Remote Sensing of Environment 114 (2010) 1312–1323

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data

I. Becker-Reshef [a,*], E. Vermote [a], M. Lindeman [b], C. Justice [a]

[a] Department of Geography, University of Maryland, College Park, MD, USA
[b] Foreign Agricultural Service (FAS), U.S. Department of Agriculture (USDA), Washington, DC, USA

ARTICLE INFO

Article history:
Received 1 October 2009
Received in revised form 14 January 2010
Accepted 16 January 2010

Keywords:
Remote sensing
Yield
Agriculture
Wheat

ABSTRACT

Wheat is one of the key cereal crops grown worldwide, providing the primary caloric and nutritional source for millions of people around the world. In order to ensure food security and sound, actionable mitigation strategies and policies for management of food shortages, timely and accurate estimates of global crop production are essential. This study combines a new BRDF-corrected, daily surface reflectance dataset developed from NASA's Moderate resolution Imaging Spectro-radiometer (MODIS) with detailed official crop statistics to develop an empirical, generalized approach to forecast wheat yields. The first step of this study was to develop and evaluate a regression-based model for forecasting winter wheat production in Kansas. This regression-based model was then directly applied to forecast winter wheat production in Ukraine. The forecasts of production in Kansas closely matched the USDA/NASS reported numbers with a 7% error. The same regression model forecast winter wheat production in Ukraine within 10% of the official reported production numbers six weeks prior to harvest. Using new data from MODIS, this method is simple, has limited data requirements, and can provide an indication of winter wheat production shortfalls and surplus prior to harvest in regions where minimal ground data is available.

1. Introduction

Wheat is the most widely-grown food crop in the world and the most important cereal crop traded on international markets (FAO, 2003; MacDonald and Hall, 1980; Wittwer, 1995). Despite the large and growing global wheat output (684 million metric tons produced in 2008/09 according to U.S. Department of Agriculture statistics), total consumption in recent years has often outweighed production. In addition, severe weather conditions in different parts of the world have affected production and contributed to the low global wheat stocks in 2008 (Trostle, 2008). In 2008 global wheat stocks reached a thirty-year low, and the U.S. wheat stocks fell to their lowest levels since the late 1940s (Vocke, 2008). As wheat is the major commodity provided as food aid, wheat shortages not only impact wheat and wheat-product prices, but also have dire implications for ensuring food security in developing countries (Mitchell and Mielke, 2005; WFP, 2009). Timely and accurate forecasts of wheat production prior to harvest at regional, national and international scales for both developing and developed countries are crucial. Such estimates can help to improve food accessibility risk management, as well as play an important role in global markets, policy and decision making (Justice and Becker-Reshef, 2007).

A range of techniques such as visual field estimates, multiple frame-based sample surveys, analog-year approaches, crop-simulation models and regression approaches have been used for forecasting pre-harvest yield estimates with varying degrees of success (Chipanshi et al., 1999; Doraiswamy et al., 2003; Maselli et al., 2001; Pinter et al., 1981; Wall et al., 2007). Earth observation data, owing to their synoptic, timely and repetitive coverage, have been recognized as a valuable tool for yield and production forecasting (Manjunath et al., 2002; Prasad et al., 2006). Agricultural monitoring from space, in particular pre-harvest assessments of crop yield and production, has been a topic of research since the early 1970s (Wall et al., 2007). In the US the Large Area Crop Inventory Experiment (LACIE), a joint effort between the USDA, NASA and the National Oceanic and Atmospheric Administration (NOAA) initiated in 1974, demonstrated that earth observations, could provide vital information on production, with unprecedented accuracy and timeliness, prior to harvest (MacDonald and Hall, 1980). This experiment spurred many agencies and researchers around the world to further develop and evaluate remote sensing technology for timely, large area, crop monitoring.

Remotely sensed satellite data offer timely, objective, economical, and synoptic information. Their utility for crop yield forecasting has been demonstrated across a wide range of scales and geographic locations (Funk and Budde, 2009; Hatfield, 1983; Hayes and Decker,

* Corresponding author.
E-mail address: ireshef@umd.edu (I. Becker-Reshef).

0034-4257/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2010.01.010

1996; Kastens et al., 2005; Labus et al., 2002; Mika et al., 2002; Mkhabela et al., 2005; Quarmby et al., 1993; Rojas, 2007; Salazar et al., 2007; Weissteiner and Kuhbauch, 2005). In particular, the Normalized Difference Vegetation Index (NDVI) has been recognized since the early 1980s for its value in monitoring crop conditions and forecasting crop yields (Boken and Shaykewich, 2002; Doraiswamy and Cook, 1995; Quarmby et al., 1993; Tucker et al., 1980). A variety of methods have been developed to estimate crop yields using remotely sensed information including biophysical crop-simulation models that retrieve crop growth parameters from remotely sensed data and which are used as inputs to calibrate and drive the models. The main drawback of such models is that they typically require numerous crop specific inputs such as soil characteristics, management practices, agro-meteorological data and planting dates, in order to simulate crop growth and development through the crop cycle (Dubey et al., 1994; Moriondo et al., 2007). Such crop-simulation models include CERES (Ritchie and Otter, 1985) WOFOST (Vandiepen et al., 1989) CROPSYST (Vanevert and Campbell, 1994) and STICS (Brisson et al., 1998). Statistical regression-based methods are the most commonly used remote sensing-based approaches (Wall et al., 2007). These are based on empirical relationships between historic yields and reflectance based vegetation indices. They are typically straightforward to implement and do not require numerous inputs. A main drawback of empirically-based approaches is that relationships between yield and reflectance are typically localized and are not easily extendable to other areas (Doraiswamy et al., 2003; Moriondo et al., 2007). Nonetheless, they are often the preferred approach owing to their limited data requirements and simplicity to implement.

The basic assumption behind the empirical statistical approaches is that measures of photosynthetic capacity, estimated from spectral vegetation indices such as NDVI, are directly related to yield. This is because many of the conditions that favorably or adversely affect plant development and ultimately yield (i.e. fertilization treatment, rust infection, drought, or precipitation-events) result in a corresponding increase or reduction of the crop's photosynthetically active biomass and this response can often be captured though spectral measures such as NDVI (Tucker, 1979). Thus, a limitation of such an empirical, NDVI-based approach is that estimates of yield are often inaccurate when photosynthetic capacity at the time of measurement is not the main determinant of grain yield. An additional limitation of this approach in highly productive crop fields, is that NDVI saturates at a LAI of approximately 4 (Wang et al., 2005).

Pioneering work carried out in this field such as by Fischer (1975) found that wheat yields could be forecast as a function of the leaf area at the onset of the reproductive stage which corresponds to the timing of maximum crop green leaf area. Pinter et al. (1981) found that wheat and barley yields could be related to accumulated NDVI over the growing season. Tucker et al. (1980) found significant linear relationships between wheat yields and time-integrated NDVI measures during the growing season, and determined that the strongest correlation of yield with NDVI occurred around the time of maximum green leaf biomass. The findings from such studies established the ground work for numerous subsequent studies relating spectral vegetation indices to crop yields at regional and national scales.

For example, in field experiments in the Punjab region of India, Mahey et al. (1993) found that NDVI measurements during maximum green crop canopy cover were highly and linearly correlated with wheat yields. Dubey et al. (1994) developed linear regression models to forecast wheat yields at the district level in Punjab using the NDVI derived from the Landsat Multi Spectral Scanner (MSS) and the Indian Remote Sensing (IRS) Satellite Linear Imagine Self Scanning Sensors (LISS-1). Similar methods were used by Sridhar et al. (1994) to forecast wheat yields in Madhya Pradesh, India. Manjunath et al. (2002) derived linear regression models to forecast wheat yields in Rajasthan utilizing NDVI data derived from the National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR), in conjunction with rainfall data and yield data.

In a study in Canada, Boken and Shaykewich (2002) enhanced a pre-existing operational, district-level wheat yield forecasting model, driven by a monthly cumulative moisture index (CMI) by incorporating daily NDVI data into the model, and found that the average NDVI during the peak of the growing season and the average NDVI of the entire growing season were the best predictor parameters for wheat yield. A subsequent study comparing the explanatory power of NDVI for wheat yield modeling versus that of CMI found that the NDVI-based model could forecast yields four weeks earlier than the CMI-based model (Wall et al., 2007). Basnyat et al. (2004) conducted field studies in the Canadian prairies and found significant correlations between NDVI and grain yield and determined that the optimal timing for obtaining NDVI measurements for spring-planted wheat was approximately one month prior to harvest.

In Senegal, Rasmussen (1997) developed a linear regression model driven by integrated NDVI measures to estimate millet yields. Maselli and Rembold (2001) used regression models utilizing multi-year NDVI data to estimate wheat yields at the national level in North African countries. In Zimbabwe, Funk and Budde (2009) found that phenologically-adjusted, crop-weighted NDVI-anomaly time-series data were correlated with crop production anomalies and could be used by the Famine Early Warning Systems Network (FEWS) to provide an early and objective evaluation of production. Other methods explored by the famine early warning community include utilizing the relationship between NDVI anomalies and the Pacific Sea Surface Temperature (SST) anomalies in Southern Africa to forecast NDVI anomalies at the onset of the growing season which can then be related to crop production (Verdin et al. 1999).

In China, Ren et al. (2008) developed a linear regression model to forecast winter wheat yields in the Shandong Province. Their model regressed spatially-accumulated NDVI measurements at the county level during the growing season with county-level production statistics. Yield was then derived from the model's predicted production divided by acreage statistics to forecast yields within 10% of official statistics. In the U.S. Doraiswamy et al. (2003) used several input parameters retrieved from satellite imagery in a crop growth model to simulate spring wheat yields at the sub county and county levels in North Dakota.

Despite extensive studies on crop yield forecasting, crop models have rarely progressed successfully into operational implementation and are typically only applicable in the region for which they are developed. Most countries and agricultural agencies, including the USDA National Agricultural Statistics Service (NASS), still rely on traditional methods for generating crop statistics. These include multiple frame-based sample surveys from farm operators, and objective yield surveys where ground measurements of crop yields are collected from randomly selected fields (Allen et al., 2002).

This study takes advantage of considerable recent improvements in sensor technology and data availability, including a new surface reflectance, BRDF-corrected, daily product from NASA's MODIS data (Vermote et al., 2009). The goal of this study was to build on the extensive body of research that has been carried out over the past several decades on crop yield and production forecasting, utilizing new remotely sensed capabilities and products to develop a generalized empirical, remotely sensed based winter wheat yield forecasting model that is simple, robust, timely, economical, widely applicable and portable between winter wheat growing regions that can be used to forecast production within 10% of official statistics. Winter wheat constitutes approximately 80% of global wheat production (internal estimate of the International Production Assessment Division (IPAD), FAS, USDA, August 2009) and therefore was selected for this study.

Winter wheat is generally planted in autumn and crop emergence and stand establishment typically occur in late autumn prior to

Table 1
Data used in the study.

Category	Data	Source	Spatial resolution
Crop statistics	NASS quick stats yield, production, area harvested, area planted	USDA NASS	State, county
	Ukraine crop statistics	State Statistical Committee of Ukraine (SSC)	Oblast, country
Crop type	Kansas: USDA NASS Cropland Data Layer (CDL)	USDA NASS	56 m
	Ukraine: crop type data layer produced using a decision tree approach and AWiFS imagery, May 2008, July 2008.	AWiFS	56 m
	16 day composite MODIS Surface Reflectance time-series (2007–2008)	MODIS	250 m
Time-series earth observations	MODIS, daily BRDF-corrected surface reflectance and NDVI (2000–2008)	MODIS	0.05°

winter. Winter wheat requires a process of vernalization (exposure to low temperatures) in order to flower the following spring (Miller, 1999). Vernalization requires exposure to temperatures of 5 to10 °C for a period of six to eight weeks, and the crop must be biologically active for at least four weeks prior to vernalization. In colder climates (Ukraine, for example), winter wheat typically enters a state of dormancy during the winter. During dormancy, vegetative growth ceases and the crop's resistance to cold weather is increased. Once the soil warms up in the spring, winter wheat resumes vegetative growth and reaches maturity in early to mid-summer. According to 2007 through 2009 statistics from the U.S. Department of Agriculture (USDA) Foreign Agricultural Service (FAS), the U.S. was the largest wheat exporter in the world, followed by the EU (EU-27), Canada, Russia, Australia, Ukraine, Kazakhstan, Argentina, Turkey and China (FAS, 2009).

2. Materials

This study combined earth observations data and derived products with reported crop statistics to develop an empirical, generalized, regression-based, winter wheat yield forecasting model. This study first developed and tested a regression-based model in Kansas, a major wheat producing region that has detailed agricultural statistics. Once developed, this Kansas regression model was directly applied to Ukraine, a major wheat growing country. According to USDA statistics, Ukraine is the sixth largest wheat exporter, on average accounting for 6% of total wheat exports between 2007 and 2009 (FAS, 2009).

Three types of data were used in this study: county-level crop statistics on winter wheat; winter wheat crop type maps; and BRDF-corrected MODIS surface reflectance data (Table 1). Crop type masks were used to identify the winter wheat areas. The Kansas yield statistics were then used with the MODIS NDVI data to develop an empirical relationship between NDVI and yield which was then applied uniformly to Kansas and Ukraine.

2.1. Study site and official crop statistics

The state of Kansas was chosen as the study site for developing the winter wheat yield forecasting model, as it is the main winter wheat producing state in the U.S., it has large areas of consecutive wheat fields which generally range from 30 to 150 ha, and it has a comprehensive and reliable county-level archive of crop statistics. In 2008, the state of Kansas produced 9.7 million tons of winter wheat, approximately one fifth of the total U.S. wheat production (NASS, 2008). Winter wheat production is concentrated in the western two-thirds of the state (Fig. 1). A reliable archive of county-level statistics on yield, area harvested, and production is available from the USDA National Agricultural Statistics Service (NASS) Quick Stats database (NASS, 2008). NASS is the agency responsible for administering the USDA's U.S. program for collecting and publishing agricultural statistics at the national, state and county levels. It is considered a world leader in agricultural statistics providing, a comprehensive, uniform,

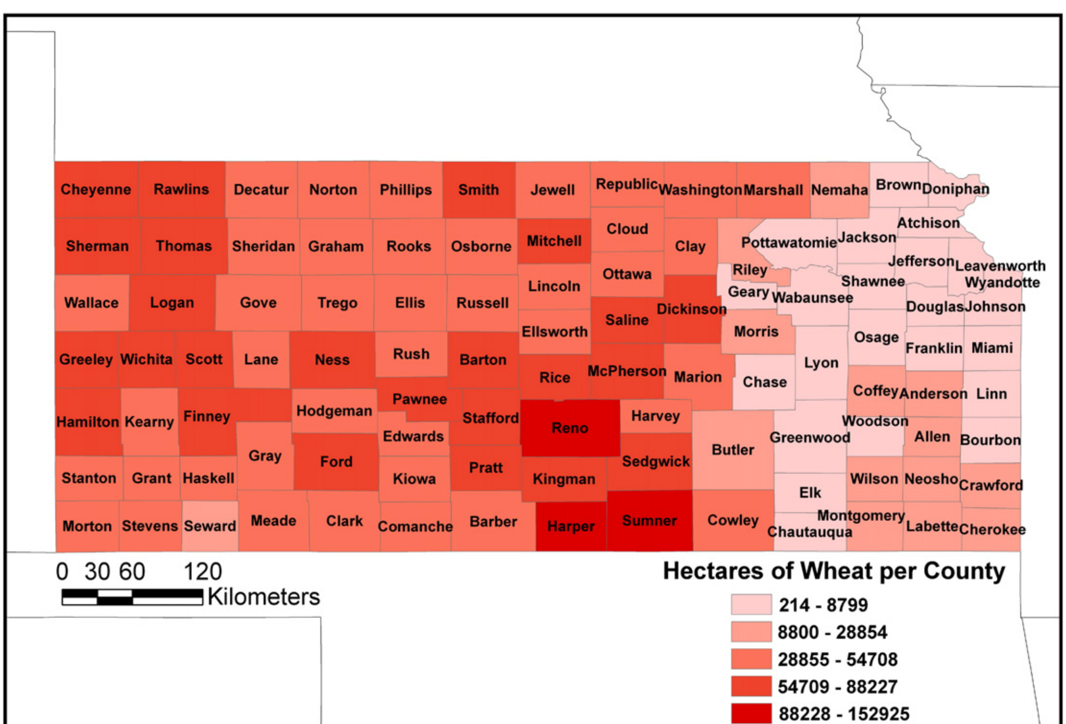

Fig. 1. Kansas winter wheat distribution based on USDA NASS average planted area (hectares) statistics from 2000 to 2006. Dark reds indicate higher per-county planted area and pinks indicate lower per-county planted area.

and reliable data set on U.S. crop statistics. The NASS crop statistics are based on data obtained from multiple frame-based sample surveys of farm operators, objective yield surveys, agribusinesses, shippers, processors and commercial storage firms (NASS, 2007d). The time-series of NASS county-level crop statistics from 2000 to 2008 were used to train and develop the basic regression model in Kansas.

For Ukraine, oblast-level crop statistics were obtained from the State Statistical Committee of Ukraine (SSC) for winter wheat area harvested and yield. (An Oblast is a sub-national unit approximately three times the size of a Kansas county.) These official statistics are based on farm surveys collected from all the agricultural enterprises (large-scale farms that produce commodities exclusively for sale) which account for over 75% of Ukraine's grain production, and from a sample of household farms (small farms and household plots that produce crops both sale and for personal consumption) which account for the remainder of the grain production (Personal communication, Oleg Prokopenko, Chief Agricultural Section, State Statistical Committee of Ukraine, April 2009). These official Ukrainian statistics were used for validation purposes only.

2.2. Crop type maps

Identification of winter wheat fields is an important component of the model development and implementation as it allows for retrieval of winter wheat specific remotely sensed parameters. For example, in a study by Genovese et al. (2001) it was found that applying a cropland mask to select NDVI values as input to a crop yield model significantly improved the accuracy of crop yield forecasts. A similar method was used in this study, applying a crop type map to select the NDVI values that were used as input to the Kansas yield regression model. For Kansas a winter wheat map was available from the Cropland Data Layer (CDL) produced by NASS. The CDL is a rasterized land cover map using field level training data from extensive ground surveys, farmer reports provided to the U.S. Farm Service Agency (FSA), and using remotely sensed data from Landsat Thematic Mapper (TM), Landsat Enhanced Thematic Mapper (ETM+) and Advanced Wide Field Sensor (AWiFS). These data are used in a decision tree classifier in order to produce a land cover classification that distinguishes between different crop types including winter wheat (NASS, 2007c). In this study, the Kansas, 2006 and 2007 CDL layers were used to identify winter wheat growing areas. A winter wheat map for Ukraine was also necessary for the application of the regression model in Ukraine. As no winter wheat maps were available for Ukraine, a rasterized winter wheat map was produced using a decision tree classification method, similar to that used to produce the NASS CDL and other land cover classifications such as those described by Hansen et al. (2000) and De Fries et al. (1998). Training data covering 37,660 ha, equivalent to 0.53% of 2008 planted wheat area, were collected from moderate resolution AWiFS images (56 m) from May, 2008 (timing of the winter wheat anthesis growth stage in Ukraine) and from June and July 2008 (timing of the maturity and harvest of winter wheat in Ukraine). The distribution of the wheat training data is presented in Fig. 2A. The classification tree was then run on one year of time-series of MODIS surface reflectance 16-day composite data (August 2007–August 2008) to produce a winter wheat map for Ukraine at the MODIS 250 m resolution. The map presented in Fig. 2B. In order to validate this map, the classified winter wheat was aggregated to the oblast scale and compared with the official Ukraine SSC winter wheat planted area statistics. When compared at the oblast level, the classified area from the 250 m mask had a bias of 28% (SSC area = Classified Mask estimate * 0.722439) and a precision accuracy of 4.5% (Fig. 2C).

2.3. MODIS daily climate model grid (CMG) time-series

Traditionally NDVI composite data from the AVHRR have been used to drive remotely sensed based crop yield forecasting models

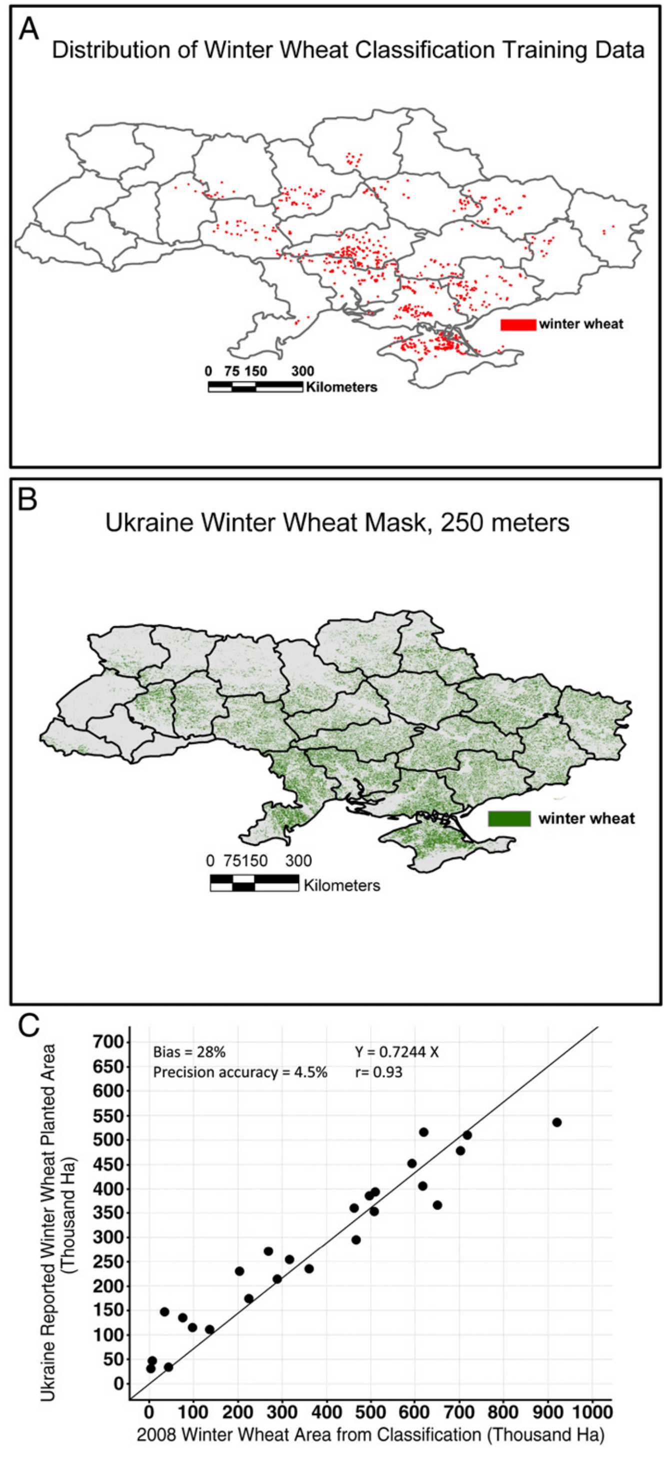

Fig. 2. Ukraine winter wheat mask. Panel (A) shows the distribution of the training data collected from AWiFS imagery for 2008 winter wheat crop type map for Ukraine. Panel (B) is the 2008 classified Ukraine winter wheat map at 250 m resolution. Panel (C) is the validation of the winter wheat classification map (*X*-axis) against the official Ukrainian SSC winter wheat oblast-level, planted area statistics (*Y*-axis). When compared at the oblast level, the area computed from the classified mask had a bias of 28% (SSC area = Classified_Mask_estimate * 0.722439) and a precision accuracy of 4.525%.

(Benedetti and Rossini, 1993; Doraiswamy et al., 2003; Ferencz et al., 2004; Hayes and Decker, 1996; Manjunath et al., 2002; Mika et al., 2002; Salazar et al., 2007; Smith et al., 1995; Weissteiner and Kuhbauch, 2005). Daily data from AVHRR had significant levels of noise due to effects of such factors as aerosols, clouds, and watervapor. To minimize noise due to these effects, temporal composite data derived using the Maximum Value Composite (MVC) method were preferred over daily data, as the highest quality pixels from each composite time-frame are selected (Holben, 1986). With time, the quality of remotely sensed data has significantly improved due to advances in sensor technology, such as with the MODIS instrument, as well as due to significant methodological enhancements to atmospheric correction, cloud detection, BRDF correction algorithms and data quality flags (Justice et al., 1998; Vermote et al., 2002). Recently, a new high quality, daily-updated, BRDF-corrected, surface reflectance data set from MODIS (at 0.05°) has been developed, enabling reliable daily monitoring of agricultural regions (Vermote et al., 2009). This is a daily-updated product which means that clear, high quality pixels are identified daily, and cloudy, low quality pixels are flagged and replaced by linearly interpolated values until high quality data become available. In this way, this dataset preserves all of the useable data rather than just a single data point selected by the compositing procedure. This daily product is similar to the 'objective analysis' scheme that was applied to AVHRR sea surface temperature (SST) (Santoleri et al., 1991). The objective analysis method was used to produce a daily SST product where cloud-free, good quality data were preserved and cloudy pixels or atmospherically contaminated pixels were removed and replaced by values that were interpolated between irregularly spaced observations. In their study Santoleri et al. (1991) compared daily objective analysis AVHRR SST maps to their corresponding composite images and found that the objective analysis images were able to more accurately characterize SST values in areas of rapid variability than in the composite images.

Similar to the objective analysis approach, the CMG MODIS daily-updated surface reflectance is able to preserve the temporal signal over cropped areas better than the corresponding standard composite products. As such it should allow for better estimation of remotely sensed based growing season parameters such as the seasonal peak NDVI than the corresponding composite product. NDVI was chosen as the RS input parameter as it has been shown to be strongly correlated during the growing season to crop condition parameters such as vigor, stress, green biomass and photosynthetic capacity (Idso et al., 1979; Jackson et al., 1986; Patel et al., 2006; Sellers, 1985; Tucker, 1979; Tucker et al., 1981; Wiegand and Richardson, 1990). The NDVI was computed from the MODIS surface reflectance time-series CMG daily product using the Red and Near Infrared (NIR) bands: NDVI = (NIR − RED)/(NIR + RED) (Tucker, 1979).

3. Methods: regression model development for Kansas

The first step in the model development was to scale up the NASS CDL to the MODIS CMG scale as a percent wheat crop type mask. A similar method to that described by Kastens et al. (2005) was used in order to produce the crop percent masks. The 2006 and 2007 CDLs had a 56 m pixel resolution. The CDLs were first reprojected to match the MODIS CMG geographic projection and the winter wheat pixels were extracted. A binary mask was then created where wheat pixels were set to 1 and all other pixels were set to 0. This binary mask was used to produce the percent wheat mask at the 0.05° resolution by computing the average wheat value from the CDL binary mask (56 m resolution) map within each 0.05° pixel footprint (Eq. (1)).

$$Pct\ Wheat_{0.05\,\mathrm{Degree}}X = \frac{1}{N}\sum\nolimits_{p=1}^{N}\left[WheatMask_p\right]*100 \tag{1}$$

where, Pct $\mathrm{Wheat}_{0.05\ \mathrm{Degree}}X$ is the percent wheat value for a given pixel X at the 0.05° scale, $\mathrm{WheatMask}_\mathrm{p}$ (1 through N) are the 56 m CDL binary wheat mask pixels (value 0 = not wheat, value 1 = wheat) within the corresponding 0.05° pixel.

Although the location of winter wheat fields varies from year to year at the CDL map scale due to crop rotations, this study found that at the CMG scale the percent wheat was relatively constant from one year to the next. This is demonstrated in Fig. 3. Fig. 3A shows the winter wheat field rotations in western Kansas at the CDL scale between 2006 (green fields) and 2007 (blue fields) with the 0.05° grid overlaid in gray. Fig. 3B shows the percent of the CDL-based, binary-mask pixels that changed between 2006 and 2007 within each 0.05° pixel. The oranges and reds indicate larger percent change (e.g. red indicates a close to 90% change of the CDL-based wheat pixels within the corresponding 0.05° pixel, and the blues and greens indicate a smaller percent change). Fig. 3C shows the percent change between the 0.05° percent wheat masks between 2006 and 2007. In this figure it is evident that at the 0.05° scale, the per pixel percent wheat remains relatively constant with the large majority of pixels varying less than 20%. This assumption, that at the coarser resolution grid the percent area of wheat becomes relatively constant, is also supported by the official NASS Kansas statistics of wheat planted area. The average percent change of wheat area relative to the mean (2000–2008) was 2% in 2006 and 4% in 2007. Therefore, the two wheat maps from 2006 and 2007 were averaged to produce one winter wheat percent mask at 0.05° resolution (Fig. 3D).

Once the percent wheat mask was produced it was used to select the top 5% purest winter wheat pixels at the 0.05° grid scale from each county in Kansas, to retrieve the NDVI values for the winter wheat crop. Using these pixels as a mask, an NDVI time-series derived from the daily CMG product was retrieved for every year of data from each county. A weighted NDVI time-series for every year of data was computed for each county, where weights were assigned according to the percent wheat of each pixel so that purer pixels had a proportionally larger weight. The result was an eight year (2000–2008) daily NDVI time-series for each county in Kansas derived from a weighted average of the purest winter wheat pixels in each county. Fig. 4 presents an example of the NDVI time-series for Harper County, one of the highest wheat producing counties in Kansas. The figure depicts the daily NDVI values from 2000 through 2008 extracted for the winter wheat areas. The numbers in blue are the final yield values for Harper County (N). As is evident from Fig. 4, the yield values co-vary with the maximum NDVI values from each season. This is apparent in all years with the exception of 2007 where the maximum NDVI value is high and the yield value is relatively low. The low yields in 2007 were due to a late spring frost which affected large portions of Kansas. The late freeze occurred in early April, when the winter wheat crop was nearing its peak vegetative state and the majority of its green biomass had been established. This freeze caused floral sterility primarily in the central and eastern parts of the state which resulted in low yields despite the high green biomass (NASS, 2007a,b; Woolverton, 2007). As a result the NDVI was relatively high, reflecting the well developed green vegetation, but the end of season yields and production were low as a result of the freeze damage which was not reflected in the NDVI. Therefore, for the purposes of establishing the generalized regression model, data from 2007 were excluded.

In this study the seasonal maximum NDVI was used as the main remotely sensed input parameter. Some studies (e.g. Labus et al., 2002; Rojas, 2007) have shown that seasonally integrated NDVI, can predict yields more accurately than measures such as the seasonal maximum NDVI, since it can capture the effect of adverse events which occur after flowering. Nevertheless, in this study the seasonal maximum NDVI was chosen since it enabled a timely prediction of production approximately a month and a half prior to harvest. To retrieve the maximum NDVI the 95th percentile seasonal NDVI value was extracted from the county, wheat specific daily time-series for each year of data (2000–2008). The maximum NDVI values were then adjusted to reduce the influence of non-wheat noise, such as soil, on the wheat NDVI signal. This was achieved through applying a method similar to that suggested by Rassmusen (1998) where the pre-growing season average low vegetation index value for each pixel was

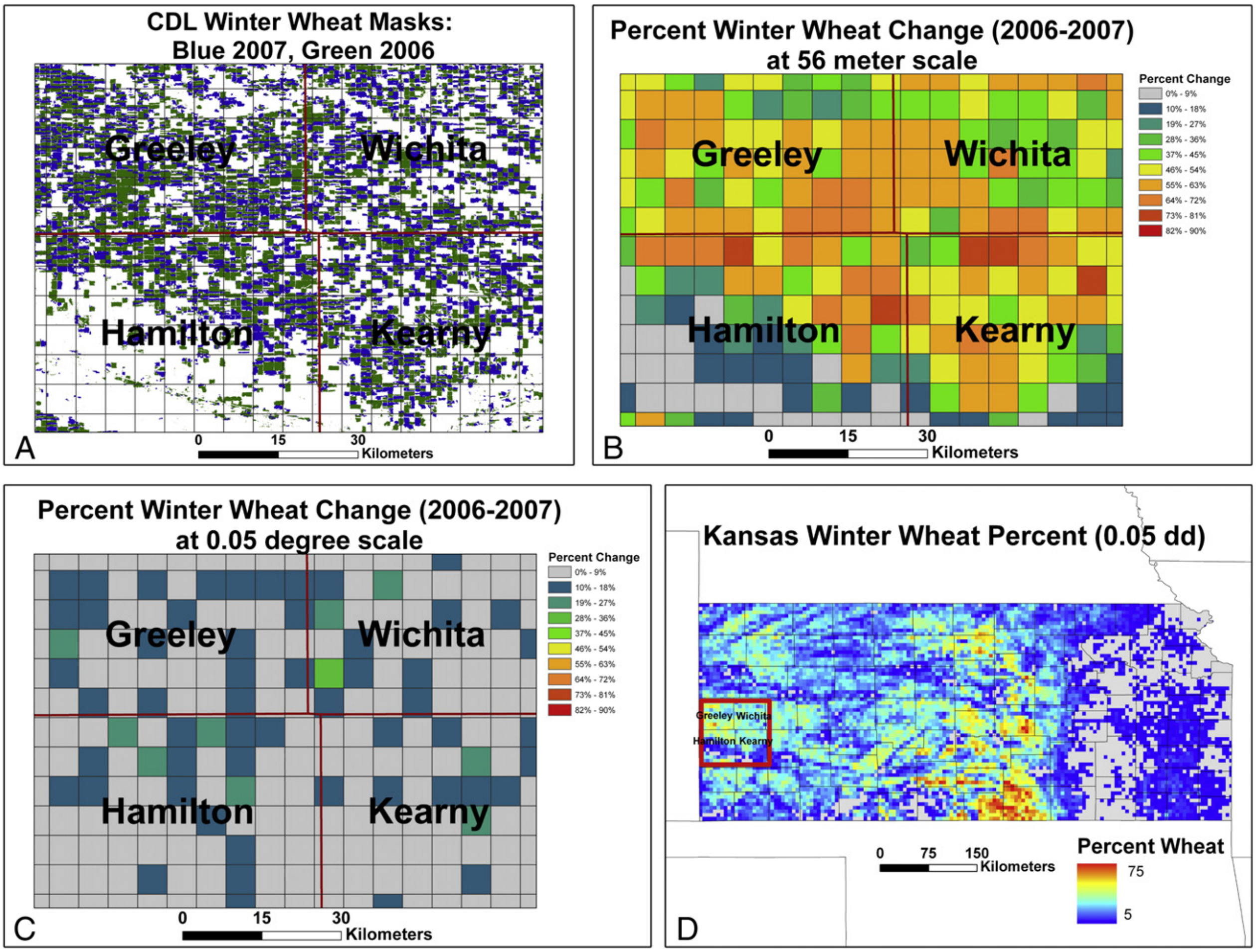

Fig. 3. Kansas winter wheat masks. Panel (A) is a zoom in to western Kansas counties showing the winter wheat field rotations at the CDL scale (56 m) between 2006 (green fields) and 2007 (blue fields) with the 0.05° grid overlaid in gray. Panel (B) shows the percent of the CDL-based, binary-mask pixels that changed between 2006 and 2007 within each 0.05° pixel within the area shown in (A). The oranges and reds indicate larger percent change (e.g. red indicates a close to 90% change of the CDL-based wheat pixels within the corresponding 0.05° pixel, and the blues and greens indicate smaller percent change). Panel (C) shows the percent change of the 0.05° percent wheat masks between 2006 and 2007 in the same region. In this figure it is evident that at the 0.05° scale, the per pixel percent wheat remains relatively constant with the large majority of pixels varying less than 20%. Panel (D) is the scaled-up 0.05° Kansas percent winter wheat mask used in this study derived from the 2007 and 2006 CDL maps. The red box highlights the area shown in (A)–(C).

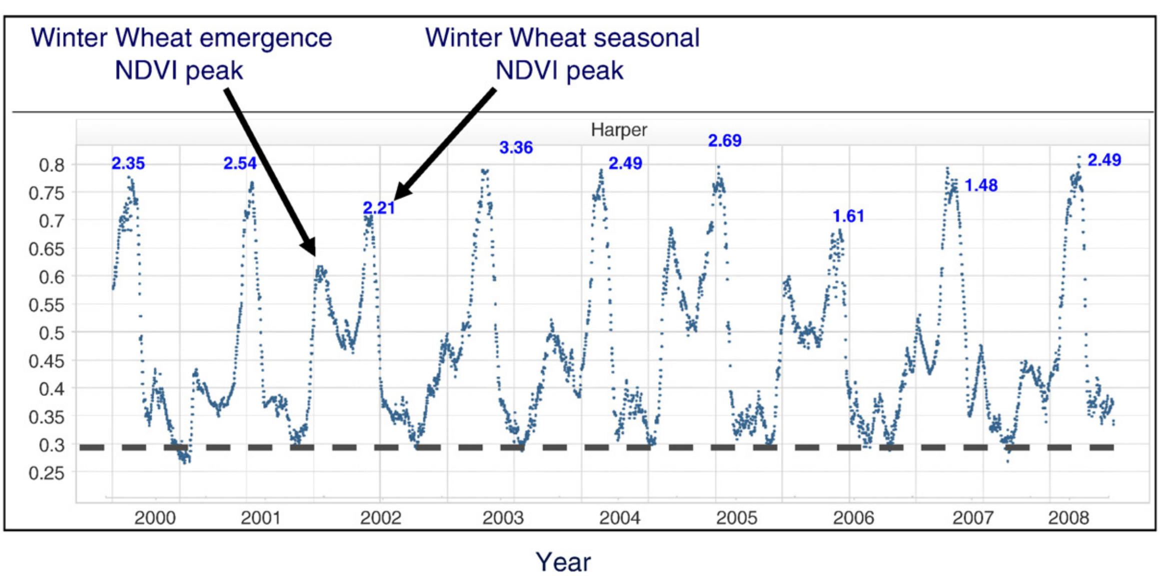

Fig. 4. This figure provides an example of the NDVI time-series for Harper County, one of the highest wheat producing counties in Kansas with time on the X-axis and NDVI values on the Y-axis. The daily NDVI values were extracted from the winter wheat areas for 2000 through 2008. The numbers in blue are the final yield values for Harper County. The figure shows that the yield values co-vary with the maximum NDVI values from each season. This is apparent in all years with the exception of 2007 where the maximum NDVI value is high and the yield value is relatively low.

subtracted from the maximum NDVI growing season values, as expressed in Eq. (2).

$$MA\_NDVI_Y = VI_{\max 95,y} - \frac{1}{N}\left[\sum_{y=1}^{N} VI_{\min 5,y}\right] \quad (2)$$

where N is the number of years (2000 through 2008), y is the year, $\mathrm{VI}_{\max 95,y}$ is the maximum 95th percentile NDVI for Year y, $\mathrm{VI}_{\min 5,y}$ is the minimum 5th percentile NDVI for Year y.

The maximum adjusted NDVI (MA-NDVI) from 2007 was excluded from the time-series due to the April frosts.

Once the winter wheat MA-NDVI values were retrieved from each county for the 8 years of data, percent wheat specific, linear relationships were derived between the retrieved MA-NDVI and the NASS reported yield statistics for each county. To derive these relationships, the maximum percent (Mpct) wheat value for each county was computed as a weighted average of the percent wheat values of the top 5% purest pixels in each county, as is shown in Fig. 5. A linear regression was then derived for each county (with intercept set to 0), regressing eight years of the averaged maximum MA-NDVI against eight years of NASS yield statistics for each county. The county-specific regression slope and maximum percent wheat values were then used in the next step for generalizing the MA-NDVI to yield relationship as a function of percent wheat.

The assumption was that the MA-NDVI value from each county is a mixed signal composed of the proportional signals from MA-NDVI of wheat and the MA-NDVI of the other land cover (Eq. (3)).

$$MA\_NDVI_{\mathrm{pixel}\_x} = (MA\_NDVI_{\mathrm{wheat}} * Pct_{\mathrm{Wheat}}) + (MA\_NDVI_{\mathrm{other}} * (1 - Pct_{\mathrm{wheat}})) \quad (3)$$

As winter wheat fields are generally located in widely cultivated areas and are surrounded primarily by fields sown with spring-planted row crops rather than by forests or other land cover classes, the MA-NDVI signal from the other land cover types is likely dominated by non-wheat cultivated lands. Winter wheat is sown in the fall and by mid-spring reaches its maximum green canopy while the other, non-wheat fields are still bare or at the start of their vegetative growth phase. It was therefore assumed that $\mathrm{MA\_NDVI}_{\mathrm{other}} < \mathrm{MA\_NDVI}_{\mathrm{wheat}}$. In this case it follows that where the percent wheat is low, the MA-NDVI signal will be lower than the MA-NDVI signal where the percent wheat is higher as the bare-ground and emerging row-crop signal will dominate the MA-NDVI signal. Given these assumptions and since the averaged maximum percent wheat (Mpct) varies by county we expected the county-specific regression slopes of yield to MA-NDVI to vary as a function of wheat percent. In other words, in cases where winter wheat percent is low we expect the slope to be proportionally larger than in cases where wheat percent is higher, in order to account for the weaker $\mathrm{MA\_NDVI}_{\mathrm{wheat}}$ signal which is due to the lower wheat percent.

At the county level the MA-NDVI was found to be a good predictor of yield and the slopes of these linear relationships varied between counties primarily as a function of the averaged maximum percent wheat. As expected, counties with lower maximum wheat percent had larger slopes than those with higher maximum percent wheat values. An example from two counties, Harper and Dickinson is presented in Fig. 6. Harper County had a maximum wheat percent of 74.25 and a MA-NDVI to Yield regression slope of 5.34. On the other hand, Dickinson County had a lower wheat percent of 41.48 and a regression slope of 7.78. Once the percent wheat dependent slopes were derived, the next step was to obtain a single generalized yield-MA-NDVI relationship that could capture the percent wheat dependent variation in slope, and be applied uniformly to the entire state. To accomplish this, a linear relationship was derived, regressing the individual county slopes computed in the previous step, against the county averaged maximum percent winter wheat. Three filters were applied to the data to select the best-suited data points for the regression: i) county average harvested area > 4000 ha ii) RMSE of yield to MA-NDVI regression < 0.43 MT/ha and iii) regression correlation coefficient > 0.6. Once the filters were applied, the remaining data points were used to derive the wheat percent to slope regression. This regression is presented in Fig. 7. The X-axis on Fig. 7 is the averaged maximum winter wheat percent value and the Y-axis is the regression slope derived from the individual counties. As shown in Fig. 7, the slope and maximum wheat percent were negatively and linearly correlated. As the percent wheat increases the slope decreases. In addition, dispersion in slope values increases at low maximum percent crop values. The derived regression is presented in Eq. (4). The regression Root Mean Square Error (RMSE) was 0.87, which is equivalent to a 10% error and the correlation coefficient was −0.7.

$$S_{\mathrm{Mpct}} = 9.61 + (-0.05 \times Mpct) \quad (4)$$

where Mpct is maximum percent wheat and S_{Mpct} is the corresponding slope of the yield to MA-NDVI regression as a function of Mpct.

Although the averaged maximum wheat percent in some counties is relatively low (below 50%), and as such the majority of the MA_NDVI signal is from $\mathrm{MA\_NDVI}_{\mathrm{other}}$, this generalized regression proportionally adjusts the MA_NDVI with a higher slope thus giving a proportionally higher weight to the $\mathrm{MA\_NDVI}_{\mathrm{wheat}}$ allowing the extraction of the MA_NDVI wheat signal even where it is proportionally low.

This linear regression was used to compute the slope of the yield to MA-NDVI regression as a function of the maximum percent wheat value. It was applied uniformly to the entire state of Kansas to obtain winter wheat yield in tons/ha for each county by multiplying the MA-NDVI value for each county by the corresponding slope derived

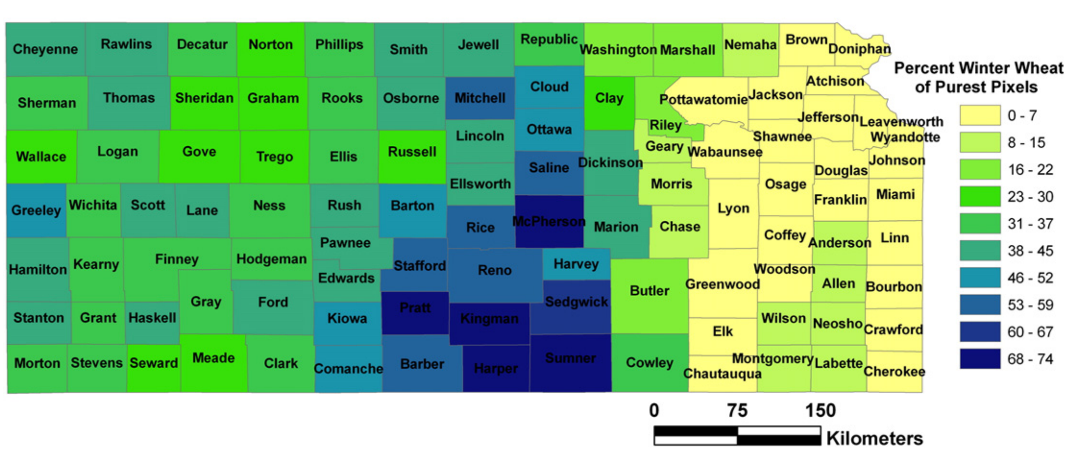

Fig. 5. This figure shows the distribution by county of the maximum percent winter wheat computed as a weighted average of the 5% purest pixels in each county in Kansas.

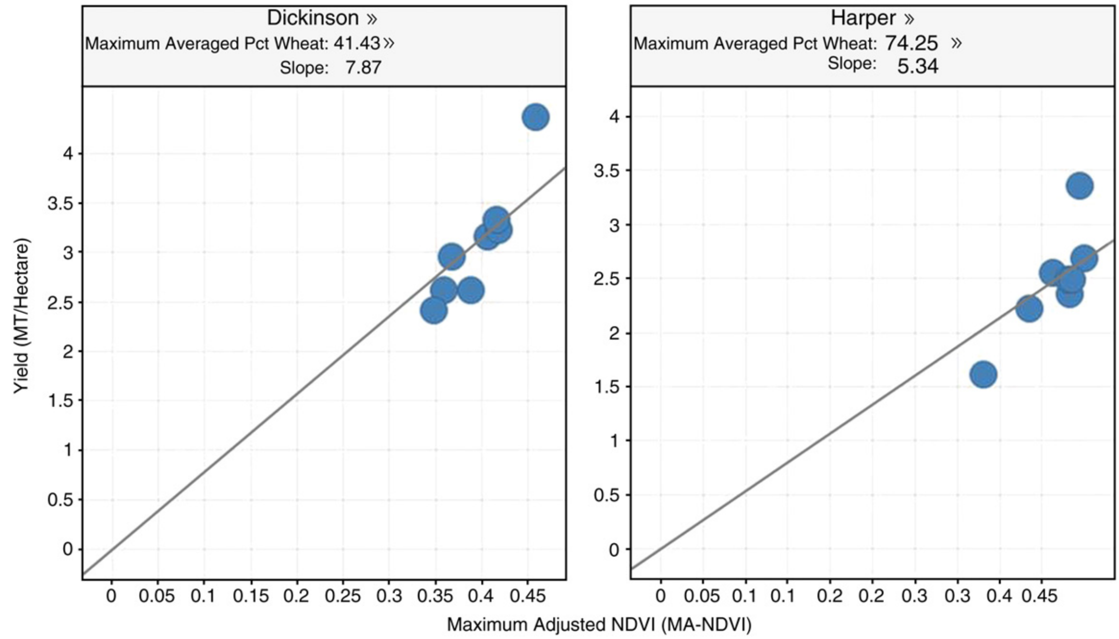

Fig. 6. This figure shows an example of two MA-NDV to Yield regressions from two Kansas counties with differing maximum percent wheat county values. Harper County which has a maximum wheat percent of 74.25 has a MA-NDVI to Yield regression slope of 5.34 whereas Dickinson County which has a lower wheat percent of 41.48 and has a higher regression slope of 7.78.

from Eq. (2). The annual per county yield was then multiplied by the NASS county area harvested statistics, to derive a statewide annual production number in million metric tons (MMT) (Eqs. (5) and (6)). Using this generalized regression model, production statistics were estimated for 2000 through 2008 for the state of Kansas.

$$\textit{Forecast Yield} = \textit{MA\_NDVI}_{\mathrm{Mpct},y} \times S_{\mathrm{Mpct}}. \tag{5}$$

$$\textit{Forecast Production} = \textit{Forecast Yield} \times \textit{Official Area Harvested}. \tag{6}$$

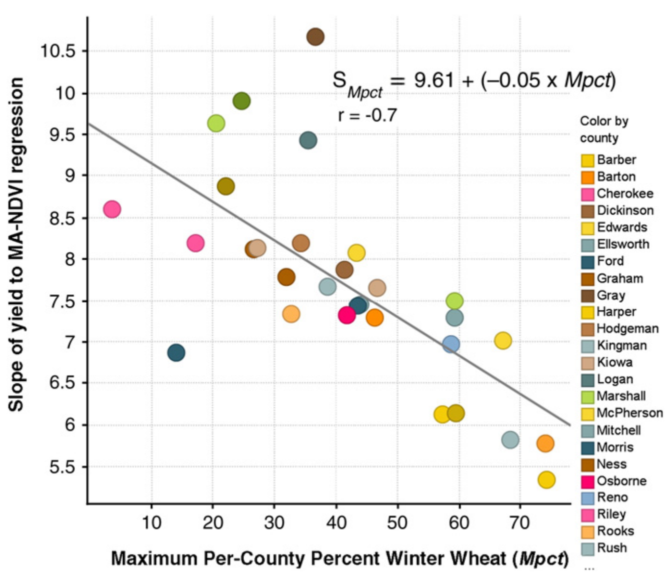

Fig. 7. This figure shows the derived linear relationship between the individual county slopes (such as those shown in Fig. 6) and the per-county maximum percent winter wheat. The *X*-axis is the county maximum winter wheat percent value and the *Y*-axis is the regression slope derived from the individual counties. It should be noted that the slope and maximum wheat percent are negatively and linearly correlated. As the percent wheat increases the slope decreases. In addition, dispersion in slope values increases at low maximum percent crop values. The regression equation is at the top of the graph. The regression Root Mean Square Error (RMSE) is 0.87, and the correlation coefficient is −0.7.

4. Regression model results for Kansas and Ukraine

The regression model described above, trained on agricultural statistics from Kansas, was applied first in Kansas and then directly applied to Ukraine, in order to evaluate its portability to another major wheat producing region of the world.

4.1. Kansas

The Kansas regression model was run for 8 years of data from 2000 through 2008 excluding 2007 to predict state level yields. Winter wheat production was predicted by multiplying the predicted yields by the NASS reported area harvested statistics. The model predictions were validated against the official yield and production NASS statistics. A scatter plot of the predicted versus reported official yield and production statistics is presented in Fig. 8. The RMSE of the official versus predicted yields was 0.18 MT/ha which is equivalent to a 7% error. The regression slope was 1.0334 with the intercept set to 0 and the regression coefficient was 0.88. The RMSE of the official versus predicted production was 0.67 MMT which is equivalent to a 7% error. The regression slope was 1.0352 with the intercept set to 0 and the regression coefficient was 0.94.

4.2. Ukraine

To directly transfer the regression model (Eq. (4)) developed in Kansas to Ukraine a percent winter wheat mask and a time-series of MA-NDVI were required. A percent wheat mask at the 0.05° resolution was produced from the MODIS-derived 250 m winter wheat map (using Eq. (1)) and is shown in Fig. 9A. Similarly to Kansas, it was assumed that at the 0.05° scale, the per pixel percent wheat remains relatively static between years. According to oblast-level planted area statistics from the Ukraine SSC there is an average 11% difference between 2008 planted area and the mean planted area (2000–2008).

The same methods used in Kansas to retrieve a MA-NDVI time-series were used in Ukraine. First the top 5% purest pixels in each oblast were identified and used as a spatial mask for NDVI retrieval. The weighted average of these purest pixels for each oblast is shown in Fig. 9B. For each year of data between 2000 and 2008 the maximum

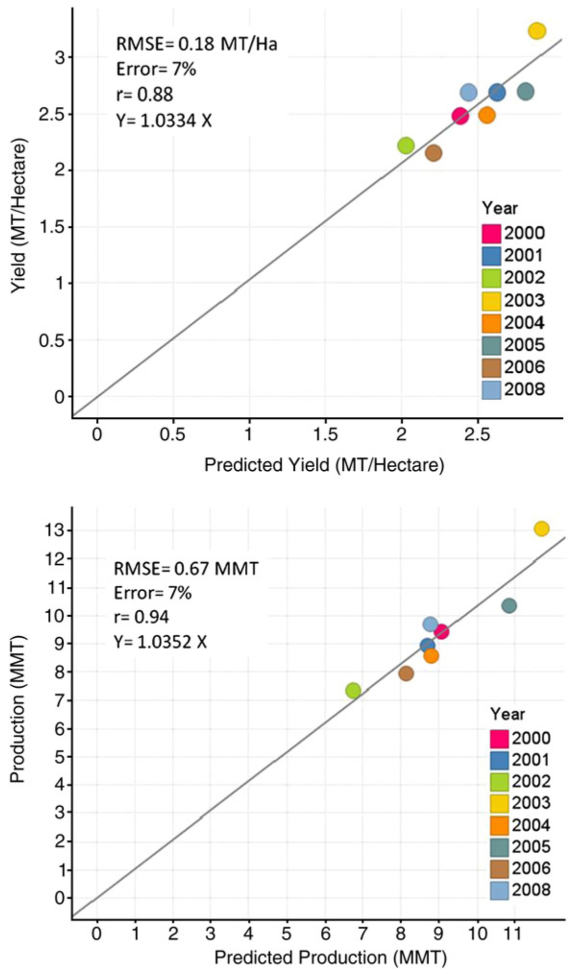

Fig. 8. This figure shows the results from running the regression model in Kansas. Panel (A) is a scatter plot of the estimated versus the official yield statistics and Panel (B) is a scatter plot of the estimated versus official production statistics. The RMSE of the official versus estimated yields estimates is 0.18 MT/ha which is equivalent to a 7% error and the regression coefficient is 0.88. The RMSE of the official versus estimated production is 0.67 MMT which is equivalent to a 7% error and the regression coefficient is 0.94.

seasonal NDVI was extracted for the purest winter wheat pixels from each oblast (the peak of the winter wheat growing season in Ukraine is between mid April and early June). The MA-NDVI was computed in the same way as it was computed in Kansas (Eq. (2)). The percent wheat dependent Kansas model, described by Eq. (4), was then directly run on the entire country. As in Kansas, the model-predicted yields were multiplied by official area harvested statistics from the Ukraine State Statistical Committee to obtain production estimates.

The regression model developed in Kansas proved to be directly applicable to Ukraine without calibration against Ukrainian yield statistics. The model results were validated against the official Ukrainian statistics and are presented in Figs. 10 and 11. In Fig. 10 a time-series graph of the official production statistics, in pink, are compared to the predicted production, in blue, where the year is on the *X*-axis and winter wheat production in Million of Metrics Tons is on the *Y*-axis. This graph shows that the regression model predictions were in good agreement with the official Ukrainian statistics capturing the fluctuations through time of winter wheat yields and production. For example, in 2003 over 60% of winter wheat in Ukraine was destroyed due to December frost damage and to a persistent ice crust that formed in February as a result of repeated cycles of thawing and refreezing (FAS, 2003). In contrast in 2008 the winter wheat crop benefited from excellent weather throughout the growing season and the yields reached a 15-year high (FAS, 2008). The Kansas model was

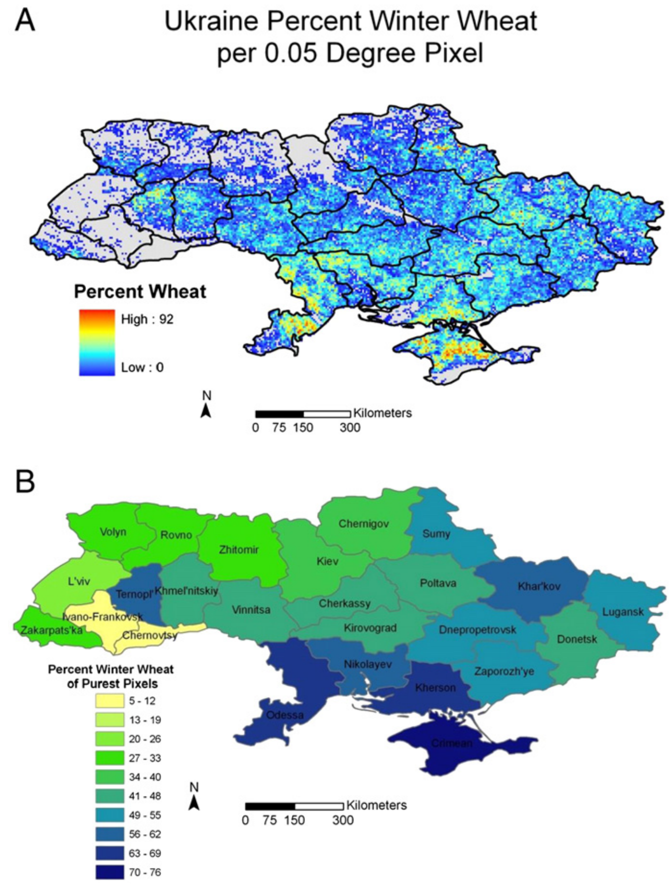

Fig. 9. Panel (A) is the percent winter wheat mask (0.05°) scaled up from the 250 m classified winter wheat mask. Panel (B) shows the distribution by oblast of the maximum percent winter wheat computed from (A), as a weighted average of the 5% purest pixels within each oblast.

able to capture both the dramatic drop in yield in Ukraine in 2003 and the sharp increase in yield in 2008, six weeks prior to harvest. The model predictions were validated against the official Ukrainian crop statistics and are presented in two scatter plots in Fig. 11. The RMSE of the official versus predicted yields was 0.44 MT/ha which is equivalent to a 15% error. The regression slope was 0.9433 with the intercept set to 0, and the regression coefficient was 0.74. The RMSE of the official versus predicted production was 1.83 MMT which is

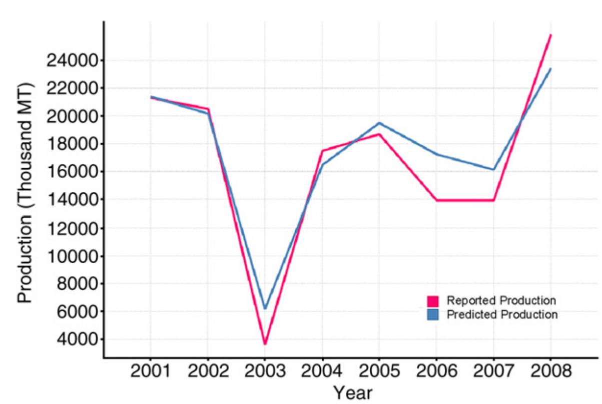

Fig. 10. This figure shows the results of directly applying the regression model developed in Kansas to Ukraine. In this time-series graph the official production statistics, in pink, are compared to the predicted production, in blue, where the year is on the *X*-axis and wheat production in thousands of metrics tons is on the *Y*-axis.

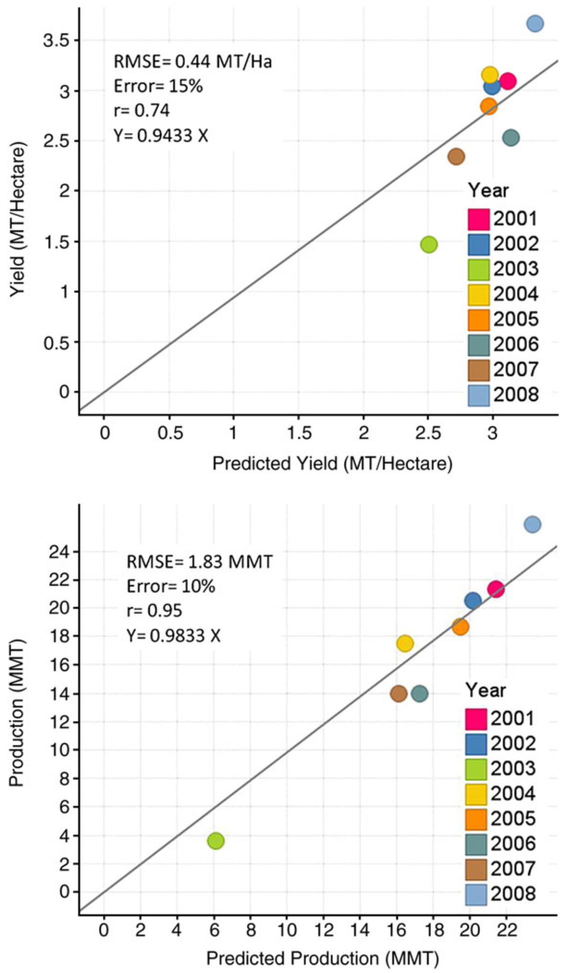

Fig. 11. This figure shows two scatter plots validating the Ukraine predictions produced by the model developed in Kansas against the official Ukrainian yield and production statistics. The RMSE of the official versus estimated yields (A) is 0.44 MT/ha which is equivalent to a 15% error and the regression coefficient is 0.74. The RMSE of the official versus estimated production (B) is 1.83 MMT which is equivalent to a 10% error and the regression coefficient is 0.95.

equivalent to a 10% error. The regression slope was 0.9983 with the intercept set to 0, and the regression coefficient was 0.95.

To further evaluate the model, it was run in real-time in May of 2009 to forecast winter wheat production in Ukraine. The model forecast 19.2522 MMT of wheat for the 2009 season. According to the official Ukrainian SSC statistics, which became available at the end of the season, winter wheat production was 20.5390 MMT in 2009, which means that the model forecasts were within 6.3% of the official 2009 statistics.

5. Discussion

The regression-based model developed in this study was empirically-based and utilized official crop statistics from Kansas to derive a relationship between winter wheat specific remotely sensed parameters and reported yield statistics. One of the main drawbacks of such remotely sensed based empirical models for estimating crop yields has been that their application is valid only for the areas they have been calibrated for (Doraiswamy et al., 2003). This study developed a single generalized-model that was applied at the state level in Kansas and was proven directly applicable to Ukraine.

The winter wheat regression-based model developed in this study assumes, like many other empirical remotely sensed based yield models (Ferencz et al., 2004) that the canopy vigor of winter wheat estimated by spectral NDVI measurements is directly related to final winter wheat yields. Specifically, NDVI measurements around the time of the maximum, which encompass the 'critical period' for grain production (Labus et al., 2002) have been found to be strongly correlated with final yields (Benedetti and Rossini, 1993; Doraiswamy and Cook, 1995; Tucker et al., 1980). The majority of such studies developed approaches which rely on NDVI measurements from composite rather than daily data sets as the data quality and data volume was an issue. In this study, owing to the new, coarse resolution, high quality daily, BRDF-corrected CMG MODIS surface reflectance data, the seasonal peak NDVI was estimated with high confidence from the daily data.

As has been well documented in the literature, this study found that yield was positively and linearly correlated to the seasonal maximum NDVI at the county scale. The goal of this study was to go beyond this finding and develop a regression model that was transferable and directly applicable at the state and national levels. Two steps were taken to generalize the linear relationships found at the county scale so that it could be widely applied and transferred. The assumption underlying the generalization was that the positive and linear relationship between the maximum NDVI signal of pure winter wheat pixels and yield is constant between locations. Therefore if the maximum $\text{NDVI}_{\text{wheat}}$ signal can be unmixed from the maximum NDVI signal it can be used directly to predict yield. The first step towards isolating the wheat signal was to minimize the noise in the maximum NDVI signal retrieved from the primarily wheat areas by removing the background signal such as the soil signal. This was accomplished by subtracting the average minimum NDVI from the maximum seasonal NDVI. This adjusted measure was termed the Adjusted Maximum NDVI (MA-NDVI). The MA-NDVI allowed for generalization and direct comparison between the maximum NDVI signals retrieved from different locations.

The second step was to un-mix the $\text{MA\_NDVI}_{\text{wheat}}$ signal from the $\text{MA\_NDVI}_{\text{other}}$ signal by accounting for wheat percent. This was accomplished by deriving a set of relationships between yield and MA-NDVI and then generalizing the slopes of these regressions as a function of wheat percent. As expected, we found that the slope of the MA-NDVI to yield regression decreased linearly as a function of increasing wheat purity. The explanation for this variation in slope as a function of winter wheat percent is based on the fact that wheat is generally located within cropped areas sown with spring-planted row crops rather than by forests or other land cover classes. In Kansas and Ukraine, where winter wheat is sown in the fall, wheat reaches its maximum green canopy by mid-spring while the other, non-wheat fields are still bare or at the start of their vegetative growth phase. Therefore, in the cases of low wheat-percentage, the wheat reflectance signal is relatively weak as it is suppressed by the bare-ground and emerging row-crop signal. Consequently the MA-NDVI signal in these circumstances is lower than the MA-NDVI of purer pixels where a larger portion of the MA-NDVI signal is from wheat. This negative linear relationship between slope and wheat percent was used to derive a single, broadly applicable regression model which enabled forecasting of wheat yields even in counties and oblasts where the majority of the MA-NDVI signal was from non-wheat.

These two steps for unmixing the maximum wheat NDVI signal, namely the maximum NDVI noise reduction and the depiction of the regression slope as a function of wheat purity, enabled the development of a robust, simple, remotely sensed based generalized-model which was applicable at the state level in Kansas, at national level in Ukraine and potentially directly applicable to other wheat growing regions in the world.

Many remote sensing studies on yield forecasting argue that moderate and high resolution data (i.e. finer than 60 m) are needed in order to improve yield models. In this study we found that the coarse resolution offered a significant advantage over higher resolution data.

This is due to the fact that at moderate and high resolution data scales any single pixel of winter wheat will likely shift between crop types from one year to the next due to crop rotations. As the regression model is based on relating the wheat specific NDVI signal to yield, it requires a-priori knowledge of the winter wheat field locations. This means that due to the widespread practice of crop rotations, running the model on moderate and high resolution data would require year-specific winter wheat crop type maps to be produced prior to the end of the season. Identifying a particular crop in the year to be predicted when only incomplete growing season information is available, is a challenge (Kastens et al., 2005).

On the other hand, at the coarse resolution (0.05°), the winter wheat percent was found to be significantly less variable than at the higher resolution. This enabled the model to run using a single wheat percent map, by deriving a yield to NDVI relationship as a function of percent wheat. This finding has two advantages, one is that the regression model did not require an annual winter wheat specific map to identify wheat pixels and the second is that the crop type map did not need to be generated prior to the end of the growing season. This finding is true for Kansas and Ukraine where the fields are large and generally range between 30 and 300 ha. In areas with smaller fields the 0.05° resolution will likely be too coarse, though a similar scaling approach could be applied by scaling up very high resolution crop type map (i.e. 10 m) to moderate resolution (i.e. 250 m) crop percent map.

There are several limitations to utilizing such an empirical, remotely sensed based regression model which relies on measurements of vegetation photosynthetic capacity to estimate yields. One of these limitations is that it cannot capture the impact of events that reduce yield but do not reduce the peak green biomass. As in the 2007 Kansas case, such events can have devastating impacts on yield though this impact is not captured by the presented regression model. One way to address this is to incorporate minimum-daily temperature along with a growing-degree-day phenology model that can detect frost events around the time of anthesis. Another challenge of this approach is that it requires a minimum of one, representative, winter wheat crop type map, which is often not available. In addition, large-scale shifts in winter wheat planted area, due to market pressures, tax laws, etc. would limit the model application since the underlying wheat percent map would no longer be valid. In such a case, a new representative crop percent map would be required.

Although this model was successfully implemented in two different locations it is not clear how it would perform in other wheat growing regions. For example, in regions that have very high yields and very dense green biomass, where the NDVI is likely to saturate prior to capturing the seasonal green biomass peak and therefore the model study peak NDVI would not serve as a good predictor of yield.

We envision that this model could potentially be well suited for regions where crop statistics and ground information on yield are limited and could serve as a good indicator of yield prior to harvest. Therefore the next step will be to evaluate the model's performance in a range of winter wheat cropping systems in order to better understand its forecasting capacity and limitations and to further evaluate its basic assumption that the positive and linear relationship between the maximum NDVI signal of pure and yield is constant between locations.

We recognize that the current model was developed based on the 8 years of data available at the time of the study. As new data from MODIS and NASS become available it will be integrated in order to enhance and calibrate the basic model.

6. Conclusions

As new remote sensing instruments and data become available their utility for improving established terrestrial monitoring tasks need to be evaluated. In this study an empirical, generalized remotely sensed based yield model was developed and successfully applied at the state level in Kansas using daily, high quality 0.05° NDVI time-series data to drive the regression model, a percent crop mask as a filter to identify the purest winter wheat pixels, and USDA NASS county crop statistics for model calibration. The model predictions of production in Kansas closely matched the USDA/NASS reported numbers with a 7% error. This empirical regression model that was developed in Kansas was successfully applied directly in Ukraine. The model forecast winter wheat production in Ukraine six weeks prior to harvest with a 10% error of the official production numbers. In 2009 the model was run in real-time in Ukraine and forecast production within 7% of the official statistics which were released after the harvest. This model is simple, has limited data requirements, and can offer an indication of winter wheat production, shortfalls and surplus prior to harvest in regions where little ground data is available. As the results of this study are promising, the performance of the developed model should now be assessed in the other main winter wheat growing regions of the world.

References

Allen, R. H., Hanuscak, G., & Graig, M. (2002). *Limited use of remotely sensed data for crop condition monitoring and crop yield forecasting in NASS.* http://www.nass.usda.gov/research/avhrr/remoteuse.pdf

Basnyat, P., McConkey, B., Lafond, G. R., Moulin, A., & Pelcat, Y. (2004). Optimal time for remote sensing to relate to crop grain yield on the Canadian prairies. *Canadian Journal of Plant Science*, *84*, 97−103.

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 Sensing of Environment*, *45*, 311−326.

Boken, V. K., & Shaykewich, C. F. (2002). Improving an operational wheat yield model using phenological phase-based Normalized Difference Vegetation Index. *International Journal of Remote Sensing*, *23*, 4155−4168.

Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M. H., Ruget, F., Nicoullaud, B., et al. (1998). STICS: A generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn. *Agronomie*, *18*, 311−346.

Chipanshi, A. C., Ripley, E. A., & Lawford, R. G. (1999). Large-scale simulation of wheat yields in a semi-arid environment using a crop-growth model. *Agricultural Systems*, *59*, 57−66.

De Fries, R. S., Hansen, M., Townshend, J. R. G., & Sohlberg, R. (1998). Global land cover classifications at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing*, *19*, 3141−3168.

Doraiswamy, P. C., & Cook, P. W. (1995). Spring wheat yield assessment using NOAA AVHRR data. *Canadian Journal of Remote Sensing*, *21*, 43−51.

Doraiswamy, P. C., Moulin, S., Cook, P. W., & Stern, A. (2003). Crop yield assessment from remote sensing. *Photogrammetric Engineering and Remote Sensing*, *69*, 665−674.

Dubey, R. P., Ajwani, N., Kalubarme, M. H., Sridhar, V. N., Navalgund, R. R., Mahey, R. K., et al. (1994). Preharvest wheat yield and production estimation for the Punjab, India. *International Journal of Remote Sensing*, *15*, 2137−2144.

FAO. (2003). *Agricultural commodities: Profiles and relevant WTO negotiating issues.* Rome: UN FAO.

FAS. (2003). *Ukraine: Extensive damage to winter wheat*. Washington DC: USDA Foreign Agricultural Service.

FAS. (2008). Ukraine 2008 commodity intelligence report. Washington DC: U.S. Department of Agriculture (USDA) Foreign Agricultural Service (FAS).

FAS. (2009). *Production supply distribution online database*. United States Department of Agriculture (USDA), http://www.fas.usda.gov/psdonline

Ferencz, C., Bognar, P., Lichtenberger, J., Hamar, D., Tarscai, G., Timar, G., et al. (2004). Crop yield estimation by satellite remote sensing. *International Journal of Remote Sensing*, *25*, 4113−4149.

Fischer, R. A. (1975). Yield potential in dwarf spring wheat and the effect of shading. *Crop Science*, *15*, 607−613.

Funk, C., & Budde, M. E. (2009). Phenologically-tuned MODIS NDVI-based production anomaly estimates for Zimbabwe. *Remote Sensing of Environment*, *113*, 115−125.

Genovese, G., Vignolles, C., Negre, 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.

Hansen, M. C., Defries, R. S., Townshend, J. R. G., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, *21*, 1331−1364.

Hatfield, J. L. (1983). Remote sensing estimators of potential and actual crop yield. *Remote Sensing of Environment*, *13*, 301−311.

Hayes, M. J., & Decker, W. L. (1996). Using NOAA AVHRR data to estimate maize production in the United States Corn Belt. *International Journal of Remote Sensing*, *17*, 3189−3200.

Holben, B. N. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, *7*, 1417−1434.

Idso, S. B., Pinter, P. J., Hatfield, J. L., Jackson, R. D., & Reginato, R. J. (1979). A remote sensing model for the prediction of wheat yields prior to harvest. *Journal of Theoretical Biology*, *77*, 217−228.

Jackson, R. D., Pinter, P. J., Reginato, R. J., & Idso, S. B. (1986). Detection and evaluation of plant stresses for crop management decisions. *IEEE Transactions on Geoscience and Remote Sensing*, *1*, 99−106.

Justice, C. O., & Becker-Reshef, I. (2007). *Report from the workshop on developing a strategy for global agricultural monitoring in the framework of Group on Earth Observations (GEO)* (pp. 1−67).

Justice, C. O., Vermote, E., Townshend, J. R. G., Defries, R., Roy, D. P., Hall, D. K., et al. (1998). The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, *36*, 1228−1249.

Kastens, J. H., Kastens, T. L., Kastens, D. L. A., Price, K. P., Martinko, E. A., & Lee, R. -Y. (2005). Image masking for crop yield forecasting using AVHRR NDVI time series imagery. *Remote Sensing of Environment*, *99*, 341−356.

Labus, M. P., Nielsen, G. A., Lawrence, R. L., Engel, R., & Long, D. S. (2002). Wheat yield estimates using multi-temporal NDVI satellite imagery. *International Journal of Remote Sensing*, *23*, 4169−4180.

MacDonald, R. B., & Hall, F. G. (1980). Global crop forecasting. *Science*, *208*, 670−679.

Mahey, R. K., Singh, R., Sidhu, S. S., Narang, R. S., Dadhwal, V. K., Parihar, J. S., et al. (1993). Preharvest state-level wheat acreage estimation using IRS-IA LISS-I data in Punjab (India). *International Journal of Remote Sensing*, *14*, 1099−1106.

Manjunath, K. R., Potdar, M. B., & Purohit, N. L. (2002). Large area operational wheat yield model development and validation based on spectral and meteorological data. *International Journal of Remote Sensing*, *23*, 3023−3038.

Maselli, F., & Rembold, F. (2001). Analysis of GAC NDVI data for cropland identification and yield forecasting in Mediterranean African countries. *Photogrammetric Engineering and Remote Sensing*, *67*, 593−602.

Mika, J., Kerenyi, J., Rimoczi-Paal, A., Merza, A., Szinell, C., & Csiszar, I. (2002). On correlation of maize and wheat yield with NDVI: Example of Hungary (1985–1998). *Advances in Space Research*, *30*, 2399−2404.

Miller, D. T. (1999). Growth stages of wheat: Identification and understanding improve crop management. *Texas Agricultural Extnsion Service*.

Mitchell, D. O., & Mielke, M. (2005). Wheat: The global market, policies and priorities. In A. M. A. & J. C. B. (Eds.), *Global agricultural trade and developing countries* (pp. 195−214). Washington DC: World Bank.

Mkhabela, M. S., Mkhabela, M. S., & Mashinini, N. N. (2005). Early maize yield forecasting in the four agro-ecological regions of Swaziland using NDVI data derived from NOAA's-AVHRR. *Agricultural and Forest Meteorology*, *129*, 1−9.

Moriondo, M., Maselli, F., & Bindi, M. (2007). A simple model of regional wheat yield based on NDVI data. *European Journal of Agronomy*, *26*, 266−274.

NASS. (2007). *Kansas Crop weather report*: USDA National Agricultural Statistics Service, Kansas Field Office, April 9, 2007.

NASS. (2007). *Kansas Crop weather report*. USDA National Agricultural Statistics Service, Kansas Field Office, April 23, 2007.

NASS. (2007). *Cropland data layer*. USDA NASS, *Research and Development Division*.

NASS. (2007). *Statistical highlights of U.S. Agriculture report*. Kansas: USDA NASS, October, 2007.

NASS. (2008). *NASS quick stats*. USDA National Agricultural Statistics Service (NASS). http://www.nass.usda.gov/QuickStats/

Patel, N. R., Bhattacharjee, B., Mohammed, A. J., Tanupriya, B., & Saha, S. K. (2006). Remote sensing of regional yield assessment of wheat in Haryana, India. *International Journal of Remote Sensing*, *27*, 4071−4090.

Pinter, P. J., Jackson, R. D., Idso, S. B., & Reginato, R. J. (1981). Multidate spectral reflectances as predictors of yield in water stressed wheat and barley. *International Journal of Remote Sensing*, *2*, 43−48.

Prasad, A. K., Chai, L., Singh, R. P., & Kafatos, M. (2006). Crop yield estimation model for Iowa using remote sensing and surface parameters. *International Journal of Applied Earth Observation and Geoinformation*, *8*, 26−33.

Quarmby, N. A., Milnes, M., Hindle, T. L., & Silleos, N. (1993). The use of multitemporal NDVI measurements from AVHRR data for crop yield estimation and prediction. *International Journal of Remote Sensing*, *14*, 199−210.

Rasmussen, M. S. (1997). Operational yield forecast using AVHRR NDVI data: Reduction of environmental and inter-annual variability. *International Journal of Remote Sensing*, *18*, 1059−1077.

Ren, J. Q., Chen, Z. X., Zhou, Q. B., & Tang, H. J. (2008). Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong, China. *International Journal of Applied Earth Observation and Geoinformation*, *10*, 403−413.

Ritchie, J. T., & Otter, S. (1985). Description and performance of CERES-Wheat: A user-oriented wheat yield model. *ARS wheat yield project* (pp. 159−175). Springfield, Virginia: USDA ARS.

Rojas, O. (2007). Operational maize yield model development and validation based on remote sensing and agro-meteorological data in Kenya. *International Journal of Remote Sensing*, *28*, 3775−3793.

Salazar, L., Kogan, F., & Roytman, L. (2007). Use of remote sensing data for estimation of winter wheat yield in the United States. *International Journal of Remote Sensing*, *28*, 3795−3811.

Santoleri, R., Marullo, S., & Bohm, E. (1991). An objective analysis scheme for AVHRR imagery. *International Journal of Remote Sensing*, *12*, 681−693.

Sellers, P. J. (1985). Canopy reflectance, photosynthesis and transpiration. *International Journal of Remote Sensing*, *6*, 1335−1372.

Smith, R. C. G., Adams, J., Stephens, D. J., & Hick, P. T. (1995). Forecasting wheat yield in a Mediterranean-type environment from the NOAA satellite. *Australian Journal of Agricultural Research*, *46*, 113−125.

Sridhar, V. N., Dadhwal, V. K., Chaudhari, K. N., Sharma, R., Bairagi, G. D., & Sharma, A. K. (1994). Wheat production forecasting for a predominantly unirrigated region in Madhya-Pradesh (India). *International Journal of Remote Sensing*, *15*, 1307−1316.

Trostle, R. (2008). Global Agricultural Supply and Demand: Factors Contributing to the Recent Increase in Food Commodity Prices. (pp. 1–30): USDA Economic Research Service.

Tucker, C. J. (1979). Red and photographic infrared linear combination for monitoring vegetation. *Remote Sensing of Environment*, *8*.

Tucker, C. J., Holben, B. N., Elgin, J. H., & McMurtrey, J. E. (1980). Relationships of spectral data to grain yield variation. *Photogrammetric Engineering and Remote Sensing*, *46*, 657−666.

Tucker, C. J., Holben, B. N., Elgin, J. H., & McMurtrey, J. E. (1981). Remote-sensing of total dry-matter accumulation in winter-wheat. *Remote Sensing of Environment*, *11*, 171−189.

Vandiepen, C. A., Wolf, J., Vankeulen, H., & Rappoldt, C. (1989). Wofost − A simulation-model of crop production. *Soil Use and Management*, *5*, 16−24.

Vanevert, F. K., & Campbell, G. S. (1994). Cropsyst − A collection of object-oriented simulation-models of agricultural systems. *Agronomy Journal*, *86*, 325−331.

Verdin, J., Funk, C., Klaver, R., & Roberts, D. (1999). Exploring the correlation between Southern Africa NDVI and Pacific sea surface temperatures: Results for the 1998 maize growing season. *International Journal of Remote Sensing*, *20*, 2117−2124.

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.

Vermote, E., Justice, C. O., & Breon, F. M. (2009). Towards a generalized approach for correction of the BRDF effect in MODIS directional reflectances. *IEEE Transactions on Geoscience and Remote Sensing*, *47*, 898−908.

Vocke, G. (2008). Wheat year in review (domestic): High U.S. exports drive record prices and low 2007/08 ending stocks. *Outlook Report* (pp. 17). USDA Economic Research Service.

Wall, L., Larocque, D., & Leger, P. M. (2007). The early explanatory power of NDVI in crop yield modeling. *International Journal of Remote Sensing*, *29*, 2211−2225.

Wang, Q., Adiku, S., Tenhunen, J., & Granier, A. (2005). On the relationship of NDVI with leaf area index in a deciduous forest site. *Remote Sensing of Environment*, *94*, 244−255.

Weissteiner, C. J., & Kuhbauch, W. (2005). Regional yield forecasts of malting barley (*Hordeum vulgare* L.) by NOAA-AVHRR remote sensing data and ancillary data. *Journal of Agronomy and Crop Science*, *191*, 308−320.

WFP. (2009). Food aid information system quantities delivered report. *World Food Program*.

Wiegand, C. L., & Richardson, A. J. (1990). Use of spectral vegetation indices to infer leaf area, evapotraspiration and yield: I. *Rationale. Journal of Agronomy*, *82*, 623−629.

Wittwer, S. H. (1995). *Food, climate, and carbon dioxide: The global environment and world food production*.

Woolverton, M. (2007). All eyes are on the Kansas wheat crop. *Grain Outlook*. Kansas State University Research and Extension.