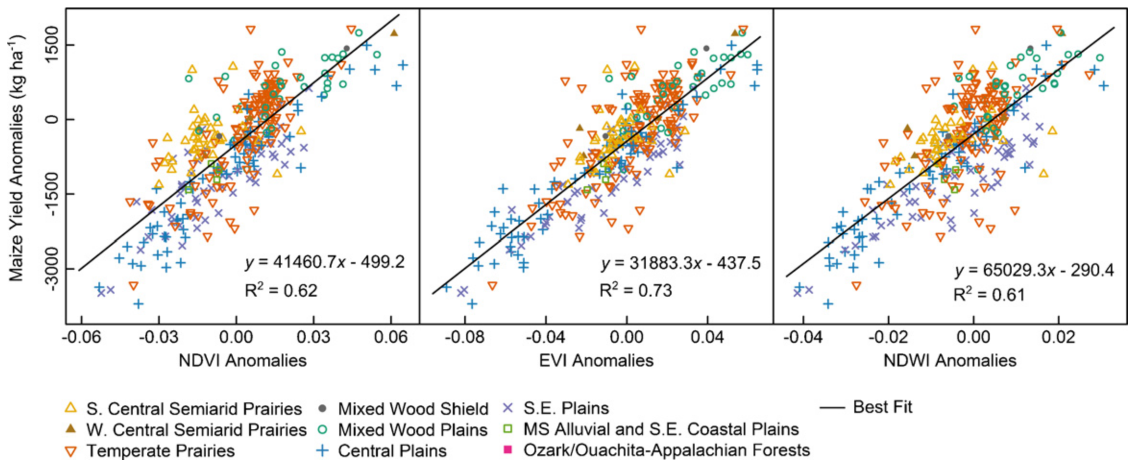

Fig. 7. Scatterplots between maize yield anomalies in 2005 and anomalies in VIs 70 days after greenup.

estimated from MODIS is strongly biased by the presence of natural vegetation and justifies our strategy of removing mixed pixels (i.e., <50% maize or soybeans) from the yield forecast models. Conversely, for maize and soybean pixels that were >90% uniform, differences between maize and soybean greenup dates were statistically significant ($p < 0.001$) in all states, and the greenup dates were in closer agreement with NASS emergence dates. Note, however, that in Wisconsin, greenup dates for pure maize and soybean pixels (>90%) were approximately 10–15 days early relative to the 50% emergence date from NASS, and greenup dates for soybean pixels consistently preceded the date of 50% crop emergence by 5–10 days. The reasons for this bias are unclear, but are probably related to the presence of sub-pixel natural vegetation and weeds in fields prior to crop emergence.

Fig. 9 shows year-to-year covariance between state-level greenup dates for highly uniform (>90%) maize and soybean pixels and the date of 50% crop emergence according to NASS from 2004 to 2009. Covariance between greenup dates and 50% emergence dates for both crops was relatively strong ($r = 0.40$–0.98), suggesting that the MODIS successfully captures inter-annual variability in crop phenology for pure maize and soybean pixels. Correlation was highest in Illinois ($r = 0.96$ and 0.98 for maize and soybeans, respectively), where inter-annual variability in crop emergence dates was largest, ranging from DOY 128 to 151 for maize and from DOY 144 to 172 for soybeans. The magnitude of these variations (~20–30 days) highlights the importance of using phenologically adjusted VIs in yield forecast models.

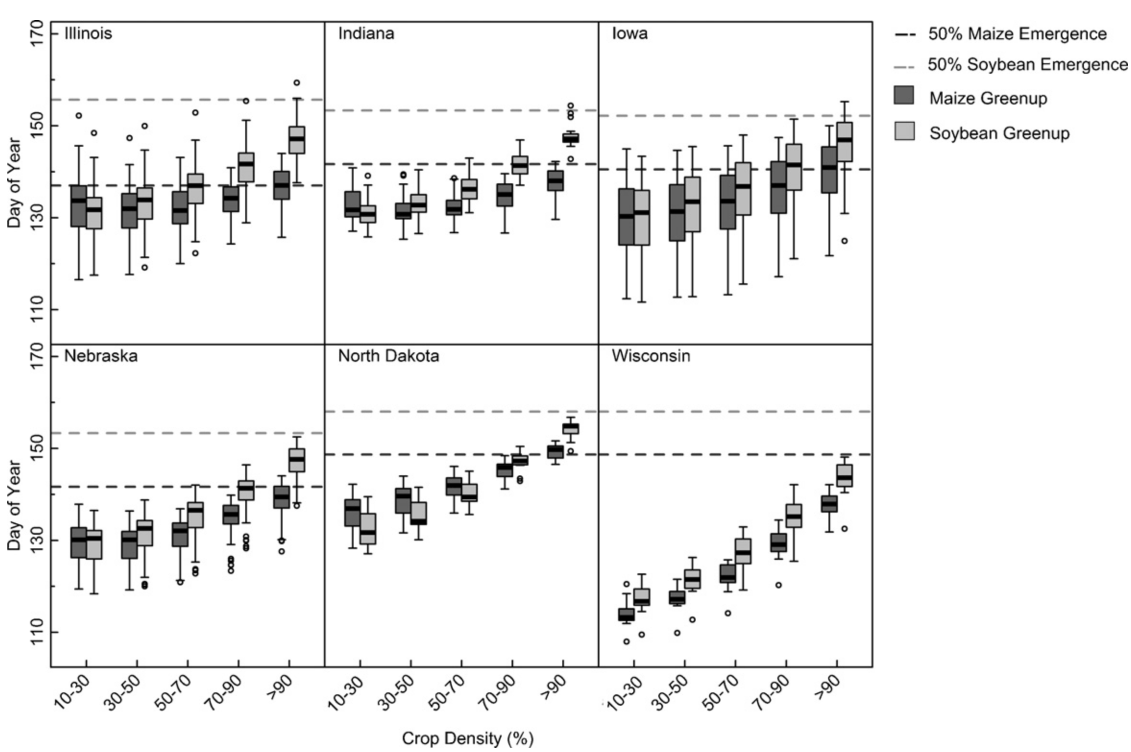

Fig. 8. Boxplots of average (2004–2009) county-level greenup dates for maize and soybeans, with pixels stratified by the sub-pixel proportion classified as maize or soybeans in the CDL. Dashed lines indicate the average (2004–2009) date on which 50% of either crop had emerged in each state following planting according to NASS.

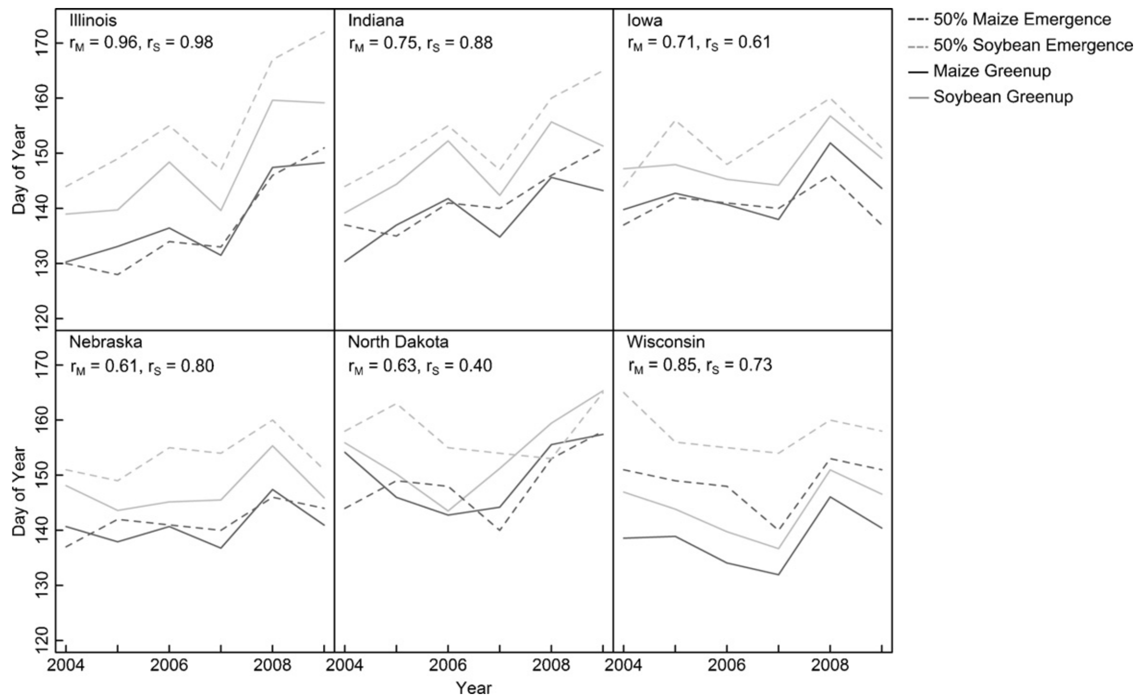

Fig. 9. Inter-annual variability in state-level greenup date and 50% crop emergence date for maize and soybeans. State-level greenup dates were calculated using only pure (>90%) maize and soybean pixels according to the CDL. r_M and r_S are correlation coefficients between greenup and 50% emergence dates for maize and soybeans, respectively.

4. Conclusions

We developed linear models to predict maize and soybean yield in the Central United States using spectral indices derived from MODIS data. To do this, we combined spectral indices with remotely sensed phenology metrics to account for geographic and inter-annual variation in crop phenology. As part of our analysis we also assessed the suitability of different spectral indices and modeling strategies for estimating yield, and examined a variety of confounding factors that affect model performance.

Our results show that including information related to crop phenology improved model predictions, and that the best times to predict crop yields were 65–75 days after greenup for maize and 80 days after greenup for soybeans. The EVI2 was a more effective predictor of maize yield and maize yield anomalies than NDVI, while the two indices performed equally well for predicting soybean yield. The NDWI was the most effective predictor of maize yield in semi-arid areas. Based on these results, we recommend a single soybean model for this study area using NDVI or EVI2, and a mixed model for predicting maize yield (i.e., NDWI in semi-arid counties and EVI2 elsewhere).

More generally, the results from this work suggest that remotely sensed information related to crop phenology is useful for agricultural monitoring and mapping applications. For relatively homogenous pixels, the timing of maize and soybeans greenup appears to be both detectable and separable. However, because MODIS acquires data at relatively coarse spatial resolution, most pixels will include mixtures of crops. Thus, separation of crops based on satellite-derived phenological information is likely to be challenging in many areas (Ozdogan, 2010).

The modeling approach described in this paper requires that a base map of land cover and land use be available to distinguish agricultural land use from non-agricultural land cover. We tested two data sources for this purpose: (1) the MODIS Land Cover Type product, which provides a relatively coarse spatial resolution representation of land cover, and (2) the Cropland Data Layer created by the USDA, which provides a much higher resolution representation. Surprisingly, correlations between VIs and crop yield for models based on the MODIS Land Cover Type product were generally comparable to those from models based on the CDL, suggesting that the MODIS product is a viable alternative for regions where high spatial resolution databases are not available. However, it is important to note that low intensity agriculture is difficult to detect at the 500-m spatial resolution of the MODIS. Therefore, the Land Cover Type product should be used with caution in areas with small field sizes. Having said this, while the analysis described in this paper focused on the agricultural heartland of the United States, the general approach we describe has potential to support yield monitoring and forecasting efforts in other regions of world.

Acknowledgements

The research described in this paper was partially supported by NASA Cooperative Agreement # NNX08AE61A and NASA grant # NNX07AW07G. The authors would like to thank Damien Sulla-Menashe for his assistance in calculating MODIS VIs.

References

Basnyat, P., McConkey, B., Lafond, G.P., Moulin, A., Pelcat, Y., 2004. Optimal time for remote sensing to relate to crop grain yield on the Canadian prairies. Can. J. Plant Sci. 84, 97–103.
Becker-Reshef, I., Vermote, E., Lindeman, M., Justice, C., 2010. A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. Remote Sens. Environ. 114, 1312–1323.
Benedetti, R., Rossini, P., 1993. On the use of NDVI profiles as a tool for agricultural statistics: The case study of wheat yield estimate and forecast in Emilia Romagna. Remote Sens. Environ. 45, 311–326.
Boryan, C., Yang, Z., Mueller, R., Craig, M., 2011. Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer program. Geocarto Int., http://dx.doi.org/10.1080/10106049.2011.562309.
Chang, J., Hansen, M.C., Pittman, K., Carroll, M., DiMiceli, C., 2007. Corn and soybean mapping in the United States using MODIS time-series data sets. Agron. J. 99, 1654–1664.
Commission for Environmental Cooperation, 1997. Ecological regions of North America: toward a common perspective. Montreal, QB, Canada. http://www.cec.org/Storage/42/3484_eco-eng_EN.pdf> (accessed August 2012).

de Boor, C., 1978. A Practical Guide to Splines. Sringer-Verlag, New York.

Doraiswamy, P.C., Hatfield, J.L., Jackson, T.J., Akhmedov, B., Prueger, J., Stern, A., 2004. Crop condition and yield simulations using Landsat and MODIS. Remote Sens. Environ. 92, 548–559.

Doraiswamy, P.C., Sinclair, T.R., Hollinger, S., Akhmedov, B., Stern, A., Prueger, J., 2005. Application of MODIS derived parameters for regional crop yield assessment. Remote Sens. Environ. 97, 192–202.

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS collection 5 global land cover: algorithm refinements and characterization of datasets. Remote Sens. Environ. 114, 168–182.

Funk, C., Budde, M.E., 2009. Phenologically tuned MODIS NDVI-based production anomaly estimates for Zimbabwe. Remote Sens. Environ. 113, 115–125.

Ganguly, S., Friedl, M.A., Tan, B., Zhang, X., Verma, M., 2010. Land surface phenology from MODIS: characterization of the Collection 5 global land cover dynamics product. Remote Sens. Environ. 114, 1805–1816.

Gao, B.-C., 1996. NDWI –A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 257–266.

Genovese, G., Vignollea, C., Nègre, T., Passera, G., 2001. A methodology for a combined use of normalised difference vegetation index and CORINE land cover data for crop yield monitoring and forecasting. A case study on Spain. Agronomie 21, 91–111.

Gu, Y., Brown, J.F., Verdin, J.P., Wardlow, B., 2007. A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great plains of the United States. Geophys. Res. Lett. 34, L06407, http://dx.doi.org/10.1029/2006GL029127.

Gu, Y., Hunt, E., Wardlow, B., Basara, J.B., Brown, J.F., Verdin, J.P., 2008. Evaluation of MODIS NDVI and NDWI for vegetation drought monitoring using Oklahoma Mesonet soil moisture data. Geophys. Res. Lett. 35, L22401, http://dx.doi.org/10.1029/2008GL035772.

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.

Jiang, Z., Huete, A.R., Didan, K., Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band. Remote Sens. Environ. 112, 3833–3845.

Justice, C.O., Vermote, E., Townshend, J.R.G., Defries, R., Roy, D.P., Hall, D.K., Salomonson, V.V., Privette, J.L., Riggs, G., Strahler, A., Lucht, W., Myneni, R.B., Knyazikhin, Y., Running, S.W., Nemani, R.R., Zhengming, W., Huete, A.R., van Leeuwen, W., Wolfe, R.E., Giglio, L., Muller, J.-P., Lewis, P., Barnsley, M.J., 1998. The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for global change research. IEEE Trans. Geosci. Remote Sens. 36, 1228–1249.

Kidwell, K.B., 1998. Polar orbiter data users' guide (TIROS-n, NOAA-6, NOAA-7, NOAA-8, NOAA-9, NOAA-10, NOAA-11, NOAA-12, and NOAA-14). Washington, DC. <http://www.ncdc.noaa.gov/oa/pod-guide/ncdc/docs/podug/index.htm> (accessed August 2012).

Knipling, E.B., 1970. Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. Remote Sens. Environ. 1, 155–159.

Ma, B.L., Morrison, M.J., Dwyer, L.M., 1996. Canopy light reflectance and field greenness to assess nitrogen fertilization and yield of maize. Agron. J. 88, 915–920.

Malingreau, J.-P., 1986. Global vegetation dynamics: satellite observations over Asia. Int. J. Remote Sens. 7, 1121–1146.

Maselli, F., Conese, C., Petkov, L., Gilabert, M.-A., 1992. Use of NOAA-AVHRR NDVI data for environmental monitoring and crop forecasting in the Sahel. Preliminary results. Int. J. Remote Sens. 13, 2743–2749.

Maselli, F., Romanelli, S., Bottai, L., Maracchi, G., 2000. Processing of GAC NDVI data for yield forecasting in the Sahelian region. Int. J. Remote Sens. 21, 3509–3523.

Maselli, F., Rembold, F., 2001. Analysis of GAC NDVI data for cropland identification and yield forecasting in Mediterranean African countries. Photogramm. Eng. Remote Sens. 67, 593–602.

Mkhabela, M.S., Bullock, P., Raj, S., Wang, S., Yang, Y., 2011. Crop yield forecasting on the Canadian prairies using MODIS NDVI data. Agric. Forest Meteorol. 151, 385–393.

Moran, M.S., Clarke, T.R., Inoue, Y., Vidal, A., 1994. Estimating crop water deficit using the relation between surface-air temperature and spectral vegetation index. Remote Sens. Environ. 49, 246–263.

Murphy, W.J., 1993. Tables for weights and measurement: crops. Agricultural publication G4020, University of Missouri extension. <http://extension.missouri.edu/p/G4020> (accessed December 2012).

Ozdogan, M., 2010. The spatial distribution of crop types from MODIS data: Temporal unmixing using independent component analysis. Remote Sens. Environ. 114, 1190–1204.

Rasmussen, M.S., 1992. Assessment of millet yields and production in northern Burkina Faso using integrated NDVI from the AVHRR. Int. J. Remote Sens. 13, 3431–3442.

R Development Core Team, 2009. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>. (accessed November 2012).

Ren, J., Chen, Z., Zhou, Q., Tang, H., 2008. Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong China. Int. J. Appl. Earth Obs. Geoinf. 10, 403–413.

Reynolds, C.A., Yitayew, M., Slack, D.C., Hutchinson, C.F., Huete, A., Petersen, M.S., 2000. Estimating crop yields and production by integrating the FAO crop specific water balance model with real-time satellite data and ground-based ancillary data. Int. J. Remote Sens. 21, 3487–3508.

Rojas, O., 2007. Operational maize yield model development and validation based on remote sensing and agro-meteorological data in Kenya. Int. J. Remote Sens. 28, 3775–3793.

Rudorff, B.F.T., Batista, G.T., 1991. Wheat yield estimation at the farm level using TM Landsat and agrometeorological data. Int. J. Remote Sens. 12, 2477–2484.

Sacks, W.J., Kucharik, C.J., 2011. Crop management and phenology trends in the U.S. corn belt: Impacts on yields, evapotranspiration and energy balance. Agric. Forest Meteorol. 151, 882–894.

Salazar, L., Kogan, F., Roytman, L., 2007. Use of remote sensing data for estimation of winter wheat yield in the United States. Int. J. Remote Sens. 28, 3795–3811.

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

Schut, A.G.T., Stephens, D.J., Stovold, R.G.H., Adams, M., Craig, R.L., 2009. Improved wheat yield and production forecasting with a moisture stress index, AVHRR and MODIS data. Crop Pasture Sci. 60, 60–70.

Thenkabail, P.S., Ward, A.D., Lyon, J.G., 1994. Landsat-5 thematic mapper models of soybean and corn crop characteristics. Int. J. Remote Sens. 15, 49–61.

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

Tucker, C.J., Vanpraet, C.L., Sharman, M.J., Van Ittersum, G., 1985. Satellite remote sensing of total herbaceous biomass production in the Senegalese Sahel: 1980–1984. Remote Sens. Environ. 17, 233–249.

Tucker, C.J., Holben, B.N., Elgin, J.H., McMurtrey III, J.E., 1981. Remote sensing of total dry-matter accumulation in winter wheat. Remote Sens. Environ. 11, 171–189.

United States Department of Agriculture, 2010. Field crop usual planting and harvesting dates. Agricultural Handbook Number 628. <http://usda01.library.cornell.edu/usda/current/planting/planting-10-29-2010.pdf>. (accessed December 2012).

Wall, L., Larocque, D., Léger, P.-M., 2008. The early explanatory power of NDVI in crop yield modelling. Int. J. Remote Sens. 29, 2211–2225.

Wardlow, B.D., Egbert, S.L., 2008. Large-area crop mapping using time-series MODIS 250 m NDVI data: an assessment for the U.S. Central Great plains. Remote Sens. Environ. 112, 1096–1116.

Zhang, X., 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 Sens. Environ. 84, 471–475.

Zhang, X., 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. J. Geophys. Res. 111, G04017, http://dx.doi.org/10.1029/2006JG000217.